Implementation of AAC Encoder for Audio Broadcasting

A.Parkavi[1], T.Kalpalatha Reddy[2].

[1]*PG Scholar,* [2]*Dean*

[1,2]*Department of Electronics and Communication Engineering*

S.K.R Engineering College, Chennai, India.

[1]*aruljothi@gmail.com,* [2]*kalpalathareddy-dean.ece@skrenggcollege.org*

Abstract*: MP3 is the popular audio coding standard. But now, a new higher quality audio coding standard Advanced Audio Coding (AAC) is proposed and widely used. The quantization/re-quantization is essential in both MP3 and AAC. It proposes a new high accuracy estimation algorithm for MP3 and MEPG-4 AAC audio coding. The algorithm can be applied not only for re-quantization process in decoder, but also for quantization in encoder. The implementation of the multichannel AAC encoder system for digital audio Broadcasting. The encoder system is based on MPEG-2/4 Advanced Audio Coding (AAC) and capable of real time encoding up to 5.1 channel audio. To give a flexible functionality, it consists of multiple DSPs, IEC61937 and TCPIIP interface and 6 channel audio input facilities. The reference AAC decoder was implemented for verification test of the encoder. The encoder system also integrated with AAC streaming system for interoperation test. Through these tests, the encoder system was verified to be a good solution for high quality audio broadcasting.*

Keywords- *Audio Codec, Bit-rate, Filter Bank, MPEG-4,MP3.*

I INTRODUCTION

AAC supports up to forty-eight audio channels. Sample rates supported range from 8kHz to 96kHz. The coder also supports three independent modes of operation: main, low complexity, and scalable sample rate profiles. The main profile has the highest quality at the expense of memory and processing power. The low complexity sacrifices some quality for much lower memory and processing power. The scalable sample rate mode is the lowest complexity of the three tools. AAC uses a combination of multiple coding tools to achieve bit rate reduction. The coding tools described below are used in the main profile configuration. The MPEG-2 AAC (Advanced Audio Coder) incorporates some very innovative technology in order to achieve low bit-rates and still retain fidelity. The coder can be manipulated for different levels of performance based on real-time constraints and supports a wide range of sample rates and data rates.MPEG audio technology has been widely applied to a variety of consumer products. Among the MPEG family standardization, MPEG Layer III (MP3) is the most popular due to the Internet and digital entertainment technology. Besides, in the recent future, MPEG-4 audio coding will be more and more important based on its higher compression ratio and audio quality. Since this application is applied based on the consumer application, it is necessary to be considered based on the low cost and low complexity requirements. No matter MP3 and MPEG-4, quantization and re-quantization process is necessary in encoder and decoder respectively. For quantization in the encoder, a non-uniform quantizer is used. Therefore the decoder must perform the inverse non-uniform quantization after the Huffman decoding of the scale factors and spectral data. The quantizer tool quantizes the data so that the quantization noise is shaped according to psychoacoustic criteria to be either completely masked or minimally audible, depending on the bit rate. Also, the number of bits for a frame must stay below a certain limit. An iterative process weighs the trade-off between noise suppression and bit starvation. Also, Huffman coding is utilized for noiseless coding of the spectral values. The Huffman strategy uses multiple codebooks and multiple dimensions to code the spectral values. The bit-stream formatter assembles the quantized and coded coefficients and the control parameters into a stream for transmission. The AAC bit-stream is composed of frames with varying size, depending on the variation caused by adaptive Huffman coding.

Although higher quality can be obtainable with the longer word length, longer word-length arithmetic results in high hardware cost and power consumption. Therefore, it is important to adapt optimized quantization/ re-quantization algorithm for short word-length arithmetic unit. It present a high quality quantization/re-quantization algorithm.

II OVERVIEW OF AAC

The decoder is essentially the inverse operations of the encoder. The details of the AAC standard and its earlier evolution are available in the literature [1,2]. However, the functionality and effectiveness of AAC can be intuitively explained as follows. The basic approach is that of an adaptive transform coder utilizing detailed psycho-acoustic models to conceal the quantization noise; however, there are several special features at each stage compared to a basic transform coder. Fig. 1 shows a block diagram of AAC encoder.
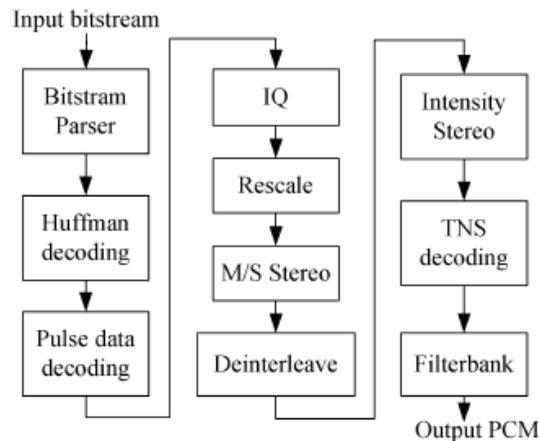

Fig.1: AAC Encoder

A high resolution, switched lapped transform is used, viz., MDCT [3], which effectively avoids the blocking effects of a transform coder. The adaptive switching provides for trading between coding gain and pre-echos that occur in transient portions of the signal. Temporal noise shaping (TNS) [4] aids this process by applying an adaptive predictor, in a novel way, in the transform domain. This is followed by a backward adaptive multi-channel predictor [SI, for each transform coefficient. This provides for exploiting signal redundancy longer than the transform width. However, this is not the same as adaptive long-term prediction (LTP of speech coders) which can exploit (spectral) periodicity within a block. The first block is the bit-stream parser, which extracts the audio frame signals and the decoding information that are used in the following decoding tools. In Huffman decoding, there are 12 Huffman codebooks. Eleven codebooks are used for spectrum coding and one for scale-factor coding. There are two in spectral Huffman decoding. Stage 1 is the Huffman decoding which unpacks the Huffman code index. Stage 2 is the De-grouping of 2- or 4-tuples of signed or unsigned code words into quantized spectral coefficients. De-grouping processing is performed using the algorithmic (division) approach. Among the 11 spectral Huffman codebooks, book 11 is a special case. It permits the encoding of quantized spectral coefficients even when their largest absolute value (LAV) is larger than 15. If the decoded value equals 16, an escape flag is used to signal the presence of a so-called escape sequence. The escape sequence consists of an escape prefix of -bits 1, an escape separator of 1-bit 0, and an escape word (N+4) of bits. The actual decoded value of the escape sequence is $2^{(N+4)}$ + escape word. Because the input maximum value of inverse quantization is 8191, the maximum length of the escape sequence is 21 bits. The next block is pulse data decoding which uses a pulse amplitude method to represent values larger than 15.When this block is utilized, one or several quantized coefficients are replaced by coefficients with smaller amplitudes in the encoder. In reconstructing the quantized spectral coefficients, these replacements are compensated by adding or subtracting amplitude from the previously decoded coefficients. Next, the quantized values are inversely quantized by the IQ tool and then scaled by the rescale tool.

The last, but most important feature of AAC, is that of quantizing transform domain information. Unlike standard transform coders, an adaptive noise-allocation strategy is used instead of adaptive bit-allocation [6,7]. To optimize quantization, several multi-dimensional entropy coders are used, which are adaptively switched to achieve least bit rate. The usual bit allocation algorithms assume a fixed bit-rate, which is not suited to entropy coding. This is circumvented by developing an iterative noise allocation scheme in which the transform coefficients are normalized and then quantized using entropy codes; the normalization factor (scale factor) itself determines the amount of quantization noise due to each coefficient. The iterative scheme tries to meet or exceed the masking threshold. In principle, this scheme can provide a fixed rate bit-stream. However, a small variation in the bit-rate is permitted and using a bit reservoir, some bits from a previous stationary segment can be passed on to the demanding transient segments. The uniqueness of the scale-factor approach is that it reduces the granularity of the standard bit-allocation methods as well as accrues some advantage due to successive quantization of transform coefficients.

III MPEG-2 AAC MULTI-CHANNEL ENCODER

3.1 Source Coding

We have seen rapid progress in source coding techniques for these signals. Linear prediction, sub-band coding, transform coding, as well as various forms of vector quantization and entropy coding techniques have been used to design efficient coding algorithms which can achieve substantially more compression than was thought possible only a few years ago. Recent results in speech and audio coding indicate that a good to excellent coding quality can be obtained with bit rates of 1 bit/sample for speech and wideband speech, 2 bit/sample for audio. Expectations over the next decade are that the rates can be reduced to at least 0.5 and 1 bit/sample, respectively. We shall show that such reductions can only be reached by employing sophisticated forms of adaptive noise shaping controlled by psychoacoustic criteria. Bit rate reduced digital representations of source signals not only allow the use of digital channels more efficiently, they can also be made less sensitive to channel impairments than analog representations if source and channel coding are implemented appropriately. Bandwidth expansion has often been mentioned as a disadvantage of digital coding and transmission, but with today's data compression and multilevel signaling techniques, one can actually reduce needed bandwidths, compared with analog systems. In broadcast systems, the reduced bandwidth requirements, together with the error robustness of the coding algorithms, will allow an efficient use of available radio and TV channels as well as "taboo" channels currently left vacant because of interference problems.

3.2 MPEG Standardization Activities

MPEG is currently working jointly with ITU's experts group on Video Coding for ATM Networks within Study Group XV in order to reach a common solution for transmitting audiovisual information over cell-based telecommunications networks. MPEG's initial effort was the MPEG-1 coding standard IS 11 172 supporting bit rates of around 1.2 Mb/s for video (with video quality comparable to that of analog video cassette recorders) and around 250 kb/s for two-channel audio (with audio quality comparable to that of today's compact discs). The coding standard IS 11172 consists of three parts: system, video, and audio. The system part defines a packet structure for multiplexing audio and video bit streams in one stream with the necessary information to keep the streams synchronized when decoding. The MPEG- 2 phase provides standards for high quality video (including High Definition TV) in bit rate ranges from 3-15 Mb/s and above. It provides new audio features including low bit rate and multichannel audio. Finally, MPEG-4 work addresses standardization of audiovisual coding at very low bit rates allowing for interactivity and universal accessibility and providing for a high degree of flexibility and extensibility 1351. In the case of audio coding, MPEG provides a three layer MPEG- 1 audio coding algorithm for stereophonic audio [33] and an MPEG-2 audio coding algorithm for multichannel audio. The three layers I, 11, or I11 of MPEG- 1 define coding algorithms with increasing complexity and performance. From a hardware and software standpoint, the higher layers incorporate the main building blocks of the lower layers.

Many companies are already offering one-chip decoders for all layers. MPEG standards will be described in detail in later sections.

IV   BIT RATE REDUCTION

Although high bit rate channels and networks become easier accessible, low bit rate coding of speech and audio signals has retained its importance. The main motivations for low bit rate coding are the need to minimize transmission costs or to provide cost-efficient storage, the demand to transmit over channels of limited capacity such as mobile radio channels, and to support variable-rate coding in packet-oriented networks. In addition, in audiovisual communications there is the need to share capacity between the audio and the video component. Basic requirements in the design of low bit rate speech or audio coders are firstly, to retain a high quality of the reconstructed signal with robustness to variations in spectra and levels. Secondly, robustness against random and burst channel bit errors and packet losses is required. Thirdly, low complexity and power consumption of the codec's are of high relevance. Table.1 shows bit rates in various storage devices. For example, in broadcast and playback applications, the complexity of audio decoders used must be low, whereas constraints on encoder complexity are more relaxed.

The algorithm of the proposed deregulated model of congestion management is given as follows.

Storage Device	Audio Rate	Overhead	Total Bit Rate
Compact Disc (CD)	1.41 Mb/s	2.91 Mb/s	4.32 Mb/s
Digital Audio Tap (DAT)	1.41 Mb/s	1.67 Mb/s	3.08 Mb/s
Digital Compact Cassette (DCC)	384 kb/s	384 kb/s	768 kb/s
MiniDisc (MD)	292 kb/s	718 kb/s	1.01 kb/s

Table.1 Bit rates in various storage devices

Additional network-related requirements are low encoder/decoder delays, robust tendering of codecs, Trans-codability, and a graceful degradation of quality with increasing bit error rates in mobile radio and broadcast applications. Finally, in professional applications, the coded bit streams must allow editing, fading, mixing, and dynamic range compression [4,5]. Coding of audiovisual signals needs an appropriate balance between audio and video bits, perhaps with a dynamic trading between them. A synchronization between these Bit-streams is required. The MPEG packet format is an important example for providing the capability to synchronize the delivery of video, audio, and auxiliary data packets in an ATM-like multiplexing environment. It will be used both in the United States Grand Alliance HDTV system and the European Hierarchical Digital Television Transmission (HDTVT) system. All of these partly conflicting factors have to be carefully considered in selecting a wideband speech or audio coding algorithm for a given audiovisual application.

V RESULTS

The bit-rate of the AAC decoder can be reduced to increase the frequency range about 20 KHz. It improves the approximation errors without multiply calculus for quantization/re-quantization but have the higher audio quality. The simulated output of MPEG-4 AAC decoder is shown in Fig.2. In this matlab output, the frequency range of the audio is increased with the help of bit-rate reduction. This gives the excellent audio quality.
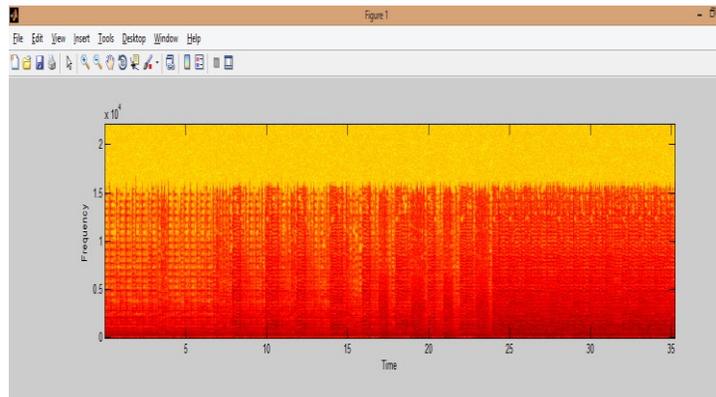
Fig. 2 simulated output of MPEG-4 AAC

VI. CONCLUSION

A new algorithm for the re-quantization/quantization was introduced. It reduced the quantization errors of re-quantization/quantization for MP3 and MPEG-4 audio coding. Compare to other algorithms, it has the lower approximation errors. If we decrease the lookup table size, it also has minor errors. So that it has the smaller memory dispenses. Broadly, the proposed algorithm improves the approximation errors without multiply calculus for quantization/re-quantization but have the higher audio quality.

REFERENCES

[1] A.M. Kondoz, "Digital Speech: Coding for Low Bit Rate Communication Systems", 2nd Ed., Wiley, 2004.

[2] R. Yu, T. Li, and S. Rahardja: "Sound Analysis using MPEG Compressed Audio*," Proceedings of IEEE International Conference Acoustic, Speech, and Signal Processing (ICASSP2000), pp. 761–764,* 2005.

[3] Matthew A.Watson and Peter Buettner*,"Design and implementation of AAC decoders,"IEEE Transactions on Consumer Electronics, Vol. 46, Issue:3, pp.819-824,* Aug. 2000.

[4] Purnhagen H. and Meine N., "*HILN-The MPEG-4 Parametric Audio Coding Tools", Proc. IEEE ISCAS 2001, pp.201-204.* May 2001.

[5] P. Srinivasan and L. H. Jamieson, *"High-Quality Audio Compression Using an Adaptive Wavelet Packet Decomposition and Psychoacoustic Modeling," IEEE Trans. on Signal Processing, vol. 46, no. 4, pp. 1085-1093,* April 2008.

[6] J. Li. Embedded audio coding (EAC) with implicit psychoacoustic masking. ACM Multimedia 2002, pages 592–601, December 2002.

[7] G. Tzanetak and P. Cook: is Perceptually enhanced bit-plane coding for scalable audio. Int. Conf. Multimedia & Expo (ICME), (July):1153–1156, 2006.