MULTI-CHANNEL SUPPORT FOR MP3

Pooja Mustapure
VLSI design and embedded systems
VTU Extension Centre,UTL Technologies Ltd.
Bangalore,India

***Abstract*-This paper presents a novel extension of the popular MP3 compression format which extends current MP3 capabilities towards the efficient and compatible representation of multi-channel audio, including the widely used 5.1 sound configuration. As previous 5.1 sound i.e. six channels are first down mixed i.e. they are converted into two channels and then given into MP3 encoder .And at the decoder side the two channels are retrieved then they are up-mixed to produce multi-channel effect sound. In this paper the six channels are directly given to the MP3 encoder in specific format. The input to the encoder will not be in stereo but it is in four channels like. After encoding ,the MP3 bitstream is in multi-channel format which will be a compressed one. Then they are retrieved back at decoder in the same manner. Here there is no necessary of up-mix or down mix. This enhances the sound quality of the main/centre channel by which we can hear the good surround sound of all six channel i.e. efficient implementation of 5.1 CH configurations can be done.**

Index terms- MP3, 5.1 CH

I INTRODUCTION

With the broad availability of the Internet and modern computer technology, a palette of perceptual audio coding schemes has found widespread use in multimedia applications. Among these codec's, the popular MP3 compression format is one of the most frequently used coding schemes.

Formally, the name MP3 refers to the Layer 3 coding scheme of the ISO/MPEG-1 and ISO/MPEG-2 Audio specifications [1] [2] that were finalized in 1992 and 1994, respectively. Compliant to these standards, MP3 capability usually supports the encoding/decoding of mono or stereo audio at a number of common sampling rates (16, 22.05, 24, 32, 44.1, 48kHz) and bit-rates up to 320 Kbit/s. Thus, the name MP3 has been synonymous to stereo (non-multichannel) audio storage and transmission for a period of more than 10 years.

More recently in the latest study, however, multi-channel sound reproductions have become more popular. Multi-channel coding methods such as MPEG surround [13]-[14] or binaural cue coding (BCC) [5] can serve as a method for the effective representation of multi-channel audio signals with backward-compatibility. Binaural Cue Coding (BCC)[4]-[5] a coding method for spatial audio, which forms part of the technological basis of the MP3 Surround format. Next, how the traditional BCC[11] approach is extended to permit the expansion of stereo signals into a multi-channel sound image.

This paper introduces the concept behind a new format that extends current MP3 capabilities towards the efficient representation of multichannel audio. Based on recent advances in multi-channel audio coding technology, it permits the representation of 5.1 sound. Here a complete compressed multi-channel MP3 format is available. This improves the sound quality in the better way.

The rest of this paper is organized as follows. In Section II ,presents an basic scheme of the multi-channel MP3 encoding and decoding process. In Section III, presents our proposed methodology. Then the experimental results in Section IV. Finally, we draw some conclusions in section V and point to possible directions for future works in Section VI.

II BASIC SCHEME

Fig 1 illustrates the general structure of a multi-channel MP3 encoder for the case of encoding a 6/6 multi-channel signal (L, R, C, Ls, Rs, LFE).It consists of parametric multichannel encoder and the MP3 encoder. As a first step, a main/center channel will be a first channel and then a two-channel compatible pairing is done (L, R) is generated from the input multichannel material by a suitable means which will be a second channel. And then again a compatible pairing is done (Ls, Rs) which will be a third channel. At last the low frequency effect channel is a fourth channel. In this way a input to the MP3 encoder will be of four channels which will be encoded by a conventional

MP3 encoder in a fully standards compliant way. These parameters are encoded and embedded as surround enhancement data into the data field of the MP3 bitstream. Then the data is retrieved back with the same format as at the encoder side.

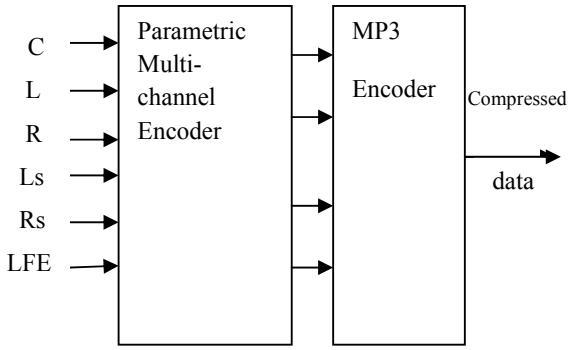

Fig 1: General structure of multi- channel MP3 encoder.

Fig 2 shows the decoder side of the transmission chain. The MP3 Surround bitstream is decoded into a compatible four channel as same as at the encoder side, that is ready for presentation over a conventional six channel reproduction setup (speakers or headphones).

While the preceding the encoding/decoding of a six channel audio signal, other multi-channel configurations can be supported in the same way with this approach.

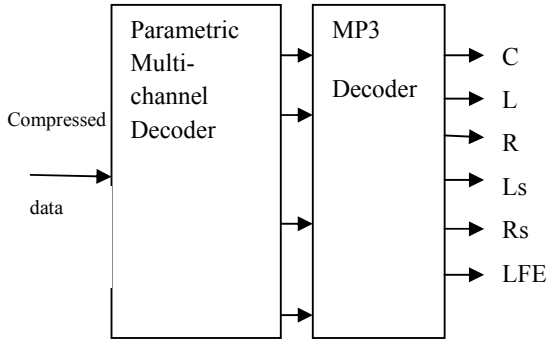

Fig 2: General structure of multi- channel MP3 decoder.

III PROPOSED METHODOLOGY

The input which is given to the proposed design is a uncompressed multi-channel audio file which consists of six channels(C,L,R,Ls,Rs,LFE). Here the arrangement of all six channels is different from stereo channel representation. In the data field ,each sample consists of all six channels. Fig 3 depicts the arrangement of all six channels in each sample. It is given to the parametric multichannel encoder. Figure 4 depicts the flowchart of the parametric multichannel encoder/decoder. Then it is encoded by the standard MP3 encoder to produce a compressed bitstream. At the decoder side it is retrieved back in the same manner as at the encoder side. In this way the proposed methodology works for multichannel MP3 sound.

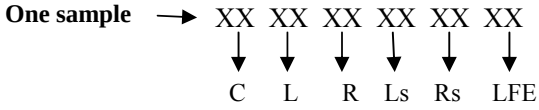

Fig 3: Arrangement of all six channels in each sample.

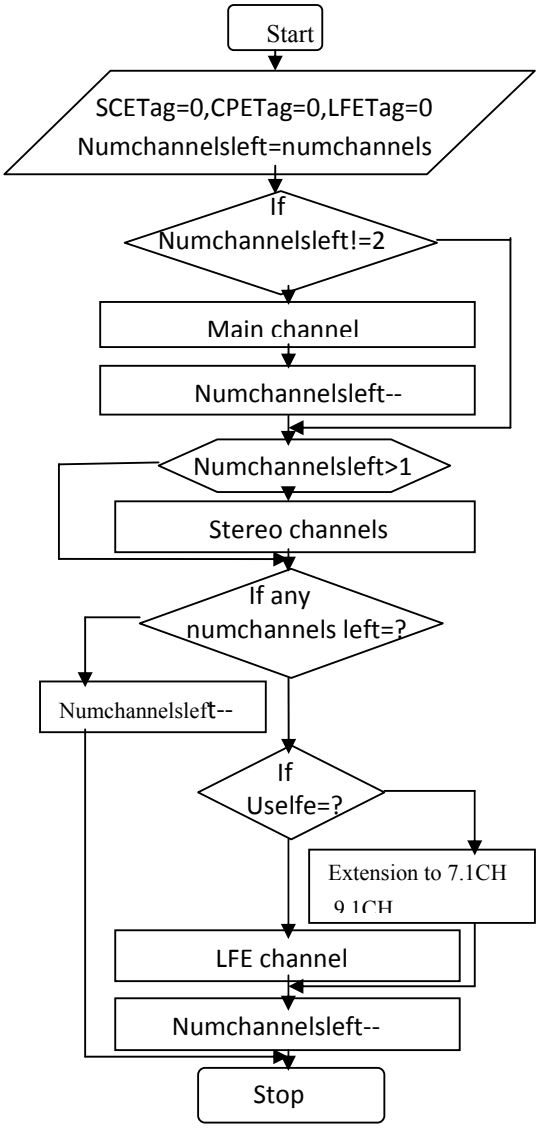

Fig 4: Flowchart of multi-channel encoder/decoder.

A. Advantages of multi-channel over stereo Representation

Downmixing: The most common approach to obtaining a stereo version of a multi-channel signal is called downmixing and involves a linear combination of certain multichannel signals to obtain the desired stereo signals. When downmixing multi-track/multichannel sound material into a stereophonic representation, a number of considerations come into play which are motivated by both psychoacoustics and production practices. On one hand, it is desired to present all parts of the multi-channel sound image also to the listener of a stereo reproduction setup. On the other hand, it is known that – by collapsing front and back channels into the front-only stereo reproduction – the listener's ability to separate the sound components diminishes due to the lack of spatial separation between front and back sound sources. Consequently, sound sources from back channels are usually attenuated within a stereo mixdown .Hence there is need for guarantee good audibility of surround sound sources. Therefore by the proposed design there is a guaranteed good audibility of all channels without attenuation.

Main channel Quality: By the proposed design there is a maintenance of better sound quality of the main channel without any attenuation i.e. the main channel quality is maintained as the original one.

lV RESULTS

The result is shown in a cool edit pro v2.1 software where the input and the output shows the presence of five channels with the sub-woofer (LFE = "Low Frequency Enhancement") channel, as it is used frequently for the representation of movie sound (5.1 configuration).Fig 5 depicts the snapshot of multi-channel at input and fig 6 depicts the snapshot of multi-channel at output side.

A. Performance Analysis: The performance analysis can be done through the RMS calculation. The formula is given as per the ISO/IEC standard.

$$RMS = \sqrt{\frac{1}{N}\sum_{k=1}^{N} diff(k)^2} \qquad (1)$$

diff (k) =(Output signal of the parametric decoder) – (input signal of the parametric encoder). (2)

If RMS is less than 1/ (2^{15}/ √ 12) and if the maximum absolute value is less than or equal to 1/ 2^{14} , then the decoder under test can be called an ISO/IEC 11172-3 audio decoder. If RMS is less than 1/ (2^{11}/√12) , then the decoder under test can be called a limited accuracy ISO/IEC 11172-3 audio decoder. RMS of the proposed decoder is 1.88×10^{-6} .

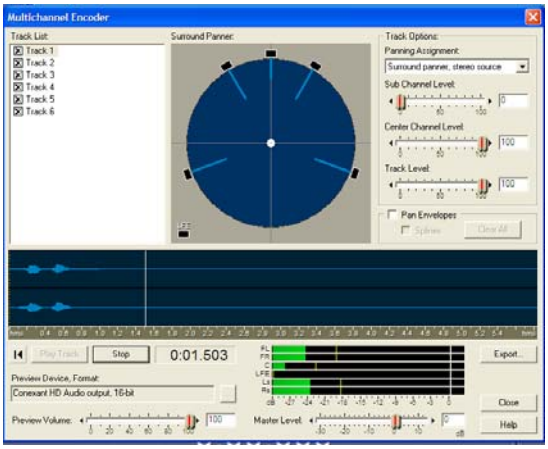

Fig 5:Snapshot of multi-channels at input

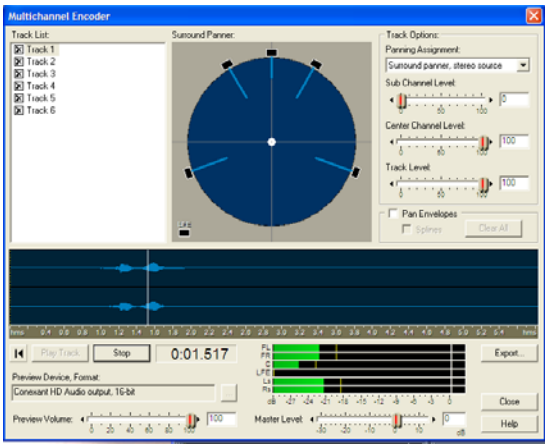

Fig 6:Snapshot of multi-channels at output

V. CONCLUSIONS

This paper introduces an extension of the popular MP3 audio coding format towards the efficient representation of multi-channel audio signals, most prominently the 5.1 channel configurations. This is achieved by parametric multi-channel encoder/decoder that serves as a pre/post processor to the MP3 encoder/decoder chain. An important feature of the resulting format is the quality of the main channel is maintained as the original and there is no attenuation of back surround channels.

Contrary to earlier approaches to parametric representation of stereo and multi-channel audio, the proposed algorithm transmits four signal channels rather than only a two/single channel.

This is a significant and important extension of the general technical paradigm because it enables the better sound quality. The MP3 Surround scheme and the underlying general ideas pave the way for using multi-channel sound for a number of attractive applications. Results of tests indicate that the proposed scheme provides an excellent sound quality which is significantly better than that of other surround formats.

VI FUTURE WORK

The resulting format does not support for the backward compatibility to the existing MP3decoders.The future work will be the proposed format should also support for stereo MP3 decoders.

REFERENCES

[1] ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 11172-3 "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s", 1993

[2] ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 13818-3 "Generic Coding of Moving Pictures and Associated Audio: Audio", 1994

[3] C. Faller, F. Baumgarte: "Efficient Representation of Spatial Audio Using Perceptual Parametrization", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York 2001

[4] C. Faller and F. Baumgarte, "Binaural Cue Coding: A novel and efficient representation of spatial audio," Proc. ICASSP 2002, Orlando, Florida, May 2002

[5] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003

[6] C. Faller: "Parametric Coding of Spatial Audio", 7th International Conference on Audio Effects (DAFX-04), Naples, Italy, October 2004

[7] F. Baumgarte, C. Faller, P. Kroon: "Audio Coder Enhancement using Scalable Binaural Cue Coding with Equalized Mixing," 116th AES Convention, Berlin 2004

[8] ITU-R Recommendation BS.775-1, "Multi-channel Stereophonic Sound System with or without Accompanying Picture", International Telecommunications Union, Geneva, Switzerland, 1992-1994

[9] S. K. Zielinski, F. Rumsey: "Effects of Down-Mix Algorithms on Quality of Surround Sound", Journal of the AES, pp. 790, September 2003

[10] D. Griesinger: "Surround from stereo", Workshop #12, 115th AES Convention, New York, 2003

[11] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications," IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003

[12] Dolby Publication, Roger Dressler: "Dolby Surround Prologic Decoder – Principles of Operation" ,http://www.dolby.com/tech/whtppr.html

[13] S. Quackenbush and J. Herre, "MPEG surround," *IEEE Multimedia*, vol. 12, no. 4, pp. 18-23, Oct. 2005.

[14] J. Herre, K. Kjorling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J.Koppens, J.Hilpert, J. Roden, W. Oomen, K. Linzmeier, and K. S. Chong,"MPEG Surround—The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding," J. Audio Eng. Soc., vol. 56,no. 11, pp. 932-955, 2008.

[15] Han-Gil Moon," A low complexity design for an MP3 multi-channel audio decoding system ", IEEE Trans. Speech, Audio and language Processing, vol. 20, pp. 314-321, jan 2012.