



MODIFIED IMDCT-DECODER BASED MP3 MULTICHANNEL AUDIO DECODING SYSTEM

Shanmuga Raju.S¹, Karthik.R², Sai Pradeep.K.P³, Varadharajan.E⁴

Assistant Professor, Dept. of ECE, Dr.NGP Institute of Technology, Coimbatore, Tamilnadu, India¹

Assistant Professor, Dept. of ECE, Dr.NGP Institute of Technology, Coimbatore, Tamilnadu, India²

Assistant Professor, Dept. of ECE, Dr.NGP Institute of Technology, Coimbatore, Tamilnadu, India³

Assistant Professor, Dept. of ECE, Dr.NGP Institute of Technology, Coimbatore, Tamilnadu, India⁴

ABSTRACT: *The MP3 multi-channel system is composed of two components, a parametric multi-channel decoder and an MP3 decoder. This paper proposes the Enhancement of the audio system by reducing the delay and increasing the audio efficiency with less consumption. This paper presents an advanced mp3 decoding system to play high quality multichannel audio by consuming low processing power. The proposed method uses effective FFT for the computation part and optimizes the IMDCT (inverse modified discrete cosine transforms). This optimization of IMDCT results in the removal of the delay of 31 samples which existed in the previous existing system. Later this optimized output is fed into FFT for computation and is passed into further blocks like phase compensator, re-ordering and synthesis filter. This proposed method uses PQMF (Pseudo code quadrature mirror filter) for the decoding purpose.*

Keywords: IMDCT, PQMF, phase compensator, synthesis filter, re-ordering

I. INTRODUCTION

The development of mobile multimedia services has changed people's daily lifestyles. It is not difficult to find people who listen to music or watch TV via their mobile devices such as mobile phones, MP3 players, and portable media players. People can have a variety of experiences with their mobile devices. However, there is a demand for something more exciting on mobile, e.g., a 3-D video service or a multi-channel audio service.

In order to cope with the demand, South Korea recently established national standards for multi-channel satellite digital multimedia broadcasting (S-DMB) service and terrestrial digital multimedia broadcasting (T-DMB) service using the parametric multi-channel audio technology. A multi-channel audio service on mobile devices, such as a T-DMB multi-channel audio service is an application that enables people to listen to high quality multi-channel audio with a well-equipped speaker system or a head-related transfer function (HRTF)-based virtual surround system.

Among the many digital music formats, MP3 is the most popular one nowadays. Therefore, multi-channel audio services that guarantee the backward compatibility to stereo MP3 format can have a great influence. On the digital music market, since we can fully support the legacy stereo MP3 devices as well as the new multichannel MP3 devices with one format.

A. Multi-Channel Audio

In order to realize a new multi-channel audio service such as the MP3 multi-channel audio service mentioned in the previous section, multi-channel audio signals must first be given a compact expression.

Parametric multi-channel coding methods such as MPEG surround or binaural cue coding (BCC) can serve as a method for the effective representation of multi-channel audio signals with backward-compatibility. As a result of the new coding concept utilized in the parametric multi-channel technique, a multichannel audio signal can be represented with a down-mixed mono or stereo signal and multi-channel side information which is used to recover a multi-channel audio signal from the down-mixed signal. The concept of the MP3 multi-channel encoding (or decoding) system is to serially connect a parametric multi-channel encoder (or decoder) with an MP3 encoder (or decoder). The parametric multi-channel encoder receives multi-channel audio, and produces down-mixed stereo audio and multi-channel side information. Then, the down-mixed stereo signal is again encoded to an MP3 bit-stream by the MP3 encoder. At the decoder side, the MP3 bit-stream is decoded to the stereo audio by the MP3 decoder.



Then, the parametric multi-channel decoder decodes the stereo audio to multi-channel audio using the multi-channel side information which is passed through joint bit-stream. In this concept, the multi-channel processing component and MP3 processing component are independent. Therefore, the MP3 multi-channel decoder can decode the legacy stereo MP3 bit-stream as well as the MP3 multi-channel bit-stream. However, this structure inevitably consumes considerable computing power.

B. Low-Complexity Design for Mp3 Multi-Channel Audio Decoder

The main idea of the proposed system is to combine the transforms used in the MP3 decoder with those of the parametric multi-channel decoder. The transform modules of both decoders consume considerable computing power. Hence, the optimization of these transform modules is helpful for reducing the total power consumption. The MP3 decoder in uses hybrid filter-banks, which are composed of the synthesis pseudoquadrature mirror filter (PQMFs) and inverse modified discrete cosine transform (IMDCTs).

The synthesis PQMF is conceptually composed of the up-sampler and 32 bandpass filters, where the prototype window and the cosine modulation matrix. A bandpass filtering operation with an up-sampled signal in a PQMF synthesis filter bank can be mathematically expressed as a convolution operation.

In order to reduce decoding complexity, the proposed decoding system performs the convolution and synthesis process on the DFT domain instead of using the synthesis PQMF, and then the multi-channel decoding module receives the convolution-synthesis process result without performing another transform. To apply this scheme, the parametric multi-channel coder should be equipped with DFT as the time–frequency transform.

C. Existing Mp3 Multichannel Decoder Audio System

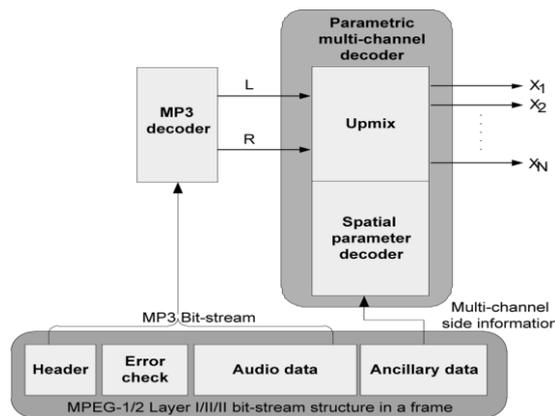


Fig.1 Mp3 decoding system block

Fig1 depicts the block diagram mp3 decoding system which has different specified blocks for early decoding with generating symbols. It has header block which has a job of error control and flow control .Error check in enhance to header do the working of making the arrived information is perceivable or not.

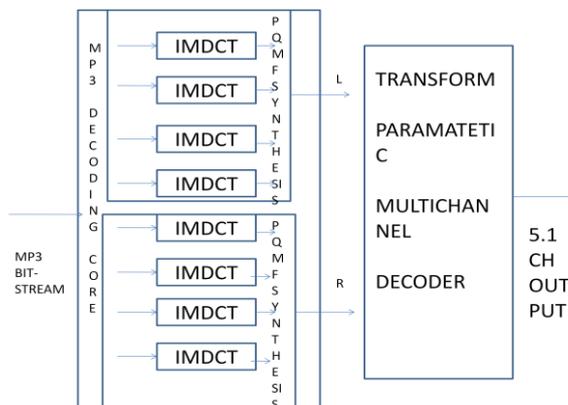


Fig 2 Block diagram of Mp3 decoding system



Here as using PQMF and transforms, output is generated. This is basically the analytical block of mp3 decoding system. The system performances can be analyzed using this system.

D. Drawbacks

The major drawback related to earlier issues is the mismatching of the systems via IMDCT and DFT .Due to that synchronization problem arises as well as the tolerance of the system deviates from the desired function as a value of exponential function.

The main difference arises when the sample number of bits of IMDCT as well as DFT does not match, for instance IMDCT has a sampling rate of 491 samples per bit and DFT has the sampling rate of 512 samples. The net rate difference between these two systems is of 31 samples and that itself causes arrival unwanted signals. Analysis performed on each sector of performance shows the complexity of each design as well as the performance quotient of each section devising the factors of system decoding audio.

II. PROPOSED METHOD

A. Design of Decoder Block

There are already systems existing for audio decoding systems which have improved the audio quality and reduced the power consumption to a large extent. With the help of the proposed system there can be further overcome the limitations in that of the existing system such as the delay and the minute audio drawbacks. Also here multichannel audio processing is possible and can have the same quality without any drawbacks. Here there are changes in the audio decoding system for improvisation.

The main blocks of the proposed systems are Optimized IMDCT, FFT Block, Phase compensator, Re-ordering block, Synthesis filter

B. Decoder Circuit

The Fig4 shows the general block diagram of the proposed system. It basically shows us the Decoder system where the modifications are made. Here we optimize the IMDCT block in the figure 2 due to which the delays are overcome which is the drawback of the existing system and the remaining blocks are followed.

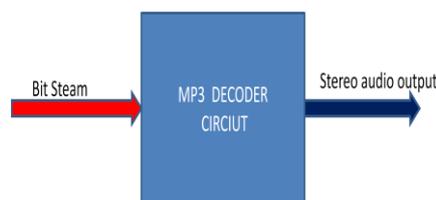


Fig 4: Basic mp3 decoder circuit

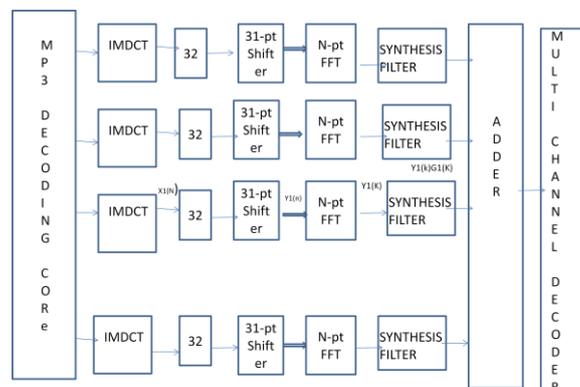


Fig5: Block diagram of the audio decoder



C. Optimization of IMDCT

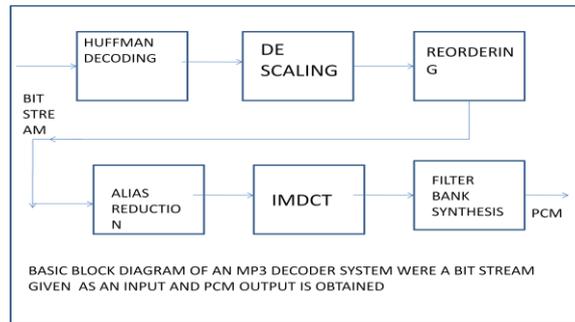


Fig 6: Block Diagram of MP3 Decoder

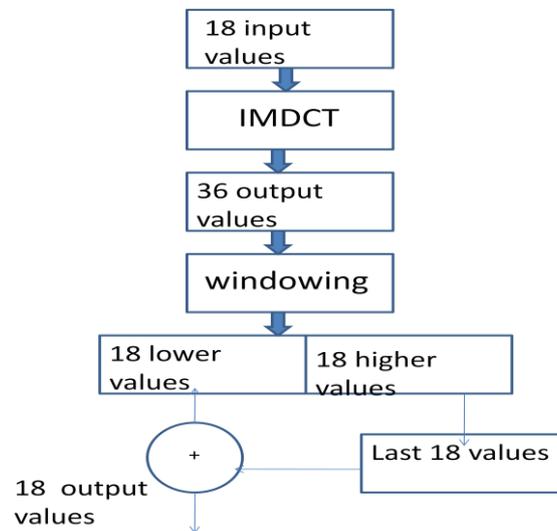


Fig 7: Flow diagram

Optimized IMDCT Equation

$$N_{ik} = \cos \left[\frac{\pi}{2M} (2i + 1)k \right], 0 \leq i \leq 31 \quad 0 \leq k \leq 31$$

The optimized version of IMDCT is given as

$$N_{ik} = \cos \left[\frac{\pi}{2M} (16 + i)(2k + 1) \right], 0 \leq i \leq 63 \quad 0 \leq k \leq 31$$

$$C_k^i = \cos \pi \left(\frac{i}{k} \right)$$

This is optimization of IMDCT matrix.

D. Implementing FFT

The FFT is an algorithm that computes the Discrete Fourier Transform and its inverse. The FFT produces the exact same result as evaluating the DFT directly, but the FFT produces an answer much faster. In general the DFT is found by using the equation

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k n/N}$$

Where $X_0 \dots X_{N-1}$ is complex numbers and $k = 0 \dots N-1$



The FFT is used as a filter bank on an audio sample. It is used to filter out unwanted data from the sample. First, incoming audio samples, $s(n)$, are normalized based the following equation $x(n) = s(n)N(2b-1)$ Where N is the FFT length of the sample and b is the number of bits in the sample.

Second, the masking threshold of the sample is found by using an estimate of the power density spectrum, $P(k)$. $P(k)$ is computed by using a 1024-point FFT.

$$[P(k) = PN + 10 \log [\sum_{n=0}^{N-1} h(n) x(n) \exp(-j 2 \pi kn)]^2, 0 \leq k \leq N-1, h(n) = 0.5(1 - \cos 2 \pi nN^{-1}), 0 \leq i \leq N-1]$$

$h(n)$ is a HannWindow denoted by: PN is the power normalization term, it is usually around 96 decibels.

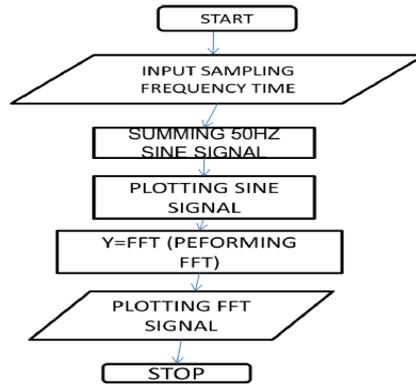


Fig8: Flowchart of FFT

E. PHASE COMPENSATOR

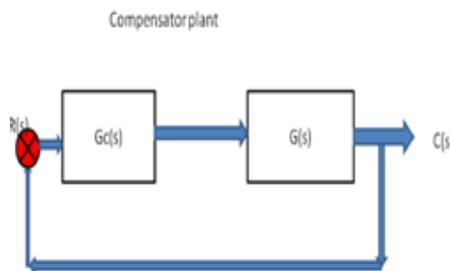


Fig 9 Compensator plant

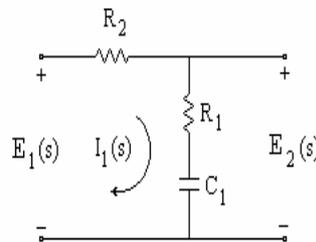


Fig 10: Compensator circuit

The output signal is proportional to the sum of the input signal and its integral. The designation lag applied to this network is based on the steady-state sinusoidal response. The sinusoidal response E_2 with a sinusoidal signal E_1

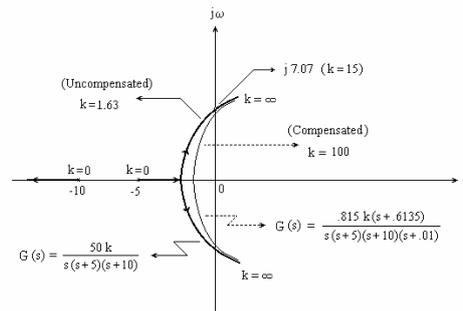


Fig.11: Graph of compensation

F. Design of Encoder Block

In order to design an effective multichannel Mp3 audio decoder, it is necessary to have a compatible multichannel encoder.

The main blocks of the proposed encoder are

1. Filtering and analyzing
2. Computation of FFT
3. Masking and encoding

Fig 10 shows the model for the proposed parametric mp3 multichannel encoder system which converts the mp3 stream into segments called frames.

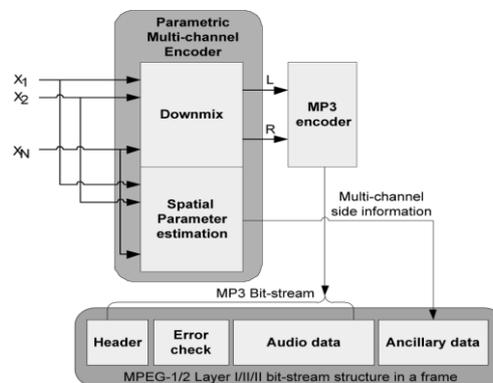


Fig 10: Parametric mp3 multichannel encoder

G. Algorithm Overview

The overall algorithm is broken up into 4 main parts.

Step1. Divides the audio signal into smaller pieces, these are called frames. An MDCT filter is then performed on the output.

Step2. Passes the sample into a 1024-point FFT, and then the psychoacoustic model is applied. Another MDCT filter is performed on the output.

Step3. Quantifies and encodes each sample. This is also known as noise allocation. The noise allocation adjusts itself in order to meet the bit rate and sound masking requirements.

Step4. Formats of the bit stream called an audio frame. An audio frame is made up of 4 parts, The Header, Error Check, Audio Data, and Ancillary Data.



H. Multichannel Encoder

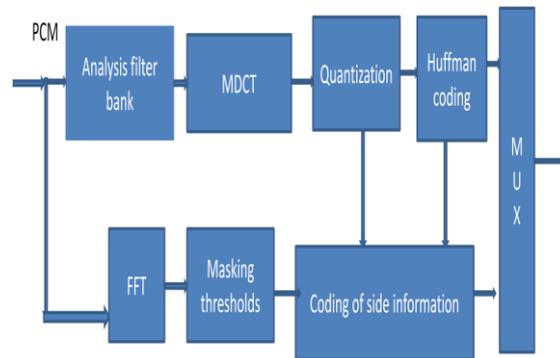


Fig 11 Block diagram of multichannel encoder

The FFT produces the exact same result as evaluating the DFT directly, but the FFT produces an answer much faster. In general the DFT is found by using the equation:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k n/N}$$

Where $X_0 \dots X_{N-1}$ are complex numbers and

$$k = 0 \dots N-1$$

The FFT is used as a filter bank on an audio sample. It is used to filter out unwanted or unneeded data from the sample. First, incoming audio samples, $s(n)$, are normalized based the following equation $x(n) = s(n)/N$. Where N is the FFT. The MP3 encoding algorithm has numerous complex parts. The FFT, DFT, and MDCT play a key role in encoding audio samples.

I. Comparison of Each Module

COMPUTATIONAL COMPLEXITY OF EACH MODULE

MODULE	NUMBER OF MULTIPLICATION	NO. OF ADDITION
N-POINT DFT	$N^2 \cdot \log_2 N$	$2N^2 \cdot \log_2 N$
M-POINT DFT	$M^2 \cdot \log_2 M$	$2M^2 \cdot \log_2 M$
SYNTHESIS FILTER BANK	592	721
PHASE COMPENSATION & RE ORDERING	2M	0
SYNTHESIS MODULE	0	N

Table 1: Comparison of each module

Table 1 shows the number of multiplication and addition operation required for each module. The synthesis filter bank module needs 592 multiplication and 721 additions, which introduces time delay. This can be reduced by implementing pipelining.



PROCESSING TIME COMPARISON OF BOTH DECODERS

INDEX	MATERIAL NAME	CONVENTIONAL METHORD	PROPOSED METHORD
A	BBC Applause	4.472	3.975
B	ARL Applause	4.699	4.254
C	Chostakovitch	4.646	4.167
D	Fountain music	4.616	4.189
E	Glock	4.128	3.774
F	Indire	3.460	3.15
G	Jackson	3.73	3.55
H	Pops	4.595	4.167
I	Poulene	4.917	4.163
J	Rock concert	4.85	4.180
K	Stomp	4.627	4.225
L	average	4.389	3.962

*The decoding time is measured with reference of audio selections

Table 2: Time comparison of decoders

NUMBER OF MODULES AND THE NUMBER OF TOTAL OPERATIONS USED IN EACH METHORD

MODULE	COVENTIONAL METHORD	PROPOSED MEHORD
N-POINT DFT	2	0
M-POINT DFT	0	64
SYNTHESIS FILTER BANK	2*N/32	0
PHASE COMPENSATION& RE ORDERING	0	64
SYNTHESIS MODULE	0	1
TOTAL MULTIPLICATIONS	2N*LOG2N+37N	2N*LOG2M+4N
TOTAL ADDITIONS	4N*LOG2N+721/16N	4N*LOG2M+N

Table 3: No of operations

The delay of 31 samples that existed in the existing system was removed. The computation was made easier using FFT. The audio quality was improved and the complexity was reduced. The power consumption was reduced and the gain was optimized.

III. CONCLUSION

There are nowadays still some drawbacks existing due to which the desired clarity and quality is not obtained in an audio decoding system thus the according to the survey done we find that there were some delay which caused the limitations. Thus the proposed system optimized the blocks and removed the delay and also made the FFT effective so that the computation could be done easily and therefore the audio is well delivered and the desired quality is obtained as in multi-channel audio decoding system.

REFERENCES

- [1]. H. G. Moon, "Backward-compatible terrestrial digital multimedia broadcasting system for multi-channel audio service," *IEEE Tans. Consumer Electron.*, vol. 56, no. 3, pp. 1556–1561, Aug. 2010.
- [2]. H. W. Kim and H. G. Moon, "The low complexity MP3 multichannel audio decoding system," in *Proc. AES 129th Conv.*, San Francisco, CA, Nov. 2010.
- [3]. "Multichannel audio services for satellite digital multimedia broadcasting," *Telecommun. Technol. Assoc., Korea, TTAS.KO-07.0069*, Jun. 2009.
- [4]. J. Hilpert and S. Disch, "The MPEG surround audio coding standard [standards in a nutshell]," *IEEE Signal Process. Mag.*, vol. 26, no.1, pp. 148–152, Jan. 2009.
- [5]. K. Ikeda and R. Sakamoto, "Convergence analyses of stereo acoustic echo cancelers with preprocessing," *IEEE Trans. Signal Process.*, vol.1, no. 5, pp. 1324–1334, May 2003.
- [6]. Y. J. Zhou and S. L. Xie, "Study on stereo acoustic echo canceller with nonlinear pre-processing units," *Dynamics of Continuous Discrete and Impulsive Systems-Series B-Applications&Algorithms*, pp. 66–72, 2003.
- [7]. D. R. Morgan, J. L. Hall, and J. Benesty, "Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 686–696, Sep. 2001.