International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 1 (2019) pp. 14-19

Audio Signal Performance Analysis using Integer MDCT Algorithm

M.Davidson Kamala Dhas [1], R.Priyadharsini [2]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Mepco Schelnk Engineering College, Sivaksi 626005, India.

[2]PG Student, Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi 626005, India.

Abstract

In Audio coding Algorithms Integer Modified Discrete Cosine Transform is mostly used. In this paper, we propose Analysis and synthesis method for perfect reconstruction of audio signal and Integer modified discrete cosine transform is mainly used to overcome Time domain aliasing cancellation problem. This works mainly depends on Integer MDCT, which is an integer approximation of a lapped transform. It inherits the properties of MDCT and it provides critical sampling, perfect reconstruction property. In both the (Analysis/Synthesis) stages Kaiser Bessel window functions are used to ensure perfect reconstruction property. The purpose of this work is to analysis and synthesis IntMDCT algorithm in terms of Lifting Scheme and to ensure its perfect reconstruction without any error. If any noise is generated in analysis stage, it will be removed by means of RLS algorithm then the signal is rewindowed and deframed in it for perfect reconstruction. Performance is measured by Mean square error and Peak Signal to noise ratio.

Keywords-Integer Modified Discrete Cosine Transform (IntMDCT), Lifting Scheme, Least Mean Square, Time Domain Aliasing Cancellation (TDAC), Inverse Integer Modified Discrete Cosine Transform.

I. INTRODUCTION

In Audio coding design and implementation two approaches are followed that is Analysis and synthesis stage, and in Analysis stage TDA is introduced and which is removed in synthesis stage. MDCT provides 50% overlapping and this overlapping filter bank is mainly used to reduce the artifacts that may help to reconstruct the original sampled signal in the receiver side. The aim of this work is to reconstruct the lossless audio signal in Integer Modified Discrete Cosine Transform using Lifting Scheme method. Critical sampling is used to achieve subsampling operation which is performed in the frequency domain therefore the aliasing is consequently canceled in it by an overlap and addition process [2].

The Lifting scheme is reversible and it maps integers to integers and rounding operation is performed. MDCT provides a good decorrelation of audio input signal is compared to number of possible input values, output values gets increases. Hence, data rate quantisation operation in needed, these operation are used to neglect the resulting error. IntMDCT can be used for lossy coding scheme and it is also used as an original floating point transforms [4].

Twiddle factors are used for fast computation it is represented in terms of real and imaginary. By cascading windowing and Time domain aliasing of DCT4 integer Modified Discrete Cosine Transform is achieved. In Integer MDCT any types of window function can be used like Triangular window, sine window, Hamming window, KBD window but in [5] the performance of sine window and KBD window is analyzed and the obtained result as compared to sine window, KBD window will provide perfect reconstruction and also in sine window frame may get lost, recovery is not possible and this type of window is rarely used in Audio coding and also while using sine window efficiency will be very low. INTMDCT transform is to bridge a gap between perceptual and lossless audio coding [6]. According to given MDCT equation that transform 2N real numbers to N real numbers are used to eliminate the blocking artifacts in analysis stage so that in synthesis stage it can able to perform without any artifacts and the signal can able to reconstruct it [7].

II. INTEGER MODIFIED COSINE TRANSFORM

In Integer Modified Discrete based on TDAC transform coding, Analysis/Synthesis stage uses most of the following properties of MDCT, including critical sampling and overlapping structure by Time domain aliasing cancellation. By cascading Integer MDCT with entropy coding scheme a lossless audio coding is builted. Integer MDCT is derived from Modified discrete cosine transform and it is performed by means of rounding operation or lifting steps.

In Integer MDCT sparse signal is used efficiently for reconstructing a signal by obtaining a solution for undetermined linear system. In audio signals input audio frames are transformed losslessly using Integer MDCT coefficient that is scaled, coded, and quantized and the MDCT ensures smooth transition between frames and good signal reconstruction.

Overlap Addition method is a fundamental technique in audio coding it decompose the sampled audio signal into simple components and recombine the processed components into output signal and it describes the nonuniform orthogonal MDCT filter bank &Time domain aliasing reduction and significantly to reduce aliasing, quantization noise is introduced. Integer MDCT is the type of filter bank, were

Integer MDCT performance is analysed in a lossy scenario it is found that noise is introduced but it doesn't affect the perceptual audio signal.

III. DEFINITION

❖ *Rounding operation or Lifting scheme:*

Integer MDCT is introduced in terms of Lifting scheme and it is decomposed into given rotations, that is planar rotation by an given angle is denoted as,

$$\begin{bmatrix} cos \propto & -sin \propto \\ sin \propto & cos \propto \end{bmatrix} \quad (1)$$

❖ *Reason for lifting scheme:*

Lifting scheme used to implement reversible integer transform and it involves floating point operation, which introduces rounding error based upon floating point arithmetic [10]. For lossless compression in lifting scheme perfect reconstruction is possible and it is easy to understand and implement. The inversion can't be able to done within the limited precision of coefficients and input values, generally it can able to solve by lifting scheme and ladder network. In the context of wavelet transform, decomposition of 2*2 matrices is used. The building blocks of lifting scheme are referred as ''*lifting steps*'' with a real value of 'a' called as lifting coefficients.it is shown in equation (2) & (3)

$$L_{a=}\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \quad (2)$$

The Transpose of the above matrix (2),

$$L_{a=}\begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix}$$

The lifting step L_a maps two values (x_1, x_2) that is,

$$L_a(x_1, x_2) = x_1, x_2 + ax_1 \quad (3)$$

By taking inverse lifting steps is given by,

$$L_a^{-1} = L_a = \begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix} \quad (4)$$

A rounding function [.] is an integer approximation of a given rotation, where R->Z can be included in each lifting steps by mapping from integers to integers in a reversible manner.

$$L_{a,[.]}\,(x_1, x_2) = (x_1, x_2 + [ax_1]) \quad (5)$$

The inverse lifting steps is given as follows,

$$L_{a,[.]}^{-1} = \left(x_{1,} x_2 + [ax_1]\right) \quad (6)$$

In a context wavelet transform 2*2 decomposition is used in lifting steps, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with b ≠ 0 and determinant 1 is given in terms of,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{d-1}{b} & 1 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{a-1}{b} & 1 \end{bmatrix} \quad (7)$$

Three basic rotations along three coordinates along X, Y, Z and the rotation angle should be positive that is counter clockwise and anticlockwise multiplies of 90 degree. The equation shown below along three coordinates

$$\begin{bmatrix} cos \propto & -sin \propto \\ sin \propto & cos \propto \end{bmatrix} = [X][Y][Z] \quad (8)$$

$$[X] = \begin{bmatrix} 1 & \frac{cos\alpha - 1}{sin\alpha} \\ 0 & 1 \end{bmatrix} \quad (9)$$

$$[Y] = \begin{bmatrix} 1 & 0 \\ sin\alpha & 1 \end{bmatrix} \quad (10)$$

$$[z] \quad = \begin{bmatrix} 1 & \frac{cos\alpha - 1}{sin\alpha} \\ 0 & 1 \end{bmatrix} \quad (11)$$

❖ *Forward Transform & Inverse Transform*

The Integer MDCT is related to MDCT and based upon three lifting steps rounding operation is implemented by rotations. For Integer Modified discrete cosine transform the butterflies have to be implemented with an angle $\pi/4$ to ensure the critical sampling, and for energy preservation, and the input signal level depends upon frequency selectivity.

The Forward/Lapped Transform is given by equation,

$$X_K = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} x_n \cos\frac{\pi}{4N}\left(n + \frac{1}{2} + N\right)\left(K + \frac{1}{2}\right)$$

$$K = 0,1, \ldots \frac{N}{2} - 1 \quad (12)$$

Multiplying equation (9) & (12) we get, equation (13)

$$X_K = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} x_n \cos\frac{\pi}{2N}\left(n + \frac{1}{2} + N\right)\left(K + \frac{1}{2}\right)$$

$$K = 0,1, \ldots \frac{N}{2} - 1 \quad (13)$$

Multiplying equation (10) & (13) we get, the matrix and rotation are performed by π values, the obtained results in equation (14)

$$X_K = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} x_n \cos\frac{\pi}{N}\left(n + \frac{1}{2} + N\right)\left(K + \frac{1}{2}\right)$$

$$K = 0,1, \ldots \frac{N}{2} - 1 \quad (14)$$

By multiplying (11) with eqn (14) the rotation and the matrix is performed by π values, results in (15)

$$X_K = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} x_n \cos\frac{2\pi}{N}\left(n + \frac{1}{2} + N\right)\left(K + \frac{1}{2}\right)$$

$$K = 0,1, \ldots \frac{N}{2} - 1 \quad (15)$$

Integer Modified Discrete Cosine Transform depends upon lifting scheme which is performed on MDCT. The integer MDCT can be inverted without introducing any error by means of lifting scheme or rounding operations. Using this lifting scheme we can able to get integer approximation of the given rotations. In the rounding operation r is odd symmetric then the inverse rounded rotation with an angle (-θ) is given as,

$$\begin{bmatrix} cos \propto & sin \propto \\ -sin \propto & cos \propto \end{bmatrix} \quad (16)$$

The *inverse transform* is given by,

$$Y_n = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} x_k \cos \frac{2\pi}{N} (2n + 1 + N)(K + 1)$$

$$n = 0,1, \ldots N - 1 \quad (17)$$

❖ *Kaiser Bessel Window*

KBD Window is mostly used for perfect Reconstruction because by using adjustable parameter β the width of the main lobe and side lobe width can be varied.

Window function W (n) is given by,

$$W(n) = \sum_{n=0}^{2M} \frac{I_0(\pi\alpha)\sqrt{1-\left(\frac{2n}{M-1}\right)^2}}{I_0(\pi\alpha)} \quad (18)$$

$$\beta = \pi\alpha \quad (19)$$

$$I_{0(X)} = \sum_{K=0}^{\infty} \frac{\left(\frac{x}{2}\right)^k}{k!} \quad (20)$$

$$I_0 \rightarrow Zeroth\ order\ Bessel\ Function$$

By using adjustable parameter α, the main lobe and the side lobe width can be varied. Kaiser Bessel window is similar to Kaiser Bessel Derived window but it ranges from 0 to 2M.

IV. PROPOSED METHOD

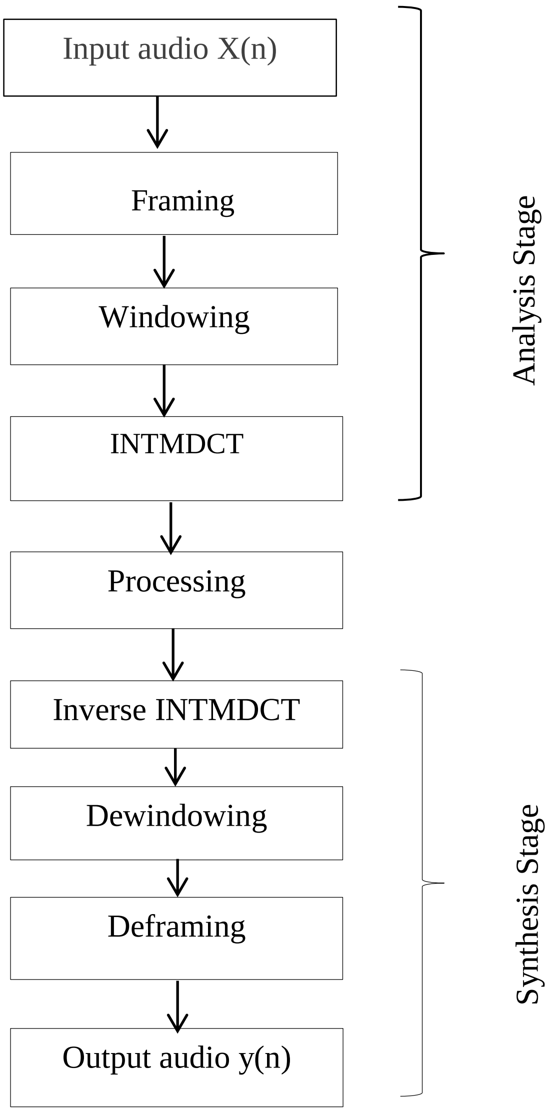

Fig.1 Block diagram of proposed method

In proposed method input audio signal is critically sampled then the signal is passes through the Analysis stage and then the signal is framed and windowed and passes through integer MDCT Algorithm [15] some aliasing may occur that can be processed using RLS filter then the aliasing can be cancelled out in synthesis stage by lifting scheme, therefore this approach ensures perfect reconstruction [6]. To subdivide the large signal to small signal Kaiser Bessel Window function is used. In Synthesis stage signal is passed through the Inverse Integer MDCT algorithm, then the signal is Rewindowed and deframed using Kaiser Bessel derived window. Here KBD window is used because it is an adjustable window and it can able to do perfect reconstruction by means of lifting scheme method. The performance is measured using Mean square error (MSE) and Peak Signal to noise ratio (PSNR).

V. RESULTS AND DISCUSSIONS

The performance measures are MSE and PSNR. MDCT and Integer MDCT performances are measured by using KBD window function. The input audio signal is in .wav format. The squared error of the original and reconstructed audio signal yields Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR).With these values the reconstructed audio signal can be predicted. Below Fig.2 shows the input audio signal of 44.1 KHz of sampling rate [12].

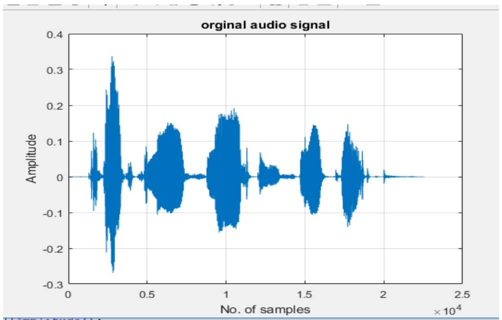

Fig.2. Input Audio Signal with the sampling rate of 44.1 KHz which is in .Wav format

In Analysis stage next step is framing, the audio signal with 50% overlapping that is shown in Fig (3).Then the next step is to multiply framed audio signal and windowed audio signal with 256 number of samples that is shown in Fig(4)

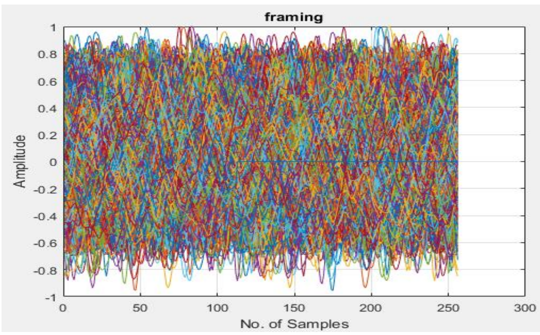

Fig.3.Framed audio signal with 256 samples per frames

Here after framing, KBD window is used for windowing technique and the waveform of Kaiser Bessel window shown below,

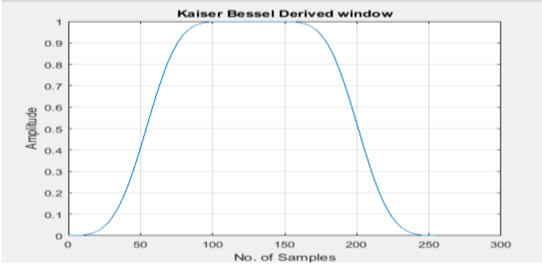

Fig.4. KBD Window function

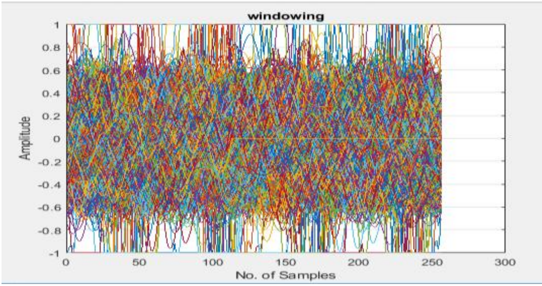

Fig.5. Windowed Audio signal

After framing the windowed audio signal is passed through the Integer MDCT Algorithm for 256 numbers of samples are shown in below fig (5)

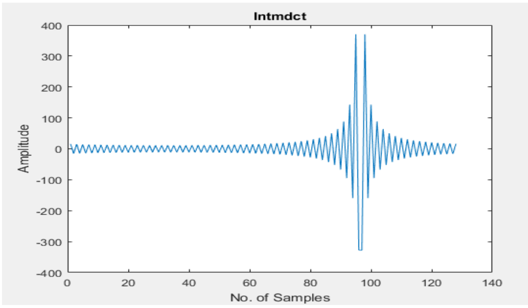

Fig.6. Integer MDCT Transformed signal

In next step the analyzed audio signals are processed by adding a random noise and the aliasing is removed by means of filtering. Various types of filtering are there to remove noise. There are many adaptive algorithms like Least Mean Square algorithm (LMS), Kalman filter, but the most commonly used is the recursive least square (RLS).

RLS ALGORITHM:

The Recursive Least Square (RLS) is an Adaptive filter, which is used to find the coefficients recursively and its main aim is to reduce the Mean Square Error.

- In next step, the processed signal is passed through the Inverse Integer MDCT transform and then the output of the audio signal is rewindowed and deframed are shown in below figure,

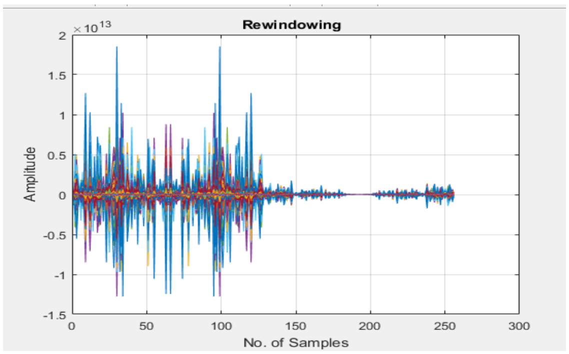

Fig.8. Rewindowed audio signal

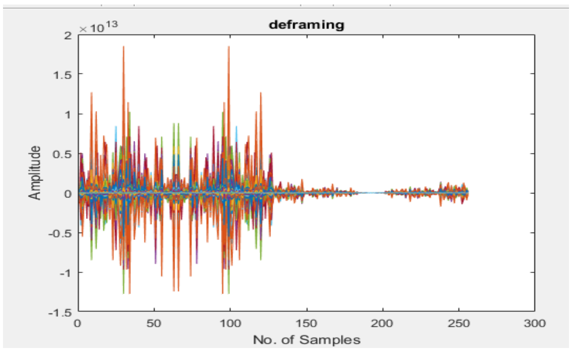

Fig. 9. Deframed audio signal

In synthesis stage after rewindowing and deframing the audio signal, reconstructed signal is obtained are shown in fig (10)

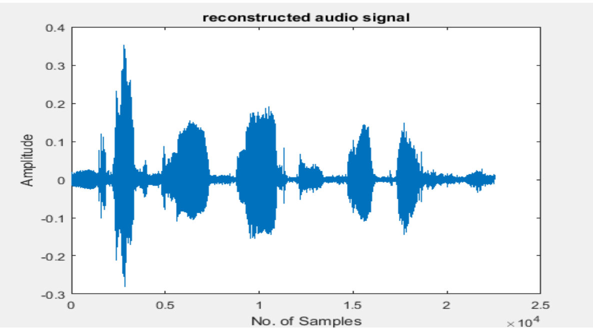

Fig.10.Reconstructed Audio signal

VI. PERFORMANCE MEASURE

To measure the average of the square of errors between original signal and the reconstructed signal Mean Square Error is predicted, and also the original and the reconstructed audio signal is analysed for different performance measure like PSNR (Peak Signal to Noise Ratio).that is shown in eqn (21)&(23)

$$MSE = \frac{1}{N}\sum_{i=0}^{N} {e_i}^2 \qquad (21)$$

$$e_i = x_i - y_i \qquad (22)$$

Table 1. Comparison of MDCT and INTMDCT is measured based on Performance Measures like PSNR (Peak Signal to Noise Ratio) and the Mean Square Error

Performance Measure	MSE	PSNR
MDCT	0.3777	**0.00788**
INTMDCT	71.2310	**75.6184**

By different duration of audio signal, the performance is measured between MDCT and IntMDCT. Hence, Integer MDCT shows less BER compared to MDCT are shown in below table.2

Table 2. Comparision of MDCT and IntMDCT based upon different duration.

Duration	*Performance Measures*	*MDCT*	*INTMDCT*
20ms	MSE	0.2767	**0.00388**
	PSNR	53.231	**59.6184**
40ms	MSE	0.3854	**0.00424**
	PSNR	63.235	**68.2345**

VII. CONCLUSION

From the above table, we have inferred that Integer MDCT is better compared to MDCT, even for different duration of audio signal. Where IMDCT is a lapped transform and it can be used for audio coding and audio processing applications, and through the integer MDCT algorithm error is reduced than that of MDCT. .Hence, by using KBD window the audio signal ensures perfect reconstruction property.

REFERENCES

[1] Yoshikazu Yokotani, Ralf Geiger and Gerald D. T. Schulle., 2006, "Lossless Audio Coding Using the IntMDCT and Rounding Error Shaping," IEEE transactions on audio, speech, and language processing, vol. 14(6).

[2] J.Princen and Bradely.,1986, "Analysis/synthesis Filter bank Design Based on Time Domain Aliasing Cancellation ", IEEE ICASSP on volume 34(5), pp, 1153-1161

[3] Yoshikazu Yokotani and Soontorn Oraintara.,2003,"lossless audio compression using integer modified discrete cosine transform", International Symposium on Intelligent Signal Processing and Communication.

[4] Shuhua Zhang and Laurent Girin.,2013" Fast and accurate direct MDCT to DFT conversion", IEEE transactions on audio, speech and language processing, institute of electrical and electronics engineers.

[5] Shuhua Zhang, Weibei Dou, Member, IEEE, and Huazhong Yang.,2013, “mdct sinusoidal analysis for audio signals analysis and processing”, IEEE transactions on audio, speech, and language processing, vol. 21(7).

[6] Te Li, Student Member, IEEE, Susanto Rahardja, Senior Member, IEEE, Rongshan Yu, Senior Member, IEEE,and Soo Ngee Koh, Senior Member, IEEE.,2017” on integer mdct for perceptual audio coding”, ieee transactions on audio, speech, and language processing, vol. 15(8).

[7] HUANG HaiBin, RAHARDJA Susanto, YU Rongshan, and LIN Xiao,” A Fast Algorithm of Integer MDCT for Lossless Audio Coding”

[8] Dong-Yan Huang and Ruihua Ma,” intmdct -a link between perceptual and lossless audio coding”, Fraunhof AEMTImenau, German2FraunhofnS-A, Erlangen, German

[9] Computation of Forward and Inverse MDCT Using Clenshaw’s Recurrence Formula

[10] Jin Li, “Reversible FFT And MDCT Via Matrix Lifting”, Microsoft Research, Communication, Collaboration and Signal Processing, One Microsoft Way, Bld. 113, Redmond, WA, 98052.

[11] MDCT PROPERTIES AND APPLICATIONs

[12] Joebert S. Jacaba, “Audio compression using modified discrete cosine transform: the mp3 coding standard” by October 2001[13] Dong-Yan Huang and Ruihua Ma, “Integer Fast Modified Cosine Transform” Institute for Infocomm Research Signal Processing Program/Video Processing Group 21 Kent Ridge Terrace Singapore 119613.

[13] Emmanuel Ravelli, Ga¨el Richard and Laurent Daudet, “Audio signal representations for indexing in the transform domain”.

[14] Mark Vinton and Grand Davidson.,2009 “Adaptive hybrid transform for signal analysis and synthesis” .

[15] Emmanuel Ravelli, Gaël Richard and Laurent Daudet.,2004 “Union of MDCT Bases for Audio Coding”.