Improved Slant and Walsh-Hadamard transforms for audio signals using watermarking as application with comparison to DCT, DWT, FWHT, LWT and Wav-Decomposition

Asst.Prof.Dr. Hala B. Abdul Wahab[1]
Asst.Prof.Dr. Abdul-Mohssen J. Abdul-Hossen[1]
Sana Ahmed Kadhom[2]

[1]Department of Computers, University of Technology, Baghdad, Iraq
[2]Department of Computer Science, Al Mammon University College, Baghdad, Iraq

Abstract

Audio signals are transmitted and recorded in time (or spatial) domain, transformed to frequency domain when compressed, watermarked or encrypted. The transformation is done by using some form of transform like Discrete cosine transform(DCT), Fast Fourier transform(FFT), Discrete wavelet transform(DWT), Fast Walsh Hadamard transform(FWHT) and many others. Slant transform(SLT) was mainly produced and used for images; it was proved that Slant had better performance measurements than Walsh-Hadamard transform, although they have many common characteristics, SLT has fast and simple computations ; In this paper a modified Slant transform that has the ability to work efficiently with audio files and modified WHT were introduced. The modified and other transforms were used especially in audio watermarking as one of the most important applications on both ‘transforms and audio’. The new improvements were tested on many audio files and compared to other transformers using SNR, BER, Time and other metrics.

Keywords: Audio, Slant transform, Walsh Hadamard transform, watermark, DCT, DWT, LWT, Wave decomposition.

تطوير محولات Slant,Walsh-Hadamard للملفات الصوتية واستخدامها في تطبيقات العلامة المائية، مع مقارنة بين الانواع المطورة وانواع اخرى مثل DCT,DWT,HWHT,LWT,wav-dec

الخلاصة

يتم تسجيل الصوت في المجال الزمني ويتم تحويله الى المجال الترددي في حالات معينة كالضغط والاخفاء والترميز باستخدام المحولات. وهناك انوع محولات عديدة معروفة ومستخدمة. يهدف البحث الى تطوير نوعين من المحولات ومقارنتها مع الانواع المعروفة من عدة نواحي وتطبيقها على اخفاء علامة مائية داخل الملف الصوتي

Introduction

Due to the explosion of internet and multimedia usage; security became a major part of technical evolution; cryptography, watermarking and other types of protection became a must. Watermarking techniques had been developed rapidly within the last years; by watermarking the media, a special data is hidden within the media without affecting the quality and clarity of the original data. Media like audio has special characteristics because of its fragile, sensitive, and complex nature. Watermarking(WM) techniques used for audio files have to deal with some important aspects like robustness, perceptual transparency, WM bit rate, security, computational complexity and cost[1]. Watermarking audio signals in time domain may not only affect the quality of the audio signal but also the watermarked signal will be affected by any signal processing techniques like amplitude compression, resampling and others, so transform methods were used to transform audio files from time domain to frequency domain such that any addition or modification on the transformed bits will be distributed on a large spectrum of samples (Spread Spectrum SS), by this the watermark bits have no effect (or tiny) on audio signals[2]. In the last years a great effort was made to modify or find new methods of transforms. In this paper we produced a modified Slantlet transform which was originally created for images, the matrices were changed in a way that is suitable for audio and gives better performance measurements. Another transform which is used frequently is Walsh Hadamard transform; it also depends on some matrices which we changed to get better results than the original once. Many other transforms like Discrete cosine transform DCT, Discrete wavelet transform DWT, Lift wavelet transform LWT, Fast Walsh Hadamard transform FWHT, and Wave decomposition and reconstruction, all these types were applied on the same file to be compared with the modified once. All transforms were used in watermarking audio files and tested with some criteria like BER, SNR, and CPU time.

I- Transforms definitions

Transforms have been used with images for long time, but for the last decade they had great effect on audio files' security and communications. Many types of transforms were used frequently:

DCT: has been generally utilized later as a part of signal compacting calculations because of its extensive capacity in packing the signal vitality even in few coefficients. In watermarking applications, DCT inserts the watermark bits into the coefficients got from the change utilizing the quantization methods. This methodology has a high SNR since the computerized watermark bits are embedded in the high vitality segments of the host sound signal producing clear and high quality watermarked sound. In addition, DCT is strong against resampling and low pass filters attack. Notwithstanding, it is defenseless against compression attack like mp3.[3]

DWT: is used in an extensive variety of digital signal processing(DSP) applications including image, audio, and video compression, information exchange over the Internet, numerical applications and pattern recognition. DWT transform can successfully represent signals particularly those have limited varieties. [4] DWT works with Haar and Daubechies family (db1,db2,…).

Haar: has two output sets (averages and differences)

$$\mathbf{X_i + X_{i-1} /2} \quad \text{.....(1)}$$

$$\mathbf{X_i - X_{i-1} /2} \quad \text{......(2)}$$

Daubechies: it works the same as Haar in finding the differences and averages, but it differs in defining the wavelets and the scaling signals. Dubechies use an overlapping window to reflect all the changes in the high frequency spectrum, that's why Daubechies were used in compressing and removing noise from audio signals[5].

Wave decomposition transform: apply a multilevel wave analysis using specific types of filters (Haar, db1,db2,sys2,….); The output of the decomposition are two coefficient vectors.

The inverse is the wave reconstruction, which rebuild the original signal from the two coefficients vectors[6].

WHT: it is a generalized form of Fourier transform; it has symmetric, linear, and orthogonal operation on complex or real numbers. [7]
The Hadamard matrix is defined as:

$$\mathbf{L}=\sum_{i=0}^{n-1} L^{i}\, 2^{i} \quad \text{.....(3)}$$

and

$$\mathbf{M}=\sum_{i=0}^{n-1} M^{i}\, 2^{i} \quad \text{.....(4)}$$

And

$$(\mathbf{H}_n)_{LM}= \frac{1}{2^{\frac{n}{2}}}(-1)^{\sum_j L_j M_j} \quad \text{.......(5)}$$

The H matrix was used in different ordering (changing the rows order) depending on the number of changes in the signs between successive rows; the familiar known orders are (0, 1, 2, 3, 4, 5, 6, and 7), (0, 1, 3, 2, 7, 6, 4, 5) and (0, 7, 3, 4, 1, 6, 2, 5) for (8*8) matrix.
FWHT: The general form of FWHT matrix is the following:

$$\mathbf{H}_M=\frac{1}{\sqrt{2}}\begin{pmatrix} H_{M-1} & H_{M-1} \\ H_{M-1} & -H_{M-1} \end{pmatrix} \quad \text{......(6)}$$, Where $1/\sqrt{2}$ is the normalization factor;

The previous matrices (H_M,H_n) are used with the following equations:
$Y=1/N * \sum_{i=1}^{N} X_i * WH(n,i)$ ……(7) , n=1,2,….,N ; Where X is the input matrix; Y is the transformed vector. For a WH matrix of size(8*8), N is equal 8;
Inverse transform:
$\mathbf{K} = \sum_{n=1}^{N} Yn * WH(n,i);$ **……(8)** , i=1,2,…..,N ; Where K should equal X.

SLT: Slant transform was originally introduced for images, monochrome and colored images. It efficiently and successfully represents the variations in linear lines brightness. Slant has orthogonal matrices, a constant first row function, a linear second row function (slant), and iteratively constructed matrices [3]. The original formula for size(4) is[8]:

$$\mathbf{T}_4=\frac{1}{2^{1/2}}\begin{bmatrix} 1 & 0 & 1 & 0 \\ a4 & b4 & -a4 & b4 \\ 0 & 1 & 0 & -1 \\ -b4 & a4 & b4 & a4 \end{bmatrix}\begin{bmatrix} T2 & 0 \\ 0 & T2 \end{bmatrix} \quad \text{.....(9)}$$, a_4, b_4 are the scaling constants

V = T_4 * U * transpose(T_4) , U is the image matrix.

For inverse Slant transform :

U = transpose(T) * V * T …..(10)

LWT: it decomposes the lifting wavelet into two coefficient vectors, the approximation and the detailed, the process depends on the wavelet type used in decomposition like Haar, db1,db2,…, the inverse lifting wavelet transform reconstructs the signal from the two coefficient vectors. [9]

II-Improved transforms

The Slant and WHT both depend on some orthogonal matrices and a normalization factor (in case of WHT); for that, the improvement depends basically on finding new matrices that are suitable for audio and make the process of transform either faster, less bit rate errors, or even easier. The matrices' elements and the normalization factors were found to get efficient, suitable and precise transformations.

New transform based on Slant (Improved Slant):

The new transform (slant improvement) works on sound with the following NT matrix:

$$\mathbf{NT=}\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & -\mathbf{1} & -\mathbf{1} \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{1} \\ \mathbf{1} & -\mathbf{1} & \mathbf{1} & -\mathbf{1} \end{bmatrix} \;;$$

This matrix will be used in transform and inverse transform in its original and transposed shape. It completely differs from the original Slant matrix. But will be used in the same methodology. No scaling constants were used.

V = NT * U * Transpose(NT) …..(11) ; where U is the audio data

For inverse transform

U = $\mathbf{\frac{1}{2^N}}$ **(Transpose(NT) * V * NT) ………(12)**

The transformed matrix will be multiplied by $1/2^N$, where N is the size of the matrix NT (N*N).

Improved WHT:

The following new matrix (NM) differs from the original WH matrix, but will be used with the same transform equation;

$$\mathbf{NM_1=}\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & -0.5 & -0.5 & -0.5 & -0.5 \\ 0.5 & 0.5 & -0.5 & -0.5 & -0.5 & -0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 & -0.5 & 0.5 & 0.5 & -0.5 & -0.5 \\ 0.5 & -0.5 & -0.5 & 0.5 & 0.5 & -0.5 & -0.5 & 0.5 \\ 0.5 & -0.5 & -0.5 & 0.5 & -0.5 & 0.5 & 0.5 & -0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 & -0.5 & 0.5 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 & 0.5 & -0.5 & 0.5 & -0.5 \end{bmatrix}$$

$$Y=1/N * \sum_{i=1}^{N} X_i * NM_1(n,i) \quad ……(12)\ ;\ n=1,2,….,N \quad ;$$

Where X is the input signal, Y is the transformed matrix and for NM_1 matrix of size(8*8), N is equal 8;

Inverse transform:

$$X = 4 * \sum_{n=1}^{N} Yn * NM_1(n,i) \quad …….(13)\ ;\quad i=1,2,…..,N;$$

The normalization factor (equal 4) used only in inverse transform.

Another NM_2 matrix proved to be workable with audio signals and gave comparative results;

$$NM_2=\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & -2 & -2 & -2 & -2 \\ 2 & 2 & -2 & -2 & -2 & -2 & 2 & 2 \\ 2 & 2 & -2 & -2 & 2 & 2 & -2 & -2 \\ 2 & -2 & -2 & 2 & 2 & -2 & -2 & 2 \\ 2 & -2 & -2 & 2 & -2 & 2 & 2 & -2 \\ 2 & -2 & 2 & -2 & -2 & 2 & -2 & 2 \\ 2 & -2 & 2 & -2 & 2 & -2 & 2 & -2 \end{bmatrix}$$

The equation of transform is the same but the inverse transform differs only with the normalization factor, where:

$$X =1/4 * \sum_{n=1}^{N} Yn * NM_2(n,i) \quad ………(14)\ ;\quad i=1,2,…..,N;$$

Where X is the inverse transformed signal. The normalization factor (equal 1/4) used only in inverse transform.

III. Implementation

The watermarking technique used to evaluate the transform process was the same of our previous work in [10]. We proposed a blind watermarking method where some features of the audio signals were extracted from specific parts of the audio, other parts of the audio were transformed using LWT and the features were hidden in the last bits of the transformed samples, inverse transform applied on the watermarked parts and returned to the original audio. The method was proved to be robust against specific attacks and the watermarked audio quality was excellent.

The improved transforms were applied on over fifteen audio file and worked perfectly; in this paper we used only one audio file and applied all mentioned types of transforms with the improved once for comparison issues. Each Figure from 1 to 10 represents one watermarked audio frame with a specific transform; Each frame was transformed, watermarked, inverse transformed and returned to audio;

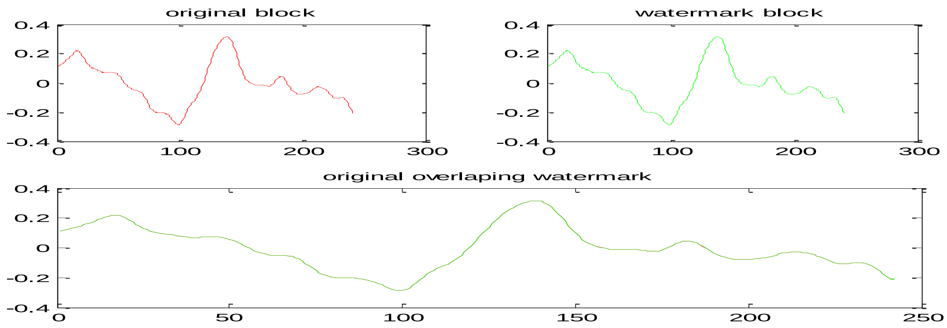

Figure 1: One block transformed using DCT, watermarked and returned to original audio

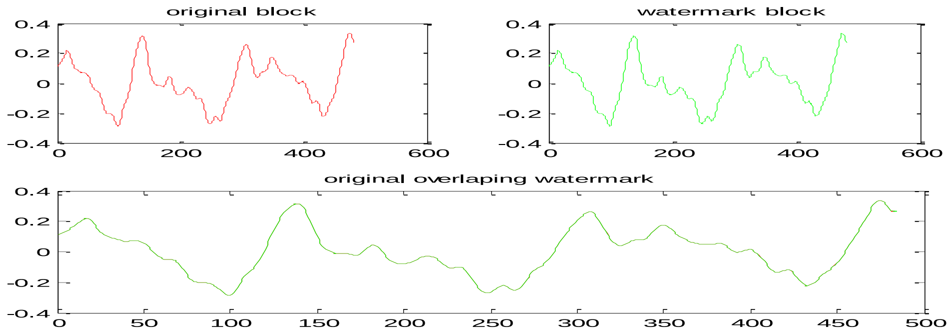

Figure 2: One block transformed using DWT (db1) , watermarked and returned to original audio

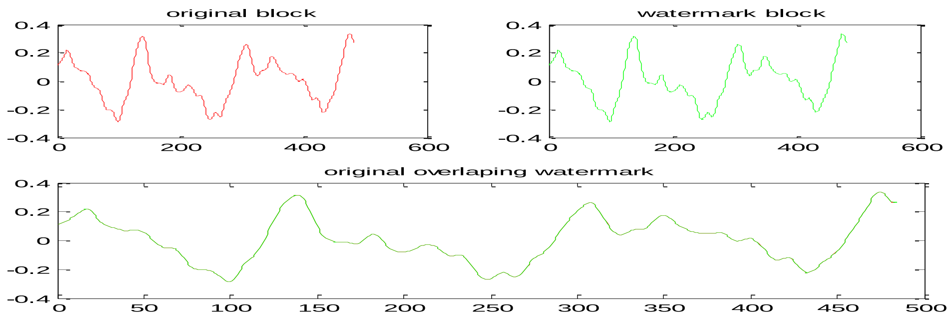

Figure 3: One block transformed using DWT (Haar) , watermarked and returned to original audio

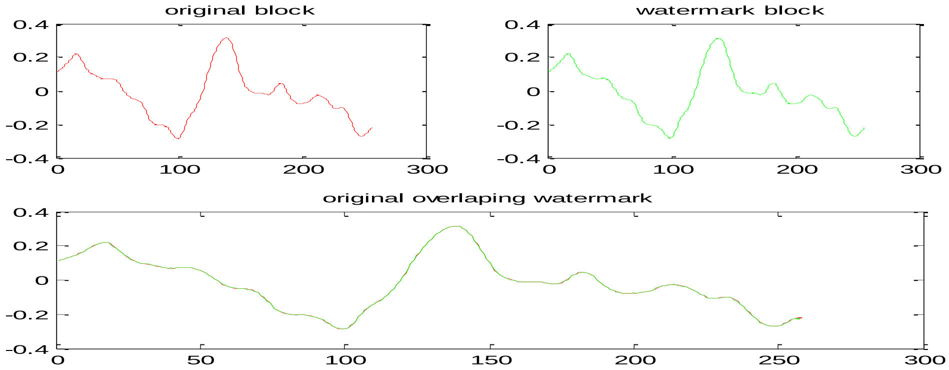

Figure 4: One block transformed using FWHT, watermarked and returned to original audio

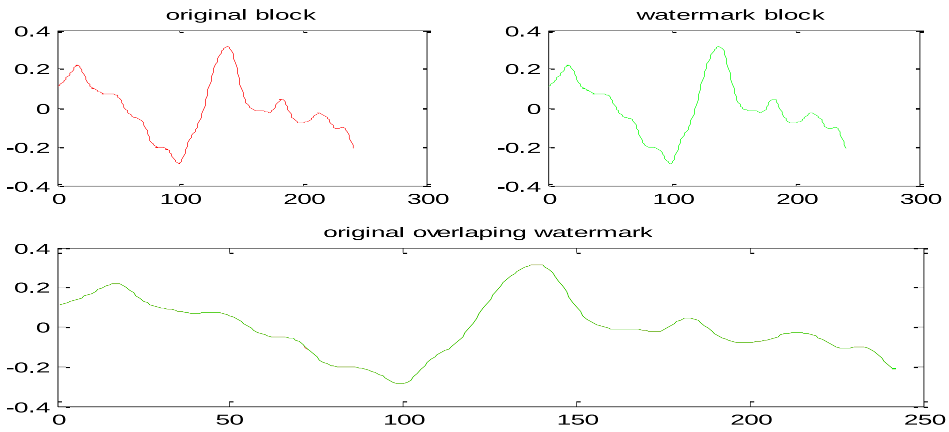

Figure 5: One block transformed using Wave-decomposition, watermarked and returned to original audio

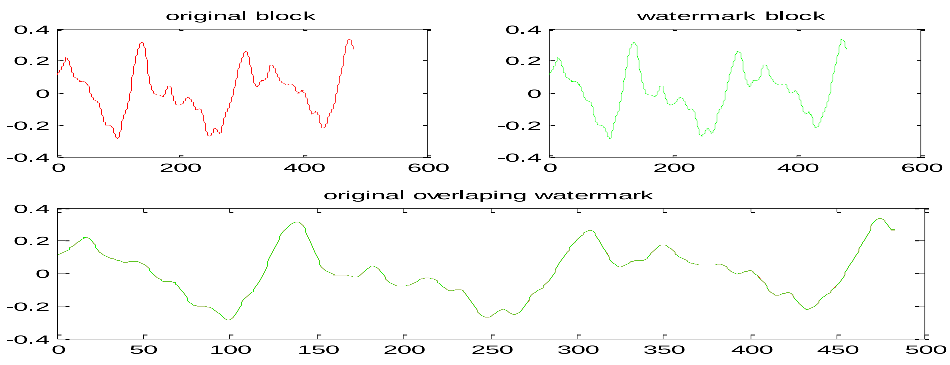

Figure 6: One block transformed using LWT, watermarked and returned to original audio

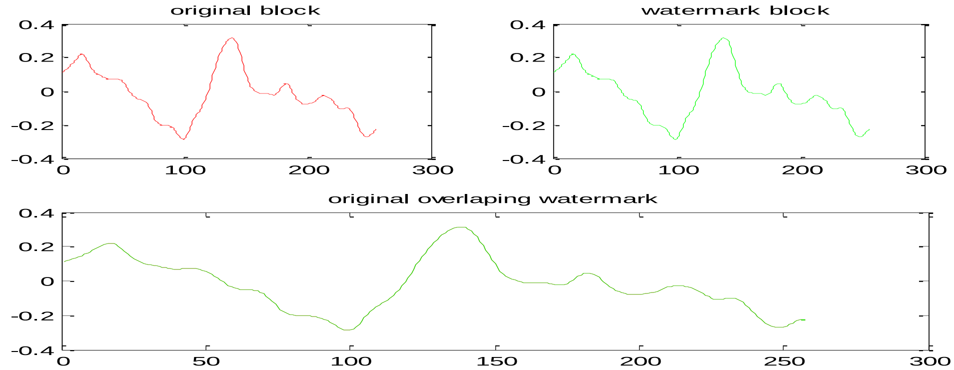

Figure 7: One block transformed using Improved Slant, watermarked and returned to original audio

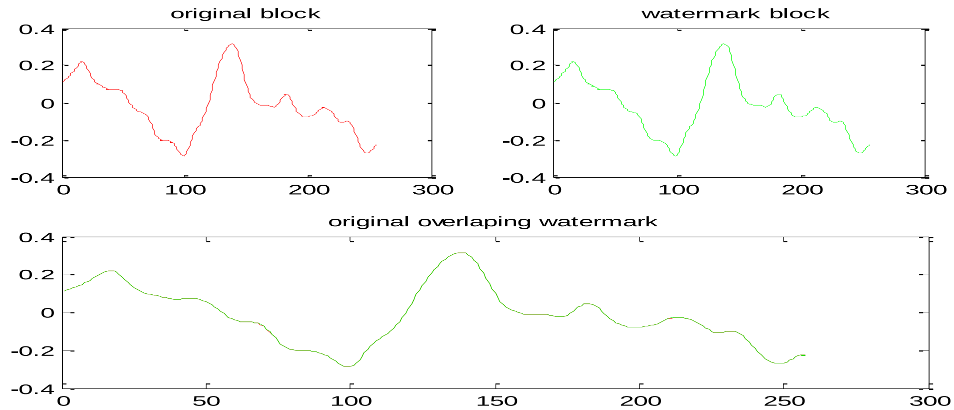

Figure 8 One block transformed using Improved WHT(NM1) , watermarked and returned to original audio

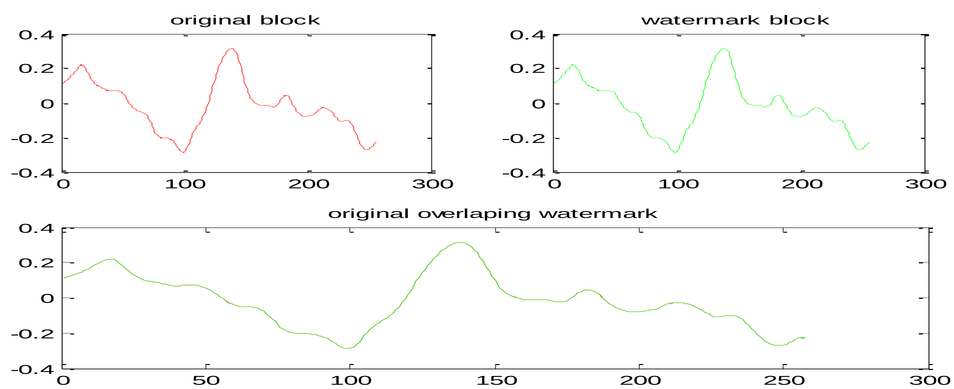

Figure 9: One block transformed using Improved WHT(NM2) , watermarked and returned to original audio

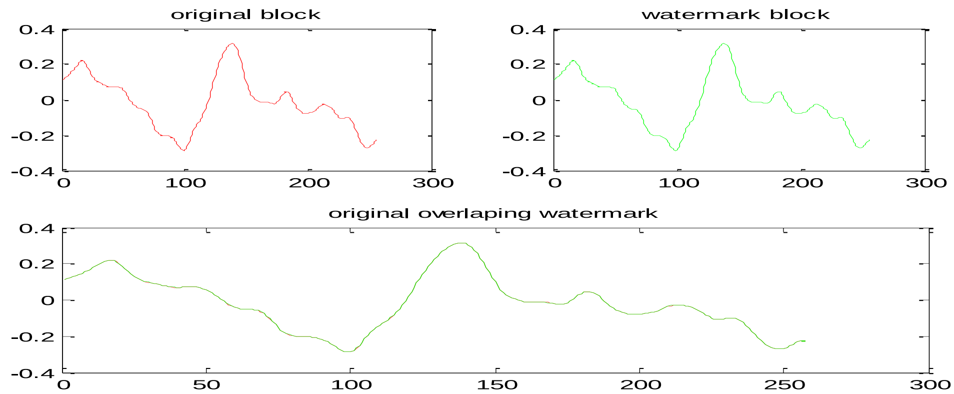

Figure 10: One block transformed using Improved WHT(NM2) , watermarked and returned to original audio

IV. Results and Evaluations

Transforms differ in their methodology and equations which used to change original audio file from time to frequency domain; outputs from those transforms also differ from one to another. Here we applied many types of transform on the same audio file, then the transformed results were used in hiding the watermarks, inverse transforms were applied and the evaluation metrics were calculated. The BER, SNRdB, PSNR, MSE, MAXERR and the overall time were calculated and compared as in table 1.

Many types of transforms were applied on the same file(w14.wav) with the same block numbers(Watermarks number), and find the measurement of performance for each type with the CPU time for the whole process (extract feature, hide feature, extract hidden feature and comparison), i.e. the process of both sides (sender and receiver). The improved Slant transform and the improved WHT were also applied on the same file. The results were as follows:

Table 1: some known transforms and the improved transforms applied on the same file

Type	CPU time	BER	SNR_{db}	PSNR	MSE	MAXERR
DCT	0.7969	0.5347	43.5888	165.5233	1.8228e-012	0.00025170
DWT(db1)	0.8438	0.4986	43.8924	145.7804	1.7181e-010	0.0192
DWT(haar)	0.8594	0.4986	43.8924	145.7804	1.7181e-010	0.0192
FWHT	1.3438	0.4986	43.8924	141.5680	4.5319e-010	0.0032
WaveDec (db1)	0.9688	0.5181	43,7263	165.0710	2.0229e-012	0.00024841
LWT	0.7969	0.4986	43.8924	165.3411	1.9010e-012	0.00010577
New-NT	0.9375	0.4861	44.0027	177.4820	1.1611e-013	0.000059424
New-NM1	0.6719	0.4667	44.1800	156.6205	1.4159e-011	0.00075171
New-NM2	0.5938	0.5208	43.7031	161.9446	4.1554e-012	0.00037354

Table(1) shows that, improved WHT(New-NM1 and New-NM2) transforms have better results than the original FWHT, the MSE for the two types of improved WHT were better than the original FWHT, the CPU time were better than the original, SNRdb were all more than 20dB and were greater than or close to the original. PSNR were all greater than the original.
For the new transform that based on Slant methodology, the results were comparative with the other types of transforms. The CPU time was acceptable and within the range of other transforms (less than Wave Decomposition and FWHT), the SNRdB was better than all the known types of transforms used for comparison, the MSE and the MAXERR were the minimum between all, the PSNR was the largest between all types.
From the previous results, the proposed new transform that base on Slant transform has the best performance compared with all types of transforms used in this research. The improved WHT has a comparative performance with other types of transform and better performance than the original FWHT.

Conclusion

The improved Slant transform was enhanced to work not only with images but it has been proved to be suitable for audio signals and has a comparative performance results. The improved WHT with the new matrices and the normalization factors used in inverse transform process has been proved to be comparative with other types of transforms and gave better performance than the original FWHT. Improved transforms were used in watermarking process by changing the shape and value of original audio signals to create a new representation in frequency domain instead of the basic time domain representation. Transforms had a great effect in watermarking audio files since the hidden data is distributed in a wide range of samples (SS) which makes the watermarking more immune against certain types of attacks.

References

[1] P.C. MALI, B.B. CHAUDHURI and D. DUTTA MAJUMDER,' Some Properties and Fast Algorithms of Slant Transform in Image Processing ', Signal Processing 9, P233-244, 233 North-Holland, 1985.

[2] Juergen Seitz, "Digital watermarking for digital media", University of Cooperative Education, Heidenheim, Germany, 2005.

[3] Syed Ali Khayam, 'The Discrete Cosine Transform: Theory and Application', Information Theory and Coding, 2003.

[4] Vimal Krishnan V.R, Babu Anto P, 'Features of Wavelet Packet Decomposition and Discrete Wavelet Transform for Malayalam Speech Recognition', International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009.

[5] Daubechies, I., Ten lectures on wavelets, CBMS-NSF conference series in applied mathematics. SIAM Ed. 1992.

[6] http://www.mathworks.com/help/wavelet/ref/wavedec.html.

[7] A.A.C.A.Jayathilake, A.A.I.Perera, M.A.P.Chamikara, ' Discrete Walsh-Hadamard Transform in Signal Processing', IJRIT International Journal of Research in Information Technology, Volume 1, Issue 1, January 2013.

[8] WILLIAM K, WEN-HSIUNG CHEN, AND LLOYD R. WELCH, "Slant Transform Image Coding", IEEE, Transaction on communications, VOL. COM-22, NO. 8, 1974.

[9] Baiying Lei, Ing Yann Soon, Feng Zhon, Zhen Li, Haijun Lei, ' A Robust watermarking scheme based on lifting wavelet transform and singular value decomposition', signal processing 92, 2012.

[10] Hala B., Abdual Mohsen J., Sana Ahmed, 'Variant length, Self-extracted audio watermark for verification using LWT and random selections', under publishing, 2016.

[11] Mallat, S., "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Pattern Anal. and Machine Intell., vol. 11, no. 7, pp 674–693, 1989.

[12] Meyer, Y. (1990), Ondelettes et opérateurs, Tome 1, Hermann Ed., English translation: Wavelets and operators, Cambridge Univ. Press. 1993.

[13] Mohamed I. Mahmoud, Moawad I. M. Dessouky, Salah Deyab, and Fatma H. Elfouly, 'Comparison between Haar and Daubechies Wavelet Transformions on FPGA Technology', World Academy of Science, Engineering and Technology International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering Vol:1, No:2, 2007.