European Scholar Journal (ESJ)
Available Online at: https://www.scholarzest.com
Vol. 2 No. 8, August 2021,
ISSN: 2660-5562

ROBUST AND FRAGILE SIGNAL AUDIO WATERMARKING IN ONE ALGORITHM

Rafidison Maminiaina Alphonse,
Telecommunication-Automatic-Signal-Image-Research,
Laboratory/Doctoral School in Science and Technology of
Engineering and Innovation/University of Antananarivo
Antananarivo 101, Madagascar
Ramafiarisona Hajasoa Malalatiana,
Telecommunication-Automatic-Signal-Image-Research,
Laboratory/Doctoral School in Science and Technology of
Engineering and Innovation/University of Antananarivo
Antananarivo 101, Madagascar
Mr. Pravin R. Choube
Assistant Professor, Department of Electrical Engineering,
Dr. D.Y.Patil Institute of Technology, Pimpri, Pune, M.S. India

Article history:

Received: 26th June 2021
Accepted: 10th July 2021
Published: 3rd August 2021

Abstract:

Inserting image in audio file is the purpose of this paper. Music listener is not able to detect the difference between original audio file and watermarked one. Songwriter can protect his artwork against illegal copy and theft of works. Discrete wavelet transform (DWT), singular value decomposition (SVD) are the mathematic approach used. They are popular on these domains but the way to operate is different for each algorithm. Frequency transformation is needed to ensure that the proposed algorithm is robust/fragile across different attack then we operate only on non-sensitive data. This criterion is not enough, so we add a module formatting the matrix to another matrix before applying singular value decomposition. Watermark image must be a binary and its extraction doesn't require original information. It means that the watermarking algorithm that we describe here is a robust/fragile and blind audio watermarking.

Keywords: Watermarking, fragile, robustness, blind extraction, discrete wavelet transform, modified discrete cosine transform, singular value decomposition, matrix formatting.

I. INTRODUCTION

With the rapid development of the internet, multimedia data stored in digital form can be easily replicated and destroyed by illegal users, so protection against intellectual property infringement increasingly becomes an important issue. There are two primary methods to overcome the above problems, which are digital signature and digital watermarking. Digital signature is a kind of number string which can be used as the secret key for both senders and receivers, and it easily stimulates the desire of illegal users to destroy the multimedia data. Digital watermarking technology conceals the watermarks into the multimedia data and later extracts such watermarks to prove the owner of multimedia data, so it is an efficient approach to protect the media contents, widely used for copyright protection, broadcast monitoring, fingerprinting, data authentication, and medical safety, and it has become a hot topic in the field of communication and information security in recent years.

According to the different application carriers, digital watermarking technology can be divided into image watermarking technology, video watermark technology, audio watermark technology, and so on. Compared with the image watermarking technology, it is harder to develop the audio watermarking technology, mainly because the human auditory system is more sensitive than the visual system. With the widespread presence of audio media on the Internet, more and more people are beginning to pay attention to the research about audio watermarking technology, and more and more research algorithms have appeared. An audio watermarking algorithm generally takes into consideration four aspects including imperceptibility, robustness, security and payload capacity. These four indexes are in conflict with each other. The increase of one index may cause a decrease in other indexes. An excellent audio watermarking algorithm should not only guarantee that the watermarked audio has good imperceptibility, but also provide enough payload capacity to accommodate necessary information. In addition, it should have a strong robustness to resist various signal processing attacks in practical application, so as to ensure the security of the extracted watermark. At present, most audio watermarking algorithms have some shortcomings, such as low capacity, poor robustness, and serious decline of the carrier audio quality. The audio watermarking algorithm for copyright protection must have good imperceptibility and strong robustness in resisting most common signal processing attacks that the audio may suffer during the transmission process to ensure that the watermark can be extracted accurately [1][4].

This paper presents an audio robust/fragile watermarking algorithm by coupling discrete wavelet transform (DWT), modified discrete cosine transform (MDCT) and singular value decomposition (SVD). To ensure a good understanding of this algorithm, we develop into the first paragraph the notion of DWT, MDCT and SVD followed by the algorithm presentation, testing, results and conclusion.

II. PRELIMINARIES

A. Discrete Wavelet Transform

Discrete Wavelet Transform is a multilevel transformation method. It decomposes an image into four wavelet sub bands as shown in Figure 1. The wavelet sub bands consist of LL, HL, LH, and HH. LL is a low frequency sub band, which is top-left side of the wavelet sub band. It obtained through Low Pass Filtering (LPF) in both row and column directions. This sub band contains approximate value of an image, which is the most significant part of the image. LL has highest robustness level among all wavelet sub bands, it able to maintains information therein (Abu et al. 2014) (Adi et al. 2015).

The top-right part of wavelet sub band represents the horizontal detail of the image called HL. It is generated from LPF in row order followed by High Pass Filtering (HPF) in column order. Conversely, LH is resulted from process of HPF in row order and LPF in column order. It depicts vertical element of the image and located at the bottom-left side of wavelet sub band. The last part of wavelet sub band is HH, which is the diagonal feature of the image. This sub band contains high wavelet coefficient, hence HH is vulnerable to attacks. DWT is able to decompose an image into multilevel wavelet sub bands as shown in Fig.2. The next level of decomposition is generally performed in LL to get the higher level of wavelet sub band (LL2, HL2, LH2, and HH2) and so on. The highest level of wavelet decomposition is when it reaches a single coefficient value.

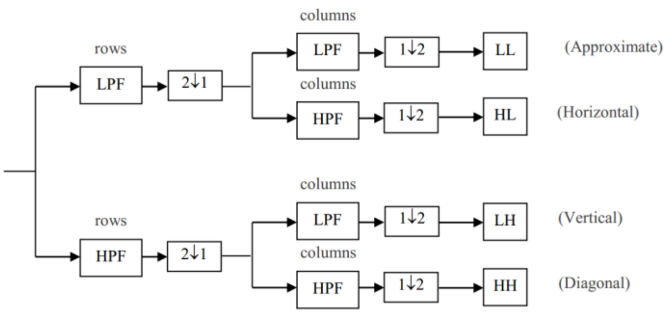

Fig. 1. Decomposition scheme DWT Level on an image

Fig. 2. DWT Level on an image

B. Modified Discrete Cosine Transform

The modified discrete cosine transform (MDCT) and inverse MDCT (IMDCT) are extensively used to realize the analysis/synthesis filter banks of time domain aliasing cancellation scheme for subband coding. This filter bank is equivalent to the modulated lapped transform (MLT) introduced by Malvar. The MDCT/IMDCT has been adopted in several international standards and commercial audio coding products such as MPEG-1, MPEG-2 and AC-3 to achieve high quality audio compression. However, the direct computation of the MDCT in MPEG audio encoding and IMDCT in MPEG audio decoding involves an extensive number of arithmetic operations. Therefore, efficient algorithms for their computation are of great importance.

The definitions and some properties of the MDCT and IMDCT are details below.

Let $\{x(n)\}$ be an input data sequence, the MDCT and IMDCT are respectively defined as

$$X(k) = \sum_{n=0}^{N-1} x(n) cos\left[\frac{\pi}{2N}(2n+1+\frac{N}{2})(2k+1)\right], k = 0,1,\dots,\frac{N}{2}-1 \quad (01)$$

$$\hat{x}(n) = \sum_{k=0}^{\frac{N}{2}-1} X(k)cos\left[\frac{\pi}{2N}(2n+1+\frac{N}{2})(2k+1)\right], n = 0,1,\dots,N-1 \quad (02)$$

Where N is the window length. In general, the recovered data sequence $\{\hat{x}(n)\}$ does not correspond to the original data sequence $\{x(n)\}$. $\{\hat{x}(n)\}$ has the following symmetries:

$$\hat{x}\left(\frac{3N}{4}+n\right) = \hat{x}(\frac{3N}{4}-1-n),\ n = 0,1,\dots,\frac{N}{4}-1 \quad (03)$$

$$\hat{x}\left(\frac{N}{2}-1-n\right) = -\hat{x}(n) \quad (04)$$

Therefore, only $\hat{x}(n)$, for $n = 0,1,\dots,\frac{N}{4}-1$ and $n = \frac{N}{2},\frac{N}{2}+1,\frac{3N}{4}-1$, need to be calculated.

$$X = [X(0),X(1),\dots,X(\frac{N}{2}-1)]^T \quad (05)$$

$$x = [x(0),x(1),\dots,x(N-1)]^T \quad (06)$$

$$\hat{x} = [\hat{x}(0),\hat{x}(1),\dots,\hat{x}(N-1)]^T \quad (07)$$

Where T denotes the transposition. Then the equation (01) and (02) can be written as

$$X = M_{(N/2)\times N}x \quad (08)$$

$$\hat{x} = M^T_{(N/2)\times N}X \quad (09)$$

Where $M_{(N/2)\times N}$ is an $(N/2)\times N$ MDCT matrix.

From (07) and (08), we know that if a realization of the MDCT is developed, then a realization for IMDCT can be obtained by transposing the signal flow graph of the MDCT [5].

C. Singular Value Decomposition

Singular Value Decomposition (SVD) is used to decompose a matrix X into matrices U, S, and V. The diagonal matrix S contains singular values of the matrix X. While, the orthogonal matrices U and V contains the left and right singular values of matrix X respectively.

$$[U\ S\ V] = SVD(X) \quad (10)$$

$$X = U*S*V^T \quad (11)$$

where, V^T is a conjugate transpose of matrix V. S is a diagonal matrix with large singular value contains in its diagonal entries. U and V are complex or real unitary matrices, such that $U*U^T$ and $V*V^T$will result in identity matrix [2].

III. RELATED WORK

A. Embedding preparation algorithm

1: Read audio file and extract sampled data in y and sample rate fs.
2: Split the audio file to different block $y = \{b_1,b_{2,},..,b_n\}$.
3: If the last block b_n doesn't have the same size of the previous blocs, we don't touch this block.
4: Read watermark image M with size $l_M \times c_M$.
5: Convert M to binary image M_b.
6: Reshape M_b to vector
7: Calculate the number of bite N in M_b to be inserted to each block:

$$N = \frac{(l_M \times c_M)}{n}\ or\ \frac{(l_M \times c_M)}{n-1} \quad (12)$$

n or $n-1$ depending on step 3

B. Embedding algorithm

1: Apply DWT Haar level 1 for robust and MDCT for fragile watermarking

$$(LL_k,HL_k,LH_k,HH_k) = DWT(b_k) \quad (13)$$

$$(HH_k) = MDCT(b_k) \quad (14)$$

with $k = \{1,2,3,n\ or\ n-1\}$

2: Calculate length l of HH_k
3: Locate the part of HH_k which cannot form a square matrix by calculating modulo l and the number of element in each square matrix $c \times c$ (step 4)

$$m = modulo(l, c \times c) \quad (15)$$

4: Format HH_k to square matrix SHH_k and don't touch the row from $l - m + 1$ to l. $(l - m)/c$ is the number of square matrix in each block.
5: Apply SVD to each square matrix SHH_k

$$(U_k, S_k, V_k) = SVD(SHH_k) \quad (16)$$

6: Insert M_b
Suppose that

$$S_k = \begin{bmatrix} S_1 & 0 & 0 & & & 0 & 0 & 0 \\ 0 & S_2 & 0 & \cdots & & 0 & 0 & 0 \\ 0 & 0 & S_3 & & & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & & \\ 0 & 0 & 0 & & S_{j-2} & 0 & & 0 \\ 0 & 0 & 0 & \cdots & 0 & S_{j-1} & & 0 \\ 0 & 0 & 0 & & 0 & 0 & & S_j \end{bmatrix} \quad (17)$$

And format M_b to many matrix as below

$$M_{b_k} = \begin{bmatrix} M_{bk_1} & 0 & 0 & & & 0 & 0 & 0 \\ 0 & M_{bk_2} & 0 & \cdots & & 0 & 0 & 0 \\ 0 & 0 & M_{bk_3} & & & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & & \\ 0 & 0 & 0 & & M_{bk_{j-2}} & 0 & & 0 \\ 0 & 0 & 0 & \cdots & 0 & M_{bk_{j-1}} & & 0 \\ 0 & 0 & 0 & & 0 & 0 & & M_{bk_j} \end{bmatrix} \quad (18)$$

$$S_{kw} = S_k + M_{b_k} \quad (19)$$
$$SHH_{kw} = U_k * S_{kw} * V_k^T \quad (20)$$

S_{kw} and SHH_{kw} are respectively the diagonal matrix watermarked and square matrix watermarked.
7: Transform to row SHH_{kw} by adding at the end the row that we don't touch described in step 4 to get diagonal watermarked HH_{kw}.
8: Apply $IDWT$ to obtain watermarked block

$$b_{kw} = IDWT(LL_k, HL_k, LH_k, HH_{kw}) \quad (21)$$
$$(b_{kw}) = IMDCT(HH_{kw}) \quad (22)$$

9: Save audio with parameters $y_w = \{b_{1w}, b_{2w,}, .., b_{nw}\}$ and fs
10: Read audio file

C. Extraction algorithm

1: Repeat step 1 to 3 of embedding preparation algorithm using watermarked audio.
2: Repeat step 1 to 5 of embedding algorithm applying the previous step result.
3: Manage diagonal matrix:

$$M_{bw} = \begin{cases} 1 & si\ S_{jw} \geq 1 \\ 0 & otherwise \end{cases} \quad (23)$$

S_{jw} presents the diagonal element watermarked and M_{bw} is mark vector watermarked.
4: Reshape M_{bw} to the same size of binary watermark image then display the mark

IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

Before going to experiment our new algorithm, let's learn three factors to qualify the performance such as Signal to Noise Ratio (SNR), Mean Square Error (MSE) and Normalized Correlation Coefficient (NCC).

A. Audio quality parameters

1) Signal to Noise Ratio

Signal to noise ratio is a parameter used to know the amount by which the signal is corrupted by the noise. It is defined as the ratio of the signal power to the noise power. Signal to noise ratio can also be calculated by equation below. y is the un-watermarked audio signal and y_w is the watermarked audio signal. Both y and y_w has M_t samples [3].

$$SNR = 10log\left(\frac{\sum_{a=1}^{M_t} y^2(a)}{\sum_{a=1}^{M_t}(y(a)-y_w(a))}\right) \quad (24)$$

2) Mean Square Error

The mean squared error (MSE) of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the "errors." The error is the amount by which the value implied by the estimator differs from the quantity to be estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate. Where y is watermarked audio and y_w is original audio. N is no. of samples [3].

$$MSE(x,y) = \frac{1}{N}\sum_{i=1}^{N}(y - y_w)^2 \quad (25)$$

3) Normalized Correlation Coefficient

In order to eliminate the subjective factors and reflect the fairness of copyright protection, we employed the normalized correlation coefficient (NNC) to estimate the similarity between the original watermark and the extracted watermark. The normalized correlation coefficient (NNC) is defined as equation [3].

$$NCC = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}[M_b(i,j) \times M_{bw}(i,j)]}{\sum_{i=1}^{N}\sum_{j=1}^{N}[M_b(i,j)]^2} \quad (26)$$

B. Performance measurement

For our experience, we select one voice file “handel.wav” which is the default voice file of Maltab in handel.mat. It will pass through our algorithm for robust and fragile watermarking. We plot the original audio in below figure in objective to see the difference with watermarked audio. One binary photo will be inserted “My Song”. SNR and MSE are the way to measure the quality of voice and NCC is for extracted watermark image.

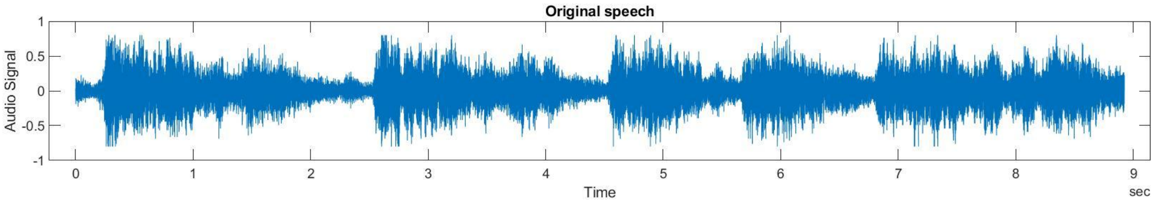

Fig. 3. Original file Audio

Fig. 4. Watermark image

After applying the algorithm, Table I is showing the performance with and no attack.

TABLE I

WATERMARKING EXPERIMENTATION

Watermarked Audio	Parameter	Value with DWT	Watermark Image	Valeur MDCT	Watermark Image
No attack	SNR	Inf dB		Inf dB	
	MSE	0.0000		0.0000	
	NCC	1		1	
White Gaussian Noise	SNR	150.7014 dB Inf dB		19.1061 dB	
	MSE	0.0000		0.1003	
	NCC	0.9801		0.9100	
Compression	SNR	199.1154 dB		17.9015 dB	
	MSE	0.0000		0.3156	
	NCC	0.9901		0.9001	
Low Pass Filtering	SNR	186.0715 dB		21.1094 dB	
	MSE	0.0000		0.0315	
	NCC	0.9891		0.9411	

The watermark can be extracted without the help from the original digital audio signal and can be easily implemented. To check the robustness SNR, MSE and NCC are calculated. After embedding watermark, we have two results:

- Robust method - DWT: The SNR of proposed method is above 20 dB which ensures the imperceptibility of the proposed system. This satisfies the IFPI requirement (20 dB). The value of MSE is always less than 1. The value of NCC was near to 1. It can be seen that using the proposed algorithm to embed watermark is inaudible and is robust to some common attacks like compression, noise and low-pass filter. We can extract watermark and it has good legibility [3].
- Fragile method - MDCT: The result in Table I is gotten over twice copy of watermarked signal. IFPI requirement is not respected any more except for low pass filtering attack but the quality of sound remains bad. MSE values increase but the extracted image is still acceptable in quality therm juged by NCC measurement.

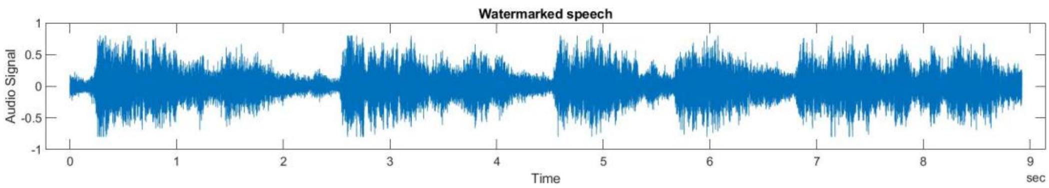

Fig. 5. Watermarked Audio file without attack using DWT

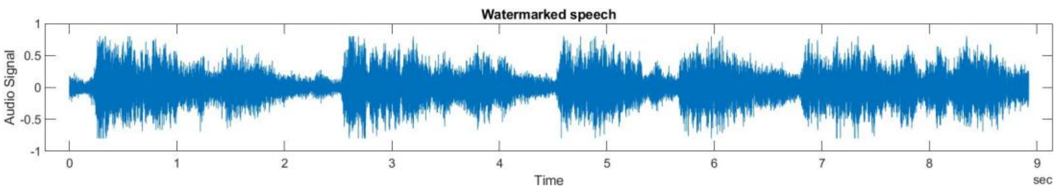

Fig. 6. Watermarked Audio file without attack using MDCT

According to Fig.5. and 6, there is no big difference between embedding with DWT and MDCT spectrum.

V. CONCLUSION

As said in introduction, Audio watermarking is remaining a pertinent hot topic in the field of communication and information security in recent years for copyright protection, broadcast monitoring, fingerprinting, data authentication, and medical safety. We can say that imperceptibility, robustness, fragility, security and payload criteria are respected by this new method. DWT combined with SVD remains a robust tool and each searcher find a method to manipulate them to get a strong algorithm. To fight against illegal copy, DWT is replaced by MDCT. After any attack, even a copy can change the music content with SNR less than 20 dB. The watermak image can be extracted, however NCC value is so high.

REFERENCES

1. W. QIULING, W. MENG, "A NOVEL ROBUST AUDIO WATERMARKING ALGORITHM BY MODIFYING THE AVERAGE AMPLITUDE IN TRANSFORM DOMAIN," APPL. SCI. 2018, 8, 723.
2. P. ADI, F. RAHMANTI, "IMAGE WATERMARKING ON LOW FREQUENCY DWT USING SINGULAR VALUE DECOMPOSITION AND DITHER QUANTIZATION," SEMINAR NASIONAL TEKNOLOGI INFORMASI DAN KOMUNIKASI 2016 (SENTIKA 2016), YOGYAKARTA, 18-19 MARET 2016.
3. MEHTA, N. SHARMA, "SECURE AUDIO WATERMARKING BASED ON HAAR WAVELET AND DISCRETE COSINE TRANSFORM," IN INTERNATIONAL JOURNAL OF COMPUTER APPLICATIONS (0975 - 8887), VOLUME 123 – NO.11, AUGUST 2015,
4. Al-Haj, "An imperceptible and robust audio watermarking algorithm," EURASIP Journal on Audio, Speech, and Music Processing 2014, 2014:37.
5. J. WU, H. SHU, L. SENHADJI, L. LU, "MIXED-RADIX ALGORITHM FOR THE COMPUTATION OF FORWARD AND INVERSE MDCT," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART 1 FUNDAMENTAL THEORY AND APPLICATIONS, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE), 2009, 56 (4), PP.784-794. FF10.1109/TCSI.2008.2002918FF. FFINSERM-00344948F.