J Electr Eng Technol.2018; 13(1): 476-484
http://doi.org/10.5370/JEET.2018.13.1.476

ISSN(Print) 1975-0102
ISSN(Online) 2093-7423

An Efficient Selective Method for Audio Watermarking Against De-synchronization Attacks

Baydaa Mohammad Mushgil*†, Wan Azizun Wan Adnan, Syed Abdul-Rahman Al-hadad** and Sharifah Mumtazah Syed Ahmad****

Abstract – The high capacity audio watermarking algorithms are facing a main challenge in satisfying the robustness against attacks especially on de-synchronization attacks. In this paper, a robust and a high capacity algorithm is proposed using segment selection, Stationary Wavelet Transform (SWT) and the Quantization Index Modulation (QIM) techniques along with new synchronization mechanism. The proposed algorithm provides enhanced trade-off between robustness, imperceptibility, and capacity. The achieved watermarking improves the reliability of the available watermarking methods and shows high robustness towards signal processing (manipulating) attacks especially the de-synchronization attacks such as cropping, jittering, and zero inserting attacks. For imperceptibility evaluation, high signal to noise ratio values of above 22 dB has been achieved. Also subjective test with volunteer listeners shows that the proposed method has high imperceptibility with Subjective Difference Grade (SDG) of 4.76. Meanwhile, high rational capacity up to 176.4 bps is also achieved.

Keywords: Stationary wavelet transform (SWT), Quantization Index Modulation (QIM), De-synchronization attacks, Jittering attack, Cropping attack, Subjective listening grade(SDG), Signal to noise ratio (SNR)

1. Introduction

Digital watermarking has been capturing the interests in programming society since several decades ago. This is highly related to the fact that development of software and programming techniques can allow illegal use and distribution of digital files. Watermarking Copyrights is one of the methods that were developed as a solution to reduce such illegal access [1].

Generally, audio watermarking is considered more challenging compared to image and video watermarking. Inserting watermarks to digital audio files is more complex compared with image and video. This is due to the fact that the sensitivity of human auditory is more than the human visual system. Hence, the invisibility for images is easier to achieve than the inaudibility. In addition, the size of hidden information data that can be embedded robustly and inaudibly in an audio file is much lower than visual media. This is because the audio files are represented by far less samples per time interval.

Several factors affect the efficiency of audio watermarking, namely, imperceptibility, robustness, payload and security. Firstly, imperceptibility of the audio file means that the embedded data should not introduce audible changes to the file. This is usually measured using signal to noise ratio where the watermarked audio should be more than 20 dB for it to be considered imperceptible. As for the robustness, the embedded data should survive signal processing attacks that manipulate the audio file which aims to remove the embedded watermark. Payload is the number of bits embedded in audio signal usually measured in bits per second (bps). For it to be considered satisfying the watermarking efficiency, it should be more than 20 bps. As for security [2, 3] requirement, it is to ensure that only the owner can extract the embedded data and no one else can find it. This issue distinctly depends on the algorithm of embedding and secret keys. However, only three factors namely imperceptibility, robustness and capacity have a conflicting relationship [3, 4]. Hence, the reliability of any algorithm depends on how good is the performance of the three factors with a tradeoff relationship among them. While imperceptibility should always be guaranteed in audio watermarking, the conflicting relationship will then be between the robustness and capacity.

Watermarking with high robustness to common attacks such as filtering, adding noise, amplitude scaling, should also cope with de-synchronization attacks such as random cropping, zero inserting, and jittering attack. De-synchronization attack is considered as a real threat to extract the embedded watermark. It restricts the capacity of embedding algorithms.

Until now, the need for an efficient system to survive

† Corresponding Author: Department of Computer and Communication, Faculty of Engineering, University Putra Malaysia UPM., Malaysia. (baydaait@yahoo.com)
* Information and Communication Engineering Department, Al-Khwarizmi College of Engineering, University of Baghdad., Iraq.
** Department of Computer and Communication, Faculty of Engineering, UPM., Malaysia. ({wawa, sar, s_mumtazah}@upm.edu.my)
Received: June 5, 2016; Accepted: September 26, 2017

the de-synchronization attacks with high data payload is still not fully satisfied [5]. The available audio watermarking algorithms lack the efficiency feature in terms of acceptable trade-off between robustness, imperceptibility and capacity. If one of these evaluation factors is achieved in these works, the others are not given the required performance for audio watermarking. Thus, an efficient audio watermarking algorithm that considers all these factors is needed to fulfill the multimedia market requirements.

In this paper, a robust, imperceptible and high capacity algorithm is proposed by using the Stationary Wavelet Transform (SWT) and the Quantization Index Modulation (QIM) technique with new synchronization method. The proposed algorithm provides an enhanced trade-off between these efficiency factors. Thus, this work overcomes the main shortcoming of the available works.

Due to its simplicity and similarity to the Discrete Wavelet Transforms (DWT), SWT has been used in the proposed audio watermarking. The SWT overcomes the issue of variant translation in DWT caused by the down-sampling vectors consideration.

Selecting QIM for watermarking is simple and needs less processing time [6]. Furthermore, the QIM method is suggested to be used in digital watermarking because of its stability against attacks.

The developed synchronization technique reduced the false alarm to zero where the embedded data is to be surrounded with synchronization codes as shown in Fig. 1. Thus, the proposed synchronization technique provides direct extracting process and then reduces the searching time by using beginning and ending synchronization codes.

Embedding a multi bit data process starts with embedding a stream of synchronization code and ends with embedding another stream of synchronization code. While embedding synchronization code process is implemented in time domain, the watermark will be hidden in the SWT to ensure the security of algorithm. The total embedding procedure is implemented in a selected segment with relatively high energy to cope with the de-synchronization attacks. Embedding technique using mean value quantization is also immuned to de-synchronization attacks. Three keys are needed for proposed embedding process namely key K1 represents segments' length, key K2 is the step size of quantization and key K3 represents the initial value for coding signal. Those keys will increase the security of the algorithm against attackers since they are needed for the extracting

The rest of this paper is organized into sections as follows: Section 2 discusses the literature review of the efficiency and the de-synchronization problem. Section 3 describes SWT transform implementation. Section 4 illustrates the methodology of proposed algorithm; while Section 5 explains the implementation of embedding and extracting processes. The evaluation of the proposed algorithm is detailed in Section 6. Finally the conclusion is summarized in Section 7.

2. Literature Review

Several algorithms were introduced for audio watermarking aiming to achieve robustness towards de-synchronization attacks as well as to cover most of the efficiency factors pointed previously with a different tradeoff between them.

Note that one of most challenging types of attacks is the de-synchronization attack which always represents an obstacle for increasing the capacity of hidden data.

Some methods were introduced to resist de-synchronization. By means of the histogram bins' relationship, Xiang et al. [7] embedded data by modifying the relation between three consecutive bins. Although this algorithm was robust to Time Scale Modification (TSM) attacks, the hidden data could not be recovered after Low Pass Filtering (LPF) attack with less than 7 KHz and MP3 compression. The maximum bit rate hiding (capacity) of the algorithm is only three bps which does not satisfied the capacity criterion.

The same authors tried to recover the hidden data after LPF [8] by implementing same histogram technique in the wavelet domain. However, the capacity of this approach is significantly degraded to only two bps. On the other hand, to cope with LPF and MP3 compression, the authors applied histogram to the "Gaussian filtered low-frequency component" [9] by modifying two consecutive bins' relationship. However, the capacity of the hidden watermark data does not improve.

Yang et al. [10] hide the data in the histogram of the un-decimated wavelet transform by introducing relationship between four consecutive bins. Their algorithm was robust

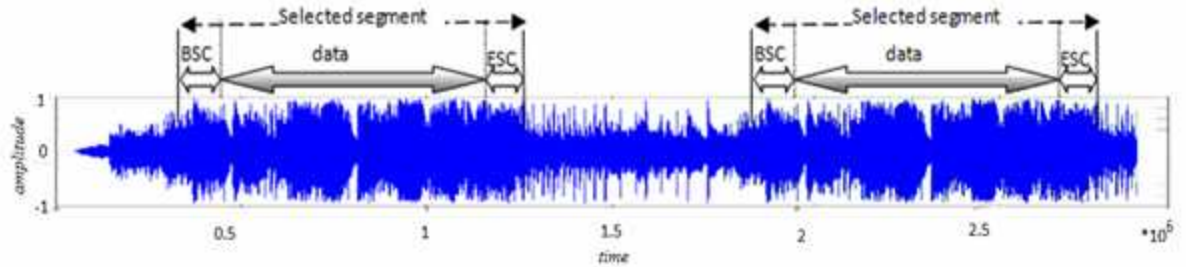

Fig. 1. Synchronization method in proposed algorithm

to de-synchronization attacks however, the capacity is still low.

It can be concluded that wherever the histogram based method is implemented either in the time domain or transform domain, it has a very low capacity which is lower than minimum accepted ratio although it shows high robustness to malicious attacks,

Fan [11] introduced a method for audio watermarking by hiding data in the low frequency in order to resolve the modification of playback speed of an audio signal. This technique adopted a relationship between three consecutive segments. However, his algorithm was not robust to low-pass filtering.

Al-Haj [5] introduced a new method to hide data in a matrix constructed from the detail coefficients of the DWT, and then embedded the data in the Singular Value Decompositions (SVD) of the constructed matrix. The author claimed to hide a huge number of bits per second which is up to 250 bps. However, his semi-blind algorithm was weak against de-synchronization attacks like cropping and jittering. This could be due to the lack of synchronization code in his algorithm.

Chen et al. [12] tried to find optimal modification for the low-frequency coefficients in wavelet domain, but his complex algorithm needs a large number of keys for extracting process with maximum capacity of 43 bps.

Wang et al. [13] could hide data in the exponent moments of frames using quantization index modulation of the exponent value. However, before calculation of exponent moments, authors needed to transform frames to the geometric center in order to acquire the invariant translation property. Nevertheless, the achieved capacity is only 20 bps.

Even though these works tried to improve the audio watermarking performance in terms of efficiency factors, they could achieve some high results but not covering all sides of efficiency like achieving robustness with low capacity or achieving capacity with low robustness towards some attacks or achieving robustness towards some attacks and fail towards the others. Therefore, an efficient audio watermarking method satisfying the conflicting imperceptible, robustness and capacity appears not yet fulfilled.

The main aim of this proposed study is to design an efficient audio watermarking algorithm that provides the best trade-off between efficiency factors of imperceptibility, robustness and capacity. Thus, all these factors are considered in the design and evaluation. In addition, the synchronization technique proposed reduces the searching time of the hidden data in watermarked audio signal.

3. Stationary Wavelet Transform (SWT):

The SWT is chosen for the proposed watermarking due to its simplicity and similarity to the DWT. The only difference is that SWT is computed without using the down-sampling vectors in order to overcome the problem of variant translation in DWT.

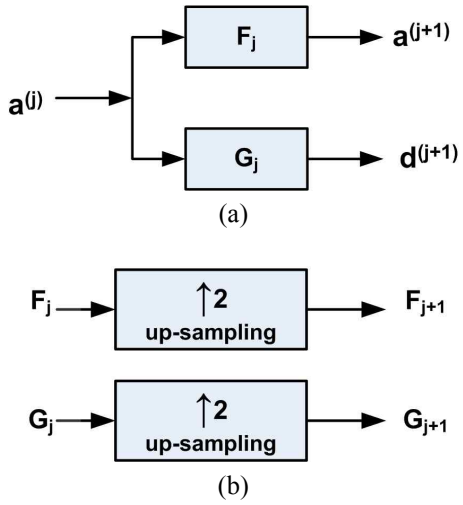

Fig. 2. Calculation of SWT: (a) Decomposition step, (b) Filter computation for next level

3.1 Calculation of SWT

To recognize the choice is made (even or odd decomposition) at level j a variable £j is defined, i.e. £j =1 if the elements were decimated with an odd index.

Then each decomposition is therefore identified by a sequence of 0 and 1 and £= £1… £j is called the £-decimated DWT. Hereafter, the SWT is defined as (£-decimated DWT) for signals with length which is divisible by 2^j where j is the maximum level of decomposition. The calculation of approximation and detailed coefficients of SWT is obtained by filtering the signal using two decomposition filters, as shown in Fig. 2. The filters are up-sampled after each stage in order to overcome the DWT variant translation problem.

where $a^{(0)}$: is the original signal.

a(j), d(j): approximation and detail coefficients at level j respectively.

Fj, Gj : are the low-pass and high-pass decomposition filters respectively.

↑2 : means up-sampling.

4. Proposed Watermarking Techniques

Practically, a successful watermarking system depends on several factors such as efficiency and computational complexity. However, the watermarking system will not be acceptable without the success of embedding and extracting processes. The embedding process depends on the algorithm designated to embed data, meanwhile the extracting process should deal with modified or tampered signal. This is due to attacks that might target the

watermarked signal intentionally to remove the hidden mark. The embedding process should use some techniques that the original signal should undergo to carry water-marking data. These techniques help increasing the robustness of hidden data to survive attacks. On the other hand, the extracting process can make use of some embedding techniques to locate and evolve hidden data again.

The algorithm of the proposed watermark was implemented by combining two audio watermarking techniques to achieve an enhanced performance compared to the other common watermarking with only one technique. The SWT and QIM embedding techniques have been chosen to design the proposed audio watermarking system. Segment selection technique is developed to provide robustness for the embedded data and assist in the localization of embedded data during extraction process [14], meanwhile the QIM adds additional robustness feature and supports easy and fast extraction [15].

The proposed algorithm of audio watermarking is implemented using MATLAB 7.12.0 (R2011a) and the code is written as M-files in the library of the program software. The open architecture of MATLAB makes it easy to explore data, implement algorithms and design new computation tools. MATLAB supports digital signal processing using both simulink and coding. For audio signal, many kind of signal processing can be implemented easily and quickly. For instant, the audio file can be loaded and analyzed using spectral analysis by transforming it to its Fast Fourier Transform (FFT) or Discrete Fourier Transform (DFT). Moreover, MATLAB supports the wavelet mathematical tools which have been a very rich area for many signal processing and signal analysis problems [16]. Wavelet transforms such as Discrete Wavelet Transform (DWT), Stationary Wavelet Transform (SWT), and Lifting Wavelet Transform (LWT) can be implemented easily using MATLAB. Furthermore, it is possible to frame the audio signal for frame analysis purposes due to the ease of loop designating and implementing. Written codes are also easy to call as functions. The M-file plays a big rule in the implementation where each specific designated state code of embedding, extracting and testing is saved. The M-file contains the programming code which can be called as a function in MATLAB.

The integrated technique that is used in our proposal gave our method the superior advantage over other methods. The proposed method works as follows:

4.1 Selecting Segments

To ensure survival of embedded data against de-synchronization attacks, the segment is chosen in a selective manner. The selection method similar to [14] depends on the amount of energy of the segment. The segments, which have relatively high energy and very small silent periods, will be selected for embedding process. This selection can guarantee the stability of segments against attacks especially de-synchronization attacks like cropping and zero inserting. The proposed selection mechanism is implemented by calculating the energy of a possible segment with a condition is that embedding segment will have energy greater than a predefined threshold. The energy of the signal will be computed depending on root mean square energy for the segment, which is computed using the following expression

$$E_{r.m.s} = \sqrt{\left(\frac{1}{L} \sum_{i=1}^{L} \left(S(i) \right)^2 \right)} \quad (1)$$

where $S(i)$ is the samples values for the corresponding segment and L is the segment length. Only segments with signal energy higher than predefined threshold value will be coded with the watermarking bits. Selecting segment must avoid long silent periods in watermarking because those silent periods are unstable against attacks and the data might be lost if silent periods are used for embedding. Moreover, selecting segment can reduce the search area in the extracting side. It is worth mentioning that we realized the energy value *Er.m.s* depends on the audio types. Thus, the threshold value might change too.

4.2 Synchronization code

The synchronization code [17, 18] can determine hidden data position after attacks especially the de-synchronization attacks like compression attacks. In addition, it can benefit in reducing the false alarm. If the chance plays a role to recover the synchronization code from somewhere it was not put in or with a little shift in segments locating The new introduced technique for splitting the synchronization code to two parts: beginning synchronization code (BSC) and ending synchronization code (ESC) to be surrounding the embedded data. This way will reduce the false alarm to zero for multi-bit embedding process and ensure full extraction of embedded data. The algorithm will allow extracting data only in case of finding both synchronization codes in the beginning of extracting process. This case will reduce searching time; also direct extracting is guaranteed instead of trial and false extracting algorithm. The synchronization code will be embedded in time domain using Quantization Index Modulation (QIM); while the watermark signal is will be embedded into SWT transform using the same QIM method (see section 4.4). Synchronization bits are simply 8-bits sequences generated from '0' and '1'. This makes the extraction of synchronization bits easy and quick; while the algorithm stills secure because of the use of secret keys.

4.3 Code generation

In the proposed work, watermarking signal is a bits stream $S(I = 1: n)$ or could be a binary logo image

transformed to one dimension array. The stream is coded with a signal generated using the chaotic, where the characteristics of chaotic signal take careful attention in the field of signal processing [19]. The following chaotic expression is used:

$$x(k+1)=\begin{cases} 2x(k) & if \quad 0 \le x(k) < 0.5 \\ 2(x(k)-1) & if \quad 0.5 \le x(k) \le 1 \end{cases} \tag{2}$$

where the initial value $x(0)$ ε $(0,1)$ which is considered as key K1 at the extraction side. Then, the coding signal will be

$$C(k)=\begin{cases} 1 & if \quad x(k) > \tau \\ 0 & otherwise \end{cases} \tag{3}$$

where τ is a predefined constant. Afterwards, water-marking bits are just the result from XORing the bit stream with the coding signal $C(k)$.

$$W(i)=S(i) \oplus C(i) \tag{4}$$

4.4 Quantization index modulation -QIM method

The QIM method is one of the most robust methods introduced in audio watermarking. Adopting this method for the mean value of a segment in the transformed domain increase the robustness against famous white Gaussian noise addition attack (AWGN). Furthermore, selecting QIM for watermarking is simple and needs less time [6]. In addition, using different quantization steps in the QIM technique offers more security to the embedding algorithm. The size of the step is going to be the second key for extraction process. The calculation of QIM embedding process to hide one bit in the mean value of a block will be as follows:

1) Initially, calculate the absolute mean value M of a block signal S with length L

$$M=\frac{1}{L}\sum_{1}^{L}\left|S(i)\right| \tag{5}$$

2) Then, the modified mean value of the block signal to embed one bit is going to be :

$$M'=\begin{cases} Round\left(\frac{M}{\Delta}\right)\times\Delta & if \quad W_k=1 \\ Floor\left(\frac{M}{\Delta}\right)\times\Delta+\frac{\Delta}{2} & if \quad W_k=0 \end{cases} \tag{6}$$

where

Δ : is the step size for quantization also can be denoted as embedding strength.

Round(x): means rounding x to the nearest integer.

Floor(x): means to round x towards minus infinity.

The extraction of hidden data can be easily obtained by following next steps:

i. Calculate M, according to Eq. (5)

ii. $$D=Round\left(\frac{M}{\Delta}\right) \tag{7}$$

iii. $$W_i=\begin{cases} 0 & if \quad D\,even \\ 1 & if \quad D\,odd \end{cases} \tag{8}$$

The rounding procedure is to solve predicted manipulations was directed to the watermarked signal.

5. Embedding and Extracting Process Implementation

The implementation of the proposed algorithm can be explained as follows.

5.1 Embedding process

The embedding procedure for proposed algorithm is illustrated in Fig. 3; which can be described by the following steps:

1. Generate the coding signal using Eqs. (2) and (3) and then code the binary stream using Eq. (4).

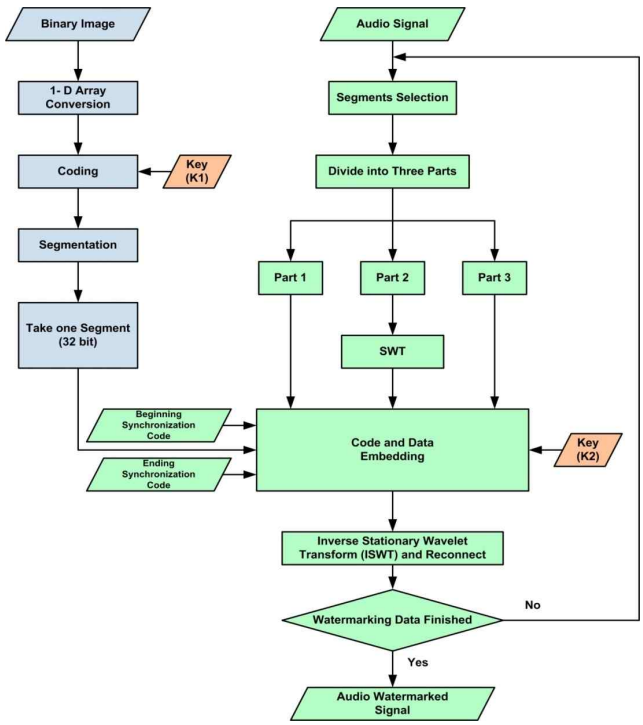

Fig. 3. Flowchart for embedding process

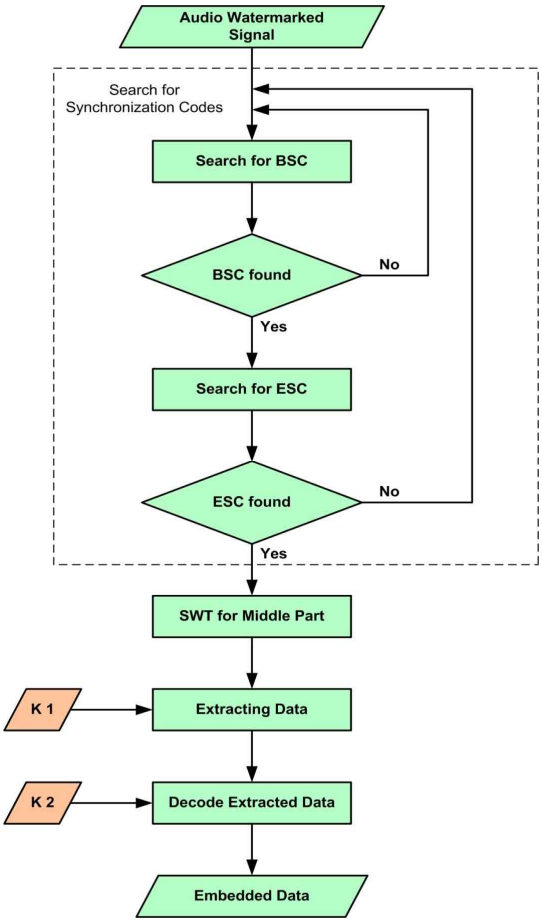

Fig. 4. Flowchart for extracting process

2. Select segment by calculating energy using Eq. (1) and the following condition
   $Er.m.s > T$
3. If this condition is achieved; then the segment will undergo the embedding process of following steps else another segment will be checked.
4. Partition the new signal into three frames: BSC, watermark, ESC frames for previously denoted sizes.
5. Divide BSC and ESC frames separately into blocks of 250 sample length.
6. Directly embed BSC code and ESC code using mean value quantization method.
7. Transform watermark frame to the first level of SWT transform.
8. Divide approximation coefficients into blocks of 250 coefficients each.
9. For each d block, perform the QIM method according to corresponding embedding bit.
10. Connect all parts together again and reconstruct the audio segment by ISWT.

5.2 Extracting process

The extracting process is described as follows:

1. Select the search area.
2. Divide into three frames.
3. Search the BSC code in the first frame using extraction of QIM method steps. If BSC found, go to next step. If not, start new search.
4. Search the ESC code for the third frame. If ESC found proceed with next step.
5. Perform SWT for the middle frame of segment.
6. Extract watermarking bits using same QIM method. Fig. 4 shows the steps of extracting process.

6. Evaluation of Algorithm

The performance of the proposed algorithm is evaluated in terms of imperceptibility, robustness and capacity, as shown in Fig. 5. The evaluation of the three important factors shows an enhanced effect on watermarking efficiency for the proposed watermarking algorithm.

6.1 Imperceptibility test

Two important factors to measure the imperceptibility: signal to noise ratio (SNR), which measures the distortion in the signal caused by embedding process, and subjective listening test. Fig. 6 shows the original and watermarked signal and the difference between them.

6.1.1 Signal to noise ratio (SNR)

The signal to noise ratio for the watermarked audio files is to be calculated as in Eq. (9). All watermarked files show high signal to noise ratio with more than 22 dB which is a very acceptable ratio. It is calculated as:

$$SNR = 10\log\left(\frac{\sum_{i=1}^{N} S(i)}{\sum_{i=1}^{N}\left(S'(i) - S(i)^2\right)}\right) \quad (9)$$

where:

$S(i)$ is the sample's value of original signal

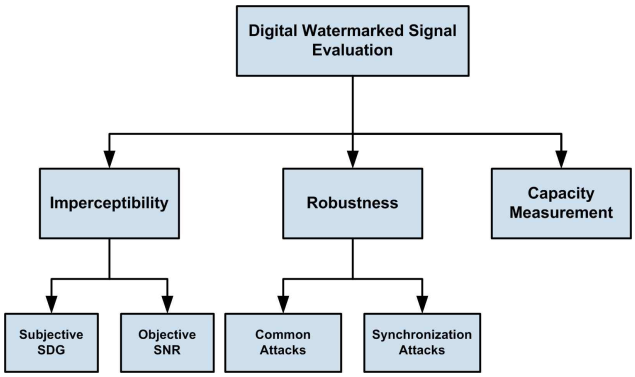

Fig. 5. Evaluation of the proposed algorithm

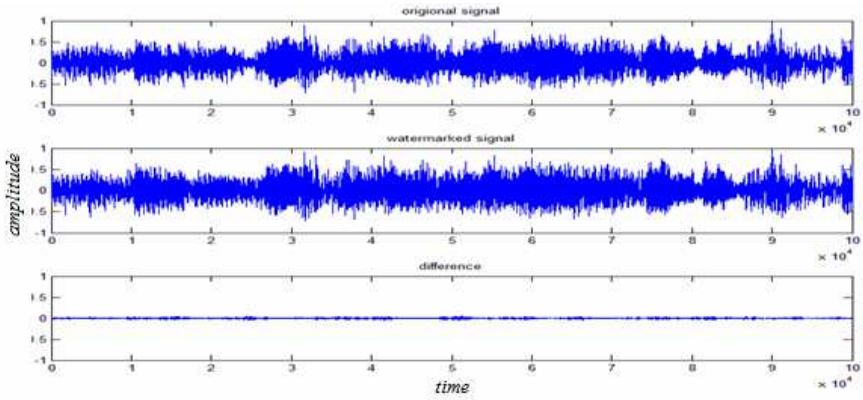

Fig. 6. Difference between original and watermarked signal

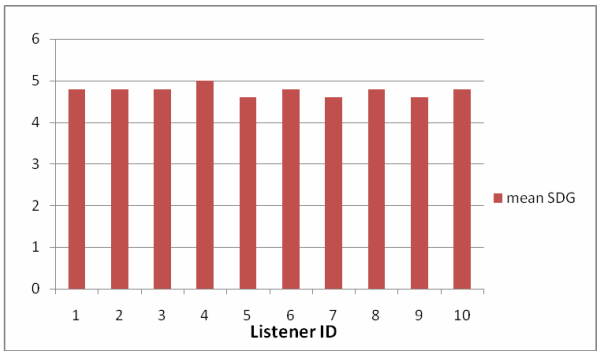

Fig. 7. Mean SDG values of ten clips for ten volunteer licensers

$S'(i)$ is the sample's value of watermarked signal and N is the length of the signal.

The high SNR is achieved due to the fact that not the whole signal is caring embedded data only selected segments will carry hidden information.

6.1.2 Listening test

The listening test is considered as a way to evaluate the watermarking algorithm subjectively [3]. Ten watermarked segments with ten volunteers and five degrees of evaluation (*imperceptible 5, barely perceptible 4, slightly annoying 3, annoying 2, and very annoying 1*) called the "Subjective Difference Grade" (SDG measurement), explained in Fig. 7. Most of our watermarked segments were imperceptible and the mean value of our subjective evaluation was more than 4.76, which is slightly higher than Al-Haj [5] who claims high imperceptibility for his algorithm with SDG 4.73.

6.2 Robustness test against attacks

The evaluation of our scheme depends on how robust is the algorithm towards signal manipulations intentional and non-intentional attacks like resampling, echo addition, and noise addition. Proposed algorithm shows high robustness against most of the common attacks.

Table 1. Robustness evaluation against various attacks

Attack	BER [16]	BER [3]	BER [5]	BER (Proposed)
Resampling 8KHZ	0.0586	0		0
Resampling 11.25 KHZ	0	0	0.000173	0
Resampling 16 KHZ	-	-	0.000591	0
Resampling 22.05 KHZ	0	0	0.000128	0
Resampling 32 KHZ	-	-	-	0
Resampling 48 KHz	-	-	-	0
Resampling 88.2KHZ	-	-	-	0
Resampling 96 KHZ	-	-	-	0
Requantization	0	0	0	0
Equalization	0.0039	0.0027	-	0
Amplitude scaling 110%	0.2363	0.158	-	0
Amplitude scaling 90%	0.2549	0.162	-	0
LPF 400HZ	-	-	-	0
LPF 4KHZ	0	0	-	0
LPF 8KHZ	-	-	0.001893	0
Jittering (1/100000)	0	0	>0.5	0
Jittering (1/50000)	0.0215	0.186	>0.5	0
Noise addition 30dB	-	-	0	0.053
Noise addition 25dB	-	-	0	0.082
Echo addition	0.0078	0.002	-	0.0088
MP3 Compressing 256	0	0	-	0.361
MP3 Compressing 128	0	0	-	0.343
Pitch shifting +1°	0.5166	0.064	>0.5	0. 425
Pitch shifting -1°	0.501	0.059	>0.5	0. 398
Zero inserting	-	-	-	0

Table 1 shows how the embedded data survived after attacks with comparison of two of new robust algorithms that was introduced in the last few years. The most common attacks as in [3, 5, 16] were tested are presented in Table 1. The robustness is measured according to BER which is calculated as in Eq. (10):

$$BER = \frac{B_e}{B_t} \tag{10}$$

where *Be* is the number of error bits and *Bt* is the number of total bits embedded.

The types of attacks used to evaluate the robustness of the algorithm are as follows:

1. Re-sampling : the signal is resampled to other frequencies, as shown in the table below, and then back to original frequency 44 KHz.
2. Re-quantization: the signal is quantized from 16 to 8bit and the embedded data was directly extracted without re-quantizing back.
3. Equalization: using an audio tool called the audacity the signal quantized according to default settings by -12,-14 dB consecutively.
4. Amplitude scaling: the values of the samples were amplified to 110% and attenuated to 90%.
5. Low pass filtering: the signal was filtered using a low pass filter with cut-off frequency 4 KHz and 400 Hz consecutively.
6. Jittering: removing a sample with small rates (1/50000) and (1/100000) consecutively.
7. Additive noise: a Gaussian signal with SNR 20 dB was added to the signal.
8. Echo addition: the signal was echoed with 10ms delay and 10% decay of the original signal.
9. MP3 compression: the signal is compressed with rates 128 and 256 consecutively using MPG1 layer III compression.
10. Pitch shifting: it is the most difficult attack because it will cause the frequency fluctuation. The attack shifted the pitch 1° higher and 1° lower.
11. Zero inserting: normally zero samples are inserted in silent periods to replace silent points to avoid perceptible noise. The selective behavior of precluding silent periods makes the algorithm immune to such attack. On the other hand, inserting one zero point or two in random places of the signal is checked out for evaluation.

6.3 Capacity of the algorithm

Although proposed method is selective where a segment is to be selected for embedding, the achieved payload can be considered high as the acceptable bit rate is just 20 bps. The achieved bit rate in the proposed method could reach 200 b/s for a high energy audio file with regular distribution energy. However, a mean value of 176.4 bps was achieved for different classical audio clips, considering an embedding rate of 250 sample/bit at least in stereo channel audio files.

7. Conclusion

This study proposed an efficient audio watermarking to overcome the limitations of previous work in satisfying all the three important factors of watermarking namely imperceptibility, robustness and capacity. A reasonable capacity has been achieved while keeping the imperceptibility and robustness at high levels.

The proposed selective behavior of audio segments guaranteed the robustness to high extent so that the proposed algorithm is able to resist common attacks as well as de-synchronization attacks. This is because the proposed algorithm selects the segments that have relatively high energy compared to the unselected one. On the other hand, choosing the mean value quantization for embedding could achieve high imperceptibility in addition to its immunity towards attacks. This adds extra robustness to the algorithm as well as increasing the security of the watermark by keeping the quantization index as a secret key. Moreover, the synchronization code was chosen to be easily hidden and extracted in order to increase the rapidity of algorithm. The combination of these techniques together with hiding data using the SWT transform have given a competitive results compared with most recent works.

The proposed algorithm has been tested and evaluated using typical common parameters to prove the fulfillment of imperceptibility, robustness and high capacity. The imperceptibility of the proposed algorithm could outstrip the performance of the most imperceptible algorithms which achieved SDG of 4.76, which is higher than the recent comparable audio watermarking algorithms. In addition, the robustness is high and the achieved results show that the proposed algorithm is robust towards most of common attacks and synchronization attacks. The high data payload obtained, practically 176.4 bps; can put the proposed algorithm in front of many recent works. In conclusion, this method can be used for copyright protection as well as for content verification. Hence, from the obtained results, it can be noted that the objectives of the proposed algorithm had been satisfied.

For future research, capacity can be further increased. A new embedding process may be developed, to resist the attacks of synchronization in order to reduce the number of samples per one bit embedding.

References

[1] Obimbo, C., & Salami, B. (2012), *Using Digital Watermarking for Copyright Protection*. INTECH Open Access Publisher. - Volume 2, Dr. Mithun Das Gupta (Ed.), ISBN: 978-953-51-0619-7, InTech, DOI: 10.5772/38184.

[2] Lin, Y., & Abdulla, W., "Objective quality measures for perceptual evaluation in digital audio watermarking," *IET signal processing*, vol. 5, no. 7, pp. 623-631, 2011.

[3] Lei, B., Soon, Y., Zhou, F., Li, Z., & Lei, H., "A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition," *Signal Processing*, vol. 92, no. 9, pp. 1985-2001, 2012.

[4] Zamani, M., & Manaf, A. B. A., "Genetic algorithm for fragile audio watermarking," *Telecom-munication Systems*, vol. 59, no. 3, pp. 291-304, 2015.

[5] Al-Haj, A., "An imperceptible and robust audio watermarking algorithm," *EURASIP Journal on*

Audio, Speech, and Music Processing, *2014*(1), pp. 1-12, 2014.
[6] Nematollahi, M. A., Akhaee, M. A., Al-Haddad, S. A. R., & Gamboa-Rosales, H., "Semi-fragile digital speech watermarking for online speaker recognition," *EURASIP Journal on Audio, Speech, and Music Processing*, *2015*(1), pp. 1-1, 2015.
[7] Xiang, S., Huang, J., & Yang, R., "Time-scale invariant audio watermarking based on the statistical features in time domain," *In Information Hiding*. Springer Berlin Heidelberg, pp. 93-108, July 2006.
[8] Xiang, S., & Huang, J., "Histogram-based audio watermarking against time-scale modification and cropping attacks," *Multimedia, IEEE Transactions on*, vol. 9, no. 7, pp. 1357-1372, 2007.
[9] Xiang, S., Kim, H. J., & Huang, J., "Audio water-marking robust against time-scale modification and MP3 compression," *Signal Processing*, vol. 88, no. 10, pp. 2372-2387, 2008.
[10] Yang, H. Y., Bao, D. W., Wang, X. Y., & Niu, P. P., "A robust content based audio watermarking using UDWT and invariant histogram," *Multimedia Tools and Applications*, vol. 57, no. 3, pp. 453-476, 2012.
[11] Fan, M. Q., & Wang, H. X., "Statistical characteristic-based robust audio watermarking for resolving playback speed modification," *Digital Signal Processing*, vol. 21, no. 1, pp. 110-117, 2011.
[12] Chen, S. T., Hsu, C. Y., & Huang, H. N., "Wavelet-domain audio watermarking using optimal modify-cation on low-frequency amplitude," *Signal Processing, IET*, vol. 9, no. 2, pp. 166-176, 2015.
[13] Wang, X. Y., Shi, Q. L., Wang, S. M., & Yang, H. Y., "A Blind Robust Digital Watermarking Using Invariant Exponent Moments," *AEU-International Journal of Electronics and Communications*, vol. 70, no. 4, pp. 416-426, 2016.
[14] Lin, Y., & Abdulla, W. H., *Audio Watermark*. Springer Science & Business Media, 2015.
[15] Li, J., & Wu, T., "Robust audio watermarking scheme via QIM of correlation coefficients using LWT and QR decomposition," *IEEE. In Informa-tive and Cybernetics for Computational Social Systems* (*ICCSS*), *2015 International Conference on*, pp. 1-6, August 2015.
[16] Debnath, L. (Ed.), Wavelets and signal processing. Springer Science & Business Media, 2012.
[17] Lei, B. Y., Soon, Y., & Li, Z., "Blind and robust audio watermarking scheme based on SVD-DCT," *Signal Processing*, vol. 91, no. 8, pp. 1973-1984, 2011.
[18] Li, W., Xue, X., & Lu, P., "Localized audio watermarking technique robust against time-scale modification," *Multimedia, IEEE Transactions on*, vol. 8, no. 1, pp. 60-69, 2006.
[19] Zhongda, T., Shujiang, L., Yanhong, W., & Yi, S., A prediction method based on wavelet transform and multiple models fusion for chaotic time series. Chaos, Solitons & Fractals, 98, pp. 158-172, 2017.

Baydaa Mohammad Mushgil received her BSc from University of Mosul in 2005 in computer engineering. Worked at University of Baghdad as an engineer at Alkawarizmi College of Engineering since 2008. Received her MSc. In computer and embedded systems from Universiti Putra Malaysia. Her main areas of research interest are signal processing and Information Security.

Wan Azizun Adnan received her BSc (Hons) in Mathematics in 1984 from Southampton University, England and obtained her Master and PhD from Universiti Malaya in 1996 and 2010 respectively. In 1999, she joined Faculty of Engineering, Universiti Putra Malaysia as a Lecturer, and in 2004 became a Senior Lecturer. She is currently an Associate Professor at the same University. Her main areas of research interest are Biometrics, Information Security and Engineering Education.

S.A. R Al-Haddad is a PhD graduated in Electrical, Electronic and Systems Engineering from National University Malaysia. His specialized in Human and Animal sound Processing, Al-Quran Sound Processing, Media Security and Biometric. Lecturer in Department of Computer and Communications Systems Engineering, Universiti Putra Malaysia since 1997 and promoted as Associate Professor in and graduate and managed to get International and national grants. Further than that, he has few patents and copyrights and actively join professional society such as Malaysia IEEE SMC as Senior Member and Vice Chair and National Islamic Calligraphy as Chairman of IT.

Sharifah Mumtazah Syed Ahmad is an Associate Professor in Faculty of Engineering of Universiti Putra Malaysia UPM. She graduated with her PhD in 2004 from University of Kent, United Kingdom. Her area of expertise is biometrics, information security, and intelligent systems. She has currently published over 70 publications in international journals and conference proceedings. She has also drafted 2 patents and numerous copyrights and published materials in newspaper articles.