American Journal of Computation, Communication and Control
2014; 1(2): 42-49
Published online August 30, 2014 (http://www.aascit.org/journal/ajccc)

Keywords
Copyright Protection,
Digital Audio Watermarking,
Discrete Wavelets Transform,
Quantization Index Modulation,
Singular Value Decomposition,
Synchronization Code

Received: August 20, 2014
Revised: August 27, 2014
Accepted: August 28, 2014

Self-synchronized digital audio watermarking using discrete wavelets transform and singular value decomposition

M. Firouzmand[1], R. Aghaei[2]

[1]Department of Electrical & Information Technology, Iranian Research Organization for Science & Technology (IROST), Tehran, Iran
[2]E-Commerce Department of Nooretouba University, Tehran, Iran

Email address
firouzmand@irost.org (M. Firouzmand), rasoulaghaei@gmail.com (R. Aghaei)

Citation
M. Firouzmand, R. Aghaei. Self-Synchronized Digital Audio Watermarking Using Discrete Wavelets Transform and Singular Value Decomposition. *American Journal of Computation, Communication and Control*. Vol. 1, No. 2, 2014, pp. 42-49.

Abstract
A self-synchronizing, robust, and inaudible audio watermarking algorithm has been proposed in this paper. The synchronization codes are embedded into the audio along with informative data, to enable self-synchronization capability of embedded data. The algorithm effectiveness brought by successive application of two powerful mathematical transforms: discrete wavelets transform (DWT) and singular value decomposition (SVD), which are then followed by quantization index modulation (QIM). The watermark can be blindly extracted without the knowledge of the original audio signal. Experimental results showed the proposed scheme is inaudible and accurate with high watermark payload (about 220 bps). It is also robust against several attacks such as common signal processing and desynchronization. Performance analysis demonstrates low error probability rates too.

1. Introduction

The advent of the digital age of the Internet revolution has made it extremely convenient for users to access, create, process, copy, or exchange multimedia data, leads to an urgent need to protect intellectual property in digital media. Digital watermarking is one such technology being developed, to embed copyright information into the host in a way that is robust to variety of intentional or unintentional attacks.

Most of the proposed watermarking algorithms over the last few years focused on digital images and video sequences. Recently, audio watermarking has become an issue of significant interest to the research community. A comprehensive survey on audio watermarking can be found in [1]. Compared to image and video watermarking techniques, embedding additional bits in an audio signal is a considerably more difficult task. This challenge arises as audio signals are represented by a much smaller number of samples per time interval compared to images and video. This indicates that the amount of information could be embedded robustly and imperceptibly in an audio media is much lower than the visual media. Moreover, the human auditory system (HAS) is much more sensitive than the human visual system (HVS), implying realization of imperceptibility for audio signals is much more difficult than realizing invisibility for images.

According to IFPI (International Federation of the Phonographic Industry) [2], audio watermarking should meet the following requirements: 1) the watermark shouldn't degrade perception of audio, 2) the algorithm should offer more than 20 dB SNR for watermarked audio versus original audio and 20 bps (bits per-second) data payload for watermark, 3) the watermark should resist most common audio processing operations and attacks, such as D/A and A/D conversions, temporal scaling (stretching by ± 10 %), additive and multiplicative noise corruption, and MP3 compression, 4) the watermark should be capable of preventing unauthorized detection, removal, and embedding, unless the quality of audio becomes very poor. These requirements present great challenges to robust audio watermarking.

Attacks against audio watermarking systems, have become more sophisticated by audio watermarking technology [3,4]. In general, the attacks on audio watermarking systems can be categorized in 2 types. First is common signal processing attacks such as MP3 compression, low-pass filtering, and noise addition. Second are desynchronization attacks such as amplitude variation, pitch shifting, and time-scale modification.

While the common signal processing reduces watermark energy, desynchronization attacks induce synchronization errors between the original and the extracted watermark during the detection process. Most of the previous audio watermarking schemes have shown robustness against common signal processing attacks, and only a few specialized watermarking methods have addressed the desynchronization attacks [5,6].

Seok, et al.[7],introduces an audio watermarking scheme by exploiting the human perceptual characteristics of the audio signal to regulate the embedding strength, but it isn't very robust to some audio signal processing such as re-sampling, re-quantization and compression. The current self-synchronization algorithm can't extract feature points steadily. Besides, it usually requires large number of threshold values which make it more difficult to be applied [8,9]. Girin, et al. [10], proposed a speech signal watermarking using the sinusoidal model base on amplitudes, phases and digital frequencies modulation of the partials. Lee, et al. [11], developed an audio watermarking algorithm through modification of ton al maskers, but these methods suffer from poor robustness against time-scale modification and pitch shifting. Wu, et al.[12],presented a self-synchronized audio watermarking algorithm employing Quantization Index Modulation (QIM). The synchronization code and the watermark data are embedded into the low-frequency sub-band in the Discrete Wavelets Transform (DWT) domain. Huang, et al.[13], choose Bark code which has better self-relativity as synchronization mark to embed it into temporal and Discrete Cosine Transform (DCT) domain sequentially. Du, et al.[14], considered an improved audio watermark detection algorithm based on HAS. It is possible to resist desynchronization attack by utilizing the advanced synchronization code technique, but the existing audio watermarking approaches have short comings as follows: 1) they choose a12-bit Barker code which is rather short and thus making it is easy to cause false synchronization, 2) they are vulnerable to re-sampling and jittering, and very few researchers have performed and published sufficient experiments involving amplitude variation, pitch shifting, time-scale modification and etc. 3) they do not completely exploit human auditory masking effects, which influences the imperceptibility and robustness capabilities of watermarking. Wang, et al.[15], proposed a digital audio watermarking algorithm based on the DWT. The watermark information is embedded into low-middle-frequency wavelet coefficients. A watermark detection scheme using linear predictive coding (LPC) also presented, which does not require the knowledge of original audio signal during watermark extraction. Chang, et al.[16],proposed a DWT-based counter-propagation neural network (CPN). The watermark embedding and extracting procedures are integrated into the proposed CPN. Li, et al.[17],suggested an audio watermarking method in which the embedding and detection regions are determined by content analysis of the music. Xiang, et al.[18], proposed a multi bit audio watermarking method based on two statistical features: the histogram shape and the modified mean value in the time domain. Moreover he developed another histogram based audio watermarking scheme, which watermark is inserted by shaping the histogram after the DWT [19]. Recently, Fan, et al.[20], introduced a novel audio watermarking scheme based on discrete fractional sine transform (DFRST). Experimental results show that the aforementioned audio watermarking methods have difficulty in obtaining favorable trade-offs among imperceptibility, robustness and data payload.

In order to address the above challenges, we proposed an audio watermarking scheme which is robust against desynchronization attacks by utilizing the powerful transforms and synchronization code technique. We choose 16-bit Barker code as synchronization mark, and embed it by QIM in DWT domain.

We have embedded the synchronization codes in audio so that the hidden data have self-synchronization capability. Synchronization codes and informative bits are both embedded by applying QIM on Singular Value Decomposition (SVD) of low frequency sub-band coefficients in DWT domain to achieve strong robustness against common signal processing procedures, noise corruption, and attacks. By exploiting the time-frequency localization capability of DWT, the proposed technique reduces computational load of searching synchronization codes dramatically and thus resolves the contending requirements between the robustness and low computational complexity.

The experimental results show that the watermark is robust against signal processing and attacks, such as

Gaussian noise corruption, resampling, requantization, cropping and MP3 compression.

The rest of this paper is organized as follows: Section 2 presents fundamental theory and synchronization, Section 3 and 4 introduces the outline of the proposed algorithm, followed by detailed descriptions. Experimental results are shown and discussed in Section 5. Finally, conclusions are drawn in Section 6.

2. Fundamental Theory and Synchronization

2.1. Fundamental Theory

In the present paper, the watermark can be embedded into the host audio in three steps. First, the host audio is segmented and DWT is performed on each segment. Second, low frequency components of each audio segment are separated into two parts, arrays are formed and then SVD is performed on each array. Finally, synchronization code is embedded into the first part and the watermark bit is embedded into the second part. The embedding model has been shown in Figure 1.

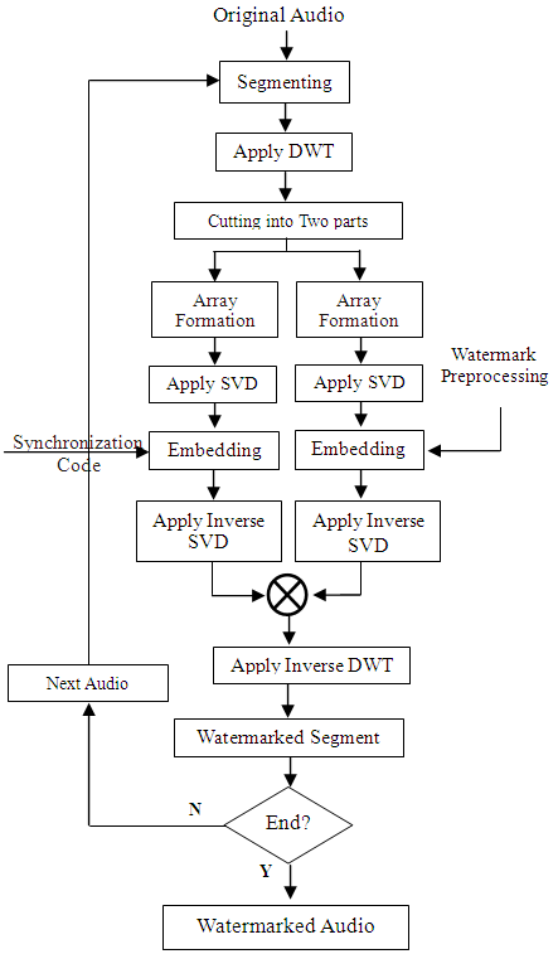

Figure 1. *Watermark embedding procedure*

In data extraction, the same DWT separates segment of incoming audio data, into two parts and then extracts binary data from SVD of DWT coefficients at low frequency sub-band. A search for synchronization codes in the first part is done. This procedure needs to be repeated by shifting the selected segment one sample at a time until a synchronization code is found. While the position of a synchronization code determined, the hidden information bits, which follow the synchronization code could be extracted. The extraction model has been shown in Figure 2.

There are several advantages for applying DWT to audio watermarking: 1) DWT is known to have the time-frequency localization capability, 2) variable decomposition levels are available. 3) DWT needs a lower computation load compared with DCT and DFT. Specifically, suppose there are N samples in an audio section, the computation load are$O(L.N)$ for DWT and$O(N.\log_2 N)$ for DCT, and DFT, where L is the length of the wavelet filter.

2.2. Synchronization Code

Synchronization is one of the key issues of audio watermarking. Watermark detection starts by alignment of watermarked block with detector. Time-scale or frequency-scale modification makes the detector lose synchronization, which causes false detection. So we need an exact synchronization algorithms based on robust synchronization code. Generally, we should avoid false synchronization during synchronization code selection. Several reasons contribute to false synchronization: 1) the style of the synchronization code, 2) the length of synchronization code, 3) the probability of “0” and “1” in synchronization code.

Among these, the length of synchronization code is particularly important. The longer it is, the more robust it becomes. We employ Barker code in front of the watermark to locate the position where the watermark is embedded.

Barker codes, which are subsets of Pseudo Number (PN) sequences, are commonly used for frame synchronization in digital communication systems. Barker codes maintain low correlation side lobes which are the correlation between a code word with a time-shifted version of itself. The correlation side lobe C_k, for a k-symbol shift of an N-bit code sequence X_j is given by (1):

$$C_k = \sum_{j=1}^{N-k} X_j X_{j+k} \quad (1)$$

Where $X_j, j = 1,2,\dots,N$ are individual code symbol taking values +1 or −1, and the adjacent symbols are assumed to be zero.

3. Watermark Embedding Procedure

Let $X = \{x(i), 0 \le i < Length\}$ represent a host digital audio signal with *Length* samples. Furthermore, $W = \{w(i,j), 0 \le i < M, 0 \le j < N\}$, represents a binary image to be embedded within the host audio signal and $F = \{f(i), 0 \le i < Lsyn\}$ stands for a synchronization code with *Lsyn*bits, where $f(i) \in \{0, 1\}$.

The main steps of the embedding procedure are described in details follows.

3.1. Watermark Preprocessing

To dispel the pixel space relationship of the binary watermark image, and improve whole digital watermark system robustness, watermarks scrambling algorithm have been used at first. In proposed watermark embedding scheme, the binary watermark image is scrambled from W to $W1$ by using Arnold transform, where:

$$W_1 = \{w_1(i,j), 0 \le i < M, 0 \le j < N\} \quad (2)$$

Then, it is transformed into a one-dimensional sequence of ones and zeros as follows:

$$W_2 = \{w_2(k) = w_1(i,j), 0 \le i < M,\ 0 \le j < N, k = i \times N + j\ , w_2(k) \in \{0,1\}\} \quad (3)$$

Finally, each bit of the watermark data is mapped into an antipodal sequence using binary phase-shift keying (BPSK) modulation according to the following equations:

$$W_3 = \{w_3(k) = 1 - 2 \times w_2(k),\ k = 0,1,\dots\ , M \times N - 1, w_3(k) \in \{-1\ ,1\}\} \quad (4)$$

3.2. Watermark Embedding

1) To improve the robustness of the proposed scheme against cropping, time-scale modification and jittering is done. Moreover, to make detector available when it loses synchronization, audio signal X segmenting is used at first.

Each segment includes $2^h \times n \times L_1 + 2^h \times m \times L_2$ samples, where L_1 and $L_2$are the length of synchronization code and watermark respectively. Constant n is chosen to be 5 samples in our experiment, m is the length of the array that forms at the next stage and h represents h-level of DWT.

2) h-level DWT is performed on each segment and then low-frequency coefficients are cut into two parts with $2^h \times n \times L_1$ and $2^h \times m \times L_2$ samples, respectively.

3) Part I and Part II are partitioned into arrays with n and m samples respectively.

4) SVD is performed on each array. (SVD is performed on the part I too).

5) According to Barker code bits, values of $S_{syn}(1,1)$ of singular values are quantized by using (5):

$$S'_{syn}(1,1) = \begin{cases} \lfloor S_{syn}(1,1)/\Delta \rfloor \times \Delta + 3\Delta/4 & if syn(i) = 1 \\ \lfloor S_{syn}(1,1)/\Delta \rfloor \times \Delta + \Delta/4 & if syn(i) = 0 \end{cases} \quad (5)$$

6) According to watermark bit, values of $S(1,1)$ of singular values are quantized by using (6):

$$S'(1,1) = \begin{cases} \lfloor S(1,1)/\Delta \rfloor \times \Delta + 3\Delta/4 & if\ w_3(i) = 1 \\ \lfloor S(1,1)/\Delta \rfloor \times \Delta + \Delta/4 & if\ w_3(i) = 0 \end{cases} \quad (6)$$

Where$\lfloor . \rfloor$indicates the floor function and Δ denotes the embedding strength (or quantization step size). The value of Δ should be as large as possible under the imperceptibility constraint.

7) After replacement of $S_{syn}(1,1)$ by $S'_{syn}(1,1)$ and $S(1,1)$ by $S'(1,1)$, the modified array is obtained by applying an inverse SVD to the modified SVs.

8) After embedding all watermark bits into each array, arrays are arranged and part I and Part II are reconstructed.

9) Then part I and part II are merged together, IDWT is performed and watermarked audio is obtained.

3.3. Robustness Improvement

In order to increase robustness against desynchronization attacks, the proposed scheme performs the same procedure as in Section 3.2 to embed synchronization code and digital watermark into every audio segment.

4. Watermark Extraction Procedure

The proposed scheme neither requires the original audio signal nor any other side information to extract the watermark. The procedure is illustrated in the block diagram shown in Figure 2. A detailed description of the algorithm follows.

1) Perform steps 1 and 2 of the embedding procedure until part I and part II are formed.

2) Before extracting the watermark, it is needed to synchronization codes be searched. By using (7), existence of synchronization code is checked in part I and go to step 3. If nothing found, segmentation should be performed again.

$$syn'(i) = \begin{cases} 1 & if S'_{syn}(1,1) - \lfloor S_{syn}(1,1)/\Delta \rfloor \times \Delta \ge s/2 \\ 0 & if S'_{syn}(1,1) - \lfloor S_{syn}(1,1)/\Delta \rfloor \times \Delta < s/2 \end{cases} \quad (7)$$

3) Part II is divided into arrays with m samples.

4) SVD is applied to each array and watermark bit is extracted using the rule below:

$$\mathrm{W'}_3(\mathrm{i}) = \begin{cases} 1 & if s'(1,1) - \lfloor s(1,1)/\Delta \rfloor \times \Delta \ge s/2 \\ 0 & if s'(1,1) - \lfloor s(1,1)/\Delta \rfloor \times \Delta < s/2 \end{cases} \quad (8)$$

5) The watermark bits are determined based on BPSK demodulation:

$$W'_2 = \{w_2(k) = (1 - \mathrm{w'}_3(k))/2, k = 0,1,\dots, M \times N - 1, w'_2(k) \in \{0\ ,1\}\} \quad (9)$$

6) All the detected watermark bits W'_2 are rearranged to form the binary watermark image W'_1.

7) Finally, the watermark image W^* can be obtained by descramblingW'_1.

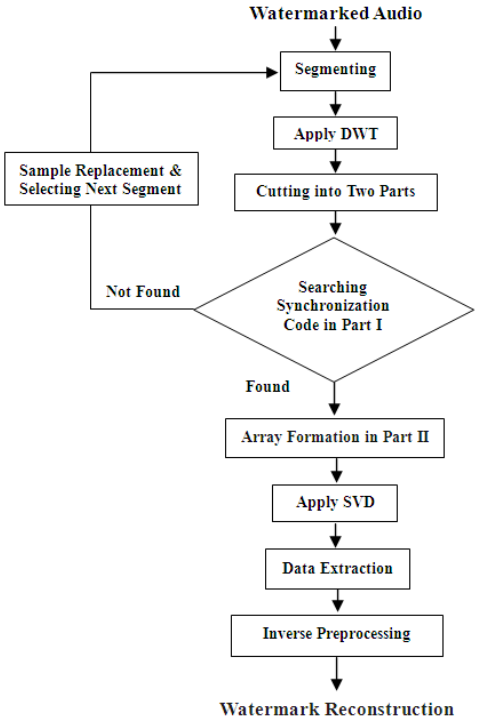

Figure 2. *Watermark extraction procedure*

5. Experimental Results

Experiments are performed in Adobe Audition 3.0 and MATLAB 7.10. Classical/Pop music and speech audio clips were used to evaluate performance of the proposed algorithm. These three audio types have different perceptual properties, characteristics and energy distribution, and thus their performances may vary from one type to another.

Each such audio signal is a 16-bit, mono file in the WAVE format and has 44.1 kHz sampling rate. We use a 30×30 bits binary image as our watermark shown in Figure3 and a 16-bit Barker code 1111100110101110 as synchronization code. Haar wavelet is applied with two decomposition levels. Array size m is 50 and the range of quantization step size Δ starts from0.15 for speech audio and goes up to 0.6 for pop audio signal.

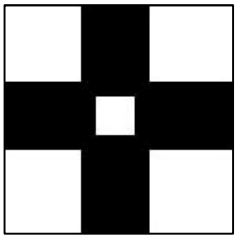

Figure 3. *Binary watermark*

Performance of audio watermarking algorithms is usually evaluated with respect to imperceptibility (inaudibility), robustness and payload. In what follows, we give a brief description of each metric.

5.1. Imperceptibility Test

Imperceptibility is related to the perceptual quality of embedded watermark data within the original audio signal. It ensures that the quality of the signal is not perceivably distorted and the watermark is imperceptible to a listener.

To measure imperceptibility, we use signal to noise ratio (SNR) as an objective measure, and a listening test as a subjective measure.

SNR is a statistical difference metric which is used to measure the similitude between the undistorted original audio signal and the distorted watermarked audio signal**.** The SNR is computed according to (10):

$$SNR = -10log_{10}\left[\frac{\sum_i f_i^2}{\left(\sum_i (\acute{f}_i - f_i)^2\right)}\right] \quad (10)$$

Where f *and* f'corresponds to the original audio signal and watermarked audio signal respectively.

Although SNR is a simple way to measure the noise introduced by the embedded watermark and can give a general idea of imperceptibility, it does not take into account the specific characteristics of the human auditory system. Therefore, we also employed the Perceptual Audio Quality Measure (PAQM) [21].

PAQM derives an estimate of the signals on the cochlea and compares the representation of the reference signal with that of the signal under test. It has been shown that the correlation between PAQM and the mean opinion score (MOS) is 0.98 [22]. Therefore, in our experiments the PAQM scores will be mapped to the grading scale of MOS which is shown in Table 1.

A listening (hearing) test was actually performed with four listeners to estimate the subjective MOS grade of the watermarked signals. Each listener was presented with the pairs of original signal and the watermarked signal and was asked to report whether any difference could be detected between the two signals. All people listened to each pair for almost 5 times, and they gave a grade for the pair. The average grade for of each pair from all listeners corresponds to the final grade for the pair. Table 2 lists the corresponding SNR values, along with MOS grades obtained by conducting the listening test.

Table 1. *MOS grading scale*

MOS Grade	Description
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very Annoying

Table 2. *SNR and MOS values for different audio*

Audio signal	SNR	MOS
Pop	25.95 dB	4.9
Classic	24.62 dB	4.7
Speech	39.14 dB	4.3

5.2. Robustness Test

Normalized correlation (NC) is used to evaluate the correlation between the extracted and the original watermark and is given by (11):

$$NC(w,w') = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} w(i,j)w'(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} w^2(i,j)}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} w'^2(i,j)}} \quad (11)$$

Where W and W' are the original and the extracted watermarks, respectively, and i, j are indices in the binary watermark image. If $NC(\mathrm{w},\mathrm{w}')$ be close to 1, then the correlation between W and W' is very high and if it is closed to zero, the correlation between W and W' is very low.

The bit error rate (BER) is used to measure the robustness of our scheme:

$$BER(w,w') = \frac{\text{Number of error bits}}{\text{Number of total bits}} \times 100\% \quad (12)$$

Various attacks are used to estimate the robustness of our proposed watermarking algorithm. According to their influence on synchronization, attacks can be divided into common audio signal processing and desynchronization attacks. Common audio signal processing may distort the perceptual quality, but not affect the synchronization structure, including requantization, re-sampling, additive noise, low-pass filtering, echo addition, equalization, MPEG compression. Desynchronization attacks introduce very little distortion to the watermarked audio, but destroy the synchronization needed by most existing audio watermarking algorithms, including random cropping, amplitude variation, pitch shifting, time-scale modification, jittering.

The following signal processing attacks are performed to assess the robustness of our scheme.

(A) Additive white Gaussian noise (AWGN): White Gaussian noise is added to the watermarked signal until the resulting signal has an SNR of 20 dB.

(B) Resampling: The watermarked signal, originally sampled at 44.1 kHz, is re sampled at 22.05 kHz, and then restored back by sampling again at 44.1 kHz.

(C) Low-pass filtering: A second-order Butterworth filter with cutoff frequency11,025 Hz is used.

(D) Requantization: The 16-bit watermarked audio signal is re-quantized down to 8 bits/sample and then back to 16 bits/sample.

(E) MP3 compression 128 kbps: The MPEG-1 layer-3 compression is applied. The watermarked audio signal is compressed at the bit rate of 128 kbps and then decompressed back to the WAVE format.

(F) MP3 compression 64 kbps: The MPEG-1 layer 3 compression is applied. The watermarked audio signal is compressed at the bit rate of 64 kbps and then decompressed back to the WAVE format.

(G) Echo addition: An echo signal with a delay of 98 ms and a decay of 41% is added to the watermarked audio signal.

(H) Denoising: The watermarked audio signal is denoised by using the "Automatic click remover" function of Adobe Audition 3.0.

(I) Invert: inverts all sample values in time domain (phase shift 180°).

(J) Amplitude variation: The watermarked signal was attenuated up to 150%and down to50%.

(K) Random cropping: In our experiment, 10% samples were cropped at each of three randomly selected positions (front, middle and back).

(L) Pitch shifting: Tempo-preserved pitch shifting is a difficult attack for audio watermarking algorithms, because it causes frequency fluctuation. In our experiment, the pitch is shifted one degree higher and one degree lower.

Experimental results for watermarks extracted after application of the various attacks on the classical, pop and speech audio signals are shown in given in Table 3.

5.3. Payload

The data payload refers to the number of bits that can be embedded into the audio signal within a unit of time and is measured in the unit of bps (bits per second) and denoted by B. Suppose that the sampling rate of audio is R (Hz), the number of wavelet decomposition levels is k and each array include m member. Then the data payload of this algorithm can be shown as:

$$B = {}^{R}/_{m \times 2^{k}}\text{bps} \quad (13)$$

The data payload of our scheme is *220 bps*.

Table 3. *Experimental results for classical, pop and speech audio signals*

Attack	NC	BER	Extracted Watermark	NC	BER	Extracted Watermark	NC	BER	Extracted Watermark
No Attack	1	0		1	0		1	0	
AWGN	1	0		1	0		0.9996	0.004	
Re-sampling	1	0		1	0		1	0	

Attack	NC	BER	Extracted Watermark	NC	BER	Extracted Watermark	NC	BER	Extracted Watermark
Re-quantization	1	0		1	0		1	0	
Low-pass filtering	1	0		1	0		1	0	
MP3 128 kbps	1	0		1	0		1	0	
MP3 64 kbps	1	0		0.9978	0.025		0.9985	0.016	
Invert	1	0		1	0		1	0	
Echo addition	1	0		1	0		1	0	
Denoising	1	0		0.9946	0.003		0.9992	0.008	
Amplitude Variation	1	0		1	0		1	0	
Random Cropping	1	0		1	0		1	0	
Pitch shifting	1	0		1	0		1	0	

5.4. Comparison

It is not straight forward to compare our algorithm with other proposed methods due to the differences of audio samples, watermark imperceptions, data payload and so on. Never the less a general comparison between our method and two competing methods [12, 23] that have high performance and payload is given in Table 4.

Our comparison is based on reported results of published methods and it is given for data payload, noise addition, resampling, low-pass filtering, cropping and MP3 compression.

In view of the comparison in Table 4, our proposed watermarking algorithm achieves high embedding capacity and low BER against attacks, such as noise addition, resampling, low-pass filtering and MP3 compression. The performance of our algorithm can be further improved by reducing the data payload (increasing wavelet decomposition level or increasing length of array) and then increasing embedding strength on the premise of the same imperceptions constrain.

We also test the performance of the proposed algorithm with different orthogonal wavelet bases, including Daubechies, Coiflets, and Symlets wavelets. The observation is that the choice of different wavelet bases has little effect on the performance of the proposed algorithm. Thus, we exploit the simplest wavelet base, Haar wavelet.

6. Conclusion

In this paper, we proposed a self-synchronized audio watermarking technique base on DWT and SVD. We have further demonstrated the robustness of this digital audio watermarking algorithm against desynchronization attacks. The robustness of the method is based on three key components of our approach: the original digital audio is segmented and the DWT is performed on each segment. Then the low frequency sub-band of each segment is cut into two parts, arrays are formed and then synchronization code and watermark are embedded into the first part and second part respectively by applying SVD on each array. Synchronization codes and watermarks are embedded into SVD of low-frequency sub-band in DWT domain, thus achieving good robust performance against common signal processing procedure and noise corruption. The time-frequency localization capability of DWT is exploited to improve the efficiency in searching for the synchronization codes.

Table 4. *Comparison of audio watermarking algorithms*

Algorithm	Our	In [12]	In [23]
Payload (bps)	220	172	196
Noise addition (BER)	0 (20dB)	0 (20 dB)	0 (20dB)
Cropping & Shifting (Robust)	Yes	Yes	No
MP3 64kbps (BER)	0.025	0.0434	0.01
Low-pass filtering (BER)	0	0	0
Re-sampling(BER)	0	0	0

Analytical and experimental findings show that the proposed watermarking method achieves robustness against both common audio signal processing and desynchronization attacks. In addition, the watermark can be extracted without the knowledge of the original digital audio signal and can be easily implemented. Moreover, the proposed scheme achieves low error probability rates. We have compared the performance of our algorithm with other recently proposed audio watermarking algorithms. Overall, our method has the high embedding capacity and achieves low BER against attacks, such as noise addition, re-sampling, low-pass filtering, and MP3 compression.

References

[1] Cvejic N, Seppanen T.Digital audio watermarking techniques and technologies: applications and benchmarks.*United States of America & United Kingdom: Information Science Reference (an imprint of IGI Global)*. 2008.

[2] Katzenbeisser S, Petitcolas FAP. Information Hiding Techniquesfor Steganography and Digital Watermarking. *Artech House*. 2000.

[3] Cox IJ, Miller ML.The first 50 years of electronic watermarking.*J. Appl. Signal Process*. 2002; 56 (2): 225–230.

[4] Barni M, Cox IJ, Kalker T.Digital watermarking, in: Fourth International Workshop, International *Workshop on Digital Watermarking, Siena, Italy, Lecture Notes in Computer Science*. 2005;3710: 260–274.

[5] Wen-Nung L, Li-Chun C.Robust and high-quality time-domain audio watermarking subject to psychoacoustic masking, *Proceedings of the IEEE International Symposium on Circuits and Systems*. 2002; 45–48.

[6] Megías D, Herrera-joancomartíJ, Minguillón J.A robust audio watermarking scheme based on MPEG 1 layer III compression, Communications and Multimedia Security—CMS.*Lecture Notes in Computer Science*. 2003; 963: 226–238.

[7] Seok J, Hong J, Kim J.A novel audio watermarking algorithm for copyright protection of digital audio.*ETRI J*. 2002;24 (3):181–189.

[8] Wu CP, SuPC,JayKuo CC.Robust audio watermarking for copyright protection, *Proceedings of the SPIE*.1999; 3807: 387–397.

[9] Li W, Xue XY.Audio watermarking based on music content analysis: robust against time scale modification.*Proceedings of the Second International Workshop on Digital Watermarking*.2003; 289–300.

[10] GirinL,Marchand S.Watermarking of speech signals using the sinusoidal model and frequency modulation of the partials.*IEEE International Conference on Acoustics, Speech, and Signal processing (ICASSP 2004)*. 2004; 633–636.

[11] Lee HS, Lee WS.Audio watermarking through modification of tonal maskers.*ETRI J*. 2005; 27 (5): 608–661.

[12] Wu SQ, HuangJW, Shi YQ.Efficiently self-synchronized audio watermarking for assured audio data transmission. *IEEE Trans Broadcast*. 2005;51(1): 69–76.

[13] Huang JW, Wang Y, Shi YQ.A blind audio watermarking algorithm with self-synchronization.*Proceedings of the IEEE International Symposium Circuits and Systems*. 2002;3: 627–630.

[14] Du CT, Wang RD.An audio watermarking detecting improved algorithm based on HAS. *J. Eng. Grap*. 2005; 26 (1): 74–79.

[15] Wang R, Xu D, ChenJ,Du C.Digital audio watermarking algorithm based on linear predictive coding in wavelet domain. *7th IEEE International conference on signal processing, (ICSP'04)*. 2004;1:2393–2396.

[16] Chang CY, Shen WC, Wang HJ. Using counter-propagation neural network for robustdigital audio watermarking in DWT domain. *IEEE International conference on systems, manand cybernetics, (SMC'06)*. 2006;2: 1214–1219.

[17] Li W, Xue X, Lu P.Localized audio watermarking technique robust against time-scale modification.*IEEE Trans Multimedia*. 2006;8(1):60–69.

[18] Xiang S, Huang J.Histogram-based audio watermarking against time-scale modification and cropping attacks. *IEEE Trans Multimedia*.2007; 9(7):1357–1372.

[19] XiangS,Kim H, Huang J.Audio watermarking robust against time-scale modification and MP3 compression. *Signal Process*. 2008;88(10):2372–2387.

[20] FanM,Wang H.Chaos-based discrete fractional Sine transform domain audio watermarking scheme. *ComputElectr Eng*.2009; 35(3):506–516.

[21] ArnoldM,Wolthusen S, Schmucker M. Techniques and applications of digital watermarking and content protection. *Artech House*. 2003.

[22] BeerendsJ,Stemerdink J.A perceptual audio quality measurement based on a psychoacoustic sound representation. *Journal of the Audio Engineering Society*.1992; 40(12): 963–972.

[23] VivekanandaKB,Sengupta I, Das A.An audio watermarking scheme using singular value decomposition and dither-modulation quantization. *Springer, Multimedia Tools and Applications*. 2010; 52(2 – 3):369–383.