

Robust audio watermarking algorithm based on DWT using Fibonacci numbers

Ali Akbar Attari¹ · Ali Asghar Beheshti Shirazi¹

Received: 26 February 2017 / Revised: 28 December 2017 / Accepted: 15 February 2018 /

Published online: 28 March 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract This paper presents a blind and robust audio watermarking algorithm developed based on Fibonacci numbers properties and the discrete wavelet transform (DWT) advantages. The method embeds watermark bits in the 6th level approximation subband of DWT at which there is less sensitivity of the human auditory system. The key idea is dividing the 6th level approximation coefficients into small frames and modifying their magnitude based on Fibonacci numbers and watermark bit values. The proposed watermarking method demonstrates a superior robustness against different common attacks (i.e., Gaussian noise addition, Low-pass filter, Resampling, Requantizing, MP3 compression, Amplitude scaling, Echo addition, Time shift, and Cropping). Compared to recently developed methods, the proposed algorithm is much more robust against the most common attacks with capacity as high as 686 bits per second. The results of PEAQ testing verify the quality of watermarked audio signal without significant perceptual distortion. The algorithm allows flexibility in audio watermark algorithm to achieve a balance between robustness and imperceptibility while the capacity is maintained constant by choosing various kinds of sequence.

Keywords Blind audio watermarking · Fibonacci numbers · Discrete wavelet transform · Human auditory masking

1 Introduction

The rapid development of computer and communication networks makes that the content of multimedia (i.e., sound, image, and video) can be easily generated, stored and distributed.

✉ Ali Asghar Beheshti Shirazi
abeheshti@iust.ac.ir

Ali Akbar Attari
A_Attari@vu.iust.ac.ir

¹ School of Electrical Engineering, Iran University of Science & Technology (IUST), Narmak, Tehran 16844, Iran

However, illegal application of multimedia has become widespread. Therefore, protection against intelligent offending is an important matter. In recent years, digital watermarking has been proposed as an approach to prevent multimedia against intelligent abusing. Originally, watermarking is a data-hiding technique inserting the owner's private information in multimedia that can be used for subsequence extraction in order to perform copyright protection, content authentication, fingerprint and broadcast monitoring [14].

For audio watermarking methods to be effective, they must satisfy basic requirements which are mentioned as follows:

- **Imperceptibility:** The quality of watermarked audio signal must be acceptable regarding the objective and subjective tests.
- **Robustness:** After encountering various signal processing and manipulating attacks, the watermarks must still be extractable.
- **Capacity:** The number of bits can embed without losing imperceptibility so called as capacity. The capacity is measured in bits per second.

These criteria are shown in the corner of the magic triangle illustrated in Fig. 1. The watermarking methods could not satisfy three requirements, simultaneously, because there is a balance among the audio watermarking requirements [17]. In another way, the watermarking methods can be classified into two main groups, blind and non-blind methods. While retrieving watermark may not need the original signal, the algorithm will be called a *blind algorithm*. Otherwise, it will be called an *aware or non-blind algorithm* [10]. In general, blind watermarking methods have more attraction in application scenarios [12].

Usually, the audio watermarking methods are categorized in the domains of time and frequency. In the time domain [5, 28], there are some methods, i.e., Spread spectrum, Echo hiding, Least significant bit modification, Phase coding and modulation, and Patchwork based. In the frequency domain, some methods such as Discrete Wavelet Transform (DWT) [9, 14, 20], Fast Fourier Transform (FFT) [10, 23], Discrete Cosine Transform (DCT) [17, 30], Cpestrum [13, 22], and Singular Value Decomposition (SVD) [2, 7, 15] can be used [14, 26]. Ordinarily, the time domain methods have a lower complexity and are easily implementable, but they demonstrate a low robustness against attacks. On the other hand, the frequency domain approaches have a better robustness and imperceptibility because of using the capability of signal character and human perception [14].

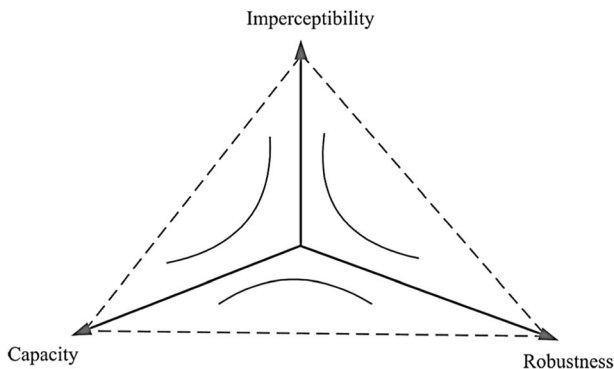


Fig. 1 The magic triangle [17]

There are three main techniques to implement watermarking. In the spread spectrum technique, the data embed different kinds of pseudo-random noise into an original signal in the coder, and subsequently, is extracted based on correlation in the decoder. In Pair Coefficient Modification technique, every watermark can be extractable based on the relation between paired groups. The Quantization Index Modulation (QIM) technique, which was proposed by Chen and Wornell, is one of most practical approaches. The main reason for its prevalence is creating balance between capacity, robustness, and imperceptibility [14].

In [14], Rational Dither Modulation performs watermarking on the vector norms of the 5th approximation coefficients of DWT by using perceptual masking. Embedding is conducted by modulating the coefficient vector using the quantization step while the quantization step is estimated from the previous watermarked vectors. Robustness and capacity are obtained by altering in the vector dimensions, and the imperceptibility is guaranteed by keeping error quantization under the auditory masking threshold. In [4], an algorithm is proposed based on SVD in the DWT domain and watermark is embedded based on QIM technique in the sv of DWT coefficients. Experimental results showed high imperceptibility and high robustness against common attacks. In [20], the authors propose a robust technique based on Lifting Wavelet Transform (LWT) in which watermark bits are embedded using SVD and QIM in the low-frequency coefficients of LWT. The results indicated resistance against common attacks and Stirmark attacks. Also the trade-off between robustness, imperceptibility and capacity was possible. In [29], the authors propose a scheme based on norm vector and DWT for blind watermarking. For more improvement in robustness and imperceptibility, the image as watermark signal is encrypted by Arnold transform and then embedded in a portion of the DWT coefficients by using QIM technique with variable quantization step. In [16], Variable-Dimensional Vector Modulations scheme is presented by norm space DWT to improve watermarking efficiency. Watermarking procedure is performed by a norm vector extracted from the low-frequency DWT coefficients. The strength of embedding can be defined by the level of quantization step, which is adjusted lower than the auditory masking threshold. This scheme provides the balance between capacity and robustness. In [12], one method based on Multilevel DWT (MDWT) is proposed to decompose host signal, and by using the first to ninth detail subbands of DWT, the watermark bits are embedded in the low-frequency DCT components using Rational Dither Modulation (RDM) and Window Vector Modulation (WVM). Also synchronous pattern is inserted into the 11th level approximation coefficients. In [9], one robust and high capacity watermarking algorithm is proposed by using high-frequency coefficients of DWT in which Human Auditory System (HAS) has a lower sensitivity to modification. The key idea is dividing high-frequency coefficients into frames, and embedding is performed by using the average of frame coefficients. Experimental results showed a high capacity without most perceptual distortion as well as robustness against common signal processing attacks.

DWT, because of its benefit in perfect reconstructing and multi-resolution specification, has gained widespread popularity. Another advantage of DWT used in audio watermarking methods is exploitation of human perceptual model. Most of the DWT-based watermarking methods utilize both of DWT and QIM for watermark embedding. However, QIM technique is damageable against some attacks such as Amplitude scaling. Therefore, we use the advantage of Fibonacci numbers with relative steps utilized in [10]. In this paper, we propose the blind watermarking method based on Fibonacci numbers, DWT, and human auditory characteristics. The relative quantization steps provided by Fibonacci numbers not only achieve maximum robustness against most common attacks but also retrieve imperceptibility.

The remainder of this paper is organized as follows: Section 2 presents the background information. In section 3, the proposed algorithm is introduced. Sections 4 and 5 present the experimental results with respect to imperceptibility and robustness, respectively. Finally, conclusion is drawn in Section 7.

2 Fibonacci numbers

The numbers $\{1, 1, 2, 3, 5, 8, 13\}$ are known as the Fibonacci numbers. For some years, mathematicians used to consider them because of their applications. In this paper, we use this sequence type to embed and extract watermarks. Fibonacci numbers were obtained by Eq. (1):

$$F_n = \begin{cases} 0, & \text{if } n < 1, \\ 1, & \text{if } n = 1, \\ F_{n-1} + F_{n-2} & \text{if } n > 1. \end{cases} \quad (1)$$

They have very interesting features. One of the most important of these features, which was also used in this study, is the relation between two successive Fibonacci numbers:

$$F_n = F_{n-1} + F_{n-2}; \quad (2)$$

$$\frac{F_n}{F_{n-1}} = \frac{F_{n-1} + F_{n-2}}{F_{n-1}} = 1 + \frac{F_{n-2}}{F_{n-1}} = 1 + \frac{1}{\frac{F_{n-1}}{F_{n-2}}}; \quad (3)$$

$$\lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{\frac{F_{n-1}}{F_{n-2}}} \right) = 1 + \frac{1}{\lim_{n \rightarrow \infty} \frac{F_{n-1}}{F_{n-2}}}; \quad (4)$$

$$\text{if } \lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}} = \varphi; \varphi = 1 + \frac{1}{\varphi}; \quad (5)$$

$$\varphi^2 - \varphi - 1 = 0; \quad (6)$$

$$\varphi = \frac{1 \pm \sqrt{5}}{2}. \quad (7)$$

If φ is positive, then $\varphi = 1.618$. In fact, φ is the Golden Ratio. The Golden Ratio is an irrational number because it is an answer of the polynomial equation. Golden Ratio is named after introducing Rectangular Golden with sides having Golden Ratio. Every Fibonacci

number can be shown by Golden Ratio. Eq. (8) demonstrates how Golden Ratio generates each Fibonacci number. $\bar{\varphi}$ is a negative form of Eq. (7) [10]:

$$\varphi = \frac{\varphi^n - \bar{\varphi}^n}{\sqrt{5}} \quad (8)$$

In the original Fibonacci set, there are two ‘1’s that we removed one of them in the proposed algorithm. In addition, in order that Fibonacci numbers become applicable in the proposed algorithm, we must add some elements between ‘0’ and ‘1’. Furthermore, it is worth to introduce $F = \{0, 0.25, 0.5, 1, 2, 3, 5, 8, 13, \dots\}$ as a new set of Fibonacci numbers.

3 Proposed scheme

Extensive works have been performed for the characterization of HAS in recent years. Figure 2 illustrates the intensity of sound to HAS response in the frequency scale. The absolute threshold is the lowest sound that can be audible by human’s ear, which depends on frequency. Sound pressure level (SPL) is measured by dB. Decibels have a logarithmic scale, such that each 6 dB increase in SPL doubles the intensity of the sound. The perceived loudness of sound has a relationship with sound intensity. According to Fig. 2, a human can hear the sound with frequency in the range of 20 Hz - 20 kHz. The human’s ear sensitivity is high in the range of 3–4 kHz and decline in lower and higher frequencies [10].

Among the used transforms, DWT has more benefits in audio processing such as multi-resolution and logarithmic subband frequency decomposition, showing a high degree of similarity to human perception [9]. In other words, DWT represents more frequency resolution in low-frequency range compared to DCT and DFT, which have equal frequency resolutions in all bands. Even DFT does not present good frequency resolution in comparison with DCT. Moreover, DWT not only has lower computational complexity but also has sharper transition edge of basis function in comparison with DCT and DFT [1]. Consequently, DWT has the lowest leakage among the subbands, which leads to better imperceptibility of watermarking.

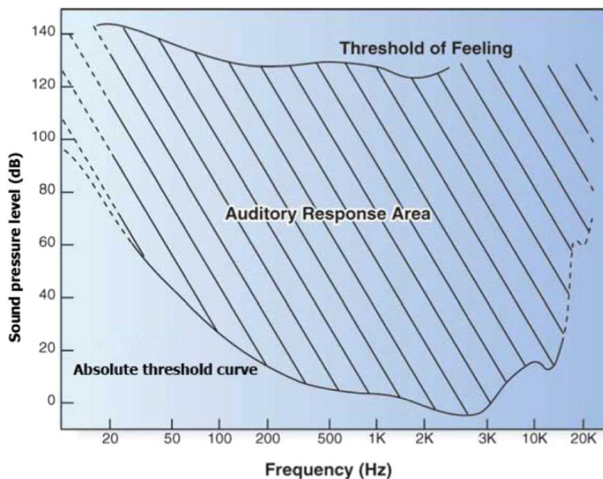


Fig. 2 Typical absolute threshold curve of the human auditory response [10]

In this paper, DWT is used for embedding watermarks because of the mentioned advantages. The 6th level approximation coefficients are proper for embedding. Due to the low sensitivity of human’s ear to low frequency as well as its high robustness against some attacks such as Mp3 compression and low-pass filter, we use the 6th level approximation coefficients equal to the frequency range of 0–344 Hz for watermark embedding. As demonstrated in Fig. 3, DWT coefficients can be presented in different levels. The horizontal axis is the decomposition level of DWT, and the vertical axis is frequency band in Hz. The 6th level approximation coefficients equals to 0–Fs/128. This band of frequency is proper for embedding, because human’s ear has low sensitivity to low frequency.

DWT separates input signal into two subbands, approximation and detail, and decimates the output of each decomposition level. The decomposition continues on the approximation part until the desired decomposition is obtained. DWT coefficients can be extracted based on Eqs. (9) and (10) as follows:

$$C^L(i) = \sum_{j=0}^{J_w-1} h(j-2i)C^{L-1}(j) \tag{9}$$

$$D^L(i) = \sum_{j=0}^{J_w-1} g(j-2i)D^{L-1}(j) \tag{10}$$

where, $C^L(i)$ and $D^L(i)$ illustrate the L^{th} level approximation and detail coefficients of DWT, respectively; $h(j)$ and $g(j)$ define the low- and high-pass filters of orthogonal wavelet, respectively; and J_w is the length of the wavelet filters. In the proposed algorithm, Daubechies-4 filter is used as a wavelet function of DWT. Most of audio signal decompositions use Daubechies filters [14, 15, 29] because of proper sharpness of frequency characteristics, causing better audio band decomposition [11], and consequently, better imperceptibility of audio watermarking.

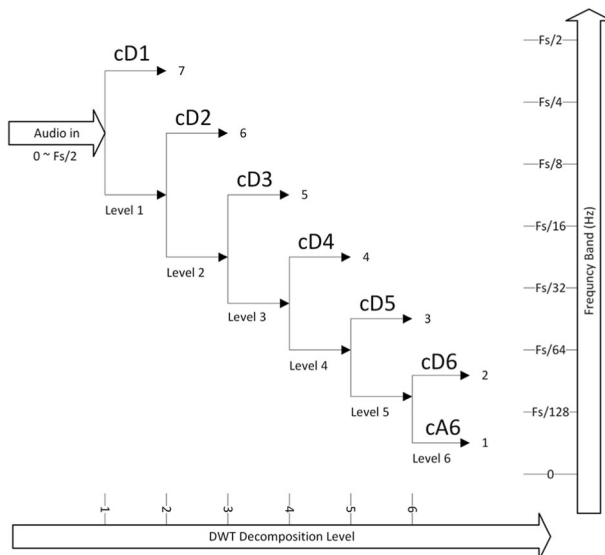


Fig. 3 DWT analysis

4 Embedding algorithm

Firstly, DWT is used to decompose the audio signal as a host signal into multiple subbands; then the 6th level approximation coefficients are modified based on Fibonacci numbers and watermark bits. Finally, the inverse DWT for regeneration of audio signal is used. Embedding steps are represented as follows:

- Step 1. Applying DWT for giving the 6th level approximation subband. We can use the frame with 30 s length.
- Step 2. Dividing the 6th level approximation coefficients of DWT into subframes with the length of d , which is used for repetitive embedding of unique watermark bits in them.
- Step 3. Finding the largest Fibonacci numbers $\{fib_{n,i}\}$, n^{th} Fibonacci numbers for i^{th} DWT coefficients, which is lower than or equal to the magnitude of i^{th} approximation coefficient in the current frame, $\{f_i\}$. $\{fib_{n,i}\}$ is determined based on Eq. (11):

$$fib_{n,i} \leq f_i \quad (11)$$

- Step 4. Obtaining the watermarked coefficients $\{f'_i\}$ by Eq. (12):

$$f'_i = \begin{cases} fib_{n,i}, & \text{if } n \bmod 2 = 0 \text{ and } w_l = 0, \\ fib_{n+1,i}, & \text{if } n \bmod 2 = 1 \text{ and } w_l = 0, \\ fib_{n+1,i}, & \text{if } n \bmod 2 = 0 \text{ and } w_l = 1, \\ fib_{n,i}, & \text{if } n \bmod 2 = 1 \text{ and } w_l = 1, \end{cases} \quad (12)$$

Where, $l = \lfloor i/d \rfloor + 1$, w_l is the l^{th} watermark bit, and $\lfloor \cdot \rfloor$ denotes the floor function. Each watermark bit embeds in the subframe repetitively; in other words, each subframe denotes a unique watermark bit.

- Step 5. Using inverse DWT for achieving the watermarked audio.

The advantage of Fibonacci numbers is that the distance between zeroes and ones automatically adapts to the magnitude of the DWT coefficient to be modified. The distance is adapted by taking into account the magnitude of the coefficient to be altered. This is one advantage of the proposed watermarking method to obtain a balance between imperceptibility and robustness, which is not given in QIM. In QIM, the distance between zeroes and ones, which equals to the quantization step, is uniform [3].

Figure 4 illustrates a flowchart of the embedding algorithm. Sequence type and subframe length (d) are two adjustable parameters to maintain the trade-off between robustness, imperceptibility, and capacity. We will illustrate the effect of adjustable parameters in the Experimental Results section.

5 Extraction algorithm

The parameter of extraction, i.e. the subframe length (d), can be sent by secure channel for the decoder. The extraction procedure shown in Fig. 5 can be summarized in the following steps:

- Step 1. Applying DWT to calculate the 6th level approximation coefficients of watermarked audio.
- Step 2. Dividing the 6th level approximation coefficients into the subframe length (d).

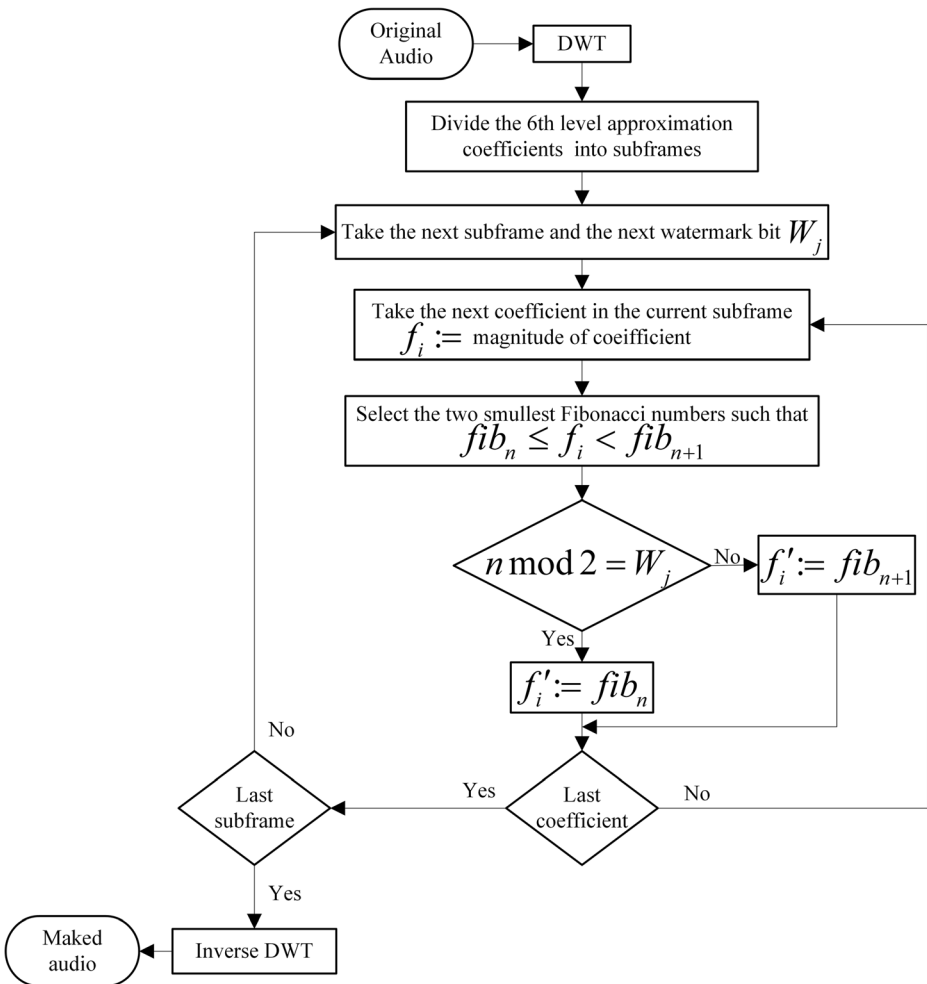


Fig. 4 Flowchart of the embedding algorithm

Step 3. Selecting the closest Fibonacci number to approximation coefficients for each coefficient in the current subframe; if DWT coefficients have an equal distance from two Fibonacci numbers, we will choose the lowest one:

$$B_i = \begin{cases} 0, & \text{if } n \bmod 2 = 0, \\ 1, & \text{if } n \bmod 2 = 1. \end{cases} \tag{13}$$

To extract each watermark bit, each sample must be checked whether ‘0’ or ‘1’ has been embedded. Based on the examination of the whole coefficients of the current subframe, the watermark can be extracted. The watermark will be extracted based on Eq. (13). B_i is the extracted watermark from each sample. After giving whole information, the extracted watermark bit can be extracted based on voting technique. According to this rule, if the number of zero samples extracted is larger than or equal to half of the subframe’s length (d), the extracted watermark bit will be ‘0’; otherwise, it will be ‘1’. For instance, if the subframe’s length is 7 and we have received three ‘0’s and four ‘1’s, then the extracted watermark bit will be ‘1’.

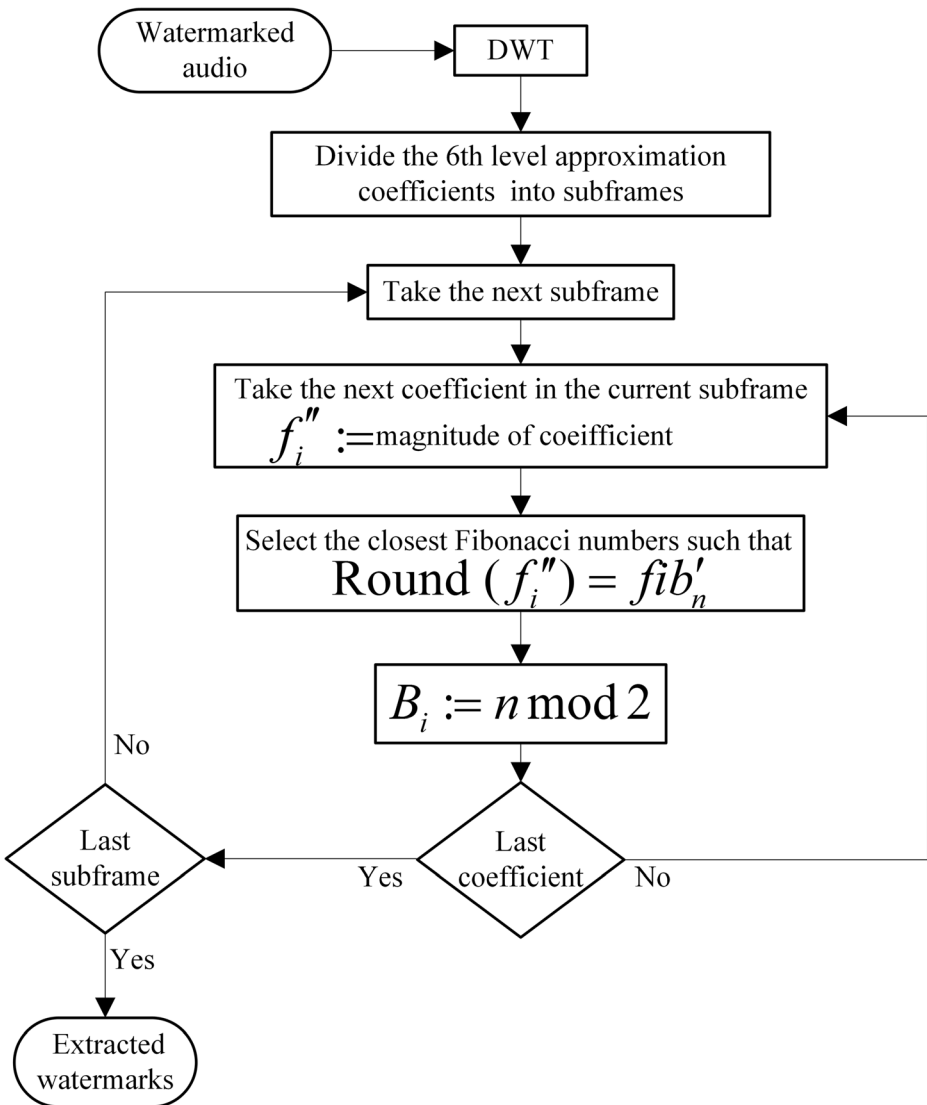


Fig. 5 Flowchart of the extracting algorithm

6 Algorithm specification

In the QIM technique (based on Dither Modulation), firstly, the host signal samples are quantized according to quantization step-size and then the residuals, $\frac{1}{4}$ and $\frac{3}{4}$ quantization step-size, are added to the quantized host signal according to the inserted watermark bit ‘0’ or ‘1’, respectively. In the extracting phase, the watermarked signal is re-quantized using the same quantization step-size and then the quantization residual is achieved. If it is larger than $\frac{1}{2}$ quantization step-size, the extracted watermark is ‘1’; otherwise, it is ‘0’. A large

quantization step-size causes to either decrease of imperceptibility or increase of robustness, and vice versa [31]. In the QIM method, the step-size quantization is constant; then the residual compared to the small magnitude of host signal samples is large, while it is small compared to larger magnitude of host signal samples. Then the modification error is affected by scale of relative residual. However, in Fibonacci numbers, the distance between two successive numbers depends to Golden Ratio. The key idea of using Fibonacci numbers is to preserve the modification error in the determined limit. The maximum distortion with assumption $r_n = \frac{F_n}{F_{n-1}}$, represented in [10], will be denoted in the following:

- (1) If the magnitude of approximation coefficient converts to F_{n+1} :

$$\text{max error rate} = \frac{\text{max error}}{F_{n+1}} = \frac{(r_n-1)F_n}{F_{n+1}} = \frac{(r_n-1)F_n}{r_n F_n} = \frac{r_n-1}{r_n} \tag{14}$$

- (2) If the magnitude of approximation coefficient converts to F_n :

$$\text{max error rate} = \frac{\text{max error}}{F_n} r_n^{-1}. \tag{15}$$

$r_3 = 2, r_4 = 2, r_5 = 2, r_6 = 1.5, r_7 = 1.66, r_8 = 1.6, r_9 = 1.625, \dots$

Furthermore, by assuming typical value, $r_n = 1.61$, the maximum error rate will be in the range of 0.36–0.61. If the magnitude of the coefficient is lower than 3, r_n will be equal to 2 or 1.5. In these cases, maximum errors for the worst cases are between 0.66 to 1; the maximum distortion is often lower than 50% of the magnitude, and it is rarely higher than 50% of the magnitude of coefficients, in which n has a low value. Then the average of maximum error rate will be 50% of the magnitude. If coefficients have a uniform distribution, the mean error rate will be 25% [10].

Use of Fibonacci numbers is not a unique method to get an ‘exponential-like spacing’ of the watermarking coefficients. We can use other sequences instead of Fibonacci numbers. To get the different sequence, we can use Eq. (16), where k is a relation between the two successive numbers.

$$F_n = \lfloor k^n \rfloor, \text{ for } n = 1, 2, \dots, \tag{16}$$

Where,

$$\frac{F_{n+1}}{F_n} \approx k, \text{ for } n = 1, 2, \dots, \tag{17}$$

By choosing different k values, we can generate various sequences.

- (1) If $|k| < 1$, the generated sequence is lower than 1; then it is not suitable for the watermarking system.
- (2) If $|k| > 2$, both the growth rate as well as modification error are high; then the generated sequence is not suitable for the watermarking system.

- (3) If $1 < |k| < 2$, the generated sequence is proper for the watermarking application. Table 1 shows the sequences generated based on various k values which are categorized into five types. The sequence constructed by $k = 1.5, 1.7$ is close to Fibonacci numbers value. If $k = \varphi$, the generated sequence is approximately equal to the Fibonacci numbers' value. Various sequences present a trade-off between robustness and imperceptibility. By increasing the distance between the numbers (k), not only the imperceptibility of watermarked audio decreases but the robustness increases too. Also by decreasing the distance between numbers (k) the imperceptibility will increase and robustness will decrease [10].

7 Synchronization

One of the most important things in audio watermarking is the synchronization problem. Most of watermarking methods have used the whole of the signal for embedding watermarks, and subsequently, the whole of the signal has been used to extract watermarks. Although this technique is simple and usable, it is not flexible. The watermarking algorithm, which uses synchronization, provides the possibility of watermark extraction by even a portion of watermarked audio. On the other hand, in the technique that uses framing for extraction synchronization causes the location of the embedded bits to be found easily with the lowest error [27].

In this paper, before embedding, 39-bit Barker codes as synchronization bits are added to watermarks, periodically. Figure 6 shows the data structure of watermark bits. Therefore, the synchronization bits are embedded at the beginning of each audio frame. The main reason of utilizing Barker code is that it has a sharp autocorrelation with a high main loop as well as a low side loop.

8 Discussion

To evaluate the performance and applicability of the proposed algorithm, we used album *Rust* by No, Really [24], from which eight audio files of Pop music were selected. These audio files are sampled at 44.1 kHz, 16 bits quantization per sample, and have two channels. The original audio clips have an MP3 format, which were changed to WAV for processing.

The proposed algorithm embeds watermarks in the 6th level approximation coefficients attained from the audio signal. The audio signal is decomposed by using the 4-coefficient Daubechies wavelet. Figure 7 shows the waveform of the original and watermarked signals for ‘Go’ clip.

Table 1 Sequence with different k values

Type	Sequence	1	2	3	4	5	6	7	8	9	a
1	$F_n(k = 1.3)$	1	2	3	4	6	8	10	13	17	23
2	$F_n(k = 1.5)$	1	2	3	5	7	11	17	25	38	57
3	Fibonacci	1	2	3	5	8	13	21	34	55	89
4	$F_n(k = 1.7)$	1	2	4	8	14	24	41	69	118	201
5	$F_n(k = 1.9)$	1	3	6	13	24	47	89	169	322	613

In the proposed algorithm, there is one adjustable parameter for applying the trade-off among watermarking requirements (i.e. robustness, imperceptibility, and capacity). By increasing the subframe’s length (d), robustness against attacks increases, and capacity decreases. We assumed the subframe’s length to be equal to $d = 7$ for evaluation in this paper.

For watermarked audio quality measuring, we can use objective and subjective tests. In subjective test, the original signal and watermarked signals are played for a group of human audience, and they are asked to distinguish the original audio from the watermarked audio. The watermarked audio will be degraded by Subjective Difference Grade (SDG) test in comparison with the original signal. Because of the complexity and non-sufficiently of SDG test, we used the Objective Difference Grade (ODG) as an objective test of SDG [21]. ODG can get a mark from -4 to 0 . If the ODG value is ‘0’, there is no quality degradation, and if it is ‘-4’, the added distortion is very annoying. An output value of ITU-R BS.1387 PEAQ standard is the ODG value. Table 2 describes the various values of ODG [17].

ODG is a perceptual test by using human model. Another objective test for measuring the modification error between the original signal ($s(n)$) and the watermarked signal ($\check{s}(n)$) is Signal to Noise Ratio (SNR), which is calculated by Eq. (18):

$$SNR = 10\log_{10} \left(\frac{\sum_n s^2(n)}{\sum_n (s(n) - \check{s}(n))^2} \right) \tag{18}$$

In this paper, we intend to evaluate imperceptibility with SNR and ODG tests. For ODG test, we use PEAQ codes, which are implemented by the TSP Lab of McGill University [18]. Table 3 shows the imperceptibility evaluation results for various audio clips with presuming embedded watermarks by different types. As shown in Table 3, the measured ODG ranges from -0.34 to -1.60 . For most of the results, the measured ODG is larger than -1 , which demonstrates watermarked audio with high quality. On the other hand, SNR shows the amount of the distortion added to the original audio during audio watermarking. However, because of less sensitivity of human’s ear in the low-frequency range, the added distortion is less audible.

One of the most important criteria to evaluate the proposed algorithm and to compare it with different watermarking algorithms is robustness against various attacks. Basically, attacks are introduced based on the proposed algorithm applications. In the decoder, watermarks are extracted from the watermarked audio after applying the attacks. For robustness evaluation, we must calculate the difference between the watermark extracted and the watermark inserted, which is measured as Bit Error Rate (BER).

We used the BER criterion, which is obtained by Eq. (19). To calculate BER, the number of the wrongly extracted watermark bits is divided by the number of inserted watermark bits:

$$BER = \frac{\text{Number of Error Watermarks}}{\text{Number of Total Watermarks}} \times 100\% \tag{19}$$

The attacks involved Gaussian noise addition, Resampling, Requantizing, Amplitude scaling, Low-pass filter, Echo addition, MP3 compression, Time shift, and Cropping. Table 4

Synchronization code	Watermark
-------------------------	-----------

Fig. 6 Watermark data structure

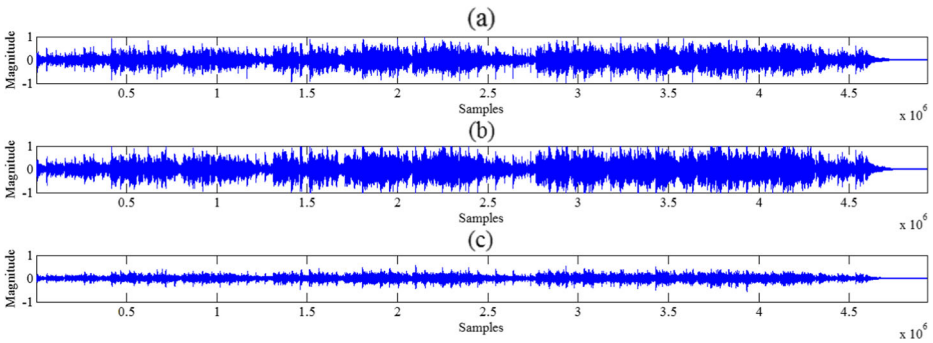


Fig. 7 a Original audio, b Watermarked audio, c Difference between original and watermarked audios

outlines the detail of common attacks. In this paper, we used the attacks implemented by MATLAB 2014b. The Mp3 coder was implemented by LAME [8] at 64 kbps and 128 kbps bit rate.

As shown in Table 5, after encountering most of the different attacks, BERs will be nearly zero. In spite of encountering attacks, the watermark bits will be extractable. The proposed watermarking method has superior robustness against Low-pass filter with a cut of frequency (4 kHz). Because the watermark bits are embedded in the 6th level approximation coefficients, which are not eliminated after applying Low-pass filter attacks. In addition, the proposed algorithm is resistant to MP3 compression at 128 Kbps and 64Kbps because of the same reasons. MP3 compression at 128 Kbps changes or even eliminates the audio signal frequency coefficients higher than about 11 kHz sampled at 44.1 kHz [6]. Furthermore, although Mp3 compression (64 bps) and Low-pass filter (4 kHz) change or eliminate high-frequency coefficients, these attacks could not affect watermark extraction since the watermarks embedded in the frequency coefficients are lower than 344 Hz. Of course, in echo addition with 5% decay and 50 ms delay, the results demonstrate high robustness.

Because of embedding the watermarks in the magnitude of coefficients, the proposed algorithm is not completely robust against the amplitude scaling of 85% in comparison with the other attacks, because the watermarks are embedded by modification of the magnitude of coefficients affecting some attacks that manipulate the amplitude of the audio signal. Table 5 demonstrates that the proposed algorithm is robust against removing the first 200 samples of the watermarked signal. It is achievable by embedding 39-bits Barker code as synchronization bits at the first of each frame.

Choosing various sequences can alter the balance between imperceptibility (ODG or SNR) and robustness against common attacks. In Fig. 8, each point states the imperceptibility and

Table 2 egradation scale of the ODG [17]

ODG	Degradation
0	Imperceptible
-1	Perceptible and not annoying
-2	Slightly annoying
-3	Annoying
-4	Very annoying

Table 3 Imperceptibility test results

Audio track	Type	ODG	SNR (dB)	Audio track	Type	ODG	SNR (dB)
Beginning of the End	1	-0.67	14.85	Breathing On Another Planet	1	-0.66	15.00
Beginning of the End	2	-0.77	13.99	Breathing On Another Planet	2	-0.74	14.59
Beginning of the End	3	-0.78	13.98	Breathing On Another Planet	3	-0.75	14.55
Beginning of the End	4	-1.32	11.31	Breathing On Another Planet	4	-1.32	12.21
Beginning of the End	5	-1.43	10.63	Breathing On Another Planet	5	-1.60	11.35
Go	1	-0.51	14.26	Floodplain	1	-0.34	15.24
Go	2	-0.52	13.83	Floodplain	2	-0.39	14.86
Go	3	-0.53	13.79	Floodplain	3	-0.39	14.90
Go	4	-0.70	11.80	Floodplain	4	-0.77	12.60
Go	5	-0.80	10.99	Floodplain	5	-0.96	11.64
Stop Payment	1	-0.95	15.42	Citizen, Go Back to Sleep	1	-0.50	17.54
Stop Payment	2	-0.94	15.16	Citizen, Go Back to Sleep	2	-0.61	16.92
Stop Payment	3	-0.97	15.13	Citizen, Go Back to Sleep	3	-0.59	16.97
Stop Payment	4	-1.28	13.07	Citizen, Go Back to Sleep	4	-0.98	14.39
Stop Payment	5	-1.49	12.09	Citizen, Go Back to Sleep	5	-1.21	13.45
Thousand Yard Stare	1	-0.50	14.97	Molten	1	-0.40	15.81
Thousand Yard Stare	2	-0.53	14.62	Molten	2	-0.43	15.57
Thousand Yard Stare	3	-0.53	14.66	Molten	3	-0.43	15.53
Thousand Yard Stare	4	-0.81	12.63	Molten	4	-0.75	13.47
Thousand Yard Stare	5	-1.01	11.69	Molten	5	-0.94	12.50

robustness results under the amplitude scaling of 85%, depending on each sequence type. In the ODG and SNR charts, their value is decreased by increasing k . Also in the Robustness

Table 4 Attack types and specifications

Attack types	Description
A	Zeros-mean Gaussian noise added to the watermarked audio signal to attain SNR = 30.
B	Zeros-mean Gaussian noise added to the watermarked audio signal to attain SNR = 20.
C	Performing down-sampling of the watermarked audio signal to 22,050 Hz and then back to 44,100 Hz.
D	Quantizing the watermarked audio signal to 8 bits/sample and then back to 16bits/sample.
E	Scaling the amplitude of the watermarked audio signal by factor 1.1.
F	Scaling the amplitude of the watermarked audio signal by factor 0.90.
G	Scaling the amplitude of the watermarked audio signal by factor 0.85.
H	Applying a Low-pass filter with 4 KHz cutoff frequency.
I	Adding an echo signal with 50 ms delay and 5% decay.
J	Compressing and decompressing the watermarked audio signal with a MPEG layer 3 coder at 128Kbps bitrate.
K	Compressing and decompressing the watermarked audio signal with a MPEG layer 3 coder at 64Kbps bitrate.
L	Purposely shifting the watermarked audio signal by one sample.
M	Removing the first 200 samples of the watermarked signal.

chart, BER is descended by increasing k excepted in type = 4 ($k = 1.7$) because of distance between the sequences. Fibonacci numbers not only have acceptable result in imperceptibility

Table 5 Robustness test results

Audio Track	Type	No Attack	B	C	D	E	F	G	H	I	K	L	M
Beginning of the End	1	0.00	0.00	0.00	0.00	0.00	0.00	5.82	0.00	0.00	0.00	0.00	0.00
Beginning of the End	2	0.00	0.00	0.00	0.00	0.00	0.00	3.76	0.00	0.01	0.00	0.00	0.00
Beginning of the End	3	0.00	0.00	0.00	0.00	0.00	0.00	3.74	0.00	0.00	0.00	0.00	0.00
Beginning of the End	4	0.00	0.00	0.00	0.00	0.00	0.00	5.48	0.01	0.00	0.01	0.00	0.00
Beginning of the End	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Go	1	0.00	0.00	0.00	0.00	0.00	0.00	4.25	0.00	0.02	0.00	0.00	0.00
Go	2	0.00	0.00	0.00	0.00	0.00	0.00	2.97	0.00	0.00	0.00	0.00	0.00
Go	3	0.00	0.00	0.00	0.00	0.00	0.02	2.92	0.02	0.02	0.02	0.00	0.00
Go	4	0.00	0.00	0.00	0.00	0.00	0.00	4.13	0.00	0.05	0.00	0.00	0.00
Go	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
Stop Payment	1	0.00	0.00	0.00	0.00	0.00	0.00	1.57	0.00	0.00	0.00	0.00	0.00
Stop Payment	2	0.00	0.00	0.00	0.00	0.00	0.00	1.34	0.00	0.00	0.00	0.00	0.00
Stop Payment	3	0.00	0.00	0.00	0.00	0.00	0.00	1.32	0.00	0.00	0.00	0.00	0.00
Stop Payment	4	0.00	0.00	0.00	0.00	0.00	0.00	1.54	0.00	0.01	0.00	0.00	0.00
Stop Payment	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Thousand Yard Stare	1	0.00	0.00	0.00	0.00	0.00	0.00	1.66	0.00	0.00	0.01	0.00	0.00
Thousand Yard Stare	2	0.00	0.00	0.00	0.00	0.00	0.00	1.26	0.00	0.00	0.00	0.00	0.00
Thousand Yard Stare	3	0.00	0.00	0.00	0.00	0.00	0.00	1.32	0.00	0.01	0.01	0.00	0.00
Thousand Yard Stare	4	0.00	0.00	0.00	0.00	0.00	0.00	1.61	0.00	0.01	0.00	0.00	0.00
Thousand Yard Stare	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Breathing On Another Planet	1	0.00	0.00	0.00	0.00	0.00	0.00	3.01	0.00	0.00	0.00	0.00	0.00
Breathing On Another Planet	2	0.00	0.00	0.00	0.00	0.00	0.00	2.38	0.00	0.01	0.01	0.00	0.00
Breathing On Another Planet	3	0.00	0.00	0.00	0.00	0.00	0.00	2.31	0.00	0.01	0.00	0.00	0.00
Breathing On Another Planet	4	0.00	0.00	0.00	0.00	0.00	0.00	2.80	0.00	0.00	0.00	0.00	0.00
Breathing On Another Planet	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Floodplain	1	0.00	0.00	0.00	0.00	0.00	0.00	1.93	0.00	0.00	0.01	0.00	0.00
Floodplain	2	0.00	0.00	0.00	0.00	0.00	0.00	1.51	0.00	0.00	0.02	0.00	0.00
Floodplain	3	0.00	0.00	0.00	0.00	0.00	0.00	1.61	0.00	0.00	0.00	0.00	0.00
Floodplain	4	0.00	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00
Floodplain	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Citizen, Go Back to Sleep	1	0.00	0.00	0.00	0.00	0.00	0.00	4.64	0.00	0.00	0.02	0.00	0.00
Citizen, Go Back to Sleep	2	0.00	0.00	0.00	0.00	0.00	0.00	3.43	0.00	0.00	0.00	0.00	0.00
Citizen, Go Back to Sleep	3	0.00	0.00	0.00	0.00	0.00	0.00	3.28	0.00	0.00	0.00	0.00	0.00
Citizen, Go Back to Sleep	4	0.00	0.00	0.00	0.00	0.00	0.00	4.45	0.00	0.00	0.02	0.00	0.00
Citizen, Go Back to Sleep	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
Molten	1	0.00	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00	0.00	0.00	0.00
Molten	2	0.00	0.00	0.00	0.00	0.00	0.00	1.26	0.00	0.00	0.00	0.00	0.00
Molten	3	0.00	0.00	0.00	0.00	0.00	0.00	1.53	0.00	0.00	0.00	0.00	0.00
Molten	4	0.00	0.00	0.00	0.00	0.00	0.00	1.53	0.00	0.01	0.00	0.00	0.00
Molten	5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

but also they have superior robustness. Figure 8 shows that the best sequence type regarding imperceptibility and robustness belongs to Fibonacci numbers.

Another requirement for evaluating a watermarked system is capacity or the number of watermark bits embedded in the host signal obtained by Eq. (20):

$$Capacity = \frac{N-S}{N} \times (Watermark\ bit\ rate) \tag{20}$$

where, N is the number of watermarks that we can embed in one frame of the host signal, and S is the number of synchronization bits in one frame. In the proposed algorithm, we embed 39-bits (per frame) as synchronization bits. Furthermore, we give 686 bps data payload.

9 Experimental results

For a better evaluation, we compare the results obtained from the proposed algorithm with the results of recently developed algorithms, which are blind. The imperceptibility and capacity results of different schemes are illustrated in Tables 6 and 7. Because of the balance between robustness and imperceptibility presented by Fibonacci numbers (Type = 3), their results are used to compare with the results of other algorithms.

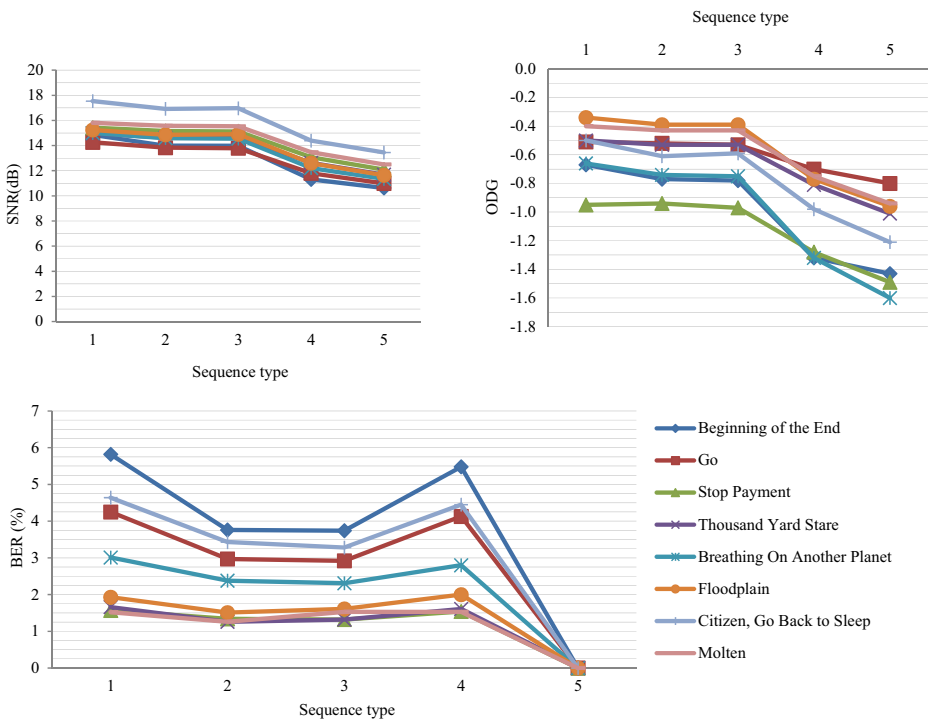


Fig. 8 Imperceptibility (ODG, SNR) and robustness (BER) tests results under the amplitude scaling of 85% for different sequence types

The imperceptibility results are depicted in Table 6. The proposed algorithm has ODG average equal to -0.623 , which is between imperceptible and perceptible but not in annoying status. The standard deviation of ODG (equal to ± 0.197) is small. Also, the average of SNR is around 15 dB with the standard deviation of ± 0.999 dB.

As shown in Table 6, the SS-ICA [25], LWT-SVD [20], DWT-norm [29], and MDWT-DCT-RDM-WVM [12] algorithms present high imperceptibility and the DWT-VDVM [16] and DWT-RDM [14] algorithms well satisfied the perceptual of the watermarked audio. While in the case of SVD-DCT [19] and DWT-SVD [4] algorithms, the perceptual quality of the watermarked audio is barely acceptable. In the SS-ICA algorithm, SNR is maintained deliberately on 20 dB and on the obtained acceptable ODG, since it uses LPC and noise shaping. In the DWT-VDVM and DWT-RDM algorithms, distortion between the original and watermarked audios is indistinguishable because the quantization steps are given based on human auditory properties, and then quantization noise maintains below the auditory masking threshold. In MDWT-DCT-RDM-WVM algorithm, although SNR is decreased due to inserting synchronous pattern, the perceptual quality is maintained because of embedding in a subband to which human hearing is insensitive. The audio watermarking algorithm using SVD-DCT does not have a good performance in imperceptibility due to embedding the watermarks in high-frequency coefficients where human's ear has a high sensitivity. Also the DWT-SVD method, because of embedding the watermark bits in the wide range of frequency coefficients, does not have a high imperceptibility.

Table 7 demonstrates a portion of the robustness results given by the proposed algorithm. Compared with other blind methods, the proposed algorithm proved to have high robustness against Gaussian noise additions (30 dB) and (20 dB), Resampling, Requantizing, Amplitude scaling, Echo addition, and Time shift. For instance, as shown in Fig. 9, although most of the algorithms have worst results in the Amplitude scaling of 85%, the proposed algorithm has a better result in comparison with the DWT-norm, DWT-SVD, and LWT-SVD algorithms. Additionally, the proposed algorithm is robust against Low-pass filter attack with a cut of the frequency of 4 kHz and Mp3 comparison attack changing or removing the high-frequency coefficients of the watermarked audio, as in the proposed algorithm, the watermark bits are embedded in the frequency coefficients lower than 344 Hz.

Table 6 Comparison of different schemes' imperceptibility and capacity

Watermarking schemes	SNR (dB)	ODG	Payload (bps)
SS-ICA	19.994 [± 0.026]	$-0.763[\pm 0.925]$	43.07
SVD-DCT	29.732 [± 2.249]	$-1.593[\pm 1.485]$	43.07
DWT-SVD	26.426 [± 2.536]	$-1.036[\pm 1.430]$	45.56
DWT-norm	24.201 [± 2.252]	$-0.594[\pm 1.122]$	102.40
LWT-SVD	24.686 [± 2.747]	$-0.381[\pm 0.971]$	170.76
DWT-VDVM	20.307 [± 0.304]	$-0.143[\pm 0.125]$	301.46
DWT-RDM	20.213 [± 0.633]	$-0.068[\pm 0.081]$	344.53
MDWT-DCT-RDM-WVM	15.954 [± 1.996]	$-0.360[\pm 0.493]$	86.13
Proposed	14.938 [± 0.999]	$-0.623[\pm 0.197]$	686.00

The results' average values are demonstrated out of the brackets, and the standard deviations are shown in the brackets

Table 7 Comparison of different schemes’ robustness

Attack type	SS-ICA	SVD-DCT	DWT-SVD	DWT-norm	LWT-SVD	DWT-VDVM	DWT-RDM	MDWT-DCT-RDM-WVM	Proposed
No Attack	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	2.86	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00
B	11.30	0.07	0.10	0.00	0.02	0.14	0.08	0.04	0.00
C	2.70	0.10	0.47	0.00	0.00	0.00	0.00	–	0.00
D	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	0.06	0.00	47.32	78.91	63.95	0.00	0.00	0.04	2.09
H	11.58	0.41	6.98	0.54	0.00	0.00	0.00	0.00	0.00
I	0.07	0.00	5.23	0.79	5.44	0.17	0.07	1.76	0.00
J	0.45	0.04	0.03	0.01	0.06	0.00	0.00	0.01	0.00
K	2.22	0.51	2.84	0.85	5.10	0.95	1.05	0.58	0.00
L	44.31	2.03	0.02	0.01	0.72	0.10	0.12	0.89	0.00

One of the problems in the SS-ICA algorithm is that it cannot correctly extract watermarks in no-attack status. The DWT-norm algorithm because of embedding the watermarks in approximation coefficients on long length interval has high resistance against Gaussian noise addition. In addition, the LWT-SVD algorithm, due to embedding in the coefficients of SVD Matrix, has high robustness against Gaussian noise addition. The MDWT-DCT-RDM-WVM algorithm, because of using combination of RDM and WVM in low-frequency components of DCT, has high robustness encountering attacks. The DWT-VDVM and DWT-RDM algorithms have high resistance against most of the attacks since they embed the watermark bits in low-frequency coefficients of DWT. The DWT-RDM and DWT-VDVM methods demonstrate moderate robustness in time shift attack, because the interval adapted for single bit watermarking in DWT-VDVM and DWT-RDM algorithms is not long enough in comparison with the interval of the proposed algorithm. However, the DWT-norm, DWT-SVD, and LWT-SVD schemes, because of fix quantization steps in watermark embedding, do not have a high robustness against Amplitude scaling attacks. As mentioned before, using Fibonacci numbers instead of fix quantization step provides a balance between robustness against attacks and imperceptibility. Moreover, embedding the watermarks in low-frequency coefficients guarantees both imperceptibility and robustness against common attacks.

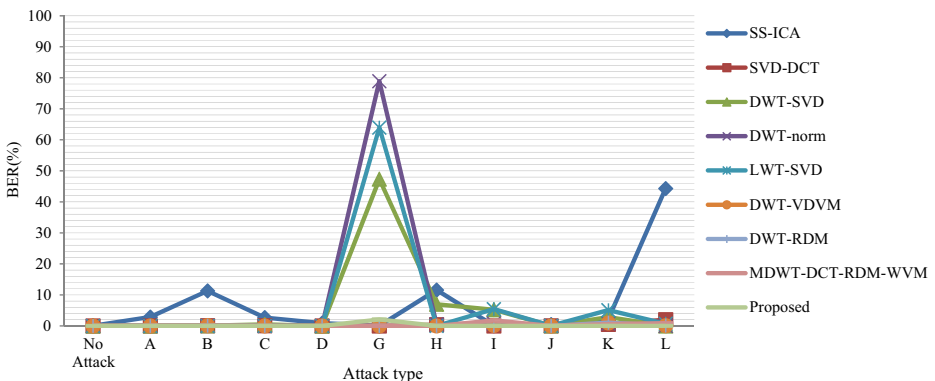


Fig. 9 Robustness test results for different watermarking schemes

10 Conclusions

This paper proposes the robust and blind audio watermarking algorithm on the 6th level approximation coefficients of DWT. Watermark embedding was conducted by modifying some magnitudes of the 6th level approximation coefficients according to Fibonacci numbers. The method is based on dividing the DWT coefficients into small frames and modifying some DWT coefficients by using Fibonacci numbers. The method provides a payload capacity of 686 bps. The experimental results demonstrated the efficiency of the proposed algorithm in retaining the perceptual quality of watermarked audio. It was also revealed that the proposed algorithm has a high robustness against most of the common attacks such as Gaussian noise addition, Resampling, Requantizing, Amplitude scaling, Echo addition, Low-pass filter, Mp3 comparison, Time shift, and Cropping. Additionally, the proposed algorithm enjoys the capability of applying a trade-off between the audio watermarking requirements such as imperceptibility, capacity, and robustness against attacks.

For future works, we are going to improve the proposed algorithm with synchronization capability. We will also check the robustness of the algorithm against more de-synchronization attacks, such as pitch scale modification, time scale modification, Jitter, etc.

References

1. Akansu AN, Richard AH (2001) Multiresolution signal decomposition: transforms, subbands, and wavelets. Academic Press, Boston
2. Al-Haj A (2014) An imperceptible and robust audio watermarking algorithm. *EURASIP J Audio, Speech, Music Process* 2014(1):37
3. Attari AA, Asghar BeheshtiShirazi A (2017) Robust and blind audio watermarking in wavelet domain. In: Proc. Int. Conf. Graph. Signal Process. - ICGSP '17. ACM Press, Singapore, Singapore, pp 69–73
4. Bhat KV, Sengupta I, Das A (2010) An adaptive audio watermarking based on the singular value decomposition in the wavelet domain. *ELSEVIER - Digit Signal Process* 20(6):1547–1558
5. Can YS, Alagoz F, Burus ME (2014) A novel spread spectrum digital audio watermarking technique. *J Adv Comput Networks* 2(1):6–9
6. Charfeddine M, El'arbi M, Amar CB (2014) A new DCT audio watermarking scheme based on preliminary MP3 study. *Multimed Tools Appl* 70(3):1521–1557
7. Dhar PK, Shimamura T (2014) Audio watermarking in transform domain based on singular value decomposition and Cartesian-polar transformation. *Int J Speech Technol* 17(2):133–144
8. Ellis D. <http://labrosa.ee.columbia.edu/matlab/>
9. Fallahpour M, Megias D (2011) High capacity audio watermarking using the high frequency band of the wavelet domain. *Multimed Tools Appl* 52(2-3):485–498
10. Fallahpour M, Megias D (2015) Audio watermarking based on Fibonacci numbers. *IEEE/ACM Trans Audio, Speech, Lang Process* 23(8):1273–1282
11. Fugal DL (2009) Conceptual wavelets in digital signal processing. Space and Signals Technical Publishing, San Diego
12. Hu HT, Chang JR (2017) Efficient and robust frame-synchronized blind audio watermarking by featuring multilevel DWT and DCT. *Clust Comput* 20(1):805–816
13. Hu HT, Chen WH (2012) A dual cepstrum-based watermarking scheme with self-synchronization. *Signal Process* 92(4):1109–1116
14. Hu HT, Hsu LY (2015) A DWT-based rational dither modulation scheme for effective blind audio watermarking. *Circuits, Syst Signal Process* 35(2):553–572
15. Hu HT, Chou HH, Yu C, Hsu LY (2014a) Incorporation of perceptually adaptive QIM with singular value decomposition for blind audio watermarking. *EURASIP J Adv Signal Process* 2014(1):1–12
16. Hu HT, Hsu LY, Chou HH (2014b) Variable-dimensional vector modulation for perceptual-based DWT blind audio watermarking with adjustable payload capacity. *Digit Signal Process A Rev J* 31:115–123

17. Jeyhoon M, Asgari M, Ehsan L, Jalilzadeh SZ (2016) Blind audio watermarking algorithm based on DCT, linear regression and standard deviation. *Multimed Tools Appl* 76(3):3343–3359
18. Kabal P (2002) An examination and interpretation of ITU-R BS. 1387: perceptual evaluation of audio quality. TSP Lab Technical Report, Dept. Electrical & Computer Engineering, McGill University, Montreal, pp 1–89
19. Lei BY, Soon Y, Li Z (2011) Blind and robust audio watermarking scheme based on SVD–DCT. *Signal Process* 91(8):1973–1984
20. Lei B, Soon Y, Zhou F, Li Z, Lei H (2012) A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition. *Signal Process* 92(9):1985–2001
21. Lin Y, Abdulla WH (2015) *Audio watermark: a comprehensive foundation using MATLAB*. Springer International Publishing
22. Liu SC, Lin SD (2006) BCH code-based robust audio watermarking in the cepstrum domain. *J Inf Sci Eng* 22(3):535–543
23. Megias D, Serra-Ruiz J, Fallahpour M (2010) Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification. *Signal Process* 90(12):3078–3092
24. No, really, Rust. <http://www.jamendo.com/en/album/7365>
25. Seok J (2012) Audio watermarking using independent component analysis. *J Inf Commun Converg Eng* 10(2):175–180
26. Tewari TK (2015) *Novel Techniques for Improving the Performance of Digital Audio Watermarking for Copyright Protection*. Ph.D. dissertation, Dept. Computer Science Engineering & Information Technology, Jaypee Institute of Information Technology, Noida
27. Wang J (2011) *New digital audio watermarking algorithms for copyright protection*. Ph.D. dissertation, Dept. Computer Science, National University of Ireland, Maynooth, Co. Kildare, Ireland
28. Wang H, Nishimura R, Suzuki Y, Mao L (2008) Fuzzy self-adaptive digital audio watermarking based on time-spread echo hiding. *Appl Acoust* 69(10):868–874
29. Wang X, Wang P, Zhang P, Xu S, Yang H (2013) A norm-space, adaptive, and blind audio watermarking algorithm by discrete wavelet transform. *Signal Process* 93(4):913–922
30. Xiang Y, Natgunanathan I, Guo S, Zhou W, Nahavandi S (2014) Patchwork-based audio watermarking method robust to de-synchronization attacks. *IEEE/ACM Trans Audio, Speech, Lang Process* 22(9):1413–1423
31. Xiang Y, Hua G, Yan B (2017) *Digital audio watermarking: fundamentals, techniques and challenges*. Springer, Singapore



Ali Akbar Attari received the B.Sc. degree in Broadcast Engineering from Islamic Republic of Iran Broadcasting University (IRIBU), Tehran, Iran in 2008 and M.Sc. degrees in Information Technology Engineering - Secure Communications from Iran University of Science and Technology (IUST), Tehran, Iran in 2017. He has been collaborating with IRIB research and development department. His research interests include Signal processing, Multimedia Watermarking, and Secure Communication.



Ali Asghar Beheshti Shirazi received the B.Sc. and M.Sc. degrees in Communication Engineering from Iran University of Science and Technology (IUST) in 1984 and 1987, respectively and Ph.D. from Okayama University, Japan in 1995. In 1995, he joined the Department of Electrical Engineering, IUST, where he currently is an Assistant Professor. His research interests include Image Processing and Coding, Signal Processing, Data Communication Networking, and Secure Communication.