A Robust Blind Audio Watermarking Scheme Based on DCT-DWT-SVD

Azadeh Rezaei and Mehdi Khalili

Abstract In this paper, Watermarking hybrid sound algorithm is presented to protect the copyright of audio files, which, in addition to the clarity and consistency of the audio signal, has increased the strength and strength. To this end, a hybrid algorithm for voice signal cryptography is presented in the three domain parser transforms, discrete cosine transform, and discrete wavelet transforms. So, after discrete cosine transformation (DCT) on the host signal, by selecting the sub-band of low frequency, which contains the highest signal energy, two discrete wavelet transform (DWT) with a random wavelet filter on the low-frequency coefficients of conversion A discrete cosine applies, after selecting the approximation coefficients, the resulting one-dimensional matrix is converted to a two-dimensional matrix, and finally the resulting matrix is applied to a single value decomposition (SVD), which results in the formation of a The diameter matrix is that the watermark bits are embedded in the first layer of the dipole matrix, so that the two bits with value S (1,1), S (2,2) of the matrix The diameter S is chosen, first compares the first and second intersections of the diameter matrix S, which is multiplied by the coefficient θ multiplied by the obtained two bits and is used as a fixed value in the embedding formula and The title of the new watermark is embedded in S (1,1). The results of the implementation show that the proposed algorithm succeeded not only in achieving transparency and resistance to general audio processing attacks, such as Gaussian white noise, quantization rates, decreasing and increasing the rate of sampling, compression and low pass filtering. But has achieved better results than other similar algorithms.

Keywords Audio cryptography · Discrete wavelet transform · Discrete cosine transform · Single value decomposition

A. Rezaei (✉) · M. Khalili
Faculty of Computer Engineering, Payame Noor University, Tehran, Iran
e-mail: azadehrezaei1369@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_8

1 Introduction

Recent advances in the Internet and digital multimedia products technology have made it possible for digital signals (audio, video, and video) to be easily distributed to different regions. This ease in Transmission, allows unauthorized copies of multimedia products to be distributed and distributed. For this reason, protecting the right to digital rights has become an important topic in the world [1]. Digital Watermark has given a lot of attention to solving this problem [2]. Information storage is a way to make information. In the form of an overlapping agent with the highest degree of security precaution, it moves between the points in question so that even if the information is accessed through unauthorized persons along the route, there is no access to the hidden data [3]. In fact, the lack of a focus on art and science is the embedding of information in a carrier medium, which is becoming increasingly widespread due to the significant advancements in digital communications [4]. In mainstream preservation, security means inability to prove the existence of a message [5]. Digital media are more popular than analogue media, and unlike them, they can be easily stored, reproduced and distributed. Thus, unauthorized reproduction of digital documents, such as audio signals, has been raised as a concern in recent years [6]. Therefore, digital sounds calling is used to protect copyright and prove ownership [7]. In this regard, several methods have been proposed in different areas of frequency and time, in which the algorithms presented in the frequency domain are more resistant [8]. The most significant transformations in the transformation area used in watermarking algorithms are the discrete cosine transform, the discrete wavelet transform, and the conversion of the parsing of unique values [9]. Discrete cosine transforms are resistant to image processing operations such as low pass filter, contrast, brightness adjustment, etc. [10]. A discrete wavelet transform is presented as an alternative to the Fourier transform of short time, and its goal is to overcome resolution problems in Fourier transforms of short time [11]. Although DCT has a combination of features of the human vision system, wavelet transformation is closer to the human vision system than DCT, and in general in some applications it is better than DCT-based methods [10]. The conversion of the unique value fragmentation is an efficient and effective tool in numerical analysis, a clear feature of this transformation of stability against attacks such as rotation and noise [11]. Considering the benefits of each of the conversion areas, various studies have been done on the combination of these conversions on Watermarking, which is, of course, much of the research done in this regard, the use of watermarking of digital images in audio signals is most commonly used. Compared to Watermarking, Video and Watermarking video are more susceptible to two major causes. First, the contents of the audio signals are one-dimensional data, so it's very difficult to add hidden data without harming the quality of the audio signal. Secondly, the human hearing system (HAS) is more sensitive to the human system of vision, so the control of degradation in quality is detected by the listener. In practical applications categories, the soundtrack method must certainly meet a number of needs [12]. The three important requirements that come with the Watermarking Voice are: Inaudible, Resistance and Security [13].

Cai et al. [14] in order to protect the copyright of digital audio and video copyright in the network, a blind audio algorithm is proposed using digital discrete wavelet transform (DWT) and the conversion of unique values (SVDs). In this algorithm, an original audio signal is segmented into blocks with a length of 1024, and each block blocks the wavelet transform into two levels. Then, the approximate coefficients of the near-correct sound of the initial sound are decomposed into the SVD conversion, and the resulting matrix of transformation is obtained. Watermarked information is embedded in a diagonal matrix. The experiments show that the proposed algorithm has a transparency of 20.7000 and its robustness against common audio signal attacks, such as re-sampling, low pass filtering, quantization, Gaussian noise, MP3 compression, on average, with a correlation coefficient of more than 0.940 and an average. The error rate is less than 0.060. The non-attack extracted signal has a mean normal correlation coefficient of 1 and an average error rate of 0 that indicates that the watermark can be extracted explicitly and without blindness without any attack. In [15], a strong blind sound market with DCT-DWT-SVD is presented. In this article, the original sound is first divided into frames of length 4096. Then DCT conversion is applied to each frame at first. The low frequency coefficients derived from the DCT conversion of the five levels of the DWT conversion are applied to obtain the approximation coefficients. Eventually, the SVD conversion is applied to obtain the three matrices S, U, V, the 32-bit binary watermark image in the matrix Diagonal S is embedded. Public sounding attacks such as sampling, low pass filtering, quantization, Gaussian white noise, and MP3 compression have been applied, the results show that the average correlation coefficient is normal at about 0.980 and the average error rate is less than 0.058. In this paper, the strength of the work has been investigated, but the transparency has not been tested and the extraction is non-blind. In the proposed paper, a method is suggested that, in addition to achieving the characteristics of Watermarking, it improves the transparency component and the resistance of the audio signal, so that by combining three discrete wavelet transforms, a discrete cosine transform, and the decomposition of single values into an algorithmic representation in To prove the right to own digital audio.

2 Background

2.1 Discrete Cosine Transform (DCT)

A discrete cosine transform is a technique for returning the signal within the elemental frequency components [16]. A discrete cosine transform transforms a signal into three sub-band frequencies down, middle, and up. Most of the input signal is accumulated in low frequency components, called DCs [17].

The relation of the discrete cosine transformation is in the form of relation (1):

$$X[K] = \mathrm{w[K]} + \sum_{n=0}^{N-1} \mathrm{X[n]}\left(\cos\frac{(2n+1)K\pi}{2N}\right), \quad 0 \ll \mathrm{K} \ll \mathrm{N} \tag{1}$$

So that

$$w[k] = \begin{cases} \sqrt{\frac{1}{N}}, & k = 0 \\ \sqrt{\frac{2}{N}}, & 0 < k \leq N-1 \end{cases}$$

2.2 Discrete Wavelet Transform (DWT)

The wavelet transform method is that with two high pass and low pass filters, the audio signal is divided into two sections: high frequency, detail (CD) and low frequency, approximation (CA), the number of samples in each of these sections Half the number of samples in the main signal [18]. The transformation of a discrete wavelet acts as follows:

$$\varphi_{jk}(t) = \frac{1}{\sqrt{a_0^j}}\varphi_{jk}\left[\frac{\frac{t-ka_0^j b_0}{j}}{a_0}\right] \tag{2}$$

$$Wnf(j,k) = \langle f(j,k), \varphi_{jk}(t)\rangle \tag{3}$$

$$= \frac{1}{\mathrm{C}_\varphi}\int_0^{+\infty}\int_{-\infty}^{+\infty} \mathrm{W_f}(\partial, b)\varphi_{\partial,b}(t)db\frac{d\partial}{|\partial|^2} \tag{4}$$

2.3 Singular Value Decomposition (SVD)

This conversion is an efficient and effective tool in numerical analysis. This conversion is due to the application of the matrix topic and the fact that a digital image is also a matrix can be effective in the process of watermarking [19]. Using this transformation on the desired matrix A Dimensions m * n are obtained by three matrices U, V, S [11] so that A = U * S * V and

$$S = diag(\sigma 1, \ldots \sigma i, \ldots \sigma r), \sigma i > 0(i = 1, \ldots, r)r = rank(\mathrm{A}) \tag{5}$$

3 The Proposed Algorithm

In order to protect the proposed algorithm, a binary image is first embedded as a watermark in an audio signal. First, the watermark image of the gray will be converted to a binary image before the embedding of the watermark, then the embedding and extraction process is described. The block diagram of embedding and extraction of watermarks is shown in Figs. 1 and 2 respectively.

3.1 Watermarking Process

1. Framing the original audio signal
2. Convert DCT to any frame
3. Performing two levels of DWT conversion on the low frequency coefficients derived from the DCT transform with a random wavelet filter for obtaining approximation coefficients (second level transformation on the approximation coefficient obtained from the first level transformation).
4. Conversion of the one-dimensional matrix obtained from the conversion of the coefficients of the DWT approximation to a two-dimensional matrix
5. Run SVD transform on a two-dimensional matrix according to the relationship and obtain three matrices U, S, V

$$R = U\,S\,V \tag{6}$$

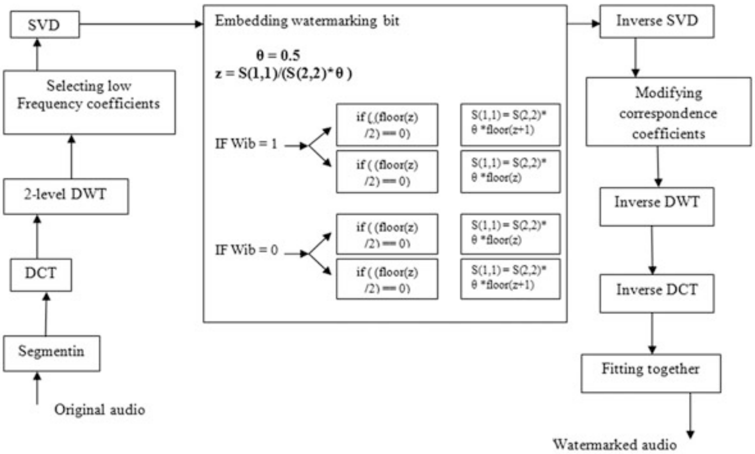

Fig. 1 Block diagram of watermark image embedding algorithm in domain DCT-DWT-SVD

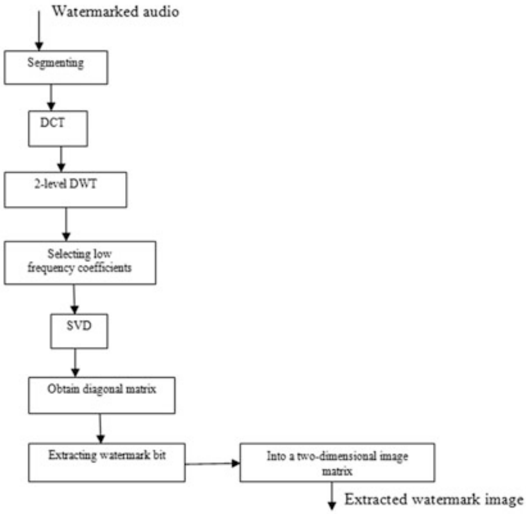

Fig. 2 Block diagram for watermarked extraction of image-sound algorithm in domain DCT-DWT-SVD

6. Two bits of value S (1,1), S (2,2) are selected from the diameter matrix S and embedded in them according to the conditions of the watermark. We first consider the comparison of the first and second intersections of the diagonal matrix S: if the remainder of the subdivision of the sub frame, including the integer S (1,1), on the product of S (2,2) in the defined coefficient, is 0.5 = θ, we consider the constant z, then If this value is a pair wise integer and a watermark bit, then it multiplies the individual value of z (i.e., z + 1) in the product of S (2,2) in the coefficient θ, and in the bit S (1,1) We install as new watermarked. Now, if the bit watermark is 0, then the z value unchanged is multiplied as a pair wise value in S (2,2) and θ, and we embed it as a new watermarked value in S (1,1).

$$\begin{aligned} &\theta = 0.5 \\ &z = S(1{,}1)/(S(2{,}2) * \theta) \end{aligned} \tag{7}$$

$$\begin{aligned} &\text{if}((\text{floor}(z)/2) == 0) \quad \text{is even} \,\&\, \text{watermark image bit} = 1 \\ &\quad S(1{,}1) = S(2{,}2) * \theta * \text{floor}(z+1)) \end{aligned} \tag{8}$$

$$\begin{aligned} &\text{if}((\text{floor}(z)/2) == 0 \quad \text{is odd} \,\&\, \text{watermark image bit} = 1 \\ &\quad S(1,1) = S(2{,}2) * \theta * \text{floor}(z)) \end{aligned} \tag{9}$$

Now if z is a certain number and the watermark bit 1 then multiplies the pair of z, (z + 1) in the product of S (2,2) in the coefficient θ, and in S (1,1) to Insert the title of the new watermark. Now, if the bit watermark is equal to 0, then the z value unchanged as an individual value in S (2,2) and θ is multiplied and embedded as a new watermarked value in S (1,1).

$$\begin{aligned} &\text{if}((\text{floor}(z)/2) == 0 \quad \text{is even} \,\&\, \text{watermark image bit} = 0 \\ &\quad S(1{,}1) = S(2{,}2) * \theta * \text{floor}(z)) \end{aligned} \tag{10}$$

$$\begin{aligned} &\text{if}((\text{floor}(z)/2) == 0 \quad \text{is odd} \,\&\, \text{watermark image bit} = 0 \\ &\quad S(1{,}1) = S(2{,}2) * \theta * \text{floor}(z+1)) \end{aligned} \tag{11}$$

In this algorithm, the bits of diameter matrix S (1,1) with the proposed solution contain watermark bits.

7. Reversal SVD conversion on watermarked matrix.
8. Convert a two-dimensional matrix to a one-dimensional matrix.
9. Inverse the two levels of DWT conversion on a one-dimensional matrix and obtaining approximation coefficients.
10. Reverse DCT conversion on approximation coefficients.
11. Finally, the addition of all the original audio signal frames and the watermark to calculate the watermarked audio signal.

3.2 Watermark Extraction Process

1. Framing the watermarked audio signal.
2. Performing DCT transformation on the watermark signal and obtaining low frequency coefficients.
3. Implementing two levels of DWT conversion on the low frequency coefficients of the DCT conversion and obtaining approximation coefficients (second level transformation on the approximation coefficient obtained from the first level transformation).
4. Conversion of a one-dimensional matrix to a two-dimensional matrix

$$R_S = U_S \, S_S \, V_S \tag{12}$$

5. Run SVD conversion on a two-dimensional matrix.
6. Isolation of watermark bits without the need for the original audio signal by the blind method.
7. Finally, sorting the extracted watermark bit in a two-dimensional matrix to extract the Watermark image.

4 Results of Implementation of the Proposed Method

In the proposed algorithm, first, a 32-bit binary image with jpg format (Fig. 3) as a watermark in a 16-bit audio signal with a sampling rate of 44,100 kHz with a size of 2,395,137 * 1 and a wav format (Fig. 4) is embedded. The main watermark

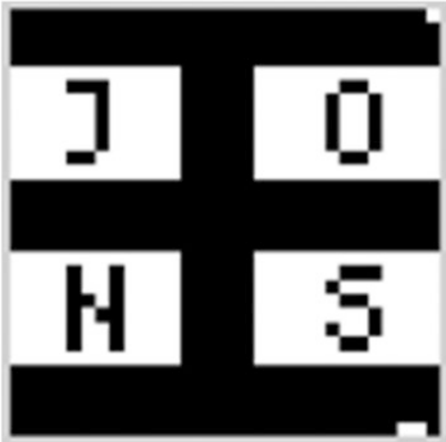

Fig. 3 Main watermark image

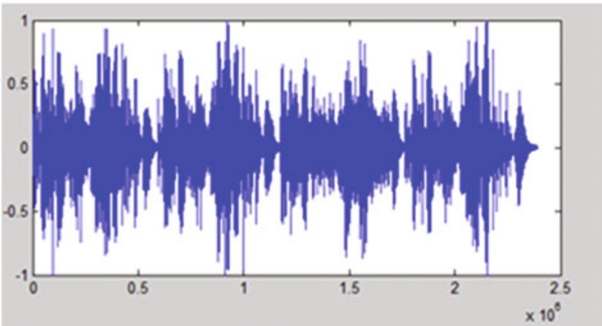

Fig. 4 Main sound signal

image and the images attacked by various attacks, including Gaussian white noise, the rate of quantization, decreasing and increasing the sampling rate, compression and low pass filter, as well as the original sound signal and the watermarked sound signal are also shown in Fig. 5. It can be seen. The test results are based on the following criteria: SNR, which represents the signal-to-noise ratio. This scale indicates the amount of noise added to the sound due to the insertion of a mark in it. The BER, which is the bit error rate between the two image images extracted with the original sign, and the NCC, which represents the normalized cross-correlation of the two signs, has been reported. The SNR value is given in Table 1, and the BER and NCC values obtained from this algorithm are shown in various tables, respectively, in Tables 2 and 3. The comparison chart for BER and NCC values is also given in Figs. 6 and 7, respectively.

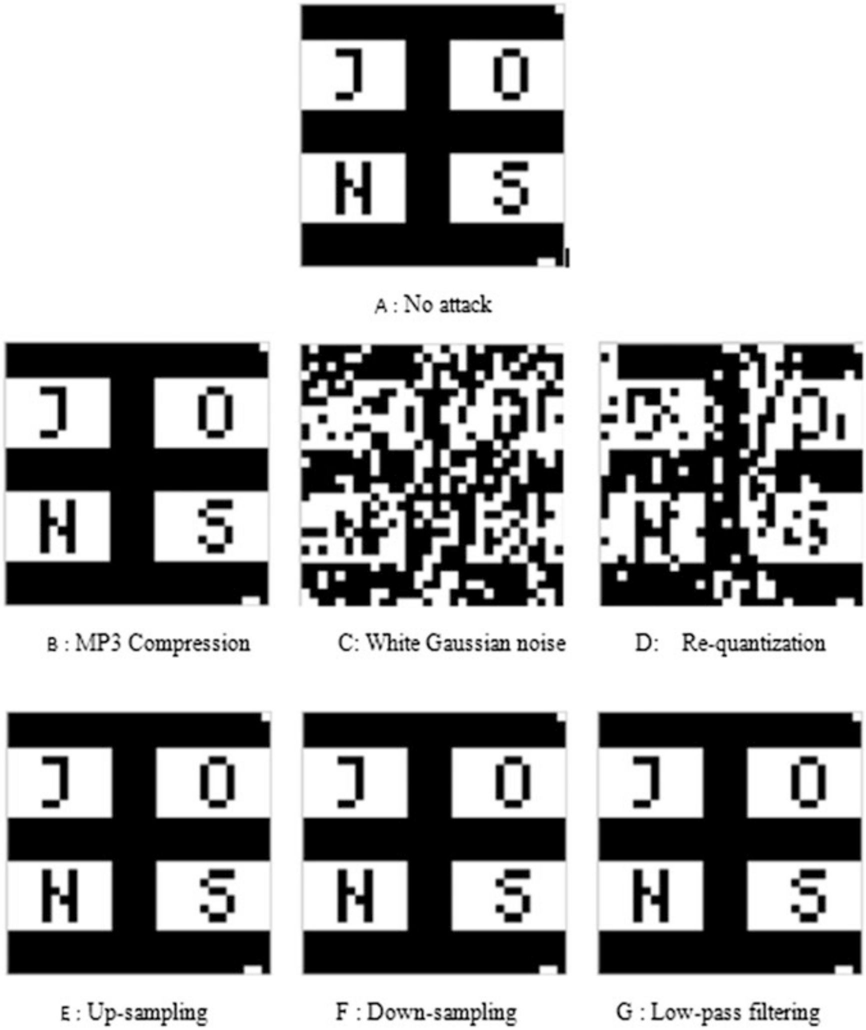

Fig. 5 Watermark image with SNR = 35.2469 with various image processing attacks

Table 1 Comparison of the proposed SNR of the proposed algorithm with similar tasks

Proposed watermarking algorithm	Method	
	Cai et al. [14]	Lei et al. [15]
35.2462	20.7000	…

Table 2 Comparison of the BER Factor of the proposed algorithm with similar tasks

Attack	Method		
	Cai et al. [14]	Lei et al. [15]	Proposed watermarking algorithm
No attack	0.0000	0.0000	0.0000
MP3 compression	0.0000	0.0000	0.0000
White Gaussian noise	0.1523	…	0.1474
Re-quantization	0.2676	0.1791	0.1033
Up-sampling	0.0000	0.0586	0.0000
Down-sampling	0.0000	…	0.0000
Low-pass filtering	0.0059	0.0586	0.0000

Table 3 Comparison of the NCC factor of the proposed algorithm with similar tasks

Attack	Method		
	Cai et al. [14]	Lei et al. [15]	Proposed watermarking algorithm
No attack	1.0000	1.0000	1.0000
MP3 compression	1.0000	1.0000	1.0000
White Gaussian noise	0.8603	…	0.8723
Re-quantization	0.7311	0.9206	0.9875
Up-sampling	1.0000	1.0000	1.0000
Down-sampling	1.0000	…	1.0000
Low-pass filtering	0.9999	0.9440	1.0000

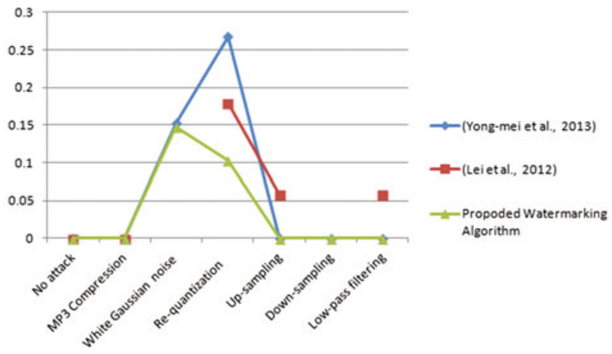

Fig. 6 Comparison of the BER Factor with the proposed method

$$SNR = 10\log_{10}\frac{\sum_{n=1}^{N} Y^2(n)}{\sum_{n=1}^{N}[Y(n) - Y'(n)]^2} \quad (13)$$

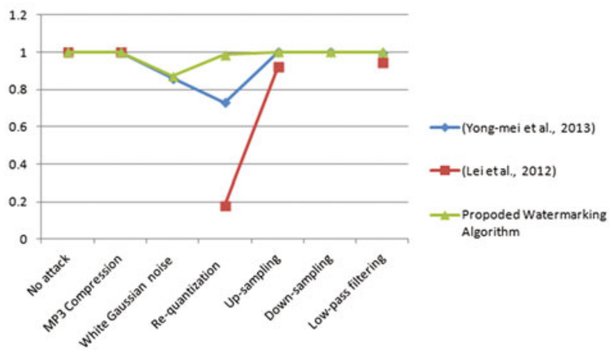

Fig. 7 Comparison of the NCC factor of the proposed method with similar tasks

$$BER = \left\{ \sum_i \sum_j \frac{\{W(i,j) \oplus EW(i,j)\}}{M \times N} \right\} \times 100\%. \tag{14}$$

$$NC(w, \hat{w}) = \frac{\sum_{i=1}^{Nw} w(i).\widehat{w}(i)}{\sum_{i=1}^{Nw} w^2(i)} \tag{15}$$

5 Conclusion

This paper is devoted to the presentation of a hybrid algorithm based on discrete cosine transformation, discrete wavelet transform and parsing single values for voice signal cryptography. The proposed algorithm has been able to prove the right to ownership of digital audio and video. On the other hand, the comparison of the proposed algorithm and the evaluation of their performance and resiliency and their transparency against various attacks has led to a rather comprehensive research on the methods of watermarking for digital audio signals.

In order to use the benefits of different domains simultaneously, a hybrid cognitive algorithm has been proposed in three discrete cosine transformation areas, discrete wavelet transform and single particle breakdowns. In the proposed cognitive auditing approach, a DCT operation layer is initially on the host audio signal is applied to produce high, low, and average frequencies, then on a low frequency coefficient, which contains the highest energy, a DWT conversion level is applied with a randomized wavelet filter to yield the frequency bands CA, CD. Then, on the sub-band CA, which is called the sub band of approximation and follows the initial shape of the signal, once again the DWT conversion is applied to the randomized wavelet filter, we transform the resulting one-dimensional matrix into a two-dimensional matrix, It converts a single value parsing (SVD), so that the two bits with S (1,1), S (2,2) are selected from the S-matrix matrix, first compare the first and second intersections of the diameter matrix S And embedded in the formula for the new watermarked value in S (1,1). Then the ISVD is converted and the two-dimensional matrix is converted to a one-dimensional matrix. By performing two conversion levels, the modified IDWT blocks are redone to the modified sub-band of the edited host audio, and then a conversion level of the IDCT is applied. Has been finally, to create a watermark, all the original audio frames and watermark bits are merged.

The watermarking process is also reversed by the embedding process, so that the host signal is first blocked and a DCT transformation level is applied to produce low-frequency coefficients and the DWT conversion factors are carried out with a randomized wavelet filter. And low frequency coefficients are approximated. Again, these approximation coefficients apply to a transformation DWT, and the one-dimensional matrix is derived from the approximation coefficients to a two-dimensional matrix, transforming the SVD to obtain a diagonal matrix The two-dimensional matrix is applied and the watermark bits are extracted without the need for the host's sound signal. The results indicate that the algorithm has a satisfactory performance in achieving the desired goals and has performed better than similar research.

References

1. Thanki R, Borisagar K, Borra S (2018) Speech Watermarking technique using the finite Ridgelet Transform, discrete wavelet transform, and singular value decomposition. Advance compression and watermarking technique for speech signals, pp 27–45
2. Zhang J, Han B (2017) Robust audio watermarking algorithm based on moving average and DCT. KYTZ201420 and scientific research project of department of education in Sichuan province under Grant, No. 16ZA0221
3. Ojha S, Sharma A, Chaturvedi R (2018) Centric-oriented novel image watermarking technique based on DWT and SVD. Soft computing: theories and applications. Advances in intelligent systems and computing, vol 583. Springer, Singapore
4. Mishra S, Yadav VK, Trivedi MC, Shrimali T (2018) Audio steganography techniques: a survey, advances in intelligent systems and computing, vol 554. Springer, Singapore
5. Seok J, Hong J, Kim J (2017) A novel audio watermarking algorithm for copyright protection of digital audio. ETRI J 24(3):181–189
6. Singh S, Singh R, Singh AK, Siddiqui TJ (2018) SVD-DCT based medical image watermarking in NSCT domain. Quantum computing: an environment for intelligent large scale real application. Studies in big data, vol 33. Springer, Cham
7. Kavitha KJ, Shan BP (2018) Video watermarking using DCT and DWT, a comparison. Eur J Adv Eng Technol 2(6):83–87
8. Saxena N, Mishra KK, Tripathi A (2018) DWT-SVD-based color image watermarking using dynamic-PSO. In: Advances in intelligent systems and computing, vol 554. Springer, Singapore
9. Tagesse Takore T, Rajesh Kumar P, Lavanya Devi G (2018) A robust and oblivious grayscale image watermarking scheme based on edge detection, SVD, and GA. In: Proceedings of 2nd international conference on micro-electronics, electromagnetics and telecommunications. Lecture notes in electrical engineering, vol 434. Springer, Singapore

10. Verma C, Tarar S (2018) Secure random sequence based frequency hoping spread spectrum audio watermarking. Int J Eng Sci Comput 6(5)
11. Jain R, Trivedi MC, Tiwari S (2018) Impact analysis of contributing parameters in audio watermarking using DWT and SVD. In: Advances in intelligent systems and computing, vol 554. Springer, Singapore
12. Gosavi CS, Mali SN (2017) Watermarking for Video using single channel block based schur decomposition. Global J Pure Appl Math 12(2):1575–1585. ISSN 0973-1768
13. Pohan N, Saragih R, Rahim R (2017) Invisible watermarking audio digital with discrete cosine transform. Sci Technol IJSRST 3(1). Print ISSN: 2395-6011, Online ISSN: 2395-602X
14. Cai Y-m, Guo W-q, Ding H-y (2013) An audio blind watermarking scheme based on DWT-SVD. J Softw 8(7)
15. Lei M, Yang Y, Guo Y, Mao J, Luo Q (2012) A robust blind audio watermarking scheme based on DCT-DWT-SVD. Int J Dig Content Technol Appl (JDCTA) 6. https://doi.org/10.4156/jdcta.vol6.issue21.29
16. Kakkirala KR, Chalamala SR, Rao G (2017) DWT-SVD based blind audio watermarking scheme for copyright protection. In: Audio, language and image processing (ICALIP)
17. Padungdit A (2018) Image watermarking using joined wavelet and time domain. In: ICT international conference, No. 3, pp 47–50
18. Csrvajal G (2017) Scaling factor for RGB image to steganography application, IEC comprehensive report on information security international engineering consortium. J Vec Rel 3:55–56
19. Darabkh KA (2018) Imperceptible and robust DWT-SVD-based digital audio watermarking algorithm. J Softw Eng Appl 7:859–871