

# Efficient audio integrity verification algorithm using discrete cosine transform

Sabry S. Nassar<sup>1</sup> · Nabil M. Ayad<sup>1</sup> · Hamdy M. Kelash<sup>2</sup> · Hala S. El-sayed<sup>3</sup> · Mohsen A. M. El-Bendary<sup>4</sup> · Fathi E. Abd El-Samie<sup>5</sup> · Osama S. Faragallah<sup>2,6</sup>

Received: 6 April 2015 / Accepted: 13 October 2015 / Published online: 12 November 2015  
© Springer Science+Business Media New York 2015

**Abstract** Audio recordings have been used as evidence for long times. Multimedia processing advancement makes it difficult to be completely sure about what is heard is the truth. This paper presents a promising approach for integrity verification of recorded audio signals using discrete cosine transform. This approach is based on self embedding concept which embeds block-based marks extracted from the same audio signal after being transformed into 2-D format into other blocks according to a specific algorithm. After the self-embedding process, the data is converted back into 1-D style which represents a marked audio signal. The 1-D audio signal is converted into a 2-D format and then converted back into a 1-D format using the popular lexicographic ordering scheme utilized in image processing. Reverse processes are executed to extract the verification marks from the audio signal throughout the

integrity verification process. Based on the extracted audio signal properties, the integrity of the marked audio signal is evaluated. Different audio processing tasks and attacks are implemented to examine the suitability of the proposed algorithm for verifying the integrity of high-confidentiality recorded audio data. The results show that the efficient ability of the proposed approach to verify integrity and detect attacks.

**Keywords** Integrity · Digital Watermarking · DCT

## 1 Introduction

With the rapid advance in multimedia processing techniques, illegal copy and modification of recorded audio have become very common. So, the content integrity of a recorded audio data has become of great concern. When dealing with audio data in medical, military, and audio applications, the integrity is very important to ensure that the original data has not been altered (Gómez et al. 2002). So, it is very important to use sensitive integrity verification schemes to detect intended or unintended modifications in the data. Recently, watermarking has gained a popular rule in this task for copyright protection, content verification, and broadcast monitoring (Zhang et al. 2012).

There are various classification criteria upon which watermarking techniques can be classified. These criteria include robustness, perceptibility, and embedding and retrieval procedures. The robustness describes the ability of the watermark to resist common data manipulations. The watermark can be robust, fragile, or semi-fragile. A robust watermark is used to protect the copyright, because it aims at resisting different types of processing (Petrovic 2005; Chang and Chang 2010; Radharani and Valarmathi 2010).

---

✉ Osama S. Faragallah  
osam\_sal@yahoo.com; o.salah@tu.edu.sa

- <sup>1</sup> Nuclear Research Center, Atomic Energy Authority of Egypt, Cairo, Egypt
- <sup>2</sup> Faculty of Electronic Engineering, Department of Computer Science and Engineering, Menoufia University, Menouf 32952, Egypt
- <sup>3</sup> Faculty of Engineering, Department of Electrical Engineering, Menoufia University, Shebin El-kom 32511, Egypt
- <sup>4</sup> Department of Communication Technology, Faculty of Industrial Education, Helwan University, Cairo, Egypt
- <sup>5</sup> Faculty of Electronic Engineering, Department of Electronics and Electrical Communications, Menoufia University, Menouf 32952, Egypt
- <sup>6</sup> Department of Information Technology, College of Computers and Information Technology, Taif University, Al-Hawiya 21974, Kingdom of Saudi Arabia

On the contrary, a fragile or semi-fragile watermark is used to verify the authenticity and content integrity. When a delicate watermark is attacked; obviously, it will be destroyed (Li and Yang 2003).

A combination of robust and fragile watermarks in audio signals can be found in (Lu et al. 2000). Of these watermarks, one was used to provide copyright protection and the other was used for tampered regions detection. The integrity verification using watermarking and fingerprinting techniques with a fragile speech recording authentication scheme based on the discrete Fourier transform (DFT) was studied by Wu and Jay Kuo (2001, 2002). In (Gómez et al. 2002), the integrity of audio recordings is verified using a mixed watermarking-fingerprinting approach with some sort of self embedding. The idea of self embedding mainly depends on extracting the fingerprint or a dependent mark from a used signal and then keeping it within the signal through a marking process. So, there is no need for additional external data during the integrity verification process. Some approaches with the same concept have been presented for image, video, and audio (Dittmann et al. 1999; Dittmann 2001; Shaw 2000).

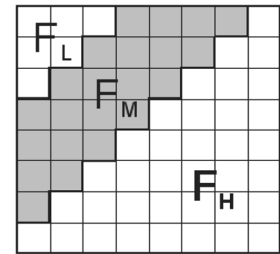
In (Liu and Chen 2004), a technique that localizes tampered regions of the speech content and recovers these regions based on LSB watermarking was proposed. In this paper, we present a self embedding mechanism based on the DCT for verifying audio content integrity. This idea of this mechanism is to work on the DCT of the signal blocks arranged in 2-D and embed a signature extracted from each block into another block in a manner that does not distort the audio signal. The rest of the paper is arranged as follows. Section 2 provides a review of the DCT. Section 3 presents the proposed content verification scheme. Section 4 presents the obtained results. Finally, the conclusions with future trends are given in Sect. 5.

## 2 Discrete cosine transform

Transform domain techniques have shown better performance in watermark embedding than time domain techniques. The majority of transform domain techniques depend on inserting the datum into the transform coefficients of the cover information, and after the modification of the coefficients, the information is converted back into the spatial domain (Shoemaker 2002; Tewfik 2000). One of the most popular transforms for watermark embedding is the DCT (Nassar et al. 2014).

The DCT is a real transform mapping a sequence of a certain length to another sequence of the same length. It has a good energy compaction property, and enables segmentation of the signal of concern into sub-bands. Type-II

**Fig. 1** Definition of DCT Regions



DCT is the most popular for DCT implementation in signal and image processing (Ahmed et al. 1974).

The DCT divides the signal into high, middle, and low frequency components (FL, FM, and FH) as shown in Fig. 1. In the low frequency components, most of the signal energy resides, which makes it improper to modify these components. On the other hand, high frequency components are usually removed through compression. So, mid-frequency components are the most appropriate for watermark embedding (Shiva Kumar et al. 2010).

In this work, we deal with audio signals with an image processing concept. We first rearrange the 1-D audio signal into a 2-D matrix, and then apply the 2-D DCT defines as: (Khayam 2003):

$$C(u, v) = a(v)a(u) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_i \cos\left(\frac{\pi u(2i+1)}{2N}\right) \times \cos\left(\frac{\pi v(2j+1)}{2N}\right) \quad (1)$$

where  $M$  and  $N$  are the 2-D matrix dimensions. Also,  $u$  and  $v = 0, 1, 2, \dots, N-1$ .

## 3 The proposed verification algorithm

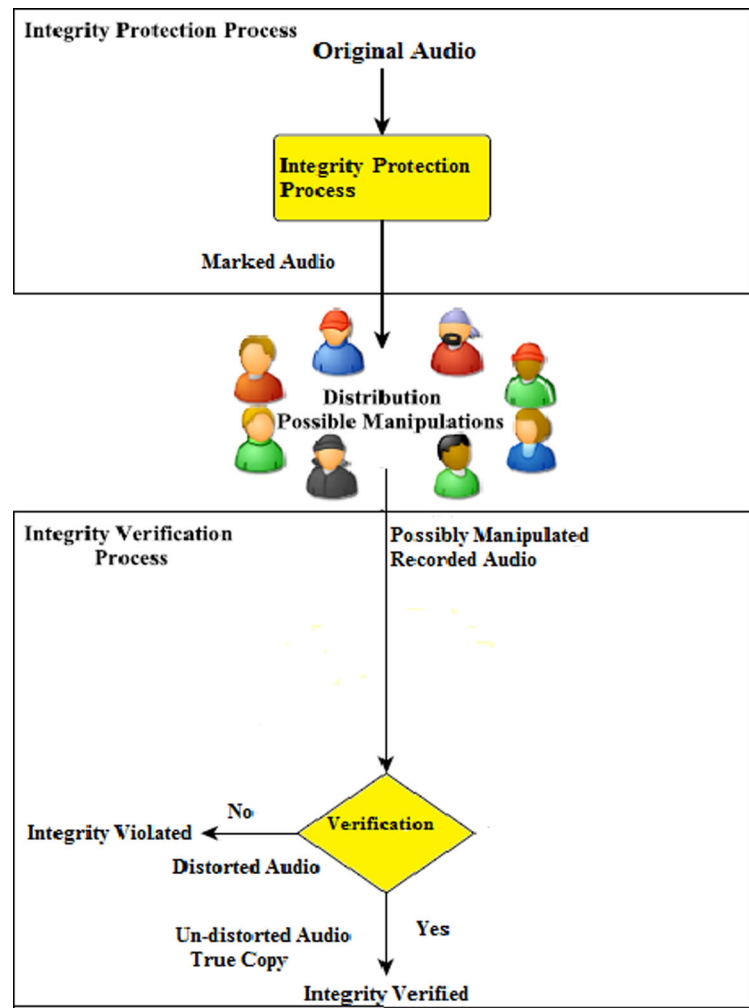
In this paper, the DCT is exploited in audio integrity verification depending on developing the concepts in (Nassar et al. 2014). After audio signal rearrangement to 2-D, we work on 16x16 blocks. A self embedding process is applied to put a mark or signature of each block into another block. The proposed verification scheme is shown in Fig. 2.

### 3.1 Integrity protection process

The steps of the suggested integrity verification method are summarized as follows:

- Step 1* The 1-D audio signal is rearranged into 2-D.
- Step 2* The obtained 2-D signal is segmented into 16x16 blocks and the DCT is applied on each block.
- Step 3* The first row and first column from each block are weighted with a small weight and embedded in the last

Fig. 2 Proposed approach



row and last column of another block, respectively. This process is performed on each pair of blocks.

*Step 4* Inverse DCT is performed on each block to obtain a modified 2-D matrix.

*Step 5* The signal is rearranged again to 1-D.

*Step 6* The obtained signal is transmitted through a communication channel.

### 3.2 Integrity verification process

The integrity verification process are summarized as follows:

*Step 1* The 1-D audio signal is rearranged in 2-D.

*Step 2* The signal is segmented into 16x16 blocks as in the embedding process, and the 2-D DCT is applied.

*Step 3* The embedded row and column in each DCT block are extracted and the correlation coefficients are estimated with the corresponding rows and columns in the source block, respectively.

*Step 4* Based on the correlation coefficient values, a decision can be made whether the signature exists or not.

*Step 5* The original signal can be recovered after signature or mark removal.

## 4 Result analysis and comparison

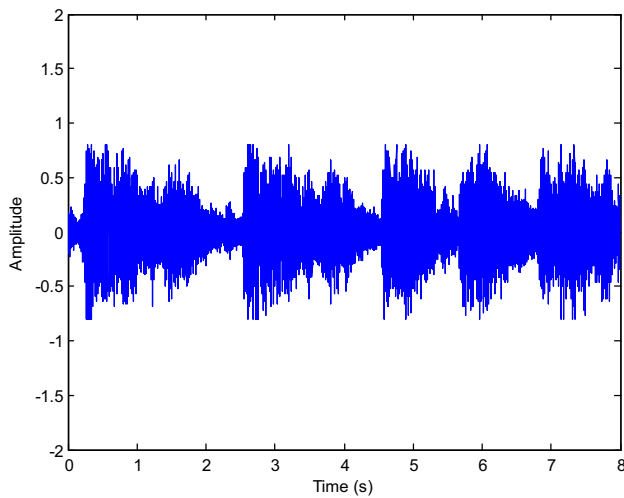
The proposed audio signal integrity verification process have been tested with MATLAB simulations. The original signal in Fig. 3 has been used.

### 4.1 Integrity protection process

Some quality metrics are presented below for the proposed scheme.

#### 4.1.1 Inaudibility

Inaudibility (audio imperceptibility) is determined by the perceptual difference between the marked audio signal and the original audio signal. It depends on the SNR defined as (Can et al. 2014; Al-Haj et al. 2009):



**Fig. 3** Original audio signal

$$SNR = 10 \log_{10} \frac{\sum_{j=1}^m x_j^2}{\sum_{j=1}^m |x_j - y_j|^2} \quad (2)$$

where  $x_j$  is the original audio signal and  $y_j$  corresponds to the marked audio signal,  $j$  is the sample index, and  $m$  is the number of samples of the output audio signal (Bassia et al. 2001).

#### 4.1.2 Audio Spectrogram

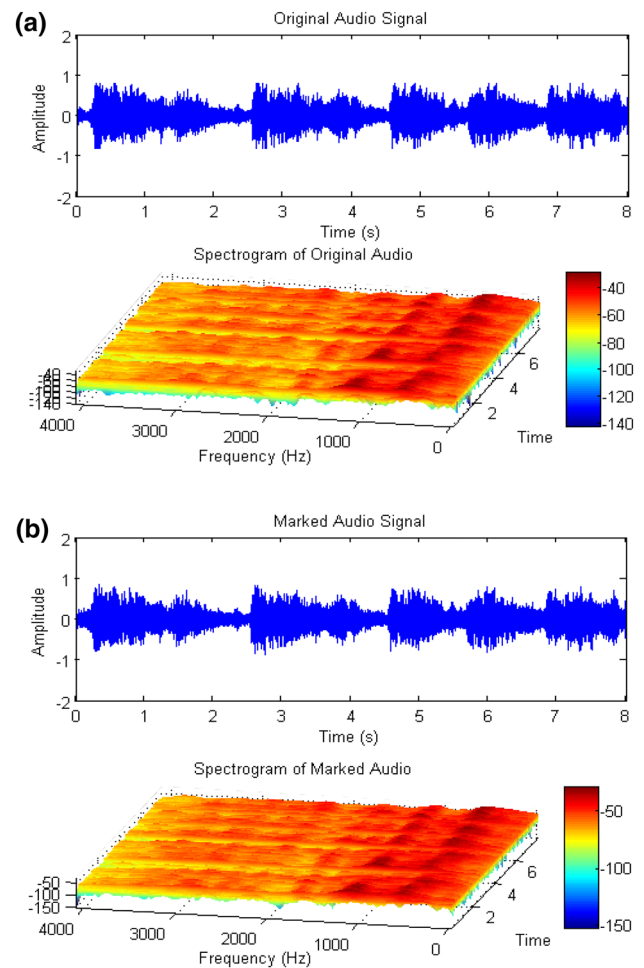
The spectrogram is simply defined as a visual representation of a frequency spectrum representing the sound or other signal as it varies with time or some other variable. It is commonly represented by a graph with two dimensions: the horizontal axis is assigned to time, and the vertical axis is assigned to frequency. In case of 3-D representations; a third dimension represents the amplitude of a dedicated frequency at a specific time which can be visually resembled by a color or intensity, or sometimes the height of a 3-D surface. The lower frequencies are more intensive than the higher frequencies (<http://en.wikipedia.org/wiki/Spectrogram>).

Figure 4 ensure that the proposed scheme does not deteriorate the marked signals.

#### 4.1.3 Correlation

Correlation coefficient can be used for signature verification. Moreover, the signal can be recovered again after signature removal and compared with the original signal showing differences close to zero.

Figures 5, 7, and 8 show some sample results.



**Fig. 4** **a** Original audio signal and its spectrogram, **b** Marked audio signal and its spectrogram

#### 4.1.4 Spectral distortion (SD)

Spectral distortion is an important metric that deserved consideration.

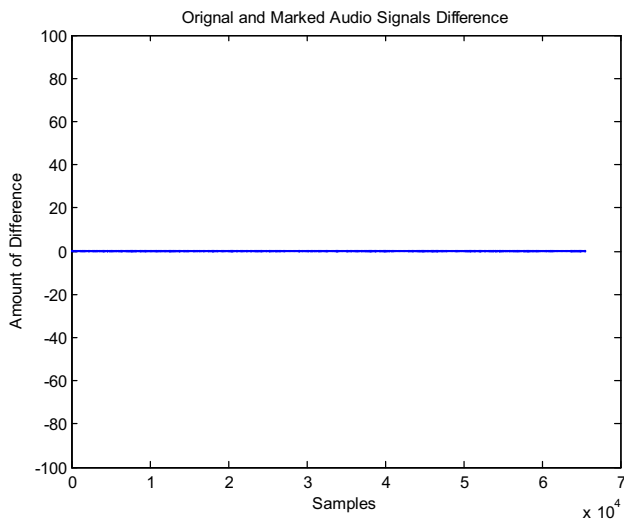
$$SD = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{i=Nm}^{Nm+N-1} |V_x(i) - V_y(i)| \quad (3)$$

where  $V_x(i)$  is the spectrum of original audio signal in  $dB$  for a certain segment and  $V_y(i)$  is a spectrum of marked audio signal in  $dB$  for the same segment. Smaller values of SD mean better quality (Kubichek 1993; Wang et al. 1992; Yang et al. 1998).

Table 1 show a summary of results.

### 4.2 Integrity verification process

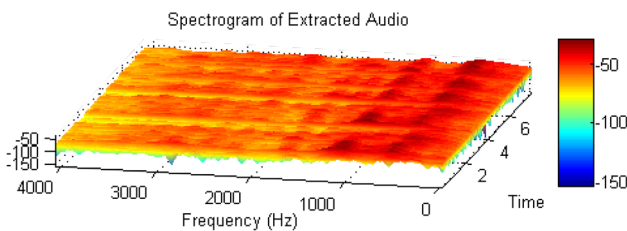
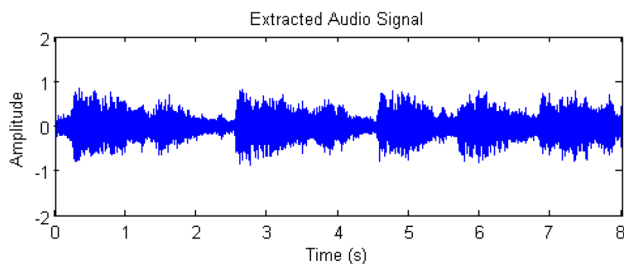
Figures 6, 7, and 8 in addition to Table 2 give good integrity verification results.



**Fig. 5** Original and marked audio signals difference

**Table 1** Numerical evaluation metrics for the integrity protection process

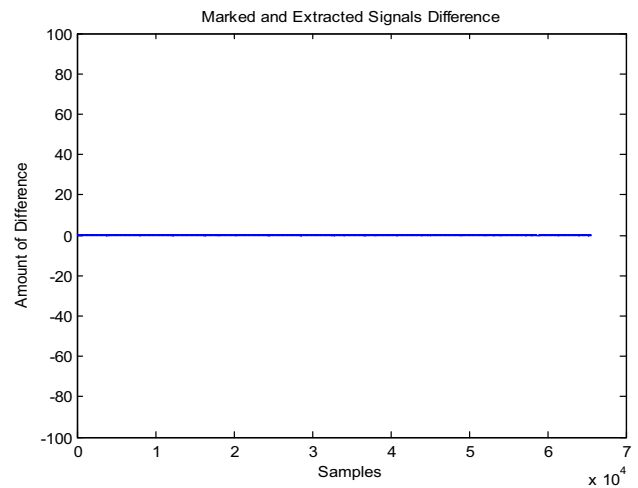
Evaluation metric	Value
Signal-to-noise ratio (SNR)	12.16 dB
Correlation coefficient (Cr)	0.97
Spectral distortion (SD)	0.4675 dB



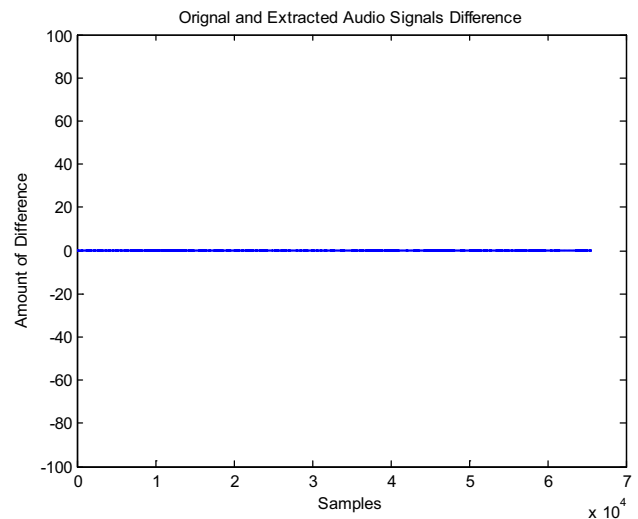
**Fig. 6** Extracted audio signal and its spectrogram

### 4.2.1 Shearing

Shearing is a common signal manipulation attack. During transmissions, some parts of transmitted audio signal may be lost intentionally or unintentionally due to channel conditions, jamming and so on. Figures 9, 10 show a sheared marked audio signal, and extracted audio signal.



**Fig. 7** Marked and extracted audio signals difference



**Fig. 8** Original and extracted audio signals difference

**Table 2** Numerical evaluation metrics for the integrity verification process

Evaluation metric	Value
Signal-to-noise ratio (SNR)	24.36 dB
Correlation coefficient (Cr)	0.999
Spectral distortion (SD)	0.3328

The results here reveal the reactivity of the proposed verification algorithm to the audio signal modification, in which any change in the marked signal even it was audibly undetected results in a great impact on the extracted audio signal. So, the extracted output audio signal quality can be considered as an integrity verification reference.

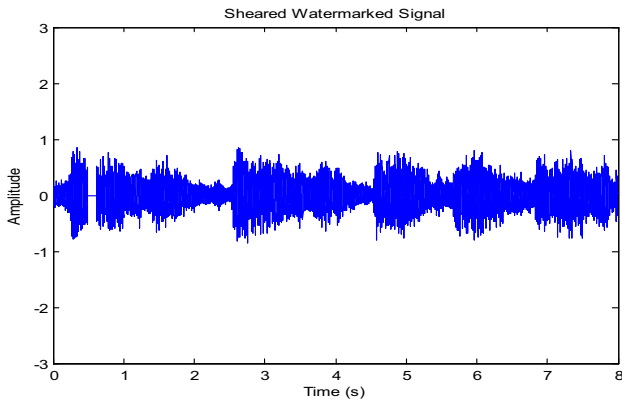


Fig. 9 Sheared marked audio signal

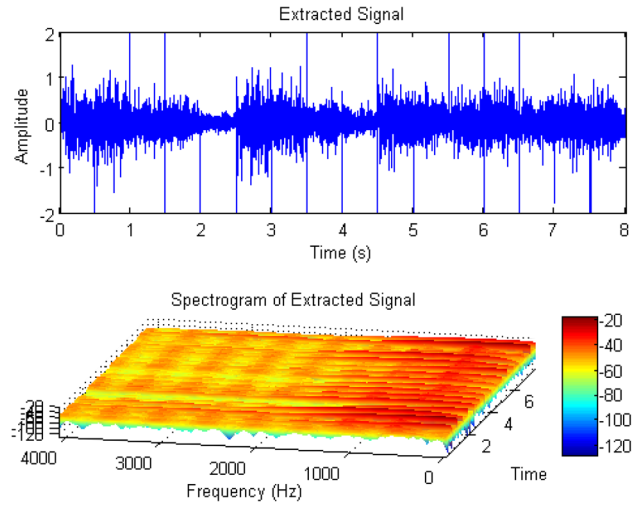


Fig. 12 Extracted distorted audio signal and its spectrogram

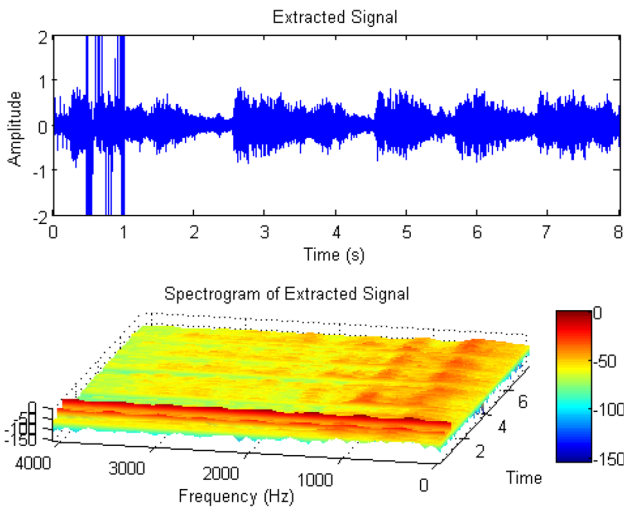


Fig. 10 Extracted distorted audio signal and its spectrogram

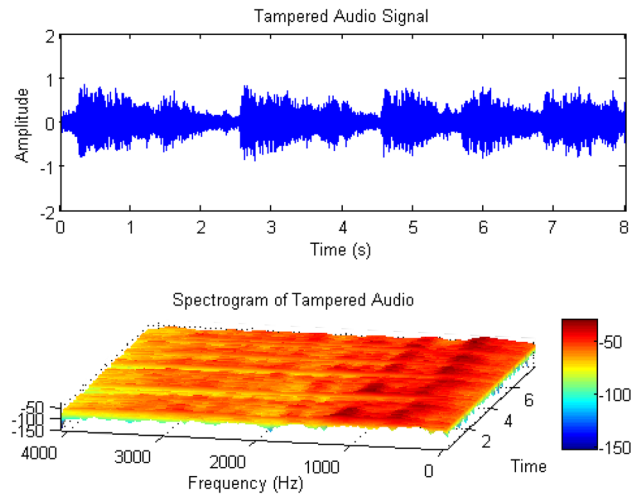


Fig. 13 Tampered audio signal and its spectrogram

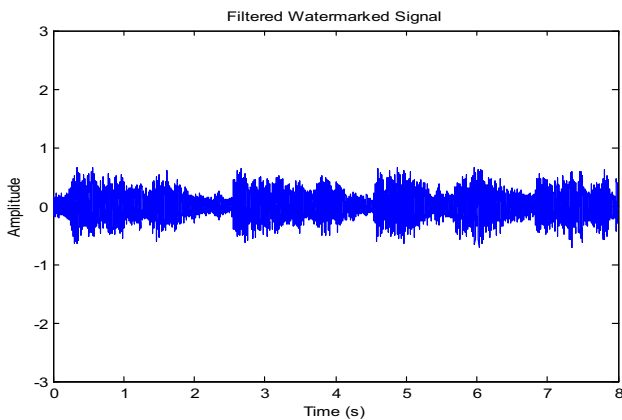


Fig. 11 Filtered watermarked audio signal

### 4.2.2 Low pass filter

The marked audio signal here is attacked by a third order Butterworth filter as shown in Fig. 11, the verification algorithm output is shown in Fig. 12, which shows the

extracted audio signal with a clear distortion revealing the impact of filtering by using verification processing.

### 4.2.3 Inaudible sample change

In this experiment, the audio signal originality is tampered by changing only two sample values at different positions throughout the audio waveform in which the tampered marked audio is intended to appear audibly and not affected as shown in Fig. 13. By verifying the tampered audio signal using the proposed integrity verification algorithm, the tampering is audibly detected and the extracted output audio signal shows larger changes than the tampered audio signal as shown in Fig. 14.

The results related to the difference between the extracted and tampered marked audio signal in case of mentioned audio manipulation attacks are shown in Figs. 15, 16 and 17.

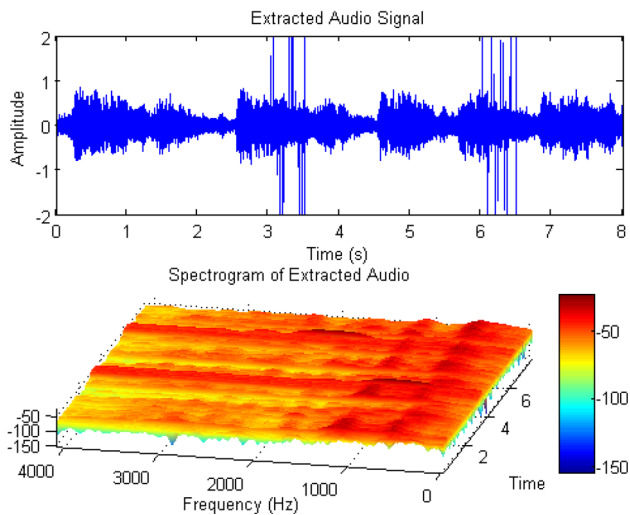


Fig. 14 Extracted audio signal and its spectrogram

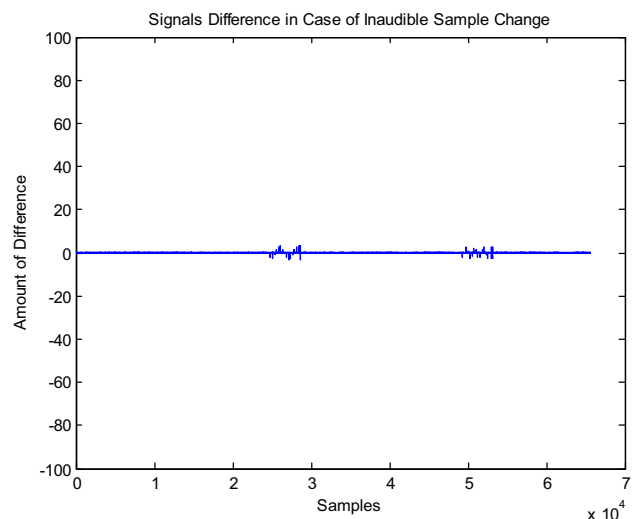


Fig. 17 Signal difference in case of inaudible sample change

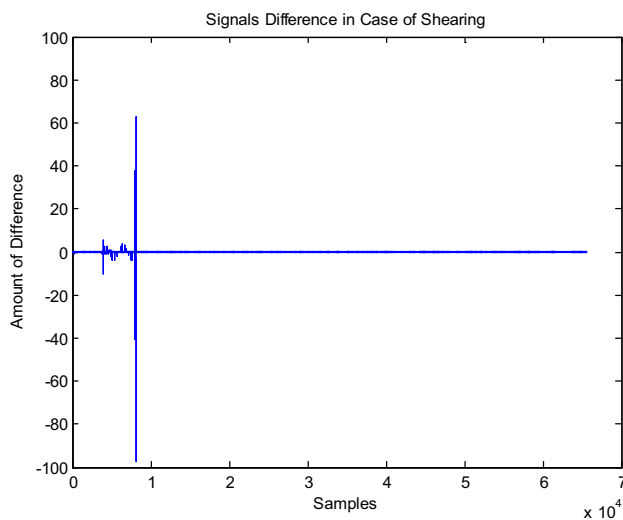


Fig. 15 Signal difference in case of shearing

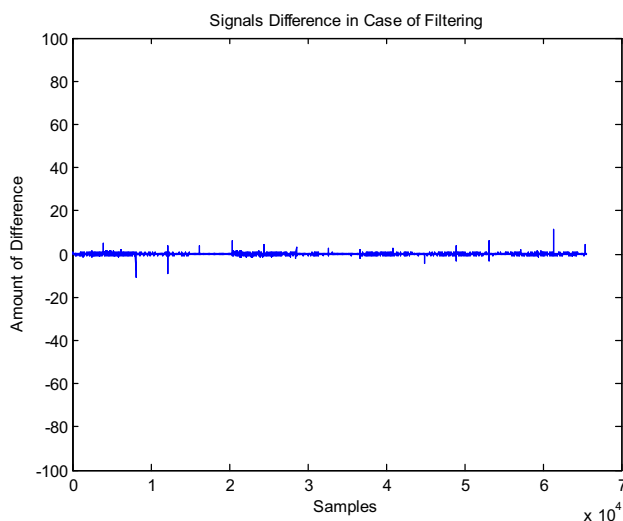


Fig. 16 Signal difference in case of filtering

Table 3 Correlation coefficient measurement in case of different attacks

Signals	Correlation coefficient value		
	Shearing	Filtering	Inaudible sample change
Marked and extracted signals	0.16	0.56	0.94

Table 4 Performance analysis in case of different audio samples without modification

Audio sample	Evaluation metric		
	Signal-to-noise ratio (SNR)	Correlation coefficient (Cr)	Spectral distortion (SD)
tenorsax.wav	23.4	0.997	1.27
Mutedtrumpet.wav	39.8	0.999	0.4661
Funky.wav	25.13	0.998	0.70
Flute.wav	22.4	0.997	0.93

Table 5 Performance analysis in case of different audio samples with shearing attack

Audio sample	Evaluation metric		
	Signal-to-noise ratio (SNR)	Correlation coefficient (Cr)	Spectral distortion (SD)
tenorsax.wav	0.14	0.18	Inf.
Mutedtrumpet.wav	0.639	0.36	Inf.
Funky.wav	4.35	0.79	Inf.
Flute.wav	5.3238	0.8390	Inf.

Also correlation coefficients are measured in the same circumstances of modification and tabulated in Table 3.

The results in Table 3 assures the ability of the proposed algorithm to detect changes in audio signals.

Tables 4 and 5 show more results on other data. They are in favor of the above-mentioned findings.

## 5 Conclusion

The paper presented a self embedding algorithm for integrity verification of digital audio signals. The concepts of image processing have been used here for signature or mark embedding. The suggested method is very simple and enables reconstruction of the original signal again. Simulation results have show that the suggested scheme is robust even in the presence of attacks. The proposed schemes is very appropriate in military as well as nuclear applications.

## References

- Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1), 90–93. doi:10.1109/T-C.1974.223784.
- Al-Haj, A., Bata, L., & Mohammad, A. (2009). Audio watermarking using wavelets. In: *Proceedings of first international conference on networked digital technologies* (pp. 398–403). NDT '09, July 28–31, 2009.
- Bassia, P., Pitas, I., & Nikolaidis, N. (2001). Robust audio watermarking in the time domain. *IEEE Transactions on Multimedia*, 3(2), 232–241.
- Can, Y. S., Alagoz, F., & Burus, M. E. (2014). A novel spread spectrum digital audio watermarking technique. *Journal of Advances in Computer Networks*, 2(1), 6–9.
- Chang, W. H., & Chang, L. W. (2010). Semi-fragile watermarking for image authentication, localization, and recovery using Tchebichef moments. In: *International symposium on IEEE communications and information technologies (ISCIT)*.
- Dittmann, J., Steinmetz, A., & Steinmetz, R. (1999). Content-based digital signature for motion pictures authentication and content-fragile watermarking. In: *International conference on multimedia computing and systems*, Florence, June 1999.
- Dittmann, J. (2001) Content-fragile watermarking for image authentication. In: *Proceedings of SPIE* (Vol. 4314). Bellingham.
- Gomez, E., Cano, P., de C. T. Gomes, L., Batlle, E., Bonnet, M. (2002) Mixed Watermarking-Fingerprinting Approach for Integrity Verification of Audio Recordings. In: *International telecommunications symposium—ITS2002*, Natal.
- Khayam, S. A. (2003). *The discrete cosine transform (DCT): Theory and application*. Michigan State University, March 2003.
- Kubichek, R. (1993). Mel-cepstral distance measure for objective speech quality assessment. In: *Proceedings of the IEEE Pacific Rim conference on communications, computers and signal processing* (pp. 125–128).
- Li, C.-T., & Yang, F.-M. (2003). One-dimensional neighborhood forming strategy for fragile watermarking. *Journal of Electronic Imaging*, 12(2), 284–291.
- Liu, C., & Chen, O. (2004). Fragile speech watermarking scheme with recovering speech contents. In: *The 2004 47th Midwest symposium on circuits and systems*. MWSCAS'04 (Vol. 2, pp. II-165–II-168).
- Lu, C., Liao, H., & Chen, L. (2000) Multipurpose audio watermarking. In: *Proceedings 15th international conference on pattern recognition* (Vol. 3, pp. 282–285), 2002.
- Nassar, S. S., Ayad, N. M., Kollash, H. M., Faragallah, O. S., M. El-Bendary, M. A., & Abd-Elsamie, F. E. (2014) Image contents verification algorithm using transform domain techniques. *International Journal of Computer Applications* (0975–8887), 97(9).
- Petrovic, R. (2005). Digital watermarks for audio integrity verification, Serbia and Montenegro, Nis, September 28–30, 2005.
- Radharani, S., & Valarmathi, M. L. (2010). A study of watermarking scheme for image authentication. *International Journal of Computer Applications*, 2(4), 24–32.
- Shaw, G. (2000) Digital document integrity. *8th ACM multimedia conference*. Los Angeles, November 2000.
- Shiva Kumar, K. B., Raja, K. B., Chhotaray, R. K., Pattnaik, S., (2010). *Coherent Steganography using Segmentation and DCT*, IEEE-978-1-4244-5967-4/10/\$26.00.
- Shoemaker, C. (2002). Hidden Bits: A Survey of Techniques for Digital Watermarking Independent Study EER-290. Prof Rudko, Spring 2002.
- Tewfik, A. H. (2000). *Digital watermarking*, San Mercury News. August 14, 2000.
- Wang, S., Sekey, A., & Gersho, A. (1992). An objective measure for predicting subjective quality of speech coders. *IEEE Journal on Selected Areas in Communications*, 10(5), 819–829.
- Wu, C.-P., & Jay Kuo, C.-C. (2001). Speech content integrity verification integrated with ITU G.723.1 speech coding. In: *IEEE international conference on information technology: Coding and computing* (pp. 680–684). Las Vegas, April 2001.
- Wu, C. & Jay Kuo, C.-C. (2002). Fragile speech watermarking for content integrity verification. In: *IEEE international symposium on circuits and systems*, 2002. ISCAS 2002 (Vol. 2, pp. II-436–II-439).
- Yang, W., Benbouchta, M., & Yantorno, R. (1998). Performance of the modified bark spectral distortion as an objective speech quality measure. In: *Proceedings of the IEEE international conference on acoustic, speech and signal processing (ICASSP)* (Vol. 1, pp. 541–544). Washington, May 1998.
- Zhang, P., Xu, S., & Yang, H. (2012) Robust and transparent audio watermarking based on improved spread spectrum and psychoacoustic masking. In: *Proceedings 2012 international conference on information science and technology (ICIST)* (pp. 640–643). March 23–25, 2012.