PAPER • **OPEN ACCESS**

Audio steganography applications using auditory features watermarking

To cite this article: Jepri Banjarnahor *et al* 2019 *J. Phys.: Conf. Ser.* **1230** 012073

Audio steganography applications using auditory features watermarking

Jepri Banjarnahor[1*], Saut Dohot Siregar[1], Oloan Sihombing[1], Mardi Turnip[1], Windania Purba[1], Siti Aisyah[1], Jaidup Banjarnahor[2]

[1]Department of Information System, Faculty of Technology and Computer Sciences, Universitas Prima Indonesia, Indonesia
[2]Department of Informatics Management, AMIK Medan Business Polytechnic (MBP) Medan, Indonesia

*jepribanjarnahor01@gmail.com

Abstract. Watermarking is a technique for hiding data or information on a digital multimedia (image, sound and video) and is not visible to the ordinary eye and is resistant to digitization (editing media, noising, blurring, etc.). While audio watermark is defined as a technique of hiding data or confidential information into an audio data for "boarding" (audio host), people do not aware of the existence of additional data on the data host. Coding phase is hiding data by exchanging the original phase of the initial segment of the sound signal with the absolute phase of the watermark signal, while maintaining the relative phase between the signal segments using different phase segments of the original signal. When the phase difference between the original signal and the modified signal is small, the difference in sound produced is not detected by human hearing. To recognize ownership of multimedia content, testing is done first by combining the cover and watermark files. Experiment result on audio files shows that audio files successfully compressed on 50% efficiency without reducing the duration of the audio. After compression process, One of our audio sample which size is 3.37MB can be reduce to 1.68MB without reducing the duration of the audio.

1. Introduction

As this technology grown up, the threat of piracy and copyright is very obvious to users. Watermarking is a process of securing the data from these threats, in which user identification (watermarking) is merged with the digital media at the sender end and at the receiver end this owner identification is used to recognize and authenticate the data. This technique can be applied to all digital media types such as image, audio, video and documents.

From many years researchers and developers worked in this area to gain the best results[1]. The main application of robust audio watermarking is to protect the developer or the author's copyright. Audio watermarking technique including three parts including extracted watermark generation and embedding of watermark[2].

Nowadays digital media such as video, audio and images have replaced the role of analogue media in various applications.The success of the application of digital media is due to several advantages possessed by digital media, such as noise-free transmission, dense storage, perfect copying, and ease of editing, but with the advantages of increasing the availability of digital media, falsifying problems also increase, causing deep matters relating to the protection of copyright ownership.

The Human Auditory System (HAS) is far more complex and less explored than the Human Visual System (HVS). HAS is very sensitive to the addition of normal sound distribution. Auditory features are filtering out hearing by deciphering input signals into the frequency of time domains using gamma tone filters [3]. The method chosen in the auditory feature is phase coding. One of the advantages of phase coding is the most effective method in terms of comparison of sound signals captured. Every beginning of the segment is design like this to maintain each segment still has a relationship and can maintain sound quality. The coding phase works based on the characteristics of the human auditory system which ignores weaker sounds if the two sounds come together.

This study has problem boundaries, namely the retrieval of data taken from the data that has been stored, the audio file used is in the form of wav format, the auditory feature method used is phase coding, media messages stored in audio files use the watermark technique in audio in wav format. Another limitation of HAS is its inability to detect the relative phase of different spectral components. It is the basis of interchanging hidden data with some particular components of the original audio signal. This method is called as phase coding and works well on the condition that changes in phase components are retained small [4]. By spreading data in the frequency domain, spread spectrum (SS) technique ensures an appropriate recovery of the watermarked data when communicated over a noise-prone channel. SS utilizes redundancy of data for degrading the error rate of data hiding. An M-sequence of code handles the data and is embedded in the cover audio. This sequence is known to sender and receiver and if some parts of these values are modified by noise, recovery of data is feasible by using other copies [5]. Pole locations of an APF are calculated from the power spectrum by several approaches. Finally, the data is hidden in some chosen APF parameters. According to calculations, all the above-mentioned techniques have higher resilience against noise additions in frequency domain (or transform domain). In order to compare and classify the audio watermarking methods, some criteria must be chosen and defined. Based on the literature, major criteria for analysis and comparison of watermarking methods are considered as robustness, security, and hiding capacity (payload).

Other parameters including the transmission environment and the application influence the evaluation criteria. For that, they should be considered for performance evaluation of every watermarking technique. In an application where multiple levels of coding and decoding are planned, evaluation of a criterion like robustness is not possible without considering the environment constraints [6].

2. Methods

Research experiment is started by inputing an audio file, through an audiotory feature compression. After these steps, user can choose to continue the watermarking process or not. If the user choose to continue the watermarking process then decompression process will be taken until the result is gained.

2.1. Coding phase

Coding phase is included in the transform audio watermarking group that works by changing the spectral content in the frequency domain of the signal. Phase Coding works based on the characteristics of the human auditory system HAS (Human Auditory System) which ignores weaker sounds if the two sounds come together [7]. Broadly speaking, the watermark data is made into noise with a amplitude that is weaker than the amplitude of the audio data, then combined [8]. An independent multi-band phase modulation is utilized for phase coding. In phase modulation method, phase alteration of the original audio signal is controlled to obtain imperceptibility of phase modifications. Phase components are determined by quantization index modulation (QIM). Then, the nearest “o” and “x” points are replaced with phase values of frequency bin to hide “0” and “1,” respectively. Therefore, phase coding achieves a higher robustness when perceptual audio compression is applied [9]. QIM was widely been used that improves the capacity of data hiding of phase coding by replacing the strongest harmonic with step size of $\pi/2n$ [10]. Phase coding has zero value of bit error rate (BER) when MP3 encoder is applied that demonstrates the high robustness of this method. As HAS is not sensitive to phase changes, an attacker simply can replace his/ her data with the real hidden data. S/he can apply frequency modulation in an inaudible way and modify the phase quantization scheme[11]

2.1.1. Encoding
The steps for inserting a watermark with phase coding techniques are as follows:

Divide the sequence of sounds into N segments, s [i], $0 \le i \le L-1$, where each segment has the same length which is equal to L.

Calculate FFT in each segment. The results of this calculation are X (k) for each segment where $0 \le k \le L-1$.

Calculate the phase value φ and amplitude A for each segment where *a k* is the real part of the FFT value and *b k* is the imaginary part $|A| = \sqrt{ak^2 + bk^2}$ $\Phi = \mathrm{Tan}^{-1}[\frac{bk}{ak}]$.

After that, calculate the phase difference between the adjacent phase segments $\Delta\phi_{N-1}(\varpi_k) = \Delta\phi_{N-1}(\varpi_k) - \Delta\phi_N(\varpi_k)$.

The absolute phase of the watermark data signal is added to the resulting phase difference. The watermark signal with length L w, w [j], $0 \le j \le Lw-1$, is presented as Φ data = $\pi / 2$ or Φ data = $-\pi / 2$ which represents bits 1 or 0.

Subscribe to the initial segment phase Φ'0 with the phase of the mark' data watermark signal.

Create a phase matrix for N & *g t*; 0 using different phases to maintain phase relativity between the sound segments. This is done to maintain signal continuity after the initial segment phase modification process.

Combine the modified segments into one.

Recalculate the new phase value and the calculated amplitude value to inverse FFT for each segment to return the signal to the time domain.

2.1.2. Decoding
A synchronization of the encoding process, Segment length and watermark must be known. In the process requires an original sound signal to do detection. The steps in watermark detection are illustrated in the flowchart diagram below

Take the first part of the sound signal where n is the known encoding segment length, s [0] . s [n-1]. The calculation is only done on the first n elements of the signal because the watermark data is inserted only in the original segment of the original sound signal.

Perform FFT on the n-1 of the signal, then find the phase value Φ.

Conversion of phase values is obtained, $\pi / 2$ become bits 1 and $-\pi / 2$ become bits 0 as long as the bit of the watermark. The phase value is compared with the phase value of the original sound signal.

Data obtained from the watermark bits in accordance with the results of the conversion. The watermark bits are compared with the original watermark bits to find out the truth.

2.2. Fast Fourier Transform
Fast Fourier Transform is the source of an algorithm to calculate the Discrete Fourier Transform (DFT) quickly, efficiently and inversely. Fast Fourier Transform (FFT) is applied in a variety of fields from digital signal processing and solving partial differential equations to algorithms for multiplying large numbers of integers. There is also the basic class of the FFT algorithm, which is decimation in time (DIT) and decimation in frequency (DIF). Generally, the fast word is interpreted because the FFT formulation is much faster than the previous fourier transform algorithm calculation method [12].

3. Results and Discussion

NO.	COVER TEST.MP3	DURATION / SIZE	COVER TEST.WAV	DURATION / SIZE
1.	Eagles - Hotel California.Mp3	20 seconds / 160 KB	Eagles - Hotel California.wav	20 seconds / 3,37 MB
2.	Coldplay - A sky full of stars.Mp3	23 seconds / 191 KB	Coldplay - A sky full of stars.wav	23 seconds / 4,03 MB
3.	Crazy Frog – I like to move it .Mp3	29 seconds / 457 KB	Crazy Frog – I like to move it .wav	29 seconds / 4,88 MB
4.	The Chainsmokers - Selfie.Mp3	30 seconds / 473 KB	The Chainsmokers - Selfie.wav	30 seconds / 5,05 MB
5.	Daft Punk - Da Funk.Mp3	40 seconds / 627 KB	Daft Punk - Da Funk.wav	40 seconds / 6,75 MB

The first audio file sample to test is "Test file6.wav". This file has duration of 7 seconds the size is 224 KB. The second sample is "Test file3.Wav" which has duration of 14 seconds and the size is 448 KB. The third sample is "Test file.wav". which has duration of 8 seconds and the size is 260 KB. The fourth sample is "Test file2.wav" which has duration of 3 seconds and the size is 104 KB and the last sample is is "Test file.wav" which has duration of 11 seconds and the size is 348 KB.

NO.	NAME OF THE TEST FILES	DURATION / SIZE
1.	Test file6.wav	7 seconds / 244 KB
2.	Test file3.wav	14 seconds / 448 KB
3.	Test file.wav	8 seconds / 260 KB
4.	Test file2.wav	3 seconds / 104 KB
5.	Test file.wav	11 seconds / 248 KB

The watermark will be inserted into the test cover. The first cover is "Eagles - Hotel California.wav" and "Test file6.wav file" as watermark. The ringtone duration is 20 seconds and the size is 3.37 MB. The watermark duration is 7 seconds and the size is 224 KB. The results of the watermark insertion into the cover are named "Test.wav" which the duration is 20 seconds and the size is 1.68 MB.

In the second test, "Coldplay - A sky full of stars.wav" is the cover and "Test file3.wav" as watermark. The ringtone duration is 23 seconds and the size is 4.03 MB. The watermark duration is 14 seconds and the size is 448 KB. The results of the watermark insertion into the cover are named "Test2.wav" which the duration is 23 seconds and the size is 4.03 MB.

In the third test, "Crazy Frog - I like to move it. Wav" is the cover and "Test file.wav" as watermark. The ringtone duration is 29 seconds and the size is 4.88 MB. The watermark duration is 8 seconds and the size is 260 KB. The results of the watermark insertion into the cover are named "Test3.wav" which the duration is 29 seconds and the size is 2.44 MB.

In the fourth test, "The Chainsmokers - Selfie.wav" is the cover and "Test file2.wav file" as watermark. The ringtone duration is 30 seconds and the size is 5.05 MB. The watermark duration is 3 seconds and the size is 104 KB. The results of the watermark insertion into the cover are named "Test4.wav" which the duration is 30 seconds and the size is 2.52 MB.

In the fifth test, "Daft Punk - Da Funk.wav" is the cover and "Test file.wav" as watermark. The ringtone duration is 40 seconds and the size is 6.75 MB. The watermark duration is 11 seconds and the size is 348 KB. The results of the watermark insertion into the cover are named "Test5.wav" which the duration is 40 seconds and the size is 6.75 MB.

NO.	COVER	FILE	ENCODE RESULT	DURATION / SIZE
1.	Eagles – Hotel California.wav	Test file6.wav	Test.wav	20 seconds / 3,37 MB
2.	Coldplay - A sky full of stars.wav	Test file3.wav	Test2.wav	23 seconds / 4,03 MB
3.	Crazy Frog – I like to move it .wav	Test file.wav	Test3.wav	29 seconds / 4,88 MB
4.	The Chainsmokers - Selfie.wav	Test file2.wav	Test4.wav	30 seconds / 5,05 MB
5.	Daft Punk – Da Funk.wav	Test file2.wav	Test5.wav	40 seconds / 6,75 MB

After the encoding results from the first test obtained, "Test.wav" will be called as the extraction target. The results of the encoding process in the first test are named "ResultTest.wav" which the duration is 20 seconds and the size is 1.68 MB.

After the encoding results from the second test obtained, "Test2.wav" will be called as the extraction target. The results of the encoding process in the second test are named "ResultTest2.wav" which the duration is 23 seconds and the size is 2.01MB.

After the encoding results from the third test obtained, "Test3.wav" will be called as the extraction target. The results of the encoding process in the third test are named "ResultTest3.wav" which the duration is 29 seconds and the size is 2.44MB.

After the encoding results from the fourth test obtained, "Test4.wav" will be called as the extraction target. The results of the encoding process in the fourth test are named "ResultTest4.wav" which the duration is 30 seconds and the size is 2.52 MB.

After the encoding results from the fifth test obtained, "Test5.wav" will be called as the extraction target. The result of the encoding process in the fifth test was named "Result Test5.wav" which the duration is 40 seconds and the size is 3.37MB.

NO.	EXTRACTION TARGET	DURATION / SIZE	DECODE RESULT	DURATION / SIZE
1.	Test1.wav	20 seconds / 3,37 MB	Result1.wav	20 seconds / 1,68 MB
2.	Test 2.wav	23 seconds / 4,03 MB	Result2.wav	23 seconds / 2,01 MB
3.	Test 3.wav	29 seconds / 4,88 MB	Result3.wav	29 seconds / 2,44 MB
4.	Test 4.wav	30 seconds / 5,05 MB	Result4.wav	30 seconds / 2,52 MB
5.	Test 5.wav	40 seconds / 6,75 MB	Result5..wav	40 seconds / 2,52 MB

4. Conclusion

The Conclusion from the research result are: The application is able to produce all encode and decode processes that can run according to auditory features, and the larger the cover file is used, the longer the extracted time will be.Experiment result on audio files shows that audio files successfully compressed on 50% efficiency without reducing the duration of the audio. After compression process, One of our audio sample which size is 3.37MB can be reduce to 1.68MB without reducing the duration of the audio.

Reference

[1] Saini LK, Shrivastava V. 2014.A Survey of Digital Watermarking Techniques and its Applications. (IJCST) – Volume 2 Issue 3, May-Jun 2014.
[2] Ren P 2014.A New Digital Audio Watermarking Algorithm.The Open Automation and Control Systems Journal, 2014, 6, 1038-1044.
[3] Al-Haj A, “A dual transform audio watermarking algorithm”, Multimedia Tools and Applications, Vol. 73, No.3, pp. 1897-1912, 2014.
[4] Ngo, NM, and Unoki M 2016. Method of audio watermarking based on adaptive phase modulation. IEICE transactions on information and systems 99(1): 92–101.
[5] Li R., Xu S, and Yang H 2016. Spread spectrum audio watermarking based on perceptual characteristic aware extraction. IET Signal Processing.
[6] Djebbar F, et al. 2012. Comparative study of digital audio steganography techniques. EURASIP Journal on Audio, Speech, and Music Processing 2012(1): 1–16.
[7] Popa R. 1998. An Analysis of steganographic Techniques, The "Politehnica" University of Timisoara, Faculty of Automatics and Computers, Department of Computer Science and Software Engineering.
[8] Wang DL and Brown GJ, Ed., 2006, Computational Auditory Scene Analysis: Principles, Algorithms, and Applications. Hoboken, NJ: Wiley-IEEE Press.
[9] Bender W, et al. 1996. Techniques for data hiding. IBM Systems Journal. 35(3.4): 313–336.
[10] Dong X., Bocko MF, and Ignjatovic Z. 2004. Data hiding via phase manipulation of audio signals. In IEEE international conference on acoustics, speech, and signal processing, 2004. Proceedings.(ICASSP’04). IEEE.
[11] Juanda K and Supangkat, Suhono H. 2000, Watermarking Sebagai teknik penyembunyian hak cipta pada data digital. Nematollahi, M.A. and Vorakulpipat, Rosales,H.H 2017. Digital Watermarking Techniques ang Trends. Springer. 203 (42).
[12] Nematollahi MA, Vorakulpipat, Rosales HH 2017. Digital Watermarking Techniques ang Trends. Springer. 203 (42).