

Modified Phase Coding Audio Watermarking Resistant to Signal Attacks

Aree A. Mohammed

Faculty of Science, Computer Dept.
University of Sulaimani Sulaimani, Iraq

Diman M. Mohammed

Faculty of Science, Computer Dept.
University of Sulaimani Sulaimani, Iraq

ABSTRACT

In this paper, a new phase coding method is used to embed any watermark data type (text, image, audio) in audio signals. The basic idea is to segment the audio signal and transform it using the Discrete Fourier Transformation (DFT) and then embedding the watermark in the initial phase coefficients for every four segments. The used cover audio files of WAV type. The watermark can be extracted semi-blindly. Performance of the proposed watermarking methods was tested by using normalized correlation (NC) and bit error rates (BER) to measure the robustness, and signal to noise ratio (SNR) to evaluate the imperceptibility of our proposed methods. Experimental results show superiority of the proposed scheme in terms of robustness, inaudibility and capacity compared with the traditional phase coding method, the best audio test sample shows that the SNR value was above (20 dB) for (250 byte) watermark capacity with high robustness against cropping attack and showed modest robustness against some type of attacks.

General Terms

Multimedia Security

Keywords

Audio watermarking; Phase coding; DFT; Semi-blind detection; Robustness; inaudibility.

1. INTRODUCTION

Digital watermarking has been proposed as a new, alternative method to enforce the intellectual property rights and protect digital media from tampering [1]. A digital watermark is a signal added to the original signal which can be later extracted or detected and it should not degrade the quality of the digital data [2]. Digital watermarking is used for numerous applications which include: copyright protection, finger printing, and tamper proofing, broadcast monitoring [3]. The simplest requirements of information hiding in digital audio is so called magic triangle: Inaudibility, Robustness to attack and watermark data rate [4]. Inaudibility is the perceptual similarity between the original audio signal and the watermarked audio signal. Data rate is a number of bits that can be embedded in to the audio signal with in unit of time, robustness stands for the resistance of the watermark against common signal processing and malicious attacks. No algorithms are known to satisfy all of the above requirements. Watermarking algorithms aim at achieving suitable trade-offs among the requirements [2]. Audio watermarking is more challenging than image and video mainly due to two reasons. Firstly, audio signals contain only one dimensional data, thus it's difficult to hide additional information without compromising the quality of audio signal. Secondly, the human auditory system (HAS) is more sensitive than the human visual system (HVS), thus a small degradation in the quality will be noticed by the listeners [5]. Several Different

methods enable watermarking in the time domain and frequency domain. Watermark is embedded in frequency domain using one or hybrid transformations of a signal by using DFT [6], DCT [7], and DWT [8]. Li et al [9] proposed a novel content dependent localized robust audio watermarking scheme to combat Synchronization attacks like random cropping and time-scale modification. The basic idea is to first select steady high-energy local regions that represent music edges like note attacks, transitions or drum sounds by using different methods, then embed the watermark in these regions. Bhaskari et al [10] presented a novel blind audio watermarking scheme to embed a meaningful data into digital audio by encoding the data using godelization technique and then the data is embedded into the digital samples based on a private key. The proposed watermarking algorithm can extract the hidden data which is embedded without the help of the original audio signals. Alrawi et al [11] used the phase coding method to embed the watermark, by using FFT method in two ways. The first way used the block size of the wave data equal to (2^2) and the second way used the block size of the wave data equal to (2^3) . In this paper a new digital audio watermarking scheme is presented by using the phase coding method. The main objective of the research is to improve the robustness and the inaudibility of the digital audio watermarking in the frequency domain. The general diagram of the proposed audio watermarking scheme is shown in Figure-1.

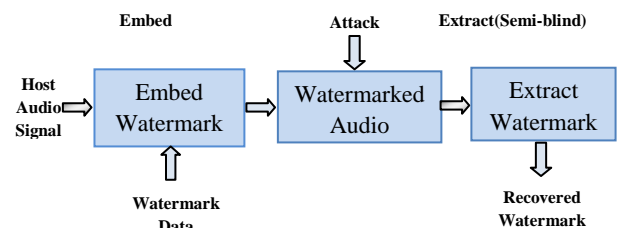


Fig-1: General diagram of the proposed watermarking scheme.

The rest of this paper is organized as follows. In section 2 the proposed watermark embedding and extraction processes are explained. Section 3 presents Performance test analysis with some discussions. Conclusions and Future Suggestions are given in section 4 and 5 respectively.

2. PROPOSED AUDIO WATERMARKING SCHEME

In the following subsections the details of the proposed scheme are described.

2.1 watermark Types

The types of the used watermark are as follow:

1. Text.
2. Image.
3. Audio WAV files.

Different watermark types are used.

2.2 Watermark Embedding Process

The main steps of the embedding algorithm are described below.

Input: the WAV audio file and the watermark. **Output:** Watermarked WAV audio file.

Step1: Select audio.wav as a carrier.

Step2: Load audio data.

Step3: Skip 44 byte of a carrier the address header part wav file.

Step4: Get audio samples as array of bytes.

Step5: Divide the sound sequence into a series of short segments with 4 byte length and create 2- dimensional matrix.

Step6: Apply a discrete Fourier transform (DFT) to the segments, and create a matrix of the phase (Ph), and magnitude (Mag).

Step7: Calculate and store the phase difference between each adjacent segment.

Step8: Read watermark file.

Step9: Let the initial phase of the first signal segment to be $\pi/2$ if the encoded bit is 0 or $-\pi/2$ if the encoded bit is 1, then jump three segment and add the 2nd bit to the initial phase of the second signal segment and so on until all the watermark bit added.

Step10: The resulting phase vector Ph' should be the addition of the phase of the preceding one and the according phase difference we stored in step 7. Re-create phase matrixes for $n > 0$ by using the phase difference as illustrated in equation (2.1)

$$\begin{bmatrix} Ph'_1 = Ph'_0 + \Delta Ph_1 \\ Ph'_n = Ph'_{n-1} + \Delta Ph_n \end{bmatrix} \quad (2.1)$$

Step11: Use the modified phase matrix Ph' and the original magnitude matrix Mag to reconstruct the sound signal by applying the inverse DFT.

Step12: Find imperceptibility measure (SNR) between original and watermarked audio. In Figure-2 the flowchart of the embedding process of the modified phase coding method is shown.

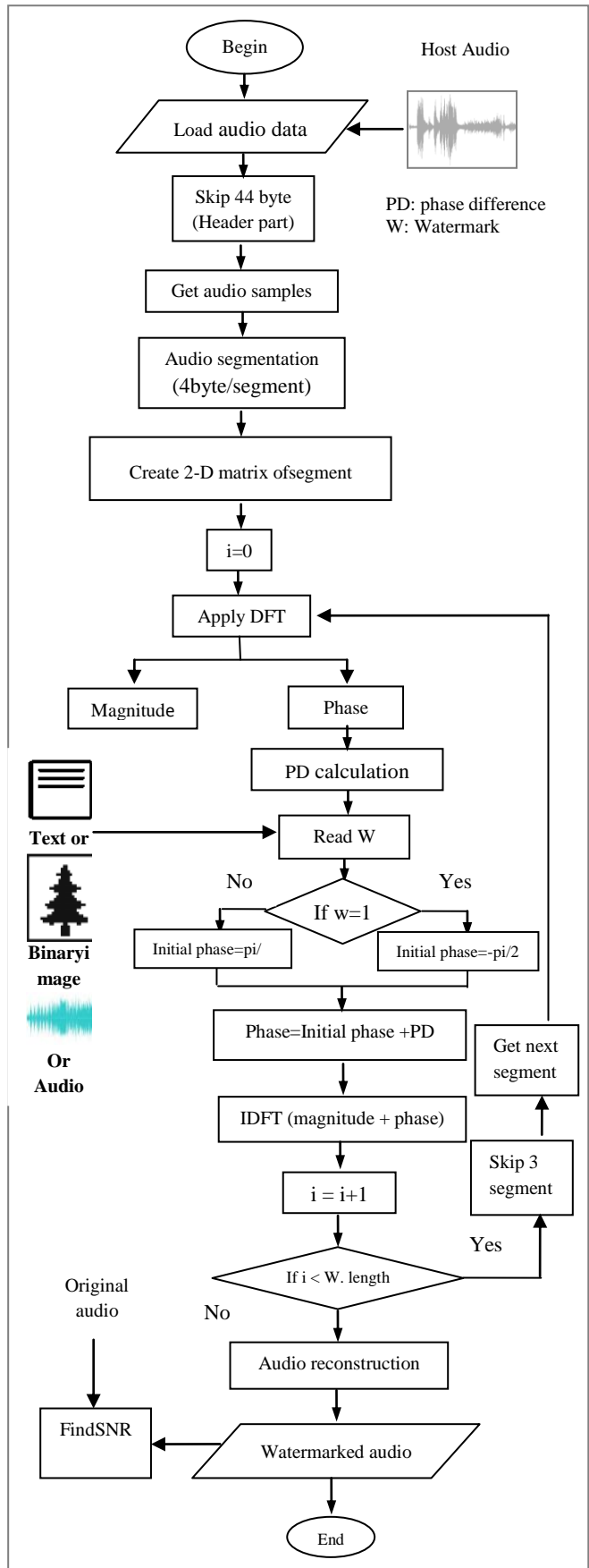


Fig-2: Modified phase coding embedding process

2.3 Watermark Extracting Process

The Extracting process is performed in a Semi- blind approach. That is, the length of the segment, the watermark length must be known. The extraction step can be summarized as follows:

Step1: Load watermarked audio wav file.

Step2: Skip 44 byte the address header part of wav file.

Step3: Get audio samples as array of bytes.

Step4: Partitioning the watermarked audio into segments and create 2 dimensional matrix, each segment has the same size used in the embedding process.

Step5: By applying the DFT the watermarked block is transformed to frequency domain to get the real and imaginary coefficients, and then phase coefficients are calculated.

Step6: Get the initial phase of the first signal segment, the value of the phase is detected as follow:

$$retrived\ bit = \begin{cases} 1 & \text{if } phase = \pi \\ 0 & \text{otherwise} \end{cases} \quad (2.2)$$

Then jump three segments and get the 2nd bit of the initial phase of the second signalsegment and so on until all the watermark bits are reconstructed.

Step7: Convert the reconstructed bit to text or image or audio as it is used in the embedding process .In Figure -3 the flow chart of extracting process of the given scheme is shown.

2.4 Audio Watermarking Attacks

The performance of the proposed algorithm against some type of attack such as(echo, cropping, smoothing filter and amplification) are performed to evaluate the robustness of our scheme by using Gold Wave (5.58) and Audacity(2.0.3) audio editing tools .Figure-4 shows the block diagram of attacking in frequency domain.

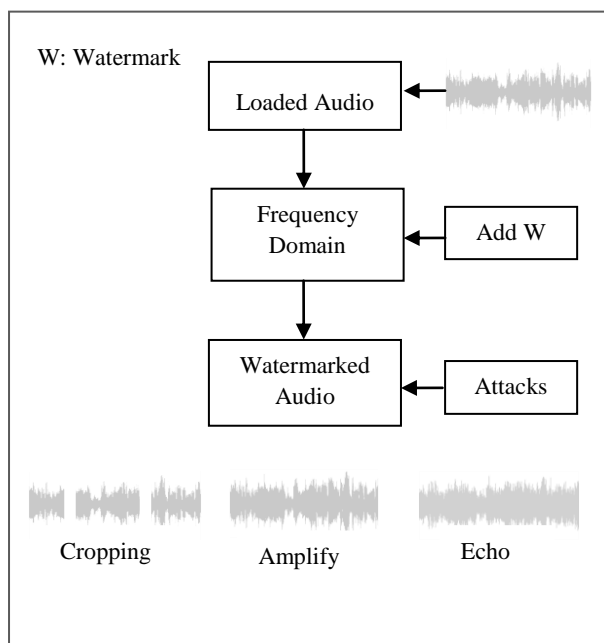


Fig-4: Attacking process diagram

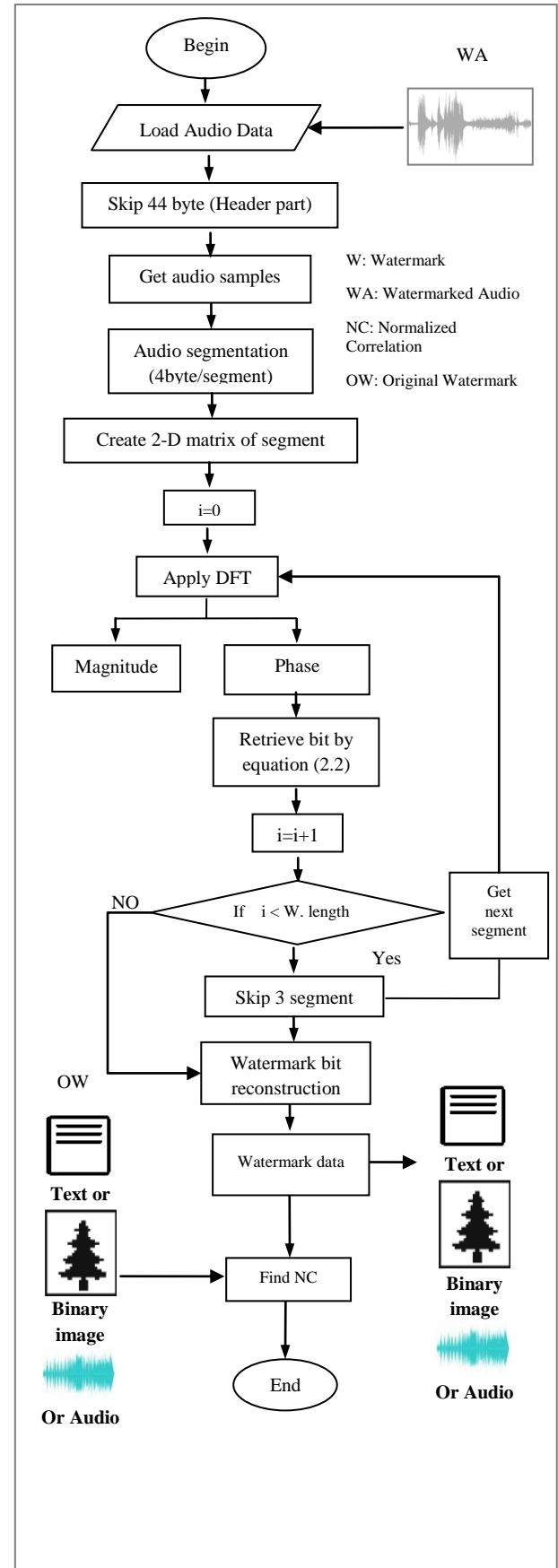


Fig-3: Modified phase coding extracting process

3. PERFORMANCE TEST ANALYSIS AND DISCUSSION

In this research, Table-1 shows the selected cover audio WAV files which are used to measure and evaluate the performance of the proposed audio watermarking algorithm.

Table-1: Wav file types

NO.	File name	Channel Type	Sample Resolution(bits)	Sampling Rate(KHz)
1	classic	Stereo	16	44.1 KHz
2	Test	Mono	16	44.1 KHz
3	pop	Stereo	16	44.1 KHz
4	Song	Stereo	16	44.1 KHz
5	Speech	Mono	16	8000 KHz

Figure 5, 6, 7 shows the original, watermarked and attacked (echo) audio wav files. The performance is evaluated with respect to imperceptibility and robustness.

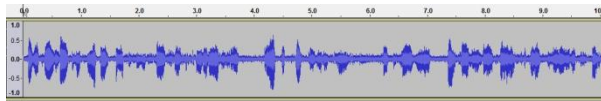


Fig-5: Original audio

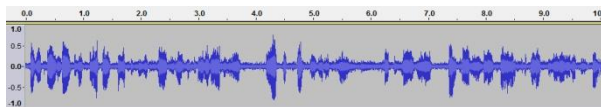


Fig-6: Watermarked audio

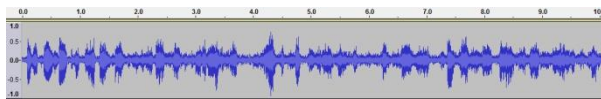


Fig-7: Attacked audio

3.1 Imperceptibility Tests

To measure the Imperceptibility, SNR (Signal-to-Noise Ratio) is used as an objective measure according to equation (3.1) below:

$$SNR(dB) = 10 \log_{10} \frac{\sum_n A_n^2}{\sum_n (A_n - A'_n)^2} \quad (3.1)$$

Where A corresponds to the original signal, and A' corresponds to the watermarked signal. The listening test (MOS) is used as a subjective measure according to Table-2 because the SNR does not take into account the specific characteristics of the human auditory system.

Table-2: MOS grading scale.

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

The SNR values and the MOS grading points for each wave file are shown in Table-3 with different watermark capacity.

Table-3: SNR and MOS values

Watermark Size (byte)	File name									
	Classic		Test		Pop		Song		Speech	
	(SNR)	MOS	(SNR)	MOS	(SNR)	MOS	(SNR)	MOS	(SNR)	MOS
12	33.6	2.5	27.6 1	4.8	33.1	4.8	32	4	23	4.8
60	26.7 6	2.3 3	22.7 5	4	27.2 2	4.8	26.9 3	4	16.9 1	2.5
90	25.2 8	2.3	21.0	3	25.5	4.5	25.1 8	3.5	15.0 1	2.4
128	23.6 7	2.2 1	19.6 6	2.5	22	4.5	23.8 2	3.5	13.3 4	2.4
256	20.6 2	2.1	16.7 1	2.3	20.7 8	4.2	20.1 5	3	10.3 4	2

For the perceptual quality of the modified phase coding, it shows an increase in SNR value above 20 dB after embedding 250 byte. Table-4 shows the difference between our proposed schemes with the traditional phase coding in [12].

Table-4: Comparison with SNR and MOS for traditional techniques

Reference	Algorithm	SNR	MOS
ULudag	DC-Level shifting	21.24	3.35
Bender	Echo	21.47	3.60
Bender	Phase	12.20	2.44
Bender	LSB	67.91	4.90
Cox	Spread spectrum	28.59	4.46
Swanson	Frequency Masking	12.87	2.93
Our Scheme	Modified Phase coding	36.41	4.5

3.2 Robustness against attack

The performance of the proposed algorithm against signal processing such as echo, cropping, smoothing filter and amplification is performed to evaluate the robustness of our scheme. To measure the similarity between the original watermark and the watermark extracted from the attacked watermark using the Normalized Correlation NC, which is computed as shown in equation (3.2):

$$NC(w, w') = \frac{\sum_{i=1}^N w_i w'_i}{\sqrt{\sum_{i=1}^N w_i^2} \sqrt{\sum_{i=1}^N w'^2_i}} \quad (3.2)$$

Where N is the number of pixels in watermark, w and w' are the original and extracted watermarks respectively. The Bit-error rate (BER) is defined as the ratio of incorrect extracted

bits to the total number of embedded bit, the BER (in percent) is given by equation (3.3) below:

$$BER = \sum_{n=0}^{B-1} \begin{cases} 1, w'(n) \neq w(n) \\ 0, w'(n) = w(n) \end{cases} \quad (3.3)$$

Where B is a watermark length bits, w_n correspond to the n th the bit of the embedded watermark and w'_n corresponds to the n th the bit of the recovered watermark. Pop and Classic audio wav files showed good robustness against the previous attacks. The robustness of the proposed method for the two

audio files with binary image watermark against attacks are showed in Table-5 and Table-6. The problem with this watermarking type that they are less robust to signal processing and malicious attacks such as audio compression, our proposed scheme can survive the cropping attack after cropping about 8000 sample in 10 different positions because the watermark distributed over the entire data set, unlike the traditional phase coding the watermark is existed in the first segment.

Table-5: Extracting image before and after attack (Classic)

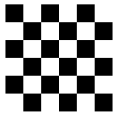
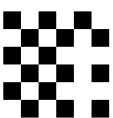
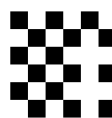



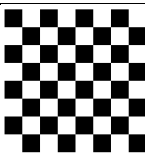
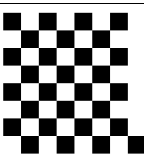
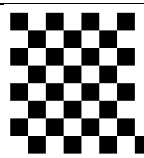
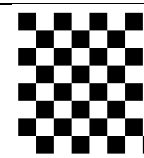
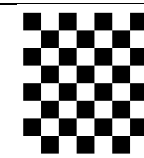
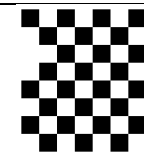
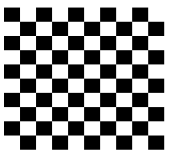





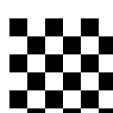
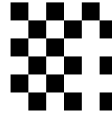
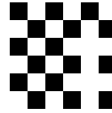
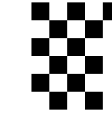

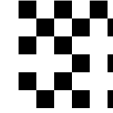
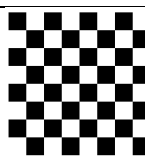
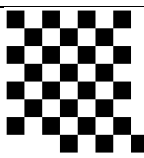
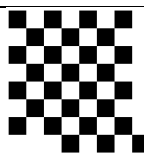
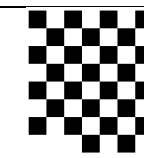
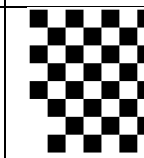
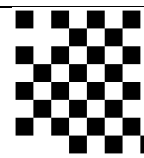
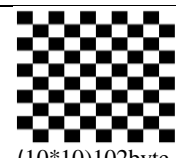
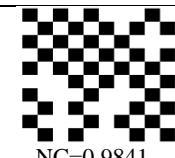
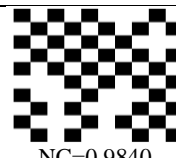
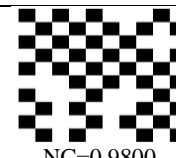
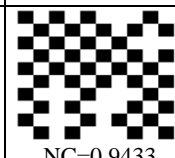
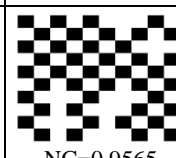
Original watermark	No attack	Echo attack	Cropping	Amplify	Smoothing
 (6*6) 86byte	 NC=0.9425 BER=0.754	 NC=0.9425 BER=0.905	 NC=0.9173 BER=1.357	 NC=0.9409 BER=0.905	 NC=0.9417 BER=1.207
 (8*8) 94byte	 NC=0.9526 BER=0.688	 NC=0.8793 BER=1.10	 NC=0.950 BER=0.688	 NC=0.8909 BER=0.825	 NC=0.950 BER=1.238
 (10*10)102byte	 NC=0.9796 BER=0.626	 NC=0.9741 BER=0.9741	 NC=0.9793 BER=0.751	 NC=0.9768 BER=0.900	 NC=0.9693 BER=1.126

Table-6: Extracting image before and after attack (Pop)

Original watermark	No attack	Echo attack	Cropping	Amplify	Smoothing
 (6*6)86byte	 NC=0.9552 BER=1.025	 NC=0.8726 BER=0.905	 NC=0.8929 BER=0.603	 NC=0.9095 BER=1.056	 NC=0.9497 BER=1.81
 (8*8)94byte	 NC=0.9632 BER=0.963	 NC=0.9631 BER=0.825	 NC=0.9003 BER=1.238	 NC=0.9488 BER=0.688	 NC=0.9167 BER=1.238
 (10*10)102byte	 NC=0.9841 BER=0.876	 NC=0.9840 BER=0.876	 NC=0.9800 BER=1.126	 NC=0.9433 BER=1.520	 NC=0.9565 BER=2.503

In Table -7, 8, comparative with respect to existing technique used DWT can also be carried out.

Table-7: Extracted watermarks after various StirMark attacks on the watermarked pop signal.





Original Watermark (108*57) gray-scale image	
Echo attack	 NC=0.06
Amplify	 NC=0.865
Smooth	 NC=0.910

Table-8: Comparison with SNR and MOS with existing technique

Reference	Algorithm	Audio Type	SNR	MOS
Ali Al-Haj [12]	DWT-Based Audio Watermarking	Pop	25.0317	5.0
Our	Modified Phase coding	Pop	20.78	4.2

4. CONCLUSIONS

The experimental results have shown that the proposed audio watermarking scheme can achieve better imperceptibility by embedding large amount of data than those used in the traditional phase coding, when an attack is applied on the watermarked audio, then normalized correlation for longer watermark code word is better than the shorter one. An improvement of robustness is obtained against cropping attack because the watermark bits are distributed over the entire data set and the type of the cover media has a direct effect on the NC between embedded and extracted watermark. As the human auditory system is more sensitive to relative phase difference, a high watermark bitrates cannot be embedded in modified phase coding method.

5. FUTURE SUGGESTIONS

In the following some suggestions are derived as future work:

1. Using the modified phase coding method with other audio file types like (mp3) which is an online and compressed audio format.
2. Improving the robustness using two different transformation techniques (i.e., hybrid DFT and DWT).

3. Adaptive quantization could be applied on the phase coefficients to decrease the distortion or noise of the watermarked audio.

4. Increasing a watermark data rates (capacity) by embedding into the magnitude coefficients of the sample in expense of (imperceptibility and robustness).

5. Proposed schemes can be adapted for mobile based applications.

REFERENCES

- [1] Adya, M. (2008). "Audio Watermark Resistant To MP3 Compression ", *M.Sc.thesis*, Indian Institute of Technology, Kharagpur, India.
- [2] Vivekanada, B.K., Sengupta, I. and Das, A. (2011) "An audio watermarking scheme using singular value decomposition and dither-modulation quantization", *Multimedia Tools Appl* (2011)52:369-383.
- [3] Smitha R, M. S., Jyothsna, A. N and Pinaka, P. R. (2012) " Digital Watermarking: Applications, Techniques and Attacks", *International Journal of Computer Applications* (0975 – 8887), Volume 44– No.7, pp.29-34.
- [4] Cvejic, N. and Seppanen, T. (2007) " Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks", Published in the United States of America and United Kingdom.
- [5] Natgunanathan, I. (2012). "Novel Watermarking Methods For Copyright Protection of Audio Signals", *Ph. D.dissertation*, Deakin University.
- [6] Smith, S.W., (1999). "Digital Signal Processing", Second Edition, California Technical Publishing, USA.
- [7] Sharma, S., Rajpurohit, J. And Dhankar, S. (2012) "Survey on Different Level of Audio Watermarking Techniques", *International Journal of Computer Application* (0975-8887), volume49-No.10.
- [8] Yassine, H., Bachir, B. and Aziz, K. (2012), "A Secure and High Robust Audio watermarking System for Copyright Protection", *International Journal of Computer Applications* (0975 – 8887) Volume 53– No.17.
- [9] Li, w., Xue, x. and Lu, P. (2006) " Localized Audio Watermarking Technique Robust Against Time-Scale Modification", *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 8, NO. 1.
- [10] Bhaskari, D., Avadhani, P. S and Damodaram, A. (2008) " A Blind Audio Watermarking Scheme for embedding text using GÖDELIZATION technique ", *IETECH Journal of Advanced Computations*, Vol: 2, No: 4, pp.209 - 213.
- [11] Alrawi, S. S., Abdulshaheed, R. and Alhadithy, A. A. (2011). " Watermarking in WAV Files Bases on Phase Coding", *Eng. & Tech. Journal*, Vol.29, No.4, pp.770-780.
- [12] Al-Haj, A., Mohammad, A. and Bata, L. (2011), " DWT–Based Audio Watermarking", *The International Arab Journal of Information Technology*, College of Electrical Engineering, Princess Sumaya University, Faculty of Engineering and Technology, University of Jordan, Jordan , Vol. 8, No. 3.