Comparison of Robustness with Two Different Keys using Time Spread Echo Method

Vaishali Passi[1], Pankaj Shukla[2]
[1,2]Department of Electronics and Communication,
University College of Engineering,
Rajasthan Technical University, Kota

***Abstract*— Audio watermarking is generally used as a multimedia Copyright protection or as a scheme that involves computing in audio signals. Audio watermarking is a method in which audio signal works as an envelope for covering information. This technology can also use the security of audio files. One of the usual methods in audio watermarking is echo hiding. Echo hiding method in which time spreading is done using (PN) sequence as a confidential key, overcomes the problem of less security which prevails in most of the echo hiding techniques. The imperceptibility and the robustness of time-spread watermarking based on echo of audio signals compared between the well-known PN sequence and the gray code of PN sequence. This paper also compares the affect of various attacks.**

Keywords*— Audio watermarking; echo hiding; time-spread echo; PN sequence; PN gray sequence.

I. INTRODUCTION

As the multimedia technology & communication networks are getting advanced day by day, it is resulting in replica and handling of digital media, without any deprivation. Devastation of copyrights is of virtually grievous concerns in the area of audio expansion of contents of digital media. So, there is a strong need of security of digital media. Every year music industry bears a loss of millions of dollars because of illegal copying and dispersion of music. As a measure to resolve this problem digital watermarking is done [1], [2]–[4], [5], [11] [16]. From the technological viewpoint, watermarking in digital form is contrived by embedding any key/ information into the media such that the owner can extract these data whenever essential to avow their copyright [11], [16], [20]. Although watermarking in digital form is suitable to multimedia information which includes audio, image, video, and text. This paper is confined to audio watermarking.

Echo hiding offers several benefits from various perspectives, for example, modest encryption and decryption, robustness and imperceptibility, etc. [4]. It's challenging to obscure information without lowering the character of audio signal because only 1-dimensional information can be enclosed in audio signals. Moreover, the human audile system is extra sensitive than human ocular system [3], [16].

Robustness, security and imperceptibility are the three vital issues which have to be conceived in watermarking of audio. Robustness points out towards the ability to recover the watermark data through the watermarked signal, equally in the circumstances of without attacks and that of with attacks. Imperceptibility means that the watermark has to be noiseless in the watermarked audio signal. Security denotes the fact, that the method of watermarking must employ a confidential key for safeguarding from an illegal user. The watermark cannot be extorted by the absence of confidential key.

In the past era, a number of watermarking methods have been suggested for audio, grounded on various techniques by means of spread spectrum [3], [4], [16], phase coding [6], [20], masking [3], [8], [23], patchwork [10], and echo hiding [1], [2], [5], [12]–[14], [20]. Echo hiding method of watermarking have a preference among these schemes, as for good perceptual quality the inserted echo induce the similar perceptual and statistical features same as the original signal.

Echo hiding is advantageous with characteristics like simple encryption & decryption processes, robustness, etc. [5], [14]. To a prominent level, in the functioning of the watermarking system, echo kernel plays a vital role.

In watermarking in which single-echo is used [14], to hide single watermark, one large echo is employed which consequences in low imperceptibility. Multiple small echoes of dissimilar delays [2] are employed to obscure a watermark, in order to overcome the disadvantage of single echo watermarking. However, this echo implanting is confined to distribution of time delay of echo. There are restrictions in time-slots for allocating various bits to preserve imperceptibility, because it is challenging to realize in simple multiple echoes.

As in that respect, no confidential key to differentiate between correct information and malicious tampering, echo hiding doesn't ensure robustness against malicious attack using multiple encoding [21], [24].

The method in which spreading of echo in time is an alternate to a single or multiple echo of echo hiding methods. This paper, compares amongst spreading an echo using pseudo noise (PN) sequence and spreading an echo using gray coding of pseudo noise (PN) sequence, the amplitude of every single echo turn out to be low, meanwhile its spectrum of power goes approximately flat . It's challenging to decrypt the implanted information without employing PN and its gray sequence employed in the encoding process, and may perhaps results in good imperceptibility over and above results in colorless and natural quality of audio after watermarking. The decoding operation of suggested method is then evaluated as functions of length and amplitude of the sequences employed

at the time of spreading. Besides, comparison is done amongst the decoding performance using PN sequence and PN gray sequence. Furthermore, evaluation is done on the basis of robustness of the suggested method against distinctive assails, like resampling, addition of noise at different decibels. Finally, investigation is done for the imperceptibility of the suggested method through a hearing test.

II. ECHO KERNEL BY TIME SPREADING USING PNGRAY SEQUENCE AND PN SEQUENCE

Where a direct sound and its reflected sound coexist, traditional hiding of echo is done. Therefore, in a direct audio signal insertion of watermark is done at a single constant dislocation. Due to multiple reflectivity's from the walls and other physical objects of the room, reverberation is perceived consisting of both direct audio signal & its various echoes.
Whenever watermarks are inserted on several dislocations from direct audio signal, perplexed echoes may provide a more natural sound tone in contrast to a single and various multiple echoes, yet if those echoes are perceptible [25], [26]. Proposed method is based on this thought.
In this method, for temporally spreading a single echo in time domain we use a PN sequence and PN gray code, which acts as confidential key at the time of decoding the inserted information from a watermarked signal. In this paper we compare the performance by using PN sequence and PN gray sequence. PN gray sequence is the sequence which is generated by gray coding of PN sequence. The kernel is created from a PN sequence as

$$f(n)=\delta(n)+\alpha P(n-\Delta),\ 0<\alpha<<1 \quad (1)$$

where P(n) is an normal PN sequence having amplitude of ±1, α is the amplitude of the P(n) and δ (n) is the Dirac delta function.
Linear convolution of the original/host audio signal and the echo kernel which is time-spreaded is done to obtain the watermarked signal means original audio signal and echoes both lies in the signal which is produced by convolving the echo which is time time-spreaded with the original/host audio signal. Thus, denotation of the watermarked signal is done as follows:

$$wm(n)=o(n)*f(n) \quad (2)$$

where o(n) denotes the host/original audio signal, f(n) is the echo kernel which is time spreaded, and * denotes linear convolution.
In the case of PN gray sequence the gray coding of PN sequence is done. Then, the proposed kernel is created from a PN gray sequence as

$$f_g(n)=\delta(n)+\alpha.P_g(n-\Delta),\ 0<\alpha<<1 \quad (3)$$

where $P_g(n)$ is PN gray sequence with amplitude is 1,0, α is the amplitude of the PN gray sequence and δ (n) is the Dirac delta function. Now, the watermarked signal is denoted as follows

$$w_gm(n)=o_g(n)*f_g(n) \quad (4)$$

where $o_g(n)$ denotes the host/original audio signal, $f_g(n)$ denotes echo kernel which is time spreaded, and * denotes linear convolution.

III. DECODING PROCEDURE

Original/host audio signal component from a watermarked signal is obtained by de-spreading using the same PN sequence and same PN gray sequence which is used at the time of encoding.
De-spreading is done as follows:

$$o(n)= wm(n)* P(n) \quad (5)$$

where o(n) is original host signal, wm(n) is watermarked signal, P(n) is PN sequence.

$$o_g(n)= w_gm(n)* P_g(n) \quad (6)$$

where $o_g(n)$ is original host signal, $w_gm(n)$ is watermarked signal, $P_g(n)$ is gray code of PN sequence.

IV. EVALUATION

For evaluating the method, the pseudo noise (PN) sequence, PN gray sequence and the decoding function of watermarking is used. Simulations are executed to give the demonstration of robustness and imperceptibility of the proposed method used for watermarking in contrast to, hiding of echo based methods in [5] and [1]. Hearing test is conducted for judging the perceptual tone of the proposed method.
Every sound clip used as host audio signal has duration of 20s. These clips are sampled at the rate of 44.1 kHz, quantized with 16bits. At various values of α, PN sequence and PN gray sequence is applied to implant watermarks on each audio clip. The size of the pseudo noise (PN) sequence and PN gray sequence is taken to be 1023.

A. Imperceptibility Test

Imperceptibility is measured using hearing test and measurement of quantity. By these tests imperceptibility of the PN gray sequence watermarking method is measured.

1) Hearing Test: A particular index of human perceptional system can only estimate the imperceptibility. It can be estimated by a hearing test. Such test is performed by utilizing original/host audio signal. For this the length of PNgray code is taken to be 1023. The value of α is taken to be 0.0001, 0.005, 0.001, 0.05, 0.01, 0.1 and 0.2. For the hearing test, three female subjects and two male subjects with normal hearing capacity took part.This test is performed for analysing the differentiation of sound signal quality between a original/host audio signal and a watermarked audio signal. Subjects were able to distinguish the sound signal quality

between the two signals beyond the value of α=0.2. This implies that our method gives better imperceptibility at α= 0.0001, 0.005, 0.001, 0.05, 0.01, and 0.1. In this test the original audio and watermarked audio signals are listened through same speakers.

2) Measurement of Quantity: For a performance index signal to noise ratio (SNR) is as follows:

$$SNR(dB) = 10\log10(\sum_n o(n)^2/(\sum_n(wm(n) - o(n))^2) \quad (7)$$

For original/host audio signals, audio clips which are used belongs to different audio groups as mentioned below:

TABLE I. ORIGINAL /HOST AUDIO SIGNALS USED FOR EVALUATION

Host Audio signals	**Genres**
A1	Famous English songs
A2	Classical
A3	Folk
A4	Jazz
A5	pop

TABLE II. IMPERCEPTIBILITY UNDER DIFFERENT VALUES OF α

α	**Value of SNR**		**Evaluation Under Different Attacks**
	Using PN Sequence	***Using PN Gray sequence***	
0.0001	SNR = 25.89870	SNR=25.912311	No Attack
	SNR1= 22.928133	SNR1=22.941767	Resampling at 22050 Hz
	SNR2=25.903686	SNR2=25.917008	AWGN at 40dB
	SNR3=25.903994	SNR3=25.917331	AWGN at 100dB
0.005	SNR=25.298204	SNR= 25.491681	No Attack
	SNR1= 22.457911	SNR1=22.269135	Resampling at 22050 Hz
	SNR2=25.302918	SNR2=25.496307	AWGN at 40dB
	SNR3=25.303222	SNR3=25.496703	AWGN at 100dB
0.001	SNR= 25.832668	SNR= 25.975918	No Attack
	SNR1= 22.887100	SNR1=22.972122	Resampling at 22050 Hz
	SNR2= 25.837351	SNR2=25.980853	AWGN at 40dB
	SNR3= 25.837689	SNR3=25.980999	AWGN at 100dB
0.2	SNR= -0.735297	SNR=12.237592	No Attack
	SNR1= -6.817549	SNR1=4.285984	Resampling at 22050 Hz
	SNR2= 0.595141	SNR2=12.246507	AWGN at 40dB
	SNR3= 0.596102	SNR3=12.247599	AWGN at 100dB

TABLE III. COMPARISON OF IMPERCEPTIBILITY UNDER DIFFERENT VALUES OF α

Different Methods	**Value of SNR for α=0.01**	**Value of SNR for α=0.1**
Method in [5]	9.91	-10.08
Method in [1]	-3.68	-23.38
Method in [19]	19.19	-0.81
Method using PNgray	26.230423	18.298103

For calculation of imperceptibility reason of utilizing SNR is described as follows. By smaller magnitudes of α, an objective index of higher SNR is obtained, where an original /host audio signal has predominant power. Hence, value of signal to noise ratio depends on the amplitude of PN sequence. Table II indicates the SNRs for the method of watermarking of PN gray sequence and the method using normal PN sequence with different kinds of attacks. It is viewed that the normal PN sequence method has the lowest SNR at α= 0.2 among the method using PN gray sequence. Perceptual quality of the watermarked audio signal is reduced because of the increase in the noise level of the watermarked audio signal in the method of normal PN sequence. Table III denotes the SNRs for the method using PN gray sequence and watermarking methods in [5], [1] and [19]. It can be seen that the method in [5], [1], and [19] has the lowest SNR at α=0.01 and at α=0.1 in comparison to PN gray method. It can be noted from Table III that the method with PN gray sequence has a good perceptual quality in comparison to the other methods, which is demoed by considerably higher SNRs

.

B. Robustness

Robustness against intentional and unintentional attacks is essential for any watermarking method. Robustness performance can be evaluated by different algorithms under attack, for this we different attacks are used: (i) resampling at 22050Hz, (ii) adding white Gaussian noise (AWGN) at 40dB, 100dB. These attacks and results are listed in Table II, where the SNR has been calculated.

V. CONCLUSION

The main thought for this proposed method is to spread an echo by utilizing PN gray sequence in the time domain. To decode the implanted data pseudo noise (PN) sequence works as a confidential key but with PN gray sequence secrecy is greater. Results show that the imperceptibility is greater with good decoding performance of PN gray sequence method. Furthermore, proposed method provides much more natural audio quality, which is supported by simulations and hearing test. Reasonable tolerance to typical attacks like AWGN and resampling is demonstrated by the proposed method.

REFERENCES

[1] S. Siahpoush and Y. Erfani, "Robust audio watermarking using improved TS echo hiding," Digit. Signal Process., vol. 19, pp. 809–814, 2009.

[2] H. Oh, J., J. W. Hong, W. Seok and D. H. Youn, "New echo embedding technique for robust and imperceptible audio watermarking," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 2001, pp. 1341–1344.

[3] N. Cvejic, T. Seppanen, and A. Keskinarkaus, "Audio watermarking using M-sequences and temporal masking," in Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics, 2001, pp. 227–230.

[4] D. Kirovski and H. Malvar, "Robust spread-spectrum audio watermarking," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 2001, pp. 1345–1348.

[5] R. Nishimura, B.-S. Ko, and Y. Suzuki, "Time-spread echo method for digital audio watermarking,"IEEE Trans. Multimedia, vol. 7, no. 2, pp.212–221, Apr. 2005.

[6] H. S. Malvar and D. Florencio, "Improved spread spectrum: A new modulation technique for robust watermarking," IEEE Trans. Signal Process., vol. 52, no. 4, pp. 898–905, Mar. 2003.

[7] I. J. Cox, T. Leighton, J. Kilian and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[8] W. Bender,, N. Morimoto, D. Gruhl and A. Lu, "Techniques for data hiding," IBM Syst. J., vol. 35, no. 3–4, 1996.

[9] M. D. Swanson, A. H. Tewfik, B. Zhu and L. Boney, "Robust audio watermarking using perceptual masking," Signal Process., vol. 66, no. 3, pp. 337–355, 1998.

[10] M. Arnold, "Audio watermarking: Feature, applications and algorithms," in Proc. IEEE Int. Conf. Multimedia Expo., 2000, vol. 2, pp. 1013–1016.

[11] W.-N. Lie and L.-C. Chang, "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification," IEEE Trans. Multimedia, vol. 8, no. 1, pp. 46–59, Feb. 2006.

[12] C. Xu, J. Wu, Q. Sun, and K. Xin, "Applications of digital watermarking technology in audio signals," J. Audio Eng. Soc., vol. 47, no. 10, pp. 805–812, Oct. 1999.

[13] H. O. Oh, H. W. Kim, J. W. Seok, J. W. Hong, and D. H. Youn, "Transparent and robust audio watermarking with a new echo embedding technique," in Proc. IEEE Int. Conf. Multimedia Expo., 2001, pp. 433–436.

[14] D. Gruhl and W. Bender, "Echo hiding," in Proc. Information Hiding Workshop, Cambridge, U.K., 1996, pp. 295–315.

[15] E. Zwicker and E. Terhardt, "Analytical expressions for critical band rate and critical bandwidth as a function of frequency," J. Acoust. Soc. Amer., vol. 68, no. 5, pp. 1523–1525, Nov. 1980.

[16] H. Wang, R. Nishimura, Y. Suzuki, and L. Miao, "Fuzzy self-adaptive digital audio watermarking based on time-spread echo hiding," Appl. Acoust., vol. 69, pp. 868–874, 2008.

[17] S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay, Speech Processing in the Auditory System. New York: Springer, 2004.

[18] A. Menezes, P. V. Oorschot, and S. Vanstone, Hand Book of Applied Cryptography. Boca Raton, FL: CRC, 1997.

[19] Yong Xiang, Dezhong Peng, Iynkaran Natgunanathan, and Wanlei Zhou, "Effective pseudonoise sequence and decoding function for imperceptibility and robustness enhancement in time-Spread echo-based audio watermarking," IEEE Transactions On Multimedia, Vol. 13, No. 1, February 2011

[20] O. T.-C. Chen and W.-C. Wu, "Highly robust, secure, and perceptual- quality echo hiding scheme," IEEE Trans. Audio, Speech, Lang. Process., vol. 16, no. 3, pp. 629–638, Mar. 2008.

[21] S. Katzenbeisser and F. A. P. Petitcolas, Eds., Information Hiding Techniques for Steganography and Digital Watermarking. Norwood, MA: Artech, 2000.

[22] Y. Yardymcy, A. E. Cetin, and R. Ansari, "Data hiding in speech using phase coding," in Proc. Eurospeech 97, 1997, vol. 3, pp. 1679–1682.

[23] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," in Proc. Eur. Signal Processing Conf., Trieste, Italy, Sep. 1996.

[24] S.-K. Lee and Y.-S. Ho, "Digital audio watermarking in the cepstrum domain," IEEE Trans. Consumer Electron., vol. 46, no. 3, pp. 744–750, Aug. 2000.

[25] B.-S. Ko and H.-Y. Kim, "Modified circulant feedback delay networks (MCFDN's) for artificial reverberator using a general recursive filter and CFDN's," J. Acoust. Soc. Korea, vol. 18, no. 4E, pp. 31–36, 1999.

[26] W. G. Gardner, "The virtual acoustic room," B.S. thesis, Dept. Comput. Sci. and Eng., MIT, Cambridge, MA, Sep. 1992.