A Survey On Audio Watermarking Methods

Nihala K. N, M.Tech student, Viswajyothi College of Engineering and Technology

---***---

Abstract - *As the use of digital multimedia data is increasing day by day, the necessity of protecting the data from illegal copyright infringement, fingerprinting and other malicious attacks is also increasing. Watermarking is the process of embedding a piece of secret data to the host data to protect from these attacks. In audio watermarking a piece of data is embedded into the original host signal without affecting the signal quality. There are different techniques available for audio watermarking. In this paper audio watermarking along with its properties is explained and different watermarking techniques are also examined and the comparison result shows the advantages and disadvantages of each technique which implies watermarking based on Fibonacci numbers satisfies most of the watermarking properties.*

***KeyWords*:** Audio watermarking, Watermark embedding and Watermark extraction.

1.INTRODUCTION

Watermarking is the technique of hiding secret digital information in a carrier signal. A digital watermark is a piece of data to be embedded in digital media such as an audio, video or image data. It is used to identify the copyright ownership of the data or to protect it against illegal application. As digital audio files can be downloaded and copied easily they are more prone to copyright infringements. Audio watermarks are secret signal data used to prevent these problems. These signals are embedded in the original signal data and extracted by detection mechanism. One of the main factor in which audio watermarking scheme depends on is the imperfection of the human auditory system. There are several methods to embed the watermark data into the original audio signals.

An efficient audio watermarking system must mainly satisfy four basic requirements. These are robustness, security, imperceptibility and payload. Some of the watermark properties are discussed below.

Robustness : Digital signal may undergo different signal processing operations. Robustness is the property of the watermark data to withstand these signal attacks and to be extracted even after these attacks.

Security : The watermark signal must not reveal about the embedded watermarked data. So the system will not be prone to unauthorized access. The security of the watermarking system lies in the security of the secret key used not on the embedding algorithm.

Imperceptibility : Imperceptibility refers to the quality of the signal after watermark embedding. The quality of the signal must be retained even after embedding data for an effective watermarking method.

Payload : Payload refers to the amount of secret data that can be added to the original signal without effecting the quality of the original signal. For audio signals, data payload refers to the number of watermark bits that can be reliably embedded within the original signal per unit of time. Payload is usually measured in bits per second (bps).

Different audio watermarking techniques are being compared in this paper based on their embedding procedure. The rest of this paper is organized as follow: Section 2 describes various watermarking techniques and Section 3 contains the comparison between different watermarking techniques and finally we have the conclusion at Section 4.

2. LITERATURE SURVEY

In [1] Mohamed F. Mansour and Ahmed H. Tewk proposes a new system for data embedding. The interval length between the audio signal's salient points are changed. Then the embedding is done in the quantization indices of the intervals. The wavelet extrema of the signal envelope are used as the salient points. The algorithm is robust against different signal processing attacks. But watermark detection depends on the selection of threshold and is not robust against pitch shifting . The salient point lengths are modified for embedding the data. The salient point interval lengths are quantized after extracting and refining them in the extraction process.

Wei Li, Xiangyang Xue, and Peizhong Lu proposes a novel content-dependent localized robust audio watermarking scheme , to combat synchronization attacks in [2]. First select the steady high-energy local regions that represent music edges by using different methods .The methods are either selects embedding regions on the original audio waveform directly, select the peaks on the audio envelope as reference to determine the embedding regions or selection based on music content analysis .After selecting the region then embed the watermark in these regions. Such regions will not be changed much for maintaining high auditory quality and in this way, the embedded watermark is robust against different distortions. The prime factor in watermark detection is accurately locating the embedding region . If the embedding region identification is inaccurate then the watermark detection is bound to be failed, because it will detect the watermark in areas where the watermark does not exist at all.

In [3] Nima Khademi Kalantari .et.al proposes a multiplicative patchwork method.In this system two subsets of the host signal features are selected and the selection is based on the secret keys inorder to embed watermark data within the host signal. The watermark data is embedded by multiplying one of the subsets .The other subset is leaved unchanged. Embedding is performed in the selected frames of the host signal which satisfies a certain condition.The wavelet domain is used for the implementation of the method and approximation coefficients are used for data embedding. T he inaudibility of the watermark insertion is controlled by Perceptual Evaluation of Audio Quality (PEAQ) algorithm .The error probability is being derived. Energy of the subset is changed according to the embedding functions and due to multiplication. In order to extract the watermark data the energy ratio of two subsets should be compared with a threshold.

Another technique was introduced by Xiangui Kang, Rui Yan and Jiwu Huang in [4],they proposes a multi-bit spread-spectrum audio watermarking scheme based on a geometric invariant log coordinate mapping (LCM) feature which is very robust to audio geometric distortions. The embedding of watermark is done in the LCM feature. Actually embedding is done in the Fourier coefficients which are mapped to the feature via LCM. the embedding is actually performed in the DFT domain without interpolation, and this will completely eliminate the severe distortion formed due to the non-uniform interpolation mapping. One of the advantage of the watermarked audio is its high auditory quality in both objective and subjective quality assessments. A mixed correlation between the LCM feature and a key-generated PN tracking sequence is proposed to align the log-coordinate mapping, thus synchronizing the watermark efficiently with only one FFT and one IFFT.

In [5] Ryouichi Nishimura, proposes a watermarking technique for audio signals, based on the spatial masking phenomenon and ambisonics. Ambisonics is a well known technique to encode and reproduce spatial information related to sound.A watermark data is created from the original sound scene by slightly rotating it. Ambisonic signals are synthesized from mono or stereo signals. Watermarks are embedded near the signal by adding a small copy of the host signal which provides reversible watermarking is possible if loudspeakers are arranged properly for playback. This will represent audio signals of first order ambisonics as mutually orthogonal four-channel signals.And this make it possible to present a larger amount of data than in the original version.

Yong Xiang.et.al proposes a novel dual-channel time-spread echo method for audio watermarking in [6], This method improves the robustness and perceptual quality of the system. At the embedding stage, the audio signal is divided into two subsignals and watermarks are added into the two subsignals simultaneously. The embedded subsignals are combined to form the watermarked signal. The watermarked signal is split up into two watermarked subsignals at the decoding stage.The similarity of the cepstra to the watermarked subsignals is used to extract the embedded watermarks. The performance of watermark extraction can be enhanced by using large peaks of pseudonoise sequences auto-correlation function.

In [7] Guang Hua.et.al, proposes a Convex optimization based finite-impulse-response(FIR) filter design.FIR is used to obtain the optimal echo filter coefficients. Proposed maximum power spectral margin (MPSM) and the absolute threshold of hearing (ATH) of human auditory system (HAS) is used to shape the desired power spectrum of the echo filter. This will ensure the optimal imperceptibility. The auto-correlation function of the echo filter coefficients will controls the robustness in terms of watermark detection. To design the echo kernel, convex optimization is used to obtain a set of filter coefficients which will replace the PN or MPN sequence.

A different method was introduced by Andrea Abrardo and Mauro Barni in [8], proposes a watermarking schema based on antipodal binary random binning, which allows the possibility of relying on simple and effective binary code constructions. And also make it easy to cope with amplitude scaling. The system outperforms previous watermarking system constructions by exhibiting very good performance also in the presence of gain attack. The encoding process is formed by an outer convolutional encoder CCo, the input to convolutional encoder is watermark w. A random interleaver permuting the outer codewords bits wc and input to the inner convolutional encoder is the permuted outer codewords wc. The outer encoder add some redundancy bits to the watermark message and the inner code embeds the watermark in the host signal by a dirty trellis mechanism. And atlast the sequences are mixed together to form the watermarking signal.

In [9] Yong Xiang.et.al proposes a patch work based audio watermarking method robust to de-synchronization attack. The watermarks are embedded into the host audio signal in the discrete cosine transform (DCT) domain and synchronization bits are implanted in the logarithmic DCT domain. At the decoding stage, find the scaling factor imposed by an attack by analyzing the received audio signal in LDCT domain. Then remove the scaling effect by modifying the received signal, together with the embedded synchronization bits. And extract the watermark from the modified signal.

Michael Arnold.et.al proposes a phase based audio watermarking system robust to acoustic path propagation in [10]. The system embeds information by modulating the phase in the WOLA domain. To improve the overall performances, three complementary modules are introduced: (i) a perceptual model to evaluate the noisiness of individual spectral bins, (ii) a resynchronization module which relies on tracking the resampling ratio over time and (iii) a statistical detector that aggregates the information from several values of the detection correlation array. The watermark modulation/demodulation framework can then be complemented by any regular channel coding technique. This channel coding is overlaid on top of the system. This can include either expanding the alphabet of symbols, introducing an error correcting code or establishing several communication channels associated to different frequency bands.

In [11] Mehdi Fallahpour and David Megas, proposes a audio watermarking system where embedding is done on the basis of Fibonacci numbers. Watermark extraction is done in a bit-exact manner by changing some of the magnitudes of the FFT spectrum. The main idea of the watermarking system is to divide the FFT spectrum into short frames and the change the magnitude of the selected FFT samples using Fibonacci numbers. First select a frequency part from FFT spectrum. Then divide the frequency band into frames. And a single secret bit is being embedded into each frame in the selected frequency band. On the basis of Fibonacci number and secret bit the magnitude of FFT coefficients in the selected frame is modified. If the bit to be embedded is 0, the coefficient value is changed to Fibonacci number with even index which is closer to the value. If the bit to be embedded is 1, the coefficient value is changed to Fibonacci number with odd index which is closer to the value.

3.COMPARISON BETWEEN THE WATERMARKING TECHNIQUES

In the above section different watermarking techniques are being analyzed. Each method has got the same aim of embedding a piece of information into the audio signal but there is difference in their watermark embedding and extraction procedure and in the piece of data to be embedded. The table comparing the advantages and disadvantages of different audio watermarking techniques are given below.

Table 1: COMPARISON OF WATERMARKING TECHNIQUES

Methods	Technique Used	Advantages	Disadvantages
Data Embedding in Audio Using Time-Scale Modification[1]	Embedding is done by changing the length of the intervals between the salient points of audio signal.	It is robust to mp3 compression,low pass filtering,TSM and other basic signal processing operations.	Detection depends on the selection of threshold. Not robust against pitch shifting.
Localized Audio Watermarking Technique Robust Against Time-Scale Modification[2]	Embedding is done at high-energy local regions that represent music edges.	High robustness against common audio signal processing and synchronization attack.	Only works well on audio with an obvious rhythm. Not robust against pitch shifting
Robust Multiplicative Patchwork Method For Audio Watermarking[3]	For the embedding process select two subset of the host signal feature using secret key. Watermark is embedded at one subset by multiplying and other is leaved unchanged.	Robust for common attacks such as noise , filtering , mp3 compression. Quality is assured	Not robust against geometric distortions such as TSM,pitch shifting.
Geometric Invariant Audio Watermarking Based on an LCM Feature[4]	Embedding is based on log coordinate mapping(LCM) feature. Embedding is done at the Fourier coefficients which are mapped to the feature via LCM.	Robustness against simultaneous geometric distortions.	Deals only with the transparency and robustness.
Audio Watermarking Using Spatial Masking and	A watermark is represented as a slightly rotated version of the original sound scene. Watermarks are embedded near the signal by adding a small copy of the host	Robust against different attacks.	Watermarks are detectable by a malicious attacker

Ambisonics[5]	signal.		
A Dual-Channel Time-Spread Echo Method for Audio Watermarking[6]	Host audio signal is divided into sub signals. Watermarks are implanted into the two sub signals simultaneously. The subsignals are combined to form the watermarked signal.	Better perceptual quality. Robust against common attacks.	Enhance perceptual quality by slightly reducing robustness.
Time-Spread Echo-Base Audio Watermarking With Optimized Imperceptibility and Robustness[7]	Convex optimization based finite-impulse-response (FIR) filter design is utilized to obtain the optimal echo filter coefficients.	Imperceptibility and robustness is achieved.	Not a quantitative approach.
A New Watermarking Scheme Based on Antipodal Binary Dirty Paper Coding[8]	Watermarking is based on based on serially concatenated convolutional codes with dirty trellis embedding.	System is immune to gain attacks. Good performance when used for the watermarking of real multimedia data.	Watermarking channel has less capacity.
Patchwork-Based Audio Watermarking Method Robust to De-synchronization Attacks[9]	Watermarks are embedded into the audio signal by modifying its DCT coefficients. A set of synchronization bits are implanted into the watermarked signal in the logarithmic DCT (LDCT) domain.	More robust to desynchronization attacks. Higher embedding capacity.	Decrease in detection rate by increase in embedding rate.
A Phase-Based Audio Watermarking System Robust to Acoustic Path Propagation[10]	Watermark embedding is done by modulating the phase in the WOLA domain	Robustness against synchronization attacks. High fidelity.	Embedded watermark suffers a slight degradation during signal reconstruction.
Audio Watermarking Based on Fibonacci Numbers[11]	Watermark embedding done by changing the magnitude of the FFT domain using the watermark data and Fibonacci number.	System has a high capacity without signicant perceptual distortion. Provides robustness against common signal processing attacks. Provides security for the watermarked data.	Collision between low frequency value and low valued Fibonacci number .Variation in the FFT value during data retrieval.

4. CONCLUSIONS

Digital watermarking schemes are used for protecting data from unauthorized access or illegal copyrighting. There are image, video and audio watermarking techniques. Studies on audio watermarking are far less than that of image watermarking or video watermarking. But nowadays audio watermarking studies have also increased considerably. This paper surveyed those techniques and presented some of the important techniques used for digital audio watermarking.

Considering the embedding domain, audio watermarking techniques can be classified into time-domain and frequency-domain methods.FFT transform is used by different systems, such as in LCM based system which proposes a multi-bit spread spectrum audio watermarking scheme based on a geometric invariant log coordinate mapping (LCM) feature. The watermark is embedded in the LCM feature, but it is actually embedded in the Fourier coefficients which are mapped to the feature via LCM. Consequently, the embedding is actually performed in the FT domain. Time-spread echo-based audio watermarking scheme with optimized imperceptibility and robustness. Specifically, convex optimization-based finite-impulse-response filter design is used to obtain the optimal echo filter coefficients. But most of these methods mainly focus on the robustness and transparency of the system and does not deal with the security of the system the Fibonacci based watermark embedding considers the security aspects of the system.

REFERENCES

[1] Mohamed F. Mansour and Ahmed H. Tewk, "Data Embedding in Audio Using Time-Scale Modifcation,"*IEEE Transactions on speech and audio processing,vol. 13, no. 3, May 2005*.

[2] Wei Li, Xiangyang Xue, and Peizhong Lu, "Localized Audio Watermarking Technique Robust Against Time-Scale Modification" *IEEE Transactions on multimedia, vol. 8, no. 1, February 2006*.

[3] Nima Khademi Kalantari, Mohammad Ali Akhaee, Seyed Mohammad Ahadi, Hamidreza Amindavar, "Robust Multiplicative Patchwork Method For AudioWatermarking," *IEEE Transactions On Speech And Audio Processing-2009.*

[4] Xiangui Kang, Rui Yan and Jiwu Huang, "Geometric Invariant Audio Watermarking Based on an LCM Feature,"*IEEE Transactions On Multimedia,vol.13, no. 2, April 2011.*

[5] Ryouichi Nishimura, "Audio Watermarking Using Spatial Masking and Ambisonics,"IEEE Transactions On Audio, Speech, And Language Processing,vol.20, no. 9, November 2012.

[6] Yong Xiang, Iynkaran Natgunanathan, Dezhong Peng,WanleiZhou andShui Yu,"A Dual-Channel Time-Spread Echo Method for Audio Watermarking, "*IEEE transactions on information forensics and security,vol. 7, no. 2, April 2012.*

[7] Guang Hua, Jonathan Goh, and Vrizlynn. L. L. Thing,"Time-Spread Echo-Based Audio Watermarking With Optimized Imperceptibility and Robustness,"*IEEE Transactions on audio, speech, and language processing,vol. 23, no. 2,February 2015.*

[8] Andrea Abrardo, Mauro Barni, "A New Watermarking Scheme Based on Antipodal Binary Dirty Paper Coding,"*IEEE Transactions on Information Forensics and Security -2013.*

[9] Yong Xiang, Iynkaran Natgunanathan, Song Guo, Wanlei Zhou and Saeid Nahavandi, "Patchwork-Based Audio Watermarking Method Robust to De-synchronization Attacks,"*IEEE Transactions on audio, speech, and language processing,vol. 22, no. 9, September 2014.*

[10] Michael Arnold, Xiao-Ming Chen, Peter Baum, Ulrich Gries, and Gwenal Dorr,"A Phase-Based Audio Watermarking System Robust to Acoustic Path Propagation,"*IEEE Transactions on information forensics and security,vol. 9, no. 3,March 2014.*

[11] Mehdi Fallahpour and David Megas, "Audio Watermarking Based on Fibonacci Numbers, " IEEE Transactions on audio, speech, and language processing, vol.23, no. 8, August 2015.