A Survey on Audio Watermarking Techniques

Nidhi Rathee¹ and Sunil Kumar²

¹Research Scholar, PDM College of Engineering, Bahadurgarh
²Asst. Professor, PDM College of Engineering, Bahadurgarh
¹ nidhirathee108@gmail.com

Abstract
Watermarking is closely related to steganography, but in watermarking the hidden information is usually related to cover object. Hence its mainly used for copyright protection and owner authentication. Image watermarking is used to hide the special information into the image and to later detect and extract that special information for the author ownership. Video watermarking is used to add watermark in the video stream to control video applications. It is extension of image watermarking. Audio watermarking is one of most popular due to internet music. Text watermarking adds watermark to the PDF, document and other text file to prevent the changes made text. The watermark is inserted in the font shape and the space between characters and line spaces. Graphic watermarking is embed the watermark to 2D computer generated graphics to indicate the copyright. This paper study the concept of audio watermarking and its related techniques.

Keywords: Watermarking, Audio Watermarking, Spread Spectrum Watermarking, Amplitude Modification.

I. Introduction
Digital watermarking is the signal embedding, secret information into the digital media such as image, audio, and video later the embedded information is detected and extracted out to identity of the digital media. Digital watermarking technology has many applications in protection, certification, distribution, anti-counterfeit of the digital media. The watermarking is embedded by two types visible watermark and invisible watermarking. In visible watermarking, the watermark that visible in the digital data likes stamping a watermark on paper. In invisible watermarking can insert information into an image which cannot be seen, but can be interrogated with right software. [1]

Digital watermarking hides the copyright information into the digital data through certain algorithm. The secret information to be embedded can be some text, author’s serial number, company logo, images with some special importance. This secret information is embedded to the digital data (images, audio, and video) to ensure the security, data authentication, identification of owner and copyright protection. The watermark can be hidden in the digital data either visibly or invisibly. For a strong watermark embedding, a good watermarking technique is needed to be applied. Watermark can be embedded either in spatial or frequency domain. Both the domains are different and have their own pros and cons and are used in different scenario [2].

II. Audio Watermarking
Digital watermarking has to embed pieces of information into a digital media for protecting it against copyright infringements and other unauthorized applications. Digital audio watermarking has to do with protecting digital audio

IJEAM
www.ijeam.com
file against illegal copying. A lot of works has been done on digital watermarking of various media such as image and video, but this particular review will focus on digital watermarking of audio file. Digital Audio files are particularly the most abused for copyright infringements because they can be downloaded and copied with ease. Audio watermarks are special signals embedded into digital audio. These signals are extracted by detection mechanisms and decoded. Audio watermarking schemes rely on the imperfection of the human auditory system. However, human ear is much more sensitive than other sensory motors. Thus, good audio watermarking schemes are difficult to design [3]. A lot of recent works have proposed solutions to this problem of copyright infringement. Some involves the use of various types of watermarks [3].

Audio watermarking is a technique that embeds information with specific meaning into the host media without perceptible interference to the original quality. As a complement to conventional encryption techniques, watermarking provides powerful tools for protecting the copyright of audio works and has become an active research area in recent years. Besides being imperceptible to human ears, the watermark must also be able to withstand audio signal processing and time-domain synchronization attacks [4].

III. Properties Of Watermark

There are about five properties that need to be satisfied for effective application of watermarking technology. These are robustness, imperceptibility, bit rate, security and computational complexity. Some of the watermark properties are discussed below [3].

a. Robustness To Signal Processing

Digital signal may undergo common signal processing operations such as Linear filtering, sample re-quantization, D/A (digital-analogue) and A/D (analogue-digital) conversion and lossy compression [5].

b. Perceptual Quality

This refers to imperceptibility of embedded watermark within the host signal. The signal-to-noise ratio of the watermark versus the host signal is the measure of the perceptual quality.

c. Bit Rate

This is the amount of watermark data that may be reliably embedded within the host signal per unit time or space. A higher bit rate may be desirable in some application to embed copyright information. Reliability is measured using BER (bit error rate).

d. Watermark Security

Watermark security refers to the inability by unauthorized users to have access to the raw watermarking channel [5].

e. Computational Complexity

This refers to the processing required to embed data into the host signal and or to extract data from the signal.

IV. Audio Watermarking Techniques

The most popular techniques for digital audio watermarking are reviewed. Specifically, the different techniques correspond to the methods for merging (or inserting) the cover data and the watermark pattern into a single signal [3].

a) Spread Spectrum Watermarking

Spread-spectrum watermarking scheme is an example of the correlation method which embeds pseudorandom sequence and detects watermark by calculating correlation between pseudo-random noise sequence and watermarked audio signal. Spread spectrum techniques for watermarking borrow most of the theory from the communications community [6]. The main idea is to embed a narrow-band signal (the watermark) into a wide-band channel (the audio file). The characteristics of both audio signal A and watermark W seem to suit the model perfectly. In addition, spread spectrum techniques offer the possibility of protecting the watermark privacy by using a secret key to control the pseudorandom sequence generator. Spread spectrum techniques allow the frequency bands to be matched before embedding the message. This is why spread spectrum techniques are valuable not only for robust communication but for watermarking as well.

There are two basic approaches to spread spectrum techniques: direct sequence and frequency hopping. In both of these approaches the idea is to spread the watermark data across a large frequency band, namely the entire audible spectrum.
b) Amplitude Modification

This method, also known as least significant bit (LSB) substitution, is both common and easy to apply in both steganography and watermarking [7] as it takes advantage of the quantization error that usually derives from the task of digitizing the audio signal. As the name states, the information is encoded into the least significant bits of the audio data. There are two basic ways of doing this: the lower order bits of the digital audio signal can be fully substituted with a pseudorandom (PN) sequence that contains the watermark message \( m \), or the PN-sequence can be embedded into the lower order bit stream using the output of a function that generates the sequence based on both \( n \)th bit of watermark message and \( n \)th sample of audio file. The major disadvantage of this method is its poor immunity to manipulation. Encoded information can be destroyed by channel noise, re-sampling, etc., unless it is encoded using redundancy techniques. In order to be robust, these techniques reduce the data rate, often by one to two orders of magnitude. Furthermore, in order to make the watermark more robust against localized filtering, a pseudorandom number generator can be used to spread the message over the cover in a random manner.

c) Replica Method

Original signal can be used as an audio watermark. Echo hiding is a good example. Replica modulation also embeds part of the original signal in frequency domain as a watermark. Thus, replica modulation embeds replica, i.e., a properly modulated original signal, as a watermark. Detector can also generate the replica from the watermarked audio and calculate the correlation. The most significant advantage of this method is its high immunity to synchronization attack.

- Echo hiding

Echo hiding embeds data into an original audio signal by introducing an echo in the time domain such that for simplicity, a single echo is added above Figure 1. Binary messages are embedded by echoing the original signal with one of two delays, either a \( d_0 \) sample delay or a \( d_1 \)sample delay. Extraction of the embedded message involves the detection of delay \( d \). Autocepstrum or cepstrum detects the delay \( d \). Cepstrum analysis duplicates the cepstrum impulses every \( d \) samples. Echo hiding is usually imperceptible and sometimes makes the sound rich. Synchronization methods frequently adopt this method for coarse synchronization. Disadvantage of echo hiding is its high complexity due to cepstrum or autocepstrum computation during detection. On the other hand, anybody can detect echo without any prior knowledge. In other words, it provides the clue for the malicious attack. This is another disadvantage of echo hiding.

d) Dither Watermarking

Dither is a noise signal that is added to the input audio signal to provide better sampling of that input when digitizing the signal [4]. As a result, distortion is practically eliminated, at the cost of an increased noise floor. To implement dithering, a noise signal is added to the input audio signal with a known probability distribution, such as Gaussian or triangular. In the particular case of dithering for watermark embedding, the watermark is used to modulate the dither signal. The host signal (or original audio file) is quantized using an associated dither quantizer. This technique is known as quantization index modulation (QIM) [9].

A graphical view of this technique is shown in Figure 4. Here, the points marked with X’s and O’s belong to two different quantizer, each with an associated index; that is, each one embedding a different value. The distance \( d_{\text{min}} \) can be used as an informal measure of robustness, while the size of the quantization cells (one is shown in the figure) measures the distortion on the audio file. If the watermark message \( m=1 \), then the audio signal is quantized to the nearest X. If \( m=2 \) then it is quantized to the nearest O.

![Figure 1: Kernels for Echo hiding][8]
The two quantizers must not intersect, as can be seen in the figure 2. Furthermore, they have a discontinuous nature. If one moves from the interior of the cell to its exterior, then the corresponding value of the quantization function jumps from an X in the cell’s interior to one X on its exterior.

e) Self-Marking Method
Self-marking method embeds watermark by leaving self-evident marks into the signal. This method embeds special signal into the audio, or change signal shapes in time domain or frequency domain. Time-scale modification method [10] and many schemes based on the salient features [11] belong to this category.

➤ Time-Scale Modification
Time-scale modification is a challenging attack and can be used for watermarking. Time-scale modification refers to the process of either compressing or expanding the time-scale of audio. Basic idea of the time-scale modification watermarking is to change the time-scale between two extrema (successive maximum and minimum pair) of the audio signal. Figure 3 explains the concept of time-scale modification. The intervals between two extrema are partitioned to N segments of equal amplitude. The slope of the signal can be changed in certain amplitude interval(s) according to the bits which is to be embedded, which changes the timescale. For example, the steep slope and gentle slope stand bits 101 and 111 or vice versa, respectively. Advanced time-scale modification watermarking scheme can survive time-scale modification attack.

f) Least Significant Bit (LSB) Coding [12]
One of the earliest techniques studied in the information hiding of digital audio (as well as other media types) is LSB coding. In this technique LSB of binary sequence of each sample of digitized audio file is replaced with binary equivalent of secret message [13].

i. Parity Coding [12]
Instead of breaking a signal down into individual samples, the parity coding method breaks a signal down into separate regions of samples and encodes each bit from the secret message in a sample region’s parity bit. If the parity bit of a selected region does not match the secret bit to be encoded, the process flips the LSB of one of the samples in the region. Advantage: The sender has more of a choice in encoding the secret bit, and the signal can be changed in a more unobtrusive manner. Disadvantage: This method like LSB coding is not robust in nature [14].

ii. Phase Coding [14]
Phase coding relies on the fact that the phase components of sound are not as perceptible to the human ear as noise is. It “works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is then adjusted in order to preserve the relative phase between segments”. Disadvantage: It is a complex method and has low data transmission rate [12].
V. Conclusion

Audio watermarking is a persistent data communication channel within an audio stream. It should survive through various format changes and manipulations (either legitimate or not) of the audio material, as long as the content retains some commercial potential. Additionally, it should do so without introducing any perceivable audio artifacts. Its naturally limited data rate suggests that this communication channel should be used in conjunction with carefully designed Watermark Data. This data is preferably small, as shorter watermarks will be repeated more frequently throughout the content, enhancing the channel’s redundancy. The main purpose of this study was to overview the concept of watermarking and the audio watermarking. It also describes the various techniques of audio watermarking.

References