LOG CO-ORDINATE MAPPING BASED AUDIO WATERMARING

Sarath Gopal
PG Scholar,.
Maharaja Prithvi Engineering college
Sarathgopal068@gmail.com

Abstract— **Watermarking is a technique, which is used in protecting digital information like images, videos and audio as it provides copyrights and ownership. Audio watermarking is more challenging than image watermarking due to the dynamic supremacy of hearing capacity over the visual field. To develop a geometric invariant audio watermarking scheme without degrading acoustical quality is more challenging. This work proposes a spread spectrum audio watermarking scheme based on a geometric invariant feature. The watermark embedding is actually performed in the DFT domain. The various audio signal distortions like pitch-shifting, random cropping, Time scale modifications, etc. are proposed to analyses in this work. The proposed method uses average Fourier magnitude over log-coordinate, which can resist most of the audio signal distortions. Through experiments we try to demonstrate that average Fourier magnitude over log coordinate is an appropriate embedding region for robust audio watermarking. A comparison between the log coordinate feature and a PN tracking sequence is proposed to retrieve the embedded watermark. The proposed watermarking scheme has been compared with a DCT based approach and analyzed. Matlab 7.8(R2009a) is used to implement the algorithms discussed in this work. The proposed algorithm may work as a tool for securing intellectual properties of the musicians and audio distribution companies because of its high performance quality and imperceptibility**.

I. INTRODUCTION

The digital data can be processed, accessed, and it can be transmitted very quickly using networks. There are numerous technical, legal, and organizational problems which arise when there is wide scale use of digital documents. Digital information can be copied any number of times from one medium to another; they can be transmitted through networks, etc., all without compromising the quality of the data. There is no way to distinguish between an original electronic documents and its copy. It is easy to change any part of an un protected electronic document. One possibility here is to replace original signatures with cryptographic methods. Digital signature is data items formed by the signatory and created from the document that is to be signed. It relates the documents to the signatory in a secure and reliable way. Digital watermarking has been proposed as one way to accomplish this.

Also advanced Internet services enabled the users to create copy and distribute multimedia products such as audio, video, and still images with much ease and less effort, minimum or no cost, and in less time. Though it encouraged trading on the Internet, but on the other hand it has created the problem of illegal copying or copyright infringement. Thus, protection of digit alights assumed a primary importance in the digital age. Digital watermarking can be defined as the process of embedding a certain piece of information (technically known as watermark) into multimedia content including text documents, images, audio or video streams, such that the watermark can be detected or extracted later to make an assertion about the data. The most important properties of digital watermarking techniques are transparency, robustness, security, capacity, inevitability (reversibility) and complexity and possibility of verification. Based

II. SPREAD-SPECTRUM WATERMARKING OF AUDIO SIGNALS

Watermarking has become a technology of choice for a broad range of multimedia copyright protection applications. Watermarks have also been used to embed format-independent metadata in audio/video signals in a way that is robust to common editing.

With the growth of the Internet, unauthorized copying and distribution of digital media has never been easier. As a result, the music industry claims a multibillion dollar annual revenue loss due to piracy, which is likely to increase due to peer-to-peer file sharing Web communities. One source of hope for copyrighted content distribution on the Internet lies in technological advances that would provide ways of enforcing copyright in client-server scenarios. Traditional data protection methods such as scrambling or encryption cannot be used since the content must be played back in the original form, at which point, it can always be rerecorded and then freely distributed. A promising solution to this problem is marking the media signal with a secret, robust, and imperceptible watermark (WM). The media player at the client side can detect this mark and consequently enforce a corresponding e-commerce policy. Recent introduction of a content screening system that uses asymmetric direct sequence spread-spectrum (SS) WMs has significantly increased the value of WMs because a single compromised detector (client player) in that system does not affect the security of the content. In order to compromise the security of such a system without any traces, an adversary needs to break in the excess of 100 000 players for a two-hour high-definition video.

A. *Watermarking Technologies*

Audio watermarking schemes rely on the imperfections of the human auditory system (HAS). Numerous data hiding techniques explore the fact that the HAS is insensitive to small amplitude changes, either in the time or frequency domains, as well as insertion of low-amplitude time-domain echoes. Information modulation is usually carried out using: SS [9] or quantization index modulation (QIM). The main advantage of both SS and QIM is that WM detection does not require the original recording and that it is difficult to extract the hidden data using optimal statistical analysis under certain conditions. However, it is important to review the disadvantages that both technologies exhibit. First, the marked signal and the WM have to be perfectly synchronized at WM detection. Next, to achieve a sufficiently small error probability, WM length may need to be quite large, increasing detection complexity and delay. Finally, the most significant deficiency of both schemes is that by breaking a single player (debugging, reverse engineering, or the sensitivity attack), one can extract the secret information (the SS sequence or the hidden quantizes in QIM) and recreate the original (in the case of SS) or create a new copy that induces the QIM detector to identify the attacked content as unmarked. While an effective mechanism for enabling asymmetric SS watermarking has been developed, an equivalent system for QIM does not exist to date.

B. Techniques for SS Watermarking of Audio Signal

Here discussing the Direct-sequence SS WMs and develop a set of technologies to improve the effectiveness of their embedding and detecting in audio. WM robustness is enabled using *i*) block repetition coding for prevention against de-synchronization attacks and *ii*) psycho-acoustic frequency masking (PAFM).We show that PAFM creates an imbalance in the number of positive and negative WM chips in the part of the SS sequence that is used for WM correlation detection and that corresponds to the audible part of the frequency spectrum. To compensate for this anomaly, we propose a *iii*) modified covariance test. In addition, to improve reliability of WM detection, we propose two techniques for reducing the variance of the correlation test *iv*) cepstrum filtering and *v*) *chess* WMs. Since we embed SS WMs in the frequency domain, the energy of a WM is distributed throughout the entire synthesis block, making SS WMs audible in blocks that contain quiet periods. We solve this problem using *vi*) a procedure that identifies blocks where SS WM may be audible to decide whether to use a particular block in the WM embedding/detection process. Finally, we propose *vii*) a technique that enables reliable covert communication over a public audio channel.

In order to investigate the security of SS WMs, we explore the robustness of such a technology with respect to watermark estimation attacks. To launch that attack, an adversary is assumed to know all the details of the WM codec, except the hidden secret. Presenting a modification to the traditional SS WM detector that *viii*) undoes the attack and, hence, forces the adversary to add an amount of noise proportional in amplitude to the recorded signal in order to successfully remove an SS WM. We have incorporate these techniques *i)-viii*) into a system capable of reliably detecting a WM in an audio clip that has been modified using a composition of attacks that degrade the original audio characteristics beyond the limit of acceptable quality. Such attacks include fluctuating scaling in the time and frequency domain, compression, addition and multiplication of noise, resampling, requantization, normalization, filtering, and random cutting and pasting of signal samples.

C. Watermarking Using PN Sequences

To overcome the limitations in watermarking due to methods like *LSB* substitution and to make the system more robust against attacks, the watermark can be spread across the cover object by using more number of bits than the minimum required. This scheme of hiding the data uses the concepts of code division multiple access (*CDMA)*. This technique ensures the survival of watermark under various attacks due to redundancy. Each of the data bits is represented by using a large number of bits out of which a significant portion may be lost without totally losing the watermark information.

Let *a, b, c, d* be the PN sequences generated for each of the elements in the watermark vector 1, 0, 0 and 1.

$$a = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} b = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} d = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Now, considering gain (k) = 2 we apply the formula I_w (x, y) = I (x, y) + k × W (x, y). From the above example we have to add PN sequences b and c multiplied with a gain factor of 2 to the cover image

$$I_w\ (x, y) = I(x,y)+(2\times b)$$
$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 3 & 4 & 0 \\ 5 & 3 & 5 & 2 \\ 1 & 2 & 3 & 3 \end{bmatrix}$$

$$I_w\ (x, y) = I(x,y)+(2\times c)$$
$$\begin{bmatrix} 1 & 4 & 0 & 3 \\ 4 & 5 & 6 & 2 \\ 5 & 5 & 7 & 2 \\ 3 & 2 & 3 & 3 \end{bmatrix}$$

This gives the resultant watermarked image after embedding the PN sequences for each black pixel in the watermark vector. For recovery of the pixels the same PN sequences are generated at the receiver and correlated with the watermarked image. The threshold is set as the mean of the correlation value for all the pixels.

corr2 (I_w (x, y), a) = -0.1041

corr2 (I_w (x, y), b) = 0.5897

corr2 (I_w (x, y), c) = 0.5897

corr2 (I_w (x, y), d) = -0.1041

Average correlation or the threshold = 0.2428

So the pixels *b* and *c* are marked as black pixels and the pixels *a* and *d* are marked as the white pixels for the recover.

Although watermarking using pseudonoise sequences is effective it has some disadvantages associated with it. The sequence period is typically greater than the image size and, therefore, the correlation at recovery is incomplete. This leads to a tradeoff between the gain and the robustness of watermarked image. As the gain is increased the recovery of the watermark improves, but at the cost of distorting the watermarked image.

III. FORMULATION OF THE PROBLEM

Globalization and internet are the main reasons for the growth of research and sharing of information. However, they have become the greatest tool for malicious user to attack and pirate the digital media. The ease of content modification and a perfect reproduction in digital domain have promoted the protection of intellectual ownership and the prevention of the unauthorized tampering of multimedia data to become an important technological and research issue. Digital watermarking has been proposed as a new method to enforce the intellectual property rights, tracing of illegal copies of digital media and protect digital media from tampering. Several audio watermarking schemes have been presented over the years. The development of a geometric invariant audio watermarking scheme without degrading acoustical quality is challenging work. The thesis proposes a geometric invariant feature, the average Fourier magnitude (AFM) over the log coordinate, which is invariant to geometric distortions and is denoted by log coordinate mapping (LCM) feature. The LCM feature is very robust to audio geometric distortions, such as time-scale modifications (TSM), tempo invariant pitch shifting, random cropping etc.

A. *Log Coordinate Transform on Frequency Index*

Geometric distortions can be described in the frequency domain as follows:

$$f' = \beta . f \qquad (3.1)$$

Where β stands for the frequency scaling factor, f and f' are frequency point of the original audio and the corresponding frequency point of a distorted audio, respectively. Frequency scaling by β can be converted into shifting by logβ in the log coordinates. Taking the logarithm of (3.1), it can be rewritten as follows:

$$\log_b f' = \log_b \beta + \log_b f \qquad (3.2)$$

where b is the logarithm base. Thus geometric distortions can be easily manipulated in the log coordinate of frequency index. As amplitude scaling is unavoidable during attacks, we select a correlation- based watermarking which can resist amplitude scaling. The host feature is the average Fourier magnitude (AFM) over the Log coordinate frequency index. Given a signal, s(n) = [s_1 .... s_N) we perform a DFT and get the Fourier magnitude *S(f)*. After selecting a portion of the normalized frequency indexes, we perform a log coordinate transform on frequency index as shown in

$$l = floor\ (\log_b \frac{f}{R})) + L/2$$

where ,R $= \sqrt{2} f_m$

$$b = 2^{1/L} \qquad (3.3)$$

where L is the number of log intervals and is specified by users. The selected frequency index f is mapped to discrete log coordinates *l* (0$\leq$ *l* < L), so that the selected frequency coefficient *S(f)* is mapped to a log coordinate mapping feature a(l) which is defined as follows:

$$a(l) = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} S(f)\, df$$

where, f_1 *= min{f(floor(*$\log_b \frac{f}{R}$*)) + L/2 = l}*

$$f2 = max\{f(floor(\log_b \frac{f}{R}) + L/2 = l\} \qquad (3.4)$$

a) *Existing Watermark Embedding*

Watermark (W_i) is a direct sequence spread-spectrum (DSSS) encoded with N_p bit bipolar PN sequence. The hidden data W consists of spread spectrum information watermark (W_i) and a tracking sequence T generated by a key. Apply DFT on the original audio signal s(n) and obtain the Fourier magnitude *S(f)* and the phase. Performing a discrete log-coordinate transform to a portion of the normalized frequency indexes fs, we obtain the discrete log-coordinate *l.*

Here several DFT magnitude coefficients are mapped to one LCM feature. One watermark bit w(l) is embedded into several DFT magnitude coefficients which are mapped to one LCM feature. Modify the Fourier magnitude *S(f)* to embed the hidden bit *w(l)* according to

$$\tilde{S(f)} = S(f)\ (1+\alpha_e\ w(l)) \qquad (3.5)$$

where *S(f)*, $\tilde{S}$ (*f*) are the Fourier magnitude coefficients before and after embedding, and α_e is the embedding strength. Finally, by performing an IDFT to the modified DFT coefficients, the watermarked audio signal $\tilde{S}$˜(*n*) is obtained.

It is observed that not the whole range of average Fourier magnitude (AFM) over the log coordinate is suitable for embedding watermarks. The amplitudes near the highest frequency are small, and this part is sensitive to low-pass filtering. So we may set the watermark embedding region to the low and middle frequency components.

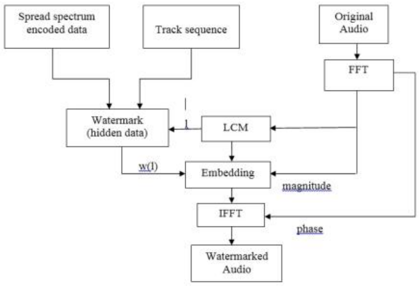

Figure 4.1: Existing watermark Embedding block

b) *Proposed Watermark Extraction*

Watermark is extracted from the average Fourier magnitude over the Log coordinate (LCM feature). The LCM feature may get translated after geometric distortion. The tracking sequence and the original PN-sequence may be known to the detector prior to extraction. First, apply the DFT on the watermarked audio signal $\tilde{s}(n)$ and obtain the magnitude coefficient $\tilde{S}(f)$.

Then perform a discrete log-coordinate transform to the frequency index f and average the entire magnitude $\tilde{S}(f)$ with the same discrete log-coordinate *l*. Then we obtain the LCM feature $\hat{a}(l)$.

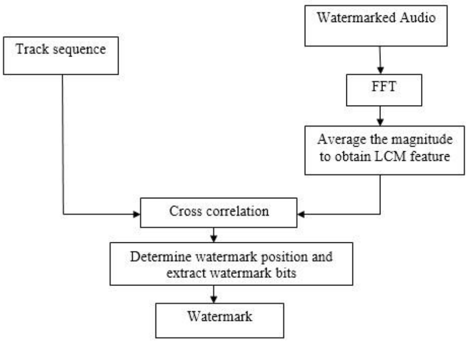
Track sequence
Watermarked Audio
FFT
Average the magnitude to obtain LCM feature
Cross correlation
Determine watermark position and extract watermark bits
Watermark

Figure 4.2: Proposed Watermark Extraction block

Computing the correlation between the tracking sequence T and the LCM features$\hat{a}$ (*l*), we can determine the watermark embedding position. One way to track the location of watermark is to perform exhaustive shifting along the log axis and to calculate the cross correlation between T and $\hat{a}$ (*l*) and then find the maximum correlation to locate the match position. However, the computation load for this exhaustive searching is heavy. A fast way to search the maximum correlation value is using correlation theorem. Append T with zeros to the same size of $\hat{a}$ (*l*) to obtain *g(l)*. Then the correlation between them is

$$c(k) = IDFT(\hat{A}(u).G^{*}(u)) \qquad (3.6)$$

where $G(u) = DFT[g(l)]$ and $\hat{A}(u) = DFT[a(l)]$.

Pick N_p elements from $\hat{a}(l)$ to form a sequence *W*(*i*), which corresponds to the embedded spread-spectrum sequence W_i from **WT**, and correlate with the original PN sequence . If the correlation value is larger than 0, the extracted bit is taken to be in favor of 1; otherwise, it is determined to be 0. The hidden watermark can thus be recovered.

B. *Effectiveness of log coordinate transform on attacks*

If we embed the watermark signal in to the Fourier magnitude of audio one to one and same scale, the original watermark will suffer the same distortion as the audio signal. When the synchronisation attacks are applied to the watermarked audio, the Fourier magnitude of the audio will fluctuate and the frequency index will be scaled. The case of the watermark is the same. Obviously the original watermark and the survival watermark are not correlative.

The case is different when applying log coordinate transform in watermark embedding and extraction. We apply log coordinate mapping to the watermark index and generate embedding positions in DFT magnitude. Now one watermark bit will be embedded in multiple DFT magnitudes. After attacks there will be some fluctuations in magnitudes, and we perform the average of the Fourier magnitudes. Before watermark extraction we utilize log coordinate transform to retrieve the watermark index. Now the watermark bits can be extracted from the survival watermark signal correctly.

a) *Under Random Cropping*

Random cropping means that a portion of the audio is lost in the time domain, but in the frequency domain, it only introduces tiny fluctuations. To resist random cropping, the watermarking strategy must be global. As the length of audio clip varies after random cropping, the frequency index must be normalised after Fourier transform. Geometric distortion by random cropping can be described by the equation (3.1). A powerful tool to deal with the scaling factor is log coordinate transform. A logarithm could convert the scaling into shifting (3.2) in the logarithm axis

b) *Under Pitch Shifting*

Pitch shifting is a very common form of processing used to change the base frequency without changing the tempo. Pitch shifting may be implemented as follows: resample a audio signal for shifting the pitch, then remove and/or insert some samples of the resampled audio signal in the time domain in order to keep the tempo invariant. Removing and inserting some samples cause only a small fluctuation in the frequency domain. Theoretically there exists a statistically positive linear correlation between pitch shifted and original audio (3.1).

c) *Under Pitch-Invariant TSM*

Pitch-invariant TSM can be considered to be removing and/or inserting some samples of audio signals while preserving the pitch. It causes only a small fluctuation to the LCM feature in the frequency domain.

C. *Methods of Analysis*

a) *Quality of Watermarked Audio*

The objective quality is measured by SNR and objective difference grade (ODG). The ODG value is mapped to the following description : 0 (insensitive),1 (audible), 2 (slightly annoying), 3 (annoying), 4 (very annoying), and 5 (catastrophic). A subjective quality evaluation of the watermarking method was done by asking 10 persons to listen to the four audio clips. In the first phase of the test, participants were presented with the pairs of the original and the watermarked audio clips in random order and asked to determine which one was the original clip and which one was not. A discrimination rate (the rate of correct discrimination) near to 50% means that the original and watermarked audio clips cannot be discriminated. In the second phase of the test, the persons are presented with the original and watermarked audio objects, and then give scores for each audio. The mean opinion score (MOS) determines the amount of distortion. The five-point impairment scale is applied, 5.0 for imperceptible, 4.0 for perceptible but not annoying, 3.0 for slightly annoying, 2.0 for annoying, and 1.0 for very annoying.

IV. SIMULATION RESULTS

In the proposed method of audio watermarking not the whole range of Fourier magnitudes AFM is suitable for embedding watermarks. The amplitudes near the high frequency components are small and this part is sensitive to low-pass filtering. So we may select the embedding region corresponding to the low and middle frequency components. In implementation, the actual watermark embedding is directly performed in the DFT domain based on the discrete log coordinate of the frequency index.

We consider audio clips, which are in .WAV format, mono, 16 bits/sample,20 s, and 44.1-kHz sampling frequency. We choose the length of watermark I=64 bits, the length of tracking sequence N_T = 320, and a total L=960 bits of hidden data are embedded. We adopt the length of DFT and IDFT being equal to the length of the host audio clip, i.e., 2044100 = 882000. The length of the hidden data is L = $I_P + N_T$, which depends on the bits of the watermark L, the length of PN sequence N_P, and the length of the track sequence NT . The value of L is selected by the user and could not be too large, because it represents the number of discrete intervals that the frequency indices are mapped to, and we must ensure that each interval contains at least one frequency index.

The experiment is conducted to implement the proposed audio watermarking algorithm using MATLAB. Audio clips of fixed duration in the. WAV format is read to the MATLAB environment, that is we get the sampled values of the original audio clip at 44.1 kHz sampling rate. Now corresponding discrete Fourier transform (DFT) of the audio clip is computed by FFT algorithm, and obtain its magnitude. The hidden data of length L = 960 bits are obtained by combining the direct sequence spread spectrum (DSSS) encoded watermark bits with a track sequence. The preliminary simulation results are shown in the following figures.

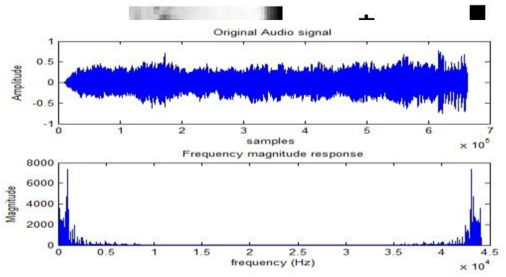

Figure 6.2: Frequency magnitude spectrum of the Audio signal

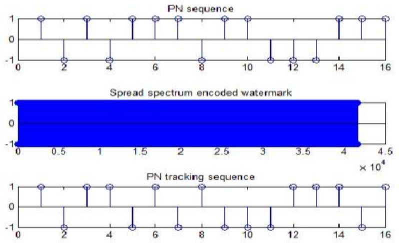

Figure 6.3: Bipolar PN sequence used for Spread spectrum encoding and Track sequence used for efficient watermark extraction

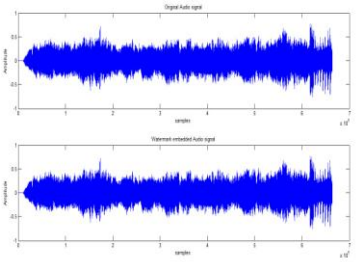

Figure 6.4: Data to be embedded in to audio signal

V. CONCLUSION

This work developed an audio watermark embedding strategy, which is actually performed in DFT domain without interpolation, to avoid completely the severe distortion due to the non uniform interpolation mapping while achieving the effect of embedding in LCM domain. The watermark is embedded in the LCM feature but is actually embedded in the Fourier are mapped to the feature via the LCM. The watermarked audio achieves 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 audio watermark efficiently with only one FFT and one IFFT.

REFERENCES

[1] X. Kang, R. Yang and J. Huang. *Geometric Invariant Audio Water-marking Based on an LCM Feature*, IEEE Trans. Multimedia, VOL. 13,NO.2, pp 181–190, APRIL 2011.

[2] S. Xiang and J. Huang *Histogram-based audio watermarking againsttime-scale modifications an cropping attacks*, IEEE Trans. Multimedia, vol. 9, no. 7, pp. 1357-1372, Nov. 2007.

[3] S.Wu, J. Huang, D. Huang, and Y. Shi *Efficiently self-synchronized audio watermarking for assured audio data transmission*, IEEE Trans. Broadcast., vol. 51, no. 1, pp. 6976, Mar. 2005.

[4] D. Kirovski and H. S. Malvar *Spread-spectrum watermarking of audio signals*, IEEE Trans. Signal Process., vol. 51, no. 4, pp. 1020-1033, Apr.2003.