IJRECE VOL. 4 ISSUE 1 JAN.-MAR. 2016 ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE)

Data Hiding Techniques for Audio Signals: A Review

Arun Kumar Arya, Priyanka
ECE Department, DCRUST Murthal, Haryana

Abstract - Digital watermarking has been proposed in recent years as a means of protecting multimedia contents from intellectual privacy. This is achieved by modifying the original content, by inserting a signature which can be extracted, when required. Information hiding technique is also a new kind of secret communication technology. Most of the today's information hiding systems uses multimedia objects such as image, audio & video. Audio watermarking is usually a difficult process than image watermarking due to dynamic supremacy of Human hearing system over the visual system. In this paper different types of available audio watermarking techniques has been presented and compared.

I. INTRODUCTION

Due to the new advancements in computer systems and telecommunication industry, multimedia files are produced, stored and shared easily across the globe. However, the ownership and copyright of multimedia files are not usually protected. Audio watermarking techniques are usually associated with content protection and digital rights management of music and to the government and law enforcement in the context of information systems security [1]. In watermarking a data message or watermark can be hided within a cover signal in the block called Embeddor using a stego key, which is a secret key of parameters of a known hiding algorithm as shown in Fig. 1. The output of the embeddor is called stego signal [2]. After compression, Transmission and other signal processing which may contaminate the stego signal, the embedded message or watermark signal is recovered using the same stego key as used in embeddor, in the extractor.

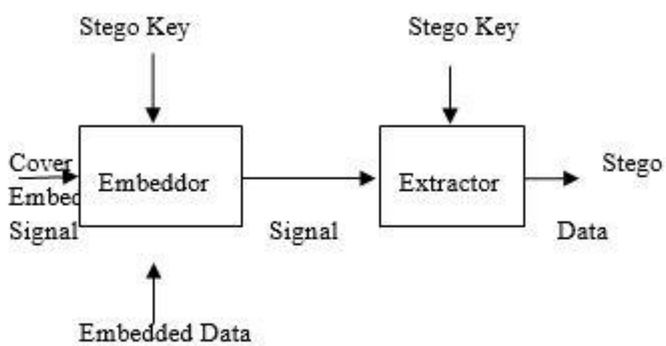


Fig.1 Block Diagram for data hiding & retrieval

II. AUDIO WATERMARKING DOMAINS

Audio watermarking techniques summarized in literature can be grouped into two types depending upon their domain of analysis: time-domain techniques and frequency domain techniques. The two watermarking domains have different characteristics and thus performances of their techniques may vary with respect to the robustness, complexity, strength and inaudibility requirements of digital watermarking. Inaudibility refers that the embedded watermark should not produce a distortion that can be easily sensed by human beings. Robustness determines the resistance of the watermark against removal and degradation. The watermark should survive malicious attacks such as random cropping and noise adding, and its removal should be impossible without perceptible signal alterations.

Time Domain Audio Watermarking is relatively easy and less complicated to implement. This requires less computational resources because there is no requirement of any domain transformation and each technique can be directly applied on the spatial domain. Thus it is less complicated to implement. But in this the strength of the watermark signal is very limited. Thus the watermark is not robust enough.

Frequency Domain Audio Watermarking techniques employ human perceptual properties and frequency masking characteristics of human auditory system for effective watermarking. In this the robustness of the watermark will increase and the watermark is more subtle and sturdy. But the computational cost in searching watermark will increase greatly at the same time.

TABLE I COMPARISON

Characteristics	Time Domain	Frequency Domain
Robustness	Low	High
Computational cost	Less	More
Watermark strength	Low	High
Watermarking Speed	More	Less

High robustness is preferable for any watermarking technique as it measures the capability of the embedded secret data to endure different types of intentional and unintentional modifications. The lesser the requirement of computations, the lower is the cost associated with it. Thus lesser computations and consequently low computational cost is desirable. The more the information we are able to hide in the cover signal, higher the technique is preferred. The imperceptibility is also a important requirement of a steganography technique, as the strength of steganography technique depends on its ability to be unnoticed by the human senses.

III. AUDIO WATERMARK TECHNIQUES

A. Least Significant Bit Substitution (LSB)

It is the time domain technique to embed information in a digital audio file.Thereceived audio file in the form of bytes is converted into the bit pattern [4]. Replace the LSB bit from audio with the message bits or the watermark bits, as shown in Fig. 2. Thus by substituting the least significant bit of each sampling point with a binary message, the data can be hided in the audio file.

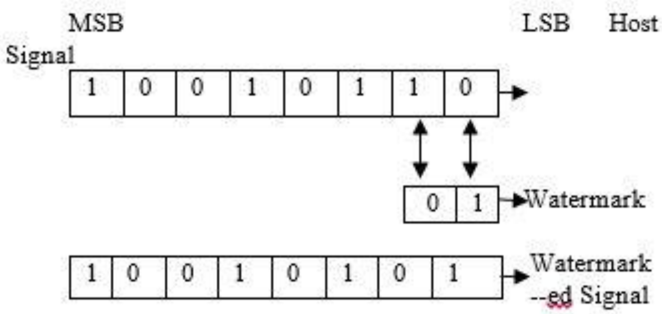


Fig. 2: Two-Bit LSB Watermarking

LSB watermarking may include one-bit watermarking or two and even three bit watermarking. As the number of watermark bits increases the robustness will increase but this will degrade the quality of original signal.

This system will not change the size of the file even after watermarking and also very fewer computations are required. The most important benefit is that this technique is suitable for any type of file format. But this will degrade the quality of the original signal to a significant extent.

B. Autocorrelation Modulation

This is also a Time domain technique of data hiding. In this a difference signal is added to the cover signal to obtain a watermarked signal, as shown in Fig. 3. The difference signal depends on cover signal, embedded data, as well as the stego key.

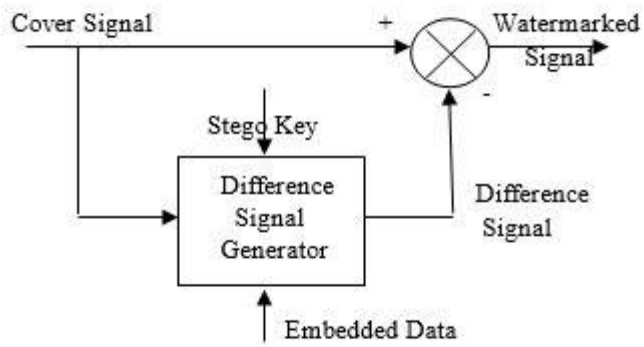


Fig. 3: Block Diagram of Embeddor

The block diagram of difference signal generator for short-term autocorrelation modulation is shown in Fig. 4. Input signal is obtained by applying a filter to the cover signal in order to obtain a portion of the cover signal that will result in minimal disturbances to the audio signal [1].

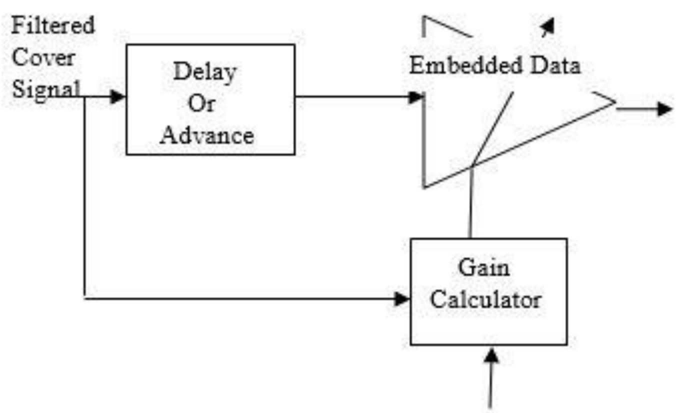


Fig. 4: Difference signal component generator

The difference signal component generator produces a difference signal component by applying a variable gain or attenuation to a delayed or advanced version of delayed or advanced version. The short term autocorrelation of the filtered cover signal can be expressed by the formula

$$R(t,\tau) = \int_{t-\tau}^{t} s(x)s(x-\tau)dx \qquad (1)$$

Where $s(t)$ is the filtered covered signal, $R(t,\tau)$ is its short term autocorrelation, τ is the delay at which correlation is evaluated, T is the integration interval and t is the time. By adding a difference signal $e(t)$ to the filtered cover signal, the short term autocorrelation of the resulting signal is

$$\mathrm{Rm}(t,\tau) = \int_{t-\tau}^{t} \{s(x)+e(x)\}\{s(x-\tau)+e(x-\tau)\}dx$$

$$= R(t,\tau) + \int_{t-\tau}^{t} \left(s(x)e(x-\tau) + e(x)s(x-\tau) + e(x)e(x-\tau)\right) dx \tag{2}$$

Where difference signal $e(t)$ is the delayed or advanced version of the existing signal multiplied by an amount of gain or attenuation. That is,

$$e(t) = g * s(t-\tau) \tag{3}$$

Substituting eq. (3) in eq. (2), we get

$$Rm(t,\tau) = R(t,\tau) + gR(t,2\tau) + gR(t-\tau,0) + g^2R(t-\tau,\tau) \tag{4}$$

The amount of gain g is small in order to keep the difference signal transparent with respect to the existing signal.
Thus g can be obtained as

$$g \approx \frac{Rm(t,\tau) - R(t,\tau)}{R(t,2\tau) + R(t-\tau,0)} \tag{5}$$

This technique is very easy to implement as inserting a delayed or advanced version of the signal itself can modify the autocorrelation. But in this the signal strength to be inserted must be very small. In this less number of computations are required, thus speed of watermarking is fast.

C. *Parity coding*

This technique introduces the watermarking in the Time Domain of the audio signal. Instead of breaking a complete signal down into individual samples, the parity coding method breaks the original audio signal 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 message bit to be encoded in the audio signal, the process complements the LSB of one of the samples in the region. Thus, the sender has more of a choice in encoding the secret bit, and the signal can be changed in a more unobtrusive fashion.

As in this technique the complete processing is done in the time domain and there is no requirement of domain transformation, the complexity and computational cost is reduced. But as one of the original message bit is changed, the original signal is considerably affected. In this technique watermarking capacity is very high.

D. *Discrete Fourier Transform (DFT)*

The DFT generally produces complex-valued frequency domain coefficients. Symmetry constraints in the DFT of real valued signals must be maintained in order to obtain a real-valued inverse transform. Fast algorithms exist for computing the DFT, such as the Fast Fourier Transform (FFT). The forward and inverse Discrete Fourier Transform of a signal having N samples in length can be written by the following transformation pair:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-jk\frac{2\pi}{N}n} \tag{6}$$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k] e^{-jk\frac{2\pi}{N}n} \tag{7}$$

Where k= 0 to (N-1).
X[n] is the signal in time domain and X[k] is the Discrete Fourier transform of x[n].
On the other hand, DCT produces real-valued coefficients that do not suffer from the symmetry constraints as in the case of DFT.
This technique is rarely used as it involves large amount of computations which leads to increase in complexity.

E. *Phase Coding*

Naturally occurring audio signals such as music or voice contain a fundamental frequency and a spectrum of overtones which are harmonic multiples of fundamental frequency, with well-defined relative phases. When the phases of the harmonic overtones are modulated to create a new composite waveform different from the original, the difference will not be easily detected [2]. Thus, the modification in the phases of the harmonics in an overtone spectrum of voice or music may be exploited as a channel for the transmission of hidden data. Consider the superposition of two sinusoidal signals that are out of phase by a constant factor:

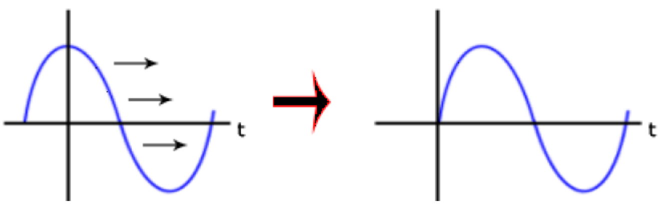
$$x(t) = \cos(2\pi f t) + \cos(2\pi f t + \emptyset) \tag{8}$$

The human auditory system is not able to discern the phase difference between the two signals. Modification in the phase must be sufficiently small and must be slowly varying in order to obtain theinaudible phase coding.
Phase coding is based on the principle that the phase components of sound are not easily perceptible to the human ear. This technique encodes the message bits as phase shifts in

the phase spectrum of a digital signal, achieving an inaudible encoding.

Phase coding is explained in the following procedure:
a. Segment the time representation of the original sound signal into a series of L frames, whose lengths equal to the size of the message to be encoded.
b. Apply Discrete Fourier Transform (DFT) to each frame of the audio data to create the matrix of the phases and fourier transform magnitudes of each frequency component.
c. Phase shifts between consecutive segments are easily detected. In other words, the absolute phases of the segments can be changed but the relative phase differences between the adjacent segments must be preserved. Therefore the secret message is inserted in the phase vector of the audio signal segment as follows:

$$Phase = \begin{Bmatrix} 0 & if\ message\ bit = 0 \\ \frac{\pi}{2} & if\ message\ bit = 1 \end{Bmatrix} \qquad (9)$$



Original Signal Encoded Signal

e. A new phase matrix is created using the new phase and the original phase differences.
f. The sound signal is reconstructed using the new phase matrix and original magnitude matrix, by applying the inverse DFT of L point and then concatenating the sound segments back together.
This is one of the most effective coding method in terms of the signal to perceived noise ratio. This technique can also easily survive MP3 and other compression and decompression techniques. Thus high degree of robustness can be achieved by this technique. In this technique watermark inserting capability is also very high. However modification of the phase must be sufficiently small in order to achieve an inaudible coding.

F. Discrete Wavelet Transform (DWT)

Wavelets are the special functions which provide powerful multi-resolution tool for analysis of signals with good time localization information. We embed the synchronization codes with the hidden informative data so that the hidden data have the self-synchronization ability [4]. Both the synchronization codes and informative bits are embedded into the low frequency subband coefficients in DWT domain to achieve strong robustness.

By exploiting the time-frequency localization capability of DWT, this technique reduces the computational complexity occurring in the searching of synchronization codes to a good extent. Thus they provide powerful multi-resolution tool for the analysis of non-stationary signals with good time localization capability. Thus resolves the contending requirements between the robustness and low computational complexity [3]. Watermarking using DWT technique is illustrated in Fig. 5

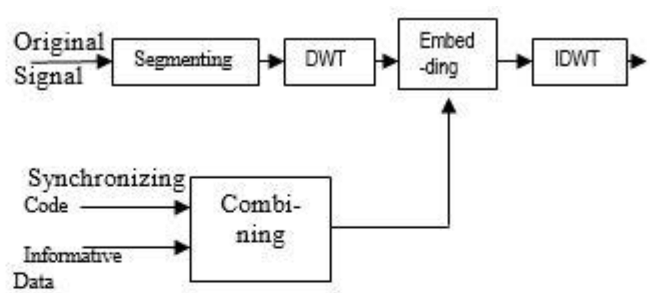


Fig. 5: DWT watermark embedding

In this the original audio signal is segmented in many sections and then one synchronization code and a portion of the informative data to be encoded is embedded in low frequency subband DWT coefficients of each section.
At the receiver end, we first perform the same DWT as in the embedding part of the audio signal and then extract binary data from the DWT coefficients of low frequency subband.

We then search for synchronization codes in the extracted data. This procedure as shown in Fig. 5, needs to be repeated by shifting the selected segment one sample at a time until a synchronization code is found. Once the synchronization code is found we can obtain the secret message bits, which are followed by the synchronization bits.

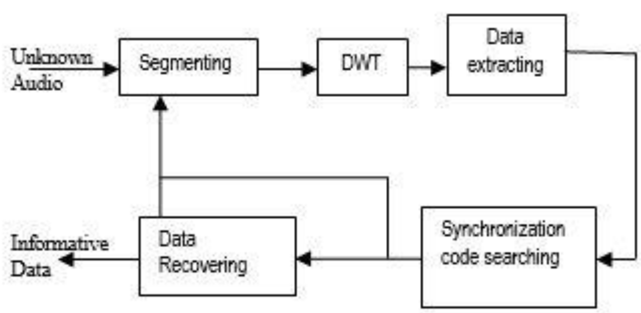


Fig. 6 Watermark Extraction

DWT is known to have the time-frequency localization capability. This characteristic can be used to improve the computational efficiency greatly in searching the synchronization codes. On the other hand, DWT itself needs a lower computation load compared with DCT and DFT. Thus by the use of DWT technique, the computational cost is reduced considerably.

The hidden data are robust against signal processing and attacks, such as Gaussian noise corruption, resampling, requantization, cropping and MP3 compression.

IV. REFERENCES

[1] Akira Nishimura: Data Hiding for Audio Signals that are Robust with respect to Air Transmission and a Speech Codec. International Conference on Intelligent Information Hiding and Multimedia Signal Processing. IEEE 2008.

[2] Rade Petrovic, Joseph M. Winograd, Kanaan jemili, Eric Metois. Data Hiding within audio Signals. IEEE 13-15 October 1999, Yugoslavia.

[3] Xiaoxiao Dong, Mark F. Bocko, Zeljko Ignjatovic. Data hiding with phase modulation of audio signals. IEEE 2004.

[4] Ali Al-Haj, Ahmad Mohammad and Lama Bata: DWT-Based Audio watermarking. The International Arab Journal of information technology, Vol. 8, No. 3, July 2011.

[5] Shaoquan Wu, Jiwu Huang, Senior Member, IEEE, Daren Huang, and Yun Q. Shi, Senior Member, IEEE: Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission. IEEE Transactions on Broadcasting, Vol. 51, No. 1, March 2005.

[6] A.R.Elshazly, M.M.Fouad: Secure and Robust High Quality DWT Domain Audio Watermarking Algorithm with binary image. IEEE 2012.

[7] Poulami Dutta, Debnath Bhattacharyya, and Tai-hoon Kim. Data Hiding in Audio Signals. International Journal of Database Theory and Application, Vol. 2, No. 2, June 2009.