Audio Watermarking Schemes: A Comparative Analysis

Jasmine Jolly[1], Mehbooba P Shareef[2]

[1]Christ College of engineering

[2] Rajagiri School of Engineering & Technology

(E-mail: jasminejparakkal@gmail.com, mehboobapshareef@gmail.com)

Abstract— The growth of internet and its reachability to all sectors of people have never been greater. Internet has become the best marketplace, the best library and may be the best guide for everything. But this revolution comes with some bigger problems. One of the most challenging problems among them would be copyright protection of digital data being transferred over internet. Digital images, videos and audios undergo illegal re-production and re-distributions, tampering and other acts of copyright violation. This is proved to have leading the film and other prominent industries to loss of millions of dollars per year. Encrypting the data provides security to it. In this case only people who pay to buy the secret key that should be used for decryption can use the data. But the problem is that once decrypted, the data can be re-produced into any number of copies and can be re-distributed without any permission from the author. Watermarking is an intelligent solution for this problem where the presence of watermark can be checked to distinguish pirated copies from the actual ones. A lot of methods have been developed for image and video watermarking, but the research on audio watermarking started a little bit later. The reason might be the fact that audio watermarking is tedious compared to image and video as Human Auditory System (HAS) is more sensitive compared to Human Visionary System. So ensuring the imperceptibility of audio watermarks is a tougher task. In this thesis various audio watermarking schemes introduced so far in the literature and their merits and demerits are studied.

Keywords— *audio watermarking; spread spectrum; transformed domain; Singular value Decomposition; psycho acoustic model; Empirical Mode Decomposition*

I. Introduction

Protection of digital media data has gained tremendous research interest in recent years. Digital watermarking is a promising technique to solve this problem. Digital watermarking aims to embed ownership information like signature, logo, id etc into the media object. Owners can extract the embedded watermarks, to declare their copyright, especially in the case of disputes. While digital watermarking research started much earlier in the areas of digital images, and videos, it took time to begin in the area of audio watermarking. The sensitivity of Human Auditory System to even a slight change in the audios was the major challenge. This paper limits its attention to audio watermarking schemes proposed in the literature so far. Audio signals are one dimensional data. And human auditory system is very sensitive to even very small changes in audio. These things make audio watermarking very difficult compared to its image and video counterparts. [2], [6]. Imperceptibility, robustness, and security are the main properties that any audio watermarking scheme should possess. Imperceptibility indicates that the watermark should be inaudible in the watermarked audio signal. Robustness points to the ability of recovering the watermark data from the watermarked signal, irrespective of whether the audio track has been attacked or not. Security implies that an unauthorized user cannot extract or delete the watermark without using a secret key. In addition, low computation complexity is also considered as an advantage since it reduces the time of watermark embedding and extraction [1]. This feature is particularly important for some time-demanding applications, such as delivering the audio data over the Internet. Furthermore, a good watermarking method should be able to extract watermarks at the decoding stage without making use of the host audio signal.

II. literature survey

During the last decade, many audio watermarking schemes have been developed. Audio watermarking can be broadly classified into Spread spectrum watermarking, Patch work based algorithms, Time domain watermarking, Transform domain watermarking, watermarking using Singular Value Decomposition and watermarking using Empirical Mode Decomposition. Still, there are ofcourse watermarking

schemes which use more than one of these approaches at a time so that a better hybrid method is resulted. This chapter reviews the literature of information hiding in audio sequences. Scientific publications included into the literature survey have been chosen in order to build a sufficient background that would help out in identifying and solving the research problems. Watermarking schemes can be blind (watermark can be extracted from the watermarked signal without the presence of original signal) or non-blind and robust(watermark signal remains unaffected by attacks) or fragile. All these kinds of watermarking schemes have been studied.

A. Spread Spectrum Audio Watermarking

Most of the existing audio watermarking techniques embed the watermarks in the time domain/ frequency domain where as there are few techniques which embed the data in cepstrum or compress domain. Spread spectrum (SS) technique is most popular technique and is used by many researchers in their implementations. Amplitude scaled Spread Spectrum Sequence is embedded in the host audio signal which can be detected via a correlation technique. Psychoacoustic models provides in-audibility limits to watermarks embedded. This is followed in watermark embedding. Here watermark is spread over a large number of co-efficient. The distortion caused due to this is kept just below the Just Noticeable Difference level. Change in each coefficient can be small enough to be imperceptible because the correlated detector output still has a high signal to noise ratio (SNR), since it dispreads the energy present in a large number of coefficients.

D. Kirovski et al [1] developed the techniques which effectively encode and decode the direct sequence spread spectrum watermark in audio signal. They have used the modulated complex lapped transform to embed the watermark. To prevent the de-synchronization attack they developed the technique based on block repetition coding. Though they have proved that they can perform the correlation test in perfect synchronization, the wow and flutter induced in watermarked signal may cause false positive/false negative detection of watermark. To improve the reliability of watermark detection they proposed the technique which uses cepstrum filtering and chess watermarks. It was observed from their study that psychoacoustic frequency masking creates an imbalance in the number of positive and negative watermark chips in the part of the Spread Spectrum sequence. The Spectrum sequence is used for correlation detection which corresponds to the audible part of the frequency spectrum. To compensate this problem they propose a modified covariance test.

B. Methods using patchwork algorithm

The patchwork technique was first presented in 1996 by Bender et al [2] for embedding watermarks in images. It is a statistical method that uses very large data sets and is based on hypothesis testing. A second of CD quality stereo audio contains 88200 samples. Because of this reason, a patchwork approach is applicable for the watermarking of audio sequences as well. Certain statistic is inserted to the host data using a pseudorandom process. The embedded watermark is extracted with the help of numerical indexes (like the mean value), describing the specific distribution. To spread the watermark in time domain and to increase robustness against signal processing modifications,the method is usually applied in a transform domain (Fourier, wavelet, etc.)

Cvejic et al [3] presented a robust audio watermarking method implemented in wavelet domain which uses the frequency hopping and patchwork method. Their scheme embeds the watermark to a mapped sub band in a predefined time period similar to frequency hopping approach in digital communication and detection method is modified patchwork algorithm. Their results show that the algorithm is robust against the mp3 compression, noise addition, re-quantization and re-sampling. This scheme uses a pseudo acoustic model for watermark to be inaudible in the host signal. For this system to be robust against the re-sampling attack it is required to find out the proper scaling parameter. The disadvantage here is that DCT,DFT patches assumes that the patches have same statistical properties which is not true.

Iyenkaaran et al [4]proposed a patchwork based watermarking scheme in DCT domain. The audio segment is converted into sub-segments and DCT coefficients of the sub-segments are found out. DCT coefficients related to specified frequency region are found and DCT coefficients are used to pair up frames. The frame pairs are chosen on the basis of certain criteria to decide whether to embed watermark in them or not. The watermark is embedded into the frame-pairs by controlling their coefficients. To improve the security, this method uses pseudo noise sequence of length 2M to sort the DCT co-efficients in a frame to two fragments. Watermark embedding is done by altering the means of the corresponding absolute-valued fragments.

C. Methods implemented in Time Domain

There are few algorithms implemented in time domain. These algorithms embed the watermarks in the host signal in time domain by modifying the selected samples.

A.N. Lemma et al [5] investigated an audio watermarking system is referred to as modified audio signal keying (MASK). In MASK, the short-time envelope of the audio signal is modified in such a way that the change is imperceptible to the human listener. In MASK, a watermark is embedded by modifying the envelope of the audio with an appropriately conditioned and scaled version of a predefined random sequence carrying some information (a payload). On the detector side, the watermark symbols are extracted by estimating the short-time envelope energy. To this end, first, the incoming audio is subdivided into frames, and then, the energy of the envelope is estimated. From this energy function, the watermark is extracted.. The MASK system can easily be customised for a wide range of applications. Experimental results show that it has a good robustness and audibility behaviour.

Paraskevi Bassia et al [6] proposed a watermarking scheme in time domain. The watermark signal is generated using a key, i.e., a single number known only to the copyright owner. Watermark embedding is done in the time domain here. As a result opf this the amplitude gets modified by a magnitude which is not detected by the Human Auditary system. Let us assume an audio signal of samples. The audio signal is divided in segments of samples each. Each of these segments is watermarked with the same bipolar Sequence. The detection procedure does not use the original signal. This watermarking scheme is statistically imperceivable and resists MPEG2 audio compression plus other common forms of signal manipulation, such as cropping, time shifting, filtering, resampling and requantization. However the method is not robust to more sophisticated attacks. For example one would not be able to detect the watermark in a signal that has been subject to a change in the time scale.

D. Methods implemented in Time Domain

X-Y. Wang *et al* [7] proposed a blind digital audio watermarking scheme against synchronization attack. They used adaptive mean quantization. The features of the method are as follows 1) a kind of more steady synchronization code and a new embedding strategy are adopted which resist the synchronization attack more effectively; 2) the multi-resolution characteristics of DWT and energy-compression characteristics of DCT are combined to improve the transparency of digital watermark 3) the watermark is embedded into the low frequency components by adaptive quantization in such a way that the HAS will not recognize the changes to the signal and 4)the scheme is blind(can extract the watermark without the help of original audio signal). The experimental result show that the technique can resist the various signal processing attacks.

Pranab Kumar Dhar et al [8] proposed a Discrete Fourier Transform based watermarking scheme .This method calculates the magnitude and phase spectrum of each frame using (DFT),finding the most prominent peak **V** from magnitude spectrum using a peak detection algorithm,and placing watermarks into the highest prominent peak of the magnitude spectrum of each frame to obtain watermarked peak **V'**. This ensures that the watermark is located at the most significant components of the audio.

M. Fallahpour et al [9] proposed a DWT–based high capacity audio watermarking scheme. They computed the third level wavelet transform of the original signal and divided the DDD samples into frames of a given length. Then based on the average of the absolute values of each frame's samples, computed the average coefficient for each frame. Each secret bit is embedded in a single suitable coefficient. After embedding the bit, the index *l* is incremented and the next secret bit is embedded in the next suitable coefficient.

Many techniques are implemented in wavelet domain [10-16,35-41]. And it is found that wavelet domain is the more suitable domain compare to the other transform domains. As the wavelet coefficients contain the multiple spectrums of multiple band frequencies, this transform is more suitable than other transform domains to select the perceptible band of frequencies for data embedding.

E. Methods implemented in Time Domain

Hamza Özer et al [17] proposed An SVD-Based Audio Watermarking Technique. They first decompose the STFT of the audio signal, organized as a time-frequency matrix, into its singular values, and then mark the singular value matrix. The watermarking pattern and bit polarity are used to modify the singular value matrix *D* of the host object. The singular value decomposition is a numerical tool, which effectively decomposes a matrix into two orthogonal matrices and its singular values. Thus a matrix *A* is decomposed into *A = U D VT*, where *A* is the *LxK* matrix that we want to summarize, *D* is *LxK* matrix with only min(L,K) diagonal elements, *U* is an *LxL* orthogonal matrix, and *V* is an *KxK* orthogonal matrix. SVD is attractive because of the property that the singular values are invariant under orthogonal transformations. The SV decomposition of the STFT matrix of each frame is calculated , and then embed the watermark bits in the *D* singular value matrix. However this method is semi-blind and the transformation matrices are to be stored which is bandwidth-consuming.

Vivekananda Bhat K et al [18] proposed an SVD based watermarking scheme in wavelet domain. Here, a watermark is embedded by applying a quantization-index-modulation process on the singular values in the SVD of the wavelet domain blocks. The watermarked signal has good PSNR values here. The algorithm is said to be robust to additive noise, re-sampling, low-pass filtering, re-quantization, MP3 compression, cropping, echo addition, and de-noising. Though the method is blind, quantization parameters used in the sender end should be sent to the receiver end.

F. Methods implemented in Time Domain

Liang Wang et al [19] proposed an EMD and psycho acoustic model based watermarking for audio. By applying the EMD, any multi-component signal is decomposed into a set of intrinsic mode functions (IMFs). The IMF can either be amplitude or frequency modulated. It is defined as a hidden oscillation mode and it is embedded in the data series, and it is allowed to be non-stationary. The watermark message is embedded into the Waveform Audio File Format (WAV) audio signal. Here the bit stream is encoded with the Pulse Code Modulation (PCM) format. For the audio and speech processing, the PCM samples are stored and processed using floating point numbers which have the zero mean (or the mean value is sufficiently small compared with the amplitude of the signal) and varies in the interval [-1.0, 1.0]. Thus,compared with the original audio signal, the amplitude of its final residue can be regarded sufficiently small which

makes is possible to embed the watermarks in the final residue of the audio signal while the watermark messages are perceptually inaudible. This method is not robust to attacks such as band-pass filtering and cropping.

A N K Zaman [20] et al proposed an EMD based watermarking scheme using Hilbert transforms. They extended the idea of Liang et al. [] by embedding watermark to the significant IMF containing highest energy for increasing robustness against different signal processing attacks. Firstly, they have divided the host signal into a number of frames. Then each frame is decomposed into a finite and often small number of intrinsic mode functions (IMFs). Because of the local characteristic time scale of the data this method is applicable to nonlinear and non-stationary process But they haven't explained the reason for choosing the IMF of highest energy. Because in practice an IMF with highest energy can be a high frequency mode and thus it is not robust to attacks[21].

Kais Khaldia et al [21] proposed a new adaptive audio watermarking algorithm based on Empirical Mode Decomposition (EMD) is introduced. The audio signal is divided into frames . Then each frame is decomposed adaptively, by EMD, into Intrinsic Mode Functions (IMFs). The watermark and the synchronization codes are embedded into the extrema of the last IMF.Last IMF is low frequency mode stable under different attacks and preserving audio perceptual quality of the host signal. The data embedding rate of the proposed algorithm is 46.9–50.3 b/s. Relying on exhaustive simulations, they show the robustness of the hidden watermark for additive noise, MP3 compression, re-quantization, filtering, cropping and re-sampling.

G. Papers studied on the performance of watermarking schemes

J. D. Gordy et al [22] in their article titled 'Performance Evaluation of Digital Audio Watermarking Algorithms' proposed a an algorithm-independent framework for rigorously comparing digital watermarking algorithms with respect to bit rate, perceptual quality, computational complexity, and robustness to signal processing. They evaluated five audio watermarking algorithms from the literature, revealing that frequency domain techniques perform well under the criteria. Four criterions were selected by authors as a part of the evaluation framework.

1) Bit rate refers to the amount of watermark data that may be reliably embedded within a host signal per unit of time or space, such as bits per second or bits per pixel. A higher bit rate may needed in some applications in order to embed more copyright information. Reliability is the measure of bit error rate (BER) of extracted watermark data.

2)Perceptual quality is the imperceptibility of embedded watermark data within the host signal. Presence of watermark should be undetectable to the listener or viewer in most of the applications. The signal-to-noise ratio (SNR) of the watermarked signal versus the host signal was used as a quality measure.

3) Computational complexity is the processing time required to embed watermark data into a host signal, and / or to extract the data from the signal. Actual CPU timings (in seconds) of algorithm implementations were collected.

4) Watermarked digital signals may undergo common signal processing operations such as linear filtering, sample requantization, D/A and A/D conversion, and lossy compression. Although these operations may not affect the perceived quality of the host signal, they may corrupt the watermark data embedded within the signal.

G.C. Rodriguez *et al* [23] presented a survey report on audio watermarking in which watermarking techniques are briefly summarized and analyzed. They have made the following observations:

• The patchwork scheme and spectrum domain scheme are robust to several signal manipulations, but for real applications authors suggest to use patchwork scheme because the spectrum domain scheme needs the original signal to determine that the host signal is marked as a consequence it needs the double storage capacity.

• The echo hiding scheme only fulfill with the inaudibility condition and is not robust to several attacks such as mp3 compression, filtering, resampling, etc.

III. conclusion

Different audio watermarking schemes have been discussed. The latest techniques for audio watermarking use Empirical Mode Decomposition. Khais Khaldhi[21] et al's method is an efficient one which is robust against attacks like filtering, cropping, resampling, re-quantization, compression etc. The use of synchronization codes and the choice of residue of every frame for embedding watermark in the account for its robustness. However, this method can be enhanced with a psycho acoustic model to ensure the inaudibility of the watermarking during the silent periods of the audio signal.

References

[1] D.Kirovski and H.S.Malvar, "*Spread spectrum watermarking of audio signals*" , IEEE Transactions on Signal Processing, Vol. 51, No. 4, April 2003 p.p. 1020-1033.

[2] W. Bender, D. Gruhl, N. Morimoto and A. Lu, "*Techniques for data hiding*", IBM system Journal, 1996,Vol. 35, p.p. 313-336.

[3] N. Cvejic and T. Seppanen, "*Robust Audio watermarking in Wavelet Domain Using Frequency Hopping and Patchwork method*", Proc. of 3rd International Symposium on Image and Signal processing and Analysis 2003, p.p. 251-255.

[4] Iynkaran Natgunanathan, Yong Xiang, Yue Rong, Wanlei Zhou, and Song Guo, Robust Patchwork-Based Embedding and Decoding Scheme for Digital Audio Watermarking *IEEE*

[5] A. N. Lemma, J. Aprea, W. Oomen and L. V. D. Kerkhof, "*A Temporal Domain Audio Watermarking Technique*", IEEE Transaction on Signal Processing, Vol. 51, No. , April 2003, p.p. 1088-1097

[6] P. Bassia and I. Pitas, "*Robust audio watermarking in the time domain*", IEEE Transactions on Multimedia,Vol. 3, No.2, June

2001 p.p.232-241.X. Y. Wang and H. Zhao, "*A Novel Synchronization Invariant Audio Watermarking Scheme based on DWT and DCT*", IEEE Transaction on signal processing, Vol.54, No.12, December 2006, pp 4835-4839.

[7] Mahesh, Bhasutkar, Maninti Venkateswarlu, and M. Raghavendra. "End-to-end congestion control techniques for Router." 2011 International Conference on Communication Systems and Network Technologies. IEEE, 2011.

[8] Mahesh, B., and K. Shyam Sunder Reddy. "Router Aided Congestion Control Techniques." Second International Conference on Information Systems and Technology.

[9] Mahesh, B. "Dynamic Update and Public Auditing with Dispute Arbitration for Cloud Data." Journal of Advanced Database Management & Systems 4.3 (2017): 14-19.

[10] Mahesh, B., et al. "A Review on Data Deduplication Techniques in Cloud." Embedded Systems and Artificial Intelligence. Springer, Singapore, 2020. 825-833.Chin-Su Ko, K.Kim, R.Hwang, Y. Kim and S.Rhee, "*Robust Audio Watermarking in wavelet domain using PN sequences",* Proc. of ICIS-2005 published by IEEE. 120

[11] R.Vieru, R. Tahboub,C. Constantinescu and V. Lazarescu, "*New results using the audio watermarking based on wavelet transform*", International Symposium on Signals, Circuits, and Systems, Kobe, Japan 2 (2005) published by IEEE 2005, p.p. 441-444.

[12] R.Wang, Dawen Xu and L Qian, "*Audio Watermarking based on wavelet packet and Psychoacoustic model",* IEEE Proc. of PDCAT-2005.

[13] Wang, Xu, Z. Hang and C. Youngrui, "*2-D digital Audio Watermarking based on Integer Wavelet Transform*", IEEE Proc. of ISCIT 2005, p.p. 877-880.

[14] S.Ratansanya, S.Poomdaeng, S. Tachaphetpiboon and T. Amornraksa, "*New Psychoacoustic models for wavelet based Audio watermarking*", IEEE Proc. of ISCIT 2005, p.p. 582-585.

[15] Hamza Özer Bülent Sankur Nasir Memon, An SVD-Based Audio Watermarking Technique

[16] Vivekananda Bhat K , Indranil Sengupta, Abhijit Das , An adaptive audio watermarking based on the singular value decomposition in the wavelet domain, Digital Signal Processing 20 (2010) 1547–1558.

[17] *Liang Wang, Sabu Emmanuel, Mohan S. Kankanhalli,* Emd And Psychoacousticmodel Based Watermarking For Audio, 978-1-4244-7493-6/10/$26.00_c 2010 IEEE

[18] A N K Zaman1, K.M. Ibrahim Khalilullah2, Md. Wahedul Islam and Md1. Khademul Islam Molla, A Robust Digital Audio Watermarking Algorithm Using Empirical Mode Decomposition

[19] Kais Khaldi and Abdel-Ouahab Boudraa, Audio Watermarking Via EMD, IEEE transactions on audio, speech, and language processing, vol. 21, no. 3, march 2013 675

[20] J. D. Gordy and L. T. Bruton , Performance Evaluation of Digital Audio Watermarking Algorithms G.C. Roddriguez, M. N. Miyatake and H. M. P. Meana, "*Analysis of Audio Watermarking Schemes*", Proc. of ICEEE 2005, p.p. 17-20.

[21] G.C. Roddriguez, M. N. Miyatake and H. M. P. Meana, "*Analysis of AudioWatermarking Schemes*", Proc. of ICEEE 2005, p.p. 17-20