DIGITAL AUDIO WATERMARKING APPLICATIONS AND TECHNIQUES

G. B. Khatri¹, D. S. Chaudhari²

¹(Department of EXTC, GCOE, Amravati, India)
²(Department of EXTC, GCOE, Amravati, India)

ABSTRACT

Digital audio watermarking is nowadays gaining popularity among the original audio content producers. It provides a way to the owner to prove ownership by using watermark which forms integral part of audio. An approach is made in this paper to collectively introduce applications and techniques up till developed for audio watermarking. The proposed algorithm is based on wavelet transform which has efficient computational load and also on cosine transform for increasing the watermark embedding capacity.

Keywords: Watermark, audio, DFT, DCT, DWT, and DSSS.

I. INTRODUCTION

Production, storage, distribution of digital multimedia data is very easy. Hence it creates the problem of protecting the intellectual copyrights. The copyrighted digital multimedia data is pirated without notification to the owner. Digital watermarking is a technique to embed the owner information as a copyright material in the digital data as a proof of ownership. Watermark as the name suggest is as transparent as water when watermark data is embedded in the original audio. Watermarking performance can be judged on parameters like imperceptibility, robustness, efficiency, and embedding capacity. A watermarking technique must achieve high performance without degrading the cover signal [1].

Digital watermarking can be applied to image, audio or video and the watermark data can be an image, audio and text. Generally there is less attention towards audio watermarking because HAS (human auditory system) is more sensitive than HVS (human visual system) and human ears can easily detect the presence of the watermark as low as one part in ten million [2].
This paper presents an overview on applications and techniques of digital audio watermarking. Section II describes the digital audio watermarking, section III briefly describes four areas of applications whereas in section IV watermarking techniques are discussed with the performances of each and in section V new algorithm is proposed.

II. DIGITAL AUDIO WATERMARKING

Digital audio watermarking is a technique of embedding watermark data such as image, audio and text in the original audio stream to create copyrighted watermarked audio. Audio watermarking application areas include such as vendor identification, evidence of proprietorship, validation of genuineness, copy protection, etc. [3]

Audio watermarking techniques can be grouped as; time-domain, frequency (transform) domain, spread-spectrum, patchwork [1,4]. The performances of the techniques are judged with respect to the robustness and imperceptibility (inaudibility) of audio watermarking. Inaudibility means the watermarked audio and original audio must be identical in nature to listen. Robustness means the resistance of the watermark against removal or degradation. The watermark should survive intentional attacks such as random cropping, noise addition, requantization, resampling, compression, filtering and its removal should degrade original audio.

III. AUDIO WATERMARKING APPLICATIONS

Watermarking serves various applications and each application puts desirable feature necessity on the watermarking technique. Hence the watermarking technique to be used depends on the area of application [3]. Thus a variety of applications are discussed below,

1) Vendor Identification –
Text form of copyright notices occurs on the packaging of copyrighted materials. This type of protection does not prove sufficient as it would be easily removed. Digital audio watermarking can be used to embed copyright notice in the audio signal itself. As notice forms an integrated part of audio one can determine the vendor of the copyrighted audio.

2) Evidence of Proprietorship –
One can prove its proprietorship in the case of copyright dispute. The original owner can prove its proprietorship by extracting the watermark copyright information from the watermarked audio, in the case when another person tries to sell the copyrighted material on behalf of his name by pirating it.

3) Validation of Genuineness –
The copyrighted audio is genuine or not, can be proved very easily by the use of watermarking. A signature or copyright watermark is embedded in the audio thus anyone trying to modify the watermarked audio, modification also applies to the watermark because watermark forms integral part of audio. Hence one can prove genuineness of copyrighted audio by extracting exact copy of the watermark.

4) Copy Protection –
The above mentioned applications do not put restriction on the illegal copying. Owner can restrict the illegal copying or the numbers of copies permitted by the use of special watermarking algorithms. One such method is modified patchwork algorithm (MPA) developed by Yeo and Kim [4].
IV. AUDIO WATERMARKING TECHNIQUES

Though audio watermarking techniques are few in numbers they can be categorised majorly into four groups viz. time domain, frequency domain, spread spectrum and patchwork. These techniques are discussed one by one along with their performances.

1) Time-domain techniques –

It includes the least significant bit (LSB) substitution, echo hiding and quantization techniques. In time-domain techniques parameters of signal samples like amplitude, masking of samples with lower amplitude by higher amplitude samples, samples LSB are varied to attain embedding of watermark. The time-domain audio watermarking is relatively easy to implement, and requires few computing resources, however, it is weak against signal processing attacks such as compression and filtering. This type of techniques suffers from problems like low watermark data embedding capacity, easily detectable by the attacker, easy to decode watermark from cover audio.

LSB technique is the simplest technique in which watermark data is embedded in the least significant bits of the audio sample values. It provides easy data embedding and extracting algorithms. As the information contained in the LSB is less, it is replaced by the data of the watermark without producing the noticeable effect in the cover audio signal. The cover audio signal degrades as the number of watermark bits is increased. A maximum of 3 watermark bits per 16 bits of audio sample is allowed for imperceptibility. If above 3 bits per audio sample is embedded distortions like noise are introduced and human auditory system begins to detect the noise introduced by the watermark. Cvejic N. and Seppanen T. have tried to increase the capacity from 3 bits/sample to 4 bits/sample without degrading the watermarked audio signal to noise ratio by using a three step technique. In this degradation in SNR of the watermarked audio is minimised by using minimum error replacement and error diffusion steps.

Echo-hiding watermarking embeds information into the original discrete audio signal by introducing a repeated version of an original sample of the audio signal with some delay and decay rate so as to make it undetectable [6]. Only binary information in the form of bits is embedded in the audio signal. Digital data is embedded by using four main parameters of echo: initial value, decay rate and different offset for 1 and 0. The offset is made so small such that the human ear cannot detect the presence of echo. Embedding is done by convolving the audio signal with the all 0 and all 1 kernel, then by using watermark data bits particular outputs are combined to form watermarked signal. Extraction is done by taking autocorrelation of cepstrum of the watermarked audio. Autocorrelation gives the power of signals at various shifts. With particular shifts in it one can easily determine the bit embedded. The watermark data embedding rate is given as 16 bps (bits per second), while it can vary in the range 2 – 64 bps and it depends on the sampling rate and the signal type to be echoed.

In the technique of quantization original sample of audio is replaced with the modified audio sample. The modified audio sample is defined as below,

\[
y = \begin{cases} 
q(x,A) + \frac{A}{4}, & \text{if data bit is 1} \\
q(x,A) - \frac{A}{4}, & \text{if data bit is 0}
\end{cases}
\]  
(1)
Where \( q(.) \) is quantization function and \( A \) is quantization step. The quantization function is given as,

\[
q(x, A) = \lfloor x/A \rfloor A
\]  

(2)

Where \( \lfloor x/A \rfloor \) is rounded to nearest integer. Thus in a single sample of audio signal one can embed only one bit of watermark. Hence a blind detection can be applied for watermark data extraction. Extraction can be done by following equation,

\[
b = \begin{cases} 
1, & \text{if } 0 < y - q(x, A) < A/4 \\
0, & \text{if } -A/4 < y - q(x, A) < 0 
\end{cases}
\]

(3)

This technique is simple and easy to implement and is robust to noise as long as the noise margin is below \( A/4 \). While the technique is easy but the watermark embedding capacity is very less.

2) Frequency domain techniques –

Frequency domain audio watermarking techniques generally include transforms like discrete Fourier transform (DFT), the discrete cosine transform (DCT), and the discrete wavelet transform (DWT). It takes the advantage of masking of different tones of human auditory system (HAS) for effective watermarking. In this the original audio signal is transformed into frequency domain by using any one of the above mentioned transforms and then the watermark is either added or replaced in the magnitude or phase response. After that the watermarked audio is applied inverse transform to obtain the watermarked audio in time domain.

Discrete Fourier transform decomposes the signal into its fundamental and harmonically related sinusoidal frequencies. The human ears sensitivity declines after the peak sensitivity around 1 kHz. Magnitude response coefficients are replaced by the watermark data in the frequency range of 2.4 – 6.4 kHz [7]. Also the human ears are insensitive to the absolute phase of the audio frequency; hence the phase difference between the phase signal coefficient and phase reference coefficient is used to modify the phase signal coefficient. Phase difference has to be added or subtracted when the watermark data bit is 1 or 0 respectively [8].

Discrete cosine transform is similar to the discrete Fourier transform except that its coefficients are real valued. Properties of DCT such as high compaction of signal energy in transform domain, highly decorrelated coefficients are used to embed data in the transform domain.

Discrete wavelet transform is nowadays gaining popularity because it can decompose the signal in time and frequency at the same time while keeping the calculations to obtain DWT coefficients small as compared to DFT and DCT. Several advantages of applying DWT on audio signal are given by Wu and Huang such as 1) It is able to localize the audio in time-frequency both with multi-resolution property, 2) variable decomposition levels are available, 3) less number of operations than DFT and DCT [9]. If there are \( N \) samples in the audio then number of operations in DFT, DCT and DWT are \( O(N \cdot \log_2(N)) \), \( O(N\cdot \log_2(N)) \) and \( O(L \cdot N) \) respectively, where \( L \) is the length of wavelet filter. A data payload capacity of 172 bps is achieved by embedding the self-synchronised watermark data in the wavelet domain without degrading the SNR too much[9].
3) Spread-spectrum technique –

It involves embedding of watermark in the original audio signal by spreading it over the bandwidth of audio signal[6]. This technique utilizes DSSS (Direct Sequence Spread Spectrum) method of spread spectrum. Encoding of watermark in the audio is shown in Figure 1.

![Diagram of Direct sequence spread spectrum encoding](image)

**Fig. 1** Direct sequence spread spectrum encoding

In DSSS PN-sequence in used to spread the watermark data in the whole frequency range of audio and then added to the audio signal by proper attenuation, so that the watermark data is treated as additive random noise. Same sequence is again used to extract the watermark data by performing correlation between watermarked audio and PN-sequence.

4) Patchwork technique –

This technique was first proposed for image watermarking and it is a pseudorandom statistical approach [6]. The idea is to select two subsets (patch) of the cover signal and in order to embed the watermark, the sample values of these two subsets are moved in opposite directions by a constant value, which defines the watermark strength and watermark bit. The imperceptibility of watermark in cover signal depends on value of $d$. Decoding is performed by taking the difference of the means of these two subsets and making decision based on the obtained value. The assumption in this method is that the difference of the means of the two patches is zero for the original cover signal and is nonzero for the watermarked cover signal. As two subsets (patch) are used this technique can extract the watermark without the original cover signal.

Yeo and Kim have proposed a modification on the patchwork technique [4]. They have embedded the watermark data bit in the two subsets taken from DCT domain of original audio signal. The use of transform domain for embedding watermark makes the technique more robust against signal processing attacks such as down-sampling, equalization, compression, filtering.

**V. PROPOSED ALGORITHM**

As observed from the above techniques that transform domain are more secure than time domain. In these models are suggested for embedding and extracting the watermark data in and out of audio signal. These models utilize the discrete wavelet transform for speedily and efficiently transforming audio in time-frequency domain, while using discrete cosine transform to decorrelate and compress watermark image.

Transformed watermark image coefficients must be normalised and multiplied with an attenuation constant before embedding. Attenuating the coefficients helps to keep noise level low in the audio signal. Since watermark image is compressed using discrete cosine
transform less number of transform coefficients are used for embedding and this improves the signal to noise ratio and also the watermark data embedding capacity.

Discrete wavelet transform is used to transform the audio signals into frequency sub bands. These sub bands are called as approximate or detail frequency sub bands. Any one of this frequency band can be used to embed watermark.

This algorithm is based on emerging wavelet transform and cosine transform for embedding and extracting watermark in audio signal, hence it will possess characteristics such as high signal-to-noise ratio (SNR) of watermarked audio, high data payload capacity i.e. number of watermark bits per second of audio signal, low computational complexity.

VI. CONCLUSION

Several algorithms are discussed for digital audio watermarking. These algorithms suffer from problems like intentional or unintentional attacks by means of signal processing. Lacunas such as unblind detection, less embedding capacity, resistance to various types of attacks also occurs. The proposed algorithm can overcome problems in existing techniques discussed while at the same time can achieve high performance with less computational cost.

REFERENCES

AUTHORS’ INFORMATION

Govind B. Khatri is pursuing M.Tech. in Electronic System and Communication from Government College of Engineering, Amravati (Autonomous Institute of M.S.). He received his B.E. in Electronics and Telecommunication from H.V.P.M.'s College of Engineering & Technology in 2011. He also received Diploma in Electronics and Communication from Government Polytechnic, Amravati (Autonomous Institute of M.S.) in year 208. His area of interest includes Digital Signal Processing, Wavelet Processing, Digital Watermarking, and HDL with FPGA.

Dr. Devendra S. Chaudhari obtained BE, ME, from Marathwada University, Aurangabad and PhD from Indian Institute of Technology, Bombay, Mumbai. He has been engaged in teaching, research for period of about 25 years and worked on DST-SERC sponsored Fast Track Project for Young Scientists. He has worked as Head Electronics and Telecommunication, Instrumentation, Electrical, Research and in charge Principal at Government Engineering Colleges. Presently he is working as Head, Department of Electronics and Telecommunication Engineering at Government College of Engineering, Amravati. He has published research papers and presented papers in international conferences abroad at Seattle, USA and Austria, Europe. He worked as Chairman / Expert Member on different committees of All India Council for Technical Education, Directorate of Technical Education for Approval, Graduation, Inspection, Variation of Intake of diploma and degree Engineering Institutions. As a university recognized PhD research supervisor in Electronics and Computer Science Engineering he has been supervising research work since 2001. One research scholar received PhD under his supervision. He has worked as Chairman / Member on different university and college level committees like Examination, Academic, Senate, Board of Studies, etc. he chaired one of the Technical sessions of International Conference held at Nagpur. He is fellow of IE, IETE and life member of ISTE, BMESI and member of IEEE (2007). He is recipient of Best Engineering College Teacher Award of ISTE, New Delhi, Gold Medal Award of IETE, New Delhi, Engineering Achievement Award of IE (I), Nashik. He has organized various Continuing Education Programmes and delivered Expert Lectures on research at different places. He has also worked as ISTE Visiting Professor and visiting faculty member at Asian Institute of Technology, Bangkok, Thailand. His present research and teaching interests are in the field of Biomedical Engineering, Digital Signal Processing and Analogue Integrated Circuits.