A Symbol Based Watermarking Approach for Spread Spectrum Audio Watermarking Methods

Bala Mallikarjunarao Garlapati, Srinivasa Rao Chalamala
TCS Innovation Labs, Hyderabad, India
balamallikarjuna.g@tcs.com, chalamala.srao@tcs.com

***Abstract*— This paper proposes an high embedding capacity symbol based audio watermarking process. The traditional watermark embedding processes require longer duration audio sequence in order to accommodate all the input watermark bits. Shorter duration audio sequences which are like advertisements in Television & Radio should allow the embedding of a complete watermark at least once to be able to correctly identify the owner or content. A symbol based approach for spread spectrum based audio watermark methods has been proposed in order to reduce the required size of the audio sequences for watermarking. Each group of watermark bits called as a symbol are mapped to a unique pseudo-random sequence. These PN sequence are used for actual embedding in Spread Spectrum based watermarking technique. This process while reducing the size of the audio sequence required, it also helps in reducing the errors in longer duration audio as it can allow redundant addition of watermark. An error correction method is also added to further reduce the symbol errors that are caused due to the intentional and unintentional watermark removal attacks.**

Keywords*– *audio watermarking; Log Coordinate Mapping; Symbol based; high embedding capacity;

I. INTRODUCTION

In a digital world it is important to ensure the integrity and copyright protection of the audio data. Audio watermarking is one such process which can be used for these purposes. In this process a unique bit pattern is added imperceptibly to be extracted later. The extracted bit pattern is used to identify the right owner. In addition, watermarking can also be used for fingerprinting, transaction tracking, content authentication, secure control using voice commands etc.

It is imperative that the watermark added to the audio must sustain several manipulations that can be done by a malicious user. If a given watermark is added repeatedly it will help in correctly identifying the owner. Even a single bit error may result in identifying a wrong user.

A new embedding process using symbols instead of bit is proposed in order to fit the scheme for small duration audio sequences in this paper. This requires a very short audio for embedding entire watermark sequece resulting in redundancy and reduced errors.

II. RELATED WORK

There has been a lot of research going on audio watermarking. But these works those are aiming to robust against geometric variations are limited. An application like second screen app requires smaller duration audio sequences for advertisements in Television and the watermark method is robust to DA-AD attacks. However the concept "log-polar mapping" was used [2] and [3] in this paper, we user a symbol to pseudo noise sequence mapping table to the watermark sequence to improve the embedding capacity and used different embedding frequency regions. These frequency regions are identified according to our application need.

A watermarking method by modifying the energy relation of three adjacent DWT coefficient sections introduced in [1]. This method is robust against TSM (Time Scale Modification) and wave magnitude distortion. Haitsma et al. [4] proposed an algorithm by slightly modifying the magnitudes of the Fourier coefficients. It was reported to be robust to 4% of pitch invariant TSM, MP3 (64 kbps), and echo addition. But it is not robust against pitch shifting.

Xiangui Kang et al [2] proposed geometric invariant audio watermarking based on an LCM feature. This is resilient against common signal processing operations, including low-pass filtering, MP3 compression, echo addition, volume change, normalization, test functions in the Stirmark benchmark, and DA/AD conversion, but also has conquered the challenging audio geometric distortion and achieves the best robustness against simultaneous geometric distortions, such as pitch invariant time-scale modification (TSM) by 20%, tempo invariant pitch shifting by 20%, re-sample TSM with scaling factors between 75% and 140%, and random cropping by 95%.

An overview and basics about audio watermarking methods studied at [5] and singular value decomposition based audio watermarking techniques explained in [6]-[9]. A blind chirp based watermarking method for tamper detection proposed by Omar Farooq [10]. The watermarking methods with different signal processing techniques have been explained in [11] and[12], which gives overview of the today's audio watermarking technology. A method which is robust to DA-AD conversion explained in [13]. Xing He. [14] proposed a spread spectrum based scheme and is robust to DA-AD conversion with line-in jack.

2166-0670/16 $31.00 © 2016 IEEE
DOI 10.1109/ISMS.2016.31

IEEE computer society

A. Background

An audio watermarking method has to be robust to several intentional and intentional watermark removal attacks. Some of the attacks are application specific. Log coordinate mapping (LCM) feature based audio watermarking method [2] has good performance against geometric variations in the signal and robust to time stretching and pitch scaling in the audio signal. These attacks are inherent when a given audio is played or recorded.

1) Log mapping feature: It is known that scaling in time domain is equivalent to inverse scaling in frequency domain. This scaling when observed in logarithmic domain results in addition. So the time scale distortion is represented as an addition in spectral magnitudes. Using these properties, a log mapping feature is defined. We apply log coordinate transform on predefined frequency regions.

Scaling in time domain is equivalent to inverse scaling in frequency domain,

$$f(qt) < -- > 1/qF(f/q) \quad (1)$$

In frequency domain, geometric distortions described as

$$f' = \lambda * f \quad (2)$$

apply logarithm on both sides, then

$$log(f') = log(\lambda) + log(f) \quad (3)$$

We can observe that the time domain scaling converted into logarithmic frequencies addition.

In this log mapping feature, AFM (Average Frequency Magnitude) of a frequency band is defined and to compute this, convert discrete time signal into frequency domain. By defining normalized frequency index, select a frequency region and divide the region into several frequency bands. Each frequency band can represent one watermark bit or symbol. The embedding capacity is equal to the number of bands used in the algorithm and this is depends on application requirements. If more watermark bits need to be embedded for an application, then the samples available to the each AFM band are low and robustness may decrease. Unlike the previous literature we map a PN sequence to each symbol of the watermark and it is add to the frequency coefficients in one AFM band. Similarly different PN sequences corresponding watermark symbols added in all the AFM bands of the selected frequency region.

2) Symbol conversion of watermark bit sequence: Watermark embedding based on single bit i.e. adding a PN sequence to the spectral components in the AFM band requires longer audio sequence in order to accommodate all the input watermark bits at least one time as bit '1' mapped to one PN squence and bit '0' represents another PN sequnce. So that this embedding scheme is not suitable for small audio sequences. Hence a new embedding scheme is required for small audio sequences. A Symbol based embedding process can help when very short audio sequece is available for watermaking, it also helps in reducing the errors in longer duration audio due to repetitive adding of the symbols, which considers all the resultant symbols over all segments after watermark detection. Where, symbol is equivalent to group of input bits. For example, in table I, each set of two binary bits represented by one symbol. Four kinds of such symbols can be used to represent the whole watermark sequence. Similarly each set of three binary bits represented by one symbol and eight different symbols will use to represent the watermark sequence. Now a watermark added based on input symbol instead of single bit. Map each group of watermark bits to a unique PN sequence and then use this PN sequence to perform embedding in AFM coefficients. The symbols are mapped to a 32 bit PN sequence in this paper. This process of using symbol mapped PN sequence embedding doesn't alter the robustness of the original watermarking method, hence it can be used with any spread spectrum based watermarking scheme.

TABLE I
4 SYMBOL MAPPING TABLE

Bit 1	Bit 2	Symbol
0	0	1
0	1	2
1	0	3
1	1	4

During watermark extraction it is required to correlate the all the PN sequences with each audio segments to find which of the PN sequcen results in better correlation. Since it involves several correlations for extracting each symbol or group of watermark bits, it is time consuming. But this method is very useful for small duration audio sequences especially for televison advertisements and radio broadcastings. We can insert double the amount of watermark information in case of two set binary bits, and triple amount of watermark information in case of three set binary bits using this process. We tested here with 16 symbol mapping and achieved 4 times the actual embedding capacity. For example if an watermarking algorithm embedding capacity is 60 bits per sec, we can embed 240 (4*60) bits per sec, hence we increased 4 times the actual embedding capacity. The same technique can be applicable for video watermarking methods also.

We chosen audio watermarking algorithm based on an LCM feature as it is a spread spectrum based method. This method shows better performance for geometric variations in audio signal. This is combined with symbol based embedding method to improve the embedding capacity in audio data. This paper is organized into different sections. In section III we gave detail description about watermark embedding and extraction procedure and also explained the implementation steps involved for this method. Section IV

TABLE II
8 SYMBOL MAPPING TABLE

Bit 1	Bit 2	Bit 3	Symbol
0	0	0	1
0	0	1	2
0	1	0	3
0	1	1	4
1	0	0	5
1	0	1	6
1	1	0	7
1	1	1	8

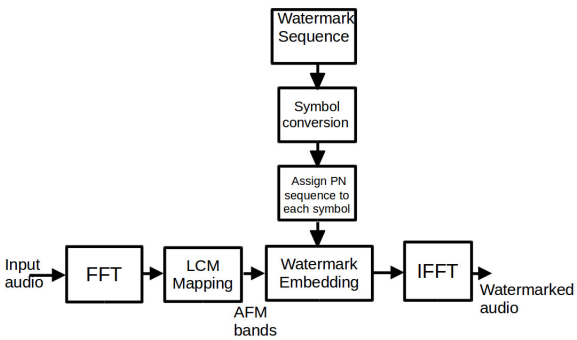

Figure 1. Watermark embedding method

provides the experimental results i.e. symbol detection rate (SDR) for different scale factors and for different signal processing attacks. Followed by conclusion in Section V.

III. IMPLEMENTATION

A. *Watermark Embedding Process*

Watermark embedding process block diagram shown in Fig. 1. Watermark embedding process adds an unique ownership information into a audio signal. This information is a series of PN sequences and add them as watermark symbols. The input audio signal is segmented into fixed size of frames with a duration of one second each. Frame size selection depends on the application and hardware system capability. The watermark symbols are mapped to a set of 32 bit PN sequences.

Select a frequency region using frequency normalization index and convert each input frame into frequency domain. Compute AFM bands using LCM feature and it's size selection depends on application requirements. Embed each PN sequence corresponding to symbols in wateramrk into AFM band coefficients, with predefined strength factor. Apply inverse DFT to get watermarked frame. The process is be repeated for all the watermark symbols and concatenating all these frames results in watermarked audio.

B. *Watermark Extraction Process*

Watermark extraction process block diagram shown in Fig. 2. As this is blind watermarking scheme, we don't need original audio. The watermarked and/or attacked audio signal is segmented into fixed size frames and converted into frequency domain as done during embedding process. Select the same frequency region, which is used in embedding process.

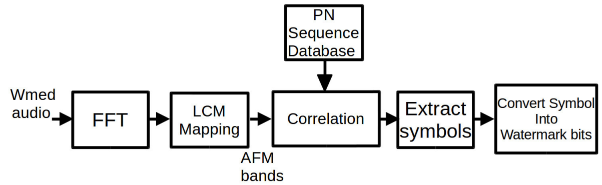

Figure 2. watermark extraction process

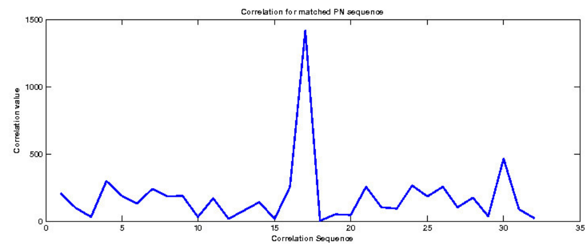

Figure 3. Correlation for matched PN sequence

After computing the AFM bands using LCM feature, these are correlated with the each of PN sequences (which is used in watermark embedding method), till a match (high correlation) occurs. Using correlation peak magnitude and it's side lobe ratio's, decide which PN sequence matched well, and hence the watermark symbol. The correlation output for a matched PN sequence and unmatched PN sequence illustrated in Fig. 3 and Fig. 4 respectively. This extracted watermark proves the ownership of the content.

C. *Steps involved in Implementation process*

Watermark symbol generation:

1) Create a watermark sequence binary bits with ones and zeros (which is ownership details).
2) Convert a set of binary bits into a symbols. For example refer table I and table II.
3) Generate PN sequence data for the entire watermark by assigning a unique PN sequences to each symbol.

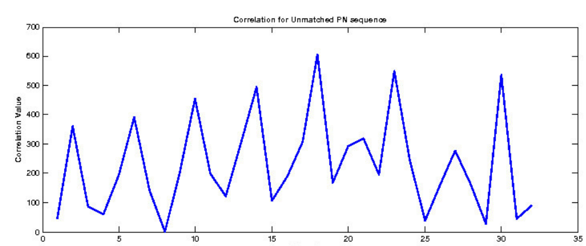

Figure 4. Correlation for unmatched PN sequence

Implementation steps involved in watermark embedding:

1) Define embedding capacity and embedding strength factor parameters, according to our application need.
2) Segment the input audio signal into (1 second) fixed size frames.
3) Apply Fourier transform on each frame.
4) Using LCM feature, compute Average Frequency Magnitude bands.
5) Watermark symbol PN sequence embedded with pre-defined scale factor, into the AFM frequency band coefficients.
6) Perform inverse DFT and concatenate all the watermarked frames.
7) Repeat the steps 2 to 6, for whole audio file.

Implementation steps involved in watermark extraction:

1) Define embedding capacity and embedding strength factor parameters, according to our application need.
2) Segment the input audio signal into (1 second) fixed size frames.
3) Apply Fourier transform on each frame.
4) Using LCM feature, compute Average Frequency Magnitude bands.
5) Correlate this magnitude coefficient array with each symbol PN sequence and detect the watermark symbol.
6) Repeat the steps 2 to 5 and extract the watermark symbols the entire frequency region.
7) Convert matched PN sequences to symbols and hence determine the watermark sequence and find SDR (symbol detection rate).

IV. Experimental Results

In this section, we presented some experimental results for the proposed algorithm. This method tested for different audio files and a speech file. The results are shown here for an audio file, having duration of 17 seconds. We measured the accuracy of the proposed algorithm using Symbol Detection Rate,

$$SDR = CDS/TS \tag{4}$$

Where CDS means Correctly Detected Symbols that are extracted from watermarked audio and TS is the Total number of Symbols embedded. Figure 3 shows correlation for matched PN sequence with magnitude coefficient array and observe the correlation peak value at center. Figure 4 shows correlation sequence for unmatched case and observe that there is no correlation peak. In this paper we given an example audio with sampling rate of 44100 and length of 749700 samples and used the watermark with 1020 symbols(2040 watermark bits).

Initially, we tested the algorithm for different embedding capacities with different embedding strength factors. Table III depicts the number of symbols in error(E_s) and the symbol detection rates against various scaling factors. We can observe by increasing the scaling (embedding strength) factor, improved the robustness but at the cost of little audibility of watermark. The SNR calculated for watermarked audio and is less than 20 db as per the standards.

After that we tested for signal processing attacks, sampling conversion and compression attacks. The algorithm showed approved performance and the results are showed in Table IV, V & VI.

TABLE III
SYMBOL DETECTION RATE (SDR) ANALYSIS OF THE PROPOSED ALGORITHM FOR DIFFERENT EMBEDDING STRENGTH FACTORS

Total no.of symbols embedded :1020	Embedding Capacity:10 symbols per sec			Embedding Capacity:20 symbols per sec		
Embed strength factor	E_s	SDR without ECC	SDR with ECC	E_s	SDR without ECC	SDR with ECC
0.5	14	98.63	100	24	97.6	100
0.8	14	98.63	100	24	97.6	100
1.0	8	99.22	100	24	97.6	100
1.2	8	99.22	100	4	99.4	100
1.5	3	99.7	100	4	99.4	100
2.0	3	99.7	100	4	99.4	100

TABLE IV
SYMBOL DETECTION RATE (SDR) ANALYSIS OF THE PROPOSED ALGORITHM FOR MP3 COMPRESSION ATTACK

Embed capacity :10 symbols per sec	MP3 Bitrate: 32 kbps		MP3 Bitrate: 64 kbps		MP3 Bitrate: 128 kbps	
Embed strength factor	SDR without ECC	SDR with ECC	SDR without ECC	SDR with ECC	SDR without ECC	SDR with ECC
0.5	89.7	100	98.6	100	99.7	100
0.8	89.7	100	98.6	100	99.7	100
1.0	97.8	100	100	100	100	100
1.2	97.8	100	100	100	100	100

TABLE V
SYMBOL DETECTION RATE (SDR) ANALYSIS OF THE PROPOSED ALGORITHM FOR AAC COMPRESSION ATTACK

Embed capacity :10 symbols per sec	AAC Bitrate: 32 kbps		AAC Bitrate: 64 kbps		AAC Bitrate: 128 kbps	
Embed strength factor	SDR without ECC	SDR with ECC	SDR without ECC	SDR with ECC	SDR without ECC	SDR with ECC
0.5	83.9	100	83.9	100	99.7	100
0.8	89.7	100	92.5	100	99.7	100
1.0	89.7	100	98.7	100	100	100
1.2	91.7	99.4	98.7	100	100	100

TABLE VI
SYMBOL DETECTION RATE (SDR) ANALYSIS OF THE PROPOSED ALGORITHM FOR SAMPLING CONVERSION ATTACK

Embedding capacity :10 symbols per sec	Up Sampling Conversion: 44.1 to 88.2 (in KHz)		Down Sampling Conversion: 44.1 to 22.05 (in KHz)	
Embed strength factor	SDR without ECC	SDR with ECC	SDR without ECC	SDR with ECC
0.5	98.7	100	91.7	100
0.8	99.4	100	92.5	100
1.0	99.4	100	98.7	100
1.2	100	100	100	100

V. CONCLUSION

In this paper, we proposed a symbol based audio watermarking process using log coordinate mapping feature, which is robust to geometric audio distortions. The simulation results proved that this process improves the embedding capacity. We can increase the number of symbols and hence the embedding capacity at the cost of more computational complexity which reduces the required audio length to embed whole watermark sequence.

This proces can also reduce the errors in longer duration audio due to repetitive adding of the symbols, by considering results from multiple segments to produce final result. The errors in symbol detection can further be improved by using error correction codes.

REFERENCES

[1] Shijun Xiang; Audio watermarking robust against D/A and A/D conversions, EURASIP Journal on Advances in Signal Processing, 2011, Volume 2011.

[2] Xiangui Kang; Rui Yang; Jiwu Huang, Geometric Invariant Audio Watermarking Based on an LCM Feature, Multimedia, IEEE Transactions, April 2011.

[3] C. Y. Lin, M. Wu, J. Bloom, I. Cox, M. Miller, and Y. Lui, Rotation, scale, and translation resilient watermarking for images, IEEE Trans. Image Process., May 2001.

[4] J. Haitsma, T. Kalker, and F. Bruekers, Audio watermarking for monitoring and copy protection, in Proc. 8th ACM Multimedia Workshop, Los Angeles, CA, 2000.

[5] Xing He Watermarking in Audio, Key techniques and technologies , Published by Cambria Press (2008-01-28).

[6] Wang, Jian, Ron Healy, and Joe Timoney. ”A robust audio watermarking scheme based on reduced singular value decomposition and distortion removal.” Signal Processing 91.8 (2011).

[7] Lei, Bai Ying, Yann Soon, and Zhen Li. ”Blind and robust audio watermarking scheme based on SVDDCT.” Signal Processing 91.8 (2011).

[8] Khalid A. Darabkh,“Imperceptible and Robust DWT-SVD-Based Digital Audio Watermarking Algorithm“, Journal of Software Engineering and Applications, 2014.

[9] Kakkirala, Krishna Rao, Srinivasa Rao Chalamala, and G. Rao. ”DWT-SVD based blind audio watermarking scheme for copyright protection.” Audio, Language and Image Processing (ICALIP), 2014 International Conference on. IEEE, 2014.

[10] Omar Farooq, S. Datta, Jonathan Blackledge, Blind Tamper detection in Audio using Chirp based Robust Watermarking, WSEAS Trans. On Signal Processing, April 2008.

[11] Xiang-Yang Wang and Hong Zhao, ”A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT“, IEEE TRANSACTIONS ON SIGNAL PROCESSING, DECEMBER 2006.

[12] Hussain, Iqtadar. ”A novel approach of audio watermarking based on S-box transformation.” Mathematical and Computer Modelling 57.3 (2013).

[13] Garlapati, Bala Mallikarjunarao, and Krishna Rao Kakkirala. ”Malicious audio source detection using audio watermarking.” Multimedia and Broadcasting (APMediaCast), 2015 Asia Pacific Conference on. IEEE, 2015.

[14] Xing He and Michael S. Scordilis Efficiently Synchronized Spread-Spectrum Audio Watermarking with Improved Psychoacoustic Model, Research Letters in Signal Processing Volume 2008.