Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp74-79)

A Robust Audio Watermarking in Frequency Domain

ROBIN JOSE AUGUSTINE
Department of Computer Science and Engineering
College of Engineering
Chengannur,Kerala
India-686 121

V.P.DEVASSIA
Principal
College of Engineering
Chengannur,Kerala
India-689 121

Abstract: We present an efficient audio watermarking algorithm in the frequency domain. Watermark for embedding is generated by a random number generator. The host signal is subjected to wavelet decomposition and the watermark is embedded in the DCT coefficients corresponding to the lowest frequency sub-band. The whole functionality of the algorithm is dependent on the seed value of the sequence generator, which itself acts as the security password for both embedding and extraction. Embedding the watermark in the place value of 10^{-2} could result in an SNR of the order of 70-80db, achieving an almost distortion less watermarking. The length of the watermark is made variable and user selectable for achieving flexibility in terms of computational complexity, security level and perceptual quality. Results are presented for watermark embedding on different classes of audio signals for different levels of decomposition and by embedding at different place values.

Key–Words: Audio Watermarking, Random Sequence Generator, Digital Watermarking

1 Introduction

Recent developments in the field of communication has helped in fast transfer and perfect copying of multimedia data. The developments in these fields have raised serious security issues, related to intellectual property right, monitoring of broadcast signals, genuineness of the multimedia data transferred etc. Digital watermarking is identified as a partial solution to related problems which allow content creator to embed hidden data such as author or copyright information into the multimedia data [1].The embedded watermark may contain information such as signature, logo, id etc which are uniquely related to the author, distributor, and/or the data being transmitted itself which helps multimedia data owner in proving his/her ownership thereby restricting others from copying the data. Within past few years several algorithms for embedding and extraction of watermark in audio sequence have been published [2-6]. Almost all audio watermarking algorithms work by exploiting the perceptual property of Human Auditory System (HAS). Some of the techniques used for embedding watermark in audio signal are *echo coding, phase coding, direct sequence and frequency hopped spread spectrum technique* etc. Echo coding works by encoding watermark as echo with an imperceptible delay period while phase coding algorithm replaces the short term phase of an audio signal with its signature [4]. Direct Sequence and Frequency Hopped Spread Spectrum (DSSS and FHSS) techniques spread watermark data using a bipolar pseudorandom sequence in the spatial or Discrete Cosine Transform (DCT) domain respectively [7-8]. Simplest visualization of the requirements of information hiding in digital audio is possible *via* a *magic triangle* [9] as given in Fig 1. In order to satisfy the requirements of *magic triangle*, watermarks are seen embedded in fourier domain [2], time domain [3], sub-band domain [5], wavelet domain [6] and by echo hiding [4].

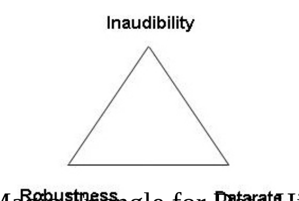

Fig 1: Magic Triangle for Data Hiding

In this paper we present a novel audio watermarking algorithm in which wavelet transform is used to divide the audio signal into low and high frequency sub-bands, and watermark is embedded in the lowest frequency sub-band in fourier domain. The uniqueness of this algorithm is that the entire functionality of the algorithm is dependant on the security password used. The computational complexity can be varied by varying the length of the watermark to be embedded, thereby achieving a flexible level of security also. The watermark is embedded by suitably modifying the DCT coefficients of the lowest frequency sub-band. In the case of extraction, the modified DCT coefficients are analyzed for detecting the presence of watermark in the signal under consideration, if any.

The paper is organized as follows. Section 2 presents outline of the new watermark embedding algorithm. In section 3 the watermark extraction algorithm is explained in detail. Section 4 gives the experimental results of the algorithm.

2 Watermark Embedding

The Proposed watermark embedding algorithm makes use of two transformations *viz* Discrete Wavelet Transform (DWT) for sub-band coding [10] and DCT for transformation into frequency domain. The scheme of watermark embedding is depicted in Fig 2.

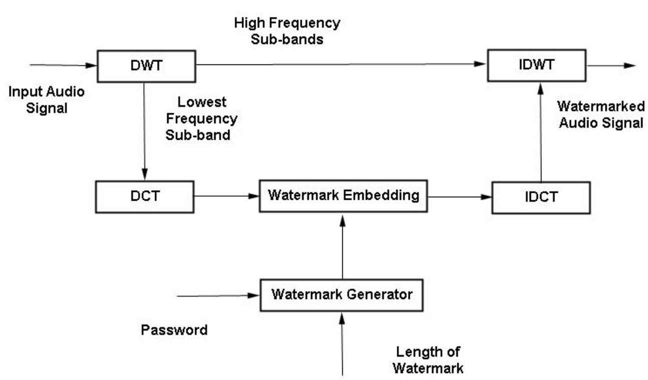

Fig 2:Watermark Embedding Scheme

2.1 Watermark Generation

Watermark is generated by using an arbitrary sequence generator [11] which acts based on two inputs. One of the inputs is the required length of the watermark signal which is made user selectable to achieve variable level of security. This gives the user an option to settle on the computational complexity depending on the use. The second input acts as the seed value for the arbitrary sequence generator which is nothing but the password itself. Alternately the watermark length can be encoded in the password itself rather than providing as a separate input. The distinct feature of this algorithm is that the password helps in providing dual protection; in addition to being the user login password for checking the correctness of the extracted watermark signal, the exact password is also needed to uniquely generate the sequence. Even though the generated sequence appears to be random, same arbitrary sequence can be generated any time by giving the same seed value. The details of the algorithm for arbitrary sequence generator is given in Fig 3.

```
Input  : Values specifying the dimension of the sequence and
seed value
Output: Arbitrary sequence with specified dimension

Function random (m, n, seed)
i =1
j=1
Begin
   Begin
      Seed = floor (seed);
      Seed = mod (seed, 2147483647);
      If (seed < 0)
          Seed = seed + 2147483647;
      End if
      k = floor (seed / 127773);
      Seed = 16807 * (seed - k * 127773) - k * 2836;
      If (seed < 0)
          seed_ = seed + 2147483647;
      End
      R (i, j) = (seed * 4.656612875E-10) +1;
    End until i=m
 End until j=n
 Return R
```

Fig 3 :Algorithm for arbitrary sequence generator

2.2 Embedding Procedure

The host audio signals are subjected to wavelet decomposition, the level of decomposition being dependent on the type of audio signal and the sampling frequency. As a trade off between robustness, computational complexity and extraction accuracy, a level 4 wavelet-decomposition is preferred. In order to improve the robustness of the method, the lowest frequency sub-band is chosen for watermark embedding. If the watermark is embedded in the high frequency sub-bands, the MPEG1 Audio Layer (MP3) compression can remove high frequency components from the input audio signal and can cause distortions in watermark signal. Moreover, the HAS is more sensitive to the change of high frequency sound when compared to the rest [12].The lowest frequency sub-band is now transformed into frequency domain by performing DCT on it. Watermark is embedded at the beginning of the signal itself by altering the 2^{nd} decimal place of the DCT coefficients. The embedding function can be represented as given in Fig 4.

When the n^{th} bit of the watermark is *one*, the 2^{nd} decimal place of the n^{th} DCT coefficient is changed to the higher odd number possible and when the watermark bit is *zero*, the coefficient is changed to the next possible even number. The watermark embed-

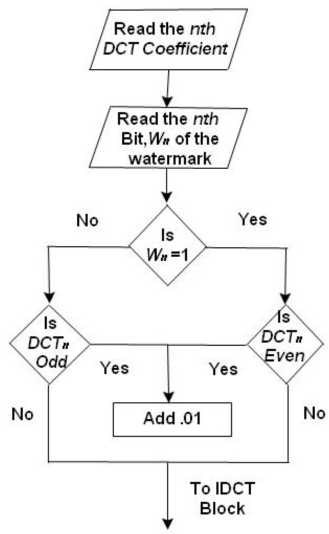

Fig 4 :Watermark Embedding Function

ded audio signal is constructed by performing the inverse operation, *viz* Inverse Discrete Cosine Transform (IDCT) and Inverse Discrete Wavelet Transform (IDWT).

3 Watermark Extraction

In the extraction phase the watermark signature is extracted from the watermarked audio signal. In order to check the genuineness of the audio signal under consideration, we need the *seed* value used for the arbitrary sequence generator at the embedding phase as the *password*, in addition to being used as the *seed* value for the sequence generator at this end. Outline of the watermark extraction algorithm is given in Fig 5.

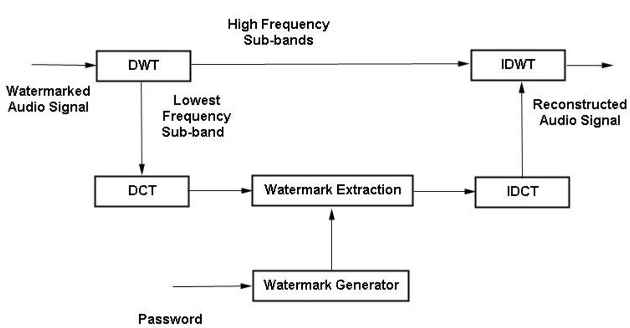

Fig 5 : Watermark Extraction Scheme

The audio signal is subjected to sub-band decomposition as done in the embedding phase and the DCT of the lowest frequency sub-band is taken. These coefficients are input to the watermark extraction unit to extract the watermark, if any. When the DCT coefficient is *even* the corresponding bit in watermark is detected as *Zero*, otherwise as *One*. This extracted bit is compared with the corresponding bit produced by the watermark generator. If both the generated and extracted sequences are same then the genuineness of test data is proved. The watermark extraction function can be represented as given in Fig 6. In order to fully recover the original audio signal, the modification introduced in the DCT coefficients are reversed. Subsequently the IDCT and IDWT are performed to reconstruct the signal devoid of any watermark.

4 Experimental results

Four classes of audio signals were used as host signals, for the study *viz* classical music, pop music, rock music and speech. These general classes were chosen because each class has different spectral properties. The spectral plot of typical segments from each class of the chosen audio signals is given in Fig 7. All the segments were amplitude normalized to ensure unambiguous visualization of the spectral characteristics. Each test signal was sampled at 44.1 kHz, represented by 16 bits resolution. The algorithm was implemented in *matlab ver 6.0* running on *Intel Pentium 4* machine at 1.8 GHz.

4.1 Perceptual Quality

In the present context the perceptual quality is taken as a measure of the ability of the watermarking algorithm to preserve the quality of the orginal audio even

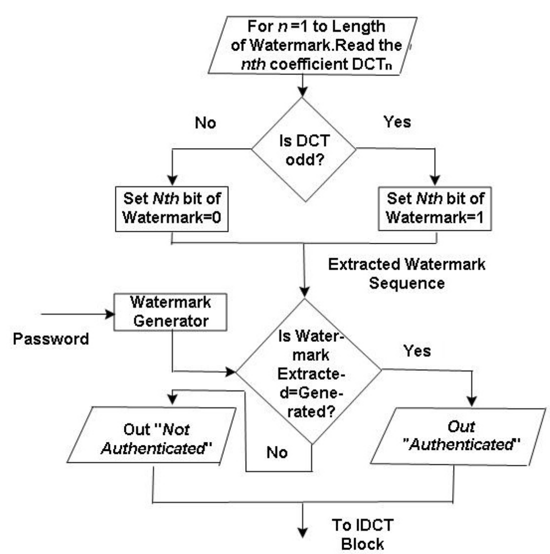

Fig 6 : Watermark Extraction Function

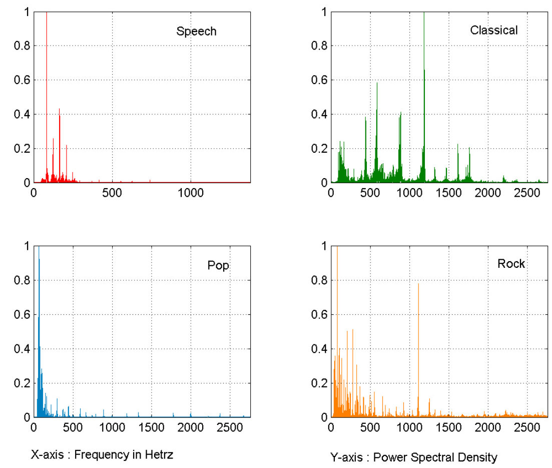

Fig 7 :Spectral plot of typical segment of the audio signals used for testing

after embedding. In most applications, it is important that the watermark is undetectable to a listener or viewer. This ensures that the quality of the host signal is not perceivably distorted, and does not indicate the presence or location of a watermark. In this study, the signal-to-noise ratio (SNR) of the watermarked signal in comparison with the original host signal was used as a quality measure. SNR in *db* is calculated using the follwing equation [1].

$$SNR = 10.log_{10}.\sum_{n=0}^{N-1}\frac{x^2(n)}{[x(n)-x'(n)]^2}$$

where,

N is the length of audio signal
*x(n)*is the orginal signal
x'(n) is the watermarked signal
n=0,1,2,3......*N*

Table 1 shows the SNR computed for different types of audio signal for different levels of decomposition and different place values of embedding. The results for a level 3 decomposition in respect of a segment from a female classical music are summarize in Fig 8. It is seen that as embedding position is changed from the most significant digit to the least significant digit, the value of SNR increases. The same result was observed irrespective of the class of signal used. The change of SNR is seen to be almost linear with respect to the embedding place value. Fig 9 shows similar results for variation in the level of sub-band decomposition of an audio signal. For practical signals, it is observed that the level of decomposition dosen't have significance on the embedding quality.

Table 1:Distortion introduced in different class of audio signal due to watermark embedding

(a)place value of embedding 10^0

Class of Audio Signal	Classical	Pop	Rock	Speech
Level	Signal to Noise Ratio(db)			
3	35.0832	37.0209	36.9925	35.2352
4	35.0834	37.0194	36.9925	36.9962
5	35.0835	37.0194	36.9928	36.0273
6	35.0840	37.0197	36.9935	36.0276

(b)place value of embedding 10^{-1}

Class of Audio Signal	Classical	Pop	Rock	Speech
Level	Signal to Noise Ratio(db)			
3	55.0883	56.2250	56.1969	56.0242
4	57.2940	55.5551	57.9618	54.5657
5	56.0513	54.4608	55.5280	54.5640
6	54.2919	55.5531	55.5318	58.2419

(c)place value of embedding 10^{-2}

Class of Audio Signal	Classical	Pop	Rock	Speech
Level	Signal to Noise Ratio(db)			
3	75.9445	77.7364	76.0638	74.3344
4	76.9899	74.9130	74.8456	75.0527
5	75.7854	76.8809	73.9129	78.1364
6	73.5184	76.8087	73.8896	73.9136

4.2 Computational Complexity

The computational complexity in watermark embedding is an important consideration, for it may influence the choice of implementation structure or DSP architecture for the algorithm. Although there are many ways to measure complexity [6], for practical applications more real life values are required. In this study, actual CPU time of implementation is taken as a measure for comparison. The computational complexity of the algorithm for both embedding and extraction for a watermark of length 12 is given in table 2.

5 Conclusion

A novel audio watermarking method in frequency domain was described. Embedding and extraction of watermark is done in low frequency sub-band of the audio signal as human ear is less sensitive to changes in low frequency bands making it more resistant to attack from mp3 compression which removes high

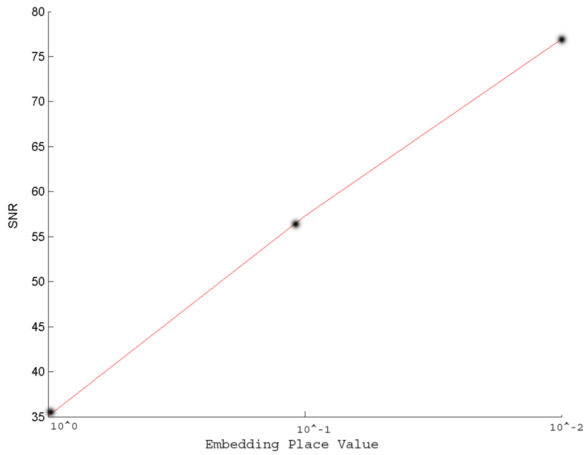

Fig 8 : Effect of changing place value of embedding on the perceptual quality of the signal

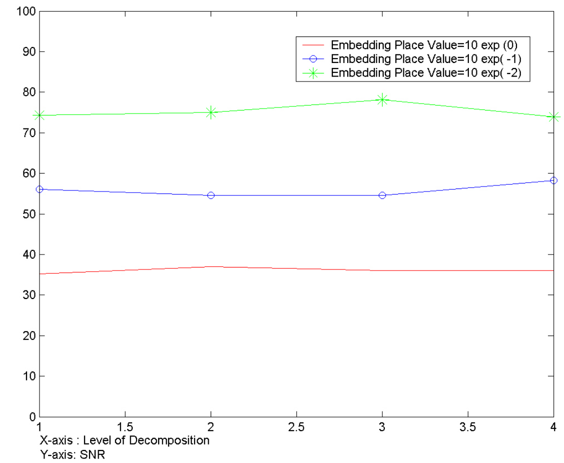

Fig 9 :Change of audio perceptual quality against level of sub-band decomposition

Table 2: Computational complexity

(a)For Watermark Embedding

Class of Audio Signal	Classical	Pop	Rock	Speech
Level of Decomposition	CPU Time in Sec			
3	2.0940	2.4060	3.5620	4.3430
4	2.7340	4.0160	2.7190	3.9840
5	3.3590	2.9070	4.7500	3.1880
6	7.5460	2.5000	4.0940	4.3910

(b)For Watermark Extraction

Class of Audio Signal	Classical	Pop	Rock	Speech
Level of Decomposition	CPU Time in Sec			
3	2.1090	2.4220	3.5160	4.3290
4	2.7340	4.0150	2.7030	3.9530
5	3.3120	2.9370	4.7500	3.2030
6	7.5780	2.5000	4.1100	4.3750

frequency sub-bands during conversion. Objective evaluation of the quality was done in terms of SNR. Study was conducted on different class of audio signals. Results show that the method is superior over state-of-the-art methods.

References:

[1] L.T.Bruton, J.D.Gordy, *Performance Evaluation of Digital Audio Watermarking Algorithms*, Proceedings of the 43rd IEEE Midwest Symposium, Michigan, Volume1, pp.456-459, 8-11 August 2000.

[2] M.D.Swanson, B.Zhu, A.H.Tewfik, *Robust audio watermarking using perceptual masking* , Signal processing, vol.66, pp.337-355, 1998.

[3] P.Bassia, I.Pitas, *Robust audio watermarking in time domain*, IEEE Trans. on Multimedia, vol.3, No.2, pp.232-241, June 2001.

[4] W.Bender, D.Grul, N.Morimoto, A.Lu, *Techniques for data hiding*, IBM System Journal, vol.35, pp.313-336, 1996.

[5] Xin Li, Hong Heather Yu, *Transparent and Robust Audio Data Hiding in Sub-band domain*, Proc. International Conference on Information Technology: coding and computing, Las Vegas, pp.74-79, March 27-29, 2000.

[6] M.F.Mansour, A.W.Tewfik, *Audio watermarking by time scale modification*, Proc International Conference on Acoustic, Speech, and Signal processing, USA, vol.3, pp.1353-1356,May 7-11, 2001.

[7] R. Pickholtz, D. Schilling, and L. Milstein, *Theory of spread-spectrum communications - a tutorial*, IEEE Trans. on Communications, vol.30, no.5, pp.855 - 884, May 1982.

[8] I.Cox, J.Kilian, T.Leighton, and T.Shamoon, *Secure spread spectrum watermarking for multimedia*, IEEE Trans. on Image Processing, vol.6, no.12, pp.1673 - 1687, December 1997.

[9]] Nedeljko Cvejic, Tapio Seppnen, *Watermark Bit Rate in Diverse Signal Domains*,

International Journal of Signal Processing, Vol.1, 2004.

[10] Martin Vetterli, Jelena Kovacevic, *Wavelets and Sub-band Coding*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[11] Dwight F. Mix, *Random Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[12] T.Furukawa, K.Konishi, S.Saito, *A digital watermarking for audio data using band division based on QMF bank*, IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002 [Online], Vol.4, pp.3473-3476, 2002.