Schemes for Evaluating Signal Processing Properties of Audio Watermarking

T.C Thanuja , Dr. R.Nagaraj

Summary
Watermarking audio files has recently become the focus of much attention. This is primarily due to faster data transmission rates on the Internet, which has allowed illegal usage of digital audio files. Watermarking may give recording companies the ability to enforce copyright protection of their products. The requirements of watermarking audio lie in preserving the file quality (imperceptibility) and remain intact after a number of file damaging operations (robustness). The main challenge in digital audio watermarking is to achieve the right tradeoff between the mutually exclusive goals of robustness and high watermark data rate. This paper gives a performance evaluation of popular audio watermarking schemes in prevalence today. A system simulation of selected schemes has been performed in MAT LAB.

Key words:
Intellectual property, audio watermarking, digital rights management.

1. Introduction

Digital watermarking is a technique of embedding a digital signal or pattern on a digital document. The digital document may be text, audio, image or video. When the digital document is in the form of an audio signal, the embedding technique is called audio watermarking. There are various purposes for audio watermarking. The original intention of watermarking is for copyright protection [1]. Therefore, the most obvious purposes are the needs for proof of ownership and the enforcement of usage policy [2]. In addition, watermarking can also be used for fingerprinting and augmenting media with additional features.

The digital media that carries the watermark is called a cover signal or host signal. The watermark is embedded into the host signal by a watermark embedder and is detected by a watermark detector. A watermark key prevents unauthorized watermark embedding and watermark detection. A watermarking scheme is said to employ Informed Detection if it requires the original host signal to be present at the watermark detector. Conversely, a watermarking scheme is said to employ Blind Detection if it does not require the original host signal to be present at the watermark detector.

According to the intention and the kind of watermark, watermarking techniques should possess certain properties [3].

Signal Processing Properties:

The watermark should be imperceptible to an observer. Differences between the original signal and the watermarked signal should be negligible to the human ear. The watermark should be robust against intentional or anticipated manipulations, e.g. compression, filtering, resampling, requantisation, cropping, scaling, etc. Error correction coding should be used to ensure data integrity. The watermark bit rate must be of a suitable value so as to embed the watermark depending on the application

2. Audio Watermarking Schemes

Several algorithms have been developed for the purpose of watermarking. Each algorithm achieves a certain tradeoff between robustness and watermark data rate for a given perceptual transparency. The choice of the algorithm depends on several factors, most important of which are, the type of cover audio, the computational complexity of the algorithm and the application, which defines the degree of robustness required [4].

Detail descriptions of some of the popular audio watermarking schemes and the steps involved in implementing are discussed.

2.1 Substitution Techniques

Redundant or non significant parts of the cover audio are substituted with the watermark message. Ex. Least significant Bit (LSB) Substitution [5] [6].

The sequences of steps implemented are as follows:

1) If watermark size exceeds number of available samples, an error message is displayed and the function exits.
2) The watermarked file is initially generated as a copy of the original audio.
3) The algorithm uses the random number generator to index into audio samples in random fashion.
4) The LSB of the sample is replaced with the watermark bit.

Manuscript received July 5, 2008.
Manuscript revised July 20, 2008.

2.2 Transform Domain Techniques

Watermark message is embedded in a Transform domain, optionally employing a psycho acoustic model to improve robustness and imperceptibility [7]. Here we perform watermarking in the Discrete Cosine Transform (DCT) domain [8] [9].

1) Split the cover audio into blocks. Each block is used to encode n message bits.
2) Blocks are chosen in a pseudorandom manner.
3) A DCT of the frame is obtained. Let v (i) represent the DCT coefficients.
4) The largest(in terms of absolute value) n DCT samples are modified using the formula

$$v_1(i) = v(i)(1 + \alpha w(i)) \quad \textbf{(1)}$$

   Where, 'α' is a scaling factor. Here a value of 0.3 is used for 'α'. w (i) is the watermark bit (0 or 1).
5) The inverse DCT is computed and samples are written back to the file.

This algorithm is an example of informed decoding. During decoding, $v_1(i)$ is read in from the watermarked sample while v(i) is read in from the original sample. By comparing their absolute values we decode the watermark bit.

2.3 Echo Hiding

Embeds a watermark by introducing an imperceptible echo into the host signal. The encoding involves addition of echoes and hence the audio signal is convolved with a series of impulses, which decay exponentially and, spaced at intervals, which are predetermined [10]. Here only one echo is added for encoding either bit '1' or bit '0'. So the delay between the original and echo signal is varied. These signals obtained are multiplied with corresponding mixer signals to obtain the encoded signal.

The decoding involves detection of spacing between the echoes. This is achieved by finding the autocorrelation of the encoded signal cepstrum. This results in an unwanted peak at time t_0, which is discarded. The magnitudes are compared only at the respective delay values. If there is a peak at delay δ_1 it is decoded as bit '1' else if there is a peak at δ_0 it is decoded as bit '0'.

To enable secure decoding, the program accepts a password from the user. This number acts as a seed to a random number generator whose output determines the relative ordering of frames to encode.

2.4 Phase Coding

The Phase Coding method works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments [11]. According to the algorithm presented in Bender et. Al [9], the watermark data is phase encoded as $\pi/2$ or- $\pi/2$ depending on the watermark bit 0 or 1, respectively. However, during the tests performed, the following difficulties were encountered,

1) The watermarked signal has a large noise component due to phase dispersion. Phase dispersion is a distortion caused by a break in the relationship of the phases between each of the frequency components.
2) The use of $\pi/2$ or $-\pi/2$ for Phase Coding destroys the magnitude component of the signal and hence the watermark cannot be decoded (since $\cos(\pi/2) = 0$). The use of other values such as $\pi/4$ or $-\pi/4$ would also not facilitate the watermark recovery since cosine of a negative angle is still positive.

In order to overcome these difficulties, the following modifications were done to the original algorithm.

1) The selection of frames for phase modification was randomized rather than selecting consecutive frames. Hence, the effect of phase dispersion is spread throughout the audio signal. This reduces the noise perceptibility.
2) The values $\pi/4$ and $3\pi/4$ are used to perform the phase modifications for watermarking. The use of these values solves both the issues in encoding mentioned above.

For the decoding process, the synchronization of the sequence is done before the decoding. The length of the segment, the DFT points, and the data interval must be known at the decoder. The value of the underlying phase of the first segment is detected as a 0 or 1, which represents the coded binary string.

2.5 DC Level Shifting

Hide watermark data in lower frequency components of the audio signal, which are below the perceptual threshold of the human auditory system. This technique involves shifting the DC level for the input audio signal to negative and positive level according to the binary watermark sequence [12][13], Eq.(2, 3) illustrate the procedure of embedding watermark.

Level 0 = DC Bias Multiplier (Frame Power) **(2)**

Level 1 = +DC Bias Multiplier (Frame Power) **(3)**

Where,
Level 0 = value of the negative level
Level 1 = value of the positive level
DC Bias Multiplier = constant for DC bias multiplier.
Frame Power = frame power to the associated frames.
The audio signal is divided into several fixed-sized frames. In order to alter the DC component of a frame, the frame is processed in the following steps;

1) The Discrete Fourier Transform (DFT) is computed for each frame, x[n].The first element of the vector thus computed represents the DC component of the frame.
2) The mean and power content of each frame is calculated as,
   Frame mean = (1/N) x[n]
   Frame power = (1/N) (x[n]) 2
   Where N=Number of samples in each frame.
3) The first element of the frame vector obtained through DFT is modified to represent watermark bit as described above with DC Bias Multiplier = 100.
4) The Inverse Discrete Fourier Transform (IDFT) of the frame vector gives the modified frame.

These steps are performed until all the watermark bits are encoded.

For the decoding process, the watermarked audio signal is divided into equal sized frames with the frame size being equal to that used during encoding. For a given frame, the frame mean is calculated and the binary watermark sequence is decoded according to the sign of the frame means.

3. Test Procedure

Two important tests have been carried out. The cover audio under consideration for evaluation is mono or stereo, 8 bit or 16 bit per sample, sampled at any sampling rate.

3.1 Perceptual Transparency

To evaluate the perceptual transparency or the imperceptibility of a watermarking scheme, watermarked samples are subjected to listening tests. The procedure followed is given below:

1) The original sample was first identified and played to a group of listeners.
2) The watermarked samples were then played in random order.
3) The listeners were asked to evaluate each sample on a scale of 1 to 5 with 5 representing the quality of the original sample.
4) An Average Score was computed for each watermarking scheme.
5) The individuals involved in the listening tests included people from different musical backgrounds (classical, rock, pop etc).

3.2 Robustness Tests

Robustness tests were carried out using an open source utility called SOX(SOund eXchange). A SOX [14] is a command line utility, which offers the ability to perform several signal processing operations and effects on a wide variety of media formats. The watermarking schemes were subjected to the following robustness tests: Additive Random Noise, Volume Filtering, Low pass Filtering, Resampling, Lossy Compression

Additive Random Noise was simulated by generating a scaled random sequence using MATLAB and saving this sequence as a noise pattern. This noise pattern was added to the watermarked sample, which was then subjected to watermark decoding.

For volume scaling, we scaled the watermarked audio over a range of 0.5 to 1.5 times the original amplitude.

Low pass Filtering was carried out by subjecting the watermarked audio to filtering up to 12 KHz.

For Resampling, the watermarked audio was subjected to a cycle of down sampling and up sampling. The audio was down-sampled to half the original sampling rate and then up-sampled back to the original sampling rate.

For the Lossy Compression test, we subjected the watermarked audio to MP3 compression technique

The results of the Perceptual Transparency tests are given in Table 1, while those of the Robustness tests are given in Table 2.

Table 1: Perceptual Transparency (On scale of 5)

Algorithm	*LSB Coding*	*Echo Hiding*	*Phase Coding*	*DC Level Shifting*	*Transform Domain*
Score	4.9	4.8	4.7	4.8	4.8

Table 2: Robustness Tests

Algorithm	*LSB Coding*	*Echo Hiding*	*Phase Coding*	*DC Level Shifting*	*Transform Domain*
Additive Noise	X	✓	✓	✓	✓
Volume Scaling	X	✓	✓	✓	✓
Low Pass Filtering	X	✓	✓	✓	✓
Resampling	X	X	✓	X	✓
Lossy Compression	X	X	X	X	X

X: Indicates test fail
✓ : Indicates test pass

4. Simulation Results

Simulation has been carried out by creating a GUI in the MATLAB. This interface allows the user to select the type of encoding scheme, appropriate type of watermark (text message, text file or image file) and also path to the original and watermarked audio files. Fig. 1-5 shows the results of the encoding process. It provides the original audio, watermarked audio as well as the difference waveform for each of the schemes under consideration.

From Fig. 1-5, we can conclude that the Mean Square Error (MSE) in LSB encoding is minimum and more for Echo Hiding. This indicates that more data can be embedded using LSB technique. The order of these schemes for higher data capacity can be structured as LSB, DC level shifting, Transform Domain Coding, Phase Coding and Echo Hiding.

From the Table 1, it is observed that LSB coding provides better perceptual transparency than other techniques.

From Table 2, we found that the watermarked data does not changed after resampling of data in case of Transform domain coding and Phase Coding techniques. From the robustness test it is found that Phase Coding and Transform coding capable of withstanding various attacks. Out of all the techniques Transform and Phase Coding are less vulnerable to the attacks, where as LSB coding is more prone to attacks. All the methods are prone to lossy compression.

From the results, we can conclude that LSB coding gives good imperceptibility and high data embedding capacity but it fails in robustness test. Good robustness can be achieved with Phase Coding whereas it lacks in providing high data rate and perceptual transparency.

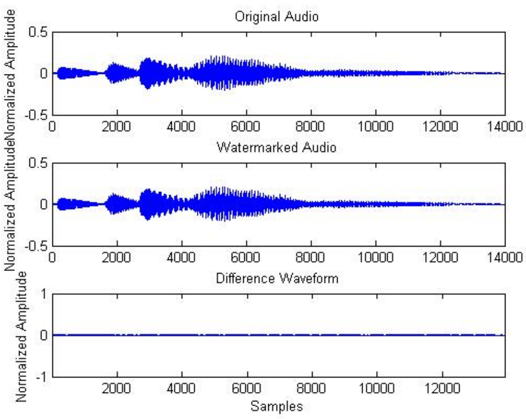


Fig 1. LSB coding

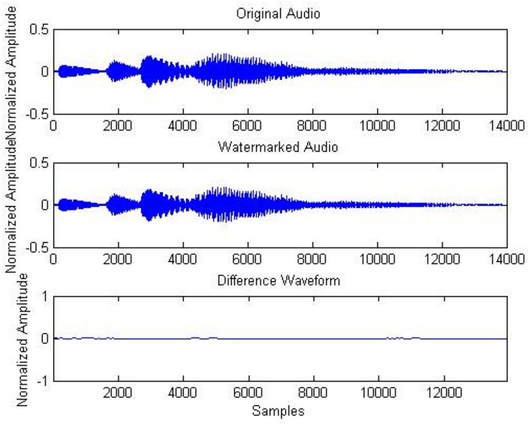


Fig 2. Transform domain coding

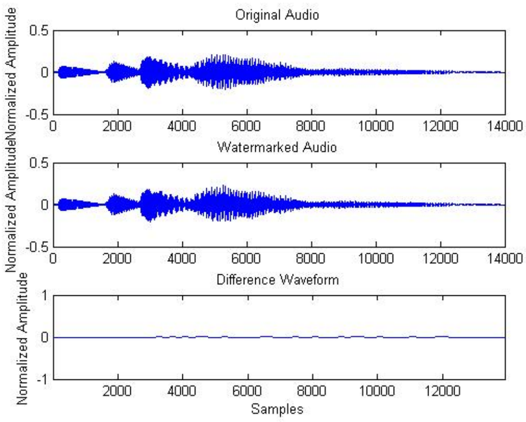


Fig 3. DC Level shifting

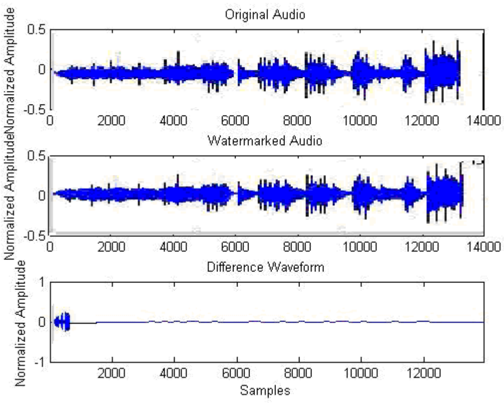


Fig 4. Phase Coding

5. Conclusion

In this paper, an overview of different properties of watermarking and several watermarking schemes, which were simulated on MATLAB, are given. These schemes are also subjected to a series of perceptual transparency and robustness tests. The performance of these schemes and the results of testing are as follows: LSB Coding works very well for a Fragile Watermarking scheme.

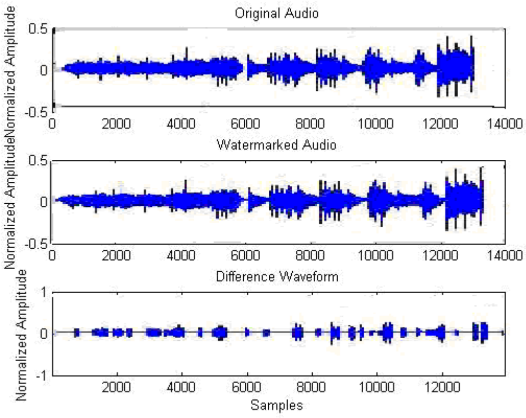


Fig 5. Echo Hiding

LSB Coding is also the least computationally intensive of all the schemes. Phase Coding and Transform Domain Coding gives the best performance with regard to robustness to common signal processing.

6. References

[1] Fabien A.P. Petitcolas Stefan Katzenbeisser, "Information Hiding Techniques for Steganography and Digital Watermarking", Artech House, Boston, 2000.

[2] Hyoung Joong Kim, Taehoon Kim, In-Kwon Yeo, "A robust audio watermarking scheme" Proceedings of the 2004 International Symposium on Circuits and Systems, 2004. ISCAS '04. Volume 5, 23-26 May 2004 Page(s):V-696 - V-699 Vol.5

[3] Yiqing Lin, Abdulla, W.H., "Perceptual evaluation of audio watermarking using objective quality measures" IEEE International Conference on Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. March 31 2008-April 4 2008 Page(s):1745 - 1748

[4] G. Cano Rodríguez, M. Nakano Miyatake, H. M. Pérez Meana, "Analysis of Audio Watermarking Schemes", 2nd International Conference on Electrical and Electronics Engineering (ICEEE) and XI Conference on Electrical Engineering (CIE 2005), September 7-9, 2005.

[5] Cvejic .N, Seppanen. T, "increasing robustness of LSB audio steganography using a novel embedding method", International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. Volume 2, 2004 Page(s):533 - 537 Vol.2

[6] Yin Xiong, Zhang Xiao ming, "Covert Communication Audio Watermarking Algorithm Based on LSB", International Conference on Communication Technology, 2006. ICCT '06. Nov. 2006 Page(s):1 – 4.

[7] A. Tefas, A.Giannoula, N.Nikolaidis, I.Pitas, "Enhanced transform-domain correlation-based audio watermarking" , IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). Volume 2, 18-23 March 2005 Page(s):ii/1049 - ii/1052 Vol. 2.

[8] Xing-Yang Wang and Hong Zhao, "A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT", IEEE Transactions on Signal Processing, Vol. 54, No. 12, December 2006.

[9] W. Bender, D. Gruhl and N. Morimoto. Techniques for data hiding. IBM Systems Journal, 35(3&4):313–336, 1996.

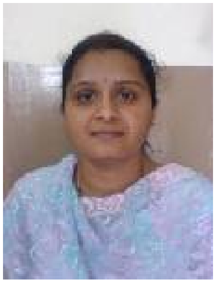
[10] Y.Erfani, M.Parviz, S.Ghanbari, "Improved time spread echo hiding method for robust and transparent audio watermarking", IEEE 15th Signal Processing and Communications Applications, 2007. SIU 2007. 11-13 June 2007 Page(s):1 – 4.

[11] Jongwon Seok, Jinwoo Hong, and Jinwoong Kim, "A Novel Audio Watermarking Algorithm for Copyright protection of Digital Audio", ETRI Journal, Volume 24, Number 3, June 2002.

[12] Umut Uludag, Dr. Levent M. Arslan, " Audio Watermarking Using DC Level Shifting", EE 683.01 Advanced Topics in Speech Processing Project Report January 2001.

[13] K.Herkiloglu, Y.Yaslan, S. Sener, B.Gunsel, "Robust audio watermarking by adaptive psychoacoustic masking", Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference, 2004. 28-30 April 2004 Page(s):29 – 32.

[14] K. Herkiloglu, Y. Yaslan, S. Sener, B. Gunsel, " Robust Audio Watermarking by Adaptive Psychoacoustic Masking", Proceedings of IEEE 2004.



T.C.Thanuja, is an Asst. Professor of Telecommunication Engineering Department, received the B.E. and M.E. degree in Electronics & Communication Engineering from Bangalore University Karnataka, India in 1995 and 2000 respectively. Currently she is perusing her Ph.D degree in Electronics and Communication Engineering, Avinashilingam university for women, Tamilnadu, India. Her research interests include digital watermarking, image processing and data security. She has guided more than 15 undergraduate and post graduate projects

Dr. R.Nagaraj, is the Director, R.V. Centre for Cognitive Technologies. He has obtained doctoral degree from the Visvesvaraya Technological University for a thesis on Automatic Flight Control. He obtained his Masters degree in Power Electronics from Gulbarga University and BE in Electronics from Bangalore University. His research interests are Aerospace Electronics, Industrial Electronics and DSP. He has guided more than 30 undergraduate projects. He is also guiding a couple of Ph.D scholars. He has several ongoing Research Projects under his supervision. Published several technical papers in International Refereed Journals & Conferences