Performance Analysis of Discrete Wavelet Transform based Audio Watermarking on Indian Classical Songs

C. M. Juli Janardhanan
Department of ECE
Government Engineering College, Wayanad
Mananthavady, Kerala

C. Sathish Kumar
Department of ECE
Government Engineering College, Wayanad
Mananthavady, Kerala

ABSTRACT

Digital audio watermarking involves the process of embedding into a host audio signal, a perceptually transparent digital signature carrying a message about the host signal. A technique of digital audio watermarking using wavelet transform is applied to watermark Indian classical songs. Investigations were performed using Haar, Daubechies and Symlet wavelets for time-frequency decomposition of the audio signal in order to embed watermark bits into the wavelet coefficients. Simulations are performed with embedding at different levels of wavelet transform and the results are encouraging with Daubechies wavelet. Robustness of the algorithm was also analysed by including additive white Gaussian noise, denoising, and resampling.

Keywords
Audio watermarking, discrete wavelet transform, Indian classical songs.

1. INTRODUCTION

Internet presents multimedia data in large amount, in the form of images, audio and video. Rapid developments in telecommunication networks, multimedia technology, and high speed computing systems have enabled the transmission and distribution of digital multimedia files, easily and efficiently to distant places. Unauthorized copying and distribution of multimedia files have been increased enormously which has prompted the need for enforcing copyright protection of digital data.

Encryption of digital multimedia necessitates the need of a proper decryption key and thus prevents access to the multimedia content to an individual. Content providers get paid for the delivery of perceivable multimedia, and each client that has paid the royalties must be able to decrypt a received file properly. Once the multimedia has been decrypted, it can be repeatedly copied and distributed without the need of any bottlenecks. But the enormous developments in multimedia technology have made data protection techniques such as encryption insufficient for protecting the intellectual properties of music industry.

In order to enforce the intellectual property rights and protect digital media from tampering, digital watermarking has been proposed as an alternative method [1, 2]. Watermarking is the process of inserting data into a multimedia element such as an image, audio, or video file. The embedded data can later be detected or extracted from the multimedia for identifying the copyright owner. Watermarking of digital media has received great interest in the literature and research community. Most watermarking schemes focus on image and video signals.

Audio watermarking is a difficult process because of the sensitivity of Human Auditory System (HAS). Achievement on imperceptibility of auditory system is more difficult than that of the visual system. In addition, the amount of data that can be embedded transparently into an audio sequence is considerably lower than the amount of data that can be hidden in video sequences as audio files normally has size much less than that of the video files.

Digital audio watermarking can be implemented in two ways, in time-domain and transform domain [1]. The former embed watermark by modifying the host's time-domain samples using typical algorithms [3]. Such methods generally can embed little watermark and have poor resistance to attack. Transform domain methods embed watermark by modifying the host signal's transform domain coefficients and have been observed to be more efficient and robust as compared to time domain techniques. The widely used transforms for audio watermarking are fast fourier transform (FFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) [4-10].

This paper presents an application of a watermarking algorithm [4] using discrete wavelet transform, with suitable changes in parameters for watermarking of Indian classical songs. Investigations have been performed using different wavelets with watermarking at different levels of decomposition, and their efficacy have been analysed. Performance of the algorithm is studied using the measures, normalized cross-correlation (NC), bit error rate (BER), peak signal to noise ratio (PSNR), mean square error (MSE), and objective difference grade (ODG). It is observed that Daubechies (db8) wavelet is most suitable for the application concerned with watermarking at A8 level. Robustness of the algorithm is established by adding AWGN, denoising and resampling. Investigations were performed with many clippings of Indian classical songs and the algorithm presents encouraging results.

2. DISCRETE WAVELET TRANSFORM BASED AUDIO WATERMARKING

The basic idea in the DWT based watermarking is time-frequency decomposition of the signal using suitable wavelet and embedding the bits corresponding to the watermark image into the wavelet coefficients. The basic scheme of digital audio watermarking using DWT is depicted in Fig. 1. Watermarking is done in either low frequency coefficients or high frequency coefficients based on the nature of the signal and applications.

Audio watermarking is the process in which watermark information is embedded into the audio signal. Watermarking problem can be mathematically represented as

$$xw = x + \propto w \qquad (1)$$

where xw is the watermarked audio signal, x is the original audio signal or its transform, w is the watermark and α is the watermark intensity.

(a) Watermarking embedding
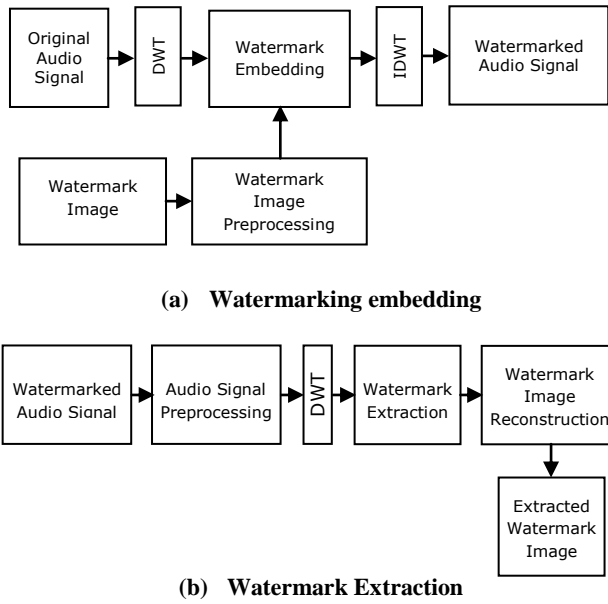

(b) Watermark Extraction

Fig. 1: Basic scheme of digital audio watermarking using DWT

The basic procedure used for watermark embedding [4] can be summarized as

a) Watermark embedding procedure

- Binary image watermark is expressed as a two dimensional matrix *(M1xM2)* which is converted into a one-dimensional watermark vector *W* of length *M1xM2*.
- Elements of *W* are multiplied by the watermark intensity *α* to obtain the final watermark vector, *Wf.*
- Audio signal is decomposed to n-levels using DWT.
- Embed the normalized watermark vector into wavelet coefficients of the selected sub-band.
- Apply the n-level inverse DWT to obtain the watermarked audio signal.

b) Watermark extraction procedure

- n-level DWT is applied to both the original and watermarked audio signals.
- Compute the final watermark vector, *Wf.*
- Divide *Wf* by the watermark intensity *α* to obtain the watermark vector *W*.
- Convert the one-dimensional image watermark vector into a two-dimensional matrix *W* of length *M1xM2*.

3. PERFORMANCE PARAMETERS

Efficiency of different schemes of watermarking is assessed with the help of a few parameters [10]. Perceptual quality assessment and watermark similarity detection are used for the assessment.

3.1 Perceptual Quality Assessment

Generally, there are two approaches to perform perceptual quality assessment of audio signals
  (i) Subjective listening tests by human acoustic perception.
  (ii) Objective evaluation tests

3.1.1 Subjective Evaluation

Blind listening tests have been performed for evaluating the subjective quality of the watermarking method. Participants listened to the original and the watermarked audio sequences and were asked to report dissimilarities between them, using a 5-grade impairment scale called Objective Difference Grade (ODG) [11] as shown in Table. 1.

Table 1: 5-grade impairment scale

Impairment description	ITU-R Grade	ODG
Imperceptible	5.0	0.0
Perceptible, but not annoying	4.0	-1.0
Slightly annoying	3.0	-2.0
Annoying	2.0	-3.0
Very annoying	1.0	-4.0

3.1.2 Objective Evaluation

Objective evaluation is performed by measuring the peak signal to noise ratio (PSNR) and mean square error (MSE).

PSNR is defined as

$$PSNR(X, {\sim}X) = 20\log10 \frac{2^B - 1}{\sqrt{MSE}} \qquad (2)$$

where X and ˜X are the original and the watermarked audio signals, B represents the bits per sample.

MSE is defined as squared norm of the difference divided by the number of elements in the signal and is given as

$$MSE(X, {\sim}X) = \frac{||X - {\sim}X||^2}{N} \qquad (3)$$

3.2 Watermark Similarity Detection

Normalized cross-correlation (NC) is used to evaluate the similarity between the original watermark and the extracted watermark and is defined as follows [10]:

$$NC(W, {\sim}W)$$
$$= \frac{\sum_{i=1}^{M1} \sum_{j=1}^{M2} W(i,j) * {\sim}W(i,j)}{\sqrt{(\sum_{i=1}^{M1} \sum_{j=1}^{M2} W(i,j) * W(i,j))} * \sqrt{(\sum_{i=1}^{M1} \sum_{j=1}^{M2} {\sim}W(i,j) * {\sim}W(i,j))}}$$
$$(4)$$

where W and ~W are the original and the extracted watermarks respectively, and i, j are indices of the binary watermark image. If NC(W, ~W) is close to 1, then the similarity between W and ~W is very high. If NC(W, ~W) is close to zero, then the similarity between W and ~W is very low.

The bit error rate (BER) is also used to evaluate the watermark detection accuracy. The BER of the watermarked signal retrieval is defined as follows [5]:

$$BER = \frac{\sum_{i=1}^{M1} \sum_{j=1}^{M2} W(i,j) \oplus {\sim}W(i,j)}{M1 * M2} \qquad (5)$$

4. SIMULATION RESULTS AND DISCUSSIONS

A sequence of experiments has been carried out for careful analysis and comparison of results. The implementation is done using MATLAB 7.11. An image shown in Fig. 2 is used for watermarking and various files of Indian classical songs were used for embedding the watermarks.

Fig. 2: Watermark image

Watermark is embedded in the audio signal with intensity 0.01, 0.02, 0.1, 0.3, 0.5, and 0.7. Table 2 shows the Normalized cross-correlation (NC), bit error rate (BER), peak signal to noise ratio (PSNR), mean square error (MSE), and objective difference grade (ODG) at different watermark intensities. In order to compromise the inaudibility of the watermark; watermark intensity of 0.01 is selected for the further simulations.

Table 2: Performance parameters for different watermark intensities

Water-mark Intensity	NC	BER	PSNR (db)	MSE	ODG
0.01	1	0	94.6733	0.000022	0.0
0.02	1	0	94.6617	0.000022	0.0
0.03	1	0	94.6424	0.000022	-1.0
0.1	1	0	94.3064	0.000022	-1.0
0.3	1	0	92.1195	0.000039	-1.0
0.5	1	0	89.5880	0.000071	-2.0
0.7	1	0	87.3802	0.000118	-3.0

With the objective of investigating and inferring about the best performing wavelet for the application, DWT decomposition of the audio signal is carried out using Haar, Daubechies and Symlets wavelets. Eight levels of decomposition of the signal are performed and the watermark bits are embedded into the eighth level approximation coefficients. Table 3 shows the performance parameters while using Haar, Daubechies and Symlets wavelets for DWT decomposition. The Daubechies wavelet db8 provides a better performance in terms of PSNR. It is observed that all wavelets provide values of ODG, NC and BER 0, 1, and 0 respectively. However, based on the values obtained for PSNR and MSE, dB8 wavelet is chosen for further investigations. .

Simulations have been performed by embedding the image at different decomposition levels in order to obtain the optimum decomposition level in which watermark is to be embedded. Best performance is provided while embedding the watermark bits at A8 level. This is in line with the results obtained by Erçelebi and Batakçı [12]. Table 4 depicts the details of results.

Table 3: Performance parameters with different wavelets

Wavelet	NC	BER	PSNR (db)	MSE	ODG
Haar	1	0	94.6733	0.0000221	0.0
db1	1	0	94.6733	0.0000221	0.0
db2	1	0	99.9284	0.0000061	0.0
db4	1	0	104.725	0.0000219	0.0
db8	1	0	107.156	0.0000012	0.0
db16	1	0	107.860	0.0000010	0.0
sym1	1	0	94.6733	0.0000221	0.0
sym2	1	0	99.9284	0.0000066	0.0
sym4	1	0	104.592	0.0000022	0.0

Table 4: Performance parameters with embedding at different decomposition levels using db8wavelet

Embedding Level	NC	BER	PSNR (db)	MSE	ODG
A8	1	0	107.860	0.0000	0.0
A7	1	0	107.1567	0.0000	0.0
A6	1	0	88.0620	0.0001	0.0
A5	1	0	82.0599	0.0004	0.0
A4	1	0	76.4470	0.0015	-1.0
A3	1	0	74.5220	0.0023	-2.0

Considering the results obtained in Tables 2, 3, and 4, dB8 wavelet with embedding at A8 level using an intensity of 0.01 is chosen for further simulations. Although simulations using only one audio signal have been shown above, it was ascertained that similar results are obtained for other samples of classical songs.

Fig. 3 shows the original audio signal. Fig. 2 shows the binary watermark image of size M1 × M2 = 64 × 64 = 4096 bits. Fig. 4 and Fig. 5 show the watermarked audio signal and the extracted watermark image respectively using Daubechies (db8) wavelet.
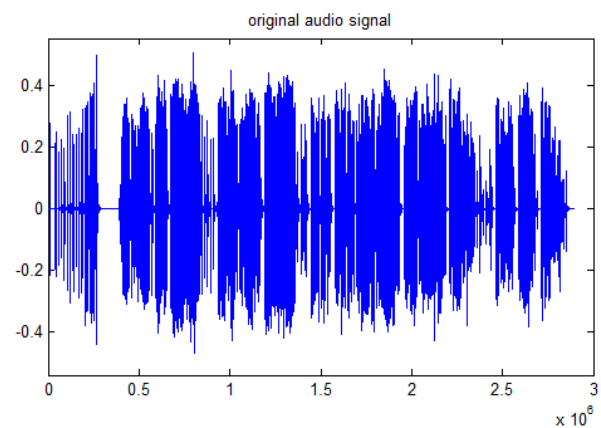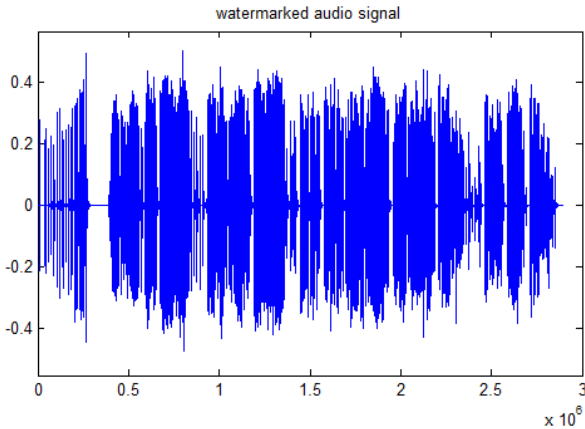

Fig. 3: Original Audio Signal

Fig. 4: Watermarked audio signal

Fig. 5: Extracted watermark image

4.1 Analysis of Robustness

Robustness refers to the ability to detect the watermark after common signal processing operations. Audio watermarking needs to be robust to noise, A/D conversion, resampling, etc. [13].

4.1.1 Effect of Additive white Gaussian noise (AWGN) and Denoising

White Gaussian noise is added to the watermarked signal till the resulting signal has an SNR of 20 dB and the performance parameters are computed. Table 5 shows the performance parameters at different SNR. Table 6 shows the performance parameters after denoising the watermarked signal.

Table 5: Performance parameters when AWGN is added

SNR(db)	NC	BER	PSNR (db)	MSE	ODG
50	1	0	106.912	0.0000013	0.0
40	1	0	105.172	0.0000019	0.0
35	0.994	0.00	102.638	0.0000035	-1.0
30	0.819	0.06	98.8403	0.0000084	-2.0
25	0.565	0.27	94.3013	0.0000241	-3.0
20	0.394	0.52	89.4507	0.0000737	-3.0

Fig. 6 shows the extracted watermark from the noised audio signal with noise added with an SNR of 30db.

Fig.6: Extracted watermark image with AWGN

Table 6: Performance parameters when watermarked signal is denoised

SNR (db)	NC	BER	PSNR (db)	MSE	ODG
50	1	0	106.911	0.0000013	0.0
40	1	0	105.168	0.0000019	0.0
35	0.998	0.00	102.637	0.0000035	0.0
30	0.829	0.05	98.8366	0.0000085	-1.0
25	0.575	0.26	94.3005	0.0000241	-1.0
20	0.427	0.51	89.4593	0.0000736	-2.0

4.1.2 Effect of Re-sampling

The watermarked signal, originally sampled at 44100Hz, is re-sampled at 5500Hz, 11000Hz, 22000Hz and 88000Hz; and then restored back by sampling again at 44.1 kHz. Table. 7 shows the performance parameters after resampling.

Table 7: Performance parameters with re-sampling.

Resampling Rate (Hz)	NC	BER	PSNR (db)	MSE	ODG
5500	1	0	107.86	0.00000106	0.0
11000	1	0	107.86	0.00000106	0.0
22000	1	0	107.86	0.00000106	0.0
88000	1	0	107.86	0.00000106	0.0

4.2 Performance Analysis for Different Classical Songs

In order to establish that the proposed scheme using watermarking with DWT is most suitable for Indian classical songs, investigations were further carried out with many such songs and very encouraging results were obtained. Results obtained in case of 3 such songs are presented in Table 8.

Table 8: Performance parameters when watermarking was applied to 3 different classical songs

Songs	NC	BER	PSNR (db)	MSE	ODG
S1	1	0	107.156	0.0000012	0.0
S2	1	0	101.788	0.0000043	0.0
S3	1	0	94.6009	0.0000225	0.0

Robustness was also verified with classical songs by adding AWGN. Table 9 shows the performance parameters when AWGN was added with an SNR of 35 dB and it demonstrate that the watermarking technique is robust against AWGN. Further, the AWGN was removed using denoising techniques and the performance parameters were estimated and the performance is depicted in Table. 10.

Table 9: Performance parameters with AWGN noise of 35 dB

Songs	NC	BER	PSNR (db)	MSE	ODG
S1	0.994	0.0015	102.6389	0.0000035	-1.0
S2	0.999	0.0002	100.2944	0.0000060	-1.0
S3	0.900	0.0310	93.7074	0.0000276	-1.0

Table 10: Performance parameters after removing AWGN with denoising

Songs	NC	BER	PSNR (db)	MSE	ODG
S1	0.998	0.00048	102.6376	0.0000	0.0
S2	1	0	100.2925	0.0000	0.0
S3	0.913	0.0259	93.7104	0.0000	0.0

5. CONCLUSIONS

In this paper, a technique for audio watermarking algorithm based on the Discrete Wavelets Transform is investigated for watermarking Indian classical songs. Comparisons using different wavelets and other parameters reveal that watermarking A8 sub-band wavelet coefficients with Daubechies wavelet, dB8 with a watermark intensity of 0.01 give promising results. Comparison of performances is carried using the parameters, peak signal to noise ratio, normalized cross-correlation, bit error rate, mean square error, and objective difference grade. Robustness of the algorithm is also ascertained by incorporating AWGN, denoising, and resampling.

6. REFERENCES

[1] Cvejic N., "Algorithms for Audio Watermarking and Steganography", Academic Dissertation, University of Oulu, 2004

[2] Alsalami M. A. T. and Al-Akaidi M. M., Digital Audio Watermarking: Survey, Proceedings 17th European Simulation Multi-conference, SCS Europe BVBA, 2003.

[3] Bassia P., Pitas I and Nikolaidis N., Robust audio watermarking in the time domain. IEEE Transactions on Multimedia 3(2), 2001, 232–241.

[4] Al-Haj A., Mohammad A. and Bata L., DWT–Based Audio Watermarking, The International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011, pp. 326-333.

[5] Vieru R., Tahboub R., Constantinescu C., and Lazarescu V., New Results Using the Audio Watermarking Based on Wavelet Transform, International Symposium on Signals, Circuits and Systems, ISSCS, IEEE, 2005, 441-444.

[6] Rehman F., Khan R., Iqbal N. and Ahmed F., Entropy based Audio Watermarking in Image using Wavelet Transform, ISECS International Colloquium on Computing, Communication, Control, and Management, IEEE, 2009, 478-481.

[7] Datta K. and Sengupta I., A Redundant Audio Watermarking Technique using Discrete Wavelet Transformation, Second International Conference on Communication Software and Networks, IEEE, 2010, pp. 27-31.

[8] Lei B. Y., Soon I. Y., and Li Z., Blind and robust audio watermarking scheme based on SVD-DCT, Signal Processing 91(8), 2011, 1973-1984

[9] Baoyuan C., Yanli H., Ruigang L. and Gang H., The audio watermarking system based on wavelet transform algorithm, Cross Strait Quad-Regional Radio Science and Wireless Technology Conference, 2011, 1274 – 1277.

[10] Bhat K. V., Sengupta I., and Das A., An adaptive audio watermarking based on the singular value decomposition in the wavelet domain", Digital Signal Processing, 20 2010, 1547–1558.

[11] Method for Objective Measurements of Perceived Audio Quality, International Telecommunications Union Union Recommendation BS.1387, 1998.

[12] Erçelebi E. and Batakçı L., Audio watermarking scheme based on embedding strategy in low frequency components with a binary image, Digital Signal Processing 19, 2009, 265–277.

[13] Yang S., Tan W., Chen Y., and Ma W., Quantization-Based Digital Audio Watermarking in Discrete Fourier Transform Domain, Journal of Multimedia, Vol. 5, No. 2, April 2010, 151-158