DWT-FFT Based Audio Watermarking Algorithm for Copyright Protection

Sheetal Shelke, Mangal Patil, J. S. Chitode

Abstract: In this proposed method a new technique is introduced to secure audio communication. Discrete Wavelet Transforms (DWT) and Fast Fourier Transform (FFT) are used in this proposed method. Separation of high frequency component and low frequency component from original audio signal is performed by applying DWT. High frequency component is then passed through Fourier Transform (FFT). Digital watermark is generated using PN sequence. The digital watermark is embedded in low amplitude high frequency region of magnitude spectrum of FFT. DWT-FFT based proposed algorithm can be used for Copyright protection of audio signals. Proposed algorithm is evaluated using SNR and NC parameters with various attacks including volume scaling, low pass filter, resampling, requantization, MP3 compression, Echo addition, time stretching and additive noise.

Keywords: Audio watermarking, Discrete Wavelet Transforms, Fast Fourier Transform

I. INTRODUCTION

Now a day, most of the data is available in digital form. A major problem faced during accessing the data is unauthorized copying. There are many technique introduced to secure audio communication [1].Many developers have introduced different watermarking techniques to avoid unauthorized copying. When some secret data is to be hidden in original file then this method is known as digital watermarking.LSB watermarking algorithm is used in this proposed method[2]. A secret data is to be hidden into the least significant bit (LSB)of magnitude spectrum of FFT. This method is easy to implement and can be used to hide larger secret messages. Hiding data in LSBs of audio samples in the time domain is one of the simplest algorithms enabling a very high data rate [3].Many algorithms have been developed to challenge the robustness of this method. The rest of the paper is organized as follows: Literature survey is described in section II. The details of Proposed Technique elaborated in section III. Section IV contains Experimental results, followed by conclusion and future scope in section V.

II. LITERATURE SURVEY

There are many techniques available for audio watermarking. Most of the algorithms are using wavelets. Few of them are reviewed as follows:

Siwar Rekik1, et.al [1] introduced an audio watermarking technique using DWT-FFT.

In this proposed method using DWT, high frequency component and low frequency component are separated for speech signal. In Low frequency component, another speech signal as watermark is embedded. In this proposed technique, experimental results are evaluated using SNR between original and watermarked speech signal (male or female voice) which is more than 20dB.

P.K.Dhar, M.L.Khan,et.al[4]has proposed method for audio watermarking using Fourier transform. Transformation based algorithms generally used to embed watermark bits by preserving the properties of the data in the representation following the transformations. Popular transformations are Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT). Some techniques, such as Quantization Index Modulation (QIM), Singular Value Decomposition (SVD), and interpolation, are often utilized to manipulate the data to embed watermark bits in the representation following transformations. Many watermarking algorithms fall into this category because the watermark bits embedded are more robust against attacks. Discrete Fourier Transform (DFT) developed fast version called as FFT. This is well known for performing frequency analysis of discrete time signals and it is powerful computational tool. There has been a variety of watermarking algorithms proposed that are based on manipulating the components contained in the FFT spectrum. Most algorithms manipulate the magnitude of the FFT components and enhance the robustness against typical audio compression systems by incorporating a model of the HAS.

Pranab Kumar Dhar,Jong-Myon Kim[5] have introduced an algorithm using FFT. FFT analysis is applied to each frame (i.e. short segment) of the original signal to derive the magnitudes of the odd bits have proposed in the method. Then the interpolated magnitudes of the even bins are derived by spine interpolation of the magnitudes of the odd bits. Firstly manipulating these spine-interpolated magnitudes of the even bits then embed the watermark bits. Finally, the watermarked signal is reconstructed by inverse FFT. In the original signal is segmented into non-overlapping frames. Watermarks are embedded by manipulating the magnitude of the highest prominent peak in the spectrum of each frame. The extraction process is an exact inverse of the watermark embedding process. In the original signal the watermark bits are embedded by swapping the magnitudes of the spectrum of the original signal at some selected frequencies. The selected frequencies and swapping of magnitude are tuned carefully, aiming at achieving a good balance between the imperceptibility and the robustness. By applying a low pass filter and a high pass filter respectively, decompose DWT signal into an approximation signal and a detail signal, both signals are then down sampled by a factor of two proposed

Revised Version Manuscript Received on December 19, 2015.

S. V. Shelke, has obtained the B.E. degree in Electronics and Telecommunication Engineering from Swami Ramanand Tirth Marathwada University, Nanded, Maharashtra, India.

M. V. Patil, has obtained the B.E degree in Electronics Engineering from Shivaji University, Kolhapur, India.

Dr. J. S. Chitode, has received the B.E. degree in Industrial Electronics Engineering from Bharati Vidyapeeth University, Pune, Maharashtra, India.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

by **K. Fitz and Haken**[6].The approximation signal can be subsequently divided into a new approximation and detail signals. The maximum number of levels is determined by the length of the signal. For each level, a high pass filtering output (“detail” coefficients) and a low pass filtering output (“approximation” coefficients) are produced. An example of one level DWT process is depicted in Figure1.The low pass filtering output is processed again to generate the subsequent level of “approximation” and “detail” coefficients.

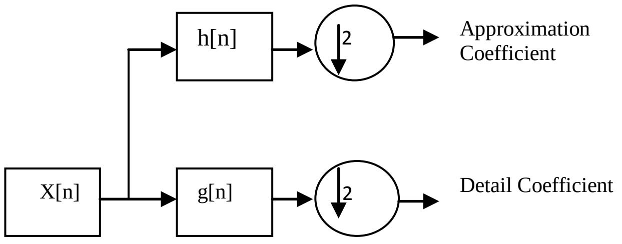

Figure 1. One level Wavelet Decomposition

III. PROPOSED TECHNIQUE

Block diagram of the proposed audio watermarking technique using DWT-FFT is shown in figure 2.The proposed algorithm is divided into two parts: first is watermark embedding process and second watermark extraction process. HARR wavelet transform applied on original audio signal[7].It decomposes audio signal into detail and approximation signal. Perform FFT operation on approximation signal. Secret key will generate PN-sequence which is used as watermark. This watermark is going to embed into LSB amplitude of magnitude spectrum of FFT[8]. Proposed method must undergo various attacks such as volume scaling, low pass filter, resampling, requantization, MP3 compression, Echo addition, time stretching and additive noise to check robustness. The main intention of this paper is to improve audio quality of watermarked audio signal, also get a satisfied SNR,PSNR of extracted audio file and to evaluate this improved quality results against different attacks like MP3 compression, Echo addition, time stretching, volume scaling ,low pass filtering, high pass filtering, additive noise. To demonstrate the application of DSP many several languages are used. But it requires high accuracy in programming. MATLAB software makes many operations easy to write.

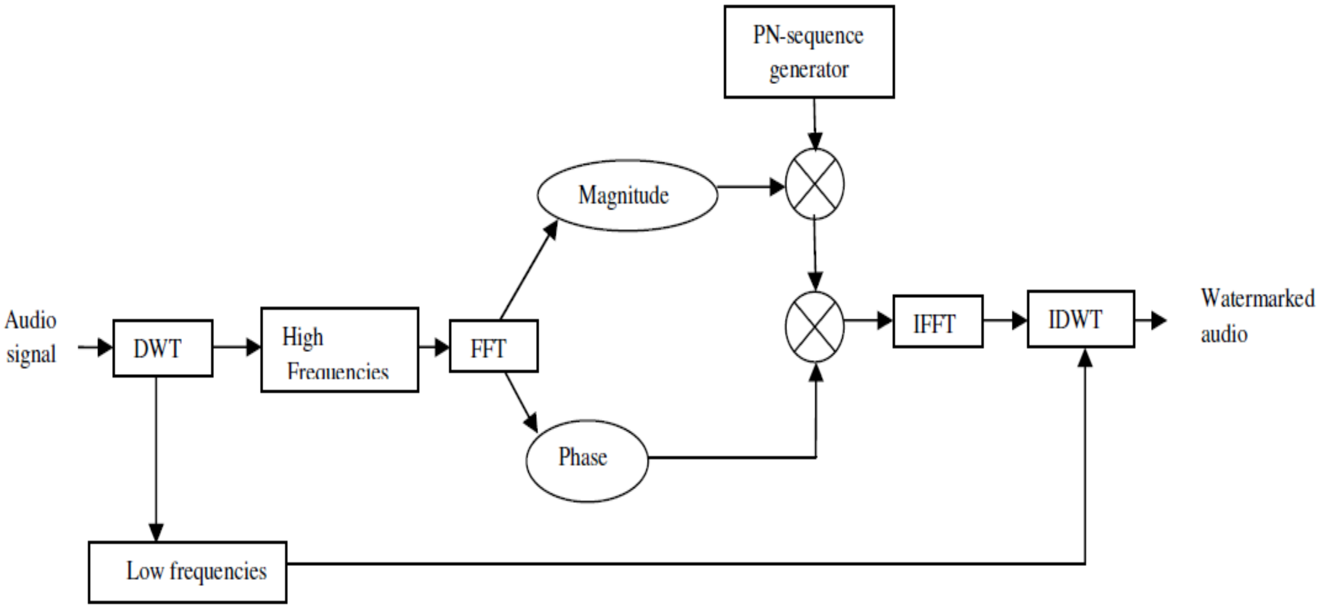

Figure 2. Block diagram of audio watermarking using DWT-FFT

Watermarking scheme can be divided into two main categories: spatial domain and Transform Domain. Spatial domain is having low frequency hiding capacity [9]. Watermark is embedded in spatial domain. It is direct method which has less computational cost, high capacity and more perceptual quality there are some advantages of spatial domain but it is less suitable for authentication application. In the transform domain desired watermark is get embedded into frequency coefficients[10].This domain having less perceptual quality but it is more robust so widely used in copyright protection. DWT is used to make separation of audio signal into high frequency component and low frequency component. Then FFT is applied to High frequency component of DWT [11]. Digital watermark is created as PN sequence and going to hide into the LSB of FFT signals means of this hide large amount of data as concern to protect original data[12]. This watermarking problem can be mathematically represented as,

$$X'(i) = X(i) + \alpha W \qquad (1)$$

Where X(i) and X’(i) are original signal and watermarked signal respectively, W considered as watermark, α denotes scaling factor.

A. Watermark Embedding procedure:

In this proposed algorithm input is given as audio signal. DWT applied to original audio signal then perform FFT operation on high frequency component [13]. Watermarked audio is obtained by applying inverse Fast Fourier Transform and Inverse Discrete Wavelet Transform [14]. The proposed watermark embedding process is shown in Figure 3. The embedding process is implemented in the following steps:

1. Original audio signal is decomposed into the high

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

frequency component and low frequency component by using DWT(discrete wavelet transform).

2. Apply FFT (fast Fourier transform) to the high frequency component.
3. Find LSB of each frame from magnitude spectrum using algorithm.
4. Use secret key to generate PN sequence which is used as digital watermark.
5. Hide digital watermark into LSB of each frame from the magnitude spectrum having low frequency in selected range. This ensures that watermark is located into LSB of each frame.
6. Apply an inverse FFT of the spectrum.
7. Add the low frequency signal which is separate out previously by using DWT.
8. Take an inverse DWT of this signal to get original audio signal.

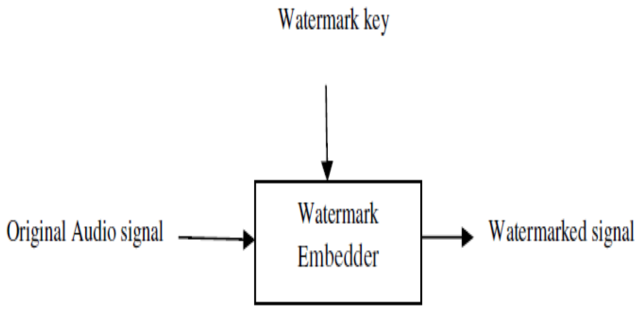

Figure 3: Watermark Embedding process

B. Watermark Extraction Process: Watermarked audio signal is taken as an input then apply inverse discrete wavelet transform. IFFT is performed on approximation coefficient of DWT. PN sequence is extracted using watermark key, this is called as extracted watermark [15]. Audio signal is retrieved as output. The proposed watermark extraction process is shown in Figure 4. This process is implemented in the following steps:

1. Give the secret key.
2. Obtain LSB of each frame from the magnitude spectrum having same frequency in selected range.
3. Extract PN sequence by using secret key.
4. Apply inverse FFT.
5. Apply inverse DWT and Extract the watermark sequence.

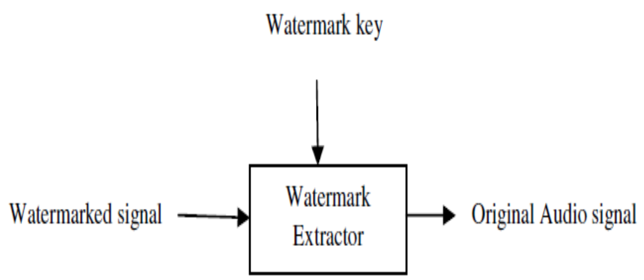

Figure 4: Watermark Extracting Process

IV. EXPERIMENTAL RESULTS

The proposed technique is evaluated using two methods:

1) Subjective evaluation and 2) objective evaluation.

This proposed algorithm is robust against various attacks such as volume scaling, additive noise,MP3 compression etc. Once attack is applied original audio file is successfully retrieved.

a) Subjective evaluation:

Mean Opinion Score (MOS) is known as Subjective evaluation measure. Table1 represent the standard MOS[16].This subjective evaluation is dependent directly on human listening so it takes more time. The result of this subjective evaluation differs from person to person. Objective evaluation is preferred for obtaining more accurate results.

Table 1: Mean Opinion Score

Grade	Description	Quality
5	Imperceptible	Excellent
4	Perceptible	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very Annoying	Bad

Table2 contains evaluation results for MOS between original audio signal and watermarked audio signal. Method for Subjective Evaluation is as follows:

1. We played audio files of various types such as classical, rock, pop, flute, speech, jazz, instrumental in front of various people.
2. After this Watermarked audio files are played.
3. People were asked to find out difference based on audio quality of original and watermarked file.
4. After that people were asked to give grade of each sample on scale 1 to 5 is shown in table 2.
5. From the given standard grade, average grade for each audio file is calculated.

Table 2: Experimental values of Subjective Measure

Audio sample	Average MOS
Classical	3.9
Rock	4.2
Pop	4.7
Flute	4.2
Speech	4.1
Jazz	4.8
Instrumental	4.6

The Proposed algorithm is designed using MATLAB10 for different audio signal such as classical, rock, pop, flute, speech etc. The proposed algorithm should give more than 20 dB SNR according to IFPI [international federation of photographic industry]. Keeping in mind that various attacks are applied on audio signal .The experimental results show the improved SNR values which satisfies IFPI requirements as well[17].

b) Objective Evaluation:

In order to determine the robustness of the proposed method, various audio files such as classical, rock, pop, flute, speech with wave format were used. The implementation and evaluation is performed using MATLAB10. The applied attacks are 1)volume scaling 2)addictive noise 3)low pass filter 4)high pass filter 5) resampling 6) requantizatin 7) Echo addition 8) MP3 compression 9) Time stretching .

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

i. **Signal to Noise Ratio (SNR):**

Signal to noise ratio is used to calculate the amount by which the signal is corrupted by the noise. The ratio of signal power to noise power is known as SNR (Signal to Noise ratio)[18] .Quality of watermarked signal is evaluated in terms of signal to noise ratio (SNR) using equation:

$$\text{SNR} = \frac{10 \log 10 \sum X(i)^2}{\sum\{X(i) - X'(i)^2\}} \quad (2)$$

Where X (i) and X' (i) are original and watermark audio signal respectively. Experimental results of SNR using various attacks are shown in Table3.In this proposed method, experimental results provide average SNR values between 35dB to 70 dB.

ii. **Normalized Correlation(NC):**

Experimental results of NC are shown in Table 4.Normalized correlation is used to evaluate the correlation between the extracted and original watermark [19] and is given by,

$$\text{NC}(W,W') = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} W(i,j)W'(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} W^2(i,j)} * \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M} W'^2(i,j)}} \quad (3)$$

Where W(i,j) and W'(i,j) are original watermark and extracted watermark respectively.

Table 3: SNR values with attack in dB:

Various Attacks	**SNR values without attack in dB**						
Audio files	**classical**	**Pop**	**Rock**	**Flute**	**Speech**	**Jazz**	**Instrumental**
SNR without attack	59.18	44.85	68.06	50.78	71.39	77.18	36.78
Payload	34.8	50.7	57.8	60.8	48.7	56.8	49.7
	SNR values with various attack in dB						
Volume scaling-0.8	58.44	42.91	66.15	48.22	69.46	76.14	36.19
Volume scaling-0.9	59.47	43.93	67.16	51.25	70.48	75.18	35.14
Volume scaling-1.1	61.21	45.68	68.87	51.01	72.22	76.45	38.49
Volume scaling-1.4	63.31	47.77	70.94	53.10	72.85	74.45	36.12
Additive noise	60.49	44.50	67.84	49.99	71.38	75.18	39.45
Low pass filter	60.38	44.83	68.07	50.17	71.39	79.14	39.48
High pass filter	61.54	45.98	67.54	50.16	70.48	69.14	38.47
Resampling	60.37	44.84	68.05	50.16	71.38	70.18	39.44
Requantization 8-bit	60.74	42.19	66.81	49.55	70.16	75.16	35.78
Requantization16-bit	59.18	44.83	68.06	50.17	66.63	79.48	37.49
Requantization 24-bit	59.18	45.85	68.06	51.59	71.39	70.14	39.99
Time stretching	40.15	48.16	30.82	49.37	47.59	65.18	35.48
Eco addition	59.89	44.86	68.14	50.17	47.59	79.45	36.19
White Gaussians noise	58.15	44.89	67.98	50.19	46.28	70.18	40.48
MP3compression 64 kbps	34.16	39.48	40.78	40.58	49.57	47.49	71.45
MP3 compression128 kbps	36.94	38.45	38.47	49.57	48.24	49.78	50.17

Table 4: NC values with attack:

Various Attacks	**NC values with attacks**						
Audio files	Classical	Pop	Rock	Flute	Speech	Jazz	Instrumental
NC without attack	1	1	1	1	1	1	1
NC with attack	**NC values with various attacks**						
Volume scaling-0.8	0.95	1	0.96	0.97	0.93	1	0.99
Volume scaling-0.9	0.95	1	096	0.97	0.90	1	0.98
Volume scaling-1.1	0.95	1	0.96	0.97	0.90	1	1
Volume scaling-1.4	0.95	1	0.96	0.97	0.91	0.99	0.93
Additive noise	0.95	1	0.77	0.97	0.96	1	0.96
Low pass filter	0.95	1	0.96	0.97	0.97	0.94	0.96
High pass filter	0.98	1	1	1	0.90	0.94	0.97
Resampling	0.95	0.92	0.83	0.97	0.90	0.98	0.98
Requantization 8-bit	0.66	1	0.96	0.62	0.90	1	0.99
Requantization 16-bit	0.85	1	0.96	0.97	0.97	0.97	0.97
Requantization 24-bit	0.85	1	0.97	0.97	0.96	0.97	0.97
Time stretching	1.41	0.45	0.56	0.47	0.5	0.47	0.48
Eco addition	0.95	1	0.96	0.97	0.90	1	0.94

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

White Gaussian noise	0.95	1	0.97	1	0.98	0.99	0.98
MP3compression64 kbps	0.98	0.97	1	1	0.98	1	0.98
MP3compresion128 kbps	1	0.97	1	0.90	1	0.97	0.97

V. CONCLUSION

A new facet in robust audio watermarking in DWT-FFT domain is proposed and analyzed in this paper. The most attracting feature of this method is the remarkable improvement in the robustness of watermark. So without disturbing quality of audio signal good authenticity is achieved. Robustness of this technique is evaluated against various attacks such as Echo addition, time stretching, MP3 compression, low pass filtering, high pass filtering, volume scaling, and additive noise. The results indicate that this technique provide better results than other methods. In this proposed method16 bit PN sequence is used as watermark. The robustness of the proposed algorithm can be improved by increasing the length of PN sequence as 32 bit, 64 bit, etc.

REFERENCES

[1] Siwar Rekik, Driss Guerchi, Sid-Ahmed Selouani and Habib Hamam, "Speech steganography using wavelet and Fourier Transforms", EURASIP Journal on Audio, Speech and music processing 2012.
[2] B.Chen and G.W.Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," IEEE Trans. Inform. Theory, vol. 47 (4), pp. 1423–1443, 2001.
[3] T.Cedric, R.Adi, and I.Mcloughlin, "Data concealment in audio using a non- linear frequency distribution of PRBS coded data and frequency-domain LSB insertion," IEEE Region 10 International Conference on Electrical and Electronic Technology, Kuala Lumpur, Malaysia, pp. 275-278, September 2000.
[4] P.K.Dhar, M.L.Khan and J.M.Kim, "A New Audio Watermarking System using Discrete Fourier Transform for Copyright Protection," International Journal of Computer Science and Network Security, vol.10 (6), June 2010.
[5] Pranab Kumar Dhar, Jong-Myon Kim, "Digital Watermarking Scheme Based on Fast Fourier Transformation for Audio Copyright Protection" International Journal of Security and Its Applications Vol. 5 No. 2, April, 2011.
[6] K.Fitz and L.Haken, "On the Use of Time-Frequency Reassignment in Additive Sound Modeling," Journal of the AES, vol. 50, no. 11, pp. 879-893, Nov. 2002.
[7] Ali Al-Haj, Ahmad Mohammad and Lama Bata, "DWT–Based Audio Watermarking", the International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011.
[8] N.Cvejic and T.Seppänen, "Increasing robustness of LSB audio steganography by reduced distortion LSB coding," Journal of University Computer Science, vol. 11(1), pp. 56-65, 2005.
[9] Y.Yoshitomi, T.Asada, Y.Kinugawa and M.Tabuse, "An Authentication Method for Digital Audio Using a Discrete Wavelet Transform," Journal of Information Security, Vol. 2 No. 2, pp. 59-68. doi: 10.4236/jis.2011.22006, 2011.
[10] Harleen Kaur, Usvir Kaur, "Blind Audio Watermarking schemes: A Literature Review", IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498, Vol.3, No.2, April 2013.
[11] J.Cooley and J. Tukey, "An algorithm for the machine calculation of the complex Fourier series," Math. Comput., vol. 19, pp. 297-301, 1965.
[12] B.P.Bogert, M.J.R.Healy and J.W.Tukey, "The Quefrency Alanysis of Time Series for Echoes: Cestrum, Pseudo-Auto covariance, Cross-cestrum and Shape Cracking," in Proceedings of the Symposium on Time Series Analysis, by M. Rosenblatt, pp. 209-243, Wile N.Y. 1963.
[13] J.Cooley and J.Tukey, "An algorithm for the machine calculation of the complex Fourier series," Math. Comput., vol. 19, pp. 297-301, 1965.
[14] S.Murata, Y.Yoshitomi and H. Ishii, "Audio Watermarking Using Wavelet Transform and Genetic Algorithm for Realizing High Tolerance to MP3 Compression," Journal of Information Security, Vol. 2 No. 3, pp. 99-112. doi: 10.4236/jis.2011.23010, 2011.
[15] ZhengLiuandAkiraInoue, "Audio Watermarking Techniques Using Sinusoidal Patterns Based on Pseudorandom Sequence", IEEE Transactions On Circuits And Systems For Video Technology,Vol.13,NO.8,pp.801-812,August2003.
[16] Komal Shedge, Mangal Patil, J.S.Chitode ,"A new facet in Robust Audio Watermarking in DWT-SVD Domain", International Journal Of Engineering Research & Technology(IJERT),ISSN:2278-0181,Vol-3 Issue 5,May-2014.
[17] Wei Li, Xiangyang Xue, and Peizhong Lu, "Localized Audio Watermarking Technique Robust Against Time-Scale Modification" IEEE transactions on multimedia, vol. 8, no. 1, February 2006.
[18] I.J.Cox, J.Kilian, and T.Leighton, "T. Shammoon, Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. Image Process, vol. 6(12), pp. 1673-1687, 1997.
[19] P.Bassia and I.Pitas, "Robust audio watermarking in the time domain", IEEE transactions onMultimedia,Vol.3,Issue:2,pp.232–241,2001.

BIBLIOGRAPHY

S. V. Shelke has obtained the B.E. degree in Electronics and Telecommunication Engineering from Swami Ramanand Tirth Marathwada University, Nanded, Maharashtra, India in 2007. Currently pursuing M-Tech.(Electronics Engg.) from Bharati Vidyapeeth Deemed University College of Engineering, Pune. She has more than 6 years of teaching experience. She has been serving as a faculty member in the Department of Electronics Engineering at Bharati Vidyapeeth College of Engineering for Women, Pune.

M. V. Patil has obtained the B.E degreein Electronics Engineering from Shivaji University, Kolhapur, India in 1999. She received M.E. degree from the Bharati Vidyapeeth Deemed University College of Engineering at Bharati Vidyapeeth Deemed University, Indiain2006. Currently pursuing Ph.D. (Electronics Engg.) from Bharati Vidyapeeth Deemed University College of Engineering, Pune. She has more than 12 years of teaching experience. She has been serving as a faculty member in the Department of Electronics Engineering at Bharati Vidyapeeth Deemed University College of Engineering, Pune.

Dr. J. S. Chitode has received the B.E. degree in Industrial Electronics Engineering from Bharati Vidyapeeth University, Pune, Maharashtra in 1991. He received M.E. degree from College of Engineering (COEP), Pune at University of Pune from Maharashtra, India in1995. He has received Ph.D. degree in Electronics from Bharati Vidyapeeth Deemed University, India in 2009. Currently he is a professor in the Bharati Vidyapeeth Deemed University College of Engineering, Pune (India).His research interest includes Signal processing, Speech Synthesis, Digital Communication, etc. He is actively participating as a member of different professional research societies, like EEE, ISTE, etc.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.