World Applied Sciences Journal 29 (Data Mining and Soft Computing Techniques): 166-172, 2014
ISSN 1818-4952
© IDOSI Publications, 2014
DOI: 10.5829/idosi.wasj.2014.29.dmsct.31

Digital Watermarking of an Audio Signal Using an Image

Pavan Kumar Kota, Mallepalli Rakesh Reddy, R. Adityan and V. Elamaran

Department of ECE, School of EEE, SASTRA University, India- 613401

Abstract: A digital watermark is a digital signal or pattern inserted into a digital document such as text, graphics or multimedia and carries information unique to the copyright owner, the creator of the document or the authorized consumer. The main applications of Digital watermarking are mainly for authentication, identification and copyright protection. Digital watermarking is a technique of embedding one signal into another signal. In this study, an implementation of digital watermarking of an audio signal using image is made with different algorithms and a comparative study is obtained. The implementation is made using both spatial and frequency domain for the task of information hiding. The algorithms use the concepts of Least significant bit (LSB) substitution, Discrete cosine transform (DCT), Discrete wavelet transform (DWT), high frequency tone and echo hiding. Peak Signal-to-Noise Ratio (PSNR), Signal-to-Noise Ratio (SNR) and Mean Square Error (MSE) are the performance metrics used for the analysis and comparison among the techniques. The capacity of each algorithm is also noted and the results are obtained using Matlab.

Key words: Audio signal · DCT · Digital Watermarking · DWT · Echo Hiding · LSB

INTRODUCTION

Digital watermarking is a technique of embedding one signal into another signal. In this paper, one of the signals is an audio signal and the other is an image. Image is embedded into the audio signal. This is performed in digital domain. The process of digital watermarking has been accomplished by different algorithms. Based on visibility, visible and invisible watermarking techniques can be applied. Based on the content to be wa termarked, text, image, audio and video watermarking can be used. The characteristics of digital watermarks are also important to grade the quality of the algorithm used for watermarking.

Large number of watermarks can be generated and all of them are distinguishable. This is referred to known as innumerable.The unambiguous of watermarking is that the identifying the owner of the data can be made without ambiguity. The owner of the data or an independent control authority can be detected easily, is being referred as readily detectable. Visible watermarking can be used to discourage theft with enough visibility. Robust watermarking is the one in which the attacker finds difficult to remove or destroy the data in order to counterfeit the copyright of the owner. Moreover, removal of it should cause a considerable degradation in the quality of the data. Unobtrusive is an another characteristic of watermarking in which it is invisible enough not to degrade the quality and to prevent an attacker from finding and deleting it.

Each algorithm has two parts. Firstly, the image is embedded into the audio signal without allowing perceivable changes in the latter. Subsequently, in the second part, the image is retrieved back from the watermarked audio signal. The image that has been retrieved from the watermarked audio signal should not differ much from the watermark image. Retrieving the watermark image from the watermarked signal is very important as this is the only evidence of presence of a watermark in the audio [1]. The audio signal is checked for distortions by analyzing it before and after watermarking, using SNR and the watermark image is analyzed before watermarking and after retrieving, using PSNR and MSE. The capacity of each algorithm is then evaluated in terms of number of bits embedded per sample [2].

Digital Watermarking Algorithms: We implement various algorithms in both spatial and frequency domains to embed the image into audio signal. LSB substitution, DCT, DWT and Echo hiding are the techniques

Corresponding Author: Pavan Kumar Kota, Department of ECE, School of EEE, SASTRA University, India- 613401.

implemented for the embedding one into another [3]. A combination of both spatial and frequency domain is also can be implanted and is being referred to as hybrid domain watermarking.

LSB Substitution: The algorithm employing LSB substitution has good capacity. The audio signal is digitized by sampling and quantizing it. In the present method, the value of each sample of the audio signal is converted into its corresponding binary value with a 10 bit resolution [4]. The color image that is used as a watermark has 3 planes, red, blue and green (RGB). Pixel vale of an image is a combination of values of corresponding pixel in all the 3 planes. The value of each pixel of each plane is converted to its corresponding binary value. The binary value of each pixel is produced as an array in such a way that, each bit of the binary value for all the pixels in the three planes is made as an element of the array. Consider an example as we had an image of 10 pixels, it can be considered as 3 planes (RGB) of 10 pixels each with every pixel having an 8 bit value [5]. So, the total number of bits for all the three planes is 3x10x8 = 240. So the total number of elements in the array obtained is 240. Value of each element of the matrix is either “1” or “0”.Let us consider any one plane and see how the image array is built. It is shown in Figure 1. According to this figure, the image array is { 1 0 1 1 1 0 1 1 1 1 1 1 1 0 0 1 1 1 1 0 0 0 0 0 …….}.

One of the least significant bits of nth sample is replaced by the nth element of the image array. Here to make a trade-off between the quality of retrieved image and watermarked signal, the third least significant bit is chosen and is operated on. The samples obtained after bit substitution are reconstructed to generate the resultant signal which is watermarked audio. The second step of the algorithm is to retrieve the watermark image from the watermarked audio signal. The watermarked audio signal is digitized by sampling and quantization [6,7]. Each sample again has a 10 bit resolution. The 3^{rd} least significant bit in each sample is extracted and an array is generated. Hence we obtain the image array and we can retrieve the watermark image from the image array. This is the retrieved image and performance metrics like, SNR of watermarked audio, PSNR and MSE are found using basic formulae. As, in this method, 1 bit can be embedded into 1 sample we can say that its capacity is 1bit/sample.

DCT Technique: This algorithm uses the properties of DCT-II and also the concept of LSB substitution. Hence this approach is known as hybrid watermarking technique.

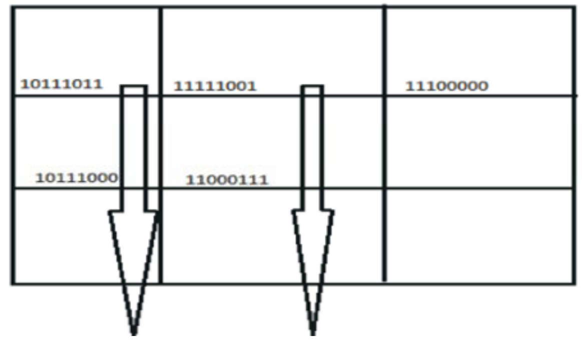


Fig. 1: Formation of image array from an image

Here initially, the original audio signal is sampled and hence we have finite discrete samples of the audio signal [1, 5]. Applying DCT to the digital samples gives as many coefficients as the number of samples. The DCT of a discrete signal is expressed [6] as in equation 1;

$$X(k) = \sum_{n=0}^{N-1} x(n)\cos\frac{\pi k}{N}(n+\frac{1}{2})$$

k = 0, 1, 2, …., N-1 (1)

LSB substitution is hence performed on the DCT coefficients by the elements of the image array obtained from the watermark image, as in Figure 1. Again we choose third least significant bit for reason stated in previous algorithm [9]. So, using the DCT we obtain the coefficients which are changed and the IDCT (inverse DCT) is applied to these changed coefficients to get the resultant signal which is watermarked audio. The second step of the algorithm is to recover back the watermark image from the above watermarked audio. Firstly the watermarked audio signal is sampled and then DCT is applied to these samples to obtain DCT coefficients [10]. Third least significant bit is taken from each coefficient which gives us the image array. Hence we obtain the image array and we can retrieve the watermark image from the image array. Again performance metrics are evaluated for the purpose of grade of algorithm and for the comparative analysis. As, in this method, 1 bit can be embedded into 1 coefficient of DCT and there are as many samples as the number of coefficients, we can say that its capacity is 1bit/sample [11].

DWT Technique: One of the most effective algorithms of digital watermarking of an audio signal is the one using discrete wavelet transform. We use the basic Haar wavelet for the task of information hiding. First the audio signal is

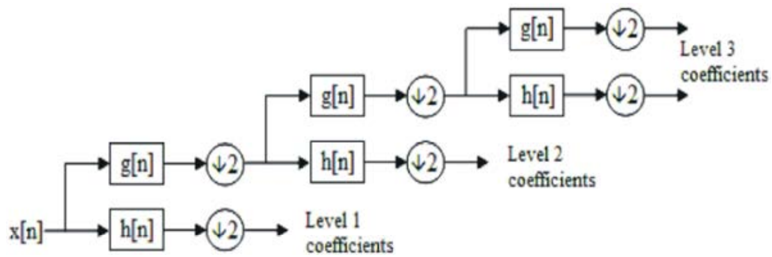


Fig. 2: Multiresolution analysis using DWT

sampled and the samples are sub-divided into frames containing equal number of samples. Applying Haar-DWT to a discrete signal of ‘2n’ samples gives us ‘n’ approximate coefficients and ‘n’ detailed coefficients. Applying DWT to a discrete signal of, say, 80 samples to obtained third level coefficients give 10 detailed coefficients and 10 approximate coefficients. Similarly, applying DWT to obtain second level coefficients of the same signal gives rise to 20 detailed and 20 approximate coefficients. A schematic view is shown below in Figure 2. Here g[n] represents detailed coefficients and h[n] represents approximate coefficients[12].

Two algorithms have been proposed based on the above concept.

DWT Algorithm-I: Third level detail coefficients of each frame are obtained using DWT [7, 8]. The image array is constructed based on the value of each pixel, as performed in Figure 1. If the nth element of the image array is 1, then each third level detail coefficient obtained from nth frame is increased by a fixed value. If the nth element of the image array is 0, then the third level detail coefficients of nth frame are not changed. After applying this for all the frames, we obtain audio signal with coefficients that have been changed. Applying IDWT (inverse DWT) to these changed coefficients gives us the watermarked audio signal [5] [13].

The second step of the algorithm is, to retrieve the watermark image from the watermarked audio signal. The watermarked audio signal is sampled, divided into frames and DWT is applied to it as done above. The third level detail coefficients are found for each frame. These coefficients are compared with the third level detailed coefficients of the original audio signal. If the coefficients of nth frame for both signals are not equal, then the nth element of the image array is taken as “1”. If the coefficients of nth frame for both signals are equal, then the nth element of the image array is taken as “0”. Hence we obtain the image array and we can retrieve the watermark image from the image array. Now, SNR of watermarked audio, PSNR and MSE of retrieved image are found using basic formulae. In this method, 1 bit is embedded into 1 frame of the original signal. Hence capacity of this algorithm is 1 bit/frame which is lesser when compared to capacities of algorithms using LSB and DCT.

DWT Algorithm-II: Second level detail coefficients of each frame are obtained using DWT. The image array is constructed based on the value of each pixel, as performed in Figure 1. The coefficients obtained from each frame are taken and formed as an array. If the nth element of the image array is 1, then the nth coefficient is increased by a certain fixed value. If the nth element of the image array is 0, then the nth coefficient is not changed. After applying this for all the frames, we obtain array with changed coefficients. These are again placed as they were, i.e., in the original frame format. Applying IDWT to this gives us the watermarked audio signal.

The second step of the algorithm is to retrieve the watermark image from the watermarked audio signal. The watermarked audio signal is sampled, divided into frames and DWT is applied to it as done above. The second level detail coefficients are found for each frame. These coefficients are compared with the second level detailed coefficients of the original audio signal. If the nth coefficients for both the signals are not equal, then the nth element of the image array is taken as “1. If the nth coefficients for both the signals are equal, then the nth element of the image array is taken as “0”. Hence we obtain the image array and we can retrieve the watermark image from the image array. The difference between DWT algorithm-I and algorithm-II is that in algorithm-I, 1 bit is embedded per 1 frame, whereas in algorithm-II 1 bit is embedded per 1 coefficient. Now, SNR of watermarked audio, PSNR and MSE of retrieved image are found using basic formulae. In this method, 1 bit is embedded into 1 coefficient of the original signal. Unlike DCT, number of coefficients is not equal to number of samples. For n samples of audio signal, there are n/4 second level detail coefficients. Hence capacity of this algorithm is 0.25 bits/sample.

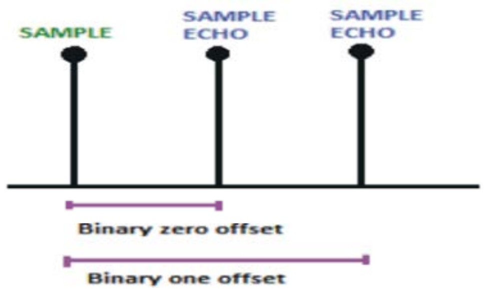


Fig. 3: Original audio signal and echo signals corresponding to "1" and "0"

Echo Hiding: Echo is a signal that has an offset from the original signal. Echoes are used in digital watermarking of an audio signal. This is a very interesting algorithm. Initially, the original audio signal is sampled. Two signals with different offsets are generated from the original audio signal. These are the echoes of the original signal. They have been shown in Figure 3.

They are sampled and divided into frames. Number of samples in one frame of echoes is same as number of samples in one frame of original signal. The image array is formed from the watermark image as shown in Figure 1. If the nth element of the image array is "1", then the echo with more offset is chosen and its nth frame is added to the nth frame of our original audio signal. If the nth element of the image array is "0", then the echo with less offset is chosen and its nth frame is added to the nth frame of our original audio signal. This process is done for all the elements of the image array. This gives rise to the watermarked audio signal.

The second step of the algorithm is to retrieve the watermark image from the watermarked audio signal. The watermarked audio signal is sampled and divided into frames. Each frame has same number of samples as the original signal had during the process of watermarking. The nth frame of the watermarked audio signal is correlated with both the echo signals. If correlation value is more for the echo with more offset, then the nth element of image array is taken as "1". If correlation value is more for the echo with less offset, then the nth element of image array is taken as "0". Hence we obtain the image array and we can retrieve the watermark image from the image array. Now, SNR of watermarked audio, PSNR and MSE of retrieved image are found using basic formulae. As, in this method, 1 bit can be embedded into 1 frame of the original audio signal, we can say that its capacity is 1bit/frame.

High Frequency Tone: Normal human perception of high frequencies is limited to around 20 KHz. That is, any signal that has a frequency in this range cannot be perceived by human ear. This property is made use in digital watermarking of an audio signal. The process of watermarking using a high frequency tone is as follows. The audio signal that is to be watermarked is sampled. The samples are divided into frames in such a way that each frame has equal number of samples. A high frequency tone is taken whose duration is equal to the duration of one frame of original audio signal and it is sampled. Now, we consider the watermark image. Since the capacity of this algorithm is low, we choose a binary image. This image is converted into its corresponding image array as shown in Figure: 1. If the nth element of the image array is "1", then the high frequency tone is added to the nth frame samples. If the nth element of the image array is "0", then the nth frame samples are not disturbed. The samples obtained after performing the above procedure for all elements of the image array are reconstructed to form the watermarked audio signal.

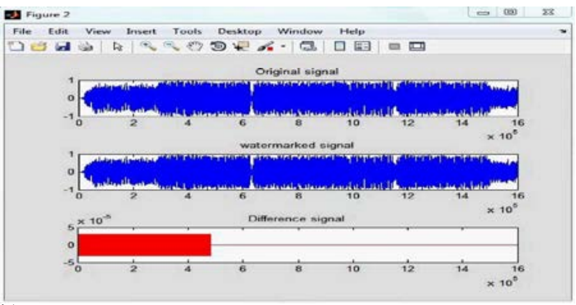
The second step of the algorithm is to retrieve the watermark image from the watermarked audio signal. It is done as follows. The original audio signal and watermarked audio signal are sampled and divided into frames as done during the process of watermarking. DWT of the original audio signal and watermarked signal is found for each frame. The detail coefficients are compared. Since detail coefficients contain information of high frequency elements, if the coefficients are unequal then it tells about the existence of high frequency tone in that particular frame. So, if we obtained that nth frame contains high frequency tone, it means that the nth element of image array is "1", otherwise the nth element of image array is "0". Hence we obtain the image array and we can retrieve the watermark image from the image array. Now, SNR of watermarked audio, PSNR and MSE of retrieved image are found using basic formulae. As, in this method, 1 bit can be embedded into 1 frame of samples, we can say that its capacity is 1bit/frame.

RESULTS

The above algorithms have been implemented using Matlab software tool and the results have been depicted below. The figures below contain watermark image, audio signal, watermarked audio signal and difference signal. Difference signal is effectively the difference between original audio signal and watermarked audio signal.

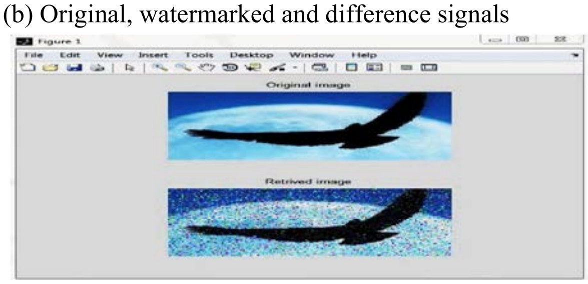
Using LSB Substitution: The process of LSB substitution has more capacity. Hence a color image can be used for watermarking. The distortions in the retrieved watermark

(a)

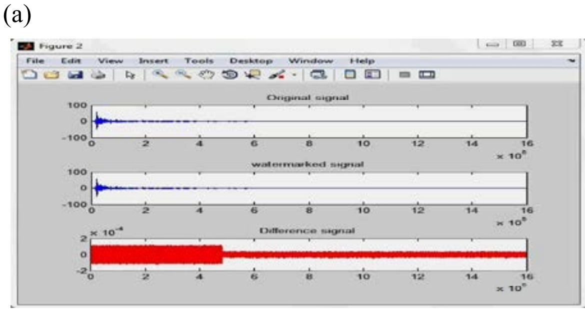


(b)

Fig. 4: (a) Original and retrieved image
(b) Original, watermarked and difference signals



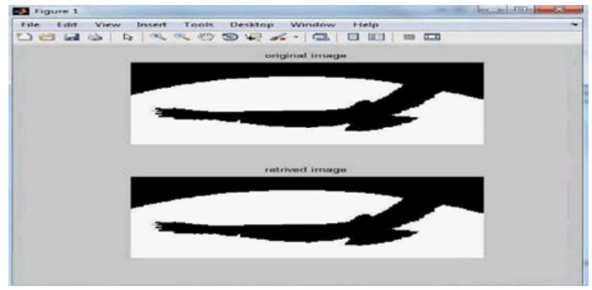
(a)



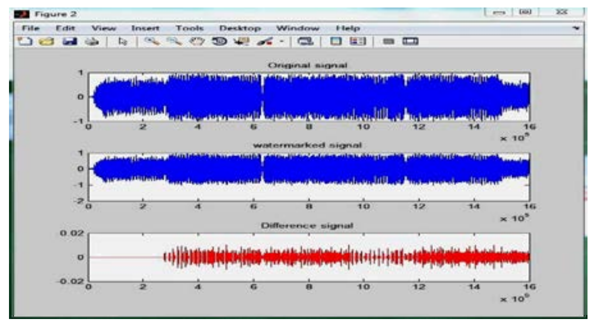
(b)

Fig. 5: (a) Original and retrieved image
(b) Original,watermarked and difference signals

are more in case of LSB substitution. To minimize the distortions and also keeping in mind the quality of watermarked audio LSB substitution has been performed in the third least significant bit. Even then the distortions can be seen in the Figure 4.



(a)



(b)

Fig. 6: (a) Original and retrieved image
(b) Original,watermarked and difference signals

Using DCT: It employs concept of DCT and LSB substitution. So the capacity is more, like in case of LSB substitution. The distortions in the retrieved watermark are more as shown in Figure 5.

Using DWT-Algorithm I: Figure 6 shows the simulation results using DWT-Algorithm I and this can be compared with DWT-Algorithm II.

Using DWT-Algorithm II: Figure 7 shows the simulation results of watermarking and dewatermarking using DWT-Algorithm II. Comparing algorithm-I and algorithm-II we can say that the capacity is more in algorithm-II. Hence we use a color image as watermark in this process whereas we use a grey scale image in algorithm-I as the capacity is more.

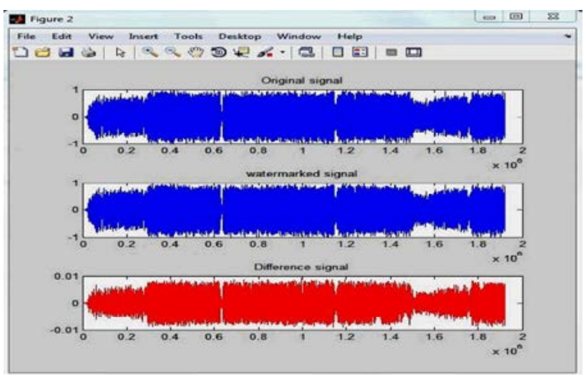
Comparing algorithm-II, LSB substitution and DCT we see that the distortions in the retrieved image are very less in algorithm-II when compared to the other two methods.

Using Echo Hiding: Figure 8 shows the simulation results of watermarking and dewatermarking using ehco hiding algorithm. This algorithm has very low capacity. Hence a binary image has been used as a watermark.

Table 1: SNR, MSE, PSNR comparison results

Algorithm used	SNR (dB)	MSE	PSNR (dB)	Capacity
LSB substitution	85.6367	1.358e-10	13.9194	1 bit/sample
DCT	77.9883	1.0484e-09	15.3834	1 bit/sample
DWT (Algorithm I)	45.3193	0	Infinite	1 bit/frame
DWT (Algorithm II)	43.5069	2.194e-06	28.0141	0.25 bits/sample
Echo hiding	47.9580	0	Infinite	1 bit/frame
High frequency tone	32.4248	0.0417	27.5973	1 bit/frame

(a)



(b)

Fig. 7: (a) Original and retrieved image
(b) Original,watermarked and difference signals

Using High Frequency Tone: Figure 8 shows the simulation results of watermarking and dewatermarking using ehco hiding algorithm. This algorithm has very low capacity. Hence a binary image has been used as a watermark.

The capacity is slightly more and hence a gray scale image can be used as a watermark. The results are tabulated and compared in Table 1. In the simulation results, the second column means the SNR of the watermarked audio signal with respect to the original audio signal. Also the results show the MSE of the retrieved image with respect to the watermark image and the PSNR of the retrieved image with respect to the watermark image.

(a)

(b)

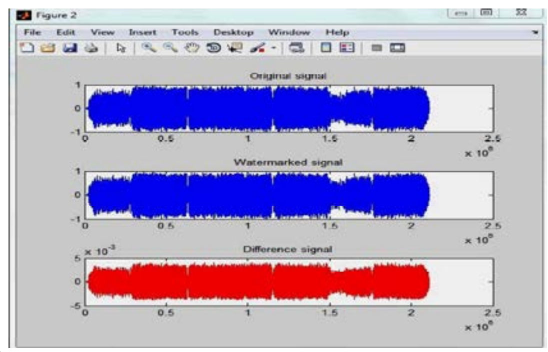
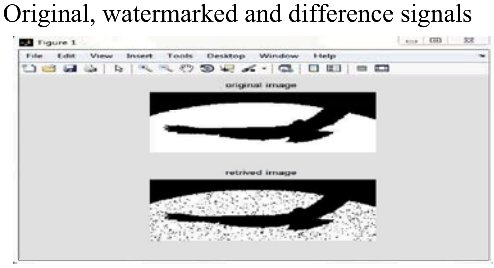


Fig. 8: (a)Original and retrieved image
(b) Original, watermarked and difference signals

(a)



(b)

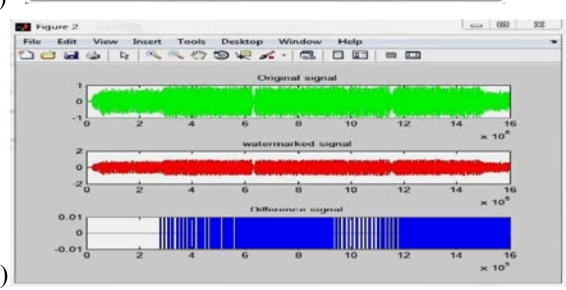


Fig. 9: (a)Original and retrieved image
(b) Original, watermarked and difference signals

CONCLUSIONS

The method that maintains a tradeoff between the quality of the watermarked audio and the retrieved image is more coveted and DWT algorithm-II is one such method. The capacity of DWT algorithm-II is more when compared to DWT algorithm-I. The capacity of algorithms using LSB substitution and DCT is more and hence a large sized image can be used as a watermark. In case of high frequency tone, the capacity is moderate and hence a grey scale image can be used. But for the algorithm using echo hiding, the capacity is very less and hence a binary image has been used for watermarking. The work can be further extended using hiding text with image, hiding image into an another image and hiding text into an audio signal using other transforms, like Discrete Hadamard Transform, Discrete Walsh Transform, etc.

REFERENCES

1. Chu, W.C., 2003. DCT-Based image watermarking using subsampling. IEEE Transactions on Multimedia, 5(1): 34-38.
2. Olanrewaju, R.F., Othman Khalifa and Khairun Nisaa' Abdul Latif, 2013. Computational Intelligence: It's Application in Digital Watermarking. Middle-East Journal of Scientific Research, 13(Mathematical Applications in Engineering): 25-30.
3. Surumbarkhuzhali, M. and S. Saravana, 2014. Verification of Parts Using Digital Image Processing in Machine Vision. World Applied Sciences Journal, Jan 2014.
4. Gopi E.S., 2007. Algorithm collections for DSP application using MATLAB. Springer Publishers.
5. Sajid Iqbal, Azhar Rauf, Saeed Mahfooz, Shah Khusro and Sajid H. Shah, 2012. Self-Constructing Fragile Watermark Algorithm for Relational Database Integrity Proof. World Applied Sciences Journal, 19(9): 1273-1277.
6. Mehdi Radmer, Seyer Mahmoud Anisheh, Mohsen Nikpour and Abbas Yaseri, 2011. Designing an Offline Method for Signature Recognition. World Applied Sciences Journal, 13(3): 438-443.
7. Harshavardhan Kayarkar, 2010. Information hiding in audio signals. International Journal of Computer Applications, 7(9): 14-19.
8. Jisna Antony, C.C. Sobin and A.P. Sherly, 2012. Audio steganography in wavelt domain – A survey. International Journal of Computer Applications, 52(13): 33-37.
9. Elamaran, V., K. Narasimhan, P.V.M. Vijayabhaskar and G. Shiva, 2013. Comparison of Wavelet Filters in Hybrid Domain Watermarking. Research Journal of Information Technology.
10. Elamaran, V., K. Narasimhan, G. Shiva and P.V.M. Vijayabhaskar, 2012. Comparison of Wavelet Filters in Image Coding and Denoising using Embedded Zerotree Wavelet Algorithm.Research Journal of Applied Sciences, Engineering and Technology, 4(24): 5449-5452.
11. Qidwai, U. and C.H. Chen, 2010. Digital Image Processing: An Algorithmic Approach with Matlab. CRC Press.
12. Asima Akber Abbasi, M.N.A. Khan and Sajid Ali Khan, 2013. A Critical Survey of Iris Based Recognition Systems. Middle-East Journal of Scientific Research, 15(5): 663-668.
13. Vafaei, M., H. Mahdavi-Nasab and H. Pourghassem, 2013. A New Robust Blind Watermarking Method Based on Neural Networks in Wavelet Transform Domain. World Applied Sciences Journal, 22(11): 1572-1580.