

# A NEW AUDIO WATERMARKING METHOD BASED ON DISCRETE COSINE TRANSFORM WITH A GRAY IMAGE

Mohammad Ibrahim Khan<sup>1</sup>, Md. Iqbal Hasan Sarker<sup>2</sup>, Kaushik Deb<sup>3</sup> and Md. Hasan Furhad<sup>4</sup>

<sup>1,2,3</sup>Department of Computer Science & Engineering, Chittagong University of Engineering & Technology, Chittagong-4349, Bangladesh.

<sup>4</sup>Department of Computer Engineering & Information Technology, EUCS Lab, University of Ulsan, South Korea.

<sup>1</sup>muhammad\_ikhan@cuet.ac.bd , <sup>2</sup>iqbal@cuet.ac.bd , <sup>3</sup>debkaushik99@cuet.ac.bd , <sup>4</sup>hfurhad@gmail.com

## ABSTRACT

*Many effective watermarking algorithms have been proposed and implemented for digital images and digital video. However, a few algorithms have been proposed for audio watermarking. This is due to the fact that, human auditory system (HAS) is far more complex and sensitive than human visual system (HVS). In this research work, a new method of embedding image data into the audio signal and additive audio watermarking algorithm based on Discrete Cosine Transformation (DCT) domain is proposed. First, the original audio is transformed into DCT domain. The DCT coefficients are divided into a fixed number of subsections and the energy of each subsection is calculated. Next, watermark is generated from image by image processing algorithm. Watermarks are then embedded into selected peaks of highest energy subsection. Experimental results demonstrate that the watermark is inaudible and this algorithm is robust to common operations of digital audio signal processing, such as noise addition, re-sampling, re-quantization and so on. To evaluate the performance of the proposed audio watermarking method, subjective and objective quality tests including Bit Error Rate (BER) and Signal to Noise ratio (SNR) are conducted.*

## KEYWORDS

*Audio watermarking, Discrete Cosine Transformation, Arnold Transformation, BER, SNR.*

## 1. INTRODUCTION

Along with the rapid growth of Internet, the transmission of audiovisual media becomes easier. It leads to the problems regarding copyright protection. Among them, copyright protection is the primary concern and the hotspot of international area in recent areas. Digital product information hiding and digital watermarking technology is generated on this basis and developed along with the protection of copyright; this technology is widely used in protecting the copyright of image, audio and video by means of extracting or detecting the watermark for its various application including copyright protection, broadcast and publication monitoring, authentication , copy control. Audio watermarking should satisfy the following features. They are as following:

- I. Imperceptibility: The digital watermark should not affect the quality of original audio signal after it is watermarked.
- II. Robustness: The embedded watermark data should not be removed or eliminated by unauthorized distributors using common signal processing operations and attacks.
- III. Capacity: Capacity refers to the numbers of bits that can be embedded into the audio signal within a unit of time.
- IV. Security: Security implies that the watermark can only be detectable by the authorized person.

In this paper, we propose a new watermarking technique using Discrete Cosine Transformation (DCT) for audio copyright protection. The watermarks are embedded into the selected peaks of the highest energy segment of DCT coefficients. Experimental results indicate that the proposed watermarking method provide strong robustness against several kinds of attacks such as noise addition, re-sampling, re-quantization and so on and achieves Bit Error Rate (BER) almost zero. In addition, our proposed method achieves SNR (signal-to-noise ratio) values ranging up to 29 dB.

The rest of the paper is organized as follows: Section 2 provides a brief description of previous works related to audio watermarking. Section 3 introduces traditional Arnold transformation. Section 4 introduces our proposed watermarking method including watermark embedding process and watermark detection process. Section 5 and 6 discusses the performance of our proposed method in terms of imperceptibility and robustness. Finally, section 7 concludes this paper.

## **2. RELATED RESEARCH**

A significant number of watermarking techniques have been reported in recent years in order to create robust and imperceptible audio watermarks. Lie et al. [1] propose a method of watermarking system for audio signals in the time domain. This proposed algorithm uses differential average-of-absolute-amplitude relations. In which samples are kept in group and each group of audio samples uses to represent one-bit information. This method also uses the low frequency amplitude modification technique so that it can scale the amplitudes in selected sections of samples. In [2], authors propose a blind audio watermarking method which inserts watermark into audio signal in the time domain. For watermark detection, the strength of the audio signal modifications is limited by the necessity for producing an output signal. The watermark is generated using a key. The watermark embedding also depends on the amplitude and frequency of audio signal so that it minimizes the audibility of the watermarked signal. In [3], authors explain a blind watermarking method which inserts watermarks into Discrete Cosine Transform (DCT) coefficients using quantization index modulation method. In [4], authors propose a blind audio watermarking method by using adaptive quantization against synchronization attack. In addition the multiresolution characteristics of discrete wavelet transform (DWT) and the energy compression characteristics of discrete cosine transform (DCT) are combined in this method. It also improves the transparency of digital watermark. Watermark is then inserted into low frequency components using adaptive quantization according to human auditory system (HAS). In [5], authors propose a watermarking method in cepstrum domain for audio signal in which a pseudo-random sequence is used as a watermark. The watermark is then weighted in the cepstrum domain according to the distribution of cepstrum coefficients and the frequency masking characteristics of human auditory system (HAS). In [6], authors propose a blind watermarking system which takes the advantages of the attack-invariant feature of the cepstrum domain and error-correction capability of BCH code to increase the imperceptibility and robustness of audio watermarking. In [7], authors propose a survey of watermarking systems

which are applied to Multimedia. In [8], watermarks are embedded into the highest  $m$  DCT coefficients of the whole sound excluding the DC component by the following equation:

$$v'_i = v_i(1 + \alpha x_i) \quad (1)$$

Where,  $m$  is the length of the watermark sequence,  $v_i$  is a magnitude coefficient into which a watermark is embedded,  $x_i$  is a watermark to be inserted into  $v_i$ ,  $\alpha$  is a scaling factor,  $v'_i$  is an adjusted magnitude coefficient. The watermark sequence is extracted by performing the inverse operation of (1) represented by the following equation:

$$x_i^* = \left(\frac{v_i^*}{v_i} - 1\right) / \alpha \quad (2)$$

In [9], the authors introduce a new method for inserting watermarks into audio signals by using components which have low frequency. This proposed method uses lifting-based wavelet transform (LBWT). The novel technology of inserting image data into the audio signal and additive audio watermarking algorithm based on DCT domain is proposed in [10]. In [11], the authors propose a watermarking method where DCT coefficients are used. In [12], the authors introduce a watermarking method which inserts synchronization signals in time domain. In [13] and [14], the authors propose a new watermarking system using Discrete Fourier Transform (DFT) and First Fourier Transform (FFT) of digital contents for copyright protection respectively. In [15], the authors propose a novel method to generate a unique personalized watermark which can act authenticating information for any watermarking algorithm. In [16], the authors describe the watermark construction process where scrambled version of watermark is obtained with the help of Arnold transform.

### 3. TRADITIONAL ARNOLD TRANSFORM

A digital image can be considered as a two unit function  $f(x,y)$  in the plane  $Z$ . It can be represented as  $Z = f(x,y)$  where  $x,y \in \{0,1,2,3,\dots,N-1\}$ . Hence,  $N$  represents order of digital image. The matrix of image can be changed into a new matrix by using the Arnold transform which results in a scrambled version to offer better security. It is a mapping function which changes a point  $(x,y)$  to another point  $(x', y')$  by using the equation (3) and (4).

$$x' = (x + y) \bmod N \quad (3)$$

$$y' = (x + 2y) \bmod N \quad (4)$$

If the original image iterative  $n$  times to come to some scrambled state, the scrambled image also need to iterative  $n$  times using new anti-Arnold transform to come to original image.

### 4. PROPOSED WATERMARKING METHOD

In this section, we propose our algorithm based on a Discrete Cosine Transform (DCT) on audio signal in which a watermark image is embedded. The algorithm consists of two procedures; watermark embedding procedure and watermark extraction procedure.

#### 4.1. Watermark Embedding Procedure

The embedding procedure performs three major operations; watermark pre-processing, DCT-based frequency decomposition of the audio signal and watermark embedding in the DCT-transformed audio signal. The operations are described in the following steps.

1. Express the gray-scale watermark image as a two-dimensional matrix whose size is  $M1 \times M2$ .

$$\text{Image} = \{ \text{Image}(k,j) , 0 \leq k < M1, \quad 0 \leq j < M2 \} \quad (5)$$

2. Convert the gray scale watermark image into binary format and also perform Arnold transform to scramble the watermark image and get scrambled image.
3. Convert two-dimensional image matrix into a one-dimensional vector  $W$  of length  $M1 \times M2$ .

$$W = \{ W(i) = \text{Image}(k,j), 0 \leq i < M1 \times M2 , 0 \leq k < M1, \quad 0 \leq j < M2 \} \quad (6)$$

4. Apply DCT to the original host signal to calculate the DCT coefficients.

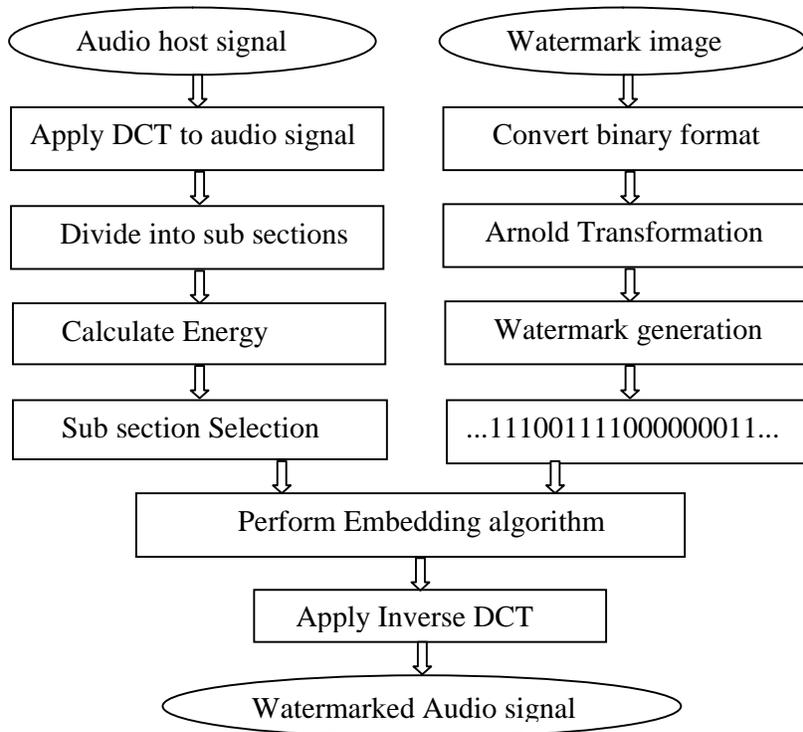


Figure 1: Watermark embedding process

5. Absolute values of the DCT coefficients are divided into subsections.
6. Energy of each section is then calculated using the following equation;

$$E = \sum_i |v(i)^2| \quad (7)$$

7. Find the most prominent peaks from the highest energy section using peak detection algorithm.
8. The watermark is then embedded into the selected N peaks of the selected section by using the equation (1) shown in [8], where N is the length of normalized one dimensional vector of image.
9. Apply the inverse Discrete Cosine Transformation (IDCT) to obtain the watermarked audio signal.

#### 4.2. Watermark Extraction Procedure

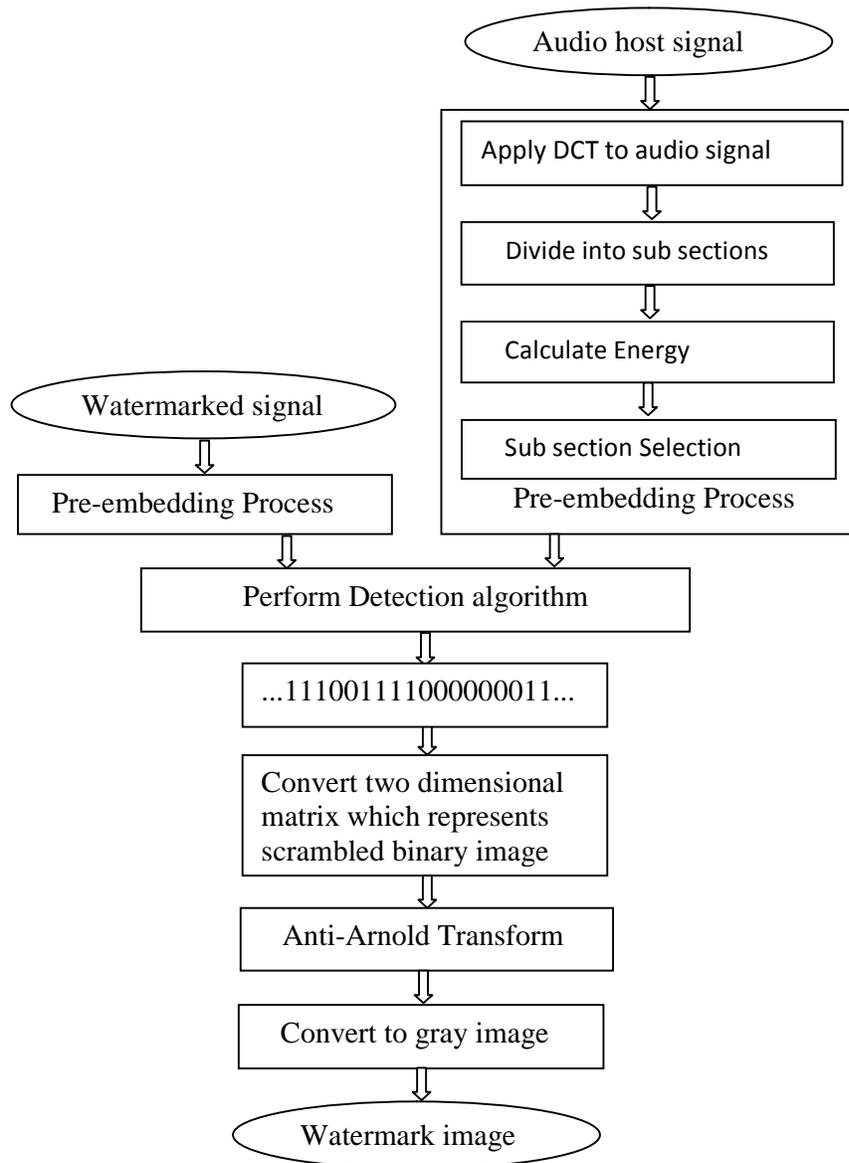


Figure 2: Watermark extraction process

The extraction process is implemented in the following steps:

1. The watermarked audio signal is transformed into DCT domain.
2. Extract the highest prominent peaks from the DCT coefficients which are located at the same position in the embedding process above.
3. The watermarked vector is then calculated by performing the extraction algorithm using the inverse operation of (1) represented by equation (2).
4. Convert extracted watermarked vector back into two dimensional matrix which represents scrambled image.
5. Apply anti-Arnold transform to get the original image of binary format.
6. Convert binary image to gray-scale image.

## 5. PERFORMANCE EVALUATION FRAMEWORK

In this section, we introduce performance evaluation of our proposed system.

### 5.1. Bit Error Rate (BER)

The robustness is measured as the BER corresponding to correctly formed binary values of the image. The BER (in present) is given by

$$BER = \frac{100}{k} \sum_{n=0}^{k-1} a \quad (8)$$

If ( $\bar{w}[n] = w[n]$ ) then  $a = 0$  otherwise  $a=1$ , where  $\bar{w}[n]$  represents the set of binary values formed from the watermarked audio host signal and  $w[n]$  indicates the set of the binary values of the original image.

### 5.2. Perceptual Quality

Perceptual quality refers to the imperceptibility of embedded watermark data within the host signal. In most applications it is important that the watermark is undetectable to a listener or a 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 paper, the signal-to-noise ratio (SNR) of the watermarked signal versus the host signal was used as a quality measure. In order to evaluate the quality of watermarked signal, the following signal-to-noise ratio (SNR) equation is used.

$$SNR = 10 \log_{10} \left\{ \frac{\sum_{n=1}^N S^2(n)}{\sum_{n=1}^N [S(n) - S'(n)]^2} \right\} \quad (9)$$

where  $S(n)$  is a host audio signal of length N samples and  $S'(n)$  be watermarked audio signal.

### 5.3. Subjective Quality Evaluation

The inaudibility of our watermarking method has been done by listening tests involving ten persons. Each listener was presented with the pairs of original signal and the watermarked signal and was asked to report whether any difference could be detected between the two signals. The ten people listed to each pair for 10 times and they have a grade for this pair, using the ITU-R BS.1284 standardized 5-point grading scale. The average grade for of each pair from all listeners is the final grade for this pair.

Table 1. Grading Scale (ITU-R)

| Grade | Quality   | Impairment                   |
|-------|-----------|------------------------------|
| 5     | Excellent | Imperceptible                |
| 4     | Good      | Perceptible but not annoying |
| 3     | Fair      | Slightly annoying            |
| 2     | Poor      | Annoying                     |
| 1     | Bad       | Very Annoying                |

## 6. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of our watermarking system for four different types of 16 bit mono audio signals sampled at 44.1 kHz: (a) an instrumental song ‘Hey Jude’ played by a Korean traditional musical instrument called the gayageum; (b) the song ‘Let it Be’ by the Beatles; (c) the beginning of Symphony No. 5 in C minor, Op.67, by Ludwig van Beethoven; (d) a human voice providing TOEIC (Test of English for International Communication) listening test instruction. A gray-scale image shown in figure 3 was used as watermark for embedding in the audio files. We choose  $15 \times 15$  gray-scale image as watermark image. The size of the watermark image was chosen after taking into consideration the length of the audio files



Figure 3: The original watermark image

Performance of audio watermarking algorithm is commonly evaluated with respect to two common metrics: (a) Imperceptibility (Inaudibility) and (b) Robustness.

### 6.1. Imperceptibility Test

Imperceptibility is related to the perceptual quality of the watermarked signal. It ensures that the quality of the signal is not perceptible to a listener. To measure imperceptibility, we use Signal-to-Noise ratio (SNR) as an objective measure and a listening test as a subjective measure.

Table 2. SNR and Average ITU-R grade value of the proposed watermarking system

| Type of Signal | Signal-to-Noise ratio (SNR) |                |                | Average Grade(ITU-R) |
|----------------|-----------------------------|----------------|----------------|----------------------|
|                | $\alpha = 0.1$              | $\alpha = 0.2$ | $\alpha = 0.3$ |                      |
| Hey Jude       | 26.06                       | 20.04          | 16.52          | 5.0                  |
| Let it be      | 26.69                       | 20.67          | 17.15          | 5.0                  |
| Symphony No 5  | 27.16                       | 21.14          | 17.62          | 5.0                  |
| Human Voice    | 29.00                       | 22.97          | 19.45          | 5.0                  |

### 6.2. Robustness Test

The robustness of a watermarking method is defined as the ability of watermark detector to extract the embedded watermark after common signal processing operations and attacks. In this paper, robustness is measured in terms of the Bit Error Rate (BER).

Table 3. Bit Error Rate (BER) of the proposed method without attack

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 0                    | 0              | 0              |
| Let it be      | 0                    | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

Table 4. Bit Error Rate (BER) of the proposed method against re-sampling

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 0                    | 0              | 0              |
| Let it be      | 0                    | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

Table 5. Bit Error Rate (BER) of the proposed method against re-quantization

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 0                    | 0              | 0              |
| Let it be      | 0                    | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

Table 6. Bit Error Rate (BER) of the proposed method against filter attacks

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 2.78                 | 0              | 0              |
| Let it be      | 1.39                 | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

Table 7. Bit Error Rate (BER) of the proposed method against random noise

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 0                    | 0              | 0              |
| Let it be      | 0                    | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

Table 8. Bit Error Rate (BER) of the proposed method against AWGN

| Type of Signal | Bit Error Rate (BER) |                |                |
|----------------|----------------------|----------------|----------------|
|                | $\alpha = 0.1$       | $\alpha = 0.2$ | $\alpha = 0.3$ |
| Hey Jude       | 0                    | 0              | 0              |
| Let it be      | 0                    | 0              | 0              |
| Symphony No 5  | 0                    | 0              | 0              |
| Human Voice    | 0                    | 0              | 0              |

## 7. CONCLUSIONS

In this paper, we have presented a new watermarking system using Discrete Cosine Transform (DCT) for copyright protection of audio sound. Experimental results indicate that our proposed watermarking system has compromised audibility and robustness better. This method also shows strong robustness against several kinds of attacks such as noise addition, re-sampling, re-quantization, filtering etc. This results obtained verify the effectiveness of audio watermarking as a reliable solution to the copyright protection problem which is facing the music industry.

## REFERENCES

- [1] W. N. Lie and L. C. Chang, (2006) "Robust and High-Quality Time-Domain Audio Watermarking Based on Low Frequency Amplitude Modification", IEEE Transactions on Multimedia, Vol. 8, No. 1, pp 46-59.
- [2] P. Bassia, I. Pitas and N. Nikolaidis, (2001) "Robust Audio Watermarking in the Time domain", IEEE Transactions on Multimedia, Vol. 3, No. 2, pp 232-241.
- [3] G. Zeng and Z. Qiu, (2008) "Audio watermarking in DCT; Embedding Strategy and Algorithm", in Proceedings of 9th International Conference on Signal Processing, (ISP'09), pp 2193-2196.
- [4] X. Y. Wang and H. Zaho, (2006) "A novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT ", IEEE Transactions on Signal Processing, Vol. 54, No. 12, pp 4835-4840.
- [5] S. K. Lee and Y. S. Ho, (2000) "Digital Audio Watermarking in the cepstrum domain", IEEE Transactions on Consumer Electronics, Vol. 46, No. 3, pp 744-750
- [6] S. C. Liu and S. D. Lin (2006) "BCH code based robust audio watermarking in the cepstrum domain", Journal of Information Science and Engineering, Vol. 22, pp 535-543.
- [7] S.J. Lee, S. H. Jung (2001) "A Survey of Watermarking Techniques Applied to Multimedia", IEEE Transactions on Industrial Electronics, Vol. 12, pp 272-277.
- [8] I. Cox, J. Killian, F. Leighton and T. Ahamoon, (1997) "Secure Spread Spectrum Watermarking for Multimedia", IEEE Transactions on Image Processing, Vol. 6, No. 12, pp 1673-1687.
- [9] E. Ercelesi. And L.Batakci., (2009) "Audio Watermarking Scheme Based on Embedding Strategy in Low Frequency Components with a Binary Image.", Digital Signal Processing, Vol. 19, No. 2, pp 265-277.
- [10] Yan Yang, Rong Huang and Mintau Xu (2009) "A Novel Audio Watermarking Algorithm for Copyright Protection Based on DCT Domain." 2nd International Symposium on Electronic Commerce and Security ,pp 184-188.
- [11] P.K Dhar, M.I.Khan and S. Ahmad (2010) "A New DCT-Based Watermarking Method for Copyright Protection of Digital Audio.", International Journal of Computer Science & Information Technology(IJCSIT), Vol. 2 No.5.
- [12] J. Huang, Y. Wang and Y.Q. Shi,(2002) "A Blind Audio Watermarking Algorithm with Self Synchronization.", IEEE International Symposium on Circuit and Systems, Vol. 3 pp 627-630.
- [13] P.K Dhar, M.I.Khan and J.M.Kim (2010) "A New Audio Watermarking System using Discrete Fourier Transform for Copyright Protection.", International Journal of Computer Science & Network Security (IJCSNS), Vol. 10 No.6.
- [14] P.K Dhar and J.M.Kim (2011) "Digital Watermarking Scheme Based on First Fourier Transformation for Audio Copyright Protection.", International Journal of Security and its Applications (IJSIA), Vol. 5 No.2.
- [15] T.K. Tewari , V.Saxena and J.P Gupta (2012) "A Novel Approach to Generate Watermarks Using Auditory Features for Audio Watermarking.", Journal of Theoretical and Applied Information Technology (JATIT), Vol. 35 No.2.
- [16] Dr. M. Sathik and S. S. Sujatha (2010) "An Improved Invisible Watermarking Technique for Image Authentication.", International Journal of Advanced Science and Technology (IJAST), Vol. 24.

## Authors

**Dr. Mohammad Ibrahim Khan** received the B.S. degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Bangladesh in 1999. He received M.S. degree in Computer Science and Engineering from the same University in 2002. He received his Ph.D. degree in Computer Science and Engineering from Jahangirnagar University in 2010. Since 1999, he has been serving as a faculty member in the Department of Computer Science and Engineering at Chittagong University of Engineering & Technology (CUET), Chittagong, Bangladesh. His research interest includes Digital Image Processing, Graph Theory, Cryptography, Digital Watermarking, Multimedia Systems, and Digital Signal Processing.



**Md. Iqbal Hasan Sarker** received the B.Sc degree in Computer Science & Engineering from Chittagong University of Engineering & Technology (CUET), Bangladesh in 2009. Currently, he is pursuing M.Sc in Computer Science & Engineering in the same University. Since 2010, he has been serving as a faculty member in the same department and university. His research interest includes Digital Watermarking, Computer Vision, Digital Image Processing, Cryptography and data mining.



**Dr. Kaushik Deb** received his B. Tech. and M. Tech. degrees from Department of Computer Science & Engineering of Tula State University, Tula, Russia, in 1999 and 2000, respectively. He received his Ph.D. degree in Electrical Engineering & Information systems from University of Ulsan, Ulsan, South Korea in 2011. Since 2001, he has been serving as a faculty member in the Department of Computer Science and Engineering of Chittagong University of Engineering & Technology (CUET), Chittagong, Bangladesh. His research interests include computer vision, pattern recognition, intelligent transportation systems (ITSS) and human-computer interaction.



**Md. Hasan Furhad** received his B. Sc degree in Computer Science & Engineering from Rajshahi University of Engineering & Technology (RUET), Bangladesh in 2008. Currently, he is pursuing MS in University of Ulsan, South Korea. He is a Lecturer in Faculty of Engineering at University of Information Technology & Sciences (UITS), Chittagong, Bangladesh. His research interests include Embedded Systems, Computer Architecture and Computer Vision.

