ISSN No. 0976-5697

Volume 7, No. 4, July-August 2016

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Steganographic Approach to Hide Information in Audio Signal using Discrete Wavelet Transforms

Krishna Bhowal
Department of Engineering Technological Studies
University of Kalyni
Kolkata, India

Debasree Chanda (Sarkar)
Department of Engineering Technological Studies
University of Kalyni
Kolkata, India

Susanta Biswas
Department of Engineering Technological Studies
University of Kalyni
Kolkata, India

Partha Pratim Sarkar
Department of Engineering Technological Studies
University of Kalyni
Kolkata, India

Abstract: This paper presents an effective data hiding technique to embed a secret image and extract them in a bit-exact manner by altering the magnitudes of the coefficient of Discrete Wavelet Transform (DWT) of audio signals. The basic concept is that the Discrete Wavelet Transform (DWT) coefficients are divided into more details frequencies and change the magnitude of the selected DWT coefficients using XOR operator. Taking advantage of XOR operation, it is possible to change the coefficients of DWT without compromising the statistical property of the Audio signals. In fact, the proposed technique confirms that the maximum change is less than 6.25% of the related DWT coefficients. The experimental results show that the technique has a high embedding capacity, without compromising the perceptual quality of the audio signals.

Keywords: Steganography, Audio Signal, Digital Wavelet Transform (DWT), Imperceptibility, Information Security, XOR operator

I. INTRODUCION

In the current digital age, with the fast growth of various communication techniques, transmitting digital media becomes more and more essential. This trend motivates industry professionals and research communities to pay a precise attention to data security. Currently, there are three main methods are used to protect data, namely, cryptography, watermarking and steganography. Cryptography techniques are based on transcription the content of a message mangled to unauthorized users or parties. Confidential data are hidden to transfer secret information through the cover medium such as ownership information and copyright information in watermarking. The main goal of steganography is to transfer hidden information secretly so that the unauthorized people cannot identify its presence. Steganography is considered a sub-discipline of security domain in data communication. Modern steganography techniques exploit the characteristics of digital media and use the multimedia content as covers to embed secret information. The multimedia contents or covers used in steganography can be of different types like text, image, audio, video, and IP datagram. Currently, to achieve better data security, steganography techniques are combined with conventional security techniques.

To minimize the dissimilarity between the cover-medium and the stego-medium, modern steganography techniques exploit natural limitations in the human auditory system and visual system. The audio steganography system exploits the masking effect of the Human Auditory System (HAS). Digital audio Steganography is a process where secret data are hidden or embedded into cover audio signals. These inserted data can later be extracted from the stego-audio signal for various purposes. There are several applications of audio Steganography including covert communication, content authentication, copyright protection, etc.

An audio Steganography system has different properties and the following basic requirements must satisfy all the applications:

Imperceptibility or inaudibility: The quality of the embedded audio or stego-audio should be retained after embedding the secret message. Imperceptibility can be measured using Objective Difference Grade (ODG), Subjective Difference Grade (SDG) and Signal-to-Noise Ratio (SNR).

Embedding Capacity: The amount of secret data that can be embedded into the cover audio signal without compromising the imperceptibility is called embedding capacity of a Steganography method. For audio Steganography, embedding rate refers to the number of secret information bits that may be embedded reliably within a host audio signal per unit of time. Generally, embedding capacity is measured in bits per second (bps).

Security: Embedded audio signals should not reveal any clues about the secret message are hidden in them. Moreover, the security of the steganography method must depend on secret keys as well as on the Steganography algorithm applied. Sometimes, Steganography and Cryptography are combined to increase the security of the Steganography method.

Robustness: Robustness is a very important requirement in data hiding technique. The robustness are measured by the ability to extract secret information from an embedded audio signal after various signal processing operations or malicious attacks.

The quality of an audio Steganographic method may be measured by various features of data hiding technique. The importance and the impact of each feature depend on the application of the system and the transmission environments. Normally, the imperceptibility and embedding capacity hardly coexist in the same Steganography system, because there is a tradeoff between these two requirements, where increased imperceptibility result in decreasing data hiding capacity [1].

The steganography methods can be categorized as follows:
Time domain substitution: Here only the least significant bits (LSBs) of the cover media is replaced by the secret message bits without significant change in the cover media.
Transform domain substitution: Here various transform domains like First Fourier Transform (FFT), Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are used to hide secret information on transformed coefficients of the cover media.
Spread Spectrum: Here the secret information is spread over a wide frequency bandwidth than the minimum required bandwidth to transmit the confidential information.
Statistical: Here the cover object is divided into different blocks and the secret message bits is hidden in each block.
Distortion: Here the secret information is hidden by making distortion in the signal. The encoder adds a sequence of modification to the cover object and the decoder checks for the deviation made between the original cover object and the distorted cover object to recover the secret information.

The LSB methods are not so robust against simple signal processing techniques, like compression, transform, etc. For this reason, we focus on transformed domain method in this study. However, this method has its advantages and disadvantages. For example, in frequency or wavelet transform domains, the capacity is higher. But, it has problem in extracting the message from the cover, due to the generated errors that result in low quality. To overcome this problem, we proposed a modified version of the substitution technique in the wavelet domain, in which the error is minimized and high quality in terms of signal to noise ratio (SNR) and Objective Difference Grade (ODG) is achieved.

A. Wavelet Transform

In this section, we briefly introduce the discrete wavelet transform (DWT). The DWT is a discipline capable of giving time frequency representation of the signal. Staring from the original audio Signal S, DWT produces two sets of coefficients. The approximation coefficients A (low frequencies) and the detail coefficients D (high frequencies). Depending on the application and also the duration of signal, the low frequency part might be further decomposed into two parts of high and low frequencies. The one level DWT decomposition is presented in figure 1.

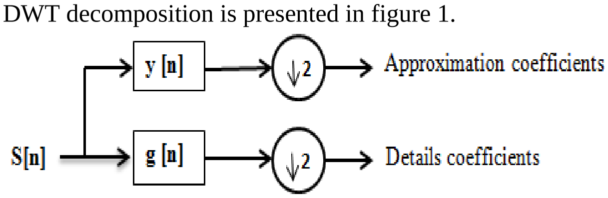

Figure 1: One-level DWT decomposition

Figure 2 shows a 3-level DWT decomposition of signal S, and the output will be 4 sub bands. Each of these belongs to a particular resolution level of wavelet transform. A complete structure of wavelet contains domain processing analysis block and a synthesis block. Analysis or decomposition block decomposes the signal into wavelet coefficients. The reconstruction process is the inverse of the decomposition process. The original signal S can be reconstructed using the inverse DWT process.

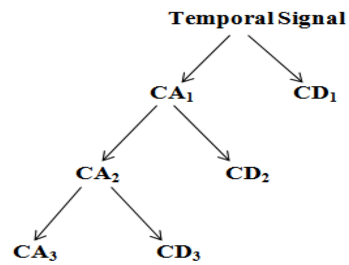

Figure 2: Three-level DWT decomposition

II. RELATED WORKS

Different methods have been proposed for hiding secret information in digital audio. Most of these methods make use of HAS weaknesses.
Many researches are based on LSB method [2-4]. When this method is used along with wavelets [3] or with a minimum error replacement technique [4], higher quality and capacity of steganography are achieved. Comparing LSB technique in time domain with wavelet and Fourier domains shows that more bits can be hidden in frequency domain for similar SNR conditions. In this way, a high bit rate LSB audio data hiding method, that reduced embedding distortion of the host audio, has been proposed in [4]. This paper proposed a two-step algorithm; in which the hidden bits have been embedded into the higher LSB layers, resulting in increased robustness against noise addition. The key idea of the algorithm is watermark bit embedding that causes minimal embedding distortion of the host audio.

In [3], a method for digital audio steganography, which encrypted covert data is embedded into the wavelet coefficients of the host audio signal, has been proposed. To avoid the extraction error, lifting wavelet transform has been used. For using the maximum capacity of audio signals, hearing threshold has been calculated in the wavelet domain. Then, according to this threshold, data bit has been embedded in the least significant bits of lifting wavelet coefficients. The Inverse lifting wavelet transform has been applied to modified coefficients to construct stego signal in time domain.

High capacity and security steganography using discrete wavelet transform has been proposed in [5]. The wavelet coefficients of both the covert and secret message have been fused into a single image. The covert and secret message has been preprocessed to reduce the pixel range to ensure the accurate recovery of the message at the destination. It has been observed that the capacity and security increased with acceptable peak SNR in the proposed algorithm as compared to the existing algorithms. As noted earlier, the capacity of the classic LSB insertion method can be increased by performing the embedding process in the wavelet domain. In this way, in [6] an algorithm has been proposed which uses perfect reconstruction filter banks and embedded additional information in the wavelet domain of the audio signal by modifying LSB of wavelet coefficients.

As a hybrid approach, simultaneous low bit rate encoding and information hiding for highly compressed audio signals has been reported in [7]. The tests with an extended MPEG4 advanced audio coding encoder confirm the robustness of the method. In this way, hiding the text message in wave files in a frequency domain format using FFT techniques has been reported in [8].

In the frequency-domain steganography [9]–[14], after applying one of the usual transforms such as the Discrete/Fast Fourier Transform (DFT/FFT) [12], [13], the Modified Discrete Cosine Transform (MDCT) or the Wavelet Transform (WT) from the signal [15] the hidden bits are embedded into the resulting transform coefficients.

In frequency-domain methods, the Fourier transform (FT) is very popular. Among different Fourier transforms, the Fast Fourier transform (FFT) is often used due to its reduced computational burden. This transform is also used by different authors, such as in [8], which proposes a multi-bit spread-spectrum audio watermarking scheme based on a geometric invariant logs coordinate mapping (LCM) feature. The watermark is embedded in the LCM feature, but it is actually embedded in the Fourier coefficients which are mapped to the feature via LCM.

In the algorithm proposed in this paper, we select a major part of the frequency of the DWT spectrum for embedding the secret bits. The selected frequency bands are divided into short subbands and a single secret bit is embedded into a selected subband.

III. PROPOSED WORK

Extensive effort has been invested over the years in understanding the features of the human auditory system and applied the relevant knowledge in audio steganography. First, we use discrete wavelet transforms in such a way that CA_3, CD_3, CD_2 and CD_1 coefficients are generated from the audio signal. The frequencies that form an audio signal range from 20Hz to 20 kHz. When wavelet transform is applied to the audio signal, it is divided into separate frequency segments by using high pass and low pass filters.

The first subband has the largest frequency range and most of the signal energy. The frequencies which are perceivable by human ear with low sound pressure level are in the range of 20Hz to 2 kHz. In this work, we called this subband1. The other three subbands are placed in 2 kHz to 20 kHz frequency range. Hearing sensitivity decreases at higher frequencies and lower frequencies. However, it increases more at higher than lower frequencies. Therefore, it is very much clear that, if secret data is embedded in the high frequency band the alteration will be mostly inaudible and therefore more transparency will be achieved. For this reason, the higher frequency subbands are used to hide secret information in this work.

In the proposed Steganographic scheme, the following algorithm is used to embed a secret image in term of bit stream into the DWT coefficients. The frequency of subband and the secret bit embedding position are two parameters that set the properties of the proposed Steganographic scheme. The selected frequency band is divided into subbands then each secret bit of the image is embedded into the best coefficients of three subbands, which makes the method more secure against attacks. The best coefficient is identified by the minimum number of alterations required during the embedding process.

A. Embedding the Secret Image:

This work presents a proficient method to achieve high quality stego-audio. Some very interesting properties of the XOR operator ($\oplus$) are convinced us to use it in this work. It is possible to embed secret bits with negligible changes in cover audio using XOR operator. On the other hand, it consumes lower embedding computational complexity.

Considering that all the DWT coefficients have 16-bit representation. Wavelet transforming up to the third level has four subbands and three subbands are used for steganography.

The basic embedding procedure is conducted by equation (1) as like below-

B(i) $\oplus$ LSB = Image-bit ……… (1)

Where B(i) means bit in position i, i = 1 to 16, of DWT coefficients. LSB means the least significant bit of DWT coefficient. Image-bit means secret image bit.

To embed secret bit using equation (1), following are the basic requirements:-

If Image-bit = 0 then LSB = 0 and B(i) = 0 for i = 2 to 16
If Image-bit = 0 then LSB = 1 and B(i) = 1 for i = 2 to 16
If Image-bit = 1 then LSB = 0 and B(i) = 1 for i = 2 to 16
If Image-bit = 1 then LSB = 1 and B(i) = 0 for i = 2 to 16

Now, there two options to fulfill the above requirements. First, we can change the bit of the LSB as required and the change will be negligible. The value of B(i) will remain unchanged. Second, we can change the bit of B(i) as required and change will not be negligible, i.e., this can increase the distortion in the audio signal. The value of the LSB will remain unchanged. On the other hand, we can search a suitable i, so that the change in B(i) is not required to fulfill the above requirement. In this case, positions i need to be sent at the receiving end to get the secret bit. In this work, we choose the first option, i.e., we will make change the bit of LSB if required to embed secret bit and is maintained the quality of the audio signal. The value of B(i) remains unchanged.

Following is the embedding algorithm:-

Let, intB = int (B(i)), i = 1 to 16, be the integer value of B(i).

First, calculate loc = intB mod 16.

If Image-bit = 0 and B(loc) = 0 and LSB = 0,
    no change is required in LSB.
If Image-bit = 0 and B(loc) = 0 and LSB = 1,
    alter LSB bit to 0 and update loc = (loc + 1).
If Image-bit = 0 and B(loc) = 1 and LSB = 1,
    no change is required in LSB.
If Image-bit = 0 and B(loc) = 1 and LSB = 0,
    alter LSB bit to 1 and update loc = (loc - 1).
If Image-bit = 1 and B(loc) = 0 and LSB = 1,
    no change is required in LSB.

If Image-bit = 1 and B(loc) = 0 and LSB = 0,
alter LSB bit to 1 and update loc = (loc - 1)
If Image-bit = 1 and B(loc) = 1 and LSB = 0,
no change is required in LSB.
If Image-bit = 1 and B(loc) = 1 and LSB = 1,
alter LSB bit to 0 and update loc = (loc + 1)
Here loc is embedded in the steganography key which is very important to extract the hidden image at the receiving end.

For embedding the image bit, first the DWT is applied to the audio signal. Then, the DWT coefficients are modified by secret bit based on the above mentioned algorithm. Finally the inverse DWT is applied to generate the stego-audio signal. Figure 3 provides the operational flowchart of the proposed embedding algorithm.

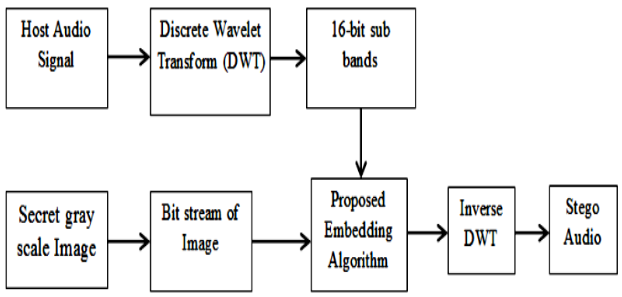

Figure 3: Operational Flow chart of embedding operation

B. *Extracting the Secret Image:*

The original audio signal is not required to extract the secret bit from stego audio signal. So, the extraction process is blind. The extraction process can be summarized as like below:
Apply the DWT to compute the DWT coefficients of the stego audio signal.
Convert DWT coefficients to 16-bit stream.
Then apply the following extracting algorithm:
Let, intB = int (B(i)), i = 1 to 16, be the integer value of B(i).
Extract loc from Steganographic key.
Calculate Message-bit = B(loc) $\oplus$ LSB
Figure 4 provides the operational flowchart of the proposed extracting algorithm.

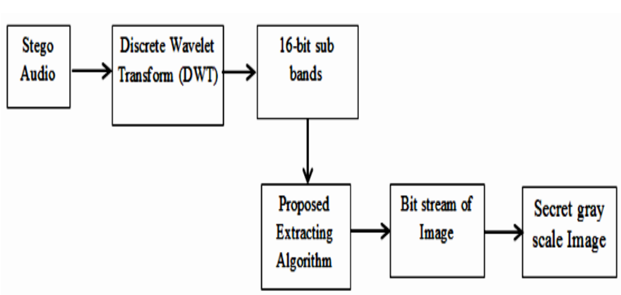

Figure 4: Operational Flow chart of extraction operation

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In order to measure the performance of our proposed method, an image, Image.jpg is embedded on 10 different types of audio clips using the proposed data hiding algorithm. Here, country, classic, pop, rock, jazz, etc., audio clips is utilized as a host audio. Audio clips are 44.1 kHz with 16 bits samples and mono types.
In this section, some experimental results are presented in terms of imperceptibility, capacity and security of our proposed scheme.

A. *Measurement of similarity between original and embedded Audio through Correlation*

The linear correlation coefficient is one of the most consistent measurements of similarity between two sets of quantities. Suppose, there is a sequence of n original audio sample X and a sequence of n stego audio samples Y. The sample correlation coefficient r can be calculated using equation 2, between X and Y.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} \quad \ldots\ldots\ldots(2)$$

Where $\bar{x}$ is the original audio sample mean of X and $\bar{y}$ is the stego audio sample means of Y, S_x and S_y are the sample standard deviations of X and Y. The correlation coefficients are calculated for ten types of audio clips in MATLAB. The value of r is about 1 for all audio clips. This confirms their absolute similarity.

B. *Subjective and Objective Audio Quality Measurements*

Signal-to-Noise Ratio and Objective Difference Grade are two relevant measurements to test the imperceptibility of an audio signal. ODG is a suitable measurement for audio distortions. It is assumed to provide a specific model of the Subjective Difference Grade (SDG) outcomes which may be originating from a group of skilled listeners. ODG = 0 means no degradation occurred in embedded audio signal and ODG = -4 means a very annoying distortion occurred in embedded audio. The ODG calculations are done using the advanced ITU-R BS.1387 standard by Thiede et al [17] and implemented by the software tool EAQUAL by A. Lerch [18]. ODG values of the embedded audio are between -0.11 and -0.47 that concludes their good qualities. The ODG values for different embedded audio are reported in Table 1.

Subjective quality measurements defined in [22], [23] are performed to measure the inaudibility of the proposed data hiding technique. Ten observers were designated for these subjective listening tests, all the participants were experts in music. They are asked to report regarding the dissimilarity between the original and stego audio signals, using five-points SDG: (5: imperceptible, 4: perceptible but not annoying, 3: slightly annoying, 2: annoying 1: very annoying). The decision of the subjective tests is an average of the quality assessments called a Mean Opinion Score (MOS). The result, reported in table 1, approves that suitable imperceptibility of the secret image in the stego audio signal is achieved.

C. *Audio Quality Evaluation by Signal-to-Noise Ratio (SNR)*

Normally, if the SNR value is higher than 50 dB, then the secret information which is hidden in the cover media are imperceptible to the human auditory system (HAS). The original signal (the host audio) is denoted x(i), i = 1 to N while the stego-audio (the embedded audio) as y(i), i = 1 to

N. The SNR is calculated using equation no. (3) and reported in Table 1 for 10 types of audio clips.

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^{N} x^2(i)}{\sum_{i=1}^{N} (x(i) - y(i))^2} \ldots\ldots (3)$$

Table 1. SDG and SNR value comparison between different audio types

Audio types	SDG	ODG	SNR(dB)
A1	5.00	-0.12	93.35
A2	4.97	-0.47	92.51
A3	4.98	-0.35	92.74
A4	5.00	-0.12	93.44
A5	5.00	-0.11	93.52
A6	5.00	-0.16	93.16
A7	4.99	-0.38	92.78
A8	4.97	-0.36	92.66
A9	5.00	-0.17	93.28
A10	5.00	-0.19	93.08

The Subjective Difference Grade (SDG) and SNR values for different audio clips are reported in Table 1. For simplicity, 10 audio clips are denoted by A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9 and A_{10}.

D. Security

In this work, the secret bit is embedded almost at random position of one of the three coefficients as described in the above algorithm. It is very difficult for an adversary to detect the position where the actual secret bit is inserted. Along with this, there are several cryptographic methods that can be used to improve the security level of the proposed system. A cryptographic method should be selected based on the requirements of the steganography system. The AES encryption method is a good choice in terms of computational complexity.

E. Comparison

The suggested method has been compared with some latest and relevant audio steganography and watermarking methods. Every method has different properties and characteristics and it is very challenging to find an impartial comparison of the proposed method with some audio steganography or watermarking methods. Here, few recent and relevant audio steganography and watermarking methods are chosen for making comparisons. Table 2 reports a comparison between the proposed steganography algorithm and several recent audio steganography and watermarking algorithms. The comparisons confirm that the improvement in both capacity and imperceptibility of the proposed method with respect to other methods in the literature. In brief, the proposed method succeeds higher embedding capacity and inaudibility.

Table 2: Comparison of Different Steganography and Watermarking Algorithms

Methods	Capacity (bps)	SNR (dB)	ODG
[9]	2	42.8 to 44.4	-1.66 <ODG<-1.88
[10]	4.3	29.5	Not reported
[12]	3k	30.55	-0.6
[13]	2k-6k	Not reported	-0.6 < ODG <-1.7
[15]	11 k	30	-0.7
[16]	64	30-45	-1< ODG
[19]	4-512	Not reported	-1 < ODG
[21]	8	Not reported	-3 < ODG < -1
Proposed	44100	93.52	-0.11 < ODG < -0.47

V. CONCLUSION AND FUTURE WORKS

In this paper, a high-capacity imperceptible steganography method for digital audio is presented. Secret image-bit is embedded in the DWT coefficient of audio signals using a new proposed algorithm to improve the security of the steganography system. The proposed method confirms that the maximum alteration of each DWT coefficient is less than 6.25%. Furthermore, the proposed method is blind, as it does not require the original audio signals for extracting the embedded bits. The experimental results show that suggested method has a high embedding capacity (44.1 kbps) without major perceptual distortion (ODG is -0.11 to -0.47).

Furthermore, there is some scope to enhance the algorithm by modifying the existing algorithm. To improve the robustness of the suggested method, secret bit can be embedded in all three numbers of DWT coefficients. Also, we can increase the level of DWT transform to improve robustness.

VI. REFERENCES

[1] W. Bender, D. Gruhl, N. Morimoto, A. Lu, “Techniques for Data Hiding”, IBM Syst. J. **35**(3 and 4), pp. 313–336 ,1996.

[2] K. Gopalan, "Audio steganography using bit modification", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '03), vol. 2, 6-10 April 2003, pp. 421-424.

[3] M. Pooyan, A. Delforouzi,. “LSB-Based Audio Steganography Method Based on Lifting Wavelet Transform”. In the Proceedings of IEEE International Symposium on Signal Processing and Information Technol., 2007, pp: 600-603.

[4] N. Cvejic, T. Seppanen, “Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method”, In the Proceedings of IEEE International Conference on Information Technology: Coding and Computing, 2:pp. 533-537, 2004.

[5] M. Reddy, H.S. and K.B. Raja, “High Capacity and Security Steganography Using Discrete Wavelet Transform. International”, J. Computer Science and Security, 3: pp. 462-472. 2009.

[6] N. Cvejic, T. Seppanen, “Increasing Capacity of LSB-Based Audio Steganography”. In the Proceedings of IEEE Workshop on Multimedia Signal Processing, pp: 336-338. 2002.

[7] Xu, S., P. Zhang, P. Wang and H. Yang, “Performance Analysis of Data Hiding in MPEG-4 AAC Audio”. Tsinghua Sci. and Technol., 14: pp- 55-61. 2009.

[8] V. Viswanathan, “Information hiding in wave files through frequency domain”. Appl. Mathematics and Computation. 2001. pp- 121-127

[9] S. Xiang, H. J. Kim, and J. Huang, “Audio watermarking robust against time-scale modification and mp3 compression,” Signal Process., vol. 88, no. 10, pp. 2372–2387, Oct. 2008.

[10] M. Mansour and A. Tewfik, “Data embedding in audio using time-scale modification,” IEEE Trans. Speech Audio Process., vol. 13, no. 3, pp. 432–440, May 2005.

[11] J. J. Garcia-Hernandez, M. Nakano-Miyatake, and H. Perez-Meana, “Data hiding in audio signal using rational dither modulation,” IEICE Electron. Express, vol. 5, no. 7, pp. 217–222, 2008.

[12] M. Fallahpour and D. Megías, “High capacity audio watermarking using fft amplitude interpolation,” IEICE Electron. Express, vol. 6, no. 14, pp. 1057–1063, 2009.

[13] M. Fallahpour and D. Megías, “High capacity method for real-time audio data hiding using the fft transform,” in Advances in Information Security and Its Application. Berlin, Germany: Springer-Verlag, 2009, pp. 91–97.

[14] W. Li and X. Xue, “Content based localized robust audio watermarking robust against time scale modification,” IEEE Trans. Multimedia, vol. 8, no. 1, pp. 60–69, Feb. 2006.

[15] M. Fallahpour and D. Megías, “High capacity audio watermarking using the high frequency band of the wavelet domain,” in Multimedia Tools and Applications. New York, NY, USA: Springer, 2011, vol. 52, pp. 485–498.

[16] X. Kang, R. Yang, and J. Huang, “Geometric invariant audio watermarking based on an LCM feature,” IEEE Trans. Multimedia, vol. 13, no. 2, pp. 181–190, Apr. 2011.

[17] T. Thiede, W. C. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. G. Beerens, C. Colomes, M. Keyhl, G. Stoll, K. Brandenburg, and B. Feiten, “PEAQ - The ITU standard for objective measurement of perceived audio quality,” J. AES, vol. 48, no. 1/2, pp. 3–29, 2000.

[18] Lerch, A. Zplane development, EAQUAL- Evaluate Audio Quality, version:0.1.3 alpha, (2002). http://www.mp3-tech.org/programmer/misc.html.

[19] M. Unoki and D. Hamada, “Method of digital-audio watermarking based on cochlear delay characteristics,” Int. J. Innovat. Comput., Inf. Control, vol. 6, no. 3(B), pp. 1325–1346, Mar. 2010.

[20] K. Kondo and K. Nakagawa, “A digital watermark for stereo audio signals using variable inter-channel delay in high-frequency bands and its evaluation,” Int. J. Innovat. Comput., Inf. Control, vol. 6, no. 3(B), pp. 1209–1220, Mar. 2010.

[21] A. Nishimura, “Audio data hiding that is robust with respect to aerial transmission and speech codecs,” Int. J. Innovat. Comput., Inf. Control, vol. 6, no. 3(B), pp. 1389–1400, Mar. 2010.

[22] M. Unoki, K. Imabeppu, D. Hamada, A. Haniu, and R. Miyauchi, “Embedding limitations with digital-audio watermarking method based on cochlear delay characteristics,” J. Inf. Hiding Multimedia Signal Process., vol. 2, no. 1, pp. 1–23, Jan. 2011.

[23] S. Wang and M. Unoki, “Speech watermarking method based on for-mant tuning,” IEICE Trans. Inf. Syst., vol. E98-D, no. 1, pp. 29–37, Jan. 2015.