High Capacity Audio Steganography Based on Contourlet Transform

Abbas Salman Hameed *

Department Electronic Engineering
College of Engineering
Diyala University
Diyala
Iraq

A B S T R A C T

The science of hiding information behind other cover media file is known steganography. Audio steganography means that a secret message is hidden by embedding it in an audio file. This paper presents a new audio steganography approach that is used the contourlet transform to hide a speech and image in an audio signal. The cover audio signal is modified to be suitable as input to contourlet transform and then secret data embed to the subbands of contourlet transform. The results showed high hiding capacity of data up to 90% of cover audio file size. In addition, performance analysis by measures factors: Normalized Correlation (NC), Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR) appears good quality results for both stego and secret data.

السعة التخزينية العالية للاخفاء الصوتي بالاعتماد على التحويل الكنتوري

الخلاصة

أخفاء المعلومات هو العلم الذي يتعامل مع إخفاء البيانات السرية في بعض الوسائط الاخرى. الاخفاء بالغطاء الصوتي يعني ان المعلومات المراد اخفاءها يتم دمجها داخل ملف صوتي وبشكل سري غير محسوس للمستمع. هذا البحث يقدم تقنية جديدة لإخفاء البيانات داخل ملف صوتي عن طريق استخدام التحويل الكنتوري. تم اخفاء نوعين من البيانات وهي بيانات صوت وبيانات صورة. في الطريقة المقترحة تم اعادة ترتيب بيانات الصوت الغطاء ليصبح مناسبا لتطبيق التحويل الكنتوري عليه ومن ثم يتم دمج البيانات السرية المراد اخفائها داخل الحزم الفرعية (subbands) الناتجة من التحويل الكنتوري للغطاء. اظهرت النتائج التي تم الحصول عليها عن طريق التقنية المقترحة قدرة عالية لاخفاء البيانات. حيث ان حجم البيانات المخفية وصل الى 90% من حجم الصوت الغطاء مع معايير جودة جيدة (عامل الارتباط و نسبة الاشارة الى الضوضاء و قمة نسبة الاشارة الى الضوضاء) لكل من البيانات المستخلصة و الغطاء الحامل لها.

1.   INTRODUCTION

Data hiding is one of the emerging techniques that provide safety of data transferring by concealing secret information into the multimedia contents such as images or audio after modifying some components in those host files.

Audio steganography is one of the familiar data hiding techniques that embeds existence secret information in audio signals. Based on the masking effect of human auditory system (HAS), audio steganography is worked by hiding a weak data signal in the presence of the large one [1-3]. Fig. (1) shows the general block diagram for the audio steganography.

There are three major performance criteria that revolve around the existing steganography algorithm: perceptual transparency, robustness and hiding capacity [4,5]. Perceptual transparency indicates that the secret message must be imperceptible from the stego signal. Robustness refers to the ability of reliable detection of embedded information after attacks. Hiding capacity is the amount of information that can be hidden successfully without a noticeable distortion in a cover media. All these requirements are supposed to produce a good steganographic system.

Many authors in the past years proposed different audio steganography methods that hid the secret information by modifying the samples of the audio signal directly and indirectly, such as

* *Corresponding author: E-mail* : abbasfuture@yahoo.com

1- Least Significant Bit (LSB): The least significant bit of each sample in the cover audio is replaced by a bit of the secret data. Even though this is a simple method, an attacker can easily extract the secret message and destroy [1,6].
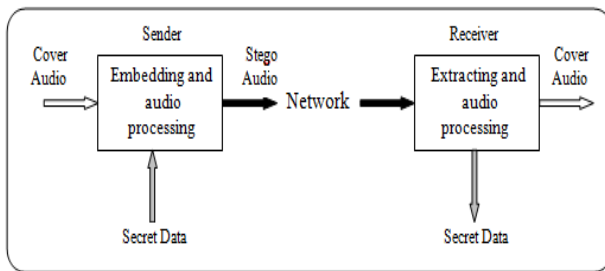
Fig. 1. The block diagram for the audio steganography [1].

2- Echo hiding: In this method, data is embedded in the echo part of the cover audio audio signal. The echo is a resonance added to the cover audio signal and hence the problem with the additive noise is avoided here but it is low embedding rate and low security [4].

3- Parity coding: To inserting message bit in parity coding method, check the parity bit of a group of samples. If the parity bit and message bit matches do nothing. Otherwise, change the LSB's of any one of the individual samples in that group to make the parity bit equal to the message bit. It is More robust than LSB, but still Easy to extract [4].

4- Tone insertion: Frequency masking property is exploited in tone insertion method. To embed information, a weak pure tone is masked in the presence of a stronger tone. Low embedding capacity is a weakness of this method [7].

5- Phase coding: This technique encodes the secret data bits as phase shifts in the phase spectrum of a digital audio signal depending on the fact that HAS can't sense the phase change in audio signal. Also, Low embedding capacity achieved with this method [3].

6- Spread spectrum: The basic spread spectrum method attempts to spread secret data across the audio signal frequency spectrum as much as possible, therefore it is occupied more bandwidth than what is indeed required for transmission [4,8].

Among these different techniques, wavelet domain [2,9,10] shows a high hiding capacity. A multi-resolution properties of wavelet transform provides an access to hide data in both approximations (significant parts) and details coefficients of the spectrum. As same in wavelet, Contourlet transform is construction a discrete domain multi-resolution and multi-direction expansion using filter banks [9,11]. Contourlet transform allows for a different number of directions at each scale by employing iterated filter banks, which makes it computationally efficient [11]. These filter banks generate a large sub-band coefficients used in this paper to hide information with a considerable hiding capacity.

A new audio steganography method based on contourlet transform is proposed in this paper. In this method, contourlet transform will apply first on cover audio file and then secret data will be hidden in proper sub-band contourlet coefficients.

Section 2, makes a review for some audio stignagrophy literatures. Contourlet transform theory is presented in section 3. The propose algorithm of audio steganography using contourlet transform will be discussed in section 4. Analysis and experimental results are depicted in section 5. Finally, section 6 draws the conclusions.

2. LITERATURE REVIEW

Shahreza and Shalmani [9] are proposed a new speech steganography in wavelet domain. The secret data hide in some of LSB detail wavelet coefficients selected with adaptive algorithm. Embedding capacity up to 20% approves with this method. However, it requires using secret key for selecting the coefficients to maximize the security. Bit modification for LSB is proposed by Gopalan and Shi [6]. In this technique, a large payload of data hiding up to 20 per cent of the cover audio size is offered using high bit indices for embedding. Shahadi and Jidin [12] proposed a paper in wavelet domain based on adoptive LSB with secret key to achieve high capacity up to 42% from the size of host audio signal with 50 dB signal to noise ratio for output stego signal. A new technique was proposed by Verma et al. [13] that hides information into audio cover based on samples comparison in wavelet domain. It offered embedding capacity up to 25% from the size of cover audio with 35 dB signal to noise ratio for output stego signal.

A novel method based on audio steganography was presented by Prasad et al. [10]. This paper implements hybrid audio steganography which hides the data by embedding it in variable LSB of some selected samples based on polynomial expression. Tayel et al. [1], present an implementation of LSB audio steganography technique using two cards of arduino.

3. CONTOURLET TRANSFORM

Contourlet transform (CT) is one of the new geometrical 2D data transforms, which can expeditiously represent 2D data elements contours and textures. It is an efficient representation of signals that requests the coefficients of functions to be sparse [11,14]. Multiresolution, Critical sampling, and Localization are many properties offered by Wavelet Transform. As well, additional properties are provided by CT such Directionality and Anisotropy [14]. In Contourlet, a discrete-domain multiresolution and multidirection expansion construct using nonseparable filter banks. Also, the high frequency sub−bands do not expose to down

sampling process as happened in wavelet that makes the high frequency scrambled [14,15].

Contourlet Transform, proposed by Do and Vetterli [12], is implemented using Pyramidal Directional Filter Bank (PDFB). Laplacian pyramid (LP) is applied for the first step to decompose the 2D data into a low pass and high pass sub-band, while to give directional shapes, directional filter banks (DFBs) are utilized in the angular decomposition stage. CT is directional representations of the signal and it allows to include different directions for different scales of that signal while achieving nearly critical sampling [14,15].

The first sub-band of this transform is the low pass region as a result of applying LP. The high pass region is split by DFB to give the frequency spectrum of the high pass sub-band [14]. CT may be decomposed into many levels and each *n*-level is efficiently divided by using DFB into 2^n sub-bands frequency partitioning for that level. Contourlet decomposition is shown in Fig. 2.

All the important feature of CT addition to its ability to capture the geometry of image edges made it used frequently in image's denoising [16], watermark detection [9], compression [17], edge detection [18] and many other applications. Here, the proposed method benefits from the huge number of coefficients that are produced by applied CT on the host audio signal to hide secret data in some of these coefficients.
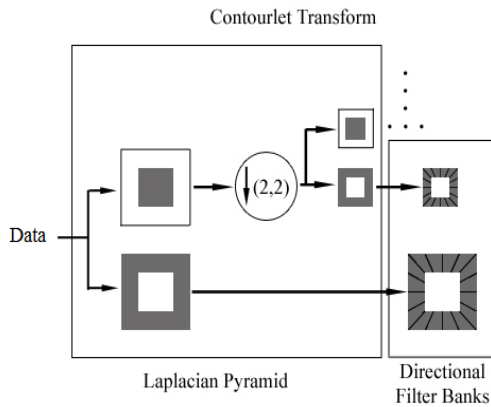

Fig. 2. Contourlet decomposition-Laplacian Pyramid followed by the directional filter bank [11].

4. PROPOSED SCHEME

The algorithm of data embedding and retrieving using audio steganography based on contourlet transform will be illustrated in this section. Fig. 3 shows the block diagram of the proposed audio steganography scheme.

Where the cover audio file is modified and rearranged by amplified and converted it into 2D initialization to be applied as input for the CT. Each sample of cover audio signal is adjusted to be set between 0 and 2 before magnifying it to produce inputting data to CT. The CT (decomposition) analyses data into many levels to generate spares structures of the 2D input data, so that the spares structures are attained and one or several CT sub-bands are

selected to embed process. After that, hiding data is prepared to embed. If the data is audio, it's embedded directly, but if it's an image, it's converted to 1D array then each of the 8-bit for two adjacent pixels are merged to generate 16-bit sample has value modified between 1 and -1 and then it's prepared to embedding process. Next step after computing the size of data that need to hide is choosing a suitable CT sub-bands coefficients from the last level (high pass region) by computing the energy for each sub-band and choose less one and so on in order to decrease the distorting that is resulted from embedding secret message in cover audio. Equation (1) [19] will be used to calculate the energy of contourlet sub-bands as:
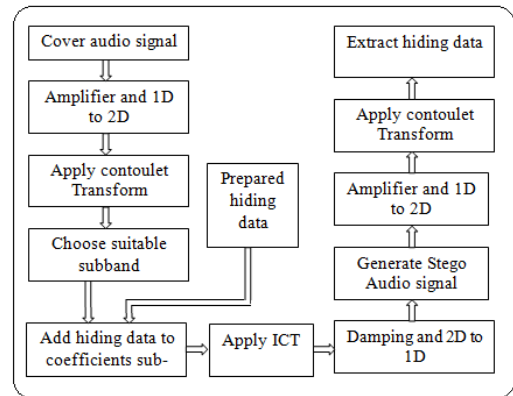

Fig. 3. Block diagram of proposed scheme.

$$E = \sum_i \sum_j |SC(i,j)|^2 \qquad (1)$$

where $0 \leq (i, j) \leq N$, $SC(i,j)$: the value of contourlet coefficient.

After identifying the suitable sub-band coefficients, the information wanted to be hidden are adding to values of these coefficients and new sub-band coefficients are resulted. The stego audio signal is prepared after reconstructing by inverse contourlet transform (ICT) then damping (by same amplified factor) and converting signal from 2D to 1D are done.

At the receiver side, the stego audio file is decomposed by CT with the same levels and filters used at the transmitter side and determine which one of the sub-band have been used to embed the information. To extract data hiding, receiver must know the original cover audio file to compare its sub bands coefficients with the sub bands coefficients for stego audio.

5. EXPERIMENTAL RESULTS

The cover audio file used here to hide data is Mobile tone with 44.1 KHz sample frequency, length of about 23 seconds, quantized by 16 bits and size of 2,097,152 bytes as shown in Fig. 4.
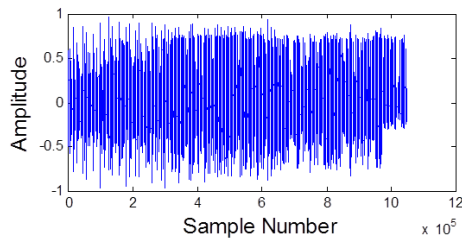
Fig. 4. Original Cover audio file.

The cover audio signal is amplified 100 times before applying CT on it and it is damped 100 times after embedding process and reconstruction of ICT. Speech and image are used as a hiding data will be applied with various sizes to illustrate the capability of hiding information in cover audio file using proposed scheme. Four level decomposition of contourlet transform is used in this work as shown in Fig. 5.
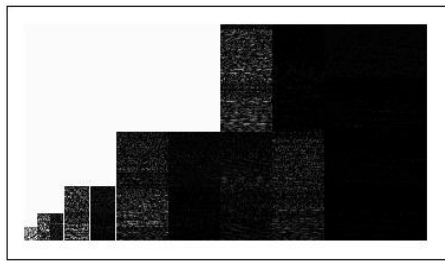

Fig. 5. Four-level Contourlet decomposition for cover audio file.

Some signal quality measures are utilized such as Normalized Correlation (NC) [20], Signal to Noise Ratio (SNR) [10], and Peak Signal to Noise Ratio (PSNR) [10], to illustrate and discuss the results for the proposed scheme.

A speech signal with 44.1 KHz sample frequency, and 16 bits per-sample is used as hidden signal. Fig. 6a shows stego audio signal with various hiding capacity for the proposed method. Fig. 6b shows original audio hidden with 209,715 bytes (10% from size of cover audio file), 524,288 bytes (25% from size of cover audio file), and 1,048,576 bytes (50% from size of cover audio file). Fig. 6(c) shows the hidden audio signals that are received with respect to these sizes.

Table 1 shows the effect of increasing data hiding capacity by embedding speech signals in cover audio file on quality of stego and extracted hidden audios. To illustrate the capability of hiding image for the proposed scheme, a gray scale Pepper image with various sizes are used in the embedding process after converting each two pixel values to 16-bit audio sample by merging the 8-bit binary form for each two pixels. Fig. 7(a) shows stego audio signal with various hiding capacity used to embed modified image signal. Fig. 7(b) shows the hidden image with different sizes (456×456 - 10%, 1024×1024 - 25%, and 1448×1448 - 50% from size of cover audio file) and Fig. 7(c) shows received hidden images that are received with respect to these sizes.

Table 2 shows the results of embedding different image size in cover audio signal. The very high similarity results between cover and stego audios, and also between original and extract secret data are illustrated by the high value of NC. Addition to that, the values of SNR and PSNR clarify a good performance of audio steganography is achieved and high hiding capacity is realized with good quality factors by using the proposed method. Several researchers have proposed algorithms to enhance the audio steganography technique. Table 3 shows the comparison between results of some those algorithms and method that is proposed in this paper.
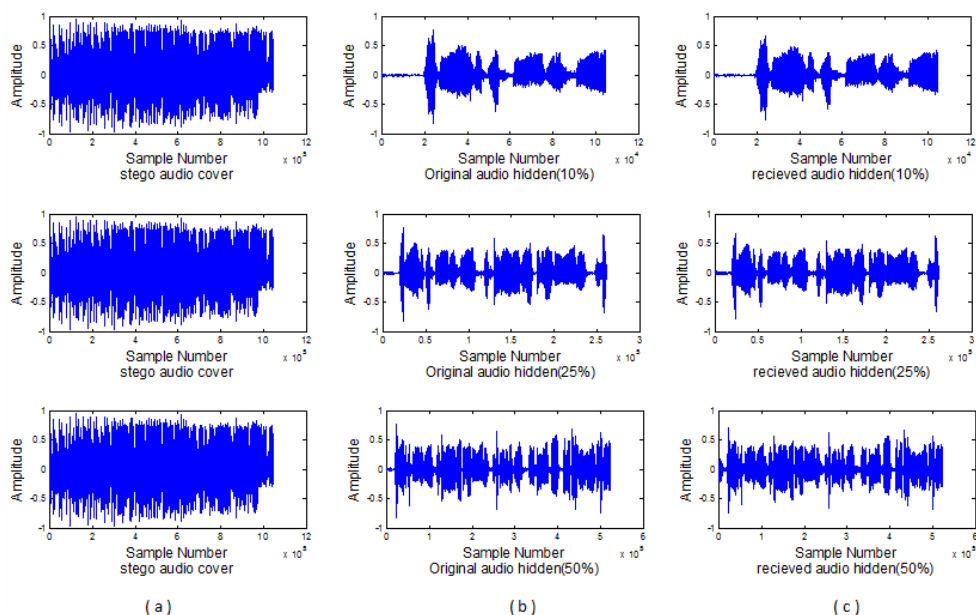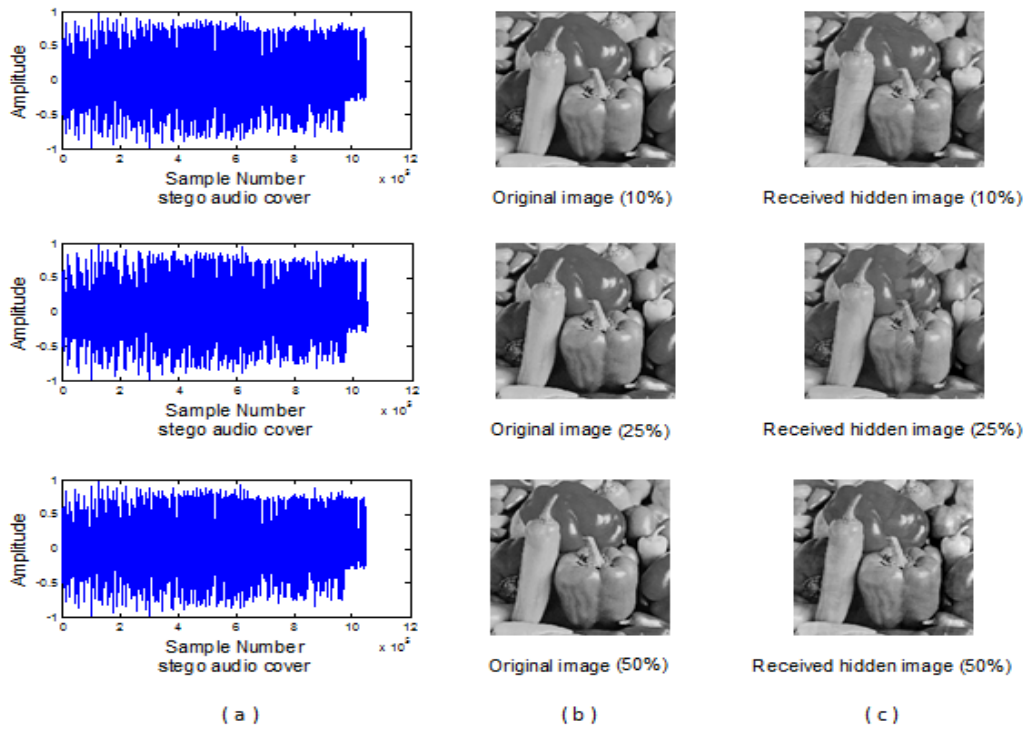

Fig. 6. Hide of audio signal in audio with a capacity 10%, 25%, and 50%. (a) shows stego audio signal, (b) original audio to be hidden, and (c) received hidden audio signal.

Fig. 7. Hide of image in audio with a capacity 10%, 25%, and 50% (a) shows stego audio signal, (b) original image to be hidden, and (c) received hidden image.

Table 1
Results of embedding different speech messages in cover audio file.

Embedding audio		Stego audio		Extracted audio	
Capacity (%)	Size (Kbyte)	NC	SNR (dB)	NC	SNR (dB)
10	204.8	0.99999	89.27	0.9954	63.72
20	409.6	0.99999	86.65	0.9951	63.57
30	614.4	0.99999	85.65	0.9941	62.51
40	819.2	0.99999	84.42	0.9935	59.48
50	1024	0.99999	83.17	0.9858	48.88
65	1331.2	0.99998	82.52	0.9787	43.49
75	1536	0.99998	81.78	0.9725	39.99
80	1638.4	0.99998	81.51	0.9622	35.31
85	1740.8	0.99998	81.13	0.9169	27.78
90	1843.2	0.99998	79.75	0.8990	23.89

Table 2
Results of embedding different image size in cover audio signal.

Embedding Image with	Stego audio		Extracted Image	
Capacity (%)	NC	SNR (dB)	NC	PSNR (dB)
10	0.99997	81.60	0.9962	60.59
20	0.99997	78.71	0.9900	54.28
30	0.99997	76.40	0.9864	52.40
40	0.99996	75.65	0.9857	51.90
50	0.99996	73.71	0.9800	51.24
65	0.99995	73.15	0.9695	46.25
75	0.99995	72.08	0.9611	43.32
80	0.99994	71.44	0.9560	41. 43
85	0.99994	70.92	0.9486	40.28
90	0.99994	70.61	0.9013	35.39

Table 3
Comparison between results of some algorithms and proposed approach.

Reference- Year	Methodology	Embedding Data capacity (%)	SNR (dB) (stego audio)
[12] -2011	Adaptive LSB in wavelet domain	Up to 42%	50
[13] -2014	Sample compression in wavelet domain	Up to 25%	35
[10] -2015	Selected LSB with Two Cryptography Levels	Up to 10%	87
[1] -2016	LSB	Up to 12.5%	58
Proposed method	Contourlet Domain	Up to 90%	79

6. CONCLUSION

In this paper, a new audio steganography method is proposed. It is applied in contourlet domain. The performance of the proposed scheme in terms of NC, SNR, and PSNR is analyzed. Very high NC and good SNR for stego audio illustrate that the secret message is imperceptible from the stego signal. As in table 3, many algorithms sought to increase the hiding data capacity with respect to good measurement's factors for stego audio and extract message signals. Using of wavelet transform in some of these algorithms achieved hiding capacity up to 42 percent of the host audio size with no less than SNR = 50 dB for stego audio. An increasing in the capability of hiding data is produced by applied the contourlet transform in audio steganography. From the results and with hiding capacity up to 90% from the size of cover audio message, the new proposed method can achieve SNR=79.75 dB and SNR=23.89 dB for stego audio and extracted hiding audio respectively when the audio file is embedded. Also, SNR=70.61 dB and PSNR=35.39 dB are produced for stego audio and extracted hiding image respectively when image file is embedded. All these results with NC near to full for stego audio and about 0.9 for extracting hiding data. These results obtain a high hiding capacity and perceptual transparency for the proposed method.

REFERENCES

[1] Tayel M, Gamal A, Shawky A. A proposed implementation method of an audio steganography technique. *The 18th International Conference on Advanced Communication Technology (ICACT2016), IEEE* 2016, 31 January-3 February; Pyeong Chang, South Korea; pp. 180 –184.

[2] Jisna VA, Sobin CC. A new proposal for audio steganography in wavelet domain. *The IEEE Fourth International Conference on Advances in Recent Technologies Communication and Computing (ARTCom2012)* 2012 October 19-20; Bangalore, India; pp. 310–312.

[3] Parab N, Nathan M, Talele KT. Audio steganography using differential phase encoding. Heidelberg, Germany: Technology Systems and Management; 2011.

[4] Djebbar F, Ayad B, AbedMeraim K, Hamam H. Comparative study of digital audio steganography techniques. *EURASIP Journal on Audio, Speech, and Music Processing* 2012; **25**: 1-16.

[5] Mohan M, Anurenjan PR. A new algorithm for data hiding in images using contourlet transform. *The Recent Advances in Intelligent Computational Systems (RAICS)* 2011 September 22-24; Trivandrum, India; pp. 411–415.

[6] Gopalan K, Shi Q. Audio steganography using bit modification-a tradeoff on perceptibility and data robustness for large payload audio embedding. *The19th International Conference on Computer Communications and Networks (ICCCN2010)*, 2010 August 2-5; Zurich, Switzerland: pp. 1–6.

[7] Gopalan K, Wenndt S, Noga A, Haddad D, Adams S. Covert speech communication via cover speech by tone insertion. *IEEE Aerospace Conference Proceedings* 2003 March 8-15; Montana, USA: pp. 1647-1653.

[8] Nugraha RM. Implementation of direct sequence spread spectrum steganography on audio data. *The International Conference on Electrical Engineering and Informatics (ICEEI)* 2011 July 17-19; Bandung, Indonesia, IEEE: pp. 1-6.

[9] Shahreza SS, Manzuri-Shalmani MT. High capacity error free wavelet domain speech steganography. *IEEE 2008 International Conference on Acoustics, Speech and Signal Processing* 2008 March 31- April 3; Las Vegas, NV, USA, IEEE: pp. 1729 -1732.

[10] Prasad GS, Varadarajan S. A novel hybrid audio steganography for imperceptible data hiding. *The International Conference on Communications and Signal Processing (ICCSP)* 2015 April 2-4; Melmaruvathur, India, IEEE: pp. 0634 – 0638.

[11] Do MN, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* 2005; **14**: 2091-2106.

[12] Shahadi HI, Jidin R. High capacity and inaudibility audio steganography scheme. *7th International Conference on Information Assurance and Security (IAS)* 2011 December 5-8; Melaka, Malaysia, IEEE: pp. 104-109.

[13] Verma SS, Gupta R, Shrivastava G. A novel technique for data hiding in audio carrier by using sample comparison in dwt domain. *Fourth International Conference on Communication Systems*

and Network Technologies 2014 April 7-9; Bhopal, India, IEEE: pp. 639-643.

[14] Do M, Vetterli M. Framing pyramids. *IEEE Transactions on Signal Processing* 2003; **51**: 2329-2342.

[15] Rabizadeh M, Amirmazlaghani M, Attari MA. A new detector for contourlet domain multiplicative image watermarking using Bessel K form distribution. *Journal of Visual Communication and Image Representation* 2016; **40**:324–334.

[16] Guo D, Chen J. The application of contourlet transform to image denoising. *Advanced in Control Engineering and Information Science* 2011, **15**, 2333-2338.

[17] Uma G, Selvi V, Nadarajan R. Coronary angiogram video compression using wavelet-based contourlet transform and region-of-interest technique. *IET Image Processing* 2012; 6: 1049–1056.

[18] Jin R, Yin J, Zhou W, Yang J. Improved multiscale edge detection method for polarimetric SAR images. *IEEE Geoscience and Remote Sensing Letters* 2016; **13**: 1104-1108.

[19] Chen X, Liu L. Contourlet-2.3 retrieval algorithm using absolute mean energy and kurtosis features. *Advanced Electrical and Electronics Engineering* 2011, **87**, 319-326.

[20] Wei SD, Lai SH. Fast template matching algorithm based on normalized cross correlation with adaptive multilevel winner update. *IEEE Transactions on Image Processing* 2008, **17**, 2227-2235.