❐ 4713

An efficient data hiding method in images

Saranya G
Department of Biomedical Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, India

Article Info	ABSTRACT
	Data Hiding has a huge range of applications in the medical field for transmission. It is helpful in securing the documentation of the patients from the violator with good storage space. The medical images of different modalities like CT, MRI, and PET with the digitized clinical information can be sent to the doctors across the world for the treatment. Due to the bandwidth and storage constraints, medical images along with the clinical information must be compressed before transmission and storage. This paper gives a technique for hiding the digitized clinical information along with the DICOM images in Complex Contourlet Transform (CCT) Domain. It also analyses the compression method by using an Entropy Encoder method. Hence, this work suggests that the data hiding method based on Complex Contourlet Transform (CCT) is efficient and also it has a high hiding capacity. The improved value of Compression Ratio (CR), Space Saving (SS), Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) shows that the new method satisfies the properties of the data hiding method.

Corresponding Author:

Saranya G,
Department of Biomedical Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, India.
Email: saranyaygk@gmail.com

1. INTRODUCTION

In the recent years, there has been a major advancement in the extent of data hiding in the medical image for various applications. Due to the properties of an image, the medical data is different from the other normal data due to its properties. Further, secretiveness and storage of medical data are important to protect it from accidental or mischievous change at the time of communication and storage. The medical data desire large extent of memory storage and communication bandwidth in the telemedicine applications. So, hospitals must provide good storage requirements [1]. This gives rise for data hiding method to reduce the storage space in the hospital digital directory.

The main specification of information hiding is to make the hidden message indistinct so that those who do not have the secret key cannot fetch the information back. Depending on the type of information, it is divided into two ways [2]; (1) Security: It is an insignificant data hiding; the hidden message is made much smoother by mixing the ability about the image noise and by using the error-correcting method. (2) Powerful: A small message bit is hidden in an efficient manner (i.e.) lossy compression, filtering, noise mixing, different geometrical information etc. And so, data hiding gives an appealing alternative by hiding rather than the appending data directly into the image. The most necessary properties of data hiding methods are Security, invisibility, complexity and high hiding capacity.

The important need of data hiding is to give a good confident level about the accuracy or control of a document [3]. In practical, the main problem in data hiding is attack. It is having the capability for unauthorized users to delete the content, to evaluate, write or redo the realistic hiding bits. So, "Attack" makes many methods inefficient. Hence, to destroy the pseudo code one should know the knowledge about

both hiding and retrieving algorithm and also about the secret key. Therefore, invisible hiding methods are designed for the better performance.

Data compression is widely used in all applications. The main aim of the data compression is to reduce the repetition for storage or to transmit the data in a better way. Hence, this results in the reduction of the file size in the hospital directory. The block diagram of the data encoding is shown in Figure 1.
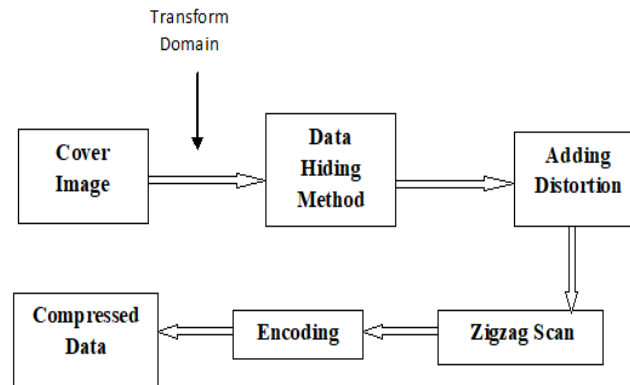

Figure 1. Block Diagram - Data Hiding

The block diagram consists of three elements namely Source, Quantizer, and Encoding. The source is commonly called as definite transforms. A quantizer is many to one aligning, which is the lossy procedure and it is the major source of compression and Encoding will compress the quantized values losslessly to provide us the improved compression. The entropy coding is generally used for discarding the coding repetition in the images. The most commonly used entropy coding are the Huffman Encoder (HE), Arithmetic Encoder (AE) and Run-Length Encoding (RLE). To provide the good compression a well-organized quantizer and encoding is very much important.

Huffman coding and Arithmetic coding are variable-length encoding method and apply the probability data of the symbols. The simplest form of quantization is commonly used in compression standard is "scalar quantization" where each signal value of scalar quantization is individually quantized. Huffman coding is easy and faster than the Arithmetic coding. It depends on probability statistics. Arithmetic works by evaluating a stream of data symbols and does not restore individual data symbols with a compressed version of large data; a number of bits are needed in the output code. Generally, it does not perform better than Huffman coding when incorrect probabilities are fed to the coder [4]. RLE removes linear correlation in the data source, but it is more suitable when large runs of the same symbol occur in the data stream. It is simple to implement and the process is fast.

The implementation of the proposed method compared with three encoding methods and the performance like Peak Signal to Noise Ratio (PSNR), Compression Ratio (CR) and Space Saving (SS) is evaluated to decide the quality of the compressed image. Compression ratio mostly ranges from 2:1 to 4:1 for natural images, but it can be much larger for the document images.

2. MATERIALS AND METHODS

An image processing application is performed in a transform domain rather than the pixel domain, often by means of an invertible definite transformation. In transform based schemes the image is first transformed before the hiding method and the watermark is concealed in the coefficients which are represented by the image. The hiding image is achieved by using a reverse transformation. The proposed method is implemented by using the Complex Contourlet Transform (CCT) for data hiding.

Origin of CCT: Initially, Wavelet Transform (WT) played a vital role in the telemedicine field, due to its better performance of multi-scale and local time-frequency. The disadvantage in the WT is shift sensitivity, fewer shapes, and its directionality. To overcome this drawback, the approaches like ridgelet, curvelet, beamlet, bandelet, contourlet etc have been introduced. Comparing to all the series of 'X-let' variance the most commonly accepted method is Contourlet Transform (CT), because of its good directionality and clear implementation. Contourlet transform was proposed by DO (2002) and so many modifications have been introduced to form new families like Non-Sub Sampled Contourlet Transform

(NSCT) and other localized Contourlet versions. NSCT was proposed by Cunha (2005), it has a good amount of shift insensitivity than the original contourlet transform but the disadvantage is it has higher repetition [5] which is given by the following (1), (i.e.)

$$Re = 1 + \sum_{s=1}^{S} 2^s \qquad (1)$$

Where, S is scale number of the transform.

Due to this disadvantage, NSCT is complex and time-consuming. Therefore, it requires a larger memory. Hence, to overcome these problems Complex Contourlet Transform (CCT) has come into existence. *Complex Contourlet Transform (CCT):* CCT is used to reduce the computational problem and also to improve the image quality. To reduce the higher repetition of the contourlet version, Dipeng chen and Qi Li introduced a CCT with the DT-CWT and NSCT.

CCT has basically two major steps, namely (1) Dual-tree complex wavelet transform (DT-CWT) for multi-resolution decomposition level which gives six directional sub-bands on each scale 2^j of the detail coefficient sub-space W_j, each sub-band has two wavelets, describe by $\vartheta \in \{1,2\}$ as real and an imaginary part of the wavelet coefficients. (2) Directional filter bank (DFB) for multi-directional decomposition to set the locally correlated coefficients which are evaluated by DT-CWT we apply l_j levels of Directional Filter Bank (DFB) to each multi-scale detail space W_j and the sub-bands are expanded to the number of 2^{l_j} [6]. The decomposition equation is given by,

$$\mu^{i,lj}_{k,n,u,v}(t) = \sum_{m \in Z^2} g_k^{lj}[m - S_k^{lj} n] \, \gamma^i_{m,u,v}(t), i = 1,2,3 \qquad (2)$$

Where, $\mu^{i,lj}_{k,n,u,v}(.)$ is a group of directional sub space $W_{j,k}^{l_j}$ at scale 2^j and decomposition direction k = $1 \dots 2^{l_j}$, $g_k^{l_j}(.)$ is impulse response of the synthesis filter, l^j is the decomposition level of the DFB.

In the performed transform, CCT combines the properties of NSCT (multi-resolution, localization, directionality, and anisotropy) and DT-CWT (translation invariant, directionality). Therefore, it is computationally more efficient than the other transform technique. The level 1 stage decomposition of CCT is shown in Figure 2.
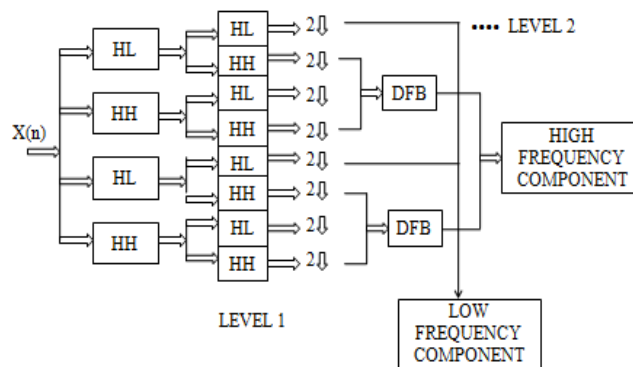

Figure 2. Decomposition Stage for Level 1 CCT

3. PROPOSED METHOD

In the proposed method, an algorithm for data hiding is explained with an example. Data Hiding: The process of data hiding method is shown in Figure 3. In the new data hiding method scaling factor/threshold value is set in such a way that no violators can pirate the patient data. If the scaling factor is very less in value then, it will give a better invisible hiding algorithm. To prove the security system of the proposed method, it is validated with various attacks. The attacks applied in the methods are Gaussian Noise, Salt & Pepper Noise, Sharpening, Rotating, Histogram Equalization, and Median Filter.
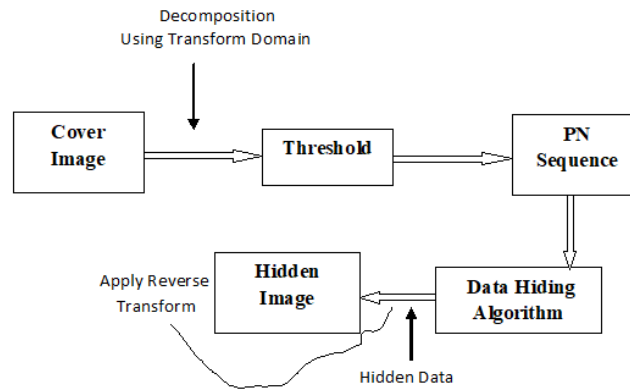
Figure 3. Block Diagram - Data Hiding

Generation of PN Sequence: Spread spectrum is one of the important techniques for transmission in which the data sequence apply the bandwidth with least requirement. It has the major advantages like strong anti-interference ability, low bit error rate, better hiding capacity, low intercept, and high confidentiality. It has two major properties (1) bandwidth of the signal transmission is higher than the bandwidth of the original data signal, (2) transmission signal bandwidth is decided by the pseudo-random code [7]. Pseudo Noise sequence generates a sequence of pseudorandom binary numbers by using the shift register. There are 'r' registers in the generator which boost their values at each instant depending on the value of the incoming pointer to the shift register. The main requirements of the pseudo noise code are good spectral characteristics and security [8].

A PN sequence is a bit stream of 1's and 0's with some idle properties. In a PN sequence the number of 1's and 0's of any amount of length differ only by one (i.e.) the number of 1's is just one more than the no of 0's (For e.g.) the PN sequence of length 15 is given by $(2^4 - 1)$ (i.e.) $(2^N - 1)$, it contains eight 1's and seven 0's (Mutagi, 1996). Therefore, a series of 1's and 0's is called 'run' and the amount of 1's and 0's is called 'run length'. A PN sequence of span $2^N - 1$ includes one run of N 1's and one run of N-1 0's. The series of other runs, N-2 to 1, of 1's and rises as the power of 2 [9] which is shown in Table 1.

Table 1. Number of runs 1's and 0's of various lengths in a PN sequence of length $2^N - 1$

Run Length	1's	0's
N	1	0
N-1	0	1
N-2	1	1
N-3	2	2
N-4	4	4
….	….	….
2	2^{N-4}	2^{N-4}
1	2^{N-3}	2^{N-3}

Data Hiding Algorithm:
1. Choose the medical cover image and decompose it by Complex Contourlet Transform (CCT)
2. Select the scaling factor or simply called threshold value for a better transmission purpose and also for data security
3. Implement the Spread spectrum hiding process by generating the Pseudo Noise Sequence to embed the message bit in the Coefficient obtained by CCT decomposition process for medical cover image
4. Hiding Data = CCT Coefficient + Scaling Factor
   * Message bit to be embed
5. Repeat the steps 3 and 4 until all the message bit is embedded
6. Apply Reverse CCT to obtain the final Embedded Image

Example for Proposed Method: Let us consider the following example for better understanding about the proposed data hiding. Initially, after the decomposition process choose any two coefficients from any direction and following process is performed to embed the entire message data.

Embedding
1. Let, A is the maximum data, B is the minimum data and C is the total message bit to be embedded.
2. Now, calculate the difference (D) between the A and B, Consider the scaling factor or the threshold value which is minimum
3. Therefore, the difference should be greater than the scaling factor. Hence, in this case, D > scaling factor
4. To embed the message bit new parameter (E) is introduced by taking difference between total message bit and D, then it is divided by 2
5. To obtain the embedding coefficient A' and B' is calculated,
   A' = A + E
   B' = B – E

4. EXPERIMENTAL RESULTS

In this paper, performance evaluation for data hiding is measured by using Peak Signal to Noise Ratio (PSNR), Space Saving (SS), Mean Square Error (MSE) and Compression Ratio (CR) [10].

a. Mean Square Error (MSE)

It is used to measure the distortion produced after hiding the data

$$MSE = \frac{1}{mn} \sum_{i=0}^{M-1} \sum_{i=0}^{N-1} (\ Cover\ Image -\ Reconstructed\ Image)^2$$

Where, M and N is the number of an image pixels.

b. PSNR:

It is used to measure the objective equality of an image which is given by,

$$PSNR = 10 \ \log_{10} \frac{Max^2}{MSE}$$

Where, Max is a maximum gray scale of an image. Better high PSNR have higher quality image (i.e.) it is close to the cover image.

c. Compression Ratio (CR)

It is given by,

$$CR = \frac{Uncompressed\ Size}{Compressed\ Size}$$

d. Space Saving (SS)

It is given by,

$$SS = 1 - \frac{Compressed\ Size}{Uncompressed\ Size}$$

5. RESULTS AND DISCUSSION

This section shows the performance evaluation for the data embedding scheme based on transform domain. Intially, the medical cover image is chosen and it is decomposed by the CCT. The medical cover image is shown in Figure 4. After the proper selection of an input image, decomposition process is performed for 3 levels. Complex Coefficients of CCT is shown in Figure 5 and Figure 6.
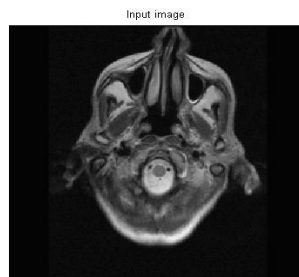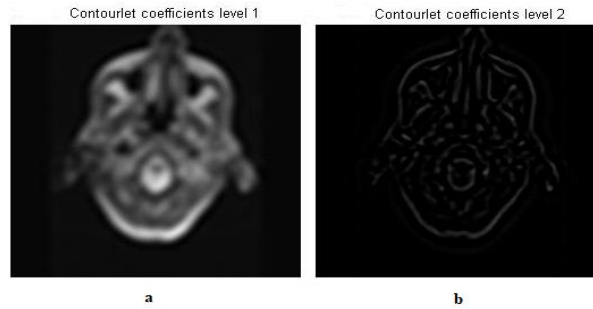

Figure 4. Medical cover image

Figure 5. Decomposition of CCT – (a) Level 1 Complex contourlet coefficient
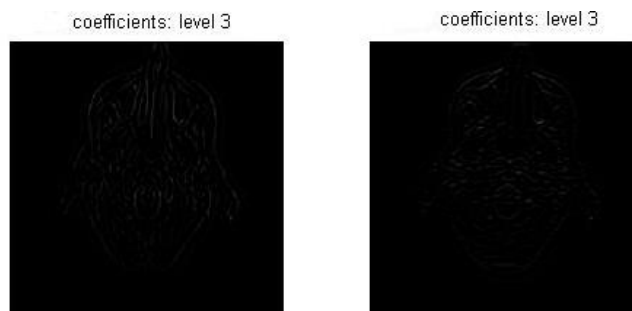(b) Level 2 Complex contourlet coefficient

Figure 6. Level 3 Complex contourlet coefficient for different direction

The medical cover image is decomposed for 3 levels, where low coefficients are displayed in black and high coefficients are displayed in white. The proposed data hiding method is applied in the decomposed cover image by selecting any two coefficients. Hidden image and the reconstructed image is shown in Figure 7 (a) and (b).
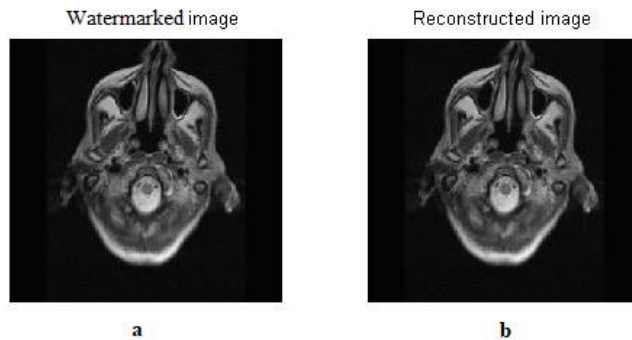
Figure 7. (a) Watermarked image (b) Reconstructed image

The reconstruction of the image is achieved by the inverse CCT. The performance evaluation for the proposed method is analysed by estimating the MSE, PSNR, CR and SS for with compression and without compression. The evaluated results are shown in Table 2 and Table 3.

Table 2. Proposed method without compression

Parameter	Proposed Method Without compression
MSE	0.000011
PSNR in dB	97.75

Table 3. Proposed method without compression and with attacks

Different Attacks Applied for Proposed Method without Compression	Parameter	
	MSE	PSNR in dB
Gaussian Noise	0.00912	68.56
Salt & Pepper Noise	0.000070	89.65
Sharpening	0.000074	89.47
Rotating	0.000099	88.20
Histogram Equalization	0.012	67.37
Median Filter	0.000889	78.67

From the Table 2 and Table 3 it is clear that the MSE and PSNR have better value for the proposed algorithm. Table 4 clearly shows how much the original data is compressed in size. It clearly shows that the compression ratio and space saving is high for Huffman encoder based Complex Contourlet Transform (CCT). The proposed algorithm is also compared with other transforms like Wavelet Transform (WT), and Non-Sub Sampled Contourlet Transform (NSCT) to evaluate the improved values for CCT which is shown in Figure 8, 9, 10 and Figure 11.

Table 4. Proposed method with encoding

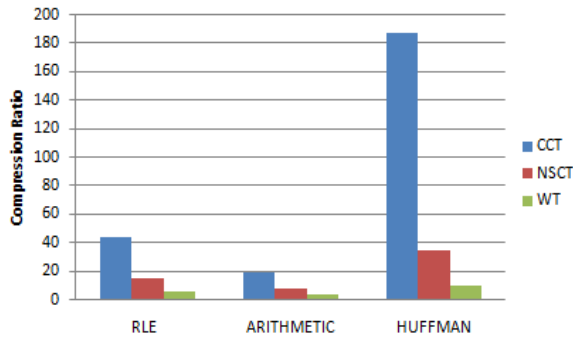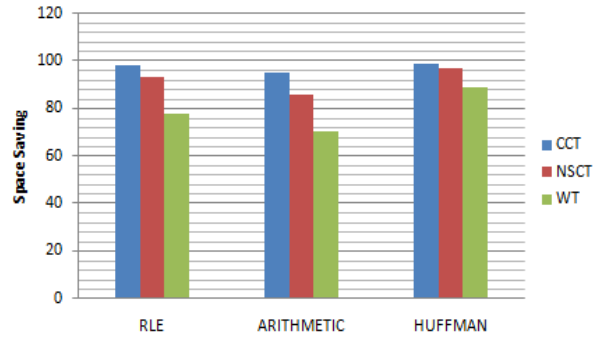| Complex Contourlet Transform (CCT)	Proposed Method With Compression		
	RLE	ARITHMETIC	HUFFMAN
Dimension	512*512	512*512	512*512
Uncompressed Size	32KB	32KB	32KB
Compressed Size	0.69KB	1.47KB	0.11KB
MSE	8.78	8.50	5.65
PSNR in dB	87.53	88.084	92.72
Compression Ratio	46:1	18:1	188:1
Space Saving	98%	95%	99%

Figure 8. Comparison of Compression Ratio

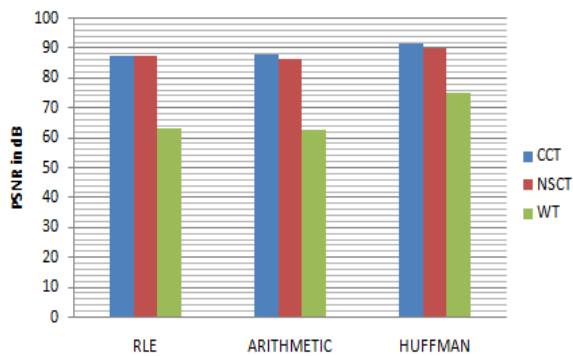Figure 9. Comparison of Space Saving
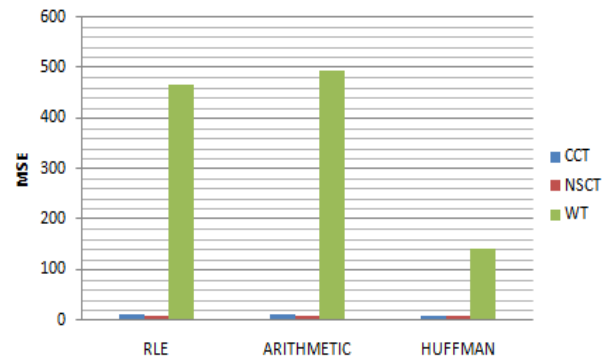

Figure 10. Comparison of PSNR

Figure 11. Comparison of MSE

The proposed method is evaluated also for compression and with attacks which is shown in the Figure 12 and Figure 13. The results of encoder are recorded on the basis of parameters such as Compression Ratio (CR), Space Saving (SS), Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) is also compared with CCT, NSCT and WT. It is tested with various attacks by evaluating Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) and from the analysis it is concluded that Huffman based CCT gives improved values than other methods.
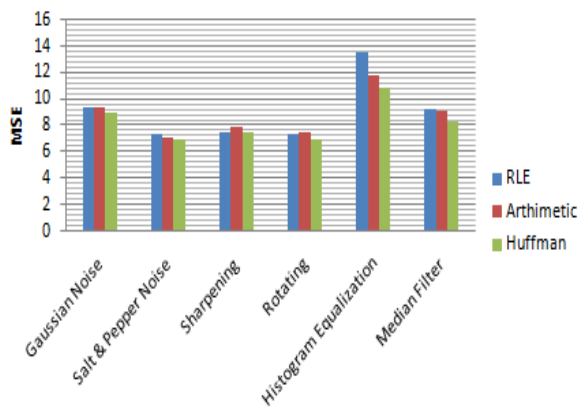

Figure 12. Proposed Method with Compression and With Attacks for MSE
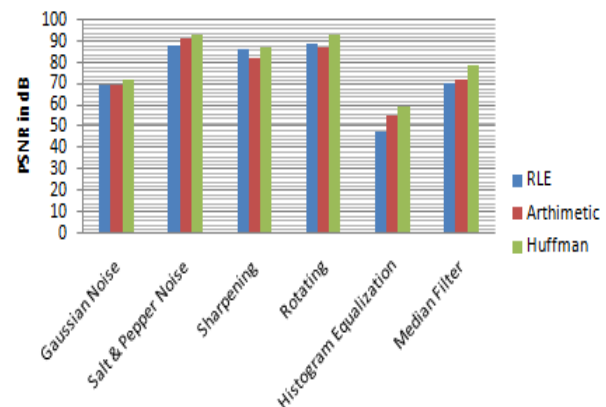
Figure 13. Proposed Method with Compression and With Attacks for PSNR

6. CONCLUSION

Provide In this paper, reversible data hiding method is used for telemedicine e healthcare application using CCT. It has the major properties of high capacity to embed the data; transparency, invisibility, and robustness to attacks are less when it is transmitted for diagnosis. From the evaluation, it is concluded that the CCT based data hiding is better than other transforms. Therefore, the proposed algorithm retrieves the original medical image and the experimental results suggest that the proposed algorithm can be efficiently used for data hiding process. The major advantages of the proposed algorithm are better PSNR, minimum MSE and large hiding capacity.

REFERENCES

[1] Saranya, G. and S. Nirmala Devi, "An approach for embedding and retrieving the data in medical image using contourlet transform" *Int. J. Scientific and Engineering Research*, 3: 769-775, 2012.

[2] Jiri Fridrich, "*Applications of data hiding in digital images*," Proceedings of the Conference on Signal Processing and Its Applications (ISSPA), IEEE, Aug 22-25, Brisbane, Qld, 1, 1999.

[3] Sunesh and Harish Kumar, "*Watermark attacks and applications in watermarking*," National Workshop-cum-Conf. Recent Trends in Mathematics and Computing (RTMC), pp: 8-10, 2011.

[4] Meghna Sharma and Rajeev Mathur, "A matlab to generation and simulation of PN sequence in matlab," *Int. J. Emerging Technology and Advanced Engineering*, 2: 143-146, 2012.

[5] Li-Wei Liu, Xin-Wu Chen and Zhi-Wei Ying, "*Texture image retrieval algorithm with dual tree complex contourlet and three statistical features,*" Int. Conf. Communication Software and networks (ICCSN), IEEE, May 27-29, Xi'an, pp: 349-352, 2011.

[6] Shao-Wei Dai, Yan-Kui Sun, Xiao-Lin Tian and Ze-Sheng Tang, 2007, "*Image denoising based on complex contourlet transform,*" Int. Conf. Wavelet Analysis and Pateern Recognition, IEEE, Nov 2-4, Beijing, China, 4, pp: 1742-1747, 2007.

[7] Mutagi, R.N., "Pseudo Noise sequences for engineers," *IET Trans., J. Electronics & Communication Engineering*, 8: 79-87, 1996.

[8] Saif Alzahir and Arber Borici, "An innovative lossless compression method for discrete-color images*,*" *Image Proc. IEEE Trans*., 2: 44-56, 2015.

[9] Ramesh Muthiya , "Gomathy Balasubramanian , and Sundararajan Paramasivam, Pseudorandom Noise Sequence of Digital Watermarking Algorithm based on Discrete Wavelet Transform using Medical Image," *The International Arab Journal of Information Technology*, 15, No. 4, 2018.

[10] Rohit Thanki, Ved Vyas Dwivedi, and Komal Borisagar, "Robust Watermarking Technique Using Different Wavelet Decomposition Levels For Signature Image Protection," *Journal of ICT*, 16, No. 1, pp: 157-174, 2017.