Asian Journal of Information Technology 15 (21): 4400-4406, 2016
ISSN: 1682-3915

Data Hiding and Retrieving Using Complex Contourlet Transform

G. Saranya and S. Nirmala Devi
Anna University, Chennai, India

Abstract: Data embedding has wide range of applications in the medical field. This method is helpful in securing information of the patients from the intruders with high storage capacity. The medical images of different modalities like CT, MRI and PET with the digitized clinical data can be sent to the physicians across the world for the diagnosis. Due to the bandwidth and storage constraints, medical images along with the clinical data must be compressed before transmission and storage. This study presents a new technique for data embedding and retrieving the digitized clinical data along with the DICOM images by using Complex Contourlet Transform (CCT). It also estimates the compression method by using entropy encoder. Hence, this study suggests that the data embedding and retrieving technique based on Complex Contourlet Transform (CCT) is more efficient and also it has high hiding capacity. The improved values of Compression Ratio (CR), Space Saving (SS), Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) show that the new technique satisfies the properties of the data embedding technique.

Key words: Data embedding, data retrieving, CCT, PN sequence, entropy encoder

INTRODUCTION

In the last few years, there has been a considerable growth in the amount of data embedding in the medical image in hospitals. Due to the properties of an image, the medical information is different from the other normal information. Moreover, confidentiality and certainty of medical data is vital to shield it from accidental or malicious change during the exchange and storage. The medical information require large amount of memory storage and exchange bandwidth in telemedicine applications. Thus, hospitals must serve with very elevated storage requirements (Saranya and Nirmala, 2012). This gives rise for data embedding technique to reduce the space in hospital digital database.

Data hiding is called steganography which is similar to the cryptography. The main requirement of the cryptography is to make the hidden message unintelligible, so that those who do not possess the secret security key cannot retrieve the message. Depending on the type of information and the form in which it is hidden in the image, it is divided by two types (Fridrich, 1999).

Robustness: It is an undetectable data hiding; the hidden message is made much more sophisticated by adding the knowledge about the image noise and by using error-correcting codes.

Robust: A small message is hidden in the image in a robust manner (i.e.) lossy compression, filtering, noise adding, geometrical variation etc. Therefore, data hiding provides an appealing alternative by hiding rather than the appending information directly into the image. By using the efficient algorithm along with security secret key will be useful to retrieve the hidden message. The most important properties of data hiding strategy are robustness invisibility, security, complexity and high hiding capacity.

The main uses of the watermarking are to give the level of certainty about the authenticity or ownership of a document (Sunesh, 2011). In practical application, the main challenging problem is water marking attacks. It is having the ability for unauthorized users like detach, discover and evaluate, write or redo the realistic watermarking bits. Hence, “Attack” makes many of the current algorithms ineffective and so invisible watermarks are designed to be an imperceptible. Therefore, the security of the watermarking is related with the secret key. To destroy the algorithm one should know the knowledge about both the algorithm and also about the security secret key.

Data compression is widely used in all applications which include multimedia, documents, video conferencing and medical imaging. The main aim of the data compression technique is to reduce the redundancy for storage or to transmit the data in an efficient way. Hence, this results in the reduction of the file size in hospital digital database. The block diagram of the data encoding and decoding is shown in Fig. 1 and 2.

The block diagram consists of three elements namely source encoder, quantizer and encoding. Source encoder is generally called as linear transforms. Quantizer is many

Corresponding Author: G. Saranya, Anna University, Chennai, India

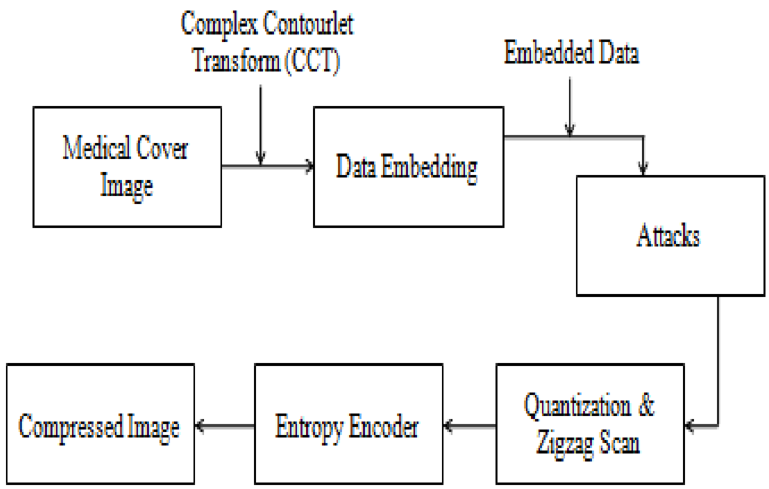

Fig. 1: Block diagram of data encoding

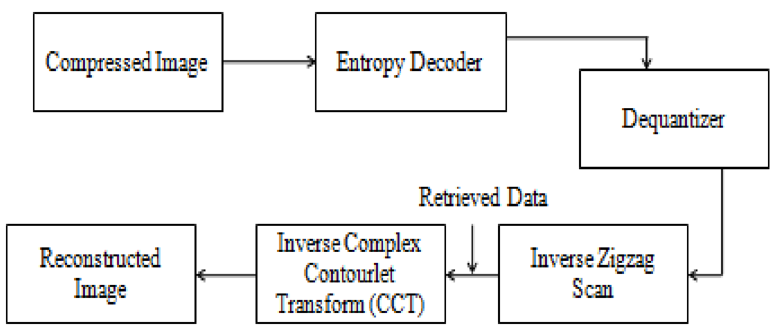

Fig. 2: Block diagram of data decoding

to one mapping which is the lossy process and it is the main source of compression in an encoder and entropy encoder will compresses the quantized values losslessly to provide us the better compression. Entropy coding technique is generally used for eliminating the coding redundancy 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 designed quantizer and entropy encoder is highly important.

Huffman coding and arithmetic coding both are variable-length encoding technique and utilise the probability information of the symbols for the coding. Simplest form of quantization mostly used in compression standard is "scalar quantization" where each signal value of scalar quantization is individually quantized. Huffman coding is the simplest and faster than the arithmetic coding. It is based on probability statistics. Arithmetic works by treating a stream of data symbols as a whole and does not replace individual data symbols with compressed versions of longer message; more number of bits is needed in the output code. Generally, it does not perform better than Huffman coding when incorrect probabilities are fed to the coder (Alzahir and Borici, 2015). RLE removes linear correlation in the data source but is more suitable when large runs of the same symbol occur in the data stream. It is easy to implement and quick to execute.

The implementation of the proposed method has been match up with all the three entropy encoder mentioned earlier and the quality measurements like Peak Signal to Noise Ratio (PSNR), Compression Ratio (CR) and Space Saving (SS) have been estimated to decide the quality of the compressed image. Compression ratio typically ranges from 2:1-4:1 for natural images but it can be much larger for the document images.

MATERIALS AND METHODS

A number of image processing functions are orderly carried out in a domain other than the pixel domain, often by means of an invertible linear transformation. In

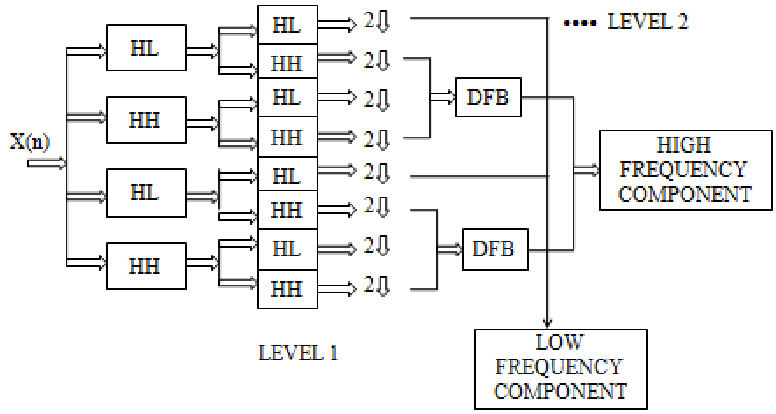

Fig. 3: Decompositions stage for level 1 CCT

transform based schemes the image is transformed prior to watermark embedding and the watermark is concealed in the coefficients representing the images. The watermarked image is obtained by using an inverse transformation. The proposed method is implemented by using the Complex Contourlet Transform (CCT) for data embedding and data retrieving.

Origin of CCT: In the past years, Wavelet Transform (WT) played an important role in the telemedicine field, due to its better characteristics of multi-scale and local time-frequency. But the major drawback in WT is shift sensitivity, lack of shapes and directionality. To overcome this drawback, many other approaches are introduced like ridgelet, curvlet, beamlet, bandelet, contourlet etc. In compare to all the series of 'X-let' families the most commonly acceptable method is Contourlet Transform (CT) because of its good directionality and easy implementation. Contourlet transform was proposed by DO and lots of modifications have been proposed and formed new families like Non-Sub sampled Contourlet Transform (NSCT) and other localized contourlet version etc. NSCT was proposed by Cunha, it has higher shift insensitivity than the original contourlet transform but the main disadvantage is it has higher redundancy (Liu *et al.*, 2011), it is given by the following equation, (i.e.,):

$$Re = 1 + \sum_{S=1}^{S} 2^{s} \tag{1}$$

where, S is scale number of the transform. Due to this disadvantage, NSCT is more time consuming and it needs larger memory. Hence, to overcome these defects Complex Contourlet Transform (CCT) has come into existence.

Complex Contourlet Transform (CCT): In this work, a new approach is performed by using CCT which reduces the computational problem and also improves the image quality. To reduce the higher redundancy of the contourlet version, Dipeng chen and Qi Li developed a CCT combining the DT-CWT and NSCT.

It has basically two major steps, namely Dual-tree complex wavelet transform (DT-CWT) for multi-resolution decomposition stage level which gives six directional sub-bands on each scale 2^j of the detail coefficient sub-space , each sub-band has two wavelets, describe by ϑ∈ as real and an imaginary part of the wavelet coefficients. Directional Filter Bank (DFB) for multi-directional decomposition to set the locally correlated coefficients which are recorded by DT-CWT we apply l_j levels of DFB to each multi-scale detail space W^j and the sub-bands are expanded to the number of 2^{lj} (Fridrich, 1999; Liu *et al.*, 2011; Dai *et al.*, 2007). The decomposition equation is given by:

$$u^{i,l,j}_{k,nu,v}(t) = \sum_{m\in Z} g_k^{lj}\left[m - S_k^{lj}\, n\right]\gamma^{i}_{m,uv}(t), i = 1,2,3 \tag{2}$$

where $u^{i,l,j}_{k,nu,v}(.)$ is a group of directional sub space $W^{lj}_{j,k}$ at scale 2^j and decomposition direction $k = 1\ldots 2^{lj}$, $g^{lj}_{k}(.)$ is impulse response of the synthesis filter, l^j is the decomposition level of the DFB.

In the obtained transform, CCT combines the characteristics of NSCT (multi-resolution, localization, directionality and anisotropy) and DT-CWT (translation invariant, directionality). Therefore, it is computationally more efficient than the other transform methods. The level 1 stage decomposition of CCT is shown in Fig. 3

Proposed method: In the proposed method, an algorithm for data embedding and data retrieving is explained with an example.

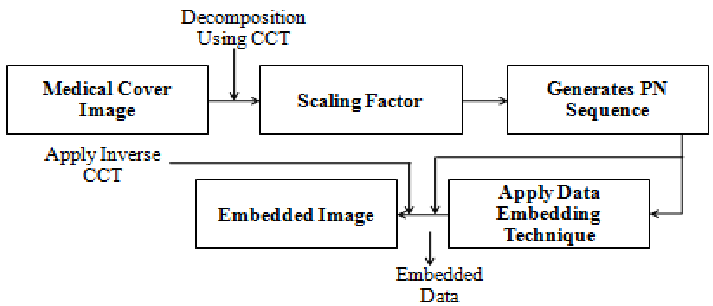

Fig. 4: Data embedding flow block

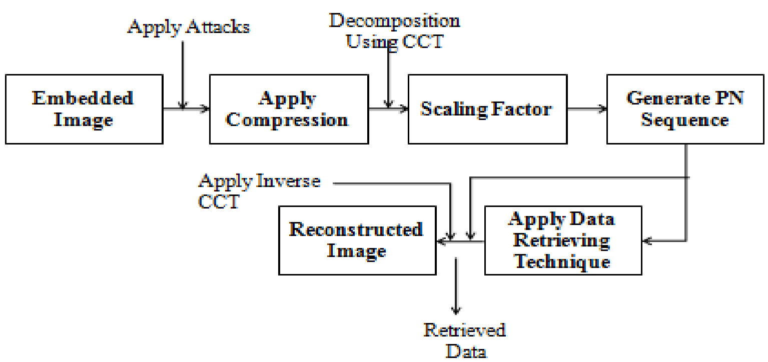

Fig. 5: Data retrieving flow block

Data embedding: The process of data embedding technique is shown in Fig. 4

Data retrieving: The process of data retrieving technique is shown in Fig. 5. In the proposed data embedding and data retrieving process scaling factor or the threshold value is chosen in such way that no intruders can pirate the patient information. Since, scaling factor is very less in value it will have a better invisible embedding process. To prove the robustness of the proposed algorithm, it is tested with various attacks. The attacks applied in the methods are Gaussian Noise, Salt and pepper noise, sharpening, rotating, histogram equalization and median filter.

PN sequence: Spread spectrum is one of the methods for transmission in which the data sequence employ the bandwidth in excess with least bandwidth which is necessary to transmit. It has advantages like strong anti-interference ability, low bit error rate, good hiding, low intercept, high confidentiality. It has two major characteristics bandwidth of the signal transmission is greater than the bandwidth of the original information signal; transmission signal bandwidth is decided by the pseudo-random code (Sharma and Mathur, 2012). Pseudo noise sequence generates a sequence of pseudorandom binary numbers by using shift register. There are 'r' registers in the generator which improve their values at each instant depending on the value of the incoming pointer to the shift register. The main properties of the pseudo noise code are good spectral characteristics and security.

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}

A PN sequence is a bit stream of 1's and 0's which is almost occurring, with some unique properties. In a PN sequence of any length the number of 1's and 0's differ only by one (i.e.,) the number of 1's is just one more than the no of 0's (e.g.) the PN sequence of length 7 is given by (i.e.) (, it contains four 1's and three 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 which is shown in Table 1.

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

Embedding: Let, X be the maximum data, Y be the minimum data and Z be the total message bit to be embedded respectively. Now, calculate the difference (D) between the X and Y, Consider the scaling factor or the threshold value which is minimum.

Therefore, the difference should be greater than the scaling factor. Hence in this case, D>scaling factor to embed the message bit new parameter is introduced by taking difference between total message bit and D and then it is divided by 2. To obtain the embedding coefficient X' and Y' is calculated:

$$X' = X + \text{new parameter}$$

$$Y' = Y - \text{new parameter}$$

Retrieving: Take X' and Y'. The difference between X' and Y' is given by D' = X'-Y'. Hence, the D'>scaling factor. Therefore, total message bit = retrieved data.

RESULTS AND DISCUSSION

In this study, performance evaluation for data embedding and retrieving is measured by using Peak Signal to Noise Ratio (PSNR), Space Saving (SS), Mean Square Error (MSE) and Compression Ratio (CR).

Mean Square Error (MSE): It is used to measure the distortion produced after the embedding and retrieving process is given by:

$$MSE = \frac{1}{mn}\sum_{i=0}^{M-1}\sum_{i=0}^{N-1}\begin{pmatrix}\text{Cover image} - \\ \text{Reconstructed image}\end{pmatrix}^2$$

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

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.

Compression Ratio (CR): It is given by:

$$CR = \text{Uncompressed size} / \text{Compressed size}$$

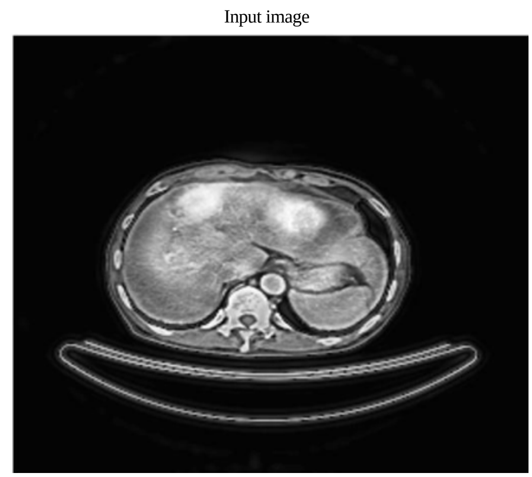

Fig. 6: Medical cover image

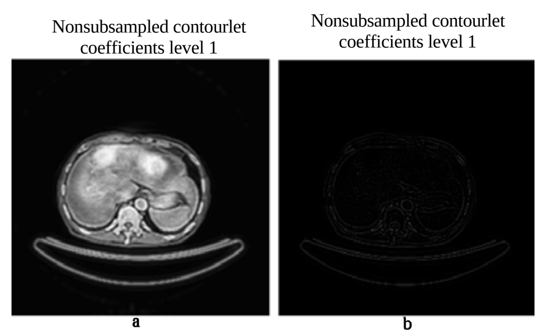

Fig. 7: Decomposition of CCT: a) Level 1 complex contourlet coefficient; b) Level 2 complex contourlet coefficient

Space Saving (SS): It is given by:

$$SS = 1 - \frac{\text{Compressed size}}{\text{Uncompressed size}}$$

This study shows the performance evaluation for the data embedding scheme based on Complex Contourlet Transform (CCT). Intially, the medical cover image is chosen and it is decomposed by the CCT. The medical cover image is shown in Fig. 6.

After the proper selection of an input image, decomposition process is performed for 3 levels. Complex Coefficients of CCT is shown in Fig. 7 and 8.

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 embedding and data retrieving method is applied in the decomposed cover image by selecting any two coefficients. Watermarked image and the reconstructed image is shown in Fig. 9a and b.

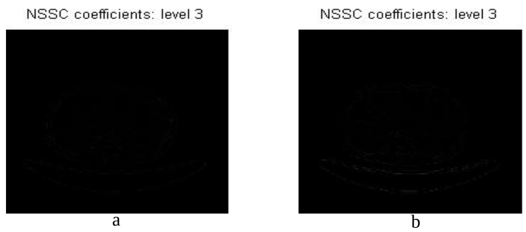

Fig. 8: Level 3 complex contourlet coefficient for different direction

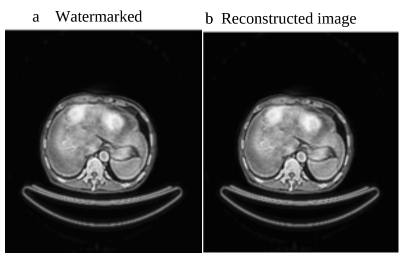

Fig. 9: a) Watermarked image; b) Reconstructed image

Table 2: Proposed method without compression

Parameter	Proposed method without compression
MSE	0.000025
PSNR in dB	94.1517

Table 3: Proposed mthod without cmpression and with attacks

Different attacks aplied for poposed mthod without cmpression	Parameter	
	MSE	PSNR in dB
Gaussian noise	9.48	69.10
Salt andpper nise	7.09	92.43
Sharpening	7.48	87.61
Rotating	7.00	93.62
Histogram eualization	12.7	51.60
Median flter	8.59	76.29

Table 4:Proposed mthod with etropy ecoding

Complex contourlet Transform (CCT)	Proposed method with compression		
	RLE	Arithmetic	Huffman
Dimension	512*512	512*512	512*512
Uncompressed size	32 kb	32 kb	32 kb
Compressed size	0.72kb	1.69 kb	0.17 kb
MSE	8.883660	8.506704	5.908911
PSNR in dB	87.65044	88.08403	91.72796
Compression ratio	44:1	19:1	188:1
Space saving (%)	98	95	99

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 3.

From Table 2 and 3, it is clear that the MSE and PSNR for proposed method without compression for CCT have better value.

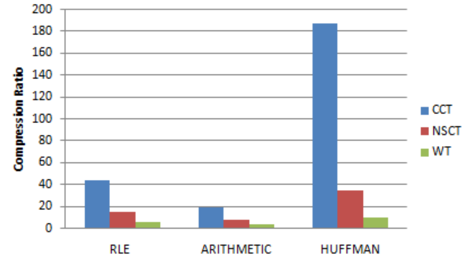

Fig. 10: Comparison of compression ratio

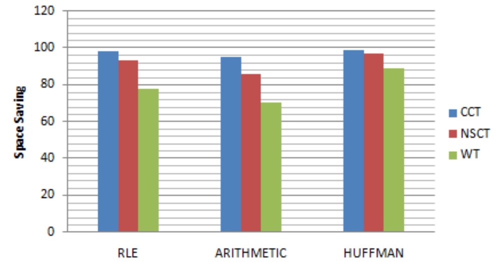

Fig. 11: Comparison of space saving

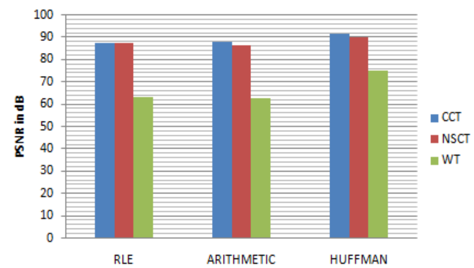

Fig. 12: Comparison of PSNR

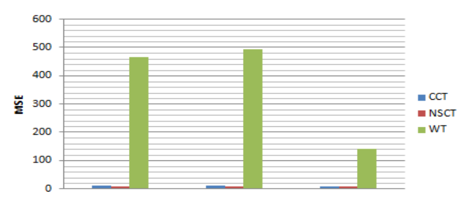

Fig. 13: Comparison of MSE

Table 4 clearly shows how much the original data is compressed in size also it clearly shows the compression ratio and space saving is high for Huffman encoder based Complex Contourlet Transform (CCT).

The proposed method 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 Fig. 10-13.

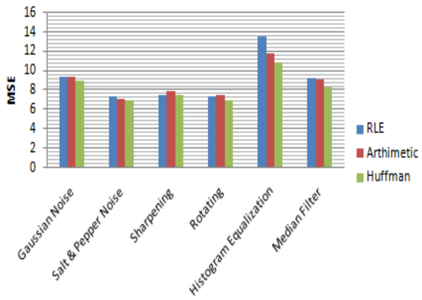

Fig. 14: Proposed method with compression and with attacks for MSE

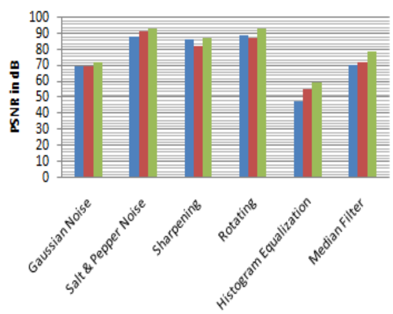

Fig. 15: Proposed method with compression and with attacks for PSNR

The proposed method is evaluated also for compression and with attacks which is shown in Fig. 14 and 15.

The results of entropy 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.

CONCLUSION

In this study, reversible data embedding and retrieving method is used for telemedicine applications using CCT. It has the major characteristics of high capacity to embed the data, imperceptibility (transparency invisibility), robustness to attacks are less when it is transmitted for diagnosis. From the results it is concluded that the CCT based data embedding and data retrieving is better than other transforms. Therefore, the proposed algorithm recovers the original medical image and the experimental results suggest that the proposed algorithm can be efficiently used for data embedding and retrieving process. The major advantages of the proposed algorithm are better PSNR, minimum MSE and large embedding capacity.

REFERENCES

Alzahir, S. and A. Borici, 2015. An innovative lossless compression method for discrete-color images. IEEE. Trans. Image Process., 24: 44-56.

Dai, S.W., Y.K. Sun, X.L. Tian and Z.S. Tang, 2007. Image denoising based on complex contourlet transform. Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, November 2-4, 2007, IEEE, Beijing, China, ISBN:978-1-4244-1065-1, pp: 1742-1747.

Fridrich, J., 1999. Applications of data hiding in digital images. Proceedings of the 5th International Symposium on Signal Processing and its Applications ISSPA'99, August 22-25, 1999, IEEE, New York, USA., ISBN:1-86435-451-8, pp: 9-9.

Liu, L.W., X.W. Chen and Z.W. Ying, 2011. Texture image retrieval algorithm with dual tree complex contourlet and three statistical features. Proceedings of the 2011 IEEE 3rd International Conference on Communication Software and Networks (ICCSN), May 27-29, 2011, IEEE, Xinyang, China, ISBN:978-1-61284-485-5, pp: 349-352.

Mutagi, R.N., 1996. Pseudo noise sequences for engineers. Electron. Commun. Eng. J., 8: 79-87.

Saranya, G. and D.S. Nirmala, 2012. An approach for embedding and retrieving the data in medical image using contourlet transform. Int. J. Sci. Eng. Res., 3: 769-775.

Sharma, M. and R. Mathur, 2012. A matlab to generation and simulation of PN sequence in matlab. Int. J. Emerging Technol. Adv. Eng., 2: 143-146.

Sunesh, H.K., 2011. Watermark attacks and applications in watermarking. Proceedings of the National Workshop-Cum-Conference on Recent Trends in Mathematics and Computing (RTMC), May 21, 2011, Conal & Gavin, India, pp: 8-10.