A Robust Image Steganography using CDF Lifting Scheme and Huffman Encoding

Suganthi Venkatachalam
Saranathan College of
Engineering
Trichy,Tamilnadu,
India

A. Shamim Banu
Saranathan College of
Engineering
Trichy,Tamilnadu,
India

M. Padmaa
Saranathan College of
Engineering
Trichy,Tamilnadu,
India

ABSTRACT

In the modern era, concerns about safeguarding the information from attacks are on rise. Here, 2-d lifting wavelet decomposition of a grayscale cover image through Cohen, Daubechies and Feauveau (2,2) lifting wavelet is performed and three secret images are shuffled and embedded in detail coefficients of the transform using approximation coefficients as indicator. Here, a novel Approximate Coefficients Value Indicator methodology is used to shuffle and embed the binary images in detail coefficients. Lossy and lossless steganographic techniques, followed by a lossless compression scheme using Huffman encoding are proposed. In lossy steganographic technique, secret images are embedded in LSB values of high frequency bands of the transform. In lossless technique, high frequency band values are expanded and images are embedded as previous method. Attacks such as Histogram attacks, statistical attacks are performed to check the robustness of the proposed algorithm. MSE and PSNR values are calculated to demonstrate the imperceptibility of the proposed method.

Keywords
DWT, CDF, Integer wavelet transform, stego image, cover image, lossy, lossless, Histogram

1. INTRODUCTION
 Steganography, which is derived from the Greek word 'Steganographia' means covered writing. Cryptography encrypts the content of the secret message where the intruder understands that something is hidden, whereas steganography conceals the existence of the secret information itself. Another form of information hiding is Image watermarking which is widely used to check the authentication of the data, to which the information is embedded into. A message is embedded in a cover digital file, preferably image and audio files because of their redundancy and the resulting stego image is transmitted through a channel to the receiver. The receiver uses extraction algorithm to retrieve the message. To prevent hackers from identifying the existence of hidden message in the cover, various image steganographic techniques, such as spatial domain techniques, transform domain techniques such as DFT DCT, FFT, DWT and statistical techniques are proposed in the literature. Both lossless and lossy methods are also proposed based on correlation of the data. Steganalysis is the identification of stego image. It uses visual, structural and statistical attacks [1] to identify the existence of hidden message. Huffman encoding is one of the lossless compression techniques, which provides efficient and lossless compression based on the frequency of the values.

2. RELATED WORKS
Image Steganography started with spatial domain approaches, predominantly spatial domain data hiding in LSB of the image pixels. [2, 3]. The major setback is that they are vulnerable to even slight attacks. Then pixel value differencing (PVD) methods [4], Side match and modified side match schemes of LSB substitution[5,6], despite of large embedding capacity, they were still ineffective against attacks.

 In Transform domain steganography methods, the coefficients of the transforms of the cover data were tweaked to store secret message. Hybrid methods combine transform with another transform or compression techniques- Based on Vector quantization and DWT, image was hidden [7]. Singular value decomposition with wavelet transform was used for watermarking [8].

 DCT based Steganographic methods on JPEG were proposed [9, 10, 11], which utilized statistical distribution of the cover data. In Wavelet-based steganographic methods, the cover image is converted into detail and approximation components and they are converted back to cover image [12, 13]. Comparatively, wavelet based methods preserves image quality and authenticity and robust against attacks [14]. Swelden proposed using the biorthogonal wavelet based on lifting scheme [15], where the discrete wavelets are polyphase decomposed.

 In this work, a novel approximation value indicator of the Transform based Lifting Steganography trailed by efficient compression is proposed. In lossy technique, three binary secret images are mixed up and distributed among the LSB values of HH, HL and LH bands, using LL band values as indicator. Instead of lossy LSB replacement process, in lossless steganographic process, one binary image is embedded and before embedding in the high frequency bands, their values are expanded. Huffman encoding of the stego DWT transform provides lossless and robust compression. Versatile images are tested and they withstand visual, structural and statistical attacks. MSE and PSNR are calculated for testing the quality of the images.

Section 3 provides brief information on Lifting based DWT, Huffman encoding and decoding. Section 4 deals with proposed Methodology with their embedding and extraction algorithm for both lossy and lossless cases. Section 5 discusses experimental results- the binary images which are embedded and retrieved. In lossy case, Original image and stego image result with its histogram before and after applying steganography, and associated MSE, PSNR and Compression ratios which proves the effectiveness of the methodology are discussed. In lossless scenario, original image, stego image, compression ratio and retrieved image are discussed. Section 6 provides conclusion of the proposed paper and future works are discussed.

3. LIFTING AND COMPRESSION
3.1 Lifting Wavelet
Wavelets are smooth basis functions. Any discrete wavelet transform with finite filters h (z) and g (z) can be decomposed

into lifting steps [15, 16]. Perfect reconstruction of filter bank and Euclidean algorithm are the keys that make the Lifting possible. Lifting has advantages over Fourier transform since it works on both time and frequency.

Wavelet filters h (z) and g (z) together forms a poly phase matrix P (z). Euclidean algorithm factorizes the resulting P (z) into elementary matrices and the elements of the decomposed matrices are used to construct Predict and Update operators in Lifting [10].

Data is divided into two smaller subsets, predict the subset based on the local correlation in the original data, and replace the detail as the difference between data and prediction. Lifting version of CDF (2, 2), also known as (5, 3) Integer wavelet transform is used here. Usually the Lifting-based DWT requires less computation compared to the convolution-based approach.

3.1.1 Forward and Inverse Transforms
Forward Transform has three stages. First, data is split into odd and even subsets. Odd subset is predicted based on local correlations in the given data and detail co-efficient are updated with difference between data and prediction. Approximation coefficients are updated with difference values.

Inverse Transform is the reverse process of Forward Transform. The original signal is recovered from its approximations and details coefficients along the wavelet decomposition tree. Three inverse lifting steps are Inverse Update, Inverse Prediction and finally merging.

3.2 Huffman Encoding and Decoding
Huffman encoding, uses less number of bits to encode the image pixels. For the given image, a codebook is generated. The codebook along with encrypted data is transmitted. The pixels which are more frequent are given smaller codes. Given the encoded message and codebook, unique decoding takes place from the codes. This provides a significant compression ratio in both lossy and lossless stego images. The ratio of number of bits in the original image to number of bits in the compressed image provides compression ratio.

4. EMBEDDING AND EXTRACTION ALGORITHM
4.1 Encoding Algorithm
(i). The original image of size M * N is transformed to low pass and high pass components- CA, CH, CV and CD which stands for Approximation, Horizontal, Vertical and Diagonal components.

(ii). In lossy technique, using last two bits of CA components, 3 binary images of size M/2 * N/2 are embedded in the LSB of high pass coefficients.

(a) If the last two bits of CA are 11 or 00, then binary image 1 is embedded in CH, image 2 in CV and image 3 in CD band.

(b) If the last two bits of CA are 10, then binary image 1 is embedded in CV, image 2 in CD and image 3 in CH band.

(c) If the last two bits of CA are 01, then binary image 1 is embedded in CD, image 2 in CH and image 3 in CV band.

This perfectly shuffles the binary images in the coefficients.

(iii). Stego image is prepared from inverse transform and it is Huffman encoded. Significant compression ratio is achieved.

In lossless technique, detail coefficients are expanded by multiplying by two and the binary image is embedded.

4.2 Extraction Algorithm
(i). Huffman decoding is applied to the stego image for decompression.

(ii). The stego image is transformed to low pass and high pass components- CA, CH, CV and CD which stands for Approximation, Horizontal, Vertical and Diagonal components.

(iii). In lossy technique, using last two bits of CA components, 3 binary images are retrieved from the LSB of high pass coefficients.

(d) If the last two bits of CA are 11 or 00, then binary image 1 is taken from CH, image 2 from CV and image 3 from CD band.

(e) If the last two bits of CA are 10, then binary image 1 is taken from CV, image 2 from CD and image 3 from CH band.

(f) If the last two bits of CA are 01, then binary image 1 is taken from CD, image 2 from CH and image 3 from CV band.

This perfectly shuffles the binary images in the coefficients. In lossless technique, binary image is retrieved from detail coefficients and the coefficients are divided by 2 to bring back to original form.Figure1 deals with embedding and encoding process during transmission of the image and Figure 2 depicts decoding process and extraction of the stego image after receiving the data.(a)(b)(c)(d)(e)(f) steps in the embedding and extraction algorithm are referred as Data hiding and data retrieval block in the figures.
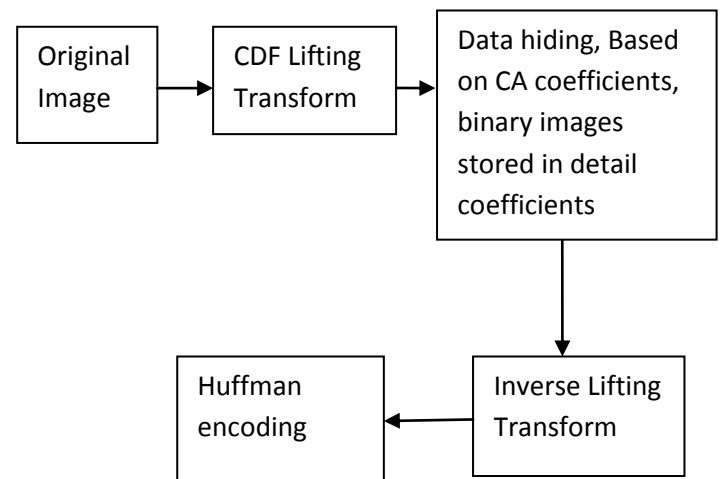

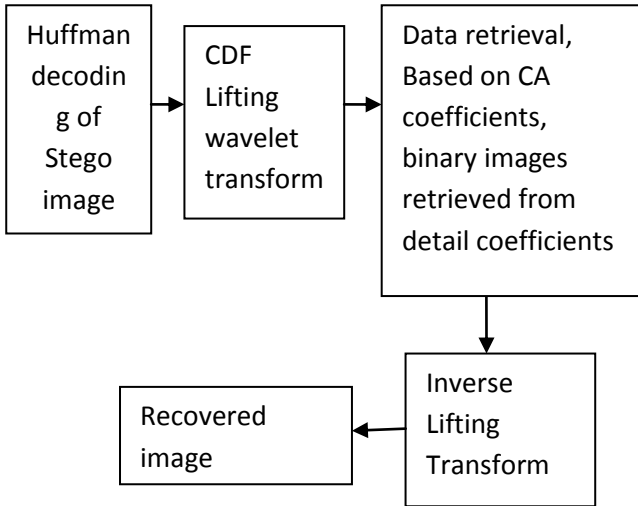
Fig 1: Steganography and compression in Transmission

Fig 2: Retrieving of images from stego image

5. RESULTS AND DISCUSSION

All the experimental simulations are performed in Matlab R2010. After embedding the binary images, compression is performed using Huffman encoding and significant compression ratio is achieved. The Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) are calculated using the formulae

$$MSE = \frac{\sum_{M,N}[OI - EI]^2}{M * N}$$

$$PSNR = 10log_{10}\left(\frac{255 * 255}{MSE}\right)$$

where OI represents Original Image and EI represents Embedded Image. The standard images which include Lena image, Barbara image and cameraman image are taken for analysis. The obtained MSE, PSNR and compression ratio are tabulated in table 1 and the corresponding bar chart is depicted in figure 3. In the chart, I1,I2 and I3 represents Lena image, Barbara image and Cameraman image respectively.

Table 1: MSE, PSNR and Compression Ratio

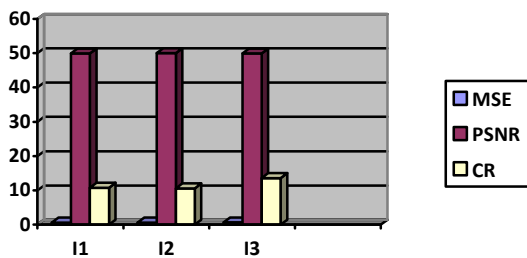| Image	MSE	PSNR	Compression ratio
Lena	0.65	49.9	10.8
Barbara	0.64	50.02	10.6
Cameraman	0.66	49.88	13.6

Fig 3: Performance of the proposed method.

5.1 Images embedded and recovered

Binary images to be embedded are shown in Figure 4. Here, Lossy method is discussed through images. Original and stego image, histograms of original and stego image corresponding to Lena grayscale image are shown in Figure 5. Original image, stego image, histograms of original and stego image corresponding to Barbara grayscale image are shown in Figure 6. Original and the stego image, histograms of the original and the stego image corresponding to cameraman grayscale image are shown in Figure 7. Cameraman image provides better compression ratio, whereas Barbara image produces high PSNR value.

Image 1   **Image 2**   **Image 3**

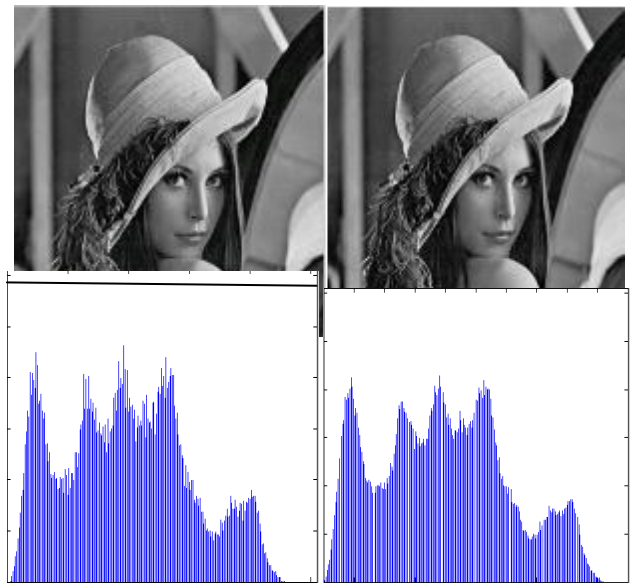**Fig 4: Three binary images which are to be embedded in detail coefficients**

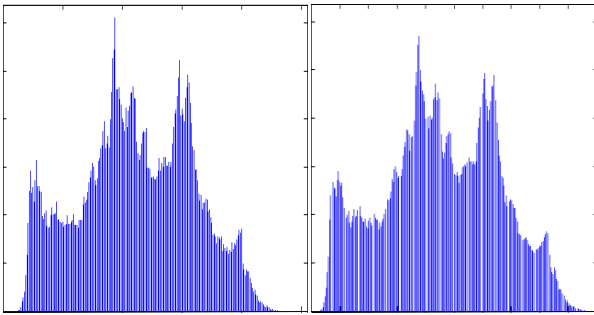**Fig 5: Original image, stego image of lena and their corresponding histograms**

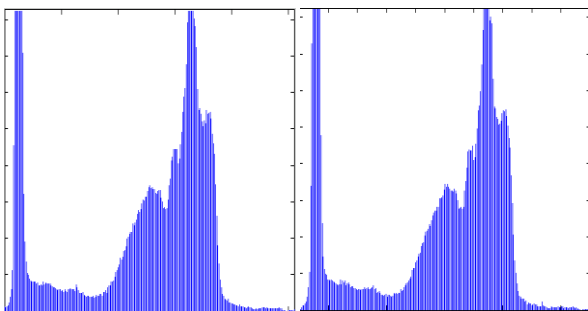**Fig 6: Original image, stego image of Barbara and their corresponding histograms**

Fig 7: Original image, stego image of cameraman and their corresponding histograms

6. CONCLUSION AND FUTURE WORK

Efficient compression ratio and PSNR are achieved using this proposed method. Also the proposed method is not vulnerable to visual attacks. In future, color image steganography is to be tried on the basis of proposed methodology. With the availability of three bands namely Red, Blue and Green in color images, the volume of information which is to be hidden can be increased. Further, different lifting wavelets such as CDF (2, 4), CDF (3, x) are to be explored and they are to be used to generate transforms in which binary images are embedded. Apart from modifying the type of lifting wavelet and incorporating color images, hiding of secret information rather than binary image are to be proposed using the lossy and lossless techniques. Medical images and data in ASCII format which are to be secured and to be transmitted clandestinely can be hidden in our proposed methodology.

7. REFERENCES

[1] Andreas Westfeld and Andreas Pfitzmann, "Attacks on Steganographic Systems Breaking the Steganographic Utilities - EzStego, Jsteg, Steganos, and S-Tool"

[2] J.Y. Hsiao. C.C. Chang. and C.-S. Chan. "Finding optimal least significant-bit substitution in image hiding by dynamic programming strategy" Pattern Recognition, 36:1583– 1595, 2003.

[3] Chan, C.K. and Cheng. L.M. "Hiding data in image by simple LB substitution" Pattern Recognition, 37: 469 – 474, 2003.

[4] D.C. Wu and W.H. Tsai "A steganographic method for Images by pixel value differencing" Pattern Recognition Letters" 24- 1613– 1626, 2003

[5] Chang, C.C and Tseng, H.W. "A Steganographic method for digital images using side match".Pattern Recognition Letters, 25: 1431 – 1437, 2004.

[6] Chen, P.Y. and Wu, W.E. "A Modified Side Match Scheme for Image Steganography" International Journal of Applied Science and Engineering, 7(1): 53 – 60, 2009.

[7] Chu, Y.P., Guo, S.W., Chan, Y.K. and Wu, H.C. "Image Hiding Based on a Hybrid Technique of VQ Compression and Discrete Wavelet Transform ",International Computer Symposium, 313-317,2004.

[8] Bao, P and Ma, X. "Image Adaptive Watermarking Using Wavelet Domain Singular Value Decomposition", IEEE Transaction on Circuits and Systems for Video Technology, 15(1):2005

[9] Andrew Westfeld " F5-a steganographic algorithm: high capacity despite better steganalysis" In Proceedings of the 4th Information Hiding Workshop, volume 2137 of LNCS, pages 289-302. Springer, 2001.

[10] J. Fridrich, M. Goljan, and D. Hogea" Attacking the outguess" In Proceedings of 2002 ACM Workshop on Multimedia and Security, pages 36. ACM Press.

[11] Chia-Chen Lin" High capacity data hiding scheme for dct-based images" Journal of Information Hiding and Multimedia Signal Processing, 1, 2010

[12] Bo Yang and Beixing Deng" Steganography in gray images using wavelet" In Proceedings of ISCCSP 2006

[13] Po-Yueh Chen and Hung-Ju Lin" A dwt based approach for image steganography" International Journal of Applied Science and Engineering, 4:275– 290, 2006

[14] V. Kumar and D. Kumar" Performance evaluation of dwt based image steganography" In Proceedings of Advance Computing Conference (IACC), 2010 IEEE 2nd International, pages 223– 228, 2010.

[15] W. Sweldens, "The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets" , Academic Press, Inc, 1996.

[16] Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps" Program for Applied and ComputationalMathematics, Princeton University, November 1997

[17] Yun Q. Shi and Huifang Sun" Image and Video Compression for Multimedia Engineering", second edition CRC Press, 2008.