International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 6, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Novel Approach of Universal Image Steganalysis using Features Derived from Gray Scale and Color Planes

Madhavi B. Desai, Prof. N. M. Patel, Vipul H. Mistry

***Abstract* — Art of hidden communication has evolved a lot with the advancement in technology. It will alter some of the media properties that may introduce few degradation or unusual characteristics. These characteristics may act as signatures that broadcast the existence of the embedded message, thus defeating the purpose of steganography.There are basically two approaches of steganalysis, one is specific approach and other is universal approach. As there are so many data hiding methods, specific steganalysis method is meaningless. That's why we need to develop a method that can defeat any steganography method and detect the existence of the hidden message. The first and most important stage of steganalysis is feature extraction. In this paper, we have used gray scale and color properties of an image to detect the existence of hidden message. Statistical moments of wavelet characteristic function, prediction error image and energy derived from R, G and B plane of image reflect the change in image properties. We have extracted the features derived from these features and proposed 39-D feature vector to train the SVM classifier. We have tested our proposed method against popular JPEG steganographic methods, spread spectrum and LSB like method. Our proposed method is able to achieve average 85% detection rate against all steganographic methods. We have used standard image database of 100 CorelDraw images of size 256x256 for our experimental purpose. This paper also compares the performance of proposed method with other existing steganalysis methods.**

***Keywords* — Universal Steganalysis, Cover-Image, Stego-Image, CF Moments, Prediction Error Image, Bayes Classifier, SVM Classifier.**

I. INTRODUCTION

In past decades digital technology has accelerated the development of multimedia system. Today so many tools are available to hide or edit the information in the media. Information hiding has been a hot research area since many years. Encryption and decryption were used in early days for secure communication. But for secure transmission only an encrypted version of information sending is not enough, and that's why hidden communication came in existence i.e. steganography (covert communication).Steganography can be possible with different cover media i.e. audio, video, text or image. Because of large amount of redundant data, Image has become most popular carrier as cover media. Large size of meaningful audio files makes them less popular compared to images. Some of the popular steganography methods used for embedding data are F5, Jsteg, LSB like method, spread spectrum method proposed by Cox et al [1], Blindhide and Steghide.

Information hiding leads to many negative effects and that's why steganalysis came in existence. Steganography may present negative effects to personal privacy, business activity, and national security also. It may be mistreated for covert communication between criminals. For example, commercial spy may thieve confidential trading or technical messages and deliver them to competitors for a great benefit by steganography. Terrorists may also use related techniques to cooperate for international attacks and prevent themselves from being traced. Some others may even think of conveying a computer virus via steganography methods. Thus it raises the concerns of developing steganalysis techniques to detect these negative effects. Even steganalysis can also serve as a measure of performance for steganographic technique.

Rest of the paper is organized as follows. Section 2 highlights previous work done in image steganalysis. The effect of steganography on various image features is discussed in section 3. Finally, section 4 describes the implementation results of proposed steganalysis method and effectiveness of various features in image Steganalysis.

II. LITERATURE SURVEY

Image can be classified in]to two classes: one with the hidden message i.e. stego image and other is without hidden message that is cover image. Image steganalysis can be considered as a task of pattern recognition to decide from which class a test image belongs to, stego or original. Two main tasks of image steganalysis are feature selection and classifier design. Both feature selection and classifier design can be evaluated by classification success rate or error rate.

In 2001, Farid [2] proposed a steganalysis method based on higher order PDF moments of wavelet Subbands. The method was tested against various steganography methods like Jsteg, EzStego & Outguess. In 2002, Avcibas et al. [3] conducted statistical analysis on the sensitivity and consistency behavior of objective IQMs, In 2003, Harmsen and Pearlman [4] proposed a method based on COM of histogram characteristic function. In 2005, steganalysis based on statistical moments of wavelet characteristic function proposed in [5], [6] and [7]. In 2009, Manjuladevi et al. [8] proposed a blind steganalysis method using histogram and DFT of an image. Ziwen Sun et. al. [9] has proposed steganalysis method based on characteristic function of wavelet subbands. In 2010, Swagota et. al. [10] proposed steganalysis of real time image by statistical attacks. A method based on neighborhood information of pixels was also proposed by Q. Guan et. al.[11] in 2011.

In this paper, we propose statistical features derived from color planes as well as gray scale. Our proposed image steganalysis method is described in next section.

International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 6, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

III. PROPOSED IMAGE STEGANALYSIS ALGORITHM

The implementation steps for the proposed steganalysis method are listed as below.

1. Convert the color RGB image into grayscale image.
2. Perform 2-level wavelet decomposition using Haar Wavelet to get 9 subbands including original image.
3. Calculate Histogram of each of the 9 subbands.
4. Calculate DFT of Histogram i.e. Characteristic Function of all the subbands.
5. Find out 1st and 2nd order moments i.e. mean and variance of all the subbands to get 18-D feature vector.
6. Using Prediction Error Image algorithm obtain prediction error image of all the subbands.
7. Now calculate 1st and 2nd order moments of each of these prediction error images to another get 18-D feature vector. This makes total of 36-D feature vector which is used further to train and test the Bayes Classifier.
8. Energy of DFT coefficients of R, G and B color planes.

This makes total of 39-D feature vector combining statistics derived from gray scale image, its prediction error image and statistics derived from its color planes. The proposed scheme uses moments derived from gray scale and color planes. The effect of data hiding on these features is discussed in next section.

IV. FEATURE SELECTION

Image steganalysis means analysis of stego images. That's why before implementing the detection method for steganography, first and most important part is feature selection. Features should be sensitive to hiding methods. In other words we can say features should act differently for original cover image and stego image. Larger the feature difference means better the feature selection. Features should be general i.e. features should be sensitive to all general data hiding methods not to specific method. Often it is very hard to achieve high recognition rate with single feature, so we need to make M-D feature vector. Feature should be independent to each other.

A. Effect of Steganography on Histogram

If the image grayscale values or the wavelet coefficient values are treated as a random variable then histogram of a digital image or a wavelet subband can be considered as the probability mass function (PMF). Characteristic function (CF) and the PDF (histogram) are similar to a Fourier transform pair (with the sign in the exponential reversed) [7]. Histogram or PDF is expected to be more flat than that of the original image [7]. This type of change in histogram can be used in steganalysis. Figure 1 shows the comparison of the histograms of original image and stego image obtained by Cox et al. [1] spread spectrum method.

We can clearly observe the difference in histogram of an image after steganography. One more thing that we can observe is the variation at the origin x=0. The peak of the histogram at origin is very much sensitive to the data hiding process. This variation in histogram can be effectively used to obtain features for steganalysis.

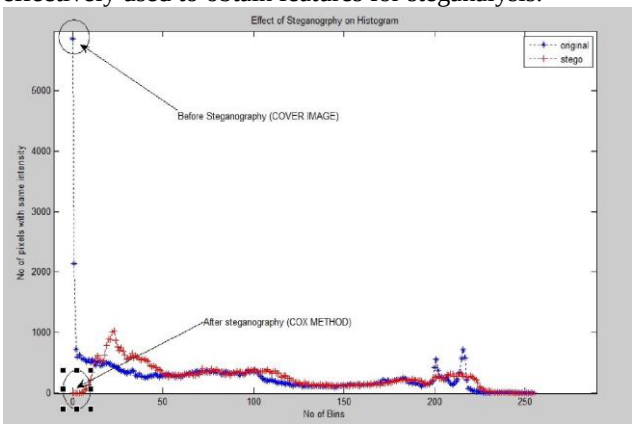


Fig.1. Variation in Histogram after Steganography

The inverse transform of characteristic function produces the PDF as follows

$$h(x) = \int_{-\infty}^{\infty} H(f) e^{-j2\pi fx} df \tag{1}$$

n-th derivative of the histogram evaluated at the origin, x=0 is, [6]

$$\left| \left(\frac{d^n}{dx^n} h(x) \Big|_{x=0} \right) \right| \le 1(2\pi) \int_0^{\infty} f^n \left| H(f) \right| df \tag{2}$$

As per equation 2, the n-th moment is proportional to the n-th derivative of the histogram at origin x=0. Figure 1 shows that after embedding data in image the histogram value becomes 0 at the origin x=0 and histogram becomes more flat. Equation 2 states that this value is propositional to moment of histogram. This makes moment as one of the effective feature in the Image Steganalysis.

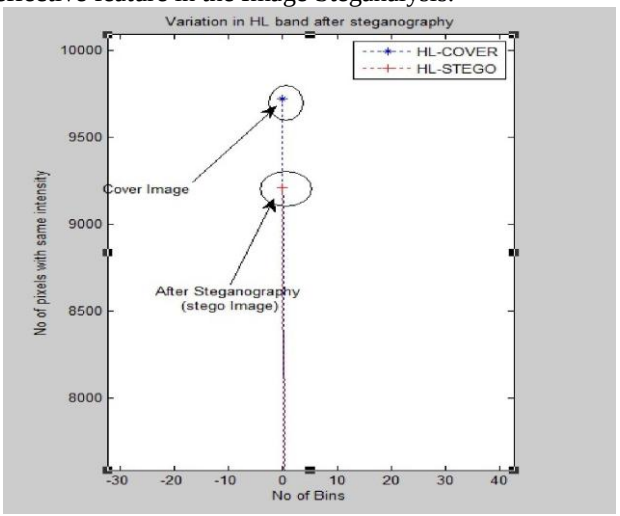


Fig.2. Variation in the pick of histogram of HL frequency band

For high Frequency bands like HH, HL and LH the DWT coefficients have mean values around at x=0.In figure 2, the histogram of HL band after 1-level DWT for original and stego image (by Cox method) is shown.

As we can see in figure 2, both the histograms are

International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 6, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

almost identical except the pick at the origin x=0. The pick of histogram after embedding message is smaller compared to the same of original. This variation in the mean of histogram (1st moment) can be considered as useful feature in image Steganalysis.

Similar type of variation is observed in other high frequency bands like HH and LH at all DWT levels of decompositions. For low frequency band LL at all the levels of DWT histogram will have similar sort of change what we observed in figure 1 which shows the variations in original and stego image. This result is true for all the images that we have taken for the experimental purpose. .

B. Effect of Steganography on CF Moments

The sensitivity of histogram against hidden messages increases if we consider the variations in DFT of a histogram i.e. Characteristic Function. The histogram and Characteristic function for low frequency band are shown in Fig. 3(a) and Fig. 3(b). Fig. 3(a) shows the histogram of LL band for original image and stego image (Cox SS Embedding method) and Fig. 3(b) shows Characteristic Function of original image and stego image.

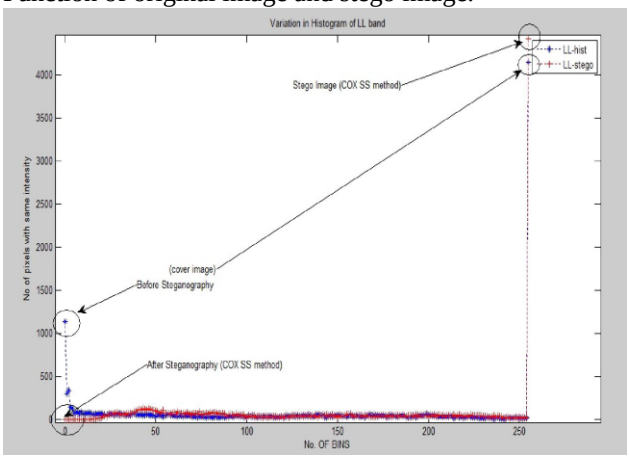


Fig.3. Variations in Histogram of LL Band

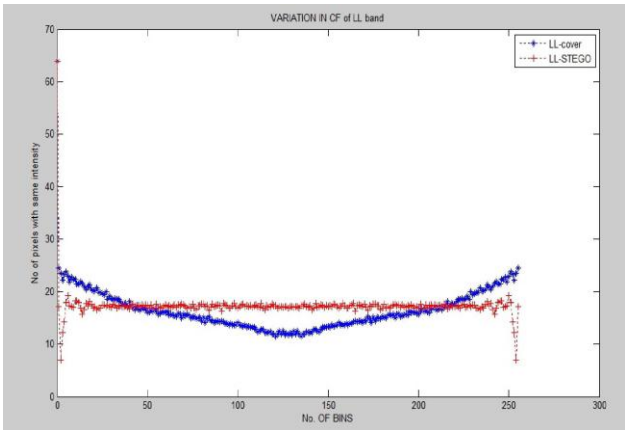


Fig.4. Variation in Characteristic Function of LL bands between original and stego images

We can observe that after embedding data in original image there is variation in entire histogram while in case of CF the pick of histogram is very much sensitive of data hiding process. We can say that characteristic function is more sensitive to data hiding process compared to histogram and that's why CF moments are more sensitive to the data hiding process compared to only histogram moments. Because of these variations in Characteristic Function of original and stego image we have used CF moments as one of the features for image steganalysis work. .

C. Prediction Error Image

Steganography method hides data in cover image, which in turn causes distortion in the original image. This type of distortion may be rather weak and covered by other types of noises, including those due to the unusual feature of the image itself. The prediction image algorithm given in [7] can be used to predict each pixel grayscale value in the original cover image by using its neighboring pixels' grayscale values, and a prediction error image is obtained by subtracting the predicted image from the test image. This prediction-error image removes information other than that caused by data hiding. Ultimate goal of prediction error image is to erase the original cover image content. The prediction algorithm is expressed below [7].

$$\hat{x} = \begin{cases} \max(a,b); c \le \min(a,b) \\ \min(a,b); c \ge \max(a,b) \\ a+b-c; otherwise \end{cases} \quad (3)$$

Where, a, b, c are is the context of the pixel x under consideration, $\hat{x}$ is the prediction value of x. The location of a, b, c can be illustrated below.

x	b
a	c

Predictor tends to predict b in cases when vertical edge exists on left of current position, a in case of horizontal edge above the current pixel position or $a + b - c$ if no edge is detected. The later one expresses the smoothness in absence of the edges. We can see that the predicted value is median of three fixed predictors: a, b and $a + b - c$. This algorithm is known as "median edge detector". We can see that the predicted value is median of three fixed predictors: a, b and $a+b-c$. This algorithm is known as "median edge detector". .

D. Impact of Image Steganography on Prediction Error Image

Prediction Image represents the correlation of pixel with neighboring pixels. When we hide something in image this correlation breaks and distortions are created. These distortions can be extracted if we obtain the error difference between original and stego images. For experimental purpose we have used one 256x256 size corel draw image and we have also generated stego image using Cox SS method. Then we applied prediction image algorithm. We have converted the original image to gray level as we will be using gray image for feature extraction. The prediction algorithm is used to generate original predicted image and stego predicted image.

International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 6, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Fig.4. (a) Original and Stego F5 grayscale image (Corel Draw Image No PH00431)

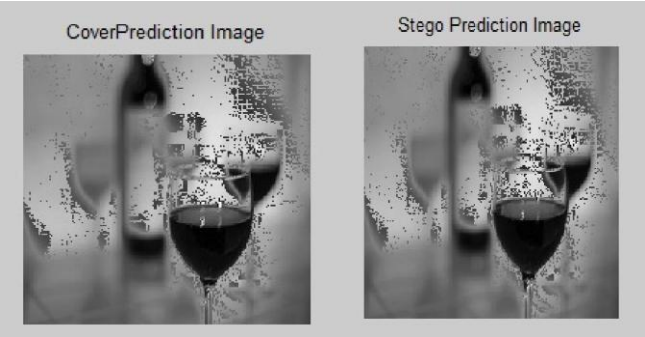


Fig.4. (b) Cover and Stego prediction images

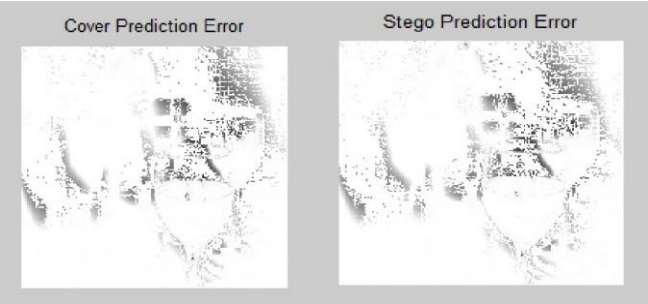


Fig.4. (c) Cover and Stego prediction error images

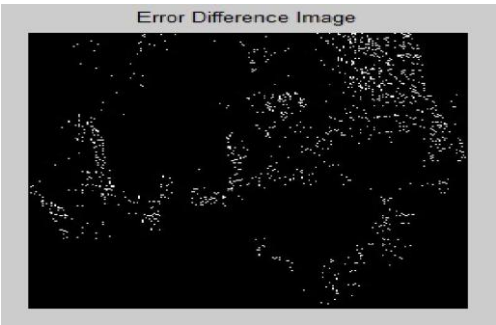


Fig.4. (d) Error difference Image

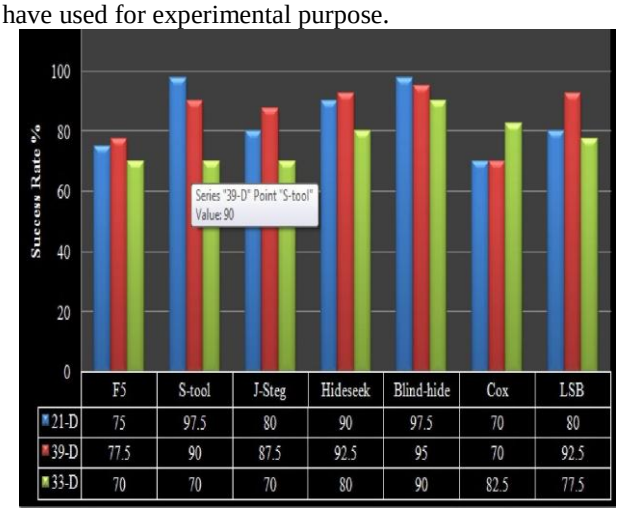
Fig. 4 (c) shows the difference in both the prediction images which indicates the presence of hidden message. As a steganalyst we can use this difference and calculate Statistical features which can be effective in the process of Image Steganalysis.

V. IMPLEMENTATION RESULTS

The experimental results are taken for the standard database of 800 Corel Draw images of size 256x256. We have taken results considering various dimensional feature vectors that are used to train and test the Bayes and SVM Classifiers. The success classification rate depends on two parameters. True Positive (TP) and True Negative (TN) image detection. True positive means detection of stego image as a stego image and True Negative means detection of cover image as a cover image. Image Steganalysis algorithm is said to be effective if it can detect both cover as well as stego image with same sort of efficiency. Average of TP and TN detection gives the success classification rate for any algorithm.

Ideally TP should be as large as possible so that one can easily detect stego image but at the same time if TN is poor then algorithm will not be efficiently differentiate between cover image and stego image. In such cases cover images are also detected as stego images which lead to erroneous results. That's why image steganalysis algorithm should have comparable detection rate for both cover as well as stego images. Generally 3/4th of the total images from the image database is used to train the classifier and remaining images are used to test the performance of the algorithm. We have also used the same approach to check the effectiveness of the steganalysis algorithms.

Fig. 5 shows the evaluation of proposed 39-D steganalysis method against various popular steganography methods. We have tried various combinations of gray scale and color plane features. Figure shows the result with some of the combinations we have used for experimental purpose.



	F5	S-tool	J-Steg	Hideseek	Blind-hide	Cox	LSB
21-D	75	97.5	80	90	97.5	70	80
39-D	77.5	90	87.5	92.5	95	70	92.5
33-D	70	70	70	80	90	82.5	77.5

Fig.5. Image Steganalysis using combined features against

Various steganography methods with SVM Classifier Description of various feature vectors are as follows:

1. 21-D :

- 18-D gray scale features – 2 level DWT and first two CF moments of each band including original image makes 18-D features.
- 3-D color features – energy of DFT coefficients of R, G and B planes

2. 33-D :

- 18-D gray scale features as mentioned above.
- 15-D selected color features
  - 2nd moment of R plane, 2nd and 4th moments of DFT of R plane which makes 3-D features.
  - 1st, 2nd and 4th moments of DFT of G plane which makes another 3-D features.
  - First four moments of DFT of B plane which makes 4-D features.
  - First order moment mean of the difference of histogram and DFT for B and G planes that makes other 2-D features.
  - Finally energy of DFT coefficients of all the three planes makes 3-D feature vector. This makes total of 15-D features.

3. 39-D:

- 18-D gray scale features – 18-D gray scale features as mentioned above.
- 18-D prediction error features – first two moments of prediction error images of 9 subbands after 2- level DWT makes another 18-D gray scale features.

4. 3-D: Energy of DFT coefficients of R, G and B planes

Table 1: Comparison of proposed 39-D feature vector steganalysis method with other existing methods

Sr. No.	Method	Classifier	Cox %	LSB %	S-Tool %	Hide Seek %	Blind Hide %	F5 %	Jsteg %
1	24-D [8]	OCSVM	-	-	96	-	-	-	-
2	78-D[7]	Neural Network	98	98	-	-	-	-	-
3	24-D Our implementation	LIBSVM	80	65	65	67.5	77.5	52.5	55
4	78-D Our Implementation	LIBSVM	**77.5**	90	90	**92.5**	90	70	85
5	Farid's Method Implemented By Ziwen[9]	BP Neural Classifier	-	-	89.9	-	-	63.3	**98.9**
6	Harmsen's method implemented By Ziwen [9]	BP Neural Classifier	-	-	79.7	-	-	57.9	83.7
7	Method by Ziwen et al [9]	BP Neural Classifier	-	-	94.4	-	-	76.4	89.4
8	**39-D Our Proposed method**	LIBSVM	70	**92.5**	**90**	**92.5**	**95**	**77.5**	**87.5**

VI. Conclusion

The proposed steganalysis scheme makes use of features derived from moments and prediction error image from gray scale and R, G and B planes. Results observed in section IV clearly demonstrates the sensitivity of these features against various data hiding methods. The steganalysis method was evaluated for standard database of 100 CorelDraw images of 256x256 sizes against various types of popular steganography methods like LSB, F5[12], Jsteg[17], Blindhide[15], Hideseek[16] and Spread spectrum method proposed by Cox et al. The average detection rate of stego and cover image is 90% for all the methods except F5 and Cox. The performance of this algorithm was also compared with other existing steganalysis methods as mentioned in table 1. Our proposed method provides result as good as 78-D feature vector method [7] with half of the feature vectors.

This method provides quite a good detection rate against DCT based spread spectrum steganography method by Cox et al [1] and LSB method but suffers to get the same detection rate against JPEG steganography method F5 [12] and S-tool [18] based steganography methods which proves the robustness of these methods.

References

[1] Ingemar J.Cox, Joe Kilian, F.Thomson Leighton and Talal Shamoon,"Secure Spread Spectrum Watermarking for Multimedia", *IEEE Transaction on image processing*, Vol. 6, No. 12, December 1997

[2] H. Farid, "Detecting steganographic messages in digital images", *Technical Report*, TR2001-412, Dartmouth College, Hanover, NH,2001.

[3] I. Avcibas, N. Memon, B. Sankur, "Statistical evaluation of image quality matrics," *In Journal of electronic imaging*, vol. 11(2), April 2002, pp. 206-223.

[4] J.J. Harmsen, W.A. Pearlman, "Steganalysis of additive noise modelable information hiding", *In: Proceedings of the SPIE, Security, Steganography, and Watermarking of Multimedia Contents V*, vol. 5020, 2003, pp. 131–142.

[5] Y.Q. Shi, G.R. Xuan, C.Y. Yang, J.J. Gao, Z.P. Zhang, P.Q. Chai, D.K. Zou,C.H. Chen, W. Chen, "Effective steganalysis based on statistical moments of wavelet characteristic function", *In Proceedings of IEEE International Conference on Information Technology: Coding and Computing*, 2005, pp. 768–773.

[6] G.R. Xuan, Y.Q. Shi, J.J. Gao, D.K. Zou, C.Y. Yang, Z.P. Zhang, P.Q. Chai,C.H. Chen, W. Chen, "Steganalysis based on multiple features formed by statistical moments of wavelet characteristic functions",*In Proceedings of seventh International Information Hiding Workshop*, Lecture Notes in Computer Science, vol. 3727, Springer, Berlin,2005, pp. 262–277.

[7] Y.Q. Shi, G.R. Xuan, D.K. Zou, "Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network", *In Proceedings of IEEE International Conference on Multimedia and Expo*, 2005, pp. 269–272.

[8] T. H. Manjula Devi1, H.S.Manjunatha Reddy,2 K. B. Raja3Venugopal K. R3 and L. M. Patnaik4, "Detecting Original Image Using Histogram, DFT and SVM", *International Journal of Recent Trends in Engineering*, Vol. 1, No. 1, May 2009.

[9] Ziwen Sun, Hui Li, Zhijian Wu, Zhiping Zhou, "A Image Steganalysis Method Based on Characteristic Function Moments of Wavelet Subbands", *In: Proceedings of IEEE International Conference on Artificial Intelligence and Computational Intelligence*, 2009.

[10] Swagota Bera, Monisha Sharma, "Steganalysis of Real time Image by Statistical Attacks", *International Journal of Engineering Science and Technology*, Vol. 2(9), 2010, pp. 4396-4405.

[11] Q.Guan, Jing Dong and Tieniu Tan, "An effective Image steganalysis method on neighbourhood information of pixels", *18th IEEE International Conference on Image Processing*, 2011, pp. 2777-2780.

[12] Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt, "Digital Image Steganography: Survey and Analysis of Current Methods".

[13] www.coreldraw.com

[14] Libsvm, http://www.csie.ntu.edu.tw/~cjlin/libsvm

[15] Blindhide, http://www.nic.funet.fi/pub/crypt/steganography

[16] Hideseek, http://www.nic.funet.fi/pub/crypt/steganography

[17] Jsteg, http://www.tiac.net/users/korejwa/jsteg.htm
[18] A.Brown, "S-tools version 4.0".
http://members.tripod.com/steganography/stego/s-tools4.html

AUTHOR'S PROFILE

Madhavi Desai
is associated with S.N.P.I.T & R.C, Umarakh Gujarat, India. She has received ME degree in Computer Engineering. Her research domain includes Image Processing, Data Mining, Soft Computing. Email: desaimadhavi30@gmail.com

Dr Narendra M. Patel
received B.E. degree in electronics engineering from M.S. University, Baroda in 1993, M.E. degree from M. S. University, Baroda in 1997 and Ph.D. from SVNIT, Surat in 2012.
He is currently Associate Professor in Computer Engineering Department, B.V.M. Engineering College, V.V.Nagar, India. His research interests include Digital Image Processing, Real time operating systems and computer graphics.
Email : bvm_nmp @yahoo.com

Vipul H. Mistry
is associated with S.N.P.I.T & R.C, Umarakh. He has received M.Tech degree from SVNIT, Surat in 2009. His research interests include Digital Image Processing, Signal Processing, VLSI design.
Email : vipul.vpl@gmail.com