Tri-Way Pixel Value Differencing Steganalysis; a Normalized Image Histogram approach

Archana O. Vyas[1], Dr. Sanjay V. Dudul[2]

[1]*Ph.D. Student, Department of Applied Electronics, S.G.B. Amravati University, Amravati, India*
[2]*Professor and Head, Department of Applied Electronics, S.G.B. Amravati University, Amravati, India*

E-mail: [1]nyasa.archana@gmail.com; [2]sanjay.dudul@gmail.com

Abstract: To enhance the hiding capacity and security of the well known pixel value differencing (PVD) steganographic method, a new method named Tri-way pixel value differencing (TPVD) was introduced. TPVD is a spatial domain steganographic method in which both horizontal and vertical edges of cover image are utilized for embedding. In TPVD, Use of three different directional edges of cover image for embedding makes it different from PVD, which utilizes only one direction. In this paper a novel method of TPVD steganalysis is proposed, based on the observation that, the difference value of the two sides of a normalized histogram of image under consideration if found to be a positive value other than "0", then it can be classified as a stego image, where as if this difference is "0" it is a non stego image. The proposed steganalyser can classify test images as stego or cover with 98% accuracy when they contain any secret image. Difference in histogram is plotted for the stego and non stego images, Parameters like, True positive rate (TPR), false positive rate (FPR) and detection rate are evaluated thereafter, and a satisfactorily high value is obtained. The algorithm can be run on two stego images simultaneously, so it can be applied on bulk image data base.

Keywords: Cover image, FPR, steganalysis, TPVD, TPR

1 Introduction

In the modern world, data security has become an important issue in many disciplines such as digital copyrights, online transactions and military communication. Steganography is concealing the secret message so that the very existence of it is not detectable. The aim in steganography is to send a secret message via the coverage of career signal. One dedicated career for this secure connection can be an image. After embedding process, a final image is obtained which is known as stego-image. The main purpose in steganography is to use the most capacity of the cover media in a way that its statistical properties have the minimum changes. In other words there should be a balance between these two terms and they can be used as evaluation factors of a steganographic technique. On the contrary, the steganalysis task is to detect the existence of hidden data. Steganalysis tries to identify the weaknesses of a steganographic method. Therefore, in order to counter it one has to improve the security [1]. In general, steganalyzers may be classified into two types: targeted and blind. Targeted steganalyzers aim to identify the existence of hidden data embedded by a specific method, while blind steganalyzers develop techniques which are independent of steganographic algorithms [2].
There are two kinds of image steganography techniques: spatial-domain and transform domain based methods. Spatial domain based methods embed messages directly in the intensity of pixels of images. For transform domain based ones, images are first transformed to another domain (such as frequency domain), and messages are then embedded in transform coefficients. A popular digital Steganography technique is so-called least significant bit (LSB) embedding. In the LSB embedding technique, The secrete message, is embedded in the LSBs of the cover object pixels directly [3]. Wu and coworkers [4] presented a steganographic method based

on Pixel_Value Differencing (PVD). They divide the cover image into a number of non-overlapping two pixel blocks. Each block is categorized according to the difference of the gray values of the two pixels in the block. A small difference value indicates that the block is in a smooth area and a large one indicates that it is in an edged area. The pixels in edged areas may tolerate larger changes of pixel values than those in the smooth areas. Therefore, it is possible to embed more data in edged areas than in the smooth areas. All possible difference values are classified into a number of ranges and the number of bits which can be embedded in a pixel pair is decided by the width of the range that the difference value belongs to. It is claimed that the changes in the resulting stego-image are unnoticeable [4]. Chang [5] proposed a novel steganographic approach using tri - way pixel-value differencing (TPVD), to increase the hiding capacity of original PVD method referring to only one direction. In their approach, three different directional edges are considered and they effectively adopted to design the tri-way differencing scheme. Also, to reduce the quality distortion of the stego-image brought from setting larger embedding capacity, an optimal approach of
Selecting the reference point and adaptive rules were presented [5].

2 Review of reported work on steganalysis

Zhang et al, [6], proposed a steganalysis method which is based on a physical quantity derived from the transition coefficients between difference histograms of an image and its processed version produced by setting all bits in the LSB plane to zero. This quantity is claimed to be a good measure of the weak correlation between successive bit planes and can be used to discriminate stego-images from cover images. In a paper by Zhang [7] PVD was successfully attacked. This was done by analysis of the histogram of the stego image. A recent steganalysis scheme against modulus PVD proposed in [8] employs steganalytic measures such as fluctuations of the histogram, asymmetry of “1” and “−1” histogram values and abnormal increasing of the “0” histogram value to detect shortcomings of the modulus PVD steganography. Another group of steganalysers are universal or blind steganalysers that classify images without any knowledge of steganographic schemes. These methods try to blindly classify an image as a cover or stego image. Goljan et al. introduced a steganalysis scheme in [9], consisted of a blind classifier whose features are selected in the wavelet domain as the higher-order absolute moments of the noise residual. Also, in [10] another blind steganalysis method was proposed which aims to detect especially spatial domain steganographic methods by employing singular value decomposition transform. Nazanin Zaker & Ali Hamzeh [11] proposed a reliable steganalysis method which detects TPVD steganography. Despite of existing steganalysis on modified versions of PVD, which are sensitive to equalization of bars in the pixel difference histogram or unusual increasing of “0” histogram value.

In the proposed method, first normalized histogram of image is plotted and then difference values between two sides of histogram up to five to six points are calculated. Depending on this difference value, it is declared as stego image or non stego image. The rest of the paper is organized as follows: in section 3, TPVD steganography scheme on two cover images is briefly explained; proposed steganalysis scheme is described in section 4. Section 5 is devoted the evaluation parameters. Experimental results and conclusion of the paper are described in sections 6 and 7 respectively.

3 TPVD Steganography on two cover images[6]

The tri-way pixel-value differencing method applied on selected two cover images is as follows.

1) Calculate four difference values $d_{i(x,y)}$ for four pixel pairs in each block of two cover images distinctly.

$$d_{0(x,y)} = p_{(x+1,y)} - p_{(x,y)} \quad (1)$$

$d_{1(x,y)} = p_{(x,y+1)} - p_{(x,y)}$ (2)

$d_{2(x,y)} = p_{(x+1,y+1)} - p_{(x,y)}$ (3)

$d_{3(x,y)} = p_{(x+1,y+1)} - p_{(x,y+1)}$ (4)

2) For both the cover images, Using $|d_{i(x,y)}|$ where (i =0,…,3) to locate a suitable $R_{k,i}$ in the designed range table, that is to compute $j = min_k(uk - d_{ix,y)}$ where $u_k \geq |d_i|$ for all $1 \leq k \leq n$. Then the located range can be represented by $R_{j,i}$.
3) Compute the amount of secret image data bits t_i that can be embedded in each pair by $R_{j,i}$ for the cover image C_1 and C_2. The value t_i can be estimated from the width $w_{j,i}$ of $R_{j,i}$, this can be defined by $t = [\log_2 w_{j,i}]$.
4) Read t_i bits from the binary secret image data and transform the bit sequence into a decimal value b_i.
5) Embed the secret image in cover image blocks according to the hiding capacity of two cover images.
6) Calculate the new difference value $d_i'_{(x,y)}$ given by $d_i' = l_{j,i} + b_i$ if $d_{i(x,y)} \geq 0$ $d_i' = -(l_{j,i} + b_i)$ if $d_{i(x,y)} < 0$ to replace the original difference $d_{i(x,y)}$ [5].
7) Modify the values of P_n and P_{n+1} by the following formula for both the cover images C_1 and C_2.

$$(P'n, P'n+1) = \left(Pn - \left[\frac{m}{2}\right], Pn + 1 + {}_{m2}\right) \quad (5)$$

Where P_n and P_{n+1} represent two pixels in P_i and $m = d'_n - d_n$. Until now, to embed the secret data into the pixel pair (P'_n ,P'_{n+1}) is done by changing the Values of P_n and P_{n+1}. Now, the new block which is constructed from all pixel pairs and embedded with secret data is generated in both the cover images depending on the hiding capacity of the two cover images C_1 and C_2, respectively. [6]

4 Proposed TPVD Steganalysis method

TPVD Steganography increases the embedding capacity and thus efficiently preserves the visual qualities of the stego-images [12]. TPVD embeds secret bits in both diagonal and vertical edges of the cover-images in addition to the horizontal edges that are used in the original PVD method. However, since TPVD also embeds secret bits in the difference values between two adjacent pixel pairs, probabilistic distribution of difference values, which have been altered during embedding procedure can be considered as valuable statistical information in TPVD for a steganalysis procedure [11].

In the proposed work, it is shown that with a view to provide a steganalysis procedure for TPVD embedding, calculating histogram difference values of two sides, of the normalized histogram of an image and then comparing this difference with a threshold value, could be a useful approach. Histogram analysis before and after embedding secret image could be used to provide valuable data for the proposed steganalysis algorithm.

4.1 Proposed Algorithm

To classify the stego images and non stego images distinctly, the algorithm is described in the following steps.

1) Read two stego-images.
2) Stego images sized n×m, are converted in image matrices of 2×2 block size.
3) For each block, four pixel values $p_{(x,y)}$, $p_{(x,y+1)}$, $p_{(x+1,y)}$, *and* $p_{(x+1,y+1)}$ are computed.
4) The relevant pixel differences *di, (i=0 to 3)* are computed using Eq.1, Eq.2, Eq.3 and Eq.4.
5) These difference values are accumulated as a vector, say V.
6) Get the histogram of accumulated difference values for minimum points.
7) Normalize the histogram.

8) Get the difference values between two sides of histogram up to nearly five to six points.
9) If the difference of both ends of histogram is "0", then no embedding done and declare the image under consideration as normal image.
10) If the difference of both ends of histogram is other than "0", but positive value, the images under consideration is detected as embedded image and declare it as stego image.
11) Calculate the True positive rate, false positive rate and classification rate and display their values on GUI.

The histogram of "V" vector values which is a graphical display of tabular frequencies can be considered as a simple tool for the statistical analysis of "V" vector values and could be used as a tool to compare the characteristics of a cover image before and after embedding secret image.

5 Performance parameters

The performance parameters that are needed to be evaluated after TPVD steganalysis are described as follows.

5.1 True Positive Rate (TPR)

It is defined as the ratio of the number of correctly classified images out of the overall test images. TPR should be as high as possible [11].

5.2 False Positive Rate (FPR)

It represents the ratio of the wrongly classified the plain images as stego ones. FPR should be as low as possible [13].

5.3 Classification Rate

Classification rate is defined as the average of positive detection *(PD)* and negative detection *(ND)*, and it is given by

$$(PD + ND) / 2 \quad (6)$$

Where, Positive Detection *(PD)* is classifying the stego images, correctly and Negative Detection *(ND)* is classifying the non stego images, correctly [14].

6 Simulation Results

The simulation is accomplished by taking three image data base of different number of cover images and corresponding stego images and algorithm is implemented in MATLAB. For all the three databases the simulation procedure followed, is comprised of the following steps,

a) Selections of two cover images and plotting their Histogram.
b) Selection of corresponding stego images and plotting their Histogram.
c) Plotting the histogram difference for the stego image and corresponding cover image.
d) Calculation of true positive rate (TPR), false positive rate (FPR), and Classification rate.

As illustrated below Fig.1 (a) shows the histogram of cover images "leena" and "girl", Fig. 1(b) depicts the histogram of the same two stego images, Fig. 1(c) gives the histogram difference of same two stego images and two cover images ('leena" and "girl"). Fig. 1(d) shows the calculated values of TPR, FPR and classification rate for the image data set 1, of which image "leena" and "girl" is a part.

Fig.2 (a) demonstrates the histogram of cover images "tulip" and "fruit", Fig. 2(b) shows the histogram of the same two stego images, Fig. 2(c) depicts the histogram difference of same two stego images and two cover images ('tulip" and "fruit"). Fig. 2(d) gives the calculated values of TPR, FPR and classification rate for the image data base 2, of which image "tulip" and "fruit" is a part. Fig. 3(a), Fig. 3(b), Fig.3(c) illustrate same kind of results for image "Modi" and "Sachin". Fig. 3(d) shows the calculated values of TPR, FPR and classification rate for the image data set 3, of which image "Modi" and "Sachin" is a part.

The TPR, FPR and classification rate values obtained after simulation are tabulated in table.1, for three different data sets of images.

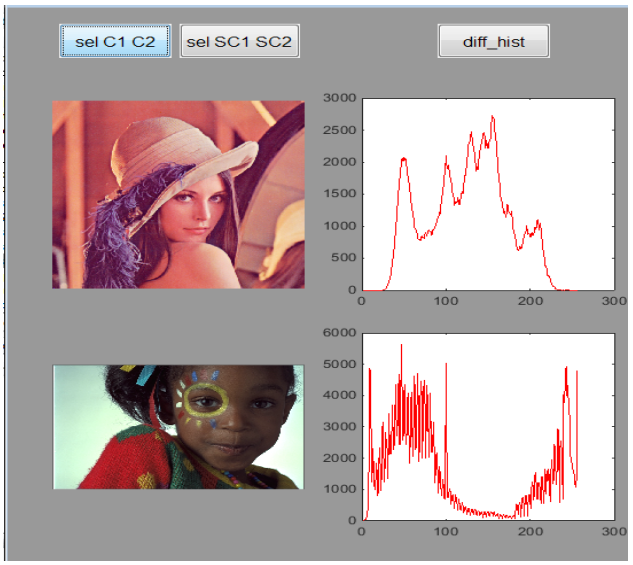

Fig.1. (a) Histogram of cover image Leena and girl

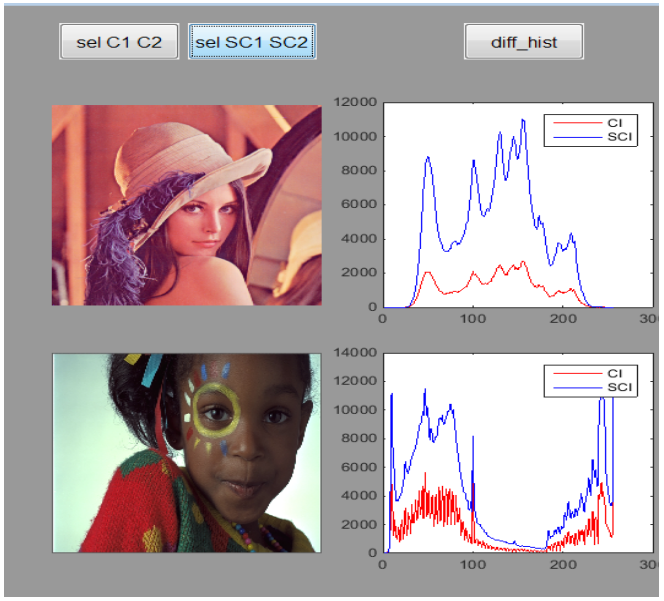

Fig.1. (b) Histogram of stego image Leena and girl (blue)

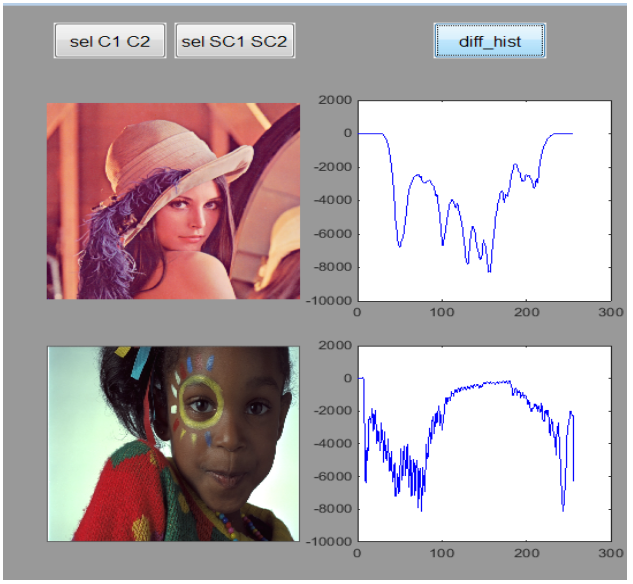

Fig.1 (c) Histogram difference two stego images and cover images

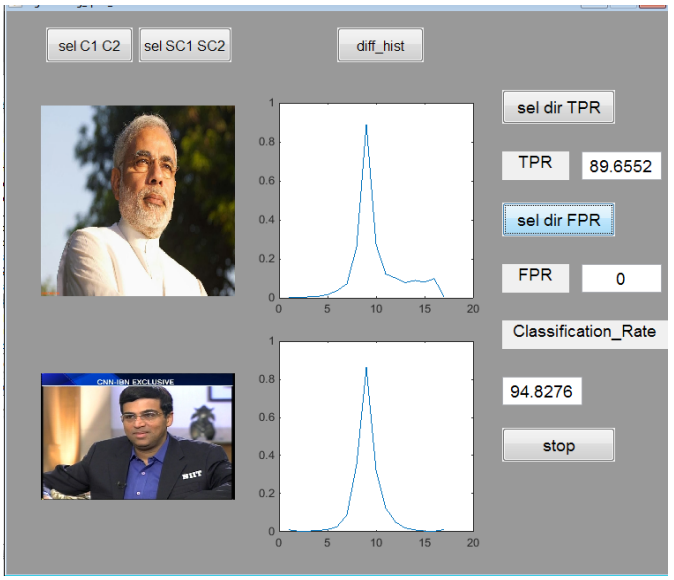

Fig.1 (d) TPR, FPR and classification rate for image database 1 (consisting of image Leena and girl)

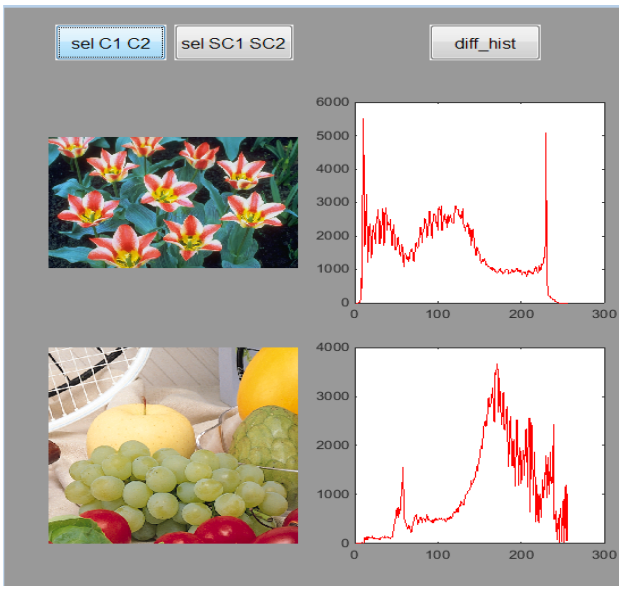

Fig.2. (a) Histogram of cover image tulip and fruit

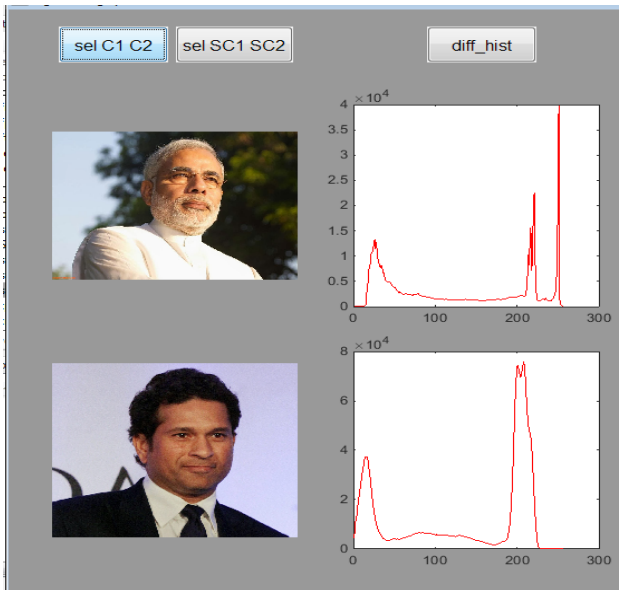

Fig.3. (a) Histogram of cover image Modi and Sachin

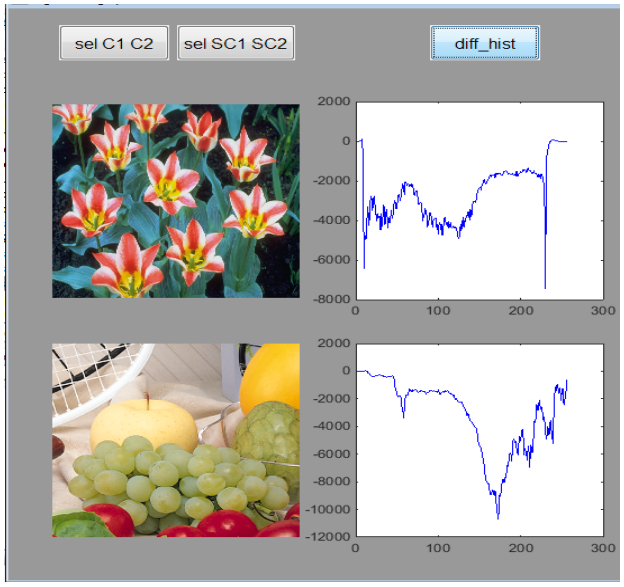

Fig.2 (c) Histogram difference two stego images and cover images

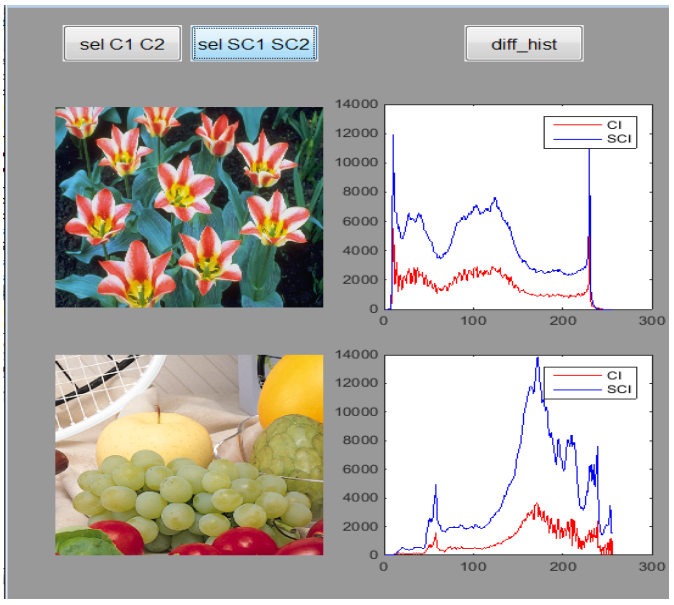

Fig.2. (b) Histogram of stego image tulip and fruit (blue)

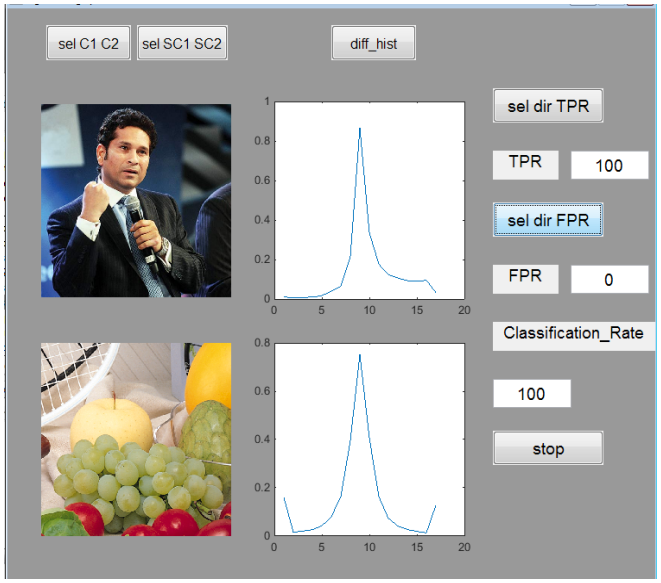

Fig.2 (d) TPR, FPR and classification rate for image database 2 (consisting of image tulip and fruit)

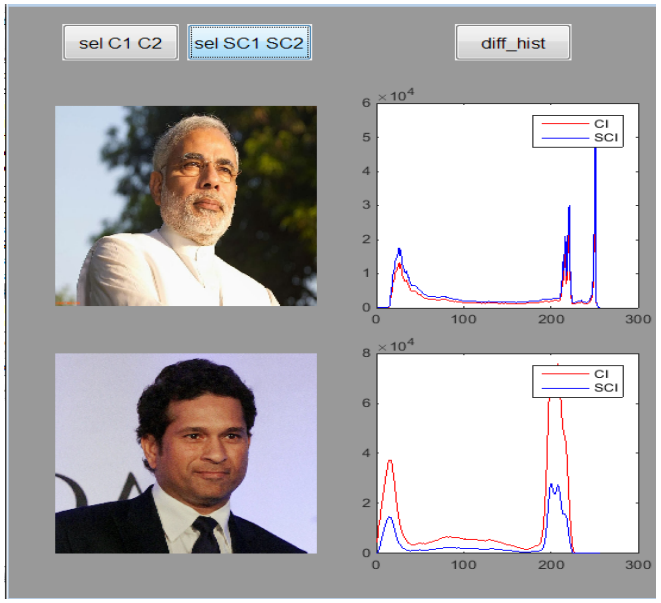

Fig.3. (b) Histogram of stego image Modi and Sachin (blue)

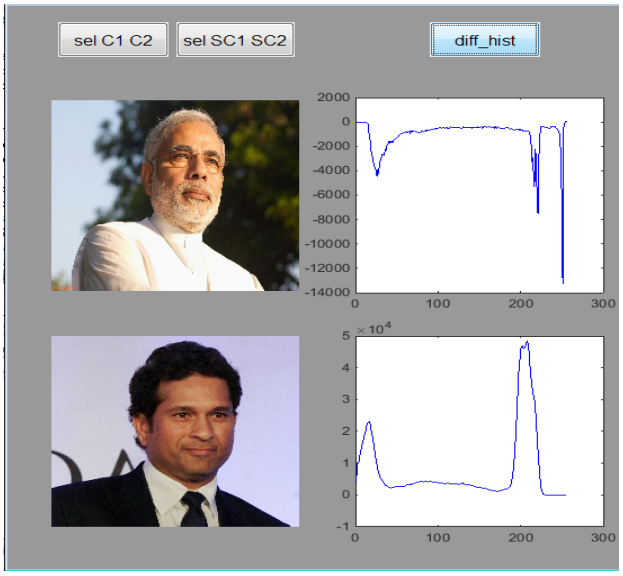

Fig.3 (c) Histogram difference two stego images and cover images

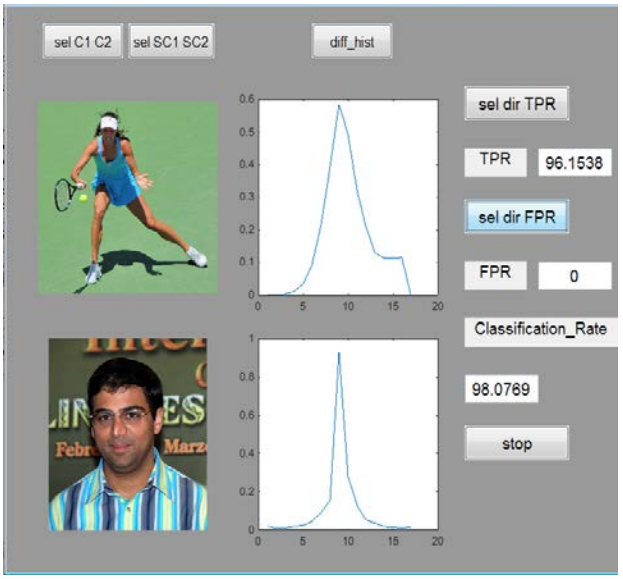

Fig.3 (d) TPR, FPR and classification rate for image database 3 (consisting of image Modi and Sachin)

Table 1. Evaluated parameters

Image data set	TPR value	FPR value	Detection Rate in %
1	89.6552	0	94.8276
2	100	0	100
3	96.1538	0	98.0769

As illustrated by above table, satisfactorily high value of true Positive rate (TPR) is obtained. Zero value of false positive rate (FPR) indicates that almost no false detection has been done by the proposed algorithm, which means that none of the stego images is detected as non embedded. The detection rate achieved is also satisfactory when tested for a large database of images.

7 Conclusion

Tri-way Pixel Value Differencing Steganographic method is an advanced PVD steganography technique. A steganalysis technique for the TPVD steganography is proposed in this paper, is based on pixel difference of normalized histogram, on two sided of the histogram up to five to six points. Depending on the value of this difference, the image under consideration can be declared as stego or non stego image. We have tested the proposed method on over 200 test images and according to the experimental results, the detection rate of proposed method is 98% in classifying image under consideration as stego images. A high value of TPR about 96.1538 is obtained and "0" value of FPR indicate that no false detection is noticed.

The proposed method founds a wide application in cyber security and computer forensic, where the steganalyzer can be applied to detect the presence of the threat to authorized documents.

References

[1] Nazanin Zaker, Ali Hamzeh, Seraj Dean Katebi, "Improving Security of Pixel Value Differencing Steganographic Method", Multimedia Tools Appication, Springer Science, Business Media, (2012) 58:147–166

[2] Liping Ji, Xiaolong Li, Bin Yang and Zhihong Liu, "A further study on a PVD-based steganography" IEEE 2010 International Conference on Multimedia Information Networking and Security.

[3] Vajiheh Sabeti, Shadrokh Samavi, Mojtaba Mahdavi, " Steganalysis of Pixel-Value Differencing Steganographic Method", IEEE 2010 International Conference on Multimedia Information Networking and Security.

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

[5] Ko-Chin Chang, Chien-Ping Chang, A Novel Image Steganographic Method UsingTri-way Pixel-Value Differencing, *Journal of Multimedia*, VOL. 3, NO. 2, JUNE 2008.

[6] T. Zhang and X. Ping, "A new approach to reliable detection of LSB steganography in natural images", Elsevier journal of Signal Processing, Vol. 83, 2003.

[7] X. Zhang and S. Wang, "Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security", Pattern Recognition Letters, 2004.

[8] Joo J, Hae-Yeoun L, Cong B, Won-Young Y, Heung-Kyu L (2008) Steganalytic measures for the steganography using pixel-value differencing and modulus function. LNCS, Springer Berlin 5353:476–485. doi:10.1007/978-3-540-89796-5_49.

[9] Goljan M, Fridrich J, Holotyak T (2006) New blind steganalysis and its implications. in Proc. of theSPIE, Security, Steganography, and Watermarking of Multimedia Contents VI, 6072:1–13.

[10] Gul G, Kurugollu F (2010) SVD-based universal spatial domain image steganalysis. IEEE Trans Inf Forensics Secur 5:349–353. doi:10.1109/TIFS.2010.2041826
[11] Nazanin Zaker & Ali Hamzeh,"A novel steganalysis for TPVD steganographic method based on differences of pixel difference histogram"
Published online: 25 January 2011 Springer Science, Business Media, LLC 2011 Multimed Tools Appl (2012) 58:147–166 DOI 10.1007/s11042-010-0714-9
[12] Chang K, Chang C, Huang PS, Tu T (2008) A novel image steganographic method using tri-way pixelvalue differencing. J Multimedia 3:37–44. doi:10.4304/jmm.3.2.37-44
[13] Archana O. Vyas Dr. Sanjay V. Dudul, "Study of Image Steganalysis Techniques", International Journal of Advanced Research in Computer Science, Volume 6, No. 8, Nov-Dec 2015.
[14] Wen-Nung Lie and Guo-Shiang Lin, "A Feature-Based Classification Technique for Blind Image Steganalysis", IEEE Transactions On Multimedia, vol. 7, no. 6, December 2005.