International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-5, Issue-3, May 2017
DOI: 10.21276/ijircst.2017.5.3.5

Practicing LSB Steganography in PCA Transform Field

Mahdi Koohi

***Abstract*—PCA conversion defines one main stream and two sub-streams for the color of homogenous areas. This subject has led to design some effective algorithms for the processing of colored images such as coloring, changing the color, compressing and Steganography. LSB is considered one of the oldest methods in secret information Steganography in images. There have been some various designs for this method. Using a compound method can lead to some new algorithms. This paper presents a new method for secret information Steganography.**

***Index Terms*— Steganography, PCA, LSB, Image processing, watermarking, noise.**

I. INTRODUCTION

Since human started communication, establishing a secret way to make it has been an important request. Internet explosion and its significant growth have pushed human being towards the digital world and digital data communication increasingly. In this regard, information security is a need which is felt considerable these days. In general, there are two methods to make a secret communication. The first is coding in which the information pieces are coded and these codes are not meaningful for the third person, however, the sender and receiver can decode them by their shared keys. Steganography is considered the second method in which the existence of this secret relation should be kept concealed.

In steganography, other's unawareness is used for hiding the message to send it in the safest way. For this reason, first, the essential information of one image which we want to send to others will be embedded in another host image in a way that others cannot understand outward discrepancy of the initial host image and the embedded host image, therefore legal receiver can extract and reconstruct initial embedded image. The most important researches conducted in this area are [1], [2], [7], [10], and [15].

Steganography is the science of sending a message by the camouflage of one medium in which the least change in that can be identified. In other words, the main objective of Steganography is to conceal the present of any relationship which leads to hide the main message from the observer. Watermarking, like Steganography, is a sub-branch of concealment which is practiced to place a watermarking in images for the purpose of identification and possession.

Manuscript received May 18, 2017

Mahdi Koohi, Electrical and computer engineering faculty, Tehran University, Tehran, Iran.

Unlike Steganography, the receiver is not obliged to get all the bits of the watermarking courier, but it is expected that the placed data can resist the attacks, to reduce the efficiency of performance, favorably. [14]

In general, there are three determining factors in information privacy systems; capacity, security, and resistance in which an improvement in one reduces the other. In Steganography methods, capacity and security play an important role while resistance is important in watermarking. The image for placing data is called cover image and the image obtained by placing a message through embedded algorithm is called stego image.

In a digital Steganography, the coverage media is called to any objects such as image, sound, film and etc. Which have extra bits for the placement of the hidden data in that there is no suspicion raised for the hidden information. Cover image is the output of steganography process which might not include the message, but it apparently looks like a coverage medium. Images are considered suitable covers for Steganography and there have been many supposed algorithms for the structure of different images. When researchers started their studies on colored images, the primary algorithms were the only paralleled issues for the processing of grey surfaced images. It was evident that they looked at an RGB image as a three-parallel image at the grey surface. This mistake is unfortunately observed these days. [14]

There have been loads of studies on Steganography and disguise breaking in images. Steganography usually happens in the fields of place and conversion. For instance PVD uses place and the methods such as DCT and F5 use conversion methods. [1] , [2] ,[3]

Either of these methods can hide the information in different ways, and one of the simplest methods is to use less-valued bits in the considered space. The ability to discover a message depends on the length of the message. Evidently, the less the information stored in one image, the less probable the identification the discoverable signs. The selection of image formats has a great influence on Steganography systems. A non-compressed format, such as BMP, provides an abundance of space for Steganography.

When working with colored images, we encounter the correlation of color vectors and the distribution of energy among them simultaneously. This situation is both promising for an efficient operation and a warning for impractical algorithms. Considering physical evidences, analysis of the main element PCA can be an effective method for non-correlation and energy density in colored images. On the same line, a suggestion to replace PCA convertors for the fixed color convertors might not sound

strange. Principal component analysis (PCA) is a standard tool in modern data analysis - in diverse fields from neuroscience to computer graphics - because it is a simple, non-parametric method for extracting relevant information from confusing data sets. With minimal effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structures that often underlie it. [12][13]

II. THE ANALYSIS OF THE MAIN ELEMENTS

The analysis of the main elements includes the decomposition of specific amounts in covariance matrix. Analysis of the main elements is defined as an orthogonal line transformation in mathematics in which the data are transformed to the first axis. In that the biggest variance is placed on the first axis, the second biggest goes on the second axis and the order continues the same way until the least data is placed on the last axis. The analysis of main elements can be practiced to reduce the dimensions so that the most effective elements are preserved. Fig 1 shows the new axis by PCA transformer. PCA is a transform in the vector

Space which is used mostly for decreasing the dimension of The data sets. At first, principle component analysis had been used by Karl Pearson in 1901 [12]. This analysis includes the decomposition of the eigenvalues of the covariance matrix.

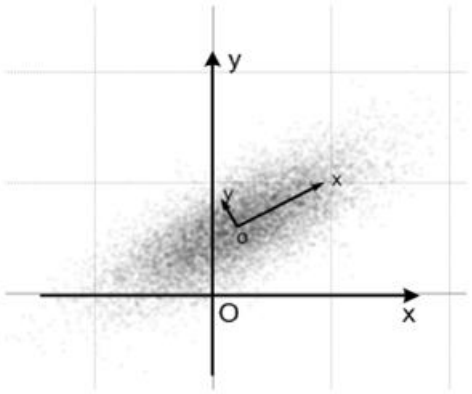

Fig.1.Axis transformation in PCA transformer

III. THE SUGGESTED ALGORITHM

The pattern is designed for color images, therefore; it shouldn't be expected to answer on binary, black and white and colored images.

IV. EMBEDDED ALGORITHM

Embedding capacity and visual quality are the two essential parameters in the stego or cover images. Embedding capacity refers maximum amount of secretive message which can be embedded in the host image, and visual quality is embedding message in the host image in a way that human eye cannot notice any difference between the new form and the original one. One criterion which is usually used for evaluating visual quality is the peak of signal to noise ratio (PSNR) between stego image and original host image and expressed in dB unit. The bigger the PSNR, the higher visual quality of estego image. In other words, it is more difficult for eye to detect stego image than to do so for host image
[5].

Principal component analysis (PCA) is a standard tool in modern data analysis - in diverse fields from neuroscience to computer graphics - because it is a simple, non-parametric method for extracting relevant information from confusing data sets. With minimal effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structures that often underlie it. [12], [4]

The first step is to select an appropriate image based on size, color and the frame. Then, the PCA cover image is taken. As a result, there will be three images and a three by three matrix which is the specific vector matrix and it is required for the PCA transformation. In order to embed data, convolutional Code is used for the selection of the number of convolutional Code in three specific images, the quality test is practiced on these images and it is compared by LSB method which will be discussed in the analysis section. After the data is embedded, and three new images are created, it is necessary to do the PCA transformation on three specific axis to obtain stego images. It sounds essential to point out that the amounts of specific images are decimal and signed so that absolute value of the true number is considered. The block of the suggested pattern is presented in fig2.

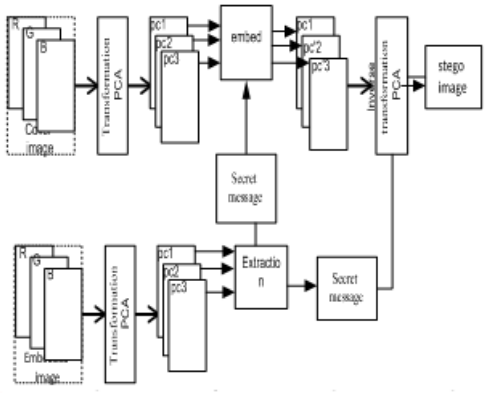

Fig.2. shows the diagram block of the proposed model.

V. EXTRACT ALGORITHM

In order to extract the information from cover images, PCA transformation needs to be obtained for the purpose of three images. Considering the selected case for the embedded data, these data will be extracted.

However, due to approximations in engineering software and computer limitations for the storage of decimal numbers, all the hidden data can't be extracted. Therefore; an error correction is required to be able to extract data thoroughly. In our proposed method, LSB is used for embedded data so that convolutional code is used. This code is easy to use and practice.

VI. CODING AND DECODING

In communication, convolutional code is an error correction code in which m bit information changes to n bit and it is sent. convolutional code, nowadays, is used to transfer huge amount of information in communication

ISSN: 2347-5552, Volume-5, Issue-3, May 2017
DOI: 10.21276/ijircst.2017.5.3.5

such as satellite, cell phones, digital radio and television. convolutional coding doesn't block the input and it reads the bits one by one and it sends them. The transferred bits are the linear function of the input bits [6].When a string of bits is received, the behavior of the coding system needs to be found to figure out the primary bit string. The probabilities of the correct bits can be achieved by the sum-product algorithm or min-sum algorithm (known as Viterbi algorithm) of the most famous ones or the other algorithms. Viterbi algorithm has the best performance to find the most probable coding path. More explanations are provided in [9], [10], [8], and [11].

VII. REVISED ALGORITHM

In order to revise the proposed algorithm, convolutional code is chosen. The message needs to be coded before embedding. The revised diagram block is shown in fig 3.

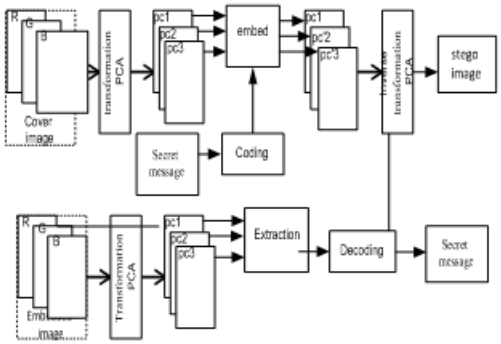

Fig .3.the revised diagram block of the proposed model

To extract the message, after the message is stored in specific images, the secret message is coded. 8-The analysis of the proposed algorithm this proposed algorithm is practiced on many images and four of them are selected as samples in fig 4. These images are color and at the same size (24 bits).

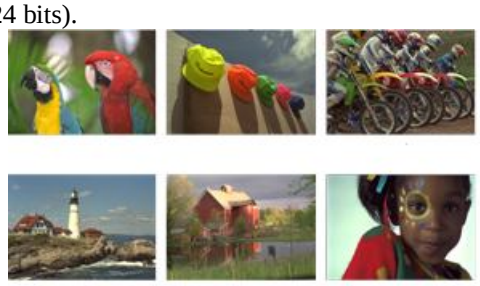

Fig .4.The images for the practice) motocross, b) hats, c) cover girl, d) rustic dream, e) light house

Different patterns can be considered for the embedded data in specific images. 19 patterns have been considered to find the best. For the LSB comparison, 6 different patterns are selected. It should be mentioned that the blue band has the least effect on the brightness of the images in LSB model. Steganography is applied by the proposed method by the 19 patterns pointed out in table 1. In addition, LSB method is considered for 6 patterns in table1. The capacity and brightness of each is the determining factor for the best case.

Regarding to table 1, the best pattern can be case 10 because it enjoys a high PSNR in the highest capacity. For LSB, the sixth case is considered with high capacity and PSNR.

These two patterns, along with PVD [1] and PVD, have been tested. The tests included noise grouse with the mean of zero and variance of 0.01, salt-and-peppernoise with the density of 0.05, the frame change from BMP to JPEG, visual [13], chi-squared [13] and size changes. The amounts of brightness and capacity have been compared in tables 2 and 3 for four methods. The stored message has been extracted and compared with the main message. BER has been used for this purpose. The results have been recorded in table 4. Visual test was practiced on one image (hats) and it shows that in the first specific image, the test has been successful. However, it fails to some extent in the second and third images. The results of these tests are illustrated in figs 5, 6 and 7.

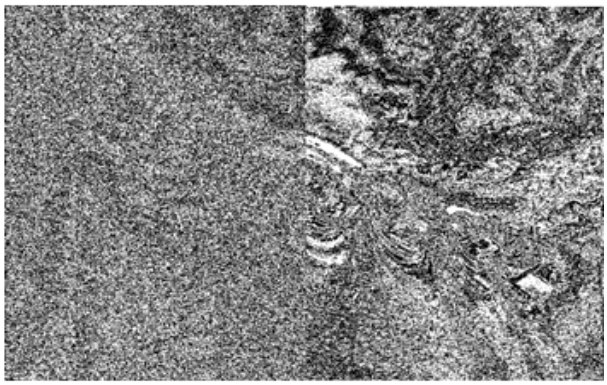

Figure 5.Visual test on the first specific image of hats

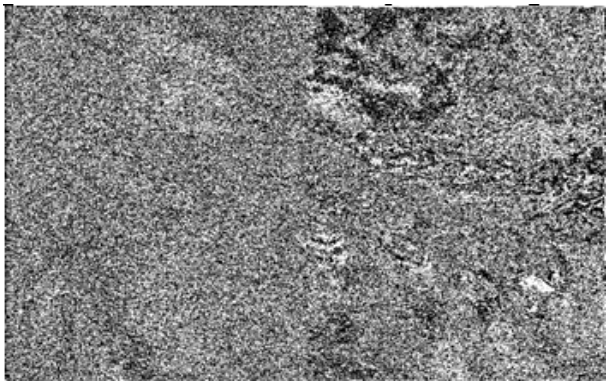

Figure 6. Visual test on the second specific image of hats

Figure 7. Visual test on the third specific image of hats

Figs 5-7 show the results of Chi-squared test. The results show that the proposed algorithm is resistant in this test.

Table 1- PSNR amount for different patterns for the proposed model compared to LSB model in different patterns

light house	rustic dream	cover girl	cockatoo	hats	motocross	Sample / Method			
						Number of storage bits			
						Third	Second	First	
58.94401	58.89474	56.23147	58.91264	55.95462	55.92741	1	0	0	
55.96228	55.97871	59.03517	55.96823	55.91859	55.91702	0	1	0	
54.29029	54.19228	56.36602	54.17276	56.10275	54.17761	0	0	1	
53.59761	53.54060	55.66982	53.84640	54.28030	53.51250	0	1	1	
54.65919	54.56567	55.30338	54.57611	54.69133	53.83650	1	0	1	
54.98521	54.97712	55.16773	54.98050	54.17792	54.17109	1	1	0	
51.79348	50.77127	52.67677	51.67250	51.82034	50.60283	2	0	0	
52.38965	52.43027	50.93106	52.39734	51.16639	52.37120	0	2	0	
.50.66160	50.64062	51.98592	50.64767	51.83404	50.62100	0	0	2	**Proposed Method**
53.36816	53.29724	54.26104	53.44348	53.27205	52.73763	1	1	1	
50.61175	50.56037	51.16934	50.40171	50.89477	50.18102	1	0	2	
51.02608	51.02878	50.40484	51.06142	50.51420	50.65049	1	2	0	
50.23705	50.20128	51.38017	50.37224	50.93351	50.44365	0	1	2	
50.74690	50.74092	50.97819	50.72239	50.59074	50.58255	0	2	1	
51.05573	50.53291	51.54193	50.95944	51.02304	50.15226	2	0	1	
50.52093	50.20289	51.22783	50.51672	50.99801	50.49459	2	1	0	
44.87915	45.24344	46.05082	45.82226	45.53546	45.16190	3	0	0	
45.81838	45.84686	45.58339	45.65742	45.62266	45.73463	0	3	0	
45.37510	45.59802	45.76145	45.57427	45.58980	45.27033	0	0	3	
						Number of storage bits in band			
						Blue	Green	Red	
59.18428	59.17369	59.21697	59.16392	59.08373	59.13255	1	0	0	**Method LSB**
52.43821	52.51514	52.50833	52.42061	52.45559	52.39101	2	0	0	
56.21183	56.18982	56.44046	56.17443	56.14266	56.19173	1	1	0	
46.46487	46.42781	46.42795	46.28687	46.34451	46.34823	3	0	0	
51.61686	51.67477	51.75558	51.58094	51.62772	51.60178	2	1	0	
54.61438	54.53027	54.69022	54.53162	53.24205	54.54522	1	1	1	

Table 2.PSNR amount for the proposed model and the compared methods

light house	rustic dream	cover girl	cockatoo	hats	motocross	Sample / Method
53.36816	53.29245	54.26104	53.44348	53.25658	52.73763	**Proposed Method**
54.61438	54.53027	54.69022	54.53162	54.51105	54.54522	**LSB**
52.07022	49.77366	49.53370	52.02471	50.84478	47.23711	**PVD**
52.87718	51.31031	51.23690	52.56360	51.12444	50.12226	**TPVD**

Table 3.The capacity for the proposed model and the compared methods

light house	rustic dream	cover girl	cockatoo	hats	motocross	Sample / Method
235929	235929	235929	235929	235929	235929	**Proposed Method**
1179648	1179648	1179648	1179648	1179648	1179648	**LSB**
589825	589825	589826	589827	589825	589825	**PVD**
294914	294914	294918	294913	294915	294924	**TPVD**

Table 4.The size of error in the test on the proposed model and compared methods

Return the size	Doubling the size	transform	Pepper-and-salt noise	Gaussian noise	Test: Image name	Test: Method
49.254	50.066	49.732	1.204	43.148	motocross	Proposed Method
49.773	50.066	49.628	1.131	43.079	hats	
48.578	50.066	49.654	1.139	43.203	cockatoo	
47.365	50.066	49.851	1.598	43.168	cover girl	
49.587	50.066	49.365	1.254	43.587	rustic dream	
49.627	50.066	49.587	1.845	43.654	light house	
0.256	48.303	49.964	2.503	49.592	motocross	LSB
0.545	48.303	49.924	2.497	49.863	hats	
0.450	48.303	49.955	2.508	49.932	cockatoo	
0.365	48.303	50.060	2.492	49.812	cover girl	
0.157	48.303	50.047	2.501	49.979	rustic dream	
0.254	48.303	50.000	2.526	49.943	light house	
3.548	43.300	46.254	49.584	51.284	motocross	PVD
3.750	43.300	46.875	49.179	51.553	hats	
3.288	43.300	46.803	49.187	51.381	cockatoo	
3.221	43.300	47.901	49.193	51.391	cover girl	
4.362	43.300	47.651	49.226	51.456	rustic dream	
3.162	43.300	45.416	49.249	51.473	light house	
48.852	43.137	46.548	49.256	50.254	motocross	TPVD
48.074	43.137	46.463	48.982	50.618	hats	
49.089	43.137	46.591	49.248	50.757	cockatoo	
48.632	43.137	47.710	49.153	50.584	cover girl	
48.771	43.137	47.926	48.779	50.652	rustic dream	
48.304	43.137	45.588	49.061	50.829	light house	

VIII. CONCLUSION

We have developed a generalized procedure for concealing diminutive information using image steganography.PCA is a standard tool in modern data analysis . It is one of the best methods for extracting related information from data sets. In our proposed method, LSB is used for embedded data so that convolutional code is used. This code is easy to use and practice. So, the proposed plan satisfies the requirement of imperceptibility and robustness for a feasible steganography plan. Experimental results show that the proposed method is robust against common image processing attacks.

REFERENCES

[1] Bas, Anil, et al. "Fitting a 3D morphable model to edges: A comparison Between hard and soft correspondences." arXiv preprint arXiv: 1602.01125 (2016).

[2] Da-Chun Wu and Wen-Hsiang Tsai, "A steganographic method for images by pixel-value differencing," Pattern Recogn. Lett., vol. 24, no. 9-10, pp. 1613-1626, 2003.

[3] Niels Provos, "Defending against statistical steganalysis," in SSYM'01: Proceedings of the 10th conference on USENIX Security Symposium. Washington: USENIX Association, 2001, pp. 323-335.

[4] Blanz, Volker, and Thomas Vetter. "A morphable model for the synthesis of 3D faces." Proceedings of the 26th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., 1999

[5] Karl Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine, vol. 2, no. 6, pp. 559-572, 1991.

[6] Chin-Chen Chang, Ju Yuan Hsiao, and Chi-Shiang Chan, "Finding optimal least-significant-bit substitution in image hiding by dynamic programming strategy," Pattern Recognition, vol. 36, no. 7, pp. 1583-1595, 2003.

[7] Mohsen Moradi ," Combining and Steganography of 3D Face Textures", arXiv:1702.01325v1 [cs.MM] 4 Feb 2017

[8] Andrew Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," EEE Transactions on Information Theory , vol. 13, no. 2, pp. 260 - 269, Apr 1967.

[9] Shu Lin and Costello Daniel J., Error Control Coding, Second Edition. USA: Prentice-Hall, Inc., 2004.

[10] M. Michelson and A. H. Levesque, Error Control Techniques for Digital Communication. New York, USA: John Wiley & Sons, 1985.

[11] Patel, Ankur, and William AP Smith. "3d morphable face models revisited." Computer Vision and Pattern Recognition, 2009. CVPR 2009.IEEE Conference on. IEEE, 2009.

[12] Rich Franzen. (2010, Feb) Kodak true color images.[Online]. http://r0k.us/graphics/kodak/index.html

[13] Pearson, Karl. "Principal components analysis." The London, Edinburgh and Dublin Philosophical Magazine and Journal 6.2 (1901): 566.

[14] Jonathon Shlens,"A Tutorial on Principal Component Analysis". arXiv:1404.1100v1 [cs.LG] 3 Apr 2014

[15] AMRUTHA ASHOK MANVI1, SHARATH YAJI2," Implementation of Secure Steganographic Scheme Using 3D Models". International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 14 Issue 2 –APRIL 2015.

Mahdi Koohi, Educated from electronic department which research in image processing and computer vision and some of papers is in edge detection Moving object tracking, Artificial Neural Networks for Image Analysis of Agricultural Products , network processor

