A Hybrid Method for Color Image Steganography in Spatial and Frequency Domain

Asghar Shahrzad Khashandarag [1], Akbar Shahrzad Khashandarag [2], Amin Rahimi Oskuei [3], Hamid Haji Agha Mohammadi [4] and Mirkamal Mirnia [5]

**[1] Young Researchers Club of Tabriz, Islamic Azad University Tabriz Branch
Tabriz, Iran**

**[2] Department of Computer Engineering, Islamic Azad University Tabriz Branch
Tabriz, Iran**

**[3] Department of Computer Engineering, Islamic Azad University Tabriz Branch
Tabriz, Iran**

Abstract

A hybrid method for color image steganography is suggested to conceal a secret data into the cover image in the spatial and frequency domain. In this method, Lempel–Ziv–Welch (LZW) compression is used to obtain a low bit rate; Also Linear Feedback Shift Register (LFSR) technique is used to enhance the security of the scheme. In the embedding process in spatial domain, a K-means clustering and EA algorithms are used for secret data embedding. Also in the embedding process in frequency domain, a Coefficients selection and frequency hopping algorithm (CSFH) and Adaptive Phase Modulation mechanism (APM) are used for secret data embedding. Abilities of the proposed method are high security, because of using a hybrid method (spatial and frequency domain) and Data Encryption Standard (DES).

Keywords: *steganography; Linear Feedback Shift Register (LFSR); Data encryption standard (DES); K-means algorithm; Discrete Fourier Transform; spatial domain; frequency domain.*

1. Introduction

Now a day, information security is very important. Information that you sent, how and under what circumstances to reach the receiver, is not clear to us. Therefore, information security will be important to you. Due to growth and progress of science in the world, computers and the Internet to send data are used. Therefore the information, as more digital data will be. Today, Image, sound, and etc are known as digital data and we can in addition to information and the digital data to hide the main recipient, we send. The first application of steganography can be predicted 440 BC by Herodotus. Data hiding methods are divided into two methods:

- Steganography
- Watermarking

The word steganography is derived from Greek and means concealed writing. Steganography techniques [1-5] are used to hide digital data in coverage media so that no one can find it. Watermarking techniques [6-10] is a process to embedding data into a coverage media that is difficult to remove. In this paper, a hybrid method for color image steganography in spatial and frequency domain is presented.

2. Related Works

Wen-Yuan Chen [11] proposed a method for color image steganography using SPHIT and DFT Sending with JPEG format (hiding secret image in cover image). His method was designed in the frequency domain. He believes that phase is suitable for secret data embedding, because the phase enjoys high noise immunity. For more information see [11]. Asghar Shahrzad Khashandarag [12] proposed a method for color image steganography using SPIHT and DFT Sending with JPEG format based on Wen-Yuan Chen works (hiding secret information in cover image). His method uses frequency domain. There are some disadvantages such as low capacity of color image for information embedding in both methods. Here we propose a hybrid method for color image steganography in the spatial and frequency domain to overcome to this problem.

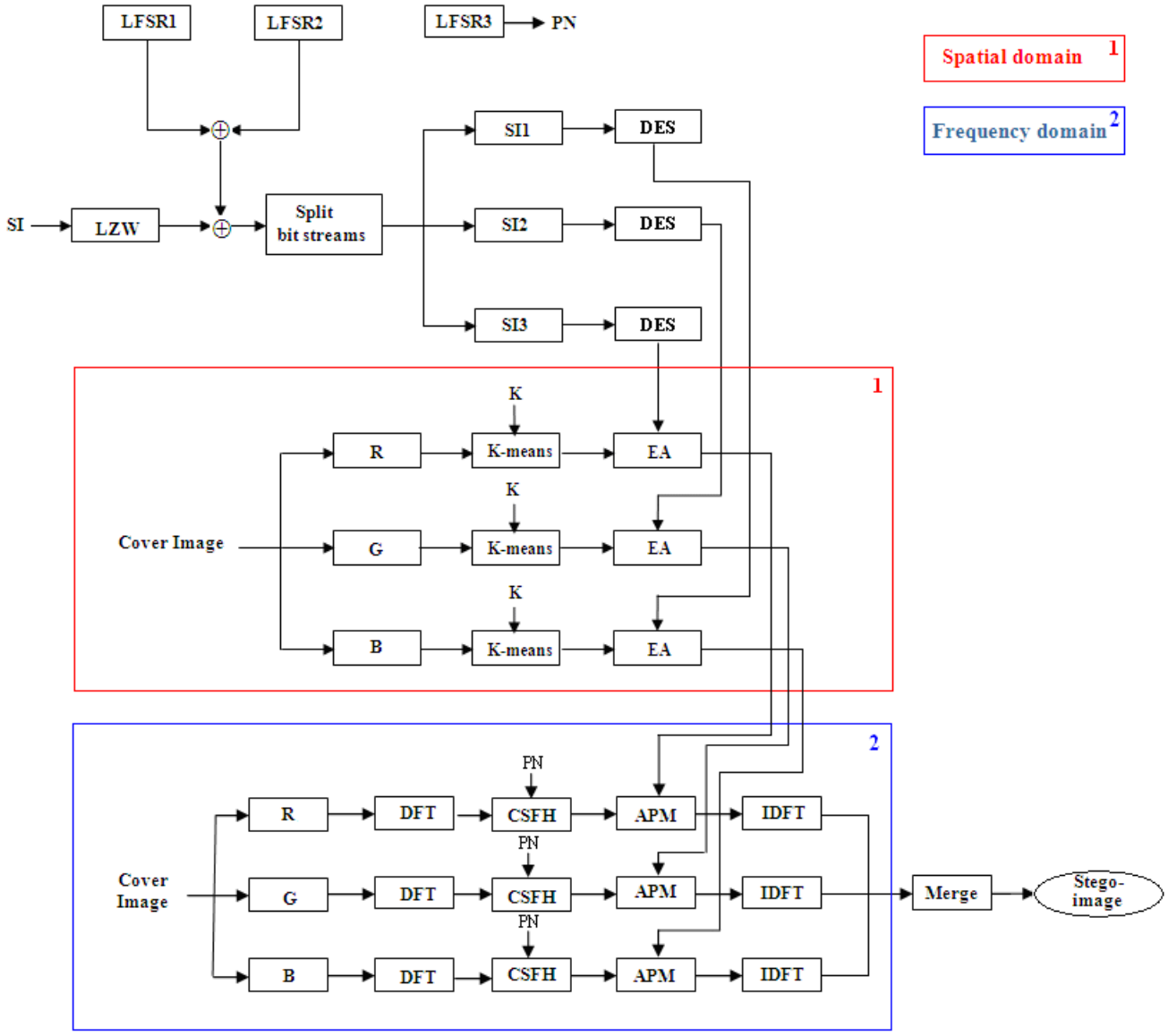

Fig. 1 The conceptual model of the secret data-embedding algorithm

3. Proposed method

3.1 Data-embedding algorithm

A steganography process must be completely secure, and does not reduce the visual quality of the cover image when the secret data is concealed. The overall concealing process of our proposed scheme is shown in Fig1. The Secret information (SI) is compression by LZW technique. The Result of compression stages, with pseudo-random number is *xor* , and then bit streams are split into three parts (SI1, SI2 and SI3) and then DES is used for security.

For more information about DES algorithm see [13]. In spatial domain the cover image is parted into R, G and B and after R, G and B clustered to k cluster by k-means algorithm where k is known. Then the result of DES embeds in clusters randomly by EA algorithm. In frequency domain the cover image also is parted into R, G and B and after by transmitting into frequency domain by DFT, hashing and modulation algorithms are applied. Then the result of the EA stage is embedded into the CSFH results. Then they are transmitted to spatial domain by

Inverse Discrete Fourier Transform (IDFT). Finally they are merged and stego-image is created.

3.1.1 LZW data compression

Lempel–Ziv–Welch (LZW) is a universal lossless data compression algorithm created by Abraham Lempel, Jacob Ziv, and Terry Welch It was published by Welch in 1984 as an improved implementation of the LZ78 algorithm published by Lempel and Ziv in 1978. The algorithm is designed to be fast to implement. LZW compression uses a code table. A common choice is to provide 4096 entries in the table. In this case, the LZW encoded data consists entirely of 12 bit codes, each referring to one of the entries in the code table. Uncompressing is achieved by taking each code from the compressed file, and translating it through the code table to find what character or characters it represents. Codes 0-255 in the code table are always assigned to represent single bytes from the input file. For example, if only these first 256 codes were used, each byte in the original file would be converted into 12 bits in the LZW encoded file, resulting in a 50% larger file size. During uncompressing, each 12 bit code would be translated via the code table back into the single bytes. Of course, this wouldn't be a useful situation [14].

3.1.2 Linear Feedback Shift Register

A linear feedback shift register (LFSR) is a shift register whose input bit is a linear function of its previous state [15]. The initial value of the LFSR is called the seed, and because the operation of the register is deterministic, the stream of values produced by the register is completely determined by its current (or previous) state. Likewise, because the register has a finite number of possible states, it must eventually enter a repeating cycle. However, an LFSR with a well-chosen feedback function can produce a sequence of bits which appears random and which has a very long cycle. In Fig2 the structure of LFSR is shown, where D_i is register for store bit. LFSR is an n-bit shift register which pseudo-randomly scrolls between 2^n -1 values, but does it very quickly because there is minimal combinational logic involved. Once it reaches its final state, it will traverse the sequence exactly as before. It has many applications you should already be familiar with if you're reading this [15]. In the equation (1), the Polynomial function for the LFSR is shown.

$$G(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_i x^i + \ldots + a_1 x^1 + a_0$$

$$a_i = 0 \text{ or } 1 \qquad (1)$$

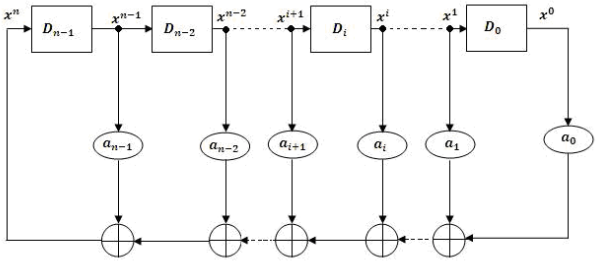

Fig. 2 The structure of LFSR technique

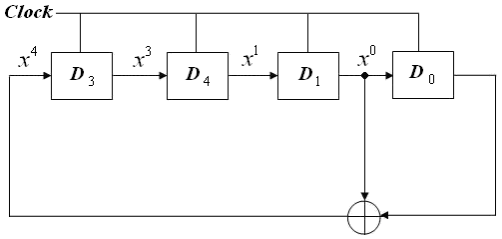

Fig. 3 The structure of LFSR for $G(x) = x^4 + x^1 + 1$.

In Fig3 the structure of LFSR for $G(x) = x^4 + x + 1$ is shown.

Table 1: The generation of random-numbers for $G(x) = x^4 + x^1 + 1$

D_3	D_2	D_1	D_0
0	0	0	1
1	0	0	0
0	1	0	0
0	0	1	0
1	0	0	1
1	1	0	0
0	1	1	0
1	0	1	1
0	1	0	1
1	0	1	0
1	1	0	1
1	1	1	0
1	1	1	1
0	1	1	1
0	0	1	1
0	0	0	1

Pattern generation again

In Table1 the generation of random-numbers for Fig3 is shown. In Table1 the period is calculated as follows:

$$Period = 2^n - 1 = 2^4 - 1 = 15$$

3.1.3 K-means clustering algorithm

Simple clustering methods use greedy interactions with existing clusters to come up with a good overall representation. K-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. Finally, this algorithm aims at minimizing an *objective function*, in this case a squared error function. The objective function in equation (2) is shown.

$$J = \sum_{j=0}^{k} \sum_{i=0}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \quad (2)$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j, is an indicator of the distance of the *n* data points from their respective cluster centres.

Algorithm1: clustering by K-Means

Choose k data points to act as cluster centers.

Until the cluster centers are unchanged

1. Allocate each data point to cluster whose center is nearest.
2. Now ensure that every cluster has at least one data point; possible techniques for doing this include supplying empty clusters with a point chosen at random from points far from their cluster center.
3. Replace the cluster centers with the mean of the elements in their clusters.

End

Fig4 shows the selected data points in Screen page. Fig5 shows the simulation of K-means clustering algorithm with k=4.

3.1.4 EA algorithm

Fig6 shows the embedding algorithm (EA) which L is number of Least Significant Bit (LSB) for embedding. The L=1 means the first LSB of R, G and B. The L=2 means the second LSB of R, G and B. The L=3 means the third LSB of R, G and B. The L=4 means the fourth LSB of R, G and B. TP is total pixel of cover image. TM is number of bits for embedding. In the condition1 (3×TP>=TM) in Fig6, number 3 means the first LSB of R, G and B. In the condition2 (6×TP>=TM) in Fig6, number 6 means the first and second LSB of R, G and B. In the condition3 (9×TP>=TM) in Fig7, number 9 means the first, second and third LSB of R, G and B. In the condition4 (12×TP>=TM) in Fig6, number 12 means the first, second, third and fourth LSB of R, G and B.

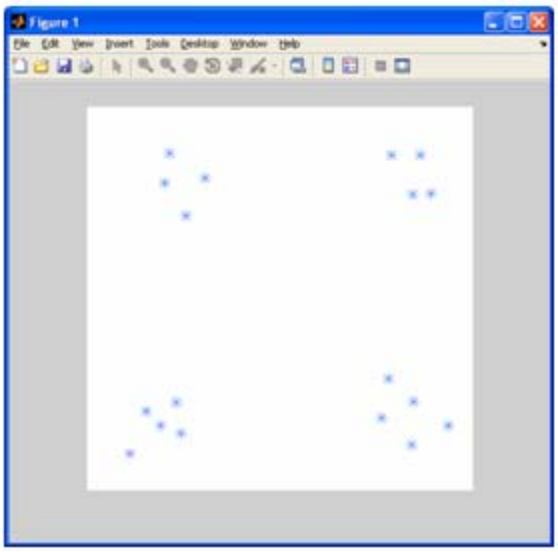

Fig. 4 The selected data points in screen page.

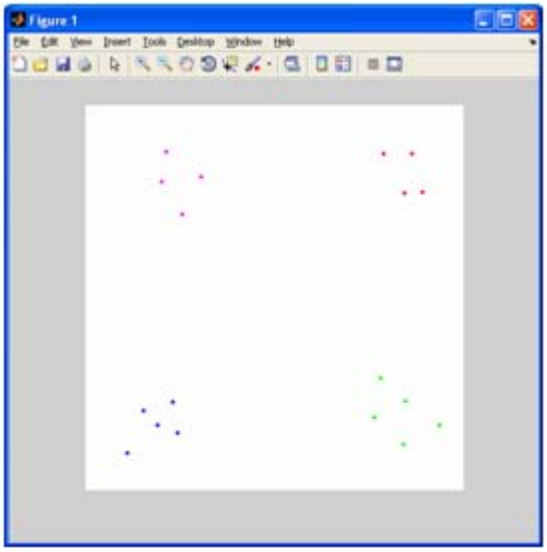

Fig. 5 The result of K-means clustering with k=4.

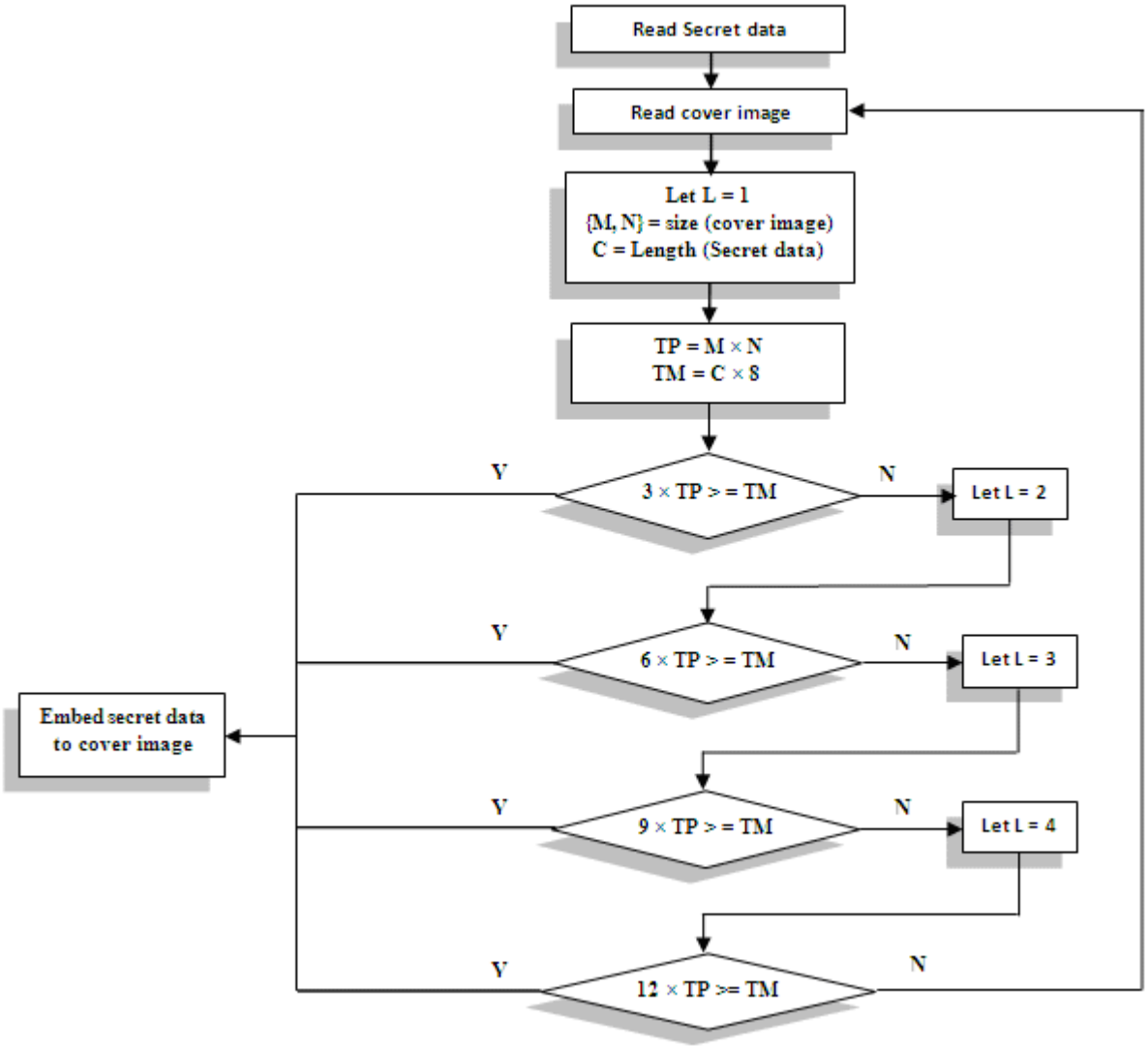

Fig. 6 The structure of EA algorithm is shown.

3.1.5 Discrete Fourier Transform (DFT)

We need DFT to transfer an image from spatial domain into frequency domain. In the equation (3) the DFT is shown. Note that $f(x,y)$ is image in spatial domain (host-image) and $F(u,v)$ is image in frequency domain (stego-image).

$$F(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

$$u=0,1,2,\ldots,M-1 \text{ and } v=0,1,2,\ldots,N-1 \qquad (3)$$

3.1.6 Inverse Discrete Fourier Transform (IDFT)

The equation (4) the DFT is shown. Note that M and N are the size of image.

$$f(x,y)=\frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

$$x=0,1,2,\ldots,M-1 \text{ and } y=0,1,2,\ldots,N-1 \qquad (4)$$

3.2 Inverse Data-embedding algorithm

The process of inverse data embedding is shown in Fig7. Note that APD and ILZW are the inverse of APM and LZW technique which explained in [13].

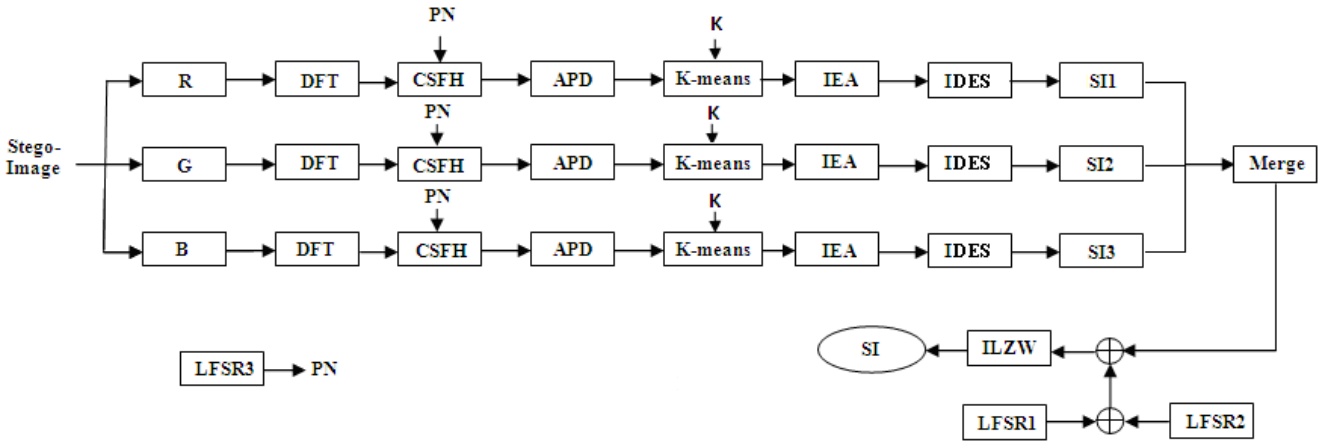

Fig. 7 The conceptual model of the inverse data-embedding algorithm

4. Experience and results

The proposed method in this paper has high capacity in comparing with previous method [12].

4.1 Previous method

For an $m \times n$ image, number of characters for embedding is calculated as follows:

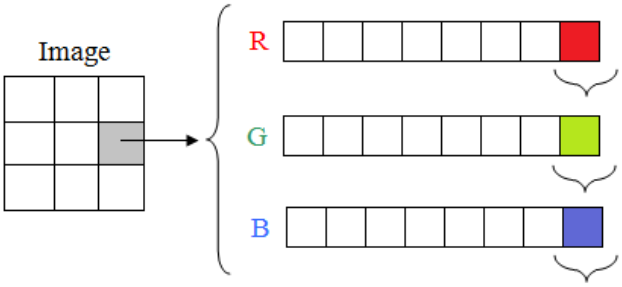

Fig. 8 Three bits for data embedding in previous method is shown.

1 Pixel	3 Bit	$\Rightarrow x_1 = 3mn$ Bit
$m \times n$ Pixel	x_1	

3 Bit	$\frac{3}{8}$ Char	$\Rightarrow y_1 = \frac{3}{8}mn$ Char
$3\ m \times n$ Bit	y_1	

x_1 : Number of characters in $m \times n$ image

y_1 : Number of bits in $m \times n$ image

4.2 Hybrid method

For an $m \times n$ image, number of characters for embedding is calculated as follows:

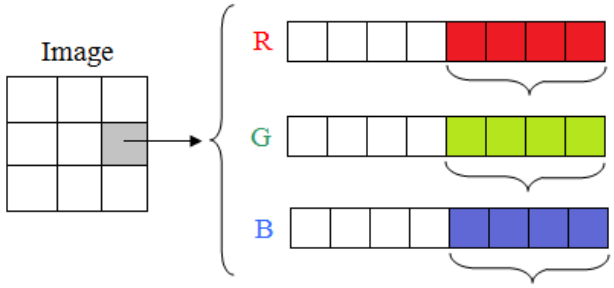

Fig. 9 twelve bits for data embedding in hybrid method is shown.

1 Pixel	12 Bit	$\Rightarrow x_2 = 12mn$ Bit
$m \times n$ Pixel	x_2	

12 Bit	$\frac{12}{8}$ Char	$\Rightarrow y_2 = \frac{12}{8}mn$ Char
$12\ m \times n$ Bit	y_2	

x_2 : Number of bits in $m \times n$ image

y_2 : Number of characters in $m \times n$ image

Here we assume $m = n$ for the simplicity of showing the efficiency; Table2 shows the deference between the

number of characters re-presentable by previous and present method in six images.

Table 2: The comparison with six images

#	*Images* ($n \times n$)	*Capacity (based on character)*		*Difference (based on character)*
		Previous method	*Hybrid method*	
1	50×50	937	3750	2813
2	100×100	3750	15000	11250
3	150×150	8437	33750	25313
4	600×600	135000	540000	405000
5	800×800	240000	960000	720000
6	1024×1024	393216	1572764	1179648

The deference is shown in Fig10 graphically.

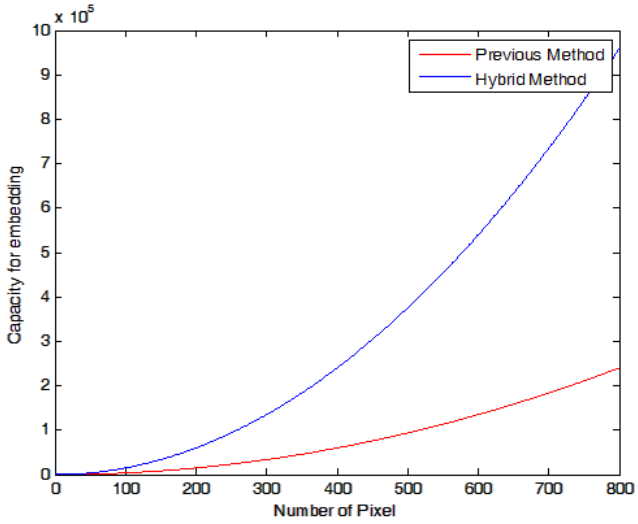

Fig. 10 The comparison of two methods is shown.

5. Conclusions

The proposed method in this paper has high capacity and security in comparing with previous method. The high capacity because of used the EA algorithm in spatial domain. Also the high security because of used the DES algorithm and frequency domain.

Acknowledgments

I would like to express my thanks to the referee for the suggestion to promote the quality of the paper to publish the paper.

References

[1] A. Cheddad, J. Condell, K. Curran, P.M. Kevitt, Digital image steganography: Survey and analysis of current methods, Signal Processing, Volume 90, Issue 3, March 2010, Pages 727-752.
[2] G.C. Kessler, C. Hosmer, An Overview of Steganography, Advances in Computers, Volume 83, 2011, Chapter Chapter 2, Pages 51-107
[3] Z. Chen, L. Huang, W. Yang. Detection of substitution-based linguistic steganography by relative frequency analysis, Digital Investigation, In Press, Corrected Proof, Available online 21 April 2011
[4] Z.G Qu, X.B Chen, M.X. Luo, X.X. Niu, Y.X. Yang, Quantum steganography with large payload based on entanglement swapping of χ-type entangled states, Optics Communications, Volume 284, Issue 7, 1 April 2011, Pages 2075-2082
[5] C.C. Chang, J.S. Lee, T.H.N. Le, Hybrid wet paper coding mechanism for steganography employing n-indicator and fuzzy edge detector, Digital Signal Processing, Volume 20, Issue 4, July 2010, Pages 1286-1307
[6] M.A. Akhaee, N.K. Kalantari, F. Marvasti, Robust audio and speech watermarking using Gaussian and Laplacian modeling, Signal Processing, Volume 90, Issue 8, August 2010, Pages 2487-2497
[7] Y. Liu, J. Zhao, A new video watermarking algorithm based on 1D DFT and Radon transform, Signal Processing, Volume 90, Issue 2, February 2010, Pages 626-639
[8] C.C. Lai, A digital watermarking scheme based on singular value decomposition and tiny genetic algorithm, Digital Signal Processing, Volume 21, Issue 4, July 2011, Pages 522-527
[9] Y.R. Wang, W.H. Lin, L. Yang, An intelligent watermarking method based on particle swarm optimization , Expert Systems with Applications, Volume 38, Issue 7, July 2011, Pages 8024-8029
[10] C.C Chang, K.N Chen, C.F Lee, L.J. Liu, A secure fragile watermarking scheme based on chaos-and-hamming code, Journal of Systems and Software, Available online 21 March 2011.
[11] W.Y Chen, Color image Steganography scheme using set partitioning in hierarchical trees coding, digital Fourier transform and adaptive phase modulation, Elsevier 2006. doi:10.1016/j.amc.2006.07.041
[12] A.S. Khashandarg, N. Ebrahimian. A new method for color image Steganography using SPIHT and DFT, sending with JPEG format, 978-0-7695-3892-1/09 $26.00 © 2009 IEEE doi:10.1109/ICCTD.2009.14
[13] http://en.wikipedia.org/wiki/Data_Encryption_Standard
[14] http://en.wikipedia.org/wiki/Lempel_Ziv_Welch
[15] http://en.wikipedia.org/wiki/Linear_Feedback_Shift_Registe

Asghar Shahrzad Khashandarag received the B.Sc. degree in computer engineering from the Payame Noor University, Bonab, Iran, and the M.Sc. degree in computer engineering from the Islamic Azad University Tabriz Branch, Tabriz, Iran, in 2008 and 2010, respectively. From 2008, he works as a researcher with the Young Researchers Club of Tabriz. His research interests include image processing, signal processing and wireless sensor network. He is a member of the IEEE and the IEEE Computer Society and Young Researchers Club of Tabriz.

Akbar Shahrzad Khashandarag received the B.Sc. degree in computer engineering from the Islamic Azad University Bonab Branch, Iran in 2009. He is currently M.Sc. student in Mechatronics engineering at the Islamic Azad University Tabriz Branch. His research interests include image processing, signal processing and wireless sensor network.

Amin Rahimi Oskuei received the B.Sc. degree in computer engineering from the Islamic Azad University Shabestar Branch, Iran in 2007. He is currently M.Sc. student in computer engineering at the Islamic Azad University Tabriz Branch.

Hamid Haji Agha Mohammadi received the B.Sc. degree in computer engineering from the Payame Noor University Bonab Branch, Iran in 2009. He is currently M.Sc. student in computer engineering at the Islamic Azad University Gazvin Branch.

Dr. Mirkamal Mirnia received the B.Sc. degree in mathematical Physics from Ferdowsi University, Mashhad, Iran in 1967, and M.Sc. degree in pure mathematics from teacher training University of Tehran in 1969 also M.Sc. degree in numerical analysis and computing from Owen university of Manchester, UK in 1975 and PhD. in applied mathematics (optimization) from university of St.Andrews, UK in 1979. He is a member of the Mathematical society of Iran, Iranian operations researches, institute of applied mathematics, UK and member of SIAM.