Multimed Tools Appl
DOI 10.1007/s11042-017-4775-x

Reversible data hiding scheme using sub-sampled image exploiting Lagrange's interpolating polynomial

Biswapati Jana[1]

Received: 27 September 2016 / Revised: 16 February 2017 / Accepted: 28 April 2017

Abstract In this paper, a new reversible data hiding scheme has been proposed using lagrange's interpolating polynomial on interpolated sub-sampled images. First, we generate sub-sampled images from original image and enlarge its size using image interpolation. Now, we convert secret message using lagrange interpolating polynomial and generate new secret message. The new secret message is divided and stored within interleaved pixel of each interpolated sub-sampled images. At the receiver end, new secret message is extracted from interleaved pixel of each sub-sampled stego images and then lagrange's interpolation is applied to generate original secret message. The security has been enhanced due to the distributive nature of hidden data within multiple images. The original pixels are not effected during data embedding which assure reversibility. The proposed scheme provides average embedding capacity with good visual quality measured by peak signal to noise ratio (PSNR) which is greater than 50 dB. It is observed that the proposed scheme provides better performance than other existing data hiding schemes in terms of data embedding capacity, visual quality and security. We have analyzed our stego images through RS analysis, calculate relative entropy, standard deviation and correlation coefficient of original and stego image to show the robustness under various steganographic attacks.

Keywords Steganography · Sub-sampled image · Reversible data hiding · Lagrange interpolation · Image interpolation · RS analysis · Relative entropy

✉ Biswapati Jana
biswapatijana@gmail.com

[1] Department of Computer Science, Vidyasagar University, Midnapore,
West Bengal, Pin-721102, India

Published online: 09 May 2017

1 Introduction

Steganography is one of the important branch of hidden data communication in the research areas of information security. The main aim of steganography is the concealment of the secret information within cover work by tweaking its properties. The intention is to concentrate on precluding the adversary from moving out the content of the confidential message by applying a variety of distortions techniques. However, data hiding is a form of covered communication that usually puts stress on simply finding the presence of a secret message. Thus, the imperceptibility becomes the most significant place for the information hiding schemes. The most common data hiding methods are based on the least significant bits (LSBs) replacement, masking, filtering, transformation [1, 2, 18] and the modulus operation [3, 15, 19]. All these methods can encode and decode the secret message successfully, but they are not reversible and a good method of practice is to keep the message data as short as possible when using data hiding schemes.

Now-a-days, there are two main kinds of data hiding approaches are used: reversible and irreversible. In reversible data hiding scheme the original image can be reconstructed after extracting the secret information from stego-image without any distortion [5, 7, 8, 14]. On the other hand, irreversible data hiding scheme can not recover original image but can hide large amount of secret data. Reversible Data Hiding (RDH) presented by Ni et al. [14] which is based on histogram shifting with zero or minimum change of the pixel gray values and the capacity is 0.3 bpp with quality is 48 dB PSNR. Multilevel reversible data hiding scheme based on histogram shifting is proposed by Lin et al. [12] and Tsai et al. [16]. In 2009, Kim et al. [9] proposed reversible data hiding scheme using histogram shifting on sub-sampled image and employ correlation among them to hide secret message. Payload and quality of Kim et al's. [9] was 20121 bits and PSNR 48.9 dB respectively for Lena image. Luo et al. [13] improve Kim's scheme by selecting the median pixel as reference of sub-image in each block. The image block are divided into four classes to preserve the median during data embedding. The embedding capacity was 0.11 bpp with 48.9 dB PSNR. Lee et al. [11] proposed two stages multilevel reversible data hiding scheme using Lagrange Interpolation. In their scheme, they first generate predicted image using Lagrange Interpolation then calculate image difference between original image and predicted image. After that they use histogram shifting algorithm to hide the secret data. In that approach, they produce only one stego image with different payload and visual quality (For example, payload 0.88 bpp with PSNR 48.32 dB for Baboon image [11]).

In data hiding schemes, achievement of reversibility and enhancement of security while maintaining good visual quality through sub-sampled images is still an important research issue. Designing a novel secure data hiding system accomplishing good visual quality, high embedding capacity, robustness and steganorgaphic protection is a technically challenging problem. After the confidential message is extracted, the requirement for the image reversibility for the entire recovery of the original image without any distortion goes high. In this context, we propose a secure reversible data hiding scheme through Lagrange's Interpolating Polynomial using sub-sampled image, where one can hide secret bits among sampled sub-images and recover successfully.

The Lagrange interpolation is very good in lower order polynomials. Higher order lead to polynomial wiggle, so only four ($t = 4$) sub-sampled images are used in this approach using Lagrange polynomial function $f(x)$ of order $(t - 1)$. The proposed scheme is suitable for some applications where high image quality is required but high data embedding capacity is not required. Also use of multiple sub-image may be strange for a normal person's point of view but for secret sharing scheme among a group of people where every members have some shared data then this scheme provides a good solution for different application areas like medical image processing, military application etc. The proposed scheme achieves good visual quality and enhance security.

Data hiding through sub-sampled image using Lagrange Interpolating Polynomial enhance security because the secret data bits are distributed among sub-sampled images and without simultaneous sub-sampled images, it is hard for eavesdroppers to retrieve secret data from stego images. This is a special case of secret sharing. Another advantage of Lagrange Polynomial is that it is possible to retrieve secret data using less number of generated sub-sampled stego image. Here, we have used four (4) sub-sampled image but during data extraction we have taken any three (3) sub- sampled stego images.

1.1 Motivation

In this paper, a new secure reversible data hiding scheme has been proposed through Lagrange Interpolating Polynomial using interpolated sub-sampled image.

- Our main motivation is to enhance security and to achieve reversibility in data hiding schemes. Data embedding through Lagrange Interpolating Polynomial was not reversible. We use image interpolation technique to achieve reversible data hiding through sub-sampled images. The pixel values of original image was unaltered within sub-sampled stego images. Hence, it is possible to recover original image successfully after extracting secret data.
- Conversion from original secret data to new secret message using lagrange interpolating polynomial and then distribution of these new secret message among sub-sampled images can enhance security. The nature of lagrange interpolating polynomial function with different unknown coefficient value and unknown order of polynomial function can also enhance security and hard to guess by the adversary.
- Due to the loss of any sub-sampled stego images, it is possible to recover original secret data successfully using less number of total sub-sampled images. It may possible to mention the required number of sub-sampled stego images to retrieve secret message, which depends on the order of lagrange polynomial function.
- So far, some data hiding techniques have been developed using single or dual images. Therefore, our motivation is to introduce data hiding approach within sub-sampled interpolated image treated as multiple images to enhance security, increase quality and achieve reversibility.

The rest of the paper is organized as follows. Proposed data hiding scheme is discussed in Section 2. Experimental results with comparisons are discussed in Section 3. The attacks are presented in Section 4. Finally, conclusion is given in Section 5.

2 Proposed scheme

The secret information which may contain ownership identification, authentication and copy right protection are embedded within multiple sub-sampled images. Sub-sampling is the process of selecting some part of cover image. Suppose that an cover image of size $(M \times N)$ pixels is denoted by $I(m, n)$, where $m = 0 \ldots (M - 1)$ and $n = 0 \ldots (N - 1)$. Two sampling factors, $\triangle u$ and $\triangle v$ determine the desired sub-sampling intervals in a row and column direction respectively. As illustrated in Fig. 1, the 2-D cover image is sub-sampled at uniform intervals. This process is called as sub-sampling. In this paper, we use only four

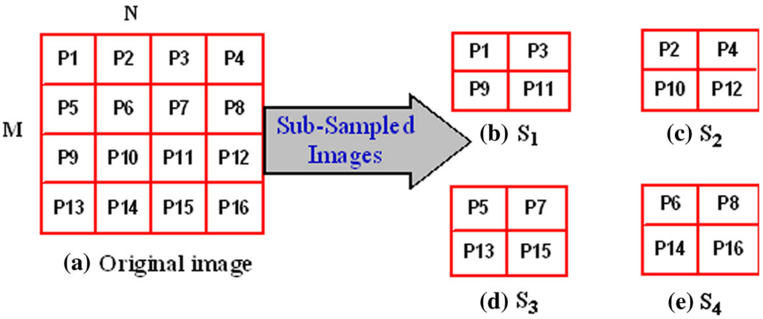


Fig. 1 Sub-sampling of a 2-D image

sub-sample images. We create sub-sampled images (S_1, S_2, S_3, S_4) of size $(m/2 \times n/2)$ from the original image using following equation.

$$\begin{aligned} S_1(i,j) &= I(2i-1, 2j-1) \\ S_2(i,j) &= I(2i-1, 2j) \\ S_3(i,j) &= I(2i, 2j-1) \\ S_4(i,j) &= I(2i, 2j) \end{aligned} \tag{1}$$

where $i = 1, 2 \ldots m/2$ and $j = 1, 2 \ldots n/2$. We enlarge the sub-sampled images using image interpolation shown below.

$$ZS_k(i,j) = \begin{cases} S_k(p,q); \\ \quad \text{where } p = 1, \ldots, m/2; \;\; q = 1, \ldots, n/2; \\ \quad \text{and } i = 1, 3, \ldots (m-1); \;\; j = 1, 3, \ldots (n-1); \\ \frac{ZS_k(i,j-1)+ZS_k(i,j+1)}{2}; \\ \quad \text{where } (i \; mod \; 2) \neq 0; \;\; (j \; mod \; 2) = 0; \\ \frac{ZS_k(i-1,j)+ZS_k(i+1,j)}{2}; \\ \quad \text{where } (i \; mod \; 2) = 0; \;\; (j \; mod \; 2) \neq 0; \\ \frac{ZS_k(i-1,j-1)+ZS_k(i-1,j+1)+ZS_k(i+1,j-1)+ZS_k(i+1,j+1)}{4}; \\ \quad \text{where } (i \; mod \; 2) = 0; \;\; (j \; mod \; 2) = 0; \end{cases} \tag{2}$$

Due to any loss of sampled image during transmission, the secret data can be lost. To recover the secret message in such situation we can set the number of sub-sampled stego image which are required to extract the secret message using the threshold value (t). Here it is chosen as three (3) out of four (4) sub-sampled image, that means any three sub-sampled stego images are sufficient to extract the secret message out of four sub-sampled stego images. It is also implies that none can extract hidden information using less than t sub-sampled images. Now, consider eight bits secret message (ASCII or pixel) (M) then apply lagrange interpolation function $f(x)$ of order $(t-1)$ is as follows:

$$f(x) = a_0 + a_{(t-2)}x^{(t-2)} + a_{(t-1)}x^{(t-1)} \tag{3}$$

where $x = 1, 2, 3, 4$. Since, to extract secret message we use t sub-sampled stego images so the degree of lagrange polynomial will be $(t-1)$. The coefficient x is chosen randomly within the specified range. Here, in the equation (3) a_0 is the ASCII value of any secret message. Then the calculated value of $f(x)$ is converted into 12 bits binary form and divided it into 4 parts each of 3 bits. Then the interpolated image (ZS_k) is divided into (4×4) non overlapping blocks (b_k) and embed 3 bits by adding with the interleaved pixel value at position $b_k(1,2), b_k(2,1), b_k(2,3), b_k(3,2)$. And the x value is added with the value at position $b_k(2,2)$. where, $(k = 1, 2, \ldots (m - 1 \times n - 1))/(4 \times 4)$.

2.1 Data embedding

The original message is converted as new secret data using lagrange interpolation polynomial and converted into binary form. The message is embedded by modifying the interleaved pixel of each sub-sampled image. Figure 2 describes an overall data embedding procedure and numerical illustration is shown in Fig. 3. The proposed data embedding algorithm is listed in Algorithm 1. At the receiver end, secret data can be extracted from sub-sampled stego images as shown in Fig. 4. The corresponding numerical illustration is shown in the Fig. 5. The algorithm of data extraction is listed in Algorithm 2.

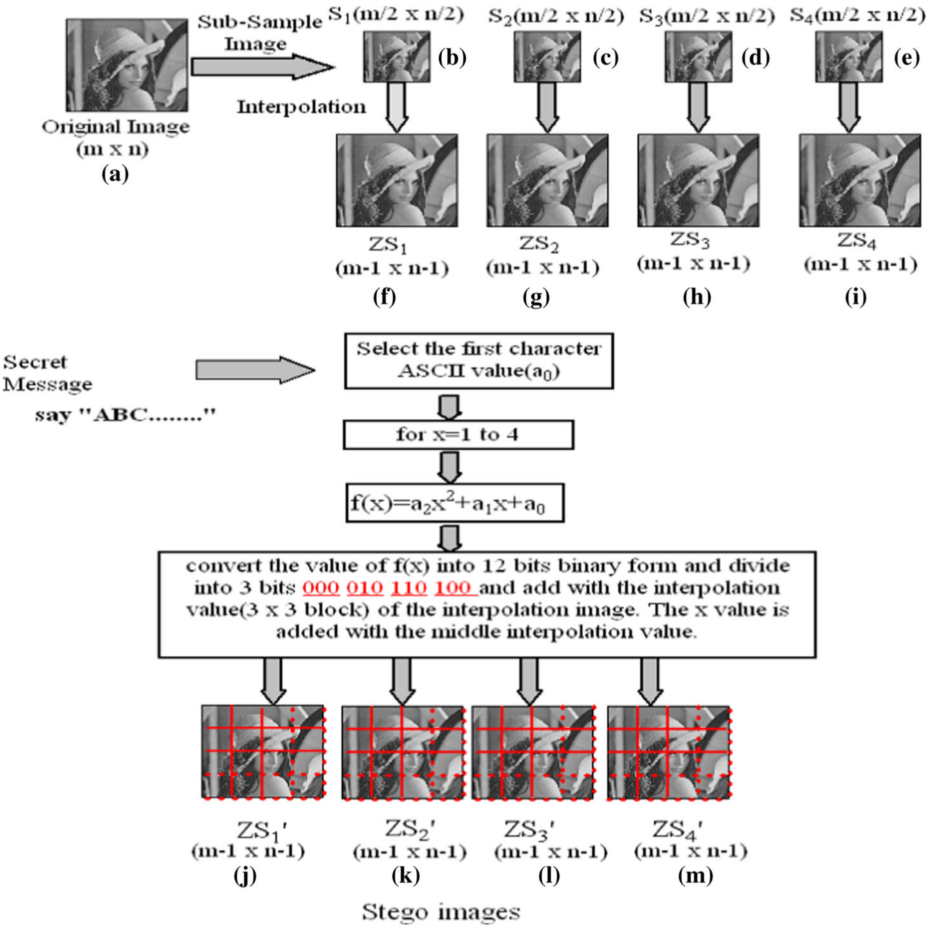


Fig. 2 Block diagram of data embedding procedure

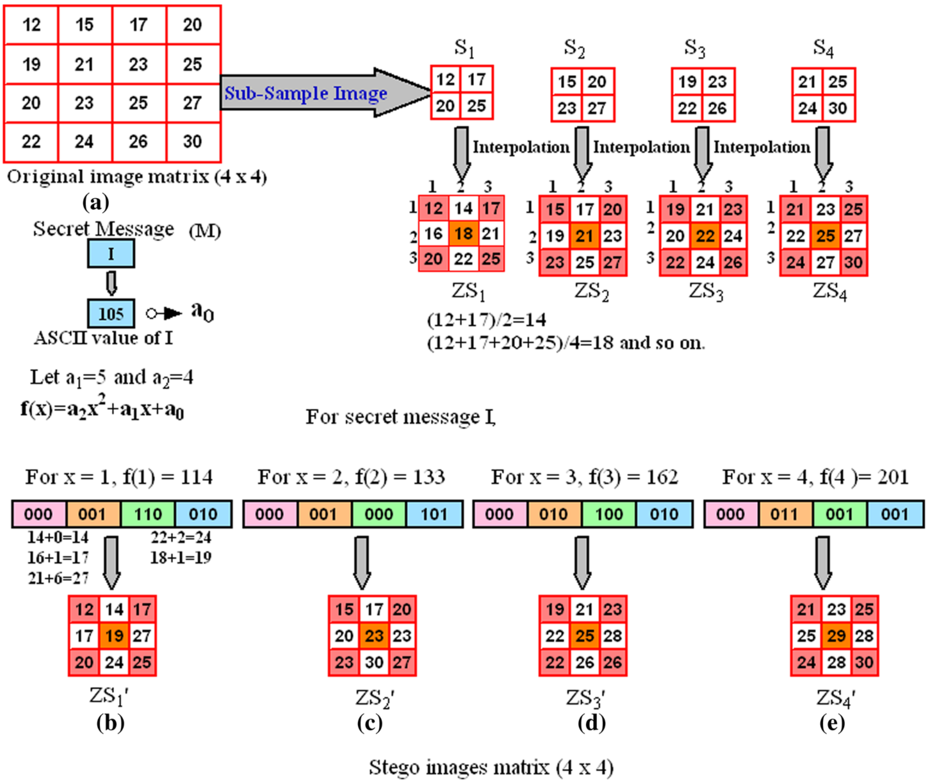


Fig. 3 Numerical illustration of data embedding

Algorithm 1 Data Embedding process of proposed scheme

```
Input: Original image ($I_{m\times n}$), Data stream M;
Output: Sub-sampled stego images, $ZS_1^{'}, ZS_2^{'}, ZS_3^{'}, ZS_4^{'}$;
Step-1: Create sub-sampling images using equation $(1)$;
Step-2: Create interpolated sub-sampled image $ZS_1, ZS_2, ZS_3, ZS_4$ using equation $(2)$ on sub-sampled images;
Step-3:
for (i=1 to length(M)) do
    Select the ASCII value $(a_0)$ of the $i^{th}$ character;
end
for x=1 to 4 do
    (1). $f(x) = a_0 + a_{(t-2)}x^{(t-2)} + a_{(t-1)}x^{(t-1)}$;
    (2). Convert the value of f(x) into 12 bits binary form by padding zero and divide into 4 parts with 3 bits each;
    (3). Take $(4 \times 4)$ block $b_k$ from $ZS_k$ and embed the 4 parts by adding with the value at position
    $b_k(1,2), b_k(2,1), b_k(2,3), b_k(3,2)$. And the value of $x$ is added with the value at position $b_k(2,2)$.
    where $(k = 1, 2, \ldots, (m - 1 \times n - 1))/(4 \times 4)$.
end
Step-4: Repeat Step-3 until all data are embedded within four sub-sampled images.
Step-5: Produce four sub-sampled stego images $ZS_1^{'}, ZS_2^{'}, ZS_3^{'}, ZS_4^{'}$.
```

Algorithm 2 Data Extraction process of proposed scheme

```
Input: Stego sub-sampled images $ZS_1^{'}, ZS_2^{'}, ZS_3^{'}, ZS_4^{'}$ of size $(m \times n)$ and data length;
Output:Cover image ($I'_{m \times n}$), Secret data M;
Note:The receiver can retrieve data using any three sub-sampled stego image out of four;
Step-1:
for (p=1 to 3) do
    (a). Select $(4 \times 4)$ block $b_k^{'}$ from $ZS_p^{'}$
    where $k = 1, 2, \ldots (m \times n)/(4 \times 4)$.
    (b). Retrieve the value $(v_1, v_2, v_3, v_4, v_5)$ from position
    $b_k^{'}(1,2), b_k^{'}(2,1), b_k^{'}(2,3), b_k^{'}(3,2), b_k^{'}(2,2)$ using the following formula.
    1. $v_1$ and $v_4 = ZS_1^{'}(i,j) - (ZS_1^{'}(i,j-1) + ZS_1^{'}(i,j+1)/2)$, where $i\%2! = 0\&j\%2 = 0$;
    2. $v_2$ and $v_3 = ZS_1^{'}(i,j) - (ZS_1^{'}(i-1,j) + ZS_1^{'}(i+1,j)/2)$, where $i\%2 = 0\&j\%2! = 0$;
    3.
    $v_5 = ZS_1^{'}(i,j) - (ZS_1^{'}(i-1,j-1) + ZS_1^{'}(i-1,j+1) + ZS_1^{'}(i+1,j-1) + ZS_1^{'}(i+1,j+1)/4)$,
    where $i\%2 = 0\&j\%2 = 0$;
    where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots n$.
    $v_5$ represent the value of $x$ means sampled number. Convert rest four value into 12 bits binary form
    $(v_1 + v_2 + v_3 + v_4)$ and convert into decimal value which represent the functional value of $v_5$.
end
Step-2: Apply Step-1 using $ZS_2^{'}(i,j)$, $ZS_3^{'}(i,j)$ and $ZS_4^{'}(i,j)$.
Step-3: Apply lagrange interpolation polynomial to extract secret data using the sampled secret which are
extracted from any three sub-sampled stego images. Choose any one combination out of these combinations of
x. (1,2,3), (2,3,4), (1,3,4), (2,1,4).
Step-4: Collect all original pixel values from four sub-sampled stego images and rearrange them to get original
cover image.
Step-5: Produce $I_{m \times n}$ and secret data M.
```

During extraction, we apply pixel value difference from stego image separately. We collect the image number and the difference value which is later combined and produce the lagrange function generated value. After extracting all the sampled values from stego images we apply lagrange formula by taking any three sets of sampled values collected from sub-sampled stego images to get the original secret data. In this way hidden data *M* are extracted as well as original cover image is retrieved.

2.2 Overflow-underflow control

We only add the three bits binary value with interleaved pixel. Over flow may occur during data embedding. For example, if any interleaved pixel (IP) is 248 and we try to add 8, then new pixel (NP) becomes $(248 + 8) = 256$ which is greater then 255. So, overflow situation will be arises but underflow situation may not arises because no subtraction operation is performed during data embedding. To overcome this overflow problem, we update the interleaved pixel (IP) before data embedding. If pixel is greater then 247 then NP is updated by

$$NP = \begin{cases} 247, & \text{if } IP > 247 \\ IP, & \text{otherwise.} \end{cases} \tag{4}$$

By this way, no overflow situation will arise. Similarly at the receiver end, receiver can easily found IP and NP. For extraction of embedded value we calculate

$$Value = \begin{cases} NP - 247, & \text{if } IP > 247 \\ NP - IP, & \text{otherwise.} \end{cases} \tag{5}$$

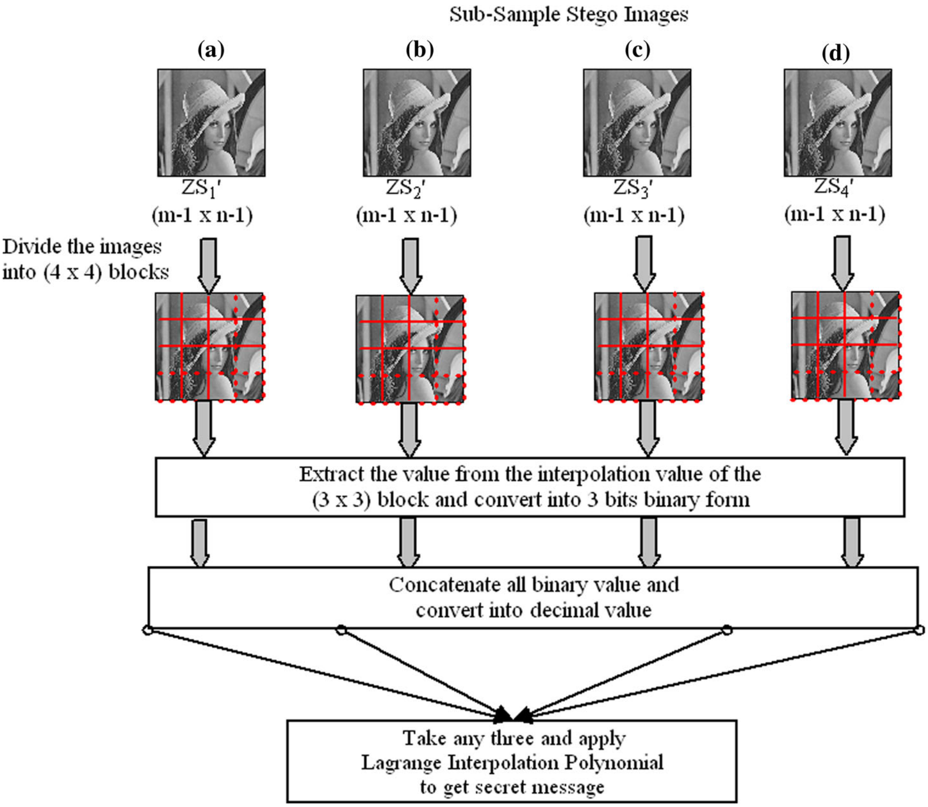

Sub-Sample Stego Images
(a)
(b)
(c)
(d)
ZS1'
(m-1 x n-1)
ZS2'
(m-1 x n-1)
ZS3'
(m-1 x n-1)
ZS4'
(m-1 x n-1)
Divide the images into (4 x 4) blocks
Extract the value from the interpolation value of the (3 x 3) block and convert into 3 bits binary form
Concatenate all binary value and convert into decimal value
Take any three and apply Lagrange Interpolation Polynomial to get secret message

Fig. 4 Block diagram of data extraction procedure

3 Experimental results and comparison

Our developed algorithms: data embedding and extraction are implemented in MATLAB Version 7.6.0.324 (R2008a). Here, the impairment is assessed by means of two factors namely, Mean Square Error (MSE) and Peak Signal to Noise Ratio ($PSNR$). The MSE is calculated as follows:

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[X(i,j) - Y(i,j)]^2}{(M \times N)}, \tag{6}$$

where M and N denote the total number of pixels in the horizontal and the vertical dimensions of the interpolated sub-sampled image respectively. $X(i, j)$ represents the pixels in the original interpolated sub-sampled image and $Y(i, j)$ represents the pixels of the sub-sampled stego image. The difference between the original and stego images is assessed by the Peak Signal to Noise Ratio ($PSNR$). The analysis in terms of PSNR of interpolated sub-sampled image and sub-sampled stego image shows reasonably good results which is shown in Table 1. PSNR is calculated using the following equation.

$$PSNR = 10\,log_{10}\frac{255^2}{MSE}, \tag{7}$$

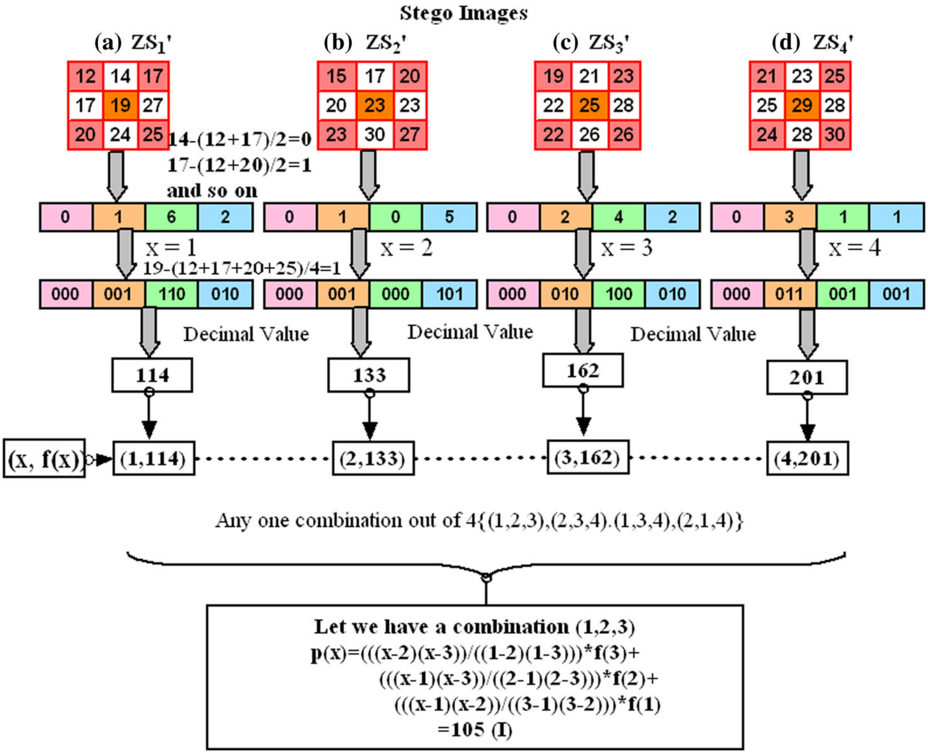


Fig. 5 Numerical Illustration of data extraction

Higher the values of PSNR between two images, better the quality of the stego image and very similar to the cover image where as low PSNR demonstrates the opposite. The payload in terms of bits per pixel (bpp) is calculated by the following formula.

$$B = \frac{\frac{m}{x} \times \frac{n}{y} \times (r)}{(m \times n \times s)}, \tag{8}$$

where m and n represent the size of the input image, that is, $I_{(m \times n)}$). x and y represent size of the block. r represents the number of bits which are hidden in each block, s represents the number of stego images. Consider $m = 512$, $n = 512$, $x = 4$, $y = 4$, $r = 8$ and $s = 4$. So, payload $B = 0.125$ bpp. Within 512×512 images we can hide $(512 \times 512)/(4 \times 4) \times 8$ bits $=$ 131072 bits. Also we can calculate the total bits which we embedded within stego images are $(512 \times 512)/(4 \times 4) \times 60$ bits $= 983040$ bits. In that case the payload is 0.94 bpp. We calculate actual bits and overhead bits and these are 131072 bits and (983040–131072) $=$ 851968 bits respectively. The experimental results are shown in Table 1.

The images are used for our experiment is shown in Fig. 6. After embedding maximum data, stego sampled image are generated which are shown in Fig. 7.

In terms of actual embedding capacity(bpp) and image quality (dB), the presented algorithm was compared with Ni et al. [14], Varsaki et al. [17], Hwang et al. [6], Kuo et al. [10], Tsai et al. [16] and Kim et al.'s [9] scheme for the Lena and Baboon images as shown in Table 2. We observed that the average PSNR of the stego images of our proposed method is

Table 1 Data embedding capacity with PSNR

Image(I)	Data (bits)	PSNR($ZS_1 vs ZS_1'$)	PSNR($ZS_2 vs ZS_2'$)	PSNR($ZS_3 vs ZS_3'$)	PSNR($ZS_4 vs ZS_4'$)
Lena	40,000	71.25	70.18	70.41	70.23
	80,000	63.79	63.17	63.78	63.88
	1,20,000	55.97	54.42	54.89	53.78
	1,30,000	50.60	50.07	49.32	50.67
Barbara	40,000	72.61	71.11	70.23	69.35
	80,000	62.68	60.10	62.5	61.36
	1,20,000	55.32	54.75	53.89	54.21
	1,30,000	49.55	48.02	50.12	50.61
Tiffany	40,000	72.69	73.21	70.23	70.45
	80,000	63.79	62.19	62.12	63.28
	1,20,000	54.99	53.44	52.24	52.34
	1,30,000	49.23	48.08	50.23	50.46
Mrittika	40,000	72.20	72.64	71.45	71.89
	80,000	62.50	61.90	62.90	61.78
	1,20,000	52.63	54.87	53.23	52.36
	1,30,000	48.97	48.44	49.32	48.56
Peppers	40,000	72.71	73.20	71.23	72.45
	80,000	64.32	64.17	63.24	62.87
	1,20,000	52.98	51.42	53.24	54.12
	1,30,000	48.61	48.06	48.23	49.56
Gold-Hill	40,000	72.71	73.20	71.21	71.24
	80,000	61.23	62.12	62.32	61.23
	1,20,000	52.95	51.42	51.64	52.34
	1,30,000	49.59	48.08	49.23	48.56

around 50 dB after embedding 1,30,000 bits. The embedding capacity and quality is higher than other existing schemes.

3.1 RS analysis

We analyze our stego images through the RS analysis scheme [4]. When the value of RS analysis is closed to zero it means that the scheme is secure. It is observed from Table 3 that the values of R_M and R_{-M}, S_M and S_{-M} are nearly equal. Thus rule $R_M \cong R_{-M}$ and $S_M \cong S_{-M}$ are satisfied for the stego image (ZS_1' in our proposed scheme. So, the proposed method is secure against RS attack. The results of stego images (ZS_1' are recoded in Table 3.

3.2 Relative entropy

To measure the imperceptibility in our proposed method, the relative entropy (R) between the probability distributions of the original image (P) and the stego image (Q) has been calculated by

$$R(Q||P) = \sum q(x) log \frac{q(x)}{p(x)} \tag{9}$$

Fig. 6 Input images of size (512×512)

When relative entropy between two probability distribution functions is zero then the system is perfectly secure. $R(Q||P)$ is a nonnegative continuous function and equals to zero if and only if $p(x)$ and $q(x)$ are coincide. Thus $R(Q||P)$ can be normally considered as a distance between the measures $p(x)$ and $q(x)$. Relative entropy of the probability distribution of the original image and the stego image varies depending upon number of bits of secret message. In our experiment, it is shown that when the number of characters in the secret message are increases, the relative entropy in stego image is also increases. The relative entropy are less in numbers which implies that the proposed scheme provides secure hidden communication. Relative entropy values are listed in Table 4 for original image and Table 5 for stego images.

4 Attack

4.1 Statistical attack

In this section, we analyze the robustness of our proposed scheme. When two variables are consider in bivariate data, we say these two variables to be correlated if the change in the value of one is related to the change in the value of other. Correlation may be of

(a1) Mrittika Sub-Sampled Stego Image (ZS_1')

(b1) Lena Sub-Sampled Stego Image (ZS_1')

(c1) Peppers Sub-Sampled Stego Image (ZS_1')

(d1) Tiffany Sub-Sampled Stego Image (ZS_1')

(a2) Mrittika Sub-Sampled Stego Image (ZS_2')

(b2) Lena Sub-Sampled Stego Image (ZS_2')

(c2) Peppers Sub-Sampled Stego Image (ZS_2')

(d2) Tiffany Sub-Sampled Stego Image (ZS_2')

(a3) Mrittika Sub-Sampled Stego Image (ZS_3')

(b3) Lena Sub-Sampled Stego Image (ZS_3')

(c3) Peppers Sub-Sampled Stego Image (ZS_3')

(d3) Tiffany Sub-Sampled Stego Image (ZS_3')

(a4) Mrittika Sub-Sampled Stego Image (ZS_4')

(b4) Lena Sub-Sampled Stego Image (ZS_4')

(c4) Peppers Sub-Sampled Stego Image (ZS_4')

(d4) Tiffany Sub-Sampled Stego Image (ZS_4')

Fig. 7 Sub-sampled stego images of size (512×512)

two types: i) If we increase the value of one variable brings on average the increase in the value of the other variable, then these two variables are said to be positively correlated. ii) If the increase in the value of one variable brings on average the decrease in the value of the other variable, then these two variables are said to be negatively correlated. In a scatter diagram, it is noticed in most of the cases that the points plotted in a diagram are more or less concentrated in the neighborhood of a curve which is called Regression curve. In the case of simple regression when the two regression curve are linear, then their degree of collinearity is measured by a quantity known as correlation coefficient and it is denoted by ρ_{xy}. Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be a set of n pairs of observations in a bivariate

Table 2 Comparison results in terms of the payload (bits) and the PSNR(dB) for Lena and Baboon

	Lena Image		Baboon Image	
	Payload (bits or bpp)	PSNR	Payload (bits or bpp)	PSNR
Ni et al. [14]	5460	48.2	5421	48.2
Varsaki et al. [17]	5460	48.2	5421	48.2
Hwang et al. [6]	5408	48.2	5208	48.2
Kuo et al. [10]	5418	48.2	5352	48.2
Tsai et al. [16]	13459	49.3	NA	48.1
Kim et al. [9]	20121	48.9	6499	48.7
Lee et al. [11]	1.2 bpp	48.16	0.88 bpp	48.32
Wang et al. [20]	53945	49.20	20592	48.87
Proposed Scheme	1,30,000	50.60	1,30,000	49.32

distribution having two variables x and y. Then the correlation coefficient between the two variables x and y is define as follows

$$\rho_{xy} = \frac{Cov(x, y)}{S_x \times S_y} \tag{10}$$

where, $cov(x, y) = \frac{1}{n}. \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ is called the co-variance between x and y and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, where $\bar{x}$ and $\bar{y}$ denote the means of the variables x and

Table 3 RS analysis for Stego images ZS_1' of size (512 × 512)

Image	Data	ZS_1'				
		R_M	R_{-M}	S_M	S_{-M}	RS value
Barbara	40000	6896	6922	3807	3817	0.0034
	80000	6403	6451	4270	4237	0.0076
	120000	6074	6138	4496	4468	0.0087
	130000	6152	6122	4527	4479	0.0073
Lena	40000	5490	5586	4142	4050	0.0195
	80000	5427	5550	4280	4149	0.0262
	120000	5422	5586	4424	4312	0.0280
	130000	5484	5535	4406	4448	0.0094
Baboon	40000	5872	5812	5010	5141	0.0176
	800000	5800	5815	5116	5112	0.0017
	120000	5851	5770	5118	5219	0.0166
	130000	5856	5757	5109	5215	0.0187

Table 4 Relative entropy of original image (512×512)

Original Image(OR)	Entropy
Lena	7.0299
Barbara	7.4429
Tiffany	7.2371

y respectively. Here, S_x and S_y are the standard deviation of x_i and $y_i, i = 1, 2, \ldots, n$. $S_x = \sqrt{\frac{1}{n}\sum x_i^2 - \bar{x}^2}$ and $S_y = \sqrt{\frac{1}{n}\sum y_i^2 - \bar{y}^2}$.

We calculate the standard deviation (SD) before and after data embedding and correlation coefficient (CC) of cover and stego images are summarized in Table 6. Minimizing parameters difference is one of the primary aims in order to get rid of statistical attacks. It observe that there is no substantial divergence between the standard deviation of the cover-image and the stego-images. This study shows that the magnitude of change in stego-images based on image parameters is small from a cover image. Since the image parameters have not changed much, the method offers a good concealment of data and reduces the chance of the secret data being detected. Thus, it indicates a perfectly secure steganographic system.

4.2 Attacks with unknown Lagrange polynomial coefficient

The proposed scheme produce four (4) sub-samples stego images which contain secret information. We have embedded secret information in a distributed manner within interleaved pixel of sub-sampled stego-images. We use predefined shared lagrange polynomial coefficient. The scheme is secure to prevent possible malicious attacks. The Fig. 8 shows the example of getting noise data when applied wrong Lagrange Polynomial Coefficient to revel the hidden message. If the malicious attacker holds the original image and stego image and

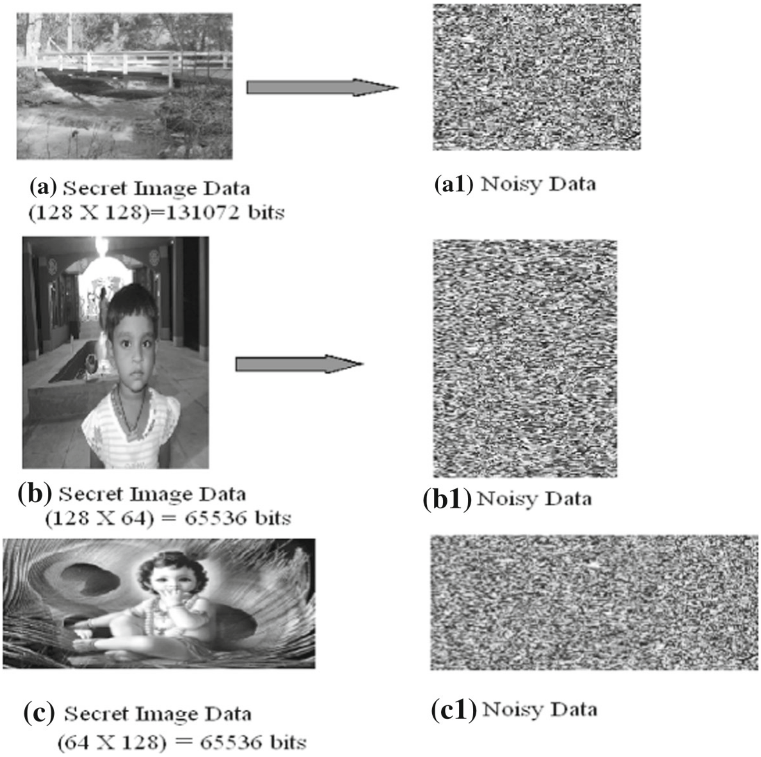
Table 5 Relative entropy of stego images ZS'_1 and ZS'_2

Image	Data	ZS'_1		ZS'_2	
		Entropy	Difference	Entropy	Difference
Lena	40000	7.0572	0.04	7.1143	0.0227
	80000	7.1220	0.0148	7.1143	0.0295
	120000	7.1547	0.1375	7.1458	0.1792
	130000	7.1555	0.1383	7.1452	0.128
Barbara	40000	7.4491	0.0224	7.4494	0.0062
	80000	7.4550	0.0283	7.4562	0.0147
	120000	7.4653	0.0386	7.4622	0.0355
	130000	7.4668	0.0401	7.4654	0.0387
Tiffany	40000	7.2393	0.0471	7.2394	0.0472
	80000	7.2438	0.0561	7.2437	0.0515
	120000	7.2471	0.0549	7.2461	0.0539
	130000	7.2469	0.0547	7.2475	0.0553

Table 6 Standard Deviation (SD) and Correlation Coefficient (CC) of proposed method

Image	SD			CC		
	I	ZS_1'	ZS_2'	I and ZS_1'	I and ZS_2'	ZS_1' and ZS_2'
Lena	61.58	61.70	61.67	0.99	0.99	0.99
Barbara	47.83	47.96	47.94	0.99	0.99	0.99
Baboon	38.37	38.48	38.50	0.99	0.99	0.99

is fully aware of the proposed scheme, the hidden message still cannot be correctly revealed without knowing the correct lagrange polynomial coefficient. Furthermore, the attacker may employ the brute force attack that tries all possible permutation to reveal the hidden message by finding correct lagrange polynomial coefficient is computationally infeasible for current computers by an adversary. The proposed scheme achieve stronger robustness against several attacks. But, the secret information can be retrieved without encountering any loss of secret data and recovered original image successfully from sub-sampled stego images.


(a) Secret Image Data (128 X 128)=131072 bits

(a1) Noisy Data

(b) Secret Image Data (128 X 64) = 65536 bits

(b1) Noisy Data

(c) Secret Image Data (64 X 128) = 65536 bits

(c1) Noisy Data

Fig. 8 Noise like secret data with wrong lagrange polynomial coefficient

5 Conclusion

A new reversible data hiding scheme through lagrange interpolation using sub-sampled image is proposed in this paper. We generate and enlarge sampled image from original image. The original secret message are converted into another secret using lagrange interpolation. We keep those secret data within sub-sampled stego images. The proposed scheme achieve secure hidden data communication because we only embed lagrange value not original secret value. Also secret data can be extracted using a specified threshold number of sub-sampled image not all sub-sampled image are required. In this scheme, we have achieved average PSNR greater than 50 dB and payload 130,000 bits. We have also tested our scheme using RS analysis, statistical analysis (such as Relative entropy, Standard Deviation and Correlation Coefficient) which have provided promising results. We have observed that the scheme is secure against several steganographic attacks.

Acknowledgements The author is grateful to the anonymous reviewers for their invaluable comments and recommendations to improve our paper. This work is supported by UGC Innovative project grant under Vidyasagar University (Grant Nos: VU/Innovative/ Sc/06/2015, Dated: 17.06.2015).

References

1. Chan C, Cheng L (2004) Hiding data in images by simple LSB substitution. Pattern Recogn 37(3): 474–496
2. Chang C, Hsiao J, Chan C (2003) Finding optimal least-significant-bits substitutionin image hiding by dynamic programming strategy. Pattern Recogn 36(7):1583–1595
3. Chang C, Chan C, Fan Y (2006) Image hiding scheme with modulus function and dynamic programming. Pattern Recogn 39(6):1155–1167
4. Fridrich J, Goljan J, Du R (2001) Invertible authentication. In: Proceedings of the SPIE, security and watermarking of multimedia contents, vol 4314. SanJose, pp 197208
5. Giri D, Jana B, Mondal SK (2016) Dual image based reversible data hiding scheme using three pixel value difference expansion. In: Information systems design and intelligent applications. Springer, India, pp 403–412
6. Hwang J, Kim JW, Choi JU (2006) A reversible watermarking based on histogram shifting. International workshop on digital watermarking, lecture notes in computer science, vol 4283. Springer, Jeju Island, p 348361
7. Jana B (2016) Dual image based reversible data hiding scheme using weighted matrix. Int J Electron Inf Eng 5(1):6–19
8. Jana B, Giri D, Mondal SK (2016) Dual image based reversible data hiding scheme using (7, 4) hamming code. Multimed Tools Appl 1–23
9. Kim KS, Lee MJ, Lee HY, Lee HK (2009) Reversible data hiding exploiting spatial correlation between sub-sampled images. Pattern Recogn 42(11):3083–3096
10. Kuo WC, Jiang DJ, Huang YC (2007) Reversible data hiding based on histogram. International conference on intelligent computing, lecture notes in artificial intelligence, vol 4682. Springer, Qing Dao, pp 1152–1161
11. Lee CF, Chang CC, Gao CY (2013) A two-staged multi-level reversible data hiding exploiting lagrange interpolation. In: 2013 Ninth international conference on intelligent information hiding and multimedia signal processing. IEEE, pp 485–488
12. Lin CC, Tai WL, Chang CC (2008) Multilevel reversible data hiding based on histogram modification of difference images. Pattern Recogn 41(35):82–91
13. Luo H, Yu FX, Chen H, Huang ZL, Li H, Wang PH (2011) Reversible data hiding based on block median preservation. Inform Sci 181(2):308328
14. Ni Z, Shi YQ, Ansari N, Su W (2006) Reversible data hiding. IEEE Trans Circ Syst Vid Technol 16(3):354–362
15. Thien C, Lin J (2003) A simple and high-hiding capacity method for hiding digitby-digit data in images based on modulus function. Pattern Recogn 36(12):2875–2881

16. Tsai PY, Hu YC, Yeh HL (2009) Reversible image hiding scheme using predictive coding and histogram shifting. Signal Process 89(11):29–43
17. Varsaki E, Fotopoulos V, Skodras AN (2006) A reversible data hiding technique embedding in the image histogram. Technical Report HOU-CS-TR-2006-08-GR Hellenic Open University
18. Wang R, Lin C, Lin J (2001) Image hiding by optimal LSB substitution andgenetic algorithm. Pattern Recogn 34(3):671–683
19. Wang SJ (2005) Steganography of capacity required using modulo operator forembedding secret image. Appl Math Comput 164(1):99–116
20. Wang XT, Chang CC, Nguyen TS, Li MC (2013) Reversible data hiding for high quality images exploiting interpolation and direction order mechanism. Digit Signal Process 23(2):569–577

Biswapati Jana is currently working as an Assistant Professor in the Department of Computer Science, Vidyasagar University, Paschim Medinipur, India. He received his B. Tech. and M. Tech. degrees in Computer Science and Engineering from University of Calcutta in 1999 and 2002 respectively. He submitted his Ph.D. thesis in July, 2016. His research interests include Image Processing, Steganography, Data Hiding, Visual Cryptography. He has published more than thirty papers in National and International Journals and Conferences.