Data Hiding Method Using Adaptive Pixel Pair Matching

Shaikh Salman[1], **Prof. S. R. Kinge**[2]

[1]PG Student, Department of E&TC, MIT College of Engineering, Savitribai Phule Pune University, Pune, India

[2]Professor, Department of E&TC, MIT College of Engineering, Savitribai Phule Pune University, Pune, India

Abstract: ***Data Hiding or Steganography has been an important communicated network since people started communicating in writing. The main aim in steganography is to hide a secret message within cover media in such a way that an observer cannot detect the presence of contents of the hidden message. Cover images are original images without secret data and after embedding secret data they are called as stego images. Steganography and cryptography are counter parts in digital security the obvious advantage of steganography over Cryptography is that messages do not attract attention to themselves, to messengers, or to recipients. The progress in steganography has also led to many serious problems such as hacking, compression, reformate, etc. Steganography finds its role in attempt to address these growing concerns. We know that, with the use of steganographic techniques, it is possible to hide secret information within images, Audio and video files which is statistically undetectable. This paper proposes a new data embedding technique which uses two pixel sequentially one after another which is kwon as pixel pair matching (PPM) technique. The basic idea of PPM is to use the values of pixel pair as a reference coordinate, and search a coordinate in the neighbourhood set of this pixel pair according to a given message digit. The two pixel is then replaced by the searched coordinate to store the digit. The APPM method offers lower distortion than DE by providing more compact neighbourhood sets and allowing embedded digits in any notational system.***

Keywords: Steganography, stego image, least significant bit (LSB), optimal pixel adjustment process (OPAP), diamond encoding (DE), Adaptive Pixel Pair Matching (APPM).

1. Introduction

Steganography/Data hiding is a secure communication method that conveys secret messages in the form of plaintexts so that the appearances of the secret messages will not draw eavesdroppers' attention while they are being transmitted through an open channel. The word steganography is originally derived from Greek words which mean ''Covered Writing''. It has been used in various forms for thousands of years. In the 5th century B.C Histaiacus shaved a slave's head, tattooed a message on his skull and the slave was dispatched with the message after his hair grew back[1].

1.1 Development

For decades peoples worked on development on methods of secret communication. Generally three methods are used for security systems steganography, cryptography & watermarking [1].

A. Steganography

Carrier used in communication for steganography can be any digital media. Main objective of steganography is for secret communication. There is no visibility in the o/p file. Any cover image can be used for communication

B. Cryptography

Carrier used in cryptography is text or image file. Main objective is data protection. There is visibility in o/p file.

C. Watermarking

Carriers are mostly image /audio files. Main objective is copyright preservation. Cover image choice is restricted.

2. Related Work

Steganography is checked on two aspects imperceptibility and embedding capacity (payload). Imperceptibility is nothing but the differences between stego image and cover image which is not visible to human eyes. This measurement is done by MSE and PSNR. Payload is maximum number of bits that can be embedded in pixel with acceptable stego image quality.

Steganography can be done in spatial domain as well as transform domain. Spatial domain has advantages over transform domain. It is simpler and faster to implement the techniques, stego image quality is under control with high embedding capacity. In this paper we shall introduce methods that embeds message in spatial domain.

2.1 LSB method

Most traditional method used for data embedding is LSB method .as we know that every pixel value is eight bit which can be represented as

$$P_x = (a_7 a_6 a_5 a_4 a_3 a_2 a_1 a_0) = \sum_{x=0}^{7} a_x 2^x$$

P_x any pixel value
a_7 MSB , a_0 LSB

Usually $a_2 a_1 a_0$ can be used to hide secret data because LSB bits carry less information than MSB .Embedding procedure goes on every pixel one by one. Let the message bits be $s_2 s_1 s_0$.

$$P_x' = (a_7 a_6 a_5 a_4 a_3 s_2 s_1 s_0)$$

We directly replace 3 LSB bits of pixels with the message bits. P' is the new pixel value with the embedded data[2].

2.2 Optimal pixel adjustment process

It was proposed by Chan *et al.* In 2004 [3].The simple LSB method is modified so that stego image quality is improved. OPAP algorithm uses LSB method as its base. OPAP gives good results than LSB replacement method [3][4]. Let p be the pixel value. Decimal value of right most n LSB be p^n. p' be the pixel value by direct embedding and m be message data to be embedded in the pixel. OPAP employs following equations to embed a data to get minimum distortion.

$$P'' \begin{cases} p' + 2^n, p^n - m > 2^{n-1}\ and\ p' + 2^n \leq 255 \\ p' - 2^n, p^n - m < -2^{n-1}\ and\ p' - 2^n \geq 0 \\ p', \quad otherwise \end{cases}$$

P'' is the result obtained after algorithm implementation[5].

Example:- let the pixel value p=160=10100000(2), Data to be embedded be m= 7=111(2). Let n(replacing bits) = 3.By direct LSB replacement p'=10100111=167. Check for the three conditions of OPAP algorithm. It satisfies $p^n - m < -2^{n-1}\ and\ p' - 2^n \geq 0$ = 0-7< -2^{3-1} .P''= p' - 2^n= 167 - 8 = 159. Hence 7 is embedded in the pixel 160 with distortion of only one unit.

Extraction:- The n LSB bits of the stego image pixels is nothing but the data embedded ,hence extraction is very simple.

3. PPM (pixel pair matching method)

In PPM method data is embedded to reduce the embedding impact by providing a simple extraction and a more compact neighbourhood set. Diamond encoding (DE) and adaptive pixel pair matching (APPM) are PPM method.

In 2009 Chao *et al.* [6] proposed a PPM based method Diamond Encoding. DE preserves acceptable stego image quality. Message is embedded in B-ary notational system, where

$$B = 2x^2 + 2x + 1\ ,\quad x > 1$$

B = payload & x is variable

The image quality obtained by these methods has less imperceptibility. The scanning method has to unique for transmitter and receiver side in communication.

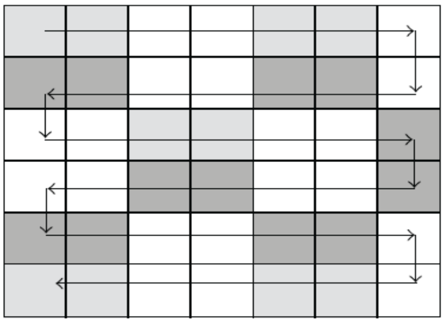

Figure 1: Sequence of non-overlapping consecutive two-pixel blocks is constructed in a cover image

3.2 Adaptive pixel pair matching

It was proposed by W.Hong and Tung-Shou Chenin, 2012 [7]. The idea behind pixel pair matching method is that message bits are embedded in pixel pair and searching a coordinate (*p', q'*) within a predefined neighbourhood set *Ψ(p, q)* such that *f(p, q)= m*, where p & q are two pixels, m be message bits and *f(p, q)* be the extraction function. Data embedding is done by replacing pixels *(p, q)* by (*p', q'*) .The range of m should be 0 to B-1. where, B is payload.

Neighbourhood Set ψ(p, q) and Extraction Function f(p, q)

ψ(p, q) and *f(p, q)* affects the stego image quality. The design of *ψ(p, q)* & *f(p, q)* have following requirements (1). all the values of *ψ(p, q)* & *f(p, q)* should be mutually exclusive (2) the summation of squared distance between the coordinates in has to be smallest. During embedding (*p, q*) is modified according to the coordinates in *ψ(p, q)* & *f(p, q)* are designed such that MSE is reduced.

$$f(p,q) = mod(p + c \times q, B)$$

The solution of $\Psi(p,q)$ & f(*p, q*) is actually a discrete optimization problem,

$$Min \sum_{i=0}^{B-1} (p_i - p)^2 + (q_{i\_}q)^2$$

$$Subject : f(p_i, q_i) \in \{0,1,2 \ldots.B-1\}$$
$$f(p_i, q_i) \neq f(p_j, q_j), \quad if\ i \neq j$$
$$for\ 0 \leq i,j \leq B-1 \qquad (1)$$

Table 2: List of Constants

	3	4		
	9	10	11	
14	15	0	1	2
	5	6	7	8
		12	13	

Figure 2: For Ψ_{16} & $c_{16} = 6$

C_2	C_4	C_8	C_{10}	C_{12}	C_{16}	C_{20}	C_{24}	C_{32}	C_{64}
1	2	3	3	4	6	8	5	7	14

Given an integer b and integer pair (*p, q*), (1) can solved to obtain a constant c and B. $\Psi(p,q)$ denote neighbourhood set of (*p, q*) .Table below lists the constants c satisfying (1) for different payloads.

Embedding and Extraction of Data

A.Steps for embedding

let the cover image be of size M ✕M

s be the message bits to be embed

find minimum size of B such that all the message bits can be embedded

$$\lfloor M \times M/2 \rfloor \geq |m|$$

Solve discrete optimization problem to find c and Ψ(*p*,*q*)
To embed a m message bit ,two pixels (*p*,*q*) are selected and modulus distance d between m and f(*p*,*q*) is calculated

$$d = mod(m - f(p,q), B)$$

Find the value d in the Ψ(*p*, *q*) and get its coordinates .modify the pixel values according to the coordinates

$$(p,q) = (p + p_d , q + q_d)$$

Repeat the steps until all the message bits are embedded

B. Steps for extraction

From stego image take two pixels in the same way as they are scanned and apply extraction function on it.

$$f(p',q') = mod(p' + c \times q', B)$$

$f(p',q')$is the hidden data in the image

Example:- let (*p*, *q*)=(20,30). let data to be embedded be 15. Cover size be 512×512 with embedding requirement of 520000 bits. The minimum B satisfying the condition is 16. So we select 16-ary notational system. By solving discrete optimization problem $c_B = c_{16} = 6$ and $\Psi_{16}(p,q)$ can be obtained. Extraction function is *f*(20,30) = mod(20+6×30,16) = 8. Modulus distance between m &extraction function d = mod (15-8, 16) = 7. Coordinates of 7 in Ψ(20,30) is p+1 & q+1. Hence $p + p_d$= 20 + 1 = 21, $q + q_d$ = 30 + 1 = 31. Hence after embedding 15 in (20, 30), pixels are modified as (21, 31).distortion in the pixels is of one unit only.

Extraction:- from extraction function $f(p',q')$ = mod(21+6×31, 16) = 15. Hence we got the embedded data.

When any stego-pixel value has the overflow or underflow problem the critical vector has to be adjusted to the appropriate value, the adjustment rules are defined as follows.

$(1)\ if\ \ p' > 255\ \ , p' = p' - B$
$(2)\ if\ \ p' < 0\ \ \ \ , p' = p' + B$
$(3)\ if\ \ q' > 255\ \ , q' = q' - B$
$(4)\ if\ \ q' < 0\ \ \ \ , q' = q' + B$

4. Experimental Results And Analysis

This section presents and analyzes the experimental results by using the proposed method. To evaluate the performance of our new scheme, in our experiments, we have used about 500 images with size 256 × 256 at their maximum payload capacity. Our evaluation starts with the six well-known images Lena, Cameraman, Bird, Baboon, House, Peppers.

In our experiments, the quality of the stego-image is measured by the Mean Square error (MSE) and Peak signal-to-noise ratio (PSNR). These two parameters are the most popular criterion to measure the distortion between the cover image and stego-image. It is defined as follows:

$$PSNR = 10 \times 10Log_{10}\left(\frac{255^2}{MSE}\right)$$

Where MSE is the mean square error between the cover image and stego-image:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j) - I(i',j')]^2$$

Here, the symbols *I* (*i*, *j*) and *I* (*i'*, *j'*) represent the pixel values of the cover image and stego-image in the position (*i*, *j*) respectively, and *M* and *N* are the width and height of the original image.

To evaluate the performance of the proposed scheme, a uncompressed images are taken. The simulation is run using MATLAB. First, LSB, OPAP, DE and APPM are evaluated for Mean Square Error (MSE) and PSNR at their maximum payload capacity. Table.2 and Table.3 presents the obtained MSE and PSNR values respectively. It is observed that APPM has minimum MSE & maximum PSNR as compared to DE, OPAP and LSB method which indicated by using APPM method we can achieve least distortion in stego image.

Table 2: MSE comparison

Images	LSB	OPAP	DE	APPM
Lena	10.37	5.44	2.05	1.32
Cameraman	10.33	5.54	2.06	1.33
Baboon	10.53	5.55	2.06	1.34
Bird	10.32	5.49	2.05	1.33
House	10.42	5.54	2.06	1.33
Pepper	10.44	5.54	3.46	1.78

Table 3: PSNR comparison

Images	LSB	OPAP	DE	APPM
Lena	37.97	40.77	45.00	46.90
Cameraman	37.98	40.80	44.98	46.86
Baboon	37.90	40.85	44.98	46.89
Bird	37.99	40.73	44.99	46.88
House	37.94	40.70	44.97	46.87
Pepper	37.94	40.60	42.72	45.60

LSB and OPAP are single pixel embedding methods. The stego image quality improves when we use OPAP method instead of LSB replacement method. Experimental result shows that PSNR value of the OPAP method is greater than LSB method. Stego image quality for 4 bits LSB replacement is acceptable for naked eyes but when embedding bits are increased beyond 4 bits it causes clear distortion. However in smooth areas even 4 bits LSB replacement method causes a noticeable distortion.

In case of PPM method they offer a high payload and acceptable stego image quality is preserved. APPM gives good results than DE method. APPM reduces the embedding impact by providing a simple extraction function and a more compact neighbourhood set. It embeds more messages per modification and increases the efficiency. The stego image quality after embedding has last MSE and is less detectable in APPM as compared to LSB, OPAP& DE.

APPM gives better results than DE because the payload of DE is determined by the selected notational system which is restricted by the parameter x,

$$B = 2x^2 + 2x + 1\ ,\quad x > 1$$

B = payload & x is variable

The neighbourhood set in DE is defined by diamond shape only which leads to unnecessary distortion, but in case of APPM there exists a better Ψ(*p*, *q*) other than diamond shape resulting in smaller embedding distortion. Four corners of diamond cause larger MSE so we select more compact region for embedding.

5. Conclusion

Steganography is the art of secret communication under the cover of digital images. In this paper we have reviewed some good data embedding schemes which have either high stego image quality or high embedding capacity. If we want to embed large amount of data and if stego image quality is not so important than use OPAP method, But when image quality is of greater importance than embedding capacity, than DE & APPM are the best choice.

Though the embedding capacity is low they give a greater image quality where distortion is invisible to human eyes. APPM allows to select digits in any notational system for data embedding and achieves a better image quality than DE & OPAP. APPM is the best for secure communication under adjustable embedding capacity.

References

[1] A. Cheddad, J. Condell, K. Curran, and P. McKevitt," Digital image Steganography: Survey and analysis of current methods," Signal Process., vol. 90, pp. 727–752, 2010.

[2] J. Fridrich, M. Goljan, and R. Du, "Reliable detection of LSB steganography in color and grayscale images" in *Proc. Int. Workshop on Multi media and Security*, pp. 27–30, 2001.

[3] C. K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," Pattern Recognition, pp. 469–474, March, 2004.

[4] C. H. Yang, "Inverted pattern approach to improve image quality of information hiding by LSB substitution," Pattern Recognition., vol. 41,no. 8, pp. 2674–2683, 2008.

[5] C. C. Thien and J. C. Lin," A Simple and high hiding capacity method for hiding digit by digit data in images based on modulus function", pattern Recognition ,vol.36. no. 12. pp 2875-2881,2003.

[6] R.M Chao, H. C. Wu , C. C. Lee and Y. P. Chu, "A novel image data hiding scheme with diamond encoding" EURASIP J Info security, vol. 2009,2009.

[7] Wien Hong and Tung-Shou Chen" A Novel Data Embedding Method Using Adaptive Pixel Pair Matching," IEEE transaction on information forensics and security vol. 7, no. 1, February 2012.