Global Journal of Researches in Engineering: J
General Engineering
Volume 14 Issue 6 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4596 & Print ISSN: 0975-5861

A Secure Steganographic Technique for Embedding Text using Adaptive Pixel Pair Matching

By K Madhuri & Ms C Padmini
Vardhaman Colloge of Engineering, India

Abstract- Steganography is a method of hiding secret messages in a cover object while communication takes place between sender and receiver. Generally data embedding is achieved in text, image, audio, video, network for the purpose of secret communication. This paper proposes a secure method for hiding text based on adaptive pixel pair matching (APPM). The basic idea of Pixel Pair Matching (PPM) is to use the values of pixel pair as a reference coordinate, and search a coordinate in the neighborhood set of this pixel pair according to a given message digit. The pixel pair is then replaced by the searched coordinate to conceal the digit. APPM allows users to select digits in any notational system for data embedding, and thus achieves a better image quality. Compared with the optimal pixel adjustment process (OPAP) Method and diamond encoding (DE), the proposed method always has lower distortion for various payloads. This paper proposes an extension to conceal text into an image for conveying secret messages confidentially.

Keywords: steganography, least significant bit method, pixel pair matching, diamond encoding.

GJRE-J Classification : FOR Code: 291899

Strictly as per the compliance and regulations of:

A Secure Steganographic Technique for Embedding Text using Adaptive Pixel Pair Matching

K Madhuri [α] & Ms C Padmini [σ]

Abstract- Steganography is a method of hiding secret messages in a cover object while communication takes place between sender and receiver. Generally data embedding is achieved in text, image, audio, video, network for the purpose of secret communication. This paper proposes a secure method for hiding text based on adaptive pixel pair matching (APPM). The basic idea of Pixel Pair Matching (PPM) is to use the values of pixel pair as a reference coordinate, and search a coordinate in the neighborhood set of this pixel pair according to a given message digit. The pixel pair is then replaced by the searched coordinate to conceal the digit. APPM allows users to select digits in any notational system for data embedding, and thus achieves a better image quality. Compared with the optimal pixel adjustment process (OPAP) Method and diamond encoding (DE), the proposed method always has lower distortion for various payloads. This paper proposes an extension to conceal text into an image for conveying secret messages confidentially.

Keywords: steganography, least significant bit method, pixel pair matching, diamond encoding.

I. Introduction

Today the growth in the information technology, especially in computer networks such as Internet, Mobile communication, and Digital Multimedia applications such as Digital camera, handset video etc. has opened new opportunities in scientific and commercial applications. But this progress has also led to many serious problems such as hacking, duplications and malevolent usage of digital information. Steganography finds its role in attempt to address these growing concerns [6]. Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. Many different carrier file formats can be used, but digital images are the most popular because of their frequency on the Internet. In image steganography, the aim is to hide information in to a given image called as cover image and the diagnosis of the hidden information will be probably difficult [10]. Every steganographic methods consist of a cover image and a stego image.

Many approaches of information hiding have been proposed for different applications, such as copyright protection, secret transmission, tampering detection, and image authentication.

Author α, σ: Vardhaman Colloge of Engineering, Hyderabad India. e-mail: madhuri.427@gmail.com, c.padmini@vardhaman.org

The most well-known data hiding scheme is the least significant bits (LSBs) substitution method [1], [13]. This method embeds fixed-length secret bits into the least significant bits of pixels by directly replacing the LSBs of cover image with the secret message bits. Although this method is simple, it generally effects noticeable distortion when the number of embedded bits for each pixel exceeds three [1]. Several methods have been proposed to reduce the distortion induced by LSBs substitution. OPAP scheme searches the minimal distortion value which LSBs equal the embedded bits and replaces stego-pixel value with it [12]. Another way of improving LSBs scheme is to reduce the amount of alterations necessary to be introduced into the cover image for data hiding when the number of secret bits is significantly less than that of available cover pixels.

Another method called optimal pixel adjustment process (OPAP) method [14] is introduced to reduce the distortion caused by LSB replacement. In LSB and OPAP methods one pixel is used as an embedding unit[11], and conceal data into the right-most LSBs. OPAP is conceptually defined as matching pixel to its optimal level. OPAP effectively reduces the image distortion compared with the traditional LSB method [12]. But in OPAP method, imbalanced embedding distortion emerges and is vulnerable to steganalysis. LSB and OPAP methods are not suitable for applications requiring high payload.

An efficient data hiding method is proposed for gray-scale images by utilizing the diamond encoding concept (DE). We first transform the secret data into a sequence of digits, and the cover image is partitioned into non-overlapping blocks of two consecutive pixels. The diamond encoding method produces a diamond characteristic value (DCV) of the pixel-pair block, and the DCV [3] is revised as the embedded secret digit after data embedding procedure. For each block, the diamond encoding technique addresses the minimal changes of two pixel values under the embedding parameter k. In other words, the difference between the cover-block and the stego-block is never more than k, and the embedding capacity of a block equals $\log_2(2k^2+2k+1)$. The payload of DE [2] is determined by the selected notational system, which is restricted by the parameter k; therefore, the notational

system cannot be arbitrarily selected. For example, when is 1, 2, and 3, then digits in a 5-ary, 13-ary, and 25-ary notational system are used to embed data, respectively. However, embedding digits in a 4-ary (i.e., 1 bit per pixel) or 16-ary (i.e., 2 bits per pixel) notational system are not supported in DE. Secondly, $\phi(x, y)$ in DE [1], [2], [4] is defined by a diamond shape, which may lead to some unnecessary distortion when k > 2. In fact, there exists a better $\phi(x, y)$ other than diamond shape resulting in a smaller embedding distortion.

a) Adaptive Pixel Pair Matching For Embedding Digits

The basic idea of the PPM-based data-hiding method is to use pixel pair (x, y) as the coordinate, and searching a coordinate (x', y') within a predefined neighborhood set Ø(x, y) such that $f(x', y') = S_B$, where *f* is the extraction function and S_B is the message digit in B-ary notational system to be concealed [1], [3]. Data embedding is done by replacing (x, y) with (x', y').

Suppose the cover image is of size *M*×*M*, *S* is the message bits to be concealed and the size of *S* is $|S|$. First we calculate the minimum *B* such that all the message bits can be embedded. Then, message digits are sequentially concealed into pairs of pixels. First minimum *B* satisfying $|M \times M / 2| \geq |S_B|$ and convert *S* into a list of digits with a *B*-ary notational system S_B. The discrete optimization problem is solved to find c_B and $\phi(x, y)$. In the region defined by $\phi(x, y)$, record the coordinate (x', y') such that $f(x', y') = i, 0 \leq i \leq B-1$

Construct a non repeating random embedding sequence *Q* using a key K. To embed a message digit S_B, two pixels (x, y) in the cover image are selected according to the embedding sequence *Q*, and calculate the modulus distance $d = (S_B - f(x, y)) \bmod B$ between S_B and $f(x, y)$, then replace (x, y) with $(x + x', y + y')$.

The rest of the paper is organized as follows. Section II deals with proposed methodology. Embedding and extraction procedures of text are given in sections IV and section V concludes the paper.

II. Proposed Methodology

As APPM is proved to offer better security against detection and lower distortion, we can take forward APPM for hiding text in an image. This method is proposed to explore a better mechanism and provide better security and lower distortion for embedding text in images. The ascii values of all the characters in the given text are converted to binary values and then they are partitioned into groups of bits. Then the embedding process is performed for embedding them into an image.

Data embedding is done by replacing (x, y) with (x', y'). These are the reference coordinate and pixel value from the neighborhood set. The concept of a PPM-based steganographic method is that, let S_B be the message bit is to be concealed and the range of S_B is between 0 and $B-1$. And there should be a coordinate (x', y') has to be found such that $f(x', y') = S_B$. That is why the range of $f(x, y)$ must be within integers between 0 and $B-1$. [2] Also here each integer must occur at least once. In APPM, consider the compact neighborhood set for reducing the distortion. The the following three conditions should be satisfied by the best ppm based data hiding method.

a) There are exactly B number of coordinates in the neighborhood set $\phi(x, y)$.

b) These coordinates and the values of extraction function must be mutually exclusive.

c) The design aspects of neighborhood set $\phi(x, y)$ and the extraction function $f(x, y)$ should be capable of embedding the message bits in least notational system.

i. *Conversion of input text into appropriate form for embedding*

First the ASCII values of all characters in the secret data are converted into an array of eight bit binary numbers. Now the array of binary numbers is divided into groups containing each containing four number of bits. The decimal value of each group is embedded into pixel pairs of a cover image.

c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	c_{17}	c_{18}
1	1	2	2	2	2	3	3	3	3	4	5	4	4	6	4	4
c_{19}	c_{20}	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	c_{26}	c_{27}	c_{28}	c_{29}	c_{30}	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}
4	8	4	5	5	5	5	10	5	5	5	12	12	7	6	6	10
c_{36}	c_{37}	c_{38}	c_{39}	c_{40}	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	c_{46}	c_{47}	c_{48}	c_{49}	c_{50}	c_{51}	c_{52}
15	6	16	7	7	6	12	12	8	7	7	7	7	14	14	9	22
c_{53}	c_{54}	c_{55}	c_{56}	c_{57}	c_{58}	c_{59}	c_{60}	c_{61}	c_{62}	c_{63}	c_{64}					
8	12	21	16	24	22	9	8	8	8	14	14					

Fig. 1 : List of c_B for 2 ≤ B ≤ 16

ii. *Finding Neighborhood Set And Extraction Function*

In this module, the extraction function is explained. By using this method, we can get a simple extraction function and compact neighborhood set. Thus the proposed method enhances the embedding efficiency. The quality of image obtained by this method is much better than the other existing data hiding method such as OPAP and DE [1]. Another two advantages of this proposed method are higher payload capability and less detectability. Compact notional system is used for data embedding which gives increased performance.

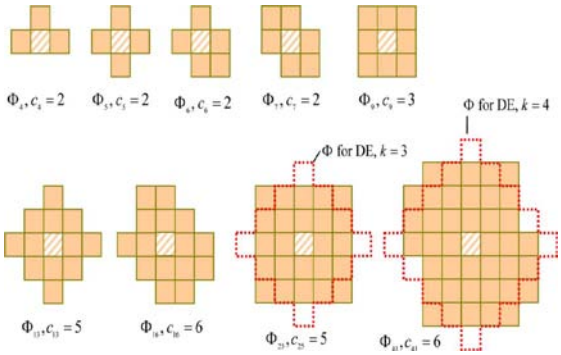

Fig. 2 : Neighborhood set for APPM

The stego image quality is significantly affected by both definitions of $\phi(x,y)$ and $f(x,y)$. All values of $f(x,y)$ in $\phi(x,y)$ must be mutually exclusive, and also the summation of the squared distances between all the coordinates present in $\phi(x,y)$ and (x,y) has to be the smallest. This is because, during embedding procedure the pixel value (x,y) is replaced by one of the pixel in the neighborhood set $\phi(x,y)$. If there is B number of coordinates in $\phi(x,y)$, then message bits a B-ary notational system are to be concealed. The averaged MSE can be calculated by averaging the summation of squared distance between the coordinates (x,y) and other coordinates in $\phi(x,y)$ the reference coordinate. Thus, MSE after embedding can be calculated the following equation by

$$\text{MSE}\,\phi(x,y) = \frac{1}{2B}\sum_{i=0}^{B-1}((\mathrm{x_i} - \mathrm{x})2 + (\mathrm{y_i} - \mathrm{y})2)$$

The adaptive pixel pair matching (APPM) data-hiding method is used to explore better $\phi(x,y)$ and $f(x,y)$. So that MSE is minimum compared with the other existing methods. In this method the extraction function is

$$f(x,y) = (x + c_B \times y) \bmod B$$

So the calculation of both neighborhood set $\phi(x,y)$ and the extraction function $f(x,y)$ can done by a discrete optimization problem. For this following conditions have to be considered.

Minimize to $\sum_{i=0}^{B-1}((\mathrm{x_i} - \mathrm{x})2 + (\mathrm{y_i} - \mathrm{y})2)$

Subject to : $f(x_i, y_i) \in \{0,1.....,B-1\}$

$$f(x_i, y_i) \neq f(x_j, y_j)$$

for $0 \le i, j \le B-1$

From figures (1) and (2) show some neighborhood sets $\phi_B(x,y)$ and their corresponding c_B values which satisfy the above condition. In the above figure 3 the shaded with lines represents the center of $\phi_B(x,y)$.

III. Data Embedding Procedure

Here the secret data has to be embedded into the given cover image. For this first we should calculate the image size and message size. If the message size exceeds size of the image, then the embedding procedure cannot be done. Consider the image size as M*M, For S message bits the size of secret message S is $|S|$. By using these, calculate the minimum B value such that all the message bits can be embedded. The message digits will be sequentially concealed into pairs of pixels.

The data embedding process is shown by a flowchart in fig 3 below. The detailed procedure is listed as follows.

a) First the ASCII values of all characters in the secret data are converted into an array of eight bit binary numbers.

b) Now the array of binary numbers is divided into groups containing each containing four number of bits.

c) The decimal value of each group is embedded into pixel pairs of a cover image.

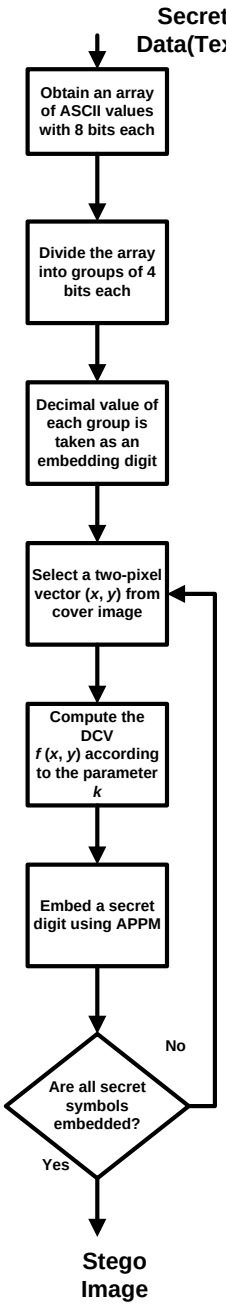

Fig. 3 : Data Embedding Process

d) Calculate the minimum B satisfying

e) $|M \times M / 2| \geq |S_B|$.

f) Convert the secret message S into the sequence of digits with a B -ary notational system.

g) Find c_B and $\phi_B(x, y)$ using the discrete optimization equations.

h) From the neighborhood region $\phi_B(0,0)$ find the coordinate positions (x_i, y_i) by satisfying the condition $f(x_i, y_i) = i, 0 \leq i \leq B-1$.

i) Create a non repeat random key Kr for embedding the secret message bits Q.

j) To embed a secret message bits B, find the two pixels (x, y) in the cover image and replace (x, y) with $(x + x_d, y + y_d)$ for the modulus distance $d = (S_B - f(x, y)) \bmod B$ between S_B and $f(x, y)$.

k) Repeat Step 6 and 7 until all the secret message bits are concealed.

IV. Data Extraction Procedure

To extract the embedded message digits, pixel pairs are scanned in the same order as done in the embedding procedure. The value of extraction function of a scanned pixel pair gives the embedded digit. Now the text input is retrieved from this output array of digits. The data extraction process is shown by a flowchart in fig 4 below.

a) Take two pixels positions (x', y') and calculate $f(x', y')$.

b) The value of $f(x', y')$ is the embedding digit.

c) The decimal value of these output digits are converted as an array of binary numbers, where each digit is represented with four bits.

d) Now the array of binary numbers is divided into groups of eight bits each.

e) The decimal value of each group of bits gives the ASCII value of the character.

f) After converting all the groups of bits into characters, we will get the input text as the output.

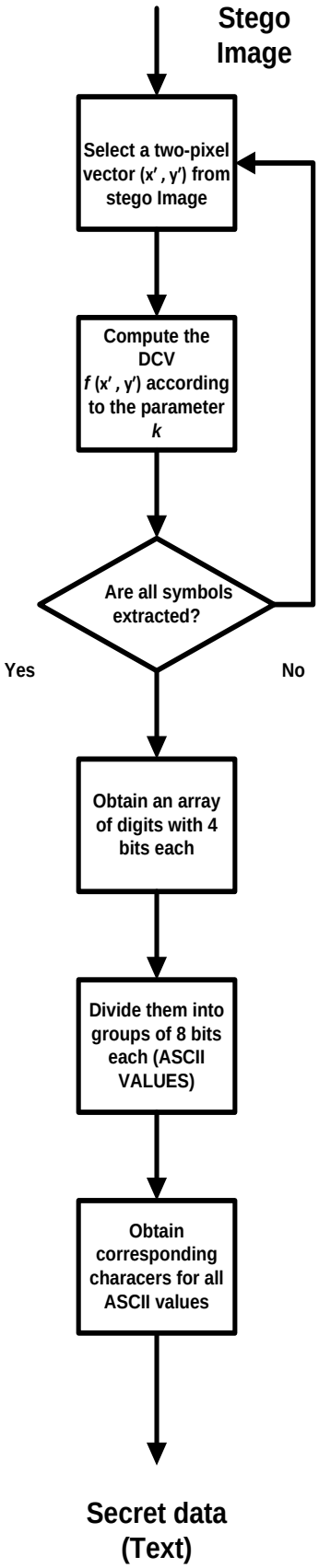

Fig. 4 : Data Extraction Process

V. Theoretical Analysis and Experimental Results

When data embed in an image, the pixel values in that image may modified and this process is known as image distortion or embedding distortion. MSE (Mean Square Error) is used to measure this distortion. MSE is calculated by the following equation

$$\frac{1}{M \times M} \sum_{i=0}^{M} \sum_{j=0}^{M} (p_{i,j} - p'_{i,j})^2$$

Where M×M is the image size, $p_{i,j}$ denotes the pixel values of original image and $p'_{i,j}$ denotes the pixel values of stego image. Here the mean square error between the cover image and stego image is represented by MSE.

Table 1 : Mean square error of APPM for embedding digits and text

Image	APPM(DIGITS) $c_B = 6$	APPM(TEXT) $c_B = 6$
Clock.tiff	0.8220	0.8167
Nature.png	0.2049	0.2041
Cartoon.png	0.2931	0.2754

The smaller MSE is for APPM which indicate the better image quality. APPM is flexible and gives less mean square error while embedding digits [1] and as well as embedding text also.

Fig. 5 : Cover Image

Fig. 6 : Stego Image

Cover image and stego image are shown above. Less detectability to the text hided in the image

is one of the desired aspect. Both the images are visually similar. Therefore the attacker cant determine whether the image is encoded or not.

VI. Conclutions

This paper proposed an efficient data embedding algorithm for hiding text in an image based on APPM. Here two pixel positions are scanned and are considered as a scanning unit. And a specially designed neighborhood set with smallest notational system is used for embedding text and hence a better image quality is achieved. The steganalysis results of stego images are similar to those of the cover images, which offer a secure communication under adjustable embedding capacity. It also contains additional features such as digital watermark and encryption of secret messages for the provision of more security. APPM technique can also be used for embedding data in audio and video also. All these various features made this APPM technique a straightforward and economical embedding method for the data hiding.

References Références Referencias

1. "A Novel Data Embedding Method Using Adaptive Pixel Pair Matching" Wien Hong and Tung-Shou Chen.
2. "A novel image hiding scheme by Optimal Pixel Pair Matching and Diamond Encoding" Rajashree Shitole, Satish Todmal.
3. "Edge Adaptive Image Steganography Based on Adaptive Pixel Pair Matching" Akhil P. V, Akbersha K. E., Mohammed Sidheeque.
4. Ruey-Ming Chao, Hsien-ChuWu, Chih-Chiang Lee, and Yen-Ping Chu "A Novel Image Data Hiding Scheme with Diamond Encoding".
5. J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
6. N. Provos and P. Honeyman, "Hide and seek: An introduction to steganography," *IEEE Security Privacy*, vol. 3, no. 3, pp. 32–44, May/Jun. 2003.
7. 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.
8. T. Filler, J. Judas, and J. Fridrich, "Minimizing embedding impact in steganography using trellis-coded quantization," in *Proc. SPIE, Media Forensics and Security*, 2010, vol. 7541, DOI: 10.1117/12.838002.
9. S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 1, pp. 111–119, Mar. 2006.
10. J. Fridrich, M.Goljan, and R.Du, "Reliable detection of LSB steganography in color and grayscale images," 2001, pp. 27–30.
11. A. D. Ker, "Steganalysis of LSB matching in grayscale images," *IEEE Signal Process. Lett.*, vol. 12, no. 6, pp. 441–444, Jun. 2005.
12. C. K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern Recognit.*, vol. 37, no. 3, pp. 469–474, 2004.
13. J. Mielikainen, "LSB matching revisited," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 285–287, May 2006.
14. X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Commun. Lett.*, vol. 10, no. 11,pp. 781–783, Nov. 2006.