Optik 127 (2016) 1762–1769

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

High capacity data hiding scheme based on multi-bit encoding function

Wen-Chung Kuo[a,*], Shao-Hung Kuo[b], Chun-Cheng Wang[b], Lih-Chyau Wuu[a]

[a] Department of Computer Science and Information Engineering, National Yunlin University of Science & Technology, R.O.C., No. 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan
[b] Graduate School of Engineering Science and Technology Doctoral Program, National Yunlin University of Science & Technology, R.O.C., No. 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan

ARTICLE INFO

Article history:
Received 26 July 2014
Accepted 18 August 2015

Keywords:
Data hiding
Cover image
Stego image
Exploiting modification direction
Steganographic method

ABSTRACT

Secret message capacity and stego image quality are the two important measures for stenography. In this paper, a high hiding capacity and acceptable stego image quality steganographic scheme is proposed based on multi-bit encoding function. Using our method, the embedding capacity is as high as 4.5 bpp (with acceptable stego image quality, i.e., PSNR better than 30 dB). There are two major contributions of this scheme. It does not need additional complex steps to embed the secret data and no additional external information is needed to recover the secret data. From our experimental results and discussion, we show that our proposed scheme achieves higher capacity than other data hiding schemes based on encoding function while maintaining suitable image quality.

1. Introduction

Data hiding techniques are important in the field of information security. One common goal is to embed secret messages in a cover image to generate a stego image that prevents others from obtaining the secret data. Other goals include enhancing embedding capacity while maintaining good image quality. However, it is not trivial to achieve both goals at the same time. Currently, there are three major digital image embedding methodologies. First, the least significant bit (LSB) replacement method is the simplest and best known way of embedding secret data directly [1,2,4,13,17,19–21]. Second, pixel value differencing (PVD) steganography, proposed in 2003, is also a method for data hiding. It uses the mean value characteristic to embed secret data [18]. The final type of data hiding is the encoding function, which uses a special equation or distinctive function to embed secret data [8–10,12,15,20,23]. In this paper, we will focus on data hiding scheme based on encoding function.

In 2006, Zhang and Wang [23] used the modulus method to modify the directional characteristics and then proposed data hiding technology based on the exploiting modification direction (EMD) method. In other words, the EMD-scheme uses the relationship of n adjacent pixels to embed the secret data. Note, only one of the adjacent n pixels is modified when the secret data want to be embedded. However, the largest data hiding bit rate is 1.16 bpp (bits per pixel) when it is 5-ary, i.e., $n = 2$. The results show that the n of the EMD-schem is inversely proportional to the rate of capacity. To enhance hiding capacity, other studies propose modifications to the EMD-type hidden data schemes [8–10,12,15,20]. For example, Lee et al.[12] proposed a high embedding capacity scheme (LWC) to improve the embedding capacity from 1.16 bpp to 1.5 bpp in 2007. Although its embedding capacity is higher, it employs an extraction function with fixed weighting which is weak against steganalysis. In order to address the security problem and maintain embedding efficiency, Kuo et al. proposed an improved high capacity EMD data hiding technique with modulus table (KKH-scheme) [8]. In 2012, to improve the stego image quality and embed capacity, Hong and Chen used adaptive pixel pair matching (APPM) to propose a novel data embedding method [7]. The advantages of the Hong-Chen scheme include offering lower distribution than the diamond encoding method [3] and allows embedding secret data in the 64-ary notation instead of binary. According to our analysis, it just only uses the relationship between two adjacent pixels to embed the secret data, i.e., it cannot be extended to use the relationship between n adjacent pixels ($n > 2$), as with the LWC, KKH and Hong–Chen schemes. Recently, to overcome this shortcoming, a new data hiding scheme based on generalized exploiting modification direction method (GEMD) was proposed by Kuo and Wang [10]. The main contribution of the Kuo–Wang scheme is that $(n + 1)$ binary bits can be embedded into n adjacent pixels directly. The simulation results of the Kuo–Wang scheme show it can embed

[*] Corresponding author. Tel.: +886 55342601.
E-mail address: simonkuo@yuntech.edu.tw (W.-C. Kuo).

http://dx.doi.org/10.1016/j.ijleo.2015.08.056
0030-4026/© 2015 Elsevier GmbH. All rights reserved.

additional data for a total of $\left(1+\frac{1}{n}\right)$ secret bits on average for each pixel. In other words, it cannot embed more than two bits of secret data for each pixel in [10].

Fu et al. [6] proposed a scheme for reversible data hiding based on prediction-error histogram shifting and EMD mechanism. In [6], they combine multi-level shifting and nonary EMD to increase capacity. Recently, Wu et al. [20] proposed a high payload hybrid data hiding scheme with LSB, EMD and modification of prediction errors (MPE) to achieve higher embedding capacity while maintaining high visual image quality of the stego-image.

In order to enhance the embedding capacity for each pixel, we propose a high capacity embedding method based on multi-bit encoding function in this paper. Two major advantages of this scheme are a modified coefficient parameter and a new extraction function approach. Therefore, the embedding capacity can be raised from 1.5 bpp to 4.5 bpp while still producing an acceptable stego image (with PSNR >30 dB). The simulation results and performance analysis show the proposed scheme not only maintains high embedding capacity but also keeps good stego image quality. In terms of security, the proposed scheme is more secure than LSB in terms of resisting visual attack and RS testing [16,22].

The rest of this paper is organized as follows: In Section 2, we will introduce the Zhang–Wang, LWC, KKH and Kuo–Wang schemes. The proposed scheme and experimental results are detailed in Section 3 and Section 4, respectively. Finally, conclusions are given in Section 5.

2. Related works

2.1. Exploiting modification direction method

In 2006, a novel data hiding scheme based on EMD method was proposed by Zhang and Wang [23]. The major embedding idea is that the secret information is injected into a group of cover image pixels. During the EMD data hiding procedure, pixels of cover image are represented as $(x_1, x_2, \ldots, x_n)$, where n represents the number of pixels in a pixel group. Before the hiding procedure, the binary secret data is converted into a $(2n+1)$-ary structure. The extraction function as Eq. (1) is provided by Zhang and Wang.

$$f_a(x_1, x_2, \ldots, x_n) = \left[\sum_{i=1}^{n}(i \times x_i)\right] \bmod (2n+1), \tag{1}$$

where x_i is the value of the pixel i. For example, when n is equal to 2, i.e., two pixels x_1 and x_2, are considered. Therefore, the extract function is $f_a(x_1, x_2) = (1 \times x_1 + 2 \times x_2) \bmod 5$.

Algorithm 1. EMD Embedding Algorithm
Input: n adjacent pixels $(x_1, x_2, \ldots, x_n)$ and $(2n+1)$-ary secret message (m)
Output: n adjacent stego-pixels $(y_1, y_2, \ldots, y_n)$

```
Step1.  Calculate $t = f_a(x_1, x_2, \ldots, x_n)$ using Eq. (1).
Step2.  Calculate the difference $d = (m - t) \bmod (2n+1)$.
Step3.  If $(d = 0)$, then $(y_1, y_2, \ldots, y_n) = (x_1, x_2, \ldots, x_n)$, else if $(n > d)$,
        then $(y_1, y_2, \ldots, y_d, \ldots, y_n) = (x_1, x_2, \ldots, x_d + 1, \ldots, x_n)$, else
        $(y_1, y_2, \ldots, y_{(2n+1)-d}, \ldots, y_n) = (x_1, x_2, \ldots, x_{(2n+1)-d} - 1, \ldots, x_n)$.
```

Therefore, for each block, the embedding procedures of Zhang–Wang scheme [23] include the following steps:

EMD-E1. Take n adjacent pixels $(x_1, x_2, \ldots, x_n)$ as a group from the cover image I_c.
EMD-E2. Adjust the group pixels $(x_1, x_2, \ldots, x_n)$ by using Algorithm 1 for all pixel groups.
EMD-E3. Get the stego image I_s.

According to their analysis, the largest hiding bit rate is 1.16 bpp for 5-ary. Unfortunately, there is a serious drawback though, since the hiding bit rate is less than 1 bpp when n is greater than 3 [23].

2.2. High embedding capacity by improving EMD

In 2007, Lee et al. [12] proposed the improving embedding idea, which the modulus in Eq. (1) is modified from 5 to 8 when $n = 2$, i.e., the data hiding capacity is embedding 3 binary bits directly for each two adjacent pixels. Therefore, a modified encoding function f_e is defined as Eq. (2)

$$f_e(x_1, x_2) = (1 \times x_1 + 3 \times x_2) \bmod 8. \tag{2}$$

Algorithm 2. LWC-embedding algorithm
Input: A pixel pair (x_1, x_2) and the secret message s
Output: A stego-pixel pair (x'_1, x'_2)

```
Step1.  Calculate $f_e(x_1, x_2)$.
Step2.  Adjust $(x_1, x_2)$ such that the secret message by using following relationships:

(A2-1)  Let $x'_1 = x_1$ and $x'_2 = x_2$, when $s = f_e(x_1, x_2)$.
(A2-2)  Let $x'_1 = x_1 + 1$ and $x'_2 = x_2$, when $s = f_e(x_1 + 1, x_2)$.
(A1-3)  Let $x'_1 = x_1 - 1$ and $x'_2 = x_2$, when $s = f_e(x_1 - 1, x_2)$.
(A2-4)  Let $x'_1 = x_1$ and $x'_2 = x_2 + 1$, when $s = f_e(x_1, x_2 + 1)$.
(A2-5)  Let $x'_1 = x_1$ and $x'_2 = x_2 - 1$, when $s = f_e(x_1, x_2 - 1)$.
(A2-6)  Let $x'_1 = x_1 + 1$ and $x'_2 = x_2 + 1$, when $s = f_e(x_1 + 1, x_2 + 1)$.
(A2-7)  Let $x'_1 = x_1 + 1$ and $x'_2 = x_2 - 1$, when $s = f_e(x_1 + 1, x_2 - 1)$.
(A2-8)  Let $x'_1 = x_1 - 1$ and $x'_2 = x_2 + 1$, when $s = f_e(x_1 - 1, x_2 + 1)$.
```

Thus, the embedding procedure in the LWC-scheme is shown as the following:

LWC-E1. Pair up all the pixels in the cover image I_c.
LWC-E2. Adjust the pixel pair value with the secret data by using Algorithm 2 for all pixel pairs.
LWC-E3. Get the stego image I_s.

Though its embedding capacity always is kept 1.5 bpp, an encoding function with fixed weighting is used in LWC scheme. In order to solve the security problem, an improved high capacity EMD data hiding technique with modulus table was proposed in 2013 [8].

2.3. Generalized exploiting modification direction method

Unfortunately, there is a major disadvantage which only uses the relationship between two adjacent pixels to embed the secret data, i.e., it cannot be extended to use the relationship between n adjacent pixels $(n > 2)$ directly in the LWC [12] or KKH [8] schemes. Therefore, Kuo and Wang proposed a novel data hiding scheme based on GEMD method to overcome this shortcoming in 2013 [10]. According to their performance analysis, the main contribution [10] is that each $(n+1)$-bit binary secret message can be hidden into n adjacent pixels in the cover image. Thus, the new extraction function $f_b(x_1, x_2, \ldots, x_n)$ is defined as Eq. (3):

$$f_b(x_1, x_2, \ldots, x_n) = \left[\sum_{i=1}^{n}(a_i \cdot x_i)\right] \bmod 2^{n+1}, \tag{3}$$

where $a_i = (2^i - 1)$ and x_i are the group pixels.