Reversible Palette Image Steganography Based on De-clustering and Predictive Coding

Hsien-Chu Wu[1], Hui-Chuan Lin[2] and Chin-Chen Chang[3]

[1]Department of Information Management, National Taichung Institute of Technology
E-mail:wuhc@ntit.edu.tw
[2] Graduate School of Computer Science and Information Technology,
National Taichung Institute of Technology
E-mail:chuan@ntit.edu.tw
[3] Department of Information Engineering and Computer Science,
Feng Chia University, Taichung, Taiwan, 40724, R.O.C.
E-mail:ccc@cs.ccu.edu.tw

Abstract

Steganographic technology acts as the important role which defends information security in the network. The information could be hidden in multimedia object. Even if transmitted in an unsafe channel, the information is invisible by human visual system to maintain security. This paper presents a reversible palette image steganography based on de-clustering and predictive coding. The information is embedded into the image during compression encoding. It does not only solve problems encountered in interception and theft over Internet. The proposed method reforms the drawback of distortion of other schemes. When retrieving the information from stego-image, the original image also can be recovered. The proposed method is robust and useful for keeping transmission secure.

Keywords: steganography, palette image, HVS, de-clustering, stego-image

1. Introduction

In past twenty years, various encryption technologies were proposed to keep information security. But it requires much computation and time to encrypt/ decrypt. Further, the encrypted code must be kept integrity. The primordial can not be rehabilitated as soon as any code changed or lost. Therefore, another protection mechanism, steganography was proposed. Steganography is a high security technology achieved message interchange purpose. The message was kept in a cover image before transmission [3-10]. The hidden message is unperceptible by human visual system (HVS). Users will not suspect something in the image. The transmission security can be achieved. So far, many applications combined these two techniques, that is, the information was encrypted by traditional cryptography, and then embedded into some media object before transmission. Thus, the information security gets double protection [3-6, 8-10].

A palette image is presented as a set of indexes and a palette. The palette is a list of colors composed of RGB values that specified from the original image. Considering a 24-bits full-color image transformed to 256 colors palette image, each pixel was presented as an 8-bits index pointing to the palette element which is the closest color instead of a 24-bits RGB value. Consequently, modifying the index will decrease image quality more than pixel value. The palette-based steganography has a common drawback which stego-image after data hiding cause more distortion [3, 6].

Over these ten years, many palette based image steganographic technologies were proposed in literature. In [6], Machado proposed EZ-stego method which is a common LSB-like method. The palette colors are sorted by luminance, and then the message were embedded to the LSB of indexes. Since the same luminance makes up with different RGB value, the colors may differ significantly. The stego-image caused serious distortion. Then, Fridrich [3] proposed another method. The secret bit was hidden by replacing the index of the cover image with the closest color which the parity bit is equal to secret bit. The capacity of both methods is one bit per pixel. The stego-image quality of Fridrich's method is higher than of EZ-stego method. Wu *et al*. [9] proposed an iterative method of palette-based image steganography to improve the stego-image quality by adding preprocessing to Fridrich's method. The method is based on a palette modification scheme. Liu *et al*. [5] proposed a high capacity distortion-free data hiding algorithm for palette image by duplicating the palette

colors. If the colors were copied 2^n-1 times, then each pixel could embed n bits. In Wu's and Liu's methods, both are easy to cause users suspicion from the changes of the palette, because of the existence of the same colors in the palette. Tzeng [8] proposed adaptive data hiding in palette images by color ordering and mapping with security protection. The indexes unsuitable for change were filtered out firstly. Then the message was embedded into suitable pixels by replaying the index with a precedent neighbor color of current pixel. The method raised the stego-image quality, but reduced the capacity much more.

In this paper, a reversible palette image steganographic scheme is proposed. The palette colors are grouped and using flags to increase the number of hiding bits. Medium edge detector (MED) predictor [2,10] is used to help the message embedding and extracting. When users extract the message, the original image will also be recovered simultaneously. The method solves the stego-image distortion problem.

The remainder of this paper is organized as follows. Section 2 introduces the proposed method. Section 3 shows the experimental results. Finally, the conclusion is provided in Section 4.

2. The proposed method

Data hiding in palette-based image is more difficult than monochromatic image relatively. It is still difficult to reform the data hiding capacity. In our approach, these two problems will be solved and the message will be embedded in compression code to avoid being looked through and intercepted. The procedure of our approach is presented in Fig. 1.

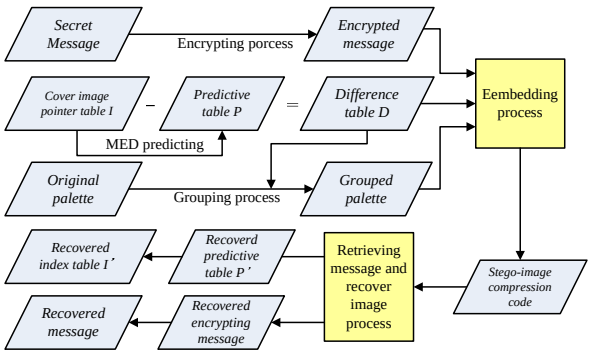

Fig. 1: The flowchart of the proposed method.

At the beginning, the cover image I is a palette image, and each pixel is denoted as an index I_{ij} to point to color of the palette C. In C, each color is presented as $c_i(r_i, g_i, b_i)$. The palette colors are divided into two groups by the de-clustering scheme described in Section 2.2. In the embedding process, a predictive table is generated by the MED predictor referring to the original index table, shown as Subsection 2.1. Then the secret message will be embedded in difference values between the original index table and the predictive table, presented in Subsection 2.3.

2.1. Median edge detector scheme

The MED predictor is used in the JPEG-LS scheme which is an image standard for lossless and near-lossless compression. In the MED predictor, each pixel was processed in left-right and top-down order. Let x represent the current pixel and a, b, c, and d be the four neighboring pixels of x as Fig. 2.

c	b	d
a	x	

Fig. 2: Using a, b, c to predict the pixel x.

The MED predictor uses the three precedent neighbors a, b, and c to check whether any vertical or horizontal edge can be found. When a vertical edge is detected in the current pixel, the MED predictor tends to pick b as the predictive value. On the other hand, it tends to pick a as the predictive value in the event a horizontal edge located above the current pixel is detected. If no edge is detected, the value $a+b-c$ is defined as the predictive value. The used Equation (1) is denoted as follows.

$$x=\begin{cases}\min(a,b) & \text{if } c\ge\max(a,b),\\ \max(a,b) & \text{if } c\le\min(a,b),\\ a+b-c & \text{otherwise.}\end{cases} \quad (1)$$

After MED predicting, the predictive table P is generated. The difference table D then produced by computing the value in table I and P.

2.2. Grouping palette colors

Chang *et al*. proposed the de-clustering concept [1], which the most dissimilar element will be matched with. In our method, the de-clustering concept is used to divide the palette colors into two groups.

It is assumed that C is the palette of a cover image. Pre-scan the values in D obtained in Subsection 2.1 and corresponding to the palette. Select 2^n $(n\ge 1)$ useless indexes from the palette to be flags and allow the flags to form a flag table. Each one of the remaining indexes in the palette will be matched with another color satisfying the distance (see the Equation (2)) of each other is the largest, and every element is not reused.

$$dis(c_i,c_j)=(R_{c_i}-R_{c_j})^2+(G_{c_i}-G_{c_j})^2+(B_{c_i}-B_{c_j})^2. \quad (2)$$

After the matching process, place an element of each pair into group G_0, and another to group G_1, respectively. And then, the palette colors are grouped into two groups G_0 and G_1. For each index c_x in G_0 is

corresponding to an index c'_x in G_1 such that c'_x is the most dissimilar to c_x. The palette is as shown in Fig. 3.

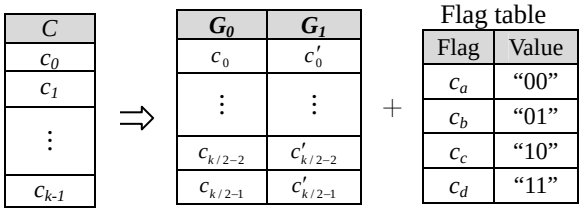

Fig. 3: The palette colors grouping process.

2.3. Embedding secret message

The flowchart of the embedding secret message is shown as Fig. 4.

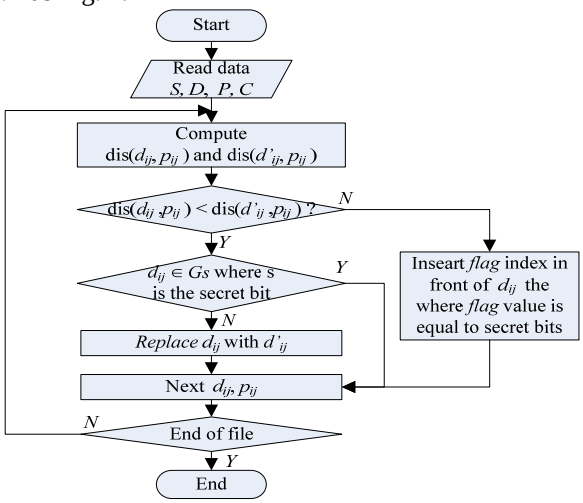

Fig. 4: Secret message embedding process.

Assume that the *ij*-th elements of the difference table *D* and the predictive table *P* are d_{ij} and p_{ij}, respectively. The symbol of d_{ij} is retained in table *D*, the absolute value d_{ij} and p_{ij} are treated as pointers to palette *C* index. Using the Equations (4), if the distance $dis(d_{ij}, p_{ij})$ is less than $dis(d'_{ij}, p_{ij})$, d_{ij} is suitable for embedding. If the index d_{ij} belongs to *Gs* where *s* is a secret bit "0" or "1", d_{ij} requires no any change. It tends to the secret bit *s* be hidden in d_{ij}; otherwise, replace the value of difference table *D* with d_{ij}. Then one bit message can be embedded into each suitable pixel. On the other hand, in case the distance $dis(d_{ij}, p_{ij})$ is greater than $dis(d'_{ij}, p_{ij})$, the d_{ij} is unsuitable to change. Then from the flag table, find the index which flag value is equal to *n* secret bits and insert the flag index in front of the difference index d_{ij}. The *n* is decided by the number of flags. Because we selected 2^n useless palette index to be flags, then *n* bits message is embedded into each of the unsuitable pixels. After the embedding process completed, we also can compress the modified difference table with some other lossless compression schemes. Then the compression codes can be sent to the receiver over the network including the initial value p_{00} of predictive table *P*, the modified difference table *D´* and the grouped palette *C*.

2.4. Extracting the message

When receivers receive the compression code, decompress and yield a difference table *D*, an initial predictive value p_{00} and the palette *C*. By using the extracting algorithm, the message can be retrieved.

In the initial, the image index table *I* and secret message *S* are empty, assume that the *ij*-th elements of *D* and *P* are presented to be d_{ij} and p_{ij}. At this time, the p_{ij} values are all empty but the initial value p_{00}. In the extracting procedure, if d_{ij} belongs to the flag table, then append the flag value to the secret bit stream *S* and take the value behind the flag to recover the image index, let $I_{ij} = d_{ij} + p_{ij}$. On the other hand, if d_{ij} does not belong to flag table then append *s* to secret bit stream where d_{ij} belongs to *Gs*. Check if the distance $dis(d_{ij}, p_{ij})$ is less than $dis(d'_{ij}, p_{ij})$, use d_{ij} to recover the image; otherwise, apply d'_{ij} to recover the image. The used equation is denoted as Equation (3).

$$I_{ij} = \begin{cases} d_{ij} + p_{ij} & \text{if } dis(d_{ij}, p_{ij}) \le dis(d'_{ij}, p_{ij}), \\ d'_{ij} + p_{ij} & \text{otherwise.} \end{cases} \quad (3)$$

Then, next p_{ij} is predicted by MED predictor like embedding part described in Subsection 2.1 by using the recovered data I_{ij}. Repeat the extracting process until the secret bit stream is retrieved and the palette image is recovered. Then the recovered image is the same as the original image if there are no attacks on the way of transmission. It means that the proposed method is reversible.

3. Experimental results

In the experiment, there are three 256 colors image were selected to test. The three images "Garfield", "Joy" and "Lena" are shown as Fig. 5.

(a) Garfield (b) Joy (c) Lena

Fig. 5: The three cover images are all 512×512 pixels.

The preliminary analysis is conducted from the experiment result of the proposed method as follows. The "Garfield" is a cartoon image which has limited colors. There were 219,840 pixels categorized to

suitable pixels, and one bit message can be hidden in each pixel. The remainder 42,304 pixels were categorized to unsuitable. The embedding process for unsuitable pixels was carried out using the flags. There are 2^5 useless colors in the palette and will be used as flags. Namely, it can be embedded five bits messages in each unsuitable pixel. Overall, the image hides 431,360 bits messages, and the average value is 1.65 bits per pixel, as shown in Table 1.

"Joy" and "Lena" both are general photographs with various colors. In these two images, there are 119,737 and 131,140 pixels are unsuitable, and 2^3 and 2^2 flags in these two palette. The capacities are 1.91 and 1.5 bits per pixel shown as Table 1. We can find the capacity is determined on the number of flags and unsuitable pixels. When considering the file length, the loads of each byte in "Garfield", "Joy" and "Lena" are 1.41, 1.31 and 1, respectively. Thus, the data hiding capacity is increased by using the proposed method evidently.

Table 1: Hiding capacities by the proposed method

Image	Flag	Capacity	Langth	Bit/Pixel	Bit/Byte
Garfie	2^5	431,360	304,448	1.65	1.41
Joy	2^3	501,618	381,881	1.91	1.31
Lena	2^2	393,284	393,284	1.50	1

Compared with other data hiding methods, the 256 colors image "Fruit" was used to be the cover image. It is a 256×256 pixel image. Using the proposed method and other methods to hide secret message, the experimental results are shown in Table 2.

Table 2: Comparison results

Method	EZ stego	Fridrich'	Wu'	Proposed
Bit per pixel	1.0	1.0	1.0	1.75
Bit per byte	1.0	1.0	1.0	1.27
RMS	21.97	7.78	2.22	0
Palette content	no change	no change	change	No change

The capacities of EZ-stego, Fridrich and Wu's methods all are one bit per pixel/byte [9]. In our method, the capacity is 1.75/1.27 per pixel/byte. It is better than the other schemes. We also compared the recovered image quality of these methods with root mean square (RMS) error between the original image and recovered image. Our method can reverse the original image, and certainly gets zero RMS errors.

In Tables 1 and 2, we did not compare with Tzeng and Liu's method. The reason is because in Tzeng's method, the secret messages were hidden in the pixels filtered out unsuitable parts. Therefore, many pixels can not be used. It only 0.2386 bit was embedded in one pixel, less than all methods shown in Table 2. In Liu's method, even though the stego-image is distortion-free, the palette index duplicated several times, and then the palette image does not only cause suspicion but also goes against compression purposes.

4. Conclusions

In this paper, the proposed technique provides four advantages as follows. (1) Information is embedded into compression code. The message will not be destroyed by compression processes. (2) The capacity of information hiding is higher than other schemes significantly. (3) The palette image is reversible and deals with the image distortion problem. (4) The stego-image could be compressed again by lossless compression method to save the storage spaces and speed up the transmission rate.

References:

[1] C.C. Chang and C. Y. Chen, "Gray Code as a De-clustering Scheme for Concurrent Disk Retrieval," *Information Science and Engineering*, Vol. 3 (2), pp. 177-188, 1987.

[2] W. J. Chen and S. C. Tai, "The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS," *IEEE Transactions on Image Processing*, pp. 1309–1324, 2000.

[3] J. Fridrich, "A New Steganographic Method for Palette Based Images," *IS&T PICS*, Vol. 25-28, pp. 285-289, 1999.

[4] I. S. Hsieh and K. Ch. Fan, "An Adaptive Clustering Algorithm for Color Quantization," *Pattern Recognition Letters*, pp. 337-346, 2004.

[5] H. Liu, Z. Zhang, J. Huang, X. Huang and Y. Q. Shi, "A High Capacity Distortion-free Data Hiding Algorithm for Palette Image," *Proc. of the ISCAS*, Vol. 2, pp. II-916-919, 2003.

[6] Machado, R., "EZ Stego, Stego Online, Stego," *Available at http://www. stego. com*, 1997.

[7] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, "Information Hiding-A Survey," *Proc. of the IEEE Special Issue on Protection of Multimedia Content*, Vol. 87 (7), pp. 1062-1078, 1999.

[8] C. H. Tzeng, Z. F. Yang and W. H. Tsai, "Adaptive Data Hiding in Palette Images by Color Ordering and Mapping with Security Protection," *IEEE Transactions on Communications*, Vol. 52 (5), 2004.

[9] M. Y. Wu, Y. K. Ho and J. H. Lee, "An Iterative Method of Palette-based Image Steganography," *Pattern Recognition Letters*, Vol. 25, pp.301-309, 2004.

[10] Y. H. Yu, C. C. Chang and Y. C. Hu, "Hiding Secret Data in Images via Predictive Coding," *Pattern Recognition*, Vol. 38, pp. 691-705, 2005.