Lossless Information Hiding Scheme Based on Neighboring Correlation

Chin-Chen Chang
Department of Information Engineering and Computer Science,
Feng Chia University, Taichung, Taiwan, 40724, R.O.C.
ccc@cs.ccu.edu.tw

Tzu-Chuen Lu*
Department of Information Management,
Chaoyang University of Technology, Taichung 41349, Taiwan, R.O.C.
tclu@cyut.edu.tw

Abstract

A lossless information hiding scheme based on neighboring correlation is proposed in this paper. The scheme hides secret information into gray-level images according to the correlation between neighboring pixels. The performance of the proposed scheme has been evaluated in terms of payload capacity, image distortion, as well as embedding rate, and the experimental results show that the proposed lossless information scheme is capable of providing a great payload capacity without causing noticeable distortion, and the embedding rate is quite high. In addition, the proposed scheme can completely restore the host image after secret data extraction.

1. Introduction

Celik et al. proposed a lossless information hiding scheme in 2002 [4]. They applied the generalized least significant bit embedding (G-LSB) scheme to hide the secret information and the residual values used to restore the original cover image in the LSBs of the pixels [2, 3, 4]. After that, Tian extended the G-LSB scheme with difference expansion and proposed a low distortion lossless information hiding algorithm in 2003 [8]. In Tian's scheme, two adjacent pixels can be used to embed one secret bit. Hence, the maximum payload capacity of Tian's scheme is approximately merely 0.5 bpp (bit per pixel).

This paper proposes a high payload capacity lossless information hiding scheme which exploits the correlation of the neighboring pixels. In the proposed scheme, any two neighboring pixels can be used to conceal one secret data bit. The payload capacity of the proposed scheme is twice that of Tian's scheme. In addition, the slight distortion caused by the embedding is acceptable. Further, after extraction, the proposed scheme can completely restore the original cover image.

2. Related Works

Many lossless information hiding approaches have been proposed [3, 4, 8, 10]. In 2002, Fridrich *et al.* proposed a lossless information hiding scheme called the RS-embedding scheme [5]. In their scheme, an image is divided into several disjoint groups. Let $G = (p_1, p_2, \ldots, p_k)$ be a group of k adjacent pixels, where p_i is the i-th pixel in the group. Next, they use a flipping function and a discrimination function to categorize G into three sets: Regular (R), Singular (S), and Unusable (U). The pixel groups categorized to the R set or the S set can be used to conceal one secret bit.

Let $F(\cdot)$ be the flipping function used to permute the pixels such that $F(F(x)) = x$, where x is the pixel value. Let $f(\cdot)$ be the discrimination function defined as

$$f(p_1, p_2, \ldots, p_k) = \sum_{i=1}^{k-1} | p_{i+1} - p_i | \quad (1)$$

to calculate the corresponding real number for each G. Let α be the categorization function defined as

$$\alpha(G) = \begin{cases} R, & \text{if } f(F(G)) > f(G), \\ S, & \text{if } f(F(G)) < f(G), \\ U, & \text{otherwise.} \end{cases} \quad (2)$$

Assume the message type of R is 1 and that of S is 0. If the secret bit does not match the message type of $\alpha(G)$, then $F(\cdot)$ is used to flip G into another set. To enable the system to restore the original pixel values, the two sets R and S are compressed by a lossless compression algorithm and concatenated with the secret data to superimpose on the cover image. Hence, the actual payload capacity of the RS-Embedding scheme is

$$Cap_{RS} = \|R\| + \|S\| - \|Com_{RS}\|, \quad (3)$$

where $\|R\|$ and $\|S\|$ are the numbers of pixel groups in the R set and the S set, respectively, and Com_{RS} is the compressed bit-stream of the R set and the S set together, and $\|Com_{RS}\|$ is the length of Com_{RS}.

In 2003, Tian proposed a high payload Type-II lossless information hiding scheme based on difference expansion. In their scheme a threshold T is set to control the distortion between the cover image and the embedded image. In Tian's scheme, two neighboring pixels can be used to conceal one secret bit such that the raw capacity is $\frac{M \times N}{2} - \|C_0\|$, where $\|C_0\|$ is the number of non-changeable pixel pairs. Similar to the two sets R and S of the RS-Embedding scheme, the location map C is losslessly compressed and concatenated with the secret data to superimpose on the cover image. Hence, the actual payload capacity of Tian's scheme is

$$Cap_{Tian} = \frac{M \times N}{2} - \|C_0\| - \|Com_C\|, \quad (4)$$

where $\|Com_C\|$ is the length of the compressed bit-stream of C.

In this paper, what we are to propose is a lossless information hiding scheme that exploits the correlation of the adjacent pixels. The capacity of the proposed scheme is greater than those of Fridrich's scheme and Tian's scheme.

3. The Proposed Scheme

In this section, we will present the proposed lossless information hiding scheme. The diagram of the scheme, which is composed of the hiding process and the extraction process, is shown in Fig. 1. The hiding process is used to hide the secret data in the cover image, while the extraction process is used to extract the secret data and restore the cover image.

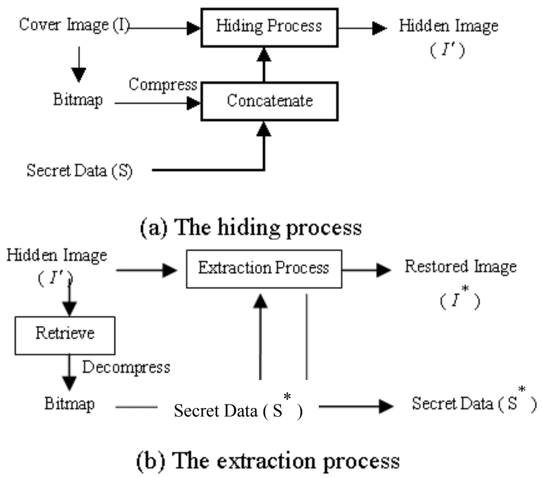

Figure 1. Diagram of the proposed scheme

3.1. The Hiding Process

Let $I = \{I_1, I_2, \ldots, I_{M\times N}\}$ be a cover image whose size is $M \times N$, where $I_i \in [0,255]$, and let $H = (h_1, h_2, \ldots, h_m)$ be the secret data to be embedded in the cover image. In order to increase the secrecy of the proposed scheme, we generate a secret sequence $K = (k_1, k_2, \ldots, k_m)$ accompanied by a secret key to manipulate it. Then the secret data is calculated by using

$$S = H \oplus K = (s_1, s_2, \ldots, s_m) = (h_1 \oplus k_1, h_2 \oplus k_2, \ldots, h_m \oplus k_m) \cdot \tag{5}$$

In order to achieve lossless information hiding, we use a bitmap $C = \{c_1, c_2, \ldots, c_{M\times N}\}$ to record whether the pixel can be used to hide information or not, where c_i is given as

$$c_i \begin{cases} 1, & \text{if } |I_i - I_{i+1}| \le T \text{ and } I_{i+1} + 2\times T \le 255, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

In the equation above, T is a predefined threshold used to control the distortion between the cover image and the hidden image. The bitmap C is then compressed by using a lossless compression method and is concatenated with the secret data S to superimpose on I. The hiding process can be described as follows:

Step 1: Calculate the correlation of the neighboring pixels to determine whether the pixels can be used to hide information or not. In our new method, the correlation is indicated by the interval between the adjacent pixels. If the pixels are judged able to be used to hide information, then the pixels are called changeable.

Step 2: For the changeable pixel pair, if the secret bit is 1, then increase the interval between the adjacent pixels for information hiding according to the value of the secret bit. Otherwise, if the secret bit is 0, the scheme retains the original pixel values.

The payload capacity of the proposed scheme is given by

$$Cap_{proposed} = M \times N - \|Com_C\|, \tag{7}$$

where $\|Com_C\|$ is the length of the compressed bit-stream of C.

3.2. The Extraction Process

In this subsection, we shall describe the extraction process. The following extraction procedure is used to extract the secret data and restore the cover image. The extraction process can be described as follows:

Step 1: Determine whether the pixel is changeable or not.
Step 2: If the pixel is changeable, then figure out the difference between the pixel and its adjacent pixel. If the difference is higher than the predefined threshold, then the hidden secret bit is 1; otherwise, the hidden secret bit is 0.
Step 3: Restore the cover image. If the pixel is changeable, then decrease the pixel value according to the secret bit.

The extraction algorithm for this process is presented in Algorithm 2. Since the receiver owns the secret key used to generate the secret sequence $K = (k_1, k_2, \ldots, k_m)$, the original secret data can be calculated by using
$H = S^* \oplus K = (h_1, h_2, \ldots, h_m) = (s_1^* \oplus k_1, s_2^* \oplus k_2, \ldots, s_m^* \oplus k_m)$.

3.3. The Compression Process

For most lossless information hiding systems, the key to higher real payload capacity is the efficiency of the lossless compression mechanism used to compress the residual information for restoring the cover image. Many lossless compression methods have been proposed, such as run-length coding, arithmetic coding [1, 11], Huffman coding, Lemple-Ziv coding and its variants (e.g. LZ77, LZ78, LZW, LZFG), and so on [7]. Among them, one of the popular approaches is arithmetic coding. The arithmetic coding method is suitable for lossless compression of bit-streams. In addition, unlike the Huffman trees in Huffman coding, it does not need to maintain complex data structures. Because of its efficient compression performance, the arithmetic coding method has drawn great attention, and many improved versions of it have been proposed, one of which is based on an adaptive finite-context statistical modeling technique - prediction by partial matching (PPM) [1]. In our new scheme, we shall adopt the arithmetic coding method based on PPM to compress the bitmap.

4. Experimental Results and Performance Evaluation

Experiments were carried out to evaluate the performance of the proposed lossless information hiding scheme. The proposed scheme was tested on a Pentium V 2.8 GHz personal computer with 512 Ram. Four 512×512 gray-level images were used as the cover images.

4.1. Compression Performance

In order to evaluate the compression performance of the arithmetic coding method we use, the first experiment was to compare on the compression rate among run-length coding, LZ77, LZ78, LZW, Huffman coding, and arithmetic coding based on PPM. The test image was Lena (512×512). Table 1 shows the bitmap comparison results with different T values on Lena by using the proposed scheme, where T stands for the threshold. It is obvious that arithmetic coding did outperform the other methods in the experiment.

4.2. Capacity-Distortion Performance

In this subsection, we compare the proposed scheme with the RS-embedding scheme and Tian's scheme. The peak signal to noise ratio (PSNR) is used to judge the quality of the hidden image, and the bit per pixel (bpp) is used to evaluate the real payload capacity. The results of how Tian's scheme, RS-embedding, and the proposed scheme compare on images Airplane, Barbara, Baboon and Lena are shown in Figures 2, 3, 4 and 5, respectively. According to the experimental results, the proposed scheme outperforms the other two.

Table 2 shows the required operations by the RS-embedding, Tian's scheme, and the proposed scheme for hiding a secret bit. Apparently, the computational effort made by the proposed scheme is less than those of the other two. In addition, only the sender and receiver own the secret key to generate the secret sequence for decoding the original secret data. That is, even if a third party who knows the hiding and extraction algorithms and can follow the procedures to extract the hidden message, he still cannot decode the original secret data correctly.

Table 1. Compression results among various lossless compression methods on Lena

T	RLC	LZ77	LZ78	LZW	Huffman Coding	Arithmetic coding
3	35,897	33,430	33,440	33,612	85,875	39,704
6	30,186	30,740	30,960	31,158	62,678	27,544
9	27,409	27,932	27,463	27,921	50,768	21,752
12	24,663	25,001	24,349	25,382	43,415	18,488
15	22,247	22,635	22,171	23,586	38,404	15,864
18	20,683	20,910	20,604	22,278	34,866	13,728

5. Conclusions

In this paper, we have proposed a simple and efficient lossless information hiding scheme based on neighboring correlation for gray level images. For T = 3, T = 6, T = 9, T = 12, T = 15, and T = 18, the average actual payload capacities of the proposed scheme are 0.37, 0.58, 0.68, 0.74, 0.78, and 0.81 (bpp), respectively. The PSNR averages are 37.2, 30.53, 26.70, 24.04, 21.98, and 20.32 (dB), respectively. The proposed scheme can not only conceal a large amount of secret information in the cover image, but also restore the cover image from the hidden image without any loss. The experimental results have shown that the proposed scheme can indeed outperform the RS-embedding scheme and Tian's scheme. So we conclude that the proposed scheme is practical in many real applications.

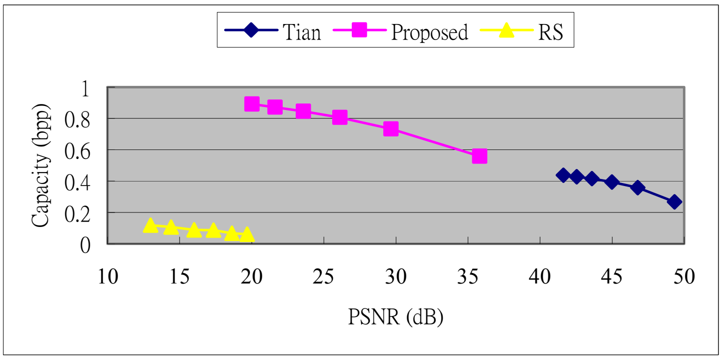

Figure 2. Capacity-distortion performance of Airplane

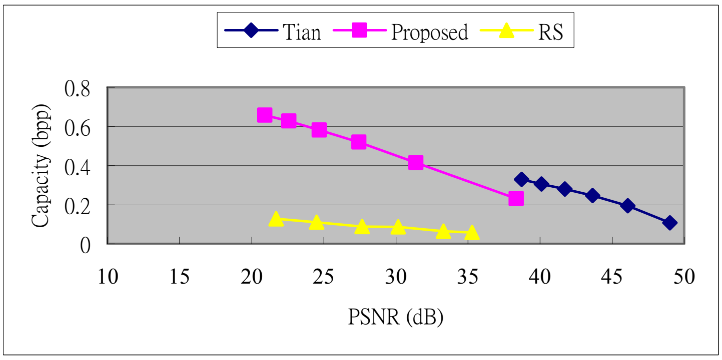

Figure 3. Capacity-distortion performance of Barbara

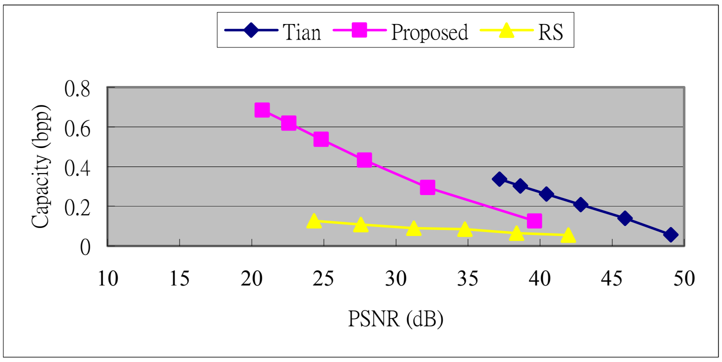

Figure 4. Capacity-distortion performance of Baboon

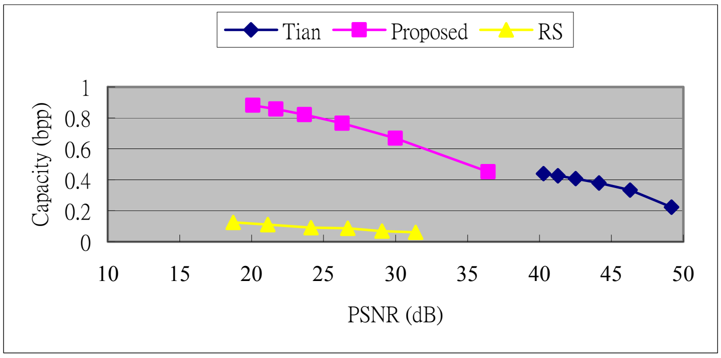

Figure 5. Capacity-distortion performance of Lena

Table 2. Required operations for hiding a secret bit

Operations	RS-embedding	Tian's scheme	The proposed scheme
Addition	k	6	2
Subtraction	k	3	1
If	3	2	2
Multiplication	0	2	2
Divide	0	5	0

6. References

[1] T.C. Bell, J.G. Cleary, and I.H. Witten, "Data Compression Using Adaptive Coding and Partial String Matching", *IEEE Transactions on Communications*, Vol. 32, No. 4, 1984, pp. 396-402.
[2] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Localized Lossless Authentication Watermark", *Proceedings of SPIE: Security and Watermarking of Multimedia Contents IV*, Vol. 5020, No. 70, 2003, pp. 689-698.
[3] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Lossless Generalized-LSB Data Embedding", Vol. 14, No. 2, 2005, pp. 253-266.
[4] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Reversible Data Hiding", *Proceedings on IEEE International Conference on Image Processing*, Vol. II, 2002, pp. 157-160.
[5] J. Fridrich, M. Goljan, and R. Du, "Lossless Data Embedding – New Paradigm in Digital Watermarking", *EURASIP Journal on Applied Signal Processing*, Vol. 2002, No. 2, 2002, pp. 185-196.
[6] J. Huang and Y.Q. Shi, "Reliable Information Bit Hiding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 10, 2002, pp. 916-920.
[7] K. Sayood, Introduction to Data Compression, 2nd Edition, Morgan Kaufmann, 2000.
[8] J. Tian, "Reversible Data Embedding and Content Authentication Using Difference Expansion", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 8, 2003, pp. 831-841.
[9] Y.C. Tseng and H.K. Pan, "Data Hiding in 2-color Images", *IEEE Transactions on Computers*, Vol. 51, No. 7, 2002, pp. 873-878.
[10] C.D. Vleeschouwer, J.F. Delaigle, and B. Macq, "Circular Interpretation of Bijective Transformation in Lossless Watermarking for Media Asset Management", *IEEE Transactions on Multimedia*, Vol. 5, No. 1, 2003, pp. 97-105.
[11] I.H. Witten, M. Radford, and J.G. Cleary, "Arithmetic Coding for Data Compression", *Communication of ACM*, Vol. 30, No. 6, 1987, pp. 520-540.
[12] M. Wu, "Joint Security and Robustness Enhancement for Quantization Based Data Embedding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 8, 2003, pp. 831-841.
[13] M.M. Yeung and F. Mintzer, "An Invisible Watermarking Technique for Image Verification", *Proceedings of IEEE International Conference on Image Processing*, Santa Barbara, CA, Vol. 2, 1997, pp. 680-683.

Authors

T. C. Lu received the B.M. degree (1999) and MSIM degree (2001) in Information Management from Chaoyang University of Technology, Taiwan. She received her Ph.D. degree (2006) in Computer Engineering from National Chung Cheng University. Her current title is Assistant Professor in Department of Information Management from Chaoyang University of Technology.

Professor C.C. Chang was born in Taichung, Taiwan on Nov. 12th, 1954. He obtained his Ph.D. degree in computer engineering from National Chiao Tung University. He's first degree is Bachelor of Science in Applied Mathematics and master degree is Master of Science in computer and decision sciences. Both were awarded in National Tsing Hua University. Dr. Chang served in National Chung Cheng University from 1989 to 2005. His current title is Chair Professor in Department of Information Engineering and Computer Science, Feng Chia University, from Feb. 2005.