High-Performance Steganographic Method Using Modulus Operation and Human Visual Model

Jiang-Lung Liu
Department of Electrical Engineering
Chung Cheng Institute of Technology
National Defense University
jlliu@ccit.edu.tw

Abstract

A secure steganographic method should basically possess the property of imperceptibility. The modulus-based steganographic methods embed a large amount of the secret message in the *k*-bit LSB of the cover-image for high-capacity applications. In this paper, we propose a method to generate adaptive modulus for each pixel of the cover-image. We first propose a generic model for constructing secure modulus-based steganographic methods. Based on this model, a simple and efficient human visual model is proposed to compute the adaptive modulus for each pixel of the cover-image so that the theoretical imperceptibility can be ensured. Experimental results show that the proposed method can efficiently adapt to the human visual system and has good performances on imperceptibility and high-capacity.

Keywords: Steganography, human visual model, information hiding.

1. Introduction

Steganography is the art of secret communication. Unlike cryptography, where the goal is to secure communications from eavesdroppers, the purpose of steganography is to hide the very presence of communications from observers. Modern steganogaphic techniques use digital images, videos, sound files, and other computer files that contain perceptually irrelevant or redundant information as cover carriers to hide secret messages. Digital images are the commonly used ones because they potentially contain a large amount of visual redundancy. The images used to carry secret messages are usually referred to as *cover-images* or *host-images*. If an image is embedded with secret messages, it is called a *stego-image*. A steganographic technique should generally possess two important properties: imperceptibility and sufficient hiding-capacity. The property of imperceptibility means that the stego-image does not contain any detectable artifacts due to message embedding. This property is critical because a third party could use such artifacts as an indication that a secret message is present. In other words, a steganographic technique is considered useless if it can not provide imperceptibility. It is clear that the less the message we embed into the cover-image, the smaller the probability of introducing detectable artifacts by the embedding process. For some applications (e.g. embedding image data in a cover-image), people may need a steganographic technique which can provide large hiding-capacity. Several works have been proposed in the literature to develop such steganographic methods [1-3].

The most common implementation for high-capacity steganography is embedding the secret bits in the *k*-bit least significant bit (LSB) of pixels of the cover-image. This kind of methods is also called the *k*-LSB replacement methods or the *k-LSB methods* because they embed the secret message by directly replacing the *k*-bit LSB of each pixel with *k* secret bits. Simple *k*-LSB methods replace the *k*-bit LSB for each pixel with fixed *k* secret bits. For a natural image, the part of LSB is regarded as the redundant information of an image so that we can embed the secret message in this part without causing detectable artifacts. Generally speaking, we can use 3-bit LSB to securely embed our secret message because the lower 3 bit-planes of a natural image contain very little information about the image. It is possible using 4-bit LSB to hide the secret message if the embedding error can be properly controlled by certain process such as the *optimal pixel adjustment process* (OPAP) proposed in [2]. It does not imply that the simple 4-LSB methods with OPAP are always secure because they can make the smooth area of the cover-image very "dirty" and reveal the existence of the secret message. That is, for security reason, a secure steganographic technique should adapt to the human visual system (HVS) so that it can use adaptive number of LSB to embed the secret message in different parts of the cover-image. In this paper, we propose such an adaptive method to meet the requirements of imperceptibility and high-capacity simultaneously.

To completely describe the proposed method, the rest of this paper is organized as follows. In Section 2, we propose a generic model to describe the modulus-based steganographic methods. Based on the proposed generic model, we proposed an adaptive steganographic method in Section 3. Several empirical results are demonstrated in Section 4 to show the effectiveness of the proposed method. The security of the proposed method is analyzed in Section 5. Section

6 concludes this work.

2. Generic Model for Modulus-Based Steganography

According to the famous Kerckhoff's principle [4], the security of a system should not depend on the concealment of its algorithm. In other words, a steganographic system should provide a theoretical security. In this section, we propose a generic model to provide such security for modulus-based steganography.

The generic model for modulus-based steganography consists of two phases: secret embedding and secret extraction. The procedure of secret embedding is shown in Figure 1. Let I be an 8-bit grayscale cover-image of size $x \times y$. The two-dimensional cover-image is first randomized using a secret key SK to obtain an 8-bit array I,

$$I = \{I_i \mid 1 \le i \le x \times y, I_i \in \{0,1,\cdots,255\}\}, \quad (1)$$

where I_i is the intensity of pixel P_i. Let m_i be the predefined modulus for pixel P_i. The original message M (may be a text or an image file) is first encrypted by a secure symmetric cryptographic system such as DES or AES [5] into the secret message S, $S = E_{SK}(M)$, where $E(\cdot)$ denotes the encryption function. The secret message S can be regarded as a bit-string with uniform distribution, which is then divided into n segments and can be represented as

$$S = \{S_i \mid 1 \le i \le n, S_i \in \{0,1,\cdots,2^{k_i}-1\}\}, \quad (2)$$

where k_i denotes the bit-length of ith segment of the secret message,

$$k_i = \lfloor \log_2 m_i \rfloor, \quad (3)$$

where $\lfloor x \rfloor$ denotes the floor operator which rounds x to the nearest integers towards minus infinity.

The n segments of the secret message are then embedded in the randomized version of the cover-image sequentially according to

$$I'_i = \begin{cases} I_i - (I_i \bmod 2^{k_i}) + S_i & \text{if } 1 \le i \le n, \\ I_i & \text{otherwise.} \end{cases} \quad (4)$$

where I' denotes the randomized version of the stego-image which is then de-randomized to a two-dimensional stego-image using the secret key SK.

Mathematically, we can extract each embedded segment of the secret message from the cover-image by computing

$$S_i = I'_i \bmod 2^{k_i}, \quad (5)$$

and concatenate all S_i to form the secret message S, $S = S_1 \| S_2 \| \cdots \| S_n$, where "||" denotes concatenation operator. The secret message can be further decrypted using the decryption function $D(\cdot)$ to obtain the original message M, $M = D_{SK}(S)$. The procedure of secret extraction is also shown in Figure 2.

In next section, we will propose a modulus-based steganographic method which can provide adaptive modulus for each pixel to meet the requirements of imperceptibility and high capacity.

3. The Proposed Method

The proposed modulus-based steganographic method is based on the model described in Section 2. To meet the requirements of imperceptibility and high capacity, the proposed method uses adaptive modulus 2^{k_i} to embed each secret segment S_i in the k_i-bit LSB of the pixel P_i. To achieve this purpose, the cover-image is first divided into 8×8 non-overlapping blocks. The value of k_i for each pixel P_i in the same block b is given the same value k_b which is derived from the characteristics of the human visual system. All the k_b's comprise the table T_k, i.e., there are $x/8 \times y/8$ elements in the table T_k. Therefore, the size of the table T_k is 1/64 the size of the cover-image and can be efficiently compressed and embedded into the cover-image for secret extraction. One of the contributions of this paper is to provide simple and efficient methods to compute the adaptive k_b for each block b and compress the table T_k. The computation of k_b and compression of T_k are detailed in the rest of this section.

3.1 Adaptive Modulus Computation

As mentioned previously, the value of k_b defines the maximum number of LSB which can be used to embed secret message for each pixel in block b. In other words, the human visual system should be taken into account for computing the value of k_b. The human visual system has been broadly studied for perceptual coding of still images [6-19]. Many works have been proposed to estimate the perceptual redundancy inherent in a still image. Perceptual redundancy has been defined as the magnitude of the stimulus at which it becomes just visible or just invisible. The visibility threshold of a particular stimulus depends primarily on two factors. One is the average background luminance behind the pixel to be tested. The other is the spatial non-uniformity of the background luminance.

Many works have been proposed to model the influence of the first factor. Figure 3 plots the results according to the model proposed in [6]. We can find that, between the values of background luminance 64 and 255, the values of visibility thresholds are lower than 8. It implies that the secret message will be visible if it is embedded in the higher 5 bit-planes of

smooth regions of the cover-image. It also implies that, if the optimal pixel adjustment process is applied, we can use at least 3 bit-planes to embed our secret message without causing significant artifacts. Therefore, in the proposed method, the lower bound of the value of k_b is defined as 3.

The second factor reflects the fact that the reduction in the visibility of stimuli is caused by the increase in the spatial non-uniformity of the background luminance. Let X_{ij} be the value of pixels in block b, $1 \le i \le 8$ and $1 \le j \le 8$. A simple and efficient way to model this factor is using the local standard deviation which is defined as

$$\begin{aligned} Std(b) &= \sqrt{Var(b)}, \\ Var(b) &= \frac{1}{64}\sum_{i=1}^{8}\sum_{j=1}^{8}(X_{ij} - A_b)^2, \end{aligned} \tag{6}$$

where A_b denotes the mean of the block b.

Based on the human visual model mentioned above, we can compute k_b for each block b and construct the table T_k according to the following steps.

Step 1. Divide the cover-image into b non-overlapping blocks of size 8×8.

Step 2. Compute the value of the standard deviation $Std(b)$ for each block b according to (8).

Step 3. Compute the value of k_b for each block b according to

$$k_b = \min(\max(round(\log_2(Std(b)+1)),3),5), \tag{7}$$

where $round(x)$ is the integer round operation which rounds the x to the nearest integers, $\max(x,y)$ and $\min(x,y)$ are the operations that take the largest element and smallest element between x and y, respectively.

Step 4. Replace k_b in the table T_k with k_b' which is computed as

$$k_b' = round\left(\frac{1}{9}\left(k_b + \sum_{r=1}^{8} k_r\right)\right), \tag{8}$$

where k_r denote the value of the 8 neighborhoods around k_b. Figure 4 shows the 8 neighborhoods around k_b.

Step 3 generates the adaptive value of k_b based on considerations of the characteristics of the human visual system. First, according to the spatial non-uniformity of the background luminance, the value of k_b is obtained from computing the base 2 logarithm of the standard deviation of the block b. Second, according to the sensitivity of the background luminance, the value 3 is assigned to the lower bound of k_b. Moreover, an upper bound 5 is also defined in Step 3 to avoid influencing the structure of the higher 3 bit-planes of the cover-image. In short, there are three kinds of values for the output of k_b: 3, 4, or 5, which represent the adaptive moduli 2^3, 2^4, and 2^5, respectively.

With the help of the table T_k, the secret message can be embedded in the cover-image using the adaptive modulus. Note that for the pixels in the same block b, the modulus used to embed the secret message is the same (i.e., $m_i = 2^{k_b'}$). It implies that the table T_k needs to be embedded in the cover-image for secret extraction. The k_b' obtained in Step 4 is to ensure that the difference between two adjacent blocks can not exceed the value 1 so that the table T_k can be efficiently encoded using the techniques of differential pulse code modulation (DPCM) and entropy coding [20]. Moreover, Step 4 also ensures that the continuity of pixel values between two adjacent blocks can be maintained.

3.2 Modulus Table Encoding and Embedding

The modulus table T_k contains $x/8 \times y/8$ elements which are appropriately arranged in a continue order so that the difference between any two adjacent elements has one of three possible values: 0, 1, or -1. For a natural image, the smooth areas and the busy areas are always clustered and are separated by edges. It means that most of the differences between any two adjacent elements of T_k are 0's. Only those across the edges have the value 1 or -1. It is ideal for us to use entropy coding to encode the differences of two adjacent elements in T_k. Let the table T_d comprise the differences of any two adjacent elements in T_k. The elements of T_d can be obtained according to

$$T_d(i,j) = \begin{cases} T_k(i,j) & \text{if } i, j = 1, \\ T_k(i,j) - T_k(i-1,j) & \text{if } i \ne 1,\ j = 1, \\ T_k(i,j) - T_k(i,j-1) & \text{otherwise.} \end{cases} \tag{9}$$

where (i,j) denote the indices of the elements in both T_k and T_d.

As mentioned above, most of the values of the elements in T_d are 0, and the others are 1 or -1. Theoretically, those with values 1 or -1 are paired, i.e., the number of elements with value 1 is equal to that with value -1. Base on this fact, an optimal static Huffman table can be constructed as Table I to encode the symbols "0", "1", and "-1". In our experiments, the table T_d can be efficiently encoded using Huffman encoding technique. That is, the Huffman-encoded result consumes ignorable capacity of the cover-image.

4. Experimental Results

To test the performance of the proposed method, several 8-bit standard grayscale images of size 512×512 were taken as the cover-images. Let C denote the average hiding capacity,

$$C = \frac{1}{x \times y} \sum_{i=1}^{x \times y} k_i \text{ bpp} . \quad (10)$$

The experiments are described in three parts. They are capacity performance, imperceptibility performance, and modulus table encoding performance.

4.1 Capacity Performance

We used the proposed method to compute the adaptive k_b for each 8×8 block of the test images. Because the test images are of size 512×512, there are 64×64 k_b's computed for each test image. The number of blocks for k_b =3, 4, and 5 are shown in Table 2. The last column of Table 2 also shows the average capacity for each test image. For most test images, the average capacities are between the ranges 3.5 to 4.0 bits per pixel (bpp). For noisy images such as "Mandrill," it is possible to obtain a capacity more than 4.0 bpp. Therefore, the proposed method is very adaptive to the human visual system.

4.2 Imperceptibility Performance

Without loss of generality, we generated enough random bits with uniform distribution to simulate the secret message and use the full capacity of the test images to embed the secret message. Take the image "Lena" as an example (the stego-image is shown in Figure 5(a)). It is clear that the secret message can be embedded in the 5-bit LSB of each pixel in the noisy blocks without causing visual distortion. Figure 5(b) shows the distribution of the used moduli, where black color represents the modulus 2^3 (k_b = 3), gray color represents the modulus 2^4 (k_b = 4), and white color represents the modulus 2^5 (k_b = 5). It should be noted that the modulus used in the smooth areas is 2^3. It is clear that the proposed method is adaptive to the human visual system and can provide good imperceptibility.

4.3 Modulus Table Encoding Performance

The proposed method uses equation (8) to reduce the difference between two adjacent blocks of the table T_k. We can find that there are no adjacent blocks with black color and white color in Figure 5(b). It ensures that the differences between two adjacent blocks only have three kinds of values, i.e., 0, 1, and -1, so that the difference table T_d can be efficiently encoded. Table 3 shows the results of using Huffman coding scheme to encode T_d. It is clear that the encoding results consume ignorable capacity of the cover-mages. Therefore, the modulus table can be efficiently encoded by the proposed method.

5. Security Analysis

As mentioned previously, a secure stegangraphic method should possess the property of imperceptibility. In other words, the secret message embedded in the cover image should not be detected by the human visual system. To meet this requirement, the proposed method incorporates a human visual model to deal with this problem. Based on the proposed human visual model, the adaptive moduli used for embedding the secret message can be appropriately defined to limit the maximum bits of LSB which can be disturbed for each pixel in the cover-image. Therefore, theoretically, the embedded secret message can be invisible using the proposed method.

Another security problem is that the attackers may try to accumulate the secret message segments to obtain the possible original secret message. The proposed generic model incorporates cryptographic systems to take care of this problem. Recall that before embedding, the cover-image is first randomized using a secret key *SK*. If an attacker wants to obtain the original secret message, he/she should know the exact order of the permuted cover-image. It is equivalent to breaking the secret key *SK*. It is shown that if the size of *SK* is 128 bits, the attacker should spend 5.4×10^{18} years to exhaust 2^{128} keys by using a system that can process 1 million keys per microsecond [5]. In other words, the incorporated cryptographic system provides a potential security for the proposed steganographic system.

6. Conclusions

In this paper, we proposed a generic model to describe the secure modulus-based steganography. We also described that a modulus-based steganographic method is insecure if it can not adapt to the human visual system. Based on the proposed model, an adaptive modulus-based steganographic method is also proposed to meet the requirements of imperceptibility and high-capacity. In our proposed method, a simple and efficient human visual model is used to compute the adaptive modulus for each pixel of the cover-image. A modulus table is appropriately generated so that it can be efficiently encoded and embedded in the cover-image for secret extraction. Experimental results show that the proposed method can efficiently adapt to the human visual system to meet both the requirements of high-capacity and imperceptibility for secure communication

References

[1] S.-J. Wang, "Steganography of capacity required using modulo operator for embedding secret image," Applied Mathematics and Computation, Vol. 164, No. 1, pp. 99-116, May 2005.

[2] C.-K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," Pattern Recognition, Vol. 37, No.3, pp. 469-474, Mar. 2004.

[3] D. C. Wu and W. H. Tsai, "Spatial-domain image

hiding using image differencing," IEE Proceedings - Vision, Image and Signal Processing, Vol. 147, No. 1, pp. 29-37, Feb. 2000.

[4] J. Seberry and J. Pieprzyk, Cryptography: An Introduction to Computer Security, New York: Prentice-Hall, 1989, p. 5.

[5] W. Stallings, Cryptography and Network Security: Principles and Practice, 3rd ed., New Jersey: Pearson Education, 2003.

[6] C.-H. Chou and Yun-Chin Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," IEEE Trans. Circuits and Systems for Video Technology, Vol. 5, No. 6, pp. 467-476, Dec. 1995.

[7] J. L. Mannos and D. J. Sakrison, "The effect of a visual fidelity criterion on the encoding of images," IEEE Trans. Information Theory, Vol. IT-20, No. 4, pp. 525-536, July 1974.

[8] C. F. Hall and E. L. Hall, "A nonlinear models for the spatial characteristics of the human visual system," IEEE Trans. System, Man and Cybernetics, Vol. SMC-7, No. 3, pp. 161-170, Mar. 1977.

[9] N. B. Nill, "A visual model weighted cosine transform for image compression and quality assessment," IEEE Trans. Communications, Vol. COM-33, No. 6, pp. 551-557, Jun. 1985.

[10] K. N. Ngan, K. S. Leonn, and H. Sinnh, "Adaptive cosine transform coding of images in perceptual domain," IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. 37 ,No. 11, pp. 1743-1749, Nov. 1989.

[11] D. L. McLaren and D. T. Nquyen, "Removal of subjective redundancy from DCT-coded images," IEE Proceedings - Communications, Speech and Vision, Vol. 138, No. 5, pp. 345-350, Oct. 1991.

[12] A. N. Netravali and B. Prasada, "Adaptive quantization of picture signals using spatial masking," Proceedings of the IEEE, Vol. 65, No. 4, pp. 536-548, Apr. 1977.

[13] J. O. Limb, "On the design of quantizer for DPCM coder: a functional relationship between visibility, probability and masking," IEEE Trans. Communications, Vol. COM-26, No. 5, pp. 573-578, May 1978.

[14] A. N. Netravali and B. G. Haskell, Digital Pictures: Representation and Compression, New York: Plenum, 1988.

[15] J. B. O. S. Martens and G. M. M. Majoor, "The perceptual relevance of scale-space image coding," Signal Processing, Vol. 17, No. 4, pp. 353-364, Aug. 1989.

[16] J. Pandel, "Variable bit-rate image sequence coding with adaptive quantization," Signal Processing: Image Communication, Vol. 3, No. 2-3, pp. 123-128, Jun. 1991.

[17] B. Gird, "Psychovisual aspects of image communication," Signal Processing, Vol. 28, No. 3, pp. 239-251, Sep. 1992.

[18] B. Giod, H. Aimer, L. Bingsson, B. Christensson, and P. Weiss, "A subjective evaluation of noise shaping quantization for adaptive intra-/interframe DPCM coding of color television signals," IEEE Trans. Communications, Vol. 36, No. 3, pp. 332-346, Mar. 1988.

[19] P. Pirsch, "Design of DPCM quantizers for video signals using subjective tests," IEEE Transactions on Communications, Vol. COM-29, No. 7, pp. 996-1000, July 1981.

[20] A. K. Jain, Fundamentals of Digital Image Processing, NJ: Prentice-Hall, 1989.

Table 1 The static huffman table used in the proposed method for encoding the difference table

Symbols	Code words
0	0
1	10
-1	11

Table 2 The hiding capacity for various test images.

Images	Number of blocks			Average (bpp)
	$k_b = 3$	$k_b = 4$	$k_b = 5$	
Lena	2365	1420	311	3.5
Mandrill	697	1262	2137	4.4
Goldhill	1598	2209	289	3.7
Boat	1726	1908	462	3.7
Barb	1363	1715	1018	3.9
F16	2068	1446	582	3.6

Table 3 The encoded results of the difference table.

Images	Encoded results (bits)				Capacity (%)
	0's	1's	-1's	Codes	
Lena	3320	387	389	4872	0.53
Mandrill	3504	294	298	4684	0.41
Goldhill	3488	283	314	4693	0.49
Boat	3480	323	293	4712	0.49
Barb	3272	394	430	4920	0.48
F16	3569	206	321	4623	0.48

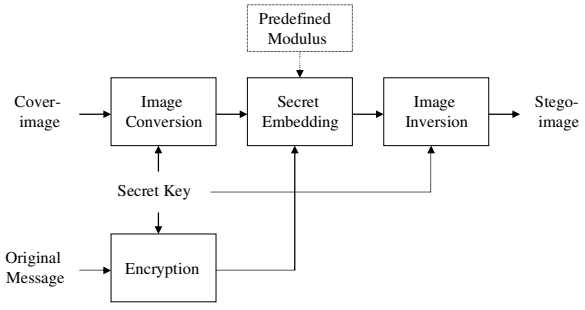

Figure 1 The embedding procedure of the proposed generic model for modulus-based steganography

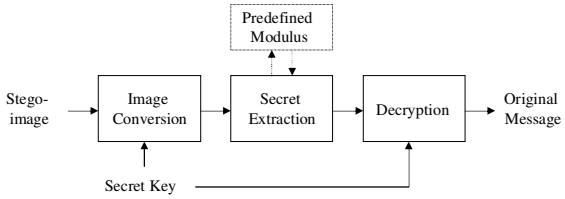

Figure 2 The extraction procedure of the proposed generic model for modulus-based steganography.

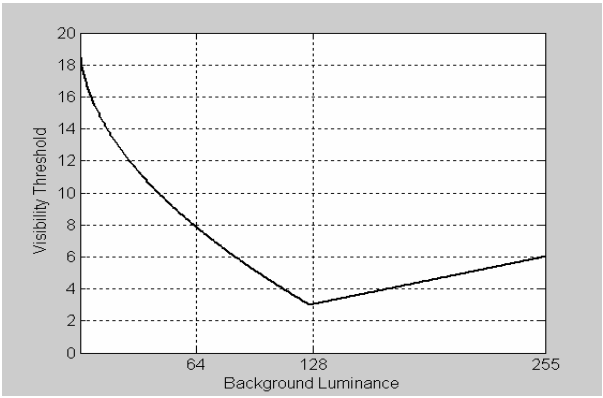

Figure 3 Visibility thresholds under different average background luminance modeled in [6].

k_1	k_2	k_3
k_4	k_b	k_5
k_6	k_8	k_7

Figure 4 The element k_b and its 8 surrounding neighborhoods

(a) (b)

Figure 5 (a) The stego-image created by the proposed method; (b) three-color presentation of the moduli used to create (a).