HIDING DATA IN VQ-COMPRESSED DOMAIN OF IMAGE

謝仕杰
國立虎尾科技大學資訊工程系
scshie@nfu.edu.tw

林信鋒
國立東華大學資訊工程學系
david@mail.ndhu.edu.tw

論文摘要

本篇論文提出了一種將資料隱藏於掩護影像之向量量化壓縮域的方法。為了提高資料的隱藏量以及維持掩護影像的位元率，所提出的方法應用了搜尋順序編碼(Search Order Coding)演算法來壓縮掩護影像的向量量化索引值。藉由所提出的方法藏入秘密資料後不會造成任何額外的影像失真。實驗結果顯示解碼端能夠在可接受位元率的情況下，同時解出秘密資訊以及掩護影像。

關鍵詞：向量量化壓縮、資訊隱藏、搜尋順序編碼演算法

Abstract

A data hiding scheme applied on the VQ-compressed domain of cover images is introduced in this article. To provide more hiding capacity for secret data and to keep an acceptable bit rate for the compressed cover images, the search-order-coding (SOC) algorithm was implemented to compress the VQ indices of cover images in the process of data hiding. During the process of data hiding, the proposed scheme adaptively embeds secret data into the compressed VQ indices of the cover image according to the amount of hidden data. In addition, the hiding process induces no extra coding distortion and adjusts the bit rate according to the size of secret data for the compressed cover image. Experiments show that the receiver can efficiently receive both the secret data and the compressed cover image with an acceptable bit rate simultaneously.

Keywords: Vector Quantization, Data Hiding, Search Order Coding.

1. Introduction

Along with the progress relating to computer hardware and software, the Internet has become the most popular channel for transmitting various forms of digital media. Since the environment of the Internet is open, the protection of digital information transmitted on the network has become an important research topic in recent years. Data hiding is a common technique to achieve the goal of data protection. It involves secretly embedding significant data into various forms of digital media such as text, audio, image and video [1]-[3]. With the rapid growth of network communication, data hiding techniques have been widely utilized in the applications of copyright protection, fingerprinting and secret communication [4]-[7].

The purpose of data hiding techniques is different from traditional cryptography [8]-[9] and watermarking techniques [10]-[12]. Cryptography encrypts meaningful messages into meaningless data while watermarking technique is utilized to protect the intellectual copyright of digital media. Data hiding technique covers secret information with the cover media as camouflage and is considered an extension of traditional cryptography. Hiding data in image involves embedding a large amount of secret data into a cover image with minimal perceptible degradation of image quality. However, the hiding capacity for secret data and the distortion of the cover image are a tradeoff since more hidden data always result in more degradation of visual quality on the cover image. Moreover, when data hiding technique is implemented on the compressed domain of image, the hiding capacity and the visual quality of cover images can be further restricted. Furthermore, the bit rate of the compressed cover image becomes another major consideration in the applications of data hiding. It should not cause apparent increase on the bit rate of the compressed cover images after the secret data have been hidden.

In the past few years, several vector quantization (VQ) based data hiding techniques have been proposed in the literature. Du *et al*. proposed a linguistic data hiding scheme to adaptively hide secret data into VQ-compressed cover images according to the size of secret data [5]. In Du *et al*.'s scheme, the secret data were hidden into the compressed cover images based on the following two phases. In the first phase, all the codewords in the codebook were rearranged into exclusive groups according to codeword's similarity. In the second phase, during the VQ encoding procedure, a codeword was chosen from the group that contains the best matched codeword of the currently encoded block to replace the best matched one itself based on the secret bits to be hidden. More details of this scheme can be found in [5]. To improve the performance of [5], Shie *et al.* proposed another VQ-based data hiding technique which takes advantages of the prediction property of side-match VQ (SMVQ) state codebook [6]. The major idea of this technique is to hide secret data into the VQ-compressed codes of cover image based on a modified encoding process of SMVQ and the concept of prediction. This technique provides better visual quality for cover images and lower

computational complexity than that of [5]. Chang *et al*. proposed a new data hiding scheme in which a limited amount of binary data can be hidden into the VQ-compressed codes of cover image [7]. The authors claimed that their scheme is the first one that technically and directly hides secret data into the VQ-compressed codes of cover image. In addition, they also announced that an acceptable compression ratio for the cover image is retained after the procedure of data hiding. In this scheme, a cover image is first compressed based on traditional VQ and an index table for the cover image is generated. After that, the search-order-coding (SOC) algorithm [13] is applied on the index table and a more compact index table is obtained. The compact index table consists of two kinds of compression codes: the search-order-coding (SOC) codes and the original-index-values (OIV) codes. Therefore, additional one bit has to be added in front of each SOC code and OIV code. The receiver can distinguish the two different codes according to the one-bit indicator. Chang *et al*. found that based on this characteristic, secret data can be hidden into the compression codes without inducing additional coding distortion. Specifically, the receiver determines that each bit of secret data is "0" or "1" based on whether the received compression code is SOC or OIV code. In the data hiding procedure of Chang *et al*.'s scheme, four types of code translation have to be taken into consideration. Two of the four types induce some additional bits for the compression codes. More details of this scheme can be found in [7]. However, the hiding capacity for secret data is restricted since the maximum hiding capacity for each block of cover image is only one bit. Moreover, Chang *et al*.'s scheme could result in a serious increase in the bit rate of the compressed cover image after secret data hiding.

In this article, we propose an adaptive data hiding scheme focusing on the hiding capacity of cover image on the VQ-compressed domain. The major goal of this scheme is to hide secret data into the VQ-compressed codes of cover image such that the interceptors will not notice the existence of secret data. In addition, the proposed scheme keeps an acceptable bit rate for the compressed cover image after a large amount of secret data is hidden into the cover image. To design a low bit rate image data hiding scheme, we incorporate the well-known VQ technique [14] into our scheme to compress the cover image. Moreover, the VQ indices of cover image are further compressed based on the search-order-coding algorithm [13] to reduce the bit rate and increase the hiding capacity of the cover image. To prevent the interceptors from being aware of the existence of secret data, the data is hidden into the compression codes of cover image. The proposed scheme introduces no extra coding distortion for the cover image after data hiding.

2. The Proposed Scheme

The goal of the proposed scheme is to secretly transmit a set of binary data via a VQ-compressed cover image at an acceptable bit rate. Assume that the cover image X is a gray-level image with $w \times h$ pixels. To compress the cover image with VQ, a codebook should be generated before image compression. Let the size of the VQ codebook be N_c and a codeword be composed of $m \times n$ elements. These N_c codewords of the codebook are generated based on the iterative LBG algorithm [14]. After a codebook is generated, X is partitioned into non-overlapping blocks of $m \times n$ pixels. Each image block of X is then encoded into a codeword index. Consequently, an index table T with $\lceil w/m \rceil \times \lceil h/n \rceil$ elements is constructed after all the image blocks of X are encoded by VQ. T is then ready to be transmitted to the receiver for the purpose of image compression. The bit rate BR_{VQ} of the ordinary VQ compression can be obtained by the following equation.

$$BR_{\mathrm{VQ}} = \log_2 N_c / (m \times n) \tag{1}$$

In order to simultaneously hide secret data into the index table T and reduce the bit rate of the VQ-compressed cover image, we incorporate the SOC algorithm into the proposed data hiding scheme. The SOC algorithm takes advantages of the high correlation among adjacent indices in T to encode the traditional VQ indices with fewer bits. Here, the high correlation means that there may be many image blocks of X which were encoded with the same VQ indices in the neighborhood. The SOC algorithm encodes each index of T one by one in raster scan order. It tries to find the same index around the current processed index in a predefined search path. If the same index is found within the search path, the current processed index will be denoted as an SOC code and replaced with a d-bits code (d is much smaller than $\log_2 N_c$). Otherwise, the currently processed index will remain unchanged and be denoted as an original index value (OIV) code. Note that a one-bit indicator is needed for each processed index in T to distinguish SOC codes from OIV codes. The performance of the SOC algorithm depends on the amount of SOC codes it can determine in the index table. Let the amount of SOC codes in T be s, therefore, the bit rate BR_{VQSOC} of the compressed cover image after applying the SOC algorithm can be computed by equation (2). The full description of the search-order-coding algorithm for compressing VQ indices of image can be found in [13]. Here, the SOC algorithm is summarized by the following steps.

$$\begin{aligned} BR_{\mathrm{VQSOC}} = \{&(1+\log_2 N_c) \times (\lceil w/m \rceil \times \lceil h/n \rceil - s) \\ &+ (1+d) \times s\} / (w \times h) \end{aligned} \tag{2}$$

Step 1. Define the number of bits d for encoding the search order.

Step 2. Input the next VQ index and use it as the

search center.

Step 3. Try to find a search point (SP) with the same VQ index value as the search center in the predefined search path on the VQ index table until the currently searched index is not a matched SP and can not be encoded with any one of the SOC codes, $(0)_2 \sim (2^d - 1)_2$.

Step 4. If a matched SP is found, the input index is encoded with a 1-bit indicator followed by the corresponding SOC code; otherwise, it is encoded with the indicator followed by its original VQ index value.

Step 5. If another VQ index has to be processed, go to *Step 2*; otherwise, output the compressed index table.

To secretly hide reasonable amount of binary data into the VQ-compressed codes of one cover image, we propose a novel idea to hide secret data while applying the SOC algorithm on *T*. Fig. 1 shows the data hiding procedure of the proposed scheme at the transmitter. In the proposed scheme, an unfixed or fixed amount of secret bits can be hidden into each SOC code in *T*. In order to hide a randomly unfixed amount of secret bits into an SOC code, a seed key *k* and an integer *i* are needed to randomly generate the amount ranging from 1 to *i*. However, if a fixed amount *p* of secret bits is hidden in each SOC code, the bit rate BR_{VQSOCDH} of the compressed cover image after data hiding can be calculated by equation (3). The embedding procedure of the proposed data hiding scheme is summarized by the following steps.

$$BR_{\text{VQSOCDH}} = \{(1+\log_2 N_c)\times(\lceil w/m \rceil \times \lceil h/n \rceil - s) + (1+d+p)\times s\} / (w \times h) \quad (3)$$

Step 1. Encode the cover image *X* into its corresponding index table *T* by applying the vector quantization algorithm.

Step 2. Apply the SOC algorithm on the index table *T*. If the current processed index is encoded as an SOC code, then extract a number of bits from the secret data and embed these secret bits into the SOC code. Otherwise, the currently processed index is encoded as an OIV code and no secret bit is embedded in the OIV code to preserve an acceptable bit rate for the compressed cover image.

Step 3. Encrypt and transmit the modified index table and the associated parameters to the receiver.

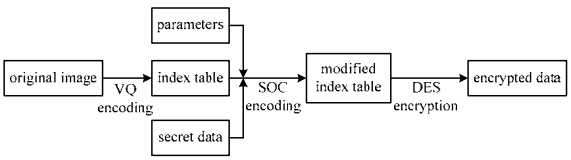

Fig. 1. Flow chart of the proposed data hiding procedure at transmitter.

A modified index table T^* for the compressed cover image *X* is obtained after the secret data has been embedded. To transmit the index table T^* to the receiver, all the SOC codes and OIV codes in T^* are merged into a binary bit stream based on the raster scan order. For more security, the bit stream can be further encrypted by the DES cryptosystem [15] to create an encrypted message. Finally, the encrypted data stream covering the secret data is generated. Note that the associated parameter set (*d*, *k* and *i* for variable length embedding or *d* and *p* for fixed length embedding) used in this scheme has to be preserved well for future use at the receiver. These parameters can be embedded into T^* or transmitted to the receiver via a secure channel.

The procedure of secret data extraction is quite simple. Fig. 2 shows the data extracting procedure of the proposed scheme at the receiver. To reconstruct the compressed codes of the cover image, the received data stream is first decrypted based on the DES decryption procedure. After the decryption process, the modified index table T^* of the compressed cover image *X* is directly obtained. The parameters used in the data hiding procedure can be extracted from T^* or obtained via a predefined secure channel. Moreover, the SOC codes in T^* can be easily specified with these parameters and the one-bit indicator that distinguishes SOC codes from OIV codes. Finally, all the secret bits are extracted and the secret data is then reconstructed.

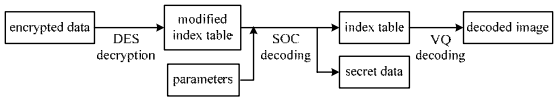

Fig. 2. Flow chart of the proposed data extracting procedure at receiver.

3. Simulation Results

In the computer experiments, we compare the proposed scheme with two earlier schemes to verify its performance. In the first experiment, the proposed scheme is performed on a set of six 256×256 images for comparing with Chang *et al.*'s scheme [7] which was introduced in 2004. In the second experiment, the proposed scheme is conducted on a set of two 512×512 images for comparing with Shie *et al.*'s scheme [6] which was introduced in 2006. All of the test images are with 256 gray levels per pixel. To simulate various types of secret data, the secret data used in our experiments are composed of randomly generated bit streams. The corresponding performance comparisons between the proposed scheme and the two earlier schemes are given in the following paragraphs.

To compare the proposed scheme with Chang *et al.*'s scheme [7], the six images, "Airplane", "Boat", "Girl", "Lena", "Peppers", and "Toys", are selected as cover images. The VQ codebook utilized in this experiment is generated from another five images,

"Airplane", "Boat", "Lena", "Sailboat" and "Toys", with size 512×512 pixels and 256 gray levels based on the LBG algorithm. In addition, the codeword size is 4×4 pixels and the codebook size is 256. Consequently, each cover image is partitioned into 4096 ((256×256)/(4×4)) image blocks with 4×4 pixels. The bit rate of the original VQ-compressed image is 0.5 bits per pixel (bpp). In order to hide secret data while keeping an acceptable bit rate for the compressed cover images, the SOC algorithm is applied on the VQ indices of cover image and the parameter *d* for the predefined search path is set to be 2 in our experiments.

To show the performance comparisons between the proposed scheme and [7], we apply equation (4) to evaluate the improvements of our scheme over Chang *et al.*'s scheme. Table 1 shows the improvements of our scheme on the hiding capacity of the compressed cover images. Note that in this experiment, the simulation environment is the same as [7]. In addition, the bit rates of the compressed cover images are adjusted to be almost the same as that of the experimental results in [7]. As shown in Table 1, the maximum improvement of our scheme over Chang *et al.*'s scheme is 146% on the cover image "Toys" whereas the minimum improvement is 55% on the cover image "Peppers". The proposed scheme outperforms Chang *et al.*'s scheme because the hiding capacity of the compressed cover images in [7] is limited by the number of image blocks partitioned from the cover image and there is no such restriction in our proposed scheme.

$$\text{Improvement} = \frac{(\text{hiding capacity of our scheme} - \text{hiding capacity of [7]})}{\text{hiding capacity of [7]}} \quad (4)$$

Table 1. Improvements of our scheme over [7].

Cover image	Airplane	Boat	Girl	Lena	Peppers	Toys
Improvement	116%	93%	74%	89%	55%	146%

Table 2 lists the bit rates of the compressed cover images hidden with different amounts of secret data, together with the experimental results of [7]. Note that the maximum hiding capacity of Chang *et al.*'s scheme is 4096 bits. However, as shown in Table 2, the secret data with size 16384 bits can be hidden into the compression codes of cover images by the proposed scheme. This is because the proposed scheme hides the secret data by technically embedding secret bits into the compression codes of cover image instead of modifying and translating these compression codes with other longer codes as proposed in [7]. Nevertheless, both of the proposed scheme and Chang *et al.*'s scheme decrease the compression rate of cover images after the secret data were hidden. As for the bit rate of compressed cover image in the proposed scheme, the increase of bit rate is steady with that of secret data size as illustrated in Table 2. In addition, the increase in the bit rate of compressed cover image is independent of the property of binary secret data. However, in Chang *et al.*'s scheme, the distribution of bits "0" and "1" in the secret data will greatly affect the bit rate of the compressed cover image. If most of the OIV codes and SOC codes in the VQ index table have to be translated into their corresponding longer codes, the increase in the bit rate can be up to 0.375 $((\log_2 N_c - d)/(w \times h))$ bpp. Such circumstances could result in a serious increase in the size of the compressed cover image after secret data hiding.

Table 2. Bit rates of the compressed cover images hidden with different amounts of secret data.

Size of secret data (in bits)	F16	Boat	Girl	Lena	Pepper	Toys
4096 in [7]	0.496	0.503	0.542	0.542	0.520	0.479
4096	0.423	0.444	0.494	0.483	0.479	0.394
8192	0.485	0.506	0.557	0.546	0.542	0.457
12288	0.548	0.569	0.619	0.608	0.604	0.519
16384	0.610	0.631	0.682	0.671	0.667	0.582

To compare the proposed scheme with Shie *et al.*'s scheme [6], the benchmark images, "Lena" and "F16", are utilized as cover images. The VQ codebook utilized in this experiment is generated from the cover images themselves. In addition, the codeword size is 4×4 pixels and the codebook size is 256. Consequently, each cover image is partitioned into 16384 ((512×512)/(4×4)) image blocks with 4×4 pixels. Note that this experimental environment is almost the same as [6], except for the codebook size. The bit rate of the original VQ compressed image is 0.5 bpp. In addition, the quality of cover images is evaluated by the peak signal-to-noise ratio (PSNR) criterion defined as equation (5).

$$\text{PSNR} = 10 \log_{10} \frac{E_{\max}^2 \times W_I \times H_I}{\sum \left(I_{m,n} - I'_{m,n}\right)^2} \quad (5)$$

where W_I and H_I are the width and height of cover image. $I_{m,n}$ is the original pixel value of the coordinate (m, n) and $I'_{m,n}$ is the altered pixel value of the coordinate (m, n). $E_{\max}$ is the largest energy of image pixels (e.g., $E_{\max}$ = 255 for 256 gray level images).

To show the performance comparisons between the proposed scheme and [6], we hide the same amounts of secret data into the compressed cover images. Before data hiding, the PSNR values of the

compressed cover images, “Lena” and “F16”, are 32.73 dB and 32.60 dB, respectively. For objective evaluation, Table 3 lists the performance comparison (in PSNR) of the proposed scheme and [6], with the amounts of hidden data ranging from 48 to 80 Kbits. Table 3 demonstrates that the proposed scheme outperforms [6] when the same amounts of data are hidden in the cover images. For subjective evaluation by the human visual system, Fig. 3 illustrates the original uncompressed cover image, the compressed cover images hidden with different amounts of secret data by [6], and the compressed cover image with secret data by the proposed scheme for test image “Lena”. Note that the proposed scheme induces no extra coding distortion on the compressed cover images. The experimental results also demonstrate that after hiding large amounts of secret data, the visual quality of cover images is quite acceptable based on the proposed scheme. To show the compression performance of the proposed scheme and [6], the bit rates with respect to the amounts of hidden data are listed in Table 4. Table 4 reveals that the compression performance of the proposed scheme is not as good as [6]. However, the increase in the bit rate is acceptable under the conditions that the visual quality is quite good and the hiding capacity is large for the compressed cover images.

Table 3. Performance comparison (in PSNR) of the proposed scheme and [6].

Size of secret data (in bits)	[6]		proposed	
	Lena	F16	Lena	F16
48 K	29.84	31.51	32.73	32.60
64 K	27.57	29.57	32.73	32.60
80 K	23.76	26.79	32.73	32.60

(a)

(b)

(c)

(d)

(e)

Fig. 3. Visual quality of cover image “Lena” by the proposed scheme and [6]. (a) the original uncompressed cover image, (b) the compressed cover image with 48K secret bits by [6], (c) the compressed cover image with 64K secret bits by [6], (d) the compressed cover image with 80K secret bits by [6], and (e) the compressed cover image by the proposed scheme.

Table 4. Performance comparison (in bits per pixel) of the proposed scheme and [6].

Capacity (in bits)	[6]		proposed	
	Lena	F16	Lena	F16
48 K	0.4632	0.4413	0.502	0.496
64 K	0.4358	0.4257	0.565	0.559
80 K	0.4195	0.4052	0.627	0.621

4. Conclusions

A data hiding scheme applied on the VQ-compressed domain of cover image has been introduced in this article. The proposed scheme provides greater hiding capacity than that of Chang *et al.*’s scheme at the same bit rates. In addition, this scheme provides much better visual quality for the compressed cover image than Shie *et al.*’s scheme. The receiver can efficiently receive both the hidden data and the compressed cover image simultaneously. The proposed scheme outperforms Chang *et al.*’s and Shie *et al.*’s schemes because of the following reasons. (i) The SOC algorithm is an efficient lossless compression technique for VQ index coding and it was applied in the proposed scheme to reduce the bit rate of cover image. (ii) The secret data are hidden by technically embedding the secret bits into the compression codes of cover image instead of modifying and translating these compression codes with other longer codes. (iii) The hiding capacity of cover image is not limited by the number of blocks partitioned from the cover image itself. And (iv) The proposed scheme does not introduce any visual distortion for the compressed cover image after secret data hiding. Therefore, we conclude that the proposed scheme is a feasible data hiding scheme.

References

[1] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding – a survey," Proc. of the IEEE, vol.87, no.7, pp.1062-1078, 1999.

[2] C. C. Chang, T. S. Chen, and L. Z. Chung, "A steganographic method based upon JPEG and quantization table modification," Information Sciences 141 (2002) 123-138.

[3] C. C. Chang, C. Y. Lin, and Y. Z. Wang, "New image steganographic methods using run-length approach," Information Sciences 176 (2006) 3393-3408.

[4] G. R. Feng, L. G. Jiang, and H. Chen, "Permutation-Based Semi-Fragile Watermark Scheme," IEICE Trans. Fundamentals of Elect., Commun. and Com. Sci., vol.E88-A, no.1, pp.374-377, 2005.

[5] W. C. Du and W. J. Hsu, "Adaptive data hiding based on VQ compressed images," IEE Proc. Vis. Image and Signal Process., vol.150, no.4, pp.233-238, Aug. 2003.

[6] S. C. Shie, S. D. Lin, and C. M. Fang, "Adaptive Data Hiding Based on SMVQ Prediction," IEICE Trans. Information and Systems, vol.E89-D, no.1, pp.358-362, 2006.

[7] C. C. Chang, G. M. Chen, and M. H. Lin, "Information hiding based on search-order coding for VQ indices," Pattern Recognition Lett., vol.25, pp.1253-1261, 2004.

[8] R. M. Davis, "The data encryption standard in perspective," Computer Security and the Data Encryption Standard, National Bureau of Standards Special Publication, February 1978.

[9] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," Communications of the ACM, vol.21, no.2, pp.120–126, 1978.

[10] M. Barni, C. I. Podilchuk, F. Bartolini, and E. J. Delp, "Watermark embedding: hiding a signal within a cover image," IEEE Communications Magazine, vol.39, no.8, pp.102–108, 2001.

[11] J. L. Liu, D. C. Lou, M. C. Chang and H. K. Tso, "A robust watermarking scheme using self-reference image," Computer Standards & Interfaces 28 (2006) 356-367.

[12] D. C. Lou, H. K. Tso and J. L. Liu, "A copyright protection scheme for digital images using visual cryptography technique," Computer Standards & Interfaces 29 (2007) 125-131.

[13] C. H. Hsieh and J. C. Tsai, "Lossless compression of VQ index with search-order coding," IEEE Trans. Image Processing, vol.5, no.11, pp.1579-1582, 1996.

[14] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun., vol.28, pp.84-95, 1980.

[15] National Bureau of Standards (U.S.), "DES Encryption Standard (DES)," Federal Information Processing Standards Publication, vol.46, National Technical Information Service, Springfield, VA, April 1997.