Signal Processing 90 (2010) 2141–2149

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

High capacity SMVQ-based hiding scheme using adaptive index

Chang-Chu Chen [a,b], Chin-Chen Chang [a,c,*]

[a] Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi 62102, Taiwan, ROC
[b] Department of Management Information System, Central Taiwan University of Science and Technology, Taichung 40601, Taiwan, ROC
[c] Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan, ROC

ARTICLE INFO

Article history:
Received 20 March 2009
Received in revised form
2 December 2009
Accepted 21 January 2010
Available online 8 February 2010

Keywords:
Secret hiding
Vector quantization
Side match vector quantization

ABSTRACT

This paper proposes a high-capacity image-hiding scheme based on an adaptive index. Data-hiding based on vector quantization (VQ) is a technique for hiding data in the VQ index code. Data-hiding based on side match vector quantization (SMVQ) has been proposed for improving the compression rate of VQ-based data-hiding schemes. However, the hiding capacity of an SMVQ-based data-hiding scheme is very low since, at most, only one secret bit is hidden in one index code. To overcome this drawback and increase the capacity, the proposed method uses an adaptive index to hide more bits in one index code. The weighted squared Euclidean distance (WSED) can also be used to increase the probability of SMVQ to get greater hiding capacity.

According to the experimental results, a higher hiding capacity was obtained and a good-quality embedded image was preserved in the adaptive index SMVQ-based data-hiding scheme. The hiding capacity of the proposed scheme was approximately twice that of relevant two hiding schemes.

1. Introduction

The Internet has become a popular channel for transmitting various data in digital form. However, two serious problems can occur with Internet communication. The first problem is secrecy, as some data, such as government documents, military photos, commercial contracts, and personal medical records, are private. The second problem is related to the flow of large transmissions. Huge documents are transmitted over the Internet every moment, which can overload the Internet and impede transmission.

A fast and safe transfer of secret data should combine compression and protection. Recently, some studies have combined the advantages of data-hiding and compression. An image watermarking technique based on VQ was proposed [12,8] in which the codebook is partitioned into two or three sub-codebooks and then, in the embedding phase, the scheme chooses a similar codeword from the proper sub-codebook according to the secret bit so the modified index of the codeword implies the secret bit. Jo and Kim's scheme [8] hides just one bit in one image block, while Chiang and Tsai [3] proposed a VQ-based scheme with higher hiding capacity using overlapping codebook partitions. To reduce the length of the VQ index code, Chang and Wu [2] proposed a data-hiding scheme based on side match vector quantization (SMVQ), where some codewords close to the image block are formed as a state codebook, and the order of the proper codewords in the state codebook is used to present the index of the image block. Since the code length of the order is less than that of the VQ index, the SMVQ-based compressed code is less than the VQ-based one. In order to improve hiding capacity, we propose a novel data-hiding scheme that uses variable-length SMVQ compressed code. In this scheme, each image block is processed by variable-length SMVQ-based hiding; then, if no proper codeword for

* Corresponding author at: Department of Information Engineering and Computer Science, Feng Chia University, 100 Wenhwa Rd., Seatwen, Taichung 40724, Taiwan, ROC. Tel.: +886 4 24517250x3790; fax: +886 4 27066495.

E-mail addresses: ccz@cs.ccu.edu.tw (C.-C. Chen), ccc@cs.ccu.edu.tw (C.-C. Chang).

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.01.018

replacement is found, improved VQ-based hiding is adopted to deal with the image block.

The rest of this paper is organized as follows. Section 2 briefly reviews work related to VQ-based and SMVQ-based compression and data-hiding methods. Section 3 presents the details of the proposed method. In order to show how the proposed method performs, Section 4 presents the experimental results and further discussions. Finally, conclusions are drawn in Section 5.

2. Related works

2.1. Vector quantization

Vector quantization [5,11] is a classical quantization technique from signal processing, also called "block quantization" or "pattern matching quantization," which was originally used for lossy data compression by using the index value to represent a signal vector. The process of VQ consists of three phases: codebook training, encoding, and decoding. In the first phase of VQ for images, the codebook training phase [1,10], a codebook that is the collection of image blocks created from training images, is developed. A good codebook can reduce the distortion between the original and the decoded images. The second phase of VQ, the encoding process, searches the codebook for the closest codeword for an image block, and the index of the codeword is used to represent the image block. The image is first divided into a series of small, non-overlapping rectangular blocks, each of which is represented as a k-dimensional vector $X=[x_1,x_2,\ldots,x_k]$. The object of encoding is to find from the codeword $Y_i=[y_{i1},y_{i2},\ldots,y_{ik}]$ that is most similar to image block X. In other words, this part of the process finds a codeword Y_i with a minimum distortion between Y_i and X. The squared Euclidean distance SED(X,Y_i) is usually used to measure the distortion between image block X and the codeword Y_i, which is expressed as

$$\mathrm{SED}(X,Y_i)=\|X-Y_i\|^2=\sum_{j=1}^{k}(x_j-y_{ij})^2. \tag{1}$$

Then the full searching (FS) process of VQ is performed to find the closest codeword Y_t in the codebook C with the minimum SED from X, calculated as $\mathrm{SED}(X,Y_t)=\min_{Y_i\in C}\mathrm{SED}(X,Y_i)$, where the codeword Y_t is called the winner, and t denotes the index. This index t, instead of X, is transmitted to the receiver as the compressed code.

In the third phase of VQ, the decoding process, the same codebook C is shared in advance with the receiver and is used for all images. When the index t is transmitted, the receiver uses the codebook C as the look-up table in order to retrieve the corresponding codeword Y_t for image block X.

Since each image block is encoded as the index, VQ is a fixed ratio compression method. The receiver retrieves an image that is only similar to the original image because the retrieved codeword Y_t is the codeword closest to image block X.

2.2. SMVQ

The main problem of VQ is that boundaries are visible between decoded image blocks. Side match vector quantization (SMVQ) [9] is a popular compression method for further improving VQ compression rate. In the SMVQ system, the image blocks in the first row and the first column are encoded by VQ, and the residual blocks are predicted using two neighboring blocks (i.e., the upper and the left adjacent blocks) of the current block in predicting its index value. SMVQ can improve the block effect of VQ as well as the compression performance of VQ. Consider the 4×4 image blocks shown in Fig. 1, where U and L are used to predict X. The pixel value in x_1 can be forecasted by computing the equation $x_1=(u_{13}+l_4)/2$. Other pixel values may be recovered by $x_2=u_4$, $x_3=u_{15}$, $x_4=u_{16}$, $x_5=l_8$, $x_9=l_{12}$, and $x_{13}=l_{16}$. After the pixel values are recovered, the closest codeword to image block X can be found by using these partially forecasted and recovered pixels. The distortion measurement between the codewords in the codebook C and the partially recovered image block X is calculated only by the differences in pixel positions 1, 2, 3, 4, 5, 9, and 13, as shown in Fig. 1.

The compression performance of SMVQ is better than that of VQ since the compressed codes are composed only of the block codes in the first row from the top and the first column from the left of the image; the other residual blocks can be restored by the neighboring and known blocks. However, the quality of the recovered image may be unacceptable as a result of the derailment problem [2]. A simple solution is to use a threshold to control the image quality of the decoded image. In threshold-based SMVQ, an indicator expresses whether the residual block is encoded by VQ or SMVQ.

2.3. VQ-based hiding schemes

Shieh et al. [12] presented an embedding method based on VQ by partitioning the codebook into two groups and selecting a codeword from the corresponding group index as a hidden secret bit. Jo and Kim [8] proposed a data-hiding scheme which used the same concept as

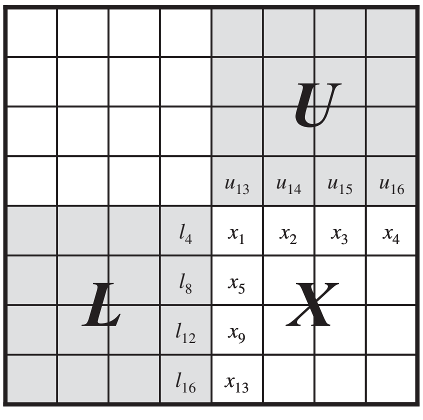

Fig. 1. Prediction process of SMVQ.