Steganalysis to Data Hiding of VQ Watermarking Upon Grouping Strategy

Ya-Ting Chang, Min-Hao Wu, Shiuh-Jeng Wang

▶ **To cite this version:**

Ya-Ting Chang, Min-Hao Wu, Shiuh-Jeng Wang. Steganalysis to Data Hiding of VQ Watermarking Upon Grouping Strategy. 2nd Information and Communication Technology - EurAsia Conference (ICT-EurAsia), Apr 2014, Bali, Indonesia. pp.633-642, 10.1007/978-3-642-55032-4_65 . hal-01397281

HAL Id: hal-01397281

https://inria.hal.science/hal-01397281

Submitted on 15 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Steganalysis to data hiding of VQ watermarking upon grouping strategy

Ya-Ting Chang [1], Min-Hao Wu[2], Shiuh-Jeng WANG *

[1,*] Department of Information Management, Central Police University,
Taoyuan, 33304, TAIWAN.
[2,] Department of Computer Science and information Engineering, National Central University,
Taoyuan, 32001, TAIWAN.
[1] bowrose@gmail.com
[2] mhwu@csie.ncu.edu.tw
* sjwang@mail.cpu.edu.tw

Abstract

This paper present a steganalysis method for the data hiding scheme based on VQ-compression. This data hiding algorithm divides the codebook into groups which contain two codewords each. The Euclidean Distance of the group is used instead of a codeword in traditional VQ-compression. A PoV-like effect of Chi-square attack is observed and used as a feature of detection. In the proposed steganalysis, we detect whether an unknown image is a VQ-compressed image or not, and then the target detection of codewords grouping type data hiding methods is proposed. We apply proposed scheme to Yang et al.'s watermarking scheme. A large amount test image database UCID (Uncompressed Colour Image Database) is utilized as various conditions, such as cover imaged, traditional VQ-compressed images, and stego images. The experimental shows that the proposed steganalysis method is able to identify the stego images among others, and the accuracy rate reaches over 90 %.

Keywords: Steganalysis, data hiding detection, vector quantization, group strategy

1 Introduction

Steganography is developed to guarantee the confidentiality of the information which is aimed to protect. One of the steganography method is to hide some information into a particular media, such as an image, a video clip, or an audio file. The algorithms are designed not to be discovered the existence of the information hidden inside.

On the opposite side, steganalysis is a novel research area which aims the opposite goal to steganography. Steganalysis is an art of identify the suspected items, in other words, detect the embedding information inside the carrier. The types of

steganalysis could be classified into blind steganalysis and targeted steganalysis according to objectives [7]. Blind steganalysis techniques detect the existence of secret messages when the steganography embedding algorithm is unknown. On the other hand, targeted Steganalysis are designed for a particular steganography algorithm.

Certain detection tools to the steganography in spatial domain are developed. One of the well-known attack to steganography, Chi-square attack [13], is proposed for LSB steganography technique [13][4]. Histogram-based and PVD-based steganography techniques are also attacked by analysis the statistical distribution of the pixel values' histogram [5][2]. But there is no research about detect the VQ-based steganography so far.

One of the VQ-based information hiding method is to re-encode the index value according to the secret bit to be embedded [10][15][6][14]. The targeted steganalysis of the method is proposed in this paper. The rest of this paper is organized as follows. Related background knowledge such as vector quantization, VQ-based steganography method, and Chi-square detection are introduced in Section 2. Our scheme is presented in Section 3. Section 4 shows the experimental results compared with related work. Conclusions are given in Section 5.

2 Related works

2.1 Vector quantization

Vector quantization (VQ) was proposed by Linde, Buzo, and Gray (LBG) in 1980 [3]. A codebook is used for VQ-compression generated by the LBG algorithm [3]. A codebook is composed of these featured blocks, or called codewords. The first step of VQ divides the cover image into several non-overlapping blocks and each block size is $m \times m$. Let k be $m \times m$, and $u_1, u_2, \ldots, u_k$ represent the pixel values of a block. The second step encodes the blocks by u_i and the codewords in codebook, for $1 \leqq i \leqq k$. If a codeword CW_x has size k, for a block in an input image, Euclidean distance is calculated by Eq. (1) comparing with every codewords in the codebook.

$$d(u, CW_x) = \sum_{j=1}^{k} (u_j - v_j)^2 , \tag{1}$$

where $v_1, v_2, \ldots, v_k$ represent the component values of the codeword CW_x which its index value is x.

For every non-overlapping blocks of the cover image, find the closest codeword with the smallest Euclidean distance. The corresponding index value of the nearest distance codeword will be the output while VQ encodes an image block. The output will be an index table. Fig. 1 shows the procedure of VQ image compression. The compressed image can be reconstructed from the index table by referencing the codewords.

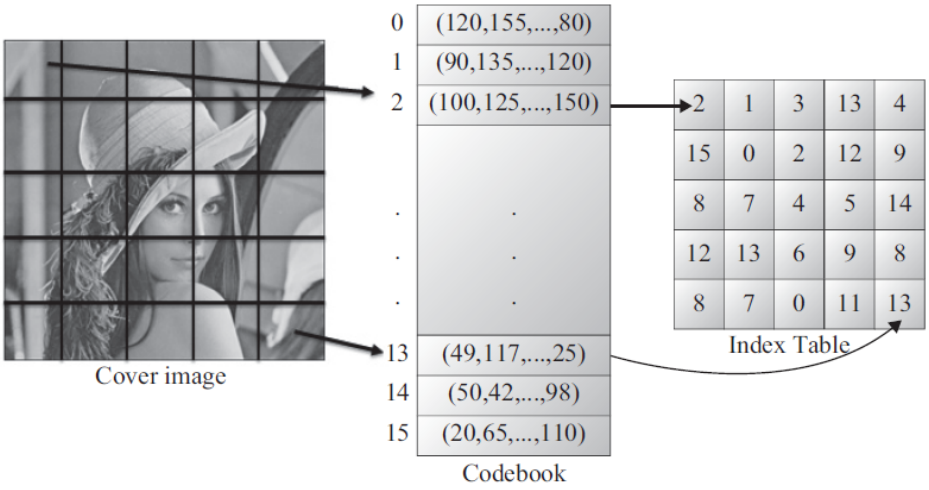

Fig. 1. The procedure diagram of VQ-compression.

2.2 Yang et al.'s data hiding method

Lu et al.'s method [6] and Wu-Chang's method [14] use codebook division to embed watermark bits. Yang et al. [15] enhanced the steganography of the algorithm of codebook division. The embedding method will be briefly presented as follows.

Codebook-sorting grouping: Before the embedding phase, a codebook is generated by the LBG algorithm [3], and the codebook division is processed. If a codebook size is *CS*, The division results are *CS*/2 groups G_x, where x is an integer from 1 to *CS*/2. For each codewords in the codebook, compute the *KEY* values by summing up the elements. For the codeword with the smallest *KEY* value, find another codeword matched to it such that their Euclidean distance is the least. The two codewords are labeled as CW_x and CW_x', forming the group G_x, The rest codewords in the codebook are divided as the same method above.

Shortest-group encoding: All the codewords are labeled as CW_x or CW_x', where x is from 1 to *CS*/2. While processing VQ-compression on a block *C*, compute the Euclidean Distance *D* between *C* and each group G_x. *D* is the summation of Euclidean Distance from *C* to all the elements in G_x. Extract one bit *b* from the secret message *S*. If G_y has the least *D*, CW_y and CW_y' are used instead of traditional VQ-compression. If *b* is '0', output the index value of CW_y; otherwise if *b* is '1', output the index value of CW_y'.

2.3 **Chi-square attack** [13]

Least significant bit (LSB) data hiding method is to embed secret messages by substitutions of the least significant bits of pixel values. Overwriting least significant bits transforms values into each other which only differ in the least significant bit. The frequencies of both values of each pairs of values (PoV), become equal. In a natural image, the least significant bits of all pixel values may not show close amount of '0' bits and '1' bits generally. But if the secret message is generated randomly with '0' and '1' bits, the appearance probabilities of '0' and '1' bits are equal. After embedding the secret message to a cover image using LSB data hiding method, the least significant bits of all the pixel values in stego image show the equal frequency of PoV.

3 Proposed steganalysis

An inference could be inducted from Chi-square attack. According to Yang's embedding method, an index of the codewords is chosen to embed a secret bit. In other words, if the group G_y is selected to embed a secret bit, either CW_y or CW_y' is chosen according to the value of the secret bit. In the presupposition that the secret message S is a bit-stream generated randomly, the composition of '0' bits and '1' bits should be similar amounts. That is, the frequency of chosen CW_x and CW_x' should be similar, too, where x is from 1 to CS/2.

The flowchart of proposed steganalysis method is shown in Fig. 2.

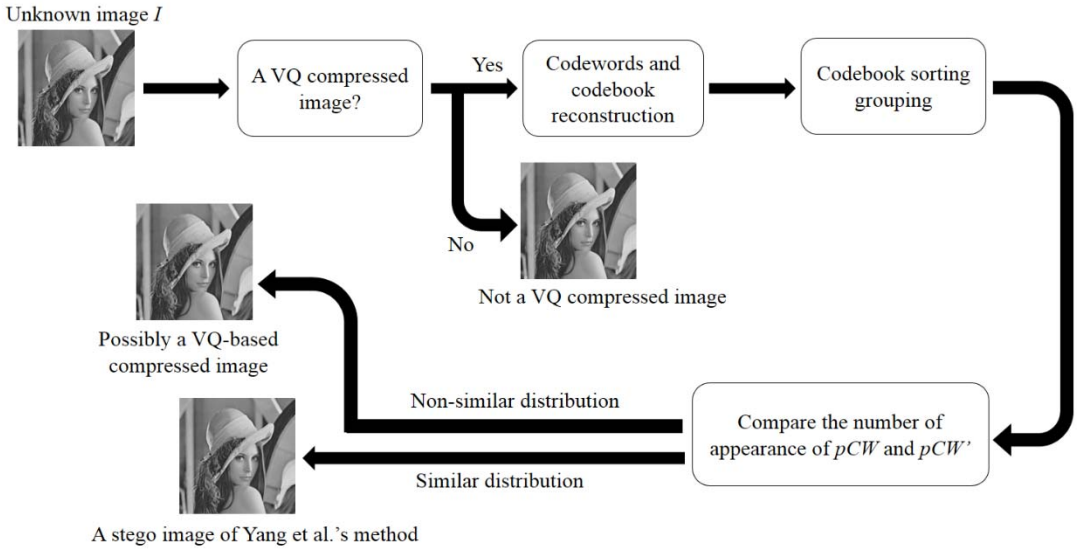

Fig. 2. The flowchart of proposed steganalysis method.

3.1 Detection of VQ-compression

The first step is to estimate an unknown image be a VQ-compression image or not. During the process of VQ-compression, all the non-overlapping blocks have the same size and are non-overlapping, we can inference that in a VQ-compression image, pixel values may have extent of diversity on the edge of non-overlapping blocks.

We first calculate the absolute difference values in vertical and horizontal directions. We can infer that a column or a row which has more intense changes of pixel values may be the edge of non-overlapping blocks of VQ-compression. We set a threshold T_F to be 1/3. If there are more than T_F of the total number of a column or a row have absolute difference values greater than its neighboring, then it is considered to be the possible edge of non-overlapping blocks. The position of the possible edges, such as i-th row or j-th column, are recorded, and their distances are calculated, dr_i and dc_j, for rows and columns, respectively. For example, if the possible edges are the 4-th, the 8-th, the 12-th, and the 20-th rows, then their edge distances are $dr_1 = 8 - 4 = 4$, $dr_2 = 12 - 8 = 4$, and $dr_3 = 20 - 12 = 8$.

If there are p edge distances between rows dr_1, dr_2,…, dr_p, and q edge distances between columns dc_1, dc_2,…, dc_q, the standard deviation of dr_i and dc_j, σ_R and σ_C, are calculated. In a natural image, σ_R and σ_C would be great values because pixel values should change smoothly in a small area and hence the "edges" do not exist. On the contrary, a small value of standard deviation represent higher centralization of the edge distances. If the non-overlapping blocks are sized 4×4, the ideal situation is that dr_i and dc_j would be all the same value 4, and σ_R and σ_C would be 0. Threshold T_s is set to be 5, and if σ_R or σ_C is greater than 5, then it is possible not compressed by VQ method.

If an unknown image is detected to be a VQ-compressed image by σ_R and σ_C, then the next step is to predict the possible non-overlapping block size. The most possible block size is the edge distance which shows up most frequently. The predicted block size $w \times h$ is defined as following Eq. (2):

$$\begin{aligned} w &= Mo(dc_j), \quad j = 1, 2, \ldots, q, \\ h &= Mo(dc_i), \quad i = 1, 2, \ldots, p, \end{aligned} \tag{2}$$

where "Mo" denotes the function which returns the mode, the value that appears most often.

3.2 Detection of stego images of Yang et al.'s method

We can reconstruct the codebook by copying the pixel values of every block. The blocks of the stego image are scanned from left to right and top to down, copied to the predicted codebook pCB. Pixel values of blocks form predicted codeword pCW_1, pCW_2, …, pCW_k, and the index values are 1, 2, …, k. The duplicated blocks would not create new codewords.

According to Chi-square attack, if the secret message is a random generated bit

stream, the number of least significant bits of a stego image of '0' bits and '1' bits should be similar. The same concept in Yang et al.'s shortest-group encoding method, the number of times choosing CW_x and CW_x' to encode should be similar for a stego image. Predicted codewords are input to Yang et al.'s codebook-sorting grouping. The output of applying Yang et al.'s codebook-sorting grouping method to *pCB* are groups pG_x which each one contains two predicted codewords, pCW_x and pCW_x'.

Correlation coefficient R is used as judgment. If the number of appearance of pCW_y and pCW_y' are close, then the value of R would approach to 1. On the other hand, if the number of appearance of pCW_y and pCW_y' vary a lot, then the value of R would approach to 0. A threshold T_R is set to be 0.7. If $R \le T_R$, the unknown image I is not a stego image using Yang et al.'s method. If $R > T_R$, then I is probably a stego image using Yang et al.'s method. If there are u groups of pG_x, the number of appearance of all pCW_x (pCW_1, pCW_1, …, pCW_u) are X_1, X_2, …, X_u, and the number of appearance of all pCW_x' (pCW_1', pCW_1', …, pCW_u') are Y_1, Y_2, …, Y_u, the correlation coefficient R is defined as Eq. (3):

$$R = \frac{\sum_{i=1}^{u}(X_i - \mu_X)(Y_i - \mu_Y)}{\sqrt{\sum_{i=1}^{u}(X_i - \mu_X)^2} \times \sqrt{\sum_{i=1}^{u}(Y_i - \mu_Y)^2}}, \tag{3}$$

where $\mu_X = \frac{1}{u}\sum_{i=1}^{u} X_i$ and $\mu_Y = \frac{1}{u}\sum_{i=1}^{u} Y_i$.

4 Experimental results

We implement our scheme using the software Marlab 2013a. Traditional VQ-compression and Yang et al.'s method are coded by following their algorithms. Different sizes of codebooks, 128, 256, 512, and 1024 are generated using the LBG algorithm [3]. The UCID (Uncompressed Colour Image Database) [17] provides a benchmark dataset for image retrieval, containing over 1,300 images. Several of them are shown in Fig. 3. The images are size 384×512 or 512×384, and are all converted to 8-bit gray level images. The non-overlapping blocks of VQ-compression and Yang et al.'s method are sized 4×4.Various conditions of images are inputted to our scheme:
Condition A: Original pure images without any processing.
Condition B: Images compressed by traditional VQ coding.
Condition C: Images applied Yang et al.'s method with full embedding.

The expected output results corresponding to the different conditions above are:
Result A: The unknown image is not a VQ-compressed image.
Result B: The unknown image may be a VQ-compressed image.
Result C: The unknown image is a stego image of Yang et al.'s method.

All the images are categorized to Condition A, B, and C as following rules: 400 images are selected as cover images, which are not processed by any compression or data hiding; other 469 images are compressed by traditional VQ with randomly selected codebook size; and the rest 469 images are applied Yang et al.'s method, also with randomly selected codebook size and full embedding. The secret message is composed of '0' bit and '1' bit generated randomly. The output result of applying our steganalysis method for each image is either Result A, Result B, or Result C. The number of various results for the UCID test images are shown in Table 1.The thresholds are set to be $T_F = 1/3$, $T_S = 5$, and $T_R = 0.7$.

Fig. 3. Examples from UCID for testing.

Table 1. The number of various results of the UCID test images.

Input \ Results	Result A (Cover image)	Result B (VQ-compressed image)	Result C (Stego image)
Condition A (400 images)	400	0	0
Condition B (469 images)	17	372	80
Condition C (469 images)	0	50	419

True positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) are statistical measures for sensitive and specificity of our steganalysis scheme. Here stego images represent "Positive", and cover images and traditional VQ-compressed images represent "Negative". According to Table 1 above, the measure values of *TP*, *FP*, *FN*, and *TN* could be calculated as follows:

TP = Result C under Condition C = 419,
FP = Summation of Result C under Condition A and Condition B = 0 + 80 = 80,
TN = Summation of Result A and Result B under Condition A and Condition B = 400 + 17 + 372 = 789,
FN = Summation of Result A and Result B under Condition C = 0 + 50 = 50.

The accuracy rate *Acc* and error rate *Err* could be obtained by the following Eq. (4) and Eq. (5):

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}. \tag{4}$$

$$Err = \frac{FP + FN}{TP + FP + TN + FN} \tag{5}$$

Substitute values of *TP*, *FP*, *FN*, and *TN*, *Acc* reaches 90.28 % but also errors occurred with *Err* 9.72 %. Error often occurs on complex image, which has many sharp lines on it, or the pixel values vary a lot from the adjacent ones. This may confuse the prediction of non-overlapping block size, which causes the wrongly reconstruction of predicted codewords and codebook. Incorrect predicted codewords and codebook lead codebook sorting grouping meaningless, and decrease the reliability of correlation coefficient.

5 Conclusions

There are many studies in respect of digital media steganalysis except for VQ-based steganography. In this paper, we take Yang et al.'s method as an example of steganalysis. We design the procedures to detect the existence of secret message in Yang et al.'s stego image. For other similar VQ-based data hiding methods such as Lu et al.'s method and Wu-Chang's method, the procedure of steganalysis in Fig. 2 is also suitable, only have to apply different algorithm of "Codebook sorting grouping". We observe the non-overlapping blocks of VQ-compression cause horizontal and vertical lines which make the pixel values change not continuously or smoothly. Hence, the proposed scheme is able to determine if an unknown image is VQ-compressed or not. Further, the PoV-like effect exists because of the same probability of '0' and '1' method of secret message, and Yang's shortest-group encoding choses either *CW* or *CW'* in a group. The distribution of predicted codewords appearance show high correlation if the unknown image is truly a stego image of Yang et al.'s method. In our experimental results, the accuracy rate reaches over 90 %. Most of the stego image could be indicated, and cover images and traditional VQ-compressed images are also classified.

Acknowledgments

This research was partially supported by the National Science Council of the Republic of China under the Grant NSC 100-2221-E-015-001-MY2- and NSC 102-2221-E-015-001-.

References

[1] E. S. M. El-Alfy and A. A. Al-Sadi, "High-capacity image steganography based on overlapped pixel differences and modulus function," *Networked Digital Technologies*, Vol. 294, No. 1, pp. 243-252, 2012.

[2] X. Li, Bin. Li, X. Luo, B. Yang, and R. Zhu, "Steganalysis of a PVD-based content adaptive image steganography," *Signal Processing*, Vol. 93, No. 9, pp. 2529-2538, 2013.

[3] Y. Linde, A. Bruzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, Vol. 28, No. 1, pp. 84-95, 1980.

[4] D. C. Lou and C. H. Hu, "LSB steganographic method based on reversible histogram transformation function for resisting statistical steganalysis," *Information Sciences*, Vol. 188, No. 1, pp. 346-358, 2012.

[5] D. C. Lou, C. L. Chou, H. K. Tso, and C. C. Chiu, "Active steganalysis for histogram-shifting based reversible data hiding," *Optics Communications*, Vol. 285, No. 10-11, pp. 2510-2518, 2012.

[6] Z. M. Lu, J. S. Pan, and S. H. Sun, "VQ-based digital image watermarking

method," *Electronics Letters*, Vol. 36, No. 14, pp. 1201-1202, 2000.

[7] X. Y. Luo, D. S. Wang, P. Wang, and F. L. Liu, "A review on blind detection for image steganography," *Signal Processing*, Vol. 88, No. 9, pp. 2138-2157, 2008.

[8] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, No. 3, pp. 354-362, 2006.

[9] H. Noda, M. Niimi, and E. Kawaguchi, "High-performance JPEG steganography using quantization index modulation in DCT domain," *Pattern Recognition Letters*, Vol. 27, No. 5, pp. 455-461, 2006.

[10] C. Qin, C. C. Chang, and Y. C. Chen, "Efficient reversible data hiding for VQ-compressed images based on index mapping mechanism," *Signal Processing*, Vol. 93, No. 9, pp. 2687-2695, 2013.

[11] V. Sachnev and H. J. Kim, "Ternary data hiding technique for JPEG steganography," *Digital Watermarking*, Vol. 6526, No. 1, pp. 202-210, 2011.

[12] S. R. Tsui, C. T. Huang, and W. J. Wang, "Image steganography using gradient adjacent prediction in side-match vector quantization," *Advances in Intelligent Systems and Applications*, Vol. 2, No. 1, pp. 121-129, 2013.

[13] A. Westfeld and A. Pfitzmann, "Attacks on steganographic systems," *Information Hiding*, Vol. 1768, pp. 61-76, 2000.

[14] H. C. Wu and C. C. Chang, "A novel digital image watermarking scheme based on the vector quantization technique," *Computers & Security*, Vol. 24, No. 6, pp. 460-471, 2005.

[15] C. H. Yang, C. Y. Weng, S. J. Wang, and H. M. Sun, "Grouping strategies for promoting image quality of watermarking on the basis of vector quantization," *Journal of Visual Communication and Image Representation*, Vol. 21, No. 1, pp. 49-55, 2010

[16] S-Tools, a steganography software that allows audio and image files to be hidden within other audio and image files, available on: http://www.cs.vu.nl/~ast/books/mos2/zebras.html

[17] Uncompressed Colour Image Database (UCID), provides a benchmark dataset for image retrieval, available on: http://homepages.lboro.ac.uk/~cogs/datasets/ucid/ucid.html