Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 6, June 2014, pg.735 – 740

RESEARCH ARTICLE

Reversible Data Hiding using Histogram Shifting Technique

K. HAZARATHAIAH[1], VAMSEE MOHAN. B[2]

[1]M.Tech Student, Dept of CSE, PBR VITS, Kavali, Nellore, A.P, India

[2]Associate Professor, Dept of CSE, PBR VITS, Kavali, Nellore, A.P, India

[1] hajju.572@gmail.com; [2] vamseemohan22@gmail.com

Abstract*— In this paper reversible data hiding (RDH), the original cover can be lossless restored after the embedded information is extracted. However, in some applications, such as medical imagery, military imagery, and law forensics, no degradation of the original cover is allowed. Kalker and Willems established a rate–distortion model for RDH, in which they proved out the rate–distortion bound and proposed a recursive code construction. In our existing base paper, we improved the recursive construction to approach the rate–distortion bound. In this paper, we generalize the method in our existing paper using a decompression algorithm as the coding scheme for embedding data and prove that the generalized codes can reach the rate–distortion bound as long as the compression algorithm reaches entropy. By the proposed binary codes, and we improve three RDH schemes that use binary feature sequence as covers, i.e., an RS scheme for spatial images, and one scheme for JPEG images, and a pattern substitution scheme for binary images. The experimental result shows that the novel codes can significantly reduce the embedding distortion. Furthermore, by modifying the histogram shift (HS) manner, we also apply these coding methods to one scheme that uses HS, and showing that the proposed codes can be also exploited to develop integer-operation-based schemes.

Keywords* — RDH, RS Scheme, Histogram Shift (HS), image encryption, Security, Difference Expansion (DE)

I. INTRODUCTION

Data hiding is a technique for embedding information into covers such as image, audio, and video files, which can be used for media notation, copyright protection, integrity authentication, and covert communication, etc. Most data hiding methods embed messages into the cover media to produce the marked media by only modifying the least significant part of the cover and, thus ensure perceptual transparency. In this the embedding process will usually introduce permanent distortion to the cover, that is, the original cover cannot at all reconstruct from the marked cover. However, in some applications, such as medical imagery, military imagery, and law forensics, etc., no degradation of the original cover is allowed. In this case, we need a special kind of data hiding method, which is referred to as reversible data hiding (RDH) or lossless data hiding, by which the original cover can be lossless restored after the embedded message is extracted.

Many RDH methods have been proposed since it was introduced. Fridrich and Goljan [1] presented a universal framework for RDH, in which the embedding process is divided into three stages. In this the first stage losslessly extracts compressible features (or portions) from the original cover. The second stage compresses the features with a lossless compression method and, thus, saves space for the payloads (messages). And the third stage embeds messages into the feature sequence and generates the marked cover. One direct reversible embedding method is to compress the feature sequence and append messages after it to form a modified feature sequence, by which restore the original features to generate the marked cover. Hence, after extracting the message, the receiver can restore the original cover by decompressing the features. Fridrich and Goljan [1] suggested features obtained by exploiting characteristics of certain image formats, e.g., texture complexity for spatial images and middle-frequency discrete cosine transform (DCT) coefficients for JPEG images. Celik *et al.* [2] extended Fridrich and Goljan’s scheme by predicting multiple least significant bit (LSB) planes. The same idea Proposed in [1] can be also used for reversible data embedding into binary images [3], [4] or videos [5], [6].

Larger embedding capacity can be achieved by constructing a longer feature sequence that can be perfectly compressed. One of such constructions is difference expansion (DE), which was first proposed by Tian [7], in which the features are the differences between two neighboring pixels. The features are compressed by expansion, i.e., the differences are multiplied by 2, and thus, the LSBs of the differences can be used for embedding messages. Alattar [8] generalized Tian’s method by applying DE to a vector of pixels. Kim *et al.* [9] improved the DE method by reducing the size of the location map used to communicate position information of expandable difference values. The methods proposed in [10] and [11] can achieve better performance by applying DE to the prediction errors.

II. RELATED WORK

Although people have hidden secrets in plain sight-now called steganography-throughout the ages, the recent growth in computational power and technology has propelled it to the front of today's security techniques. And essentially, the information-hiding process in a steganographic system starts by identifying a cover medium's redundant bits (those that can be modified without destroying that medium's integrity). Here the embedding process creates a stego medium by replacing these redundant bits with data from the hidden message. This article discusses existing steganographic systems and as well as presents recent research in detecting them via statistical steganalysis. Here, finally we present recent research and discuss the practical application of detection algorithms and the mechanisms for getting around them.

Steganography is the science that involves communicating secret data in an appropriate multimedia carrier, e.g., image, audio, and video files. It comes under the assumption that if the feature is visible, the point of attack is evident, thus the goal here is always to conceal the very existence of the embedded data. Steganography has various useful applications. However, like any other science it can be used for ill intentions. It has been propelled to the forefront of current security techniques by the remarkable growth in computational power, the increase in security awareness by, e.g., individuals, groups, agencies, government and through intellectual pursuit. Steganography's ultimate objectives, which are undetectability, robustness (resistance to various image processing methods and compression) and capacity of the hidden data are the main factors that separate it from related techniques such as watermarking and cryptography. This paper provides a state-of-the-art review and analysis of the different existing methods of steganography along with some common standards and guidelines drawn from the literature. This paper concludes with some recommendations and advocates for the object-oriented embedding mechanism. Steganalysis, which is the science of attacking steganography, is not the focus of this survey but nonetheless will be briefly discussed.

In this paper, we propose a practical approach to minimizing embedding impact in steganography based on syndrome coding and trellis-coded quantization and contrast its performance with bounds derived from appropriate rate-distortion bounds.

We assume that each cover element can be assigned a positive scalar expressing the impact of making an embedding change at that element (single-letter distortion). The main problem is to embed a given payload with minimal possible average embedding impact. This task, which can be viewed as a generalization of matrix embedding or writing on wet paper, has been approached using heuristic and suboptimal tools in the past. So here, we propose a fast and very versatile solution to this problem that can theoretically achieve performance arbitrarily close to the bound. It is based on syndrome coding using linear convolution codes with the optimal binary quantizer implemented using the Viterbi algorithm run in the dual domain, and here the complexity and memory requirements of the embedding algorithm are linear w.r.t. the number of cover elements. For practitioners, we include detailed algorithms for finding good codes and their implementation. Finally, we report extensive experimental results for a large set of relative payloads and for different distortion profiles, including the wet paper channel also.

III. Histogram Shifting Technique

Data hiding methods embed messages into the cover media to generate the marked media by only modifying the least significant part of the cover and, thus, ensure perceptual transparency. The embedding process will usually introduce permanent distortion to the cover, that is, the original cover can never be reconstructed from the marked cover. However, in some applications, such as medical imagery, military imagery, and law forensics, no degradation of the original cover is allowed.

In data hiding there is no mechanism for getting our original image back. To over this problem we went for reversible data hiding schemes. The existing methods for RDH introducing distortion while embedding and retrieving hidden data. We can reduce this distortion by using optimal codes. In existing systems, there are still limitations in the past three aspects.

- First, the recursive code construction is close to but cannot reach the rate–distortion bound.
- Second, the codes in existing systems are restricted to some discrete embedding rates and cannot approach the maximum embedding rate at the least admissible distortion.
- Third, the codes are restricted to improve Type-I RDH for spatial images, and how to improve Type-II RDH by binary codes is still a problem.

Many RDH methods have been proposed since it was introduced. Fridrich and Goljan presented a universal framework for RDH, in which the embedding process is divided into three stages. The first stage losslessly extracts compressible features (or portions) from the original cover. The second stage compresses the features with a lossless compression method and, thus, saves space for the payloads (messages). The third stage embeds messages into the feature sequence and generates the marked cover. One direct reversible embedding method is to compress the feature sequence and append messages after it to form a modified feature sequence, by which replace the original features to generate the marked cover. Therefore, after extracting the message, the receiver can restore the original cover by decompressing the features. Fridrich and Goljan suggested features obtained by exploiting characteristics of certain image formats, e.g., texture complexity for spatial images and middle-frequency discrete cosine transform (DCT) coefficients for JPEG images. Celik extended Fridrich and Goljan's scheme by predicting multiple least significant bit (LSB) planes.

Lossless Compression Technique

The reversible data hiding schemes based on lossless compression were proposed. The key point of these schemes is to find a subset B in the original. The first methodology is based on lossless compression of subsets or features of the samples comprising the digital object X. If the object X contains a subset B, or a set of features B from X Such that B can be losslessly compressed and be randomized without causing perceptible quality of or object X. The extraction of the hidden message proceeds by extracting the subset B and joining with the bit stream consisting of the compressed bit stream and the message. Replace the set B with its compressed form C(B) and the secret data M , showed in Figure 1. X and X′ indicate the cover image and stego image, respectively.

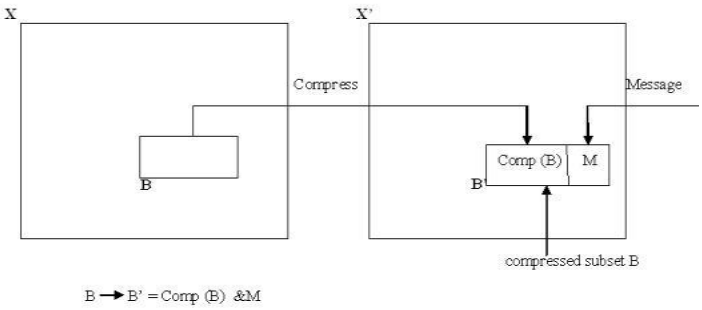

Fig3.1.A Reversible data hiding using the lossless compression technique

Histogram Shifting Technique

The reversible data hiding schemes based on histogram shifting were proposed. In these schemes, peak point in the histogram of the cover image is used to select the embedding area for the secret data, then the part [Peak point +1, Zero point] is shifted to get the embedding area. These schemes were improved by using the histogram of the difference image or predict error image instead of the original image to get a higher peak point. If the peak point is high, the hiding capacity will be large.

Histogram Modification Technique

In histogram modification technique [14], the differences between adjacent pixels instead of simple pixel value is considered. Since image neighbor pixels are strongly correlated, the difference is expected to be very close to zero. At the sending side, first scan the image in an inverse s-order and calculate the pixel difference di between pixels xi-1 and xi by di = xi , if i = 0, | xi-1 - xi |, otherwise. (7)

Determine the peak point P from the histogram of pixel differences. Then again scan the whole image in the same inverse s-order and if di > P, shift xi by 1 unit as follow

xi , if i = 0 or di < P,

yi = xi + 1, if di > P and xi >= xi-1

xi – 1, if di > P and xi < xi-1 (8)

where yi is the watermarked value of pixel i. If di = P, modify xi according to the message bit b as follow

yi = xi + b, if di = P and xi >= xi-1

xi – b, if di = P and xi < xi-1 (9)

At the receiving end, the recipient extracts message bits from the watermarked image by scanning the image in the same order as during the embedding. The message bit b can be extracted by

b = 0, if |yi - xi-1| = P

1, if |yi - xi-1| = P + 1

Where xi−1 denotes the restored value of yi−1. The original pixel value can be restored by

yi + 1, if | yi − xi−1| > P and yi < xi−1

xi = yi – 1, if | yi − xi−1| > P and yi > xi−1

yi , otherwise. (11)

Thus, an exact copy of the original host image is obtained. These steps complete the data hiding and extraction process in which only one peak point is used. Large hiding capacities can be obtained by repeating the data hiding process. However, recipients may not be able to retrieve both the embedded message and the original host image without knowledge of the peak points of every hiding pass. A binary tree structure used to deal with communication of multiple peak points. Modification of a pixel may not be allowed if the pixel is saturated (0 or 255). In this to prevent overflow and underflow, histogram shifting technique is used that narrows the histogram from both sides.

Performance Evaluation

In this module we compare the performance of the existing and the proposed system shows embedding rate and retrieval accuracy of the image and messages. PSNR is most easily defined via the mean squared error (*MSE*). Given a noise-free $m \times n$ monochrome image *I* and its noisy approximation *K*, *MSE* is defined as:

$$\text{MSE}=1/mn\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[\text{I}(i,j)-\text{K}(i,j)]^2$$

The PSNR is defined as:

$$\text{PSNR}=10.\log_{10}(\text{MAX}_I^2/\text{MSE})$$

IV. CONCLUSION

In this paper, most state-of-the-art RDH schemes use a strategy with separate processes of feature compression and message embedding. Kalker and Willems [12] noted that a higher embedding using joint encoding of feature compression and message embedding and, therefore, proposed the recursive code construction. In this paper, we improve the recursive construction by using not only the joint encoding above but also a joint decoding of feature decompression and message extraction. The proposed code construction significantly outperforms previous codes [12], [13] and is proved to be optimal when the compression algorithm reaches entropy.

The current codes are designed for binary covers and, thus, can significantly improve Type-I schemes based on binary feature sequences. In this by slightly modifying the HS manner, we found that the proposed binary codes can be also partly applied to Type-II schemes and improve their performance, but the improvement is not significant as that for Type-I schemes. Note that we only use two simple methods to modify HS, and therefore, one interesting problem is whether there exists other more effective modifying methods or not. Another problem is how to design recursive codes for gray scale covers. We will pay our attention to these problems in further works.

REFERENCES

[1] Weiming Zhang,Biao Chen, and Nenghai Yu, “Improving Various Reversible Data Hiding Schemes Via Optimal Codes for Binary Cover” IEEE *Trans. Image Process.*, vol. 21, no. 6, June. 2013.

[2] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, “Lossless generalized- LSB data embedding,” *IEEE Trans. Image Process.*, vol. 14, no. 2, pp. 253–266, Feb. 2005.

[3] S. Li and A. C. Kot, “Privacy protection of fingerprint database using lossless data hiding,” in *Proc. IEEE Int. Conf.Multimedia Expo.*, 2010, pp. 1293–1298.

[4] Y.-A. Ho, Y.-K. Chan, H.-C. Wu, and Y.-P. Chu, “High-capacity reversible data hiding in binary images using pattern substitution,” *Comput. Standards Interfaces*, vol. 31, no. 4, pp. 787–794, Jun. 2009.

[5] R. Du and J. Fridrich, “Lossless authentication of MPEG-2 video,” in *Proc. IEEE Int. Conf. Image Process.*, 2002, vol. 2, pp. II-893–II-896.

[6] K. Wong, K. Tanaka, K. Takagi, and Y. Nakajima, “Complete video quality-preserving data hiding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 10, pp. 1499–1512, Oct. 2009.

[7] J. Tian, “Reversible data embedding using a difference expansion,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.

[8] A. M. Alattar, “Reversible watermark using difference expansion of a generalized integer transform,” *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1147–1156, Aug. 2004.

[9] H.-J. Kim, V. Sachnev, Y. Q. Shi, J. Nam, and H.-G. Choo, “A novel difference expansion transform for reversible data embedding,” *IEEE Trans. Inf. Forensic Security*, vol. 3, no. 3, pp. 456–465, Sep. 2008.

[10] D. M. Thodi and J. J. Rodriguez, “Expansion embedding techniques for reversible watermarking,” *IEEE Trans. Image Process*, vol. 16, no. 3, pp. 721–730, Mar. 2007.

[11] Y. Hu, H.-K. Lee, and J. Li, “DE-based reversible data hiding with improved overflow location map,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 250–260, Feb. 2009.

[12] T. Kalker and F.M.Willems, “Capacity bounds and code constructions for reversible data-hiding,” in *Proc. 14th Int. Conf. DSP*, 2002, pp. 71–76.

[13] D. Maas, T. Kalker, and F.M.Willems, “A code construction for recursive reversible data-hiding,” in *Proc. Multimedia Security Workshop ACM Multimedia*, Juan-les-Pins, France, Dec. 6, 2002.

[14] W. Zhang, B. Chen, and N. Yu, “Capacity-approaching codes for reversible data hiding,” in *Proc 13th IH*, 2011, vol. 6958, LNCS, pp. 255–269, Springer-Verlag.

SHORT BIOGRAPHY

Mr. K. Hazarathaiah received the **B.Tech** Degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, Anantapur, in **2011.**He currently pursing **M.Tech (CSE) in Dept of Computer Science and Engineering** in PBR VITS Engg college, kavali, Nellore**,** under JNTUA University, Anantapur**.**

Mr. B. Vamsee Mohan. B.Tech., M.Tech**.,** Ph.D.**, His areas of interest** Cloud Computing, Big Data, Network Security, Data mining, Mobile Communication. He has got total 13 years of teaching experience and currently he is working as Assoc.Professor in PBR VITS Engg College, Kavali, Nellore, A.P. He published 12 papers in various conferences and journals.

