Improving Reversible Data Hiding Schemes in Encrypted Images before Encryption

Amol L. Deokate[1], Harshal S. Sangle[2], Kishor P. Jadhav[3]

Abstract: ***This paper describe reversible data hiding in encrypted images, it maintains the excellent property that the original cover can be losslessly recovered after embedded data is extracted while protecting the image content's confidentiality. The drawback of previous methods embeds data by reversibly vacating room from the encrypted images, which may be subject to some errors on data extraction and/or imag restoration. In this paper, we propose method by reserving room before encryption with a traditional RDH algorithm, and so it is easy for the data hider to reversibly embed data in the encrypted image. The proposed method can achieve real reversibility, that is, data extraction and image recovery are free of any error. Experiments show that this method can embed more than 10 times as large payloads for the same image quality as the previous methods.***

Keywords: Image encryption, image extraction, data embedding, reversible data hiding, private security

1. Introduction

The reversible data hiding in images is a technique, by which the original cover can be losslessly recovered after the embedded message is extracted so digital images has increased rapidly on the Internet. Image security becomes increasingly important for many applications, for example confidential transmission, video surveillance, military and medical applications. For example, the necessity of fast and secure diagnosis is vital in the medical world. Nowadays, the transmission of images is a daily routine and it is necessary to find an efficient way to transmit them over networks. To decrease the transmission time, the data compression is necessary. The protection of this multimedia data can be done with encryption or reversible data hiding algorithms. Since few years, a problem is to try to combine compression, encryption and data hiding in a single step. For example, some solutions were proposed in to combine image encryption and compression. Two main groups of technologies have been developed for this purpose. The first one is based on content protection through encryption. There are several methods to encrypt binary images or gray level images. The second group bases the protection on data hiding, aimed at secretly embedding a message into the data. Nowadays, a new challenge consists to embed data in encrypted images. Previous work proposed to embed data in an encrypted image by using an irreversible approach of data hiding or data hiding, aimed at secretly embedding a message into the data. A new idea is to apply reversible data hiding algorithms on encrypted images by wishing to remove the embedded data before the image decryption. Recent reversible data hiding methods have been proposed with high capacity, but these methods are not applicable on encrypted images.

In theoretical aspect, Kalker and Willems [1] established a rate-distortion model for RDH, through which they proved the rate-distortion bounds of RDH for memoryless covers and proposed a recursive code construction which, however, does not approach the bound. Zhang *et al.* [2], [3] improved the recursive code construction for binary covers and proved that this construction can achieve the rate-distortion bound as long as the compression algorithm reaches entropy, which establishes the equivalence between data compression and RDH for binary covers.

In practical aspect, many RDH techniques have emerged in recent years. Fridrich *et al.* [4] constructed a general framework for RDH. By first extracting compressible features of original cover and then compressing them losslessly, spare space can be saved for embedding auxiliary data. A more popular method is based on difference expansion (DE) [5], in which the difference of each pixel group is expanded, e.g., multiplied by 2, and thus the least significant bits (LSBs) of the difference are all-zero and can be used for embedding messages. Another promising strategy for RDH is histogram shift (HS) [6], in which space is saved for data embedding by shifting the bins of histogram of gray values. The state-of-art methods [7]–[11] usually combined DE or HS to residuals of the image, e.g., the predicted errors, to achieve better performance.

With regard to providing confidentiality for images, encryption [12] is an effective and popular means as it converts the original and meaningful content to incomprehensible one. Although few RDH techniques in encrypted images have been published yet, there are some promising applications if RDH can be applied to encrypted images. In [13], Hwang *et al.* advocated a reputation-based trust-management scheme enhanced with data coloring (a way of embedding data into covers) and software watermarking, in which data encryption and coloring offer possibilities for upholding the content owner's privacy and data integrity. Obviously, the cloud service provider has no right to introduce permanent distortion during data coloring into encrypted data. Thus, a reversible data coloring technique basedon encrypted data is preferred. Suppose a medical image database is stored in a data center, and a server in the data center can embed notations into an encrypted version of a medical image through a RDH technique. With the notations, the serve can manage the image or verify its integrity without having the knowledge of the original content, and thus the patient's privacy is protected. On the other hand, a doctor, having the cryptographic key, can decrypt and restore the image in a reversible manner for the purpose of further diagnosing .

Some attempts on RDH in encrypted images have beenmade. In [16], Zhang divided the encrypted image into several blocks. By flipping 3 LSBs of the half of pixels in each block, room can be vacated for the embedded bit. The data extraction and image recovery proceed by finding which part has been flipped in one block. This process can be realized-with the help of spatial correlation in decrypted image. Hong *et al.* [17] ameliorated Zhang's method at the decoder side by further exploiting the spatial correlation using a different estimation equation and side match technique to achieve much lower error rate. These two methods mentioned above rely on spatial correlation of original image to extract data. That is, the encrypted image should be decrypted first before data extraction.

To separate the data extraction from image decryption, Zhang [18] emptied out space for data embedding following the ide of compressing encrypted images [14], [15]. Compression of encrypted data can be formulated as source coding with side information at the decoder [14], in which the typical method is to generate the compressed data in lossless manner by exploiting the syndromes of parity-check matrix of channel codes. The method in [18] compressed the encrypted LSBs to vacate room for additional data by finding syndromes of a parity-check matrix, and the side information used at the receiver side is also the spatial correlation of decrypted images.

All the three methods try to vacate room from the encrypted images directly. However, since the entropy of encrypted images has been maximized, these techniques can only achieve small payloads [16], [17] or generate marked image with poor quality for large payload [18] and all of them are subject to some error rates on data extraction and/or image restoration. Although the methods in [16], [17] can eliminate errors by errorcorrecting codes, the pure payloads will be further consumed.

In the present paper, we propose a novel method for RDH in encrypted images, for which we do not "vacate room after encryption" as done in [16]–[18], but "reserve room before encryption" In the proposed method, we first empty out room by embedding LSBs of some pixels into other pixels with a traditional RDH method and then encrypt the image, so the positions of these LSBs in the encrypted image can be used to embed data. Not only does the proposed method separate data extraction from image decryption but also achieves excellent performance in two different prospects:

- Real reversibility is realized, that is, data extraction and image recovery are free of any error.
- For given embedding rates, the PSNRs of decrypted image containing the embedded data are significantly improved; and for the acceptable PSNR, the range of embedding rates is greatly enlarged.

2. Background

For instance, when the secret data to be transmitted are encrypted, a channel provider without any knowledge of the cryptographic key may tend to compress the encrypted data due to the limited channel resource [1], a lossless compression method for encrypted gray image using progressive decompose and rate-compatible turbo codes is developed in [2]. With the lossy compression method presented in [3], an encrypted gray image can be efficiently compressed by discarding the excessively rough and fine information of coefficients generated from orthogonal transform. When having the compressed data, a receiver may reconstruct the principal content of original image by retrieving the values of coefficients. The computation of transform in the encrypted domain has also been studied. Based on the homomorphic properties of the underlying cryptosystem, the discrete Fourier transform in the encrypted domain can be implemented [4]. In [5], a composite signal representation method packing together a number of signal samples and processing them as a unique sample is used to reduce the complexity of computation and the size of encrypted data. There are also a number of works on data hiding in the encrypted domain. In a buyer–seller watermarking protocol [6], the seller of digital multimedia product encrypts the original data using a public key, and then permutes and embeds an encrypted fingerprint provided by the buyer in the encrypted domain. After decryption with a private key, the buyer can obtain a watermarked product. This protocol ensures that the seller cannot know the buyer's watermarked version while the buyer cannot know the original version.

By introducing the composite signal representation mechanism, both the computational overhead and the large communication bandwidth due to the homomorphic public key encryption are also significantly reduced [8]. For example [9], the intraprediction mode, motion vector difference and signs of DCT coefficients are encrypted, while a watermark is embedded into the amplitudes of DCT coefficients. In [10], the cover data in higher and lower bit-planes of transform domain are respectively encrypted and watermarked. In [11], the content owner encrypts the signs of host DCT coefficients and each content-user uses a different key to decrypt only a subset of the coefficients, so that a series of versions containing different fingerprints are generated for the users. The reversible data hiding in encrypted image is investigated in [12]. Most of the work on reversible data hiding focuses on the data embedding/extracting on the plain spatial domain [13]–[14]. In other words, the additional data must be extracted from the decrypted image, so that the principal content of original image is revealed before data extraction, and, if someone has the data-hiding key but not the encryption key, receiver cannot extract any information from the encrypted image containing additional data.

3. Proposed Method

The proposed scheme is made up of image encryption, data embedding and data-extraction/image-recovery phases. This paper proposes a separable reversible data hiding in encrypted image. In the proposed scheme, the original image is encrypted using an encryption key and the additional data are embedded into the encrypted image using data-hiding key. With an encrypted image containing additional data, if the receiver has only the data-hiding key, he can extract the additional data though receiver does not know the image content. If receiver has only the encryption key, he can decrypt the received data to obtain an image similar to the original one, but cannot extract the embedded additional data.

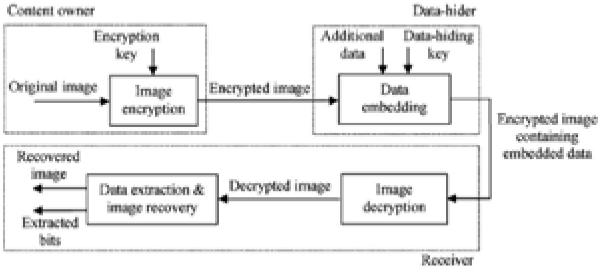

Figure 1: Non-separable reversible data hiding in encrypted Image

The content owner encrypts the original uncompressed image using an encryption key to produce an encrypted image. Then, the data hider compresses the least significant bits (LSB) of the encrypted image using a data-hiding key to create a sparse space to accommodate the additional data. At the receiver side, the data embedded in the created space can be easily retrieved from the encrypted image containing additional data according to the data-hiding key. Since the data embedding only affects the LSB, a decryption with the encryption key can result in an image similar to the original version. Fig. 2 shows the three cases at the receiver side.

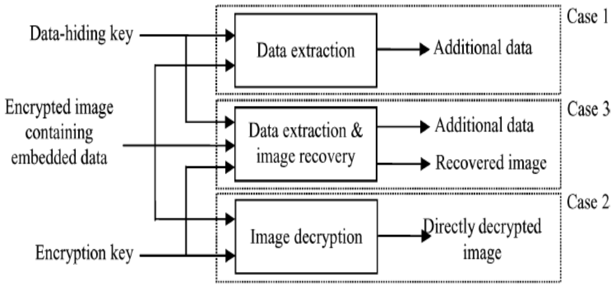

Figure 2: Three cases at receiver side of the proposed separable scheme

When using both of the encryption and data-hiding keys, the embedded additional data can be successfully extracted and the original image can be perfectly recovered by exploiting the spatial correlation in natural image.

4. Implementation Details

A. Image Encryption

While an encrypted binary image can be compressed with a lossless manner by finding the syndromes of low-density parity-check codes, a lossless compression method for encrypted gray image using progressive decomposition and rate-compatible punctured turbo codes is developed in .With the lossy compression method presented in, an encrypted gray image can be efficiently compressed by discarding the excessively rough and fine information of coefficients generated from orthogonal transform. When having the compressed data, a receiver may reconstruct the principal content of original image by retrieving the values of coefficients. The content owner encrypts the original uncompressed image using an encryption key to produce an encrypted image. Then, the data-hider compresses the least significant bits (LSB) of the encrypted image using a data-hiding key to create a sparse space to accommodate the additional data. At the receiver side, the data embedded in the created space can be easily retrieved from the encrypted image containing additional data according to the data-hiding key. Since the data embedding only affects the LSB, a decryption with the encryption key can result in an image similar to the original version. When using both of the encryption and data-hiding keys, the embedded additional data can be successfully extracted and the original image can be perfectly recovered by exploiting the spatial correlation in natural image.

B. Data Embedding

In the data embedding phase, some parameters are embedded into a small number of encrypted pixels, and the LSB of the other encrypted pixels are compressed to create a space for accommodating the additional data and the original data at the positions occupied by the parameter.

According to the data hiding key, the data hider pseudo randomly selects NP encrypted pixels that will be used to carry the parameters for data hiding. Here NP is a small positive integer, for example NP=20.The other encrypted pixels are pseudo-randomly permuted and divided into number of groups, each of which contain L pixels. The permutation way is also determined by the data-hiding key. For each pixel-group, collect the M least significant bits of the L pixels, and denote them as B (k,1) , B (k,2) …… B(k,M*L) where k is a group index within [1,(N-Np)/L] and M is a positive integer less than 5. The data-hider also generates a matrix G sized (M*L – S) * M*L, which is composed of two parts. The left part is the identity matrix and the right part is pseudo-random binary matrix derived from the data-hiding key. For each group, which is product with the G matrix to form a matrix of size (M * L-S). Which has sparse bits of size S, in which the data is embedded and arrange the pixels into the original form and permutated to form a original image.

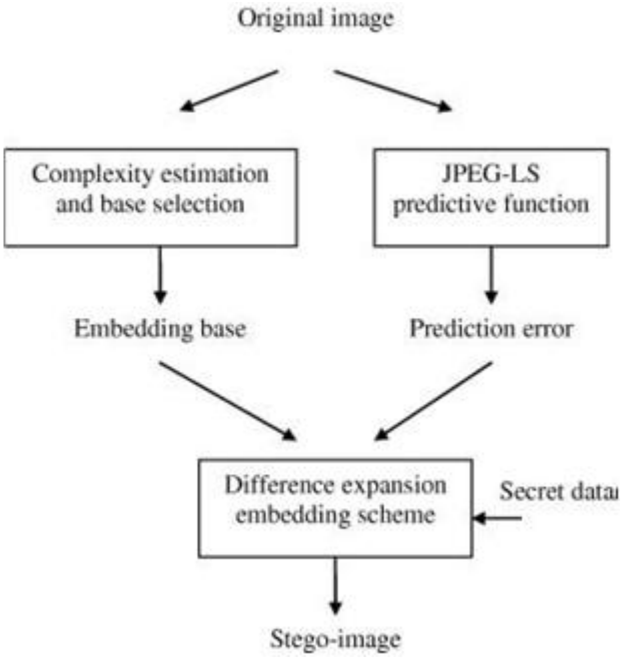
Original image
Complexity estimation and base selection
JPEG-LS predictive function
Embedding base
Prediction error
Difference expansion embedding scheme
Secret data
Stego-image

Figure 3: Data Embedding

C. Image Decryption

When having an encrypted image containing embedded data, a receiver firstly generates ri,j,k according to the encryption key, and calculates the exclusive-or of the received data and ri,j,k to decrypt the image. We denote the decrypted bits as b1i,j,k . Clearly, the original five most significant bits (MSB) are retrieved correctly. For a certain pixel, if the embedded bit in the block including the pixel is zero and the pixel belongs to S1, or the embedded bit is 1 and the pixel belongs to S0 , the data-hiding does not affect any encrypted bits of the pixel. So, the three decrypted LSB must be same as the original LSB, implying that the decrypted gray value of the pixel is correct. On the other hand, if the embedded bit in the pixel's block is 0 and the pixel belongs to S0, or the embedded bit is 1 and the pixel belongs to S1 , the decrypted LSB.That means the three decrypted LSB must be different from the original LSB.In this case: **b'i,j,k + bi,j,k = 1** On the other hand, if the embedded bit in the pixel's block is 0 and the pixel belongs to S0, or the embedded bit is 1 and the pixel belongs to S1, the decrypted LSB **D. Data Extraction**

The receiver has both the data hiding, he may aim to extract the embedded data according to the data hiding key. The values of M, Land S, the original LSB of the Np selected encrypted pixels, and the (N-Np) * S/L - Np additional bits can be extracted from the encrypted image containing embedded data. By putting the Np LSB into their original positions, the encrypted data of the Np selected pixels are retrieved, and their original gray values can be correctly decrypted using the encryption keys. In the following, it will recover the original gray values of the other (N-Np) pixels. Consider the case that the receiver has the encryption key but does not know the data-hiding key. Clearly, he cannot obtain the values of parameters and cannot extract the embedded data. However, the original image content can be roughly recovered.

5. Experimental Results

The proposed reversible data hiding algorithm has been applied to many different types of images, including some commonly used images, medical images, texture images, aerial images, and all of the 1096 images in the Corel DRAW database, and has always achieved satisfactory results, thus demonstrating its general applicability. The proposed reversible data hiding technique is able to embed about 5–80 kb into a 512* 512 8 grayscale image while guaranteeing the PSNR of the marked image versus the original image to be above 48 dB. In addition, this algorithm can be applied to virtually all types of images. In fact, it has been successfully applied to many frequently used images, medical images, texture images, aerial images, Furthermore, this algorithm is quite simple, and the execution time is rather short. Therefore, its overall performance is better than many existing reversible data hiding algorithms. It is expected that this reversible data hiding technique will be deployed for a wide range of applications in the areas such as secure medical image data systems, and image authentication in the medical field and law enforcement, and the other fields where the rendering of the original images is required or desired.

Additionally, the quality index Q works in spatial domain, as a combination of correlation loss, luminance distortion and contrast distortion. Higher PSNR, lower Watson metric or higher Q means better quality. In these figures, while the abscissa represents the embedding rate, the ordinate is the values of PSNR, Watson metric or quality index Q. The curves are derived from different L, M and S under a condition that the original content can be perfectly recovered using the data hiding and encryption keys. Since the spatial correlation is exploited for the content recovery, the rate-distortion performance in a smoother image is better. The performance of the non separable method is also given in Figure. It can be seen that the performance of the proposed separable scheme is significantly better. It also compared the proposed scheme with the non-separable method over 100 images sized 2520 * 3776, which were captured with a digital camera and contain landscape and people. When meeting the perfect recovery condition, the proposed scheme has an average 203% gain of embedded data amount with same PSNR value in directly decrypted image, or an average gain of 8.7 dB of PSNR value in directly decrypted image with same embedded data amount.

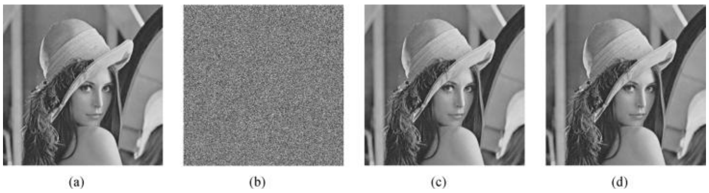

Figure 5: (a) Original image (b) Encrypted version (c) Encrypted image with message (d) Decrypted image

6. Conclusion

This paper presents reversible data hiding scheme for encrypted image which consists of image encryption, data embedding and data extraction or image recovery phases. The data of original image are entirely encrypted by a stream cipher. Although a data hider does not know the original content, he can embed additional data into the encrypted image by modifying a part of encrypted data. With an encrypted image containing embedded data, a receiver may firstly decrypt it using the encryption key, and the decrypted version is similar to the original image. According to the data hiding key, with the aid of spatial correlation in natural image, the embedded data can be correctly extracted while the original image can be perfectly recovered. Although someone with the knowledge of encryption key can obtain a decrypted image and detect the presence of hidden data using LSB steganalytic methods, if he does not know the data hiding key, it is still impossible to extract the additional data and recover the original image. For ensuring the correct data extraction and the perfect image recovery, It may let the block side length be a big value or introduce error correction mechanism before data hiding to protect the additional data with a cost of payload reduction. The implemented a reversible method can be enhanced in future by using the following provisions and MLSB technique can also be applied after embedding when there is lot of change in the pixel to retain nearest to the original value. It can be applied in networking and the keys are sent and received securely. The image produced by the reversible data hiding using two key has distortion. In order to remove distortion and to produce the image in a high quality using 3 key.

7. Acknowledgment

It is with the greatest pride that we publish this paper. At this moment, it would be unfair to neglect all those who helped me in the successful completion of this paper. I would also like to thank all the faculties who have cleared all the major concepts that were involved in the understanding of techniques behind my paper.

References

[1] T. Kalker and F.M.Willems, “Capacity bounds and code constructions for reversible data-hiding,” in *Proc. 14th Int. Conf. Digital Signal Processing (DSP2002)*, 2002, pp. 71–76.

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

[3] W. Zhang, B. Chen, and N. Yu, “Improving various reversible data hiding schemes via optimal codes for binary covers,” *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 2991–3003, Jun. 2012.

[4] J. Fridrich and M. Goljan, “Lossless data embedding for all image formats,” in *Proc. SPIE Proc. Photonics West, Electronic Imaging, Security and Watermarking of Multimedia Contents*, San Jose, CA, USA, Jan. 2002, vol. 4675, pp. 572–583.

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

[6] Z. Ni, Y. Shi, N. Ansari, and S. Wei, “Reversible data hiding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 354–362, Mar. 2006.

[7] 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.

[8] X. L. Li, B. Yang, and T. Y. Zeng, “Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection,” *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3524–3533, Dec. 2011.

[9] P. Tsai, Y. C. Hu, and H. L. Yeh, “Reversible image hiding scheme using predictive coding and histogram shifting,” *Signal Process.*, vol. 89, pp. 1129–1143, 2009.

[10] L. Luo *et al.*, “Reversible imagewatermarking using interpolation technique,” *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 187–193, Mar. 2010.

[11] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi, “Reversible watermarking algorithm using sorting and prediction,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 989–999, Jul. 2009.

[12] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL, USA: CRC, 1996.

[13] K. Hwang and D. Li, “Trusted cloud computing with secure resources and data coloring,” *IEEE Internet Comput.*, vol. 14, no. 5, pp. 14–22, Sep./Oct. 2010.

[14] M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran, “On compressing encrypted data,” *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2992–3006, Oct. 2004.

[15] M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran, “On compressing encrypted data,” IEEE Trans. Signal Process., vol. 52, no. 10, pp. 2992–3006, Oct. 2004.

[16] W. Liu, W. Zeng, L. Dong, and Q. Yao, “Efficient compression of encrypted grayscale images,” IEEE Trans. Image Process., vol. 19, no. 4, pp. 1097–1102, Apr. 2010.

[17] X. Zhang, “Lossy compression and iterative reconstruction for encrypted image,” IEEE Trans. Inform. Forensics Security, vol. 6, no. 1, pp. 53–58, Feb. 2011.

[18] T. Bianchi, A. Piva, and M. Barni, “On the implementation of the discrete Fourier transform in the encrypted domain,” IEEE Trans. Inform. Forensics Security, vol. 4, no. 1, pp. 86–97, Feb. 2009.

[19] T. Bianchi, A. Piva, and M. Barni, “Composite signal representation for fast and storage-efficient processing of encrypted signals,” IEEE Trans. Inform. Forensics Security, vol. 5, no. 1, pp. 180– 187, Feb. 2010.

[20] N. Memon and P. W. Wong, “A buyer-seller watermarking protocol,” IEEE Trans. Image Process., vol. 10, no. 4, pp. 643–649, Apr. 2001.

[21] M. Kuribayashi and H. Tanaka, “Fingerprinting protocol for images based on additive homomorphic property,” IEEE Trans. Image Process., vol. 14, no. 12, pp. 2129–2139, Dec. 2005.