Improve The Recursive Construction for Decompression and Message Extraction by using Joint Encoding and Decoding

Ms. M. Magishasini.C[1] Mr. C. Prakash Narayanan[2] B. Sakthivel[3]
[1]M. E. [2]Associate Professor [3]HOD
[1, 2, 3]Computer Science Engineering Department
[1, 2, 3] P.S.V College of Engineering and Technology, Krishnagiri, Tamilnadu, India

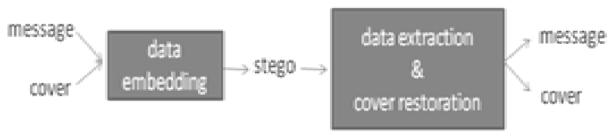
Abstract---In reversible data hiding (shortly RDH), the original cover can be lossless restored after the embedded information is extracted. A famous author Kalker and Willems established a rate–distortion model for RDH, in which they proved out the rate–distortion bound and proposed a recursive code construction. In previous concept, improved the recursive construction to approach the rate distortion bound. But In this approach, I will use 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 proposed binary codes, I improve three RDH schemes that use binary feature sequence as covers, i.e., an RS scheme for spatial images, one scheme for JPEG images, and a pattern substitution for binary images. By modifying the histogram shift (HS) manner, I also apply this coding method to one scheme that uses HS, showing that the proposed codes can be also exploited to improve integer-operation-based schemes.

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, covert communication, etc. Most 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 these cases, we need a special kind of data hiding method, which is referred to as reversible data hiding (shortly called as RDH) or lossless data hiding, by which the original cover can be lossless restored after the embedded message is extracted.

What is reversible data hiding?

The original cover can be losslessly restored after the embedded information is extracted.



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 *et al.* extended Fridrich and Goljan's scheme by predicting multiple least significant bit (LSB) planes. The same idea proposed in can be also used for reversible data embedding into binary images or videos.

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, 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 generalized Tian's method by applying DE to a vector of pixels. Kim improved the DE method by reducing the size of the location map used to communicate position information of expandable difference values.

The methods proposed can achieve better performance by applying DE to the prediction errors. Another well-known strategy for RDH is histogram shift (shortly HS), in which the histogram of the image is used as the compressible features because the distribution of the pixel values of an image is usually uneven. To compress the histogram, proposed to select a peak bin and a zero bin and shift the bins between them toward the zero bin by one step.

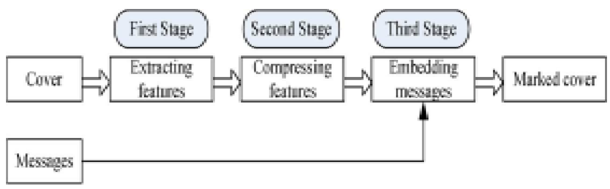


Fig. 1: Diagram for the framework of RDH at the sender side.

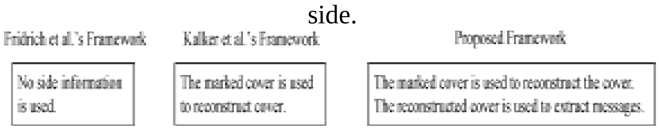


Fig. 2: Side information used at the receiver side in three frameworks.

Therefore, the peak bin's neighboring bin, which is now emptied out, and the peak bin can be used to present "1" and "0," respectively. It is easy to see that a steeper histogram implies larger capacity, and, usually, the histogram of residuals is quite steep. Thus, most state-of-the-art methods apply HS to residuals of the image. Both DE- and HS-based schemes use integer features and special methods to compress the features. As for DE, the features (differences) are compressed by expansion operation, and as for HS, the features (histogram) are compressed by shifting operation. There is a common character in both these schemes, that is, the distortion to the original cover is mainly introduced by the special compressing manners. By contrast, Fridrich and Goljan's schemes use a binary feature sequence and a generic compression algorithm, e.g., the arithmetic coder, and no distortion must be introduced by the compression. According to such differences, we divide RDH into two types as follows.

Type I. The features can be formulated as a binary sequence and can be compressed by using a generic compression algorithm. The methods in belong to *Type I*.

Type II. The features are no binary and compressed in some specific manners. Both DE-based and HS-based methods belong to Type II.

For Type-I RDH, the problem is formulated as how to reversibly embed data into a compressible binary sequence with good performance. The performance is measured by embedding rate versus distortion, which is a special rate–distortion coding problem.

In, the authors obtained the rate–distortion function, i.e., the upper bound of the embedding rate under a given distortion constraint, and, by dividing the cover into disjoint blocks, they proposed a recursive code construction, which consists of a nonreversible data embedding code and a conditional compression code.

In fact, Kalker and Willems noted that the receiver can reconstruct the cover with the help of the marked cover, and thus, the sender can compress the cover under the condition of the marked cover. That is why the recursive construction is efficient. In our previous paper, we improved the recursive construction by using not only conditional compression but also conditional embedding, which enables us to design an efficient embedding algorithm and a perfect compressing method to approach the rate–distortion bound. In fact, we noted that the receiver could extract messages from the marked cover with the help of the reconstructed cover because of reversibility. In Fig. 2, the side information exploited at the receiver side in the proposed framework is compared with those used in two previous frameworks.

However, there are still limitations in three aspects first, we construct embedding codes by improving the decompression algorithm of run-length coding, by which the recursive code construction is close to but cannot reach the rate–distortion bound. Second, the codes 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.

In this paper, we generalize the code construction by using a general decompression algorithm as the embedding code and extend the applications to *Type-II RDH*. Compared with our preliminary paper [17], the new contributions of this paper are as follows.

- We prove that the recursive code construction can reach the rate–distortion bound when the decompression/compression algorithms used in the code are optimal, which establishes equivalence between source coding and RDH for binary covers.
- With the decompression of the adaptive arithmetic coder (shortly AAC) as the embedding code, the proposed codes realize continuous embedding rates and reach the maximum embedding rate at the least admissible distortion.
- A method is presented to improve integer-operation-based RDH *(Type II)* by the proposed binary codes, which are also applied to Type-I RDH for JPEG and binary images.

II. MODULES

- Coding Model
- Recursive Construction
- Optimality
- Improving the Scheme for JPEG Images

III. MODULES DESCRIPTION

A. *Coding Model*

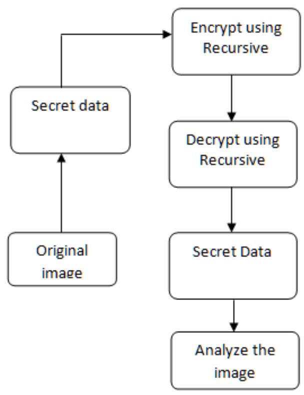
Throughout this paper, I denote matrices and vectors by boldface fonts and use the same notation for the random variable and its realization, for simplicity. To do RDH, a compressible feature sequence should be first extracted from the original cover. For Type-I schemes, the features can be usually represented by a binary sequence. Therefore, I directly take the binary feature sequence as the cover to discuss the coding method and follow the notation established.

B. *Recursive Construction*

This recursive construction performs better than the simple method because of two key points: 1) The data is embedded by an efficient nonreversible embedding code, and 2) the cover block is compressed under the condition of the marked block. However, the above recursive construction cannot approach the upper bound.

C. *Optimality*

The next theorem shows that the proposed code construction is optimal as long as the compression algorithm is optimal.



D. Improving the Scheme for JPEG Images

To improving the JPEG images, I apply the codes to the reversible embedding scheme for JPEG images proposed by Firdrich and Goljan. In the method in, quantized DCT coefficients that are equal to 0 and 1 at middle or high frequency are selected to form a compressible binary sequence. In our experiments, the test images are generated by compressing test images in Fig. 7 into a JPEG format with quality factor 80.

I will construct the binary cover by extract 0–1 coefficients from 11 positions, such as (3, 3), (2, 4), (4, 2), (1, 5), (5, 1), (3, 4), (4, 3), (2, 5), (5, 2), (1, 6), and (6, 1), from every 8 8 block of quantized DCT coefficients. Random messages are embedded into the binary cover by using Fridrich and Goljan's method and the proposed codes with several kinds of embedding rates.

IV. LITERATUER SURVEY

A. A Novel Data Hiding Method by Using Chaotic Map and Histogram

Data hiding is to conceal the existence of secret data and it is considered for more protection of multimedia data. A reversible data hiding method can extract the cover image without any distortion from the stego image after the hidden data have been extracted. This study tackles a chaotic based reversible data hiding. In this paper first image histogram is employed for detect the pixels which are selected for hiding a bit of secret data, then after a sequence of hiding a bit stream is determined by logistic chaotic map. Experimental results show that WICA not only demonstrates superior hiding effect, but also resists various typical attacks. The obtained PSNR of the proposed method is approximately 54 which is proven our method excellence.

1) The chaotic model

Chaotic signals seem like noise, but they are completely definite: if the initial values and the mapping function are known, the same values can be accurately reproduced. The advantages of these signals are studied under the following three headings:

a) Sensitivity to the Initial Conditions

This means that any slight change in the initial values will cause huge changes in the subsequent values of the function i.e., if there is a small change in the initial values of the signal, the resultant signal will be very different from the initial one.

b) The Apparently Random Behavior

Compared to the producers of the natural random numbers in which the string of the random numbers produced cannot be reproduced, the methods used in producing random numbers in algorithms based on chaotic models allow the reproduction of the same random numbers, provided that the initial values and the mapping function are known.

c) Definite Operation

Although chaotic models appear to be random, yet they are completely definite: if the mapping function and the initial values are known, a set of values can be produced (apparently without any order in their production) in order to be used in the reproduction of those same initial values.

2) The proposed method

In the proposed method, the most frequent gray surface in the image, which can be seen at the tip of the histogram diagram, is used for hiding the bits of the encrypted data. In the process of hiding information, first the most frequent gray surface and the gray surface with zero frequency are found.

B. DE-Based Reversible Data Hiding With Improved Overflow Location Map

Difference-expansion (shortly DE)-based reversible data hiding, the embedded bit-stream mainly consists of two parts: one part that conveys the secret message and the other part that contains embedding information, including the 2-D binary (overflow) location map and the header file. The first part is the payload while the second part is the auxiliary information package for blind detection. To increase embedding capacity, I have to make the size of the second part as small as possible.

Tian's classical DE method has a large auxiliary information package. The mitigated the problem by using a payload-independent overflow location maps. However, the compressibility of the overflow location map is still undesirable in some image types. In this paper, I focus on improving the overflow location map. I design a new embedding scheme that helps us construct an efficient payload-dependent overflow location map. Such an overflow location map has good compressibility. Our accurate capacity control capability also reduces unnecessary alteration to the image. Under the same image quality, the proposed algorithm often has larger embedding capacity. It performs well in different types of images, including those where other algorithms often have difficulty in acquiring good embedding capacity and high image quality.

Reversible data hiding was first proposed for authentication. Early reversible algorithms often have small embedding capacity and poor image quality. With the improvement of embedding capacity and image quality, this technique is being considered not only for the whole spectrum of fragile watermarking, such as authentication watermarks or watermarks protecting the image integrity but also for covert communication, even for some unprecedented applications like image/video coding.

V. CONCLUSION

Most state-of-the-art RDH schemes use a strategy with separate processes of feature compression and message embedding. Kalker and Willems noted that a higher embedding using joint encoding of feature compression and message embedding and, thus, proposed the recursive code construction. In this paper, I will 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 and is proved to be optimal when the compression algorithm reaches entropy.

VI. FUTURE WORK

The current codes are designed for binary covers and, thus, can significantly improve Type-I schemes based on binary feature sequences. By slightly modifying the HS manner, I found that the proposed binary codes can be also partly applied to Type-II schemes and improve their performance, but the improvement is not as significant as that for Type-I schemes. Note that I will 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. I will pay our attention to these problems in further work.

REFERENCES

[1] Biao Chen, Nenghali Yu and Weiming Zhang (2012) „Improving Various Reversible Data Hiding Schemes Via Optimal Codes for Binary Covers‟ IEEE Trans. on Image Processing, vol.21,no.6.
[2] J. Fridrich and M. Goljan, “Lossless data embedding for all image formats,” in Proc. EI SPIE, Security Watermarking Multimedia Contents IV, San Jose, CA, 2002, vol. 4675, pp. 572–583.
[3] 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.
[4] 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.
[5] H.J Kim,V.Sachnev,Y.Q.Shi,J.Nam,and H.G.Choo,“A novel difference expansion transform for reversible data embedding,”