ISSN 2186-5647

2-10

Reversible Data Hiding Exploiting Variance in Wavelet Coefficients

China University of Technology ○ Te-Lung Yin

ABSTRACT

In this paper, we present a new reversible data hiding scheme that utilizes the wavelet transform and better exploits the large wavelet coefficient variance to achieve high capacity and imperceptible embedding. Our scheme differs from those of previous studies in that the wavelet coefficients histogram rather than the gray-level histogram is manipulated. In addition, we design intelligent histogram-shifting rules to avoid the decimal problem in grayscale pixel values after recovery process to achieve reversibility. Small changes in the wavelet coefficients after embedding process are important factors contributing to low visual distortion in the marked image. Furthermore, an important property of our scheme is that the use of threshold differs greatly from previous schemes. The experimental results show that our scheme outperforms other reversible data hiding schemes.

Keywords: reversibility, marked media, wavelet transform, wavelet coefficient, distortion, histogram.

1. Introduction

Reversible data hiding, or so-called invertible, distortion-free data hiding, is a branch of fragile technique mainly used for quality-sensitive applications such as multimedia content authentication, medical imaging systems, law enforcement, and military imagery, etc. One of the most important requirements in these fields is to recover the original media exactly during analysis to enable the right decisions. The other significant necessities of reversible data hiding are the embedding capacity and visual quality of the marked media, since they are critically essential to achieving satisfactory performance in various applications [1, 2].

The scheme we present in this paper is an attempt to achieve high-performance reversible data hiding, in which the embedding and recovering processes are devised in the frequency domain. The particularities of large variance in wavelet coefficients and minor changes in wavelet coefficients following from the embedding process in wavelet coefficients are exploited to achieve high capacity and imperceptibility. The rest of this paper is organized as follows. In Section 2, previous reversible data hiding schemes and their characteristics will be briefly reviewed in terms of embedding capacity and visual quality. Our proposed scheme is introduced in Section 3. Experimental results and comparative analyses are presented in Section 4. Finally, some conclusions are drawn in Section 5.

2. Related studies

Nowadays, a number of research works in this field can be classified into two major categories according to the embedding strategies. Category-I reversible data hiding schemes work on the transform domain [3]. In category-II, schemes are performed in the spatial domain [4-7].

For all of the above reversible data hiding schemes, the requirement of additional overhead is on of the thorniest problems in the restore process. This paper presents a novel method in the frequency domain to achieve high-performance lossless data hiding.

3. Proposed scheme

The proposed scheme combines the two-level Haar discrete wavelet transform (HDWT) algorithm and a new histogram

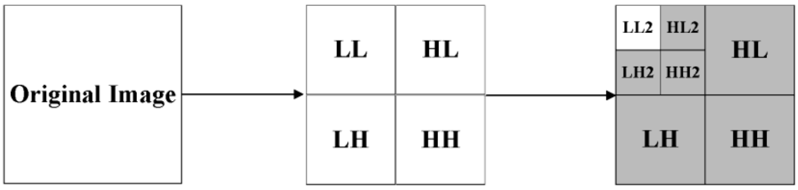

Fig. 1. The two-level segmentation process.

Fig. 2. A tow-level HDWT four-band split of "Lena".

shifting technique to achieve reversible data hiding. In our scheme, a given image is first transformed into a frequency domain and sub-bands in the middle- and high-frequency ranges are then used to create sub-band differences. Each histogram of these sub-band differences is then shifted according to a selected threshold. Message bits can then be embedded in the empty space of the shifted histograms. Finally, the marked image is reconstructed with the sub-bands carrying and non-carrying hidden message by performing the inverse HDWT algorithm to complete the embedding process. As to the extracting process, the corresponding inverse operations can be performed to recover the hidden information and the original image.

3.1. Segmentation algorithm

The two-level HDWT algorithm utilizes the four-band sub-band coding system to decompose an image into a set of different frequency sub-bands. As illustrated in Figs. 1 and 2, the size of each sub-band is one eighth of the original image in the spatial domain. The eight different sub-bands can be classified into the low-, middle-, and high-frequency sub-bands. Since the low-frequency sub-band of an image incorporates more energy than the other sub-bands, its coefficients are the most fragile that if any of them are manipulated, a suspect can visibly detect the changes on the spatial domain image. In contrast, if the coefficients in the middle- and/or high-frequency sub-bands are altered, changes in the spatial domain image are imperceptible to human eyes. As a result, this feature is generally exploited to conceal secret messages.

3.2. Data embedding algorithm

We assume that the embedded message is a random binary sequence. The histograms of the sub-band differences between the reference sub-band and the other destination sub-bands are shifted to embed the secret message. Fig. 3 depicts the overall data embedding process, which is described in detail below.

3.3 Data extracting and recovering algorithm

Before extracting the hidden message, receiver needs to verify whether or not the marked image has been modified. If there is more than one occurrence at , we can conclude that the marked image has been tampered with. The proposed scheme then stops the following extraction steps immediately. The extraction and recovery process is schematized in Fig. 4.

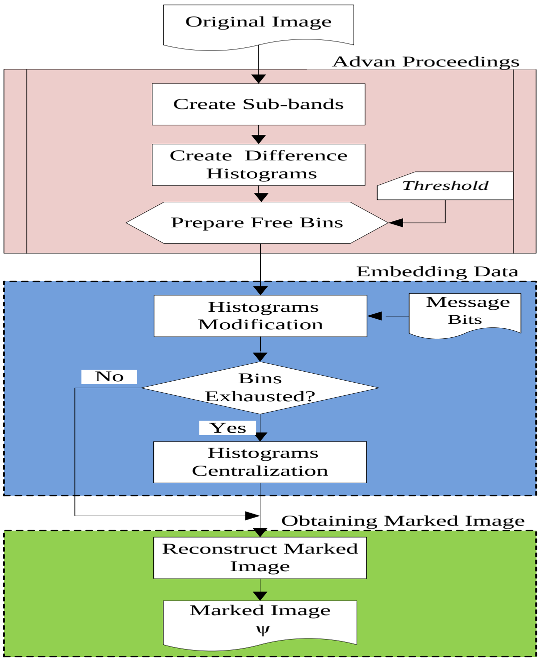

Fig. 3. Flowchart of data embedding process

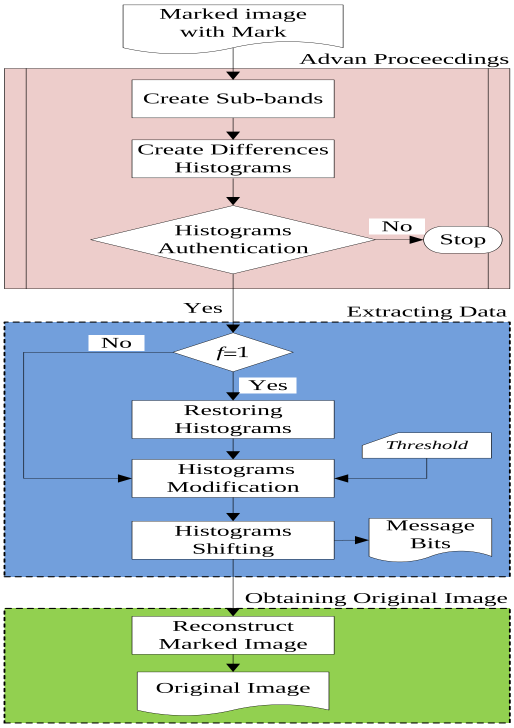

Fig. 4. Flowchart of extraction and recovery process

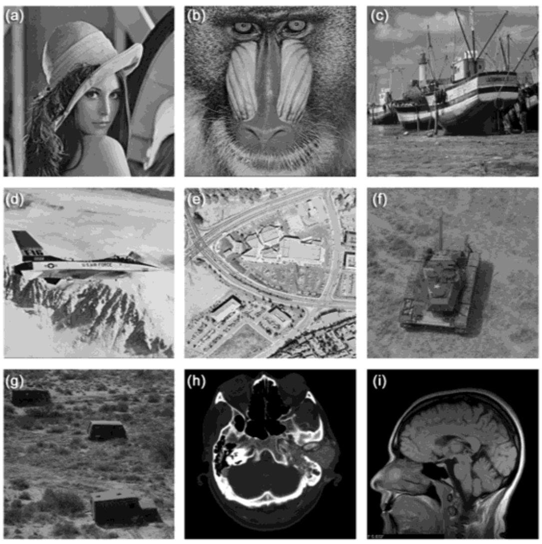

Fig. 5. 8-bit 512×512 images(a)Lena, (b)Baboon, (c)Boat, (d)Airplane, (e)Aerial, (f)Tank, (g)Trucks, (h)Medical image1, (i) Medical image2 .

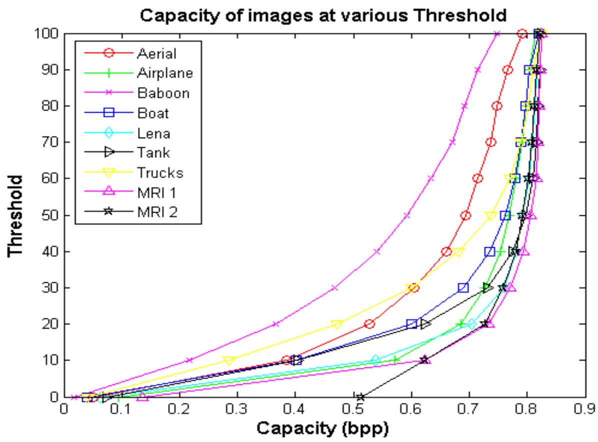

Fig.6. Embedding capacity at various thresholds.

4. Experimental results and comparison

In this section, a series of experiments are performed to evaluate of our scheme. For these experiments, we used many different types of images, including some commonly used ones and two medical images (Fig. 5). The message bits to be embedded in our experiments are randomly generated by a pseudo-random binary generator. The threshold ranges from 0 to 100.

Capacity versus Threshold

The relationship between the capacity (in bpp) and the threshold is presented in Fig. 6. The embedding rate almost reaches 0.8 bpp at threshold 100 for most the test images. As expected the capacity is nearly proportional to the threshold at the beginning and saturates when the threshold is sufficiently high.

5. Conclusion

In this paper, a reversible data hiding scheme exploiting the large variance of wavelet coefficients and clever histogram shifting rules is presented. The proposed scheme, compared with previous ones, can obtain better visual quality of the marked image given the same payload. The main reason is that the visual quality of our scheme does not decay with increasing threshold as in the other schemes. In addition, our scheme provides the greatest embedding capacity. It may be of interest for future research that the threshold predictions, multi-round schemes, and fast algorithms will be explored to meet real-time application requirements.

References

[1] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding", IBM Systems Journal, vol.35, no.3, pp.313–336, 1996.

[2] M. Awrangjeb, "An overview of reversible data hiding", Proc. Sixth International Conf. on Computer and Information Technology, Jahangirnagar University, Bangladesh, pp. 75–79, December 2003.

[3] J. Tian, "Reversible data embedding using a difference expansion", IEEE Trans. on Circuits and Systems for Video Technology vol.13, no.8, pp.890–896, 2003.

[4] M.U. Celik, G. Sharma, and A.M. Tekalp, "Reversible data hiding", Proc. IEEE International Conf. on Image Processing, Rochester, NY, pp.157–160, 2002.

[5] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Lossless generalized-LSB data embedding", IEEE Trans. on Image Proc. vol.14, no.2, pp.253–266, February 2005.

[6] 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, March 2006.

[7] K.-S. Kim, M.-J. Lee, H.-Y. Lee, and H.-K. Lee, "Reversible data hiding exploiting spatial correlation between sub-sampled images", Pattern Recognition, vol.42, pp.3083-3096, 2009.