Robust Lossless Data Hiding: Analysis and Evaluation

Lingling An, Xinbo Gao, Cheng Deng, and Feng Ji
School of Electronic Engineering, Xidian University, Xi'an 710071, China
an.lingling@gmail.com

ABSTRACT

Robust lossless data hiding (LDH) methods have attracted more and more attentions for copyright protection of multimedia in lossy environment. One of the important requirements of the robust LDH methods is the reversibility, that is, the host images can be recovered without any distortion after the hidden messages are removed. The reversibility is often guaranteed by the embedding model, which affects the performance of the methods greatly. Another requirement is to have a better robustness, which allows the LDH methods to be well adaptable to the lossless and lossy environment, e.g., JPEG compression. In this paper, we firstly categorize the existing robust LDH methods according to the embedding model, and then make a theoretical analysis on the performance in terms of capacity, invisibility and the robustness. Finally, experimental comparisons are carried out to summarize the advantages and dis-advantages of each kind of method.

KEYWORDS: Lossless data hiding, robust, reversibility, histogram.

1. INTRODUCTION

Lossless data hiding (LDH) has been widely studied as a popular and powerful technique to protect copyright in many sensitive scenarios, e.g., medical diagnosis, remote sensing and law enforcement [1]. However, most LDH methods are strictly constrained to a lossless environment to convey the hidden messages, and thus they are not robust for the real-world applications. In other words, a slight change of the stego multimedia, e.g., JPEG compression, can make the correct recovery of the messages impossible. The critical problem can be well solved by recently proposed robust LDH methods [3] that have attracted more and more attentions for copyright protection and content authentication of multimedia.

The key issues of the robust LDH methods are how to design a lossless embedding mechanism to ensure the reversibility, and how to construct the invariant features to achieve the robustness against the attacks, which influence the performance of the methods significantly. According to the embedding mechanism, the existing robust LDH methods can be classified into two categories: 1) histogram rotation (HR)-based; 2) histogram distribution constrained (HDC) embedding methods. In [3], De Vleeschouwer et al. proposed the first robust LDH method by utilizing the grayscale histogram rotation and modulo-256 operation. Experimental results show that this method achieves a good robustness against JPEG compression, the salt-and-pepper noise in the stego images, however, is recognized as the major disadvantage of it, as pointed out by De Vleeschouwer himself. Aiming to remedy this drawback, Zou et al. [9][10] and Ni et al. [7][8] designed a novel embedding mechanism by applying the HDC embedding strategy in the wavelet and spatial domain, respectively. Thereafter, Gao et al. [6] developed an improved version of the HDC embedding aiming at the incomplete reversibility of Ni's method. Although the HDC mechanism is helpful to improve the visual quality of the stego images, it also limits the capacity of this kind of method. To sum up, the aforementioned methods provide a good foundation for the research on the robust LDH methods; however, there is scarcely any evaluation and comparison of them in detail up to now. Therefore, we target the problem and make a comprehensive analysis and evaluation of the above two kinds of robust LDH methods. The contribution of this paper focuses on the following aspects:

1) Based on the efforts for studying the typical robust LDH methods, we firstly summarize the embedding models of different kinds of methods with the illustrations to make the explanation visual and explicit;
2) Secondly, the capacity, invisibility and robustness, which are the main criterions of the robust LDH methods, are analyzed theoretically;

978-1-4244-6829-4/10/$26.00 ©2010 IEEE

3) Thirdly, the important part of this paper is providing the empirical comparisons to indicate the advantages and disadvantages of different methods experimentally.

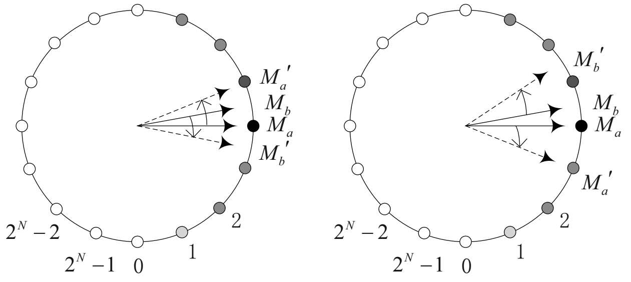
The rest of this paper is organized as follows. Section 2 and Section 3 discuss the first and second kinds of robust LDH methods in detail, respectively. In Section 4, the experimental results and analysis are presented. Finally, Section 5 is a conclusion.

2. HR-BASED EMBEDDING METHOD

As mentioned above, the robust LDH methods consist of the HR-based and HDC embedding methods. In this Section, we focus on the HR-based method and expatiate on the work proposed by De Vleeschouwer et al. [3], which is recognized as the start-up in the field of the robust LDH.

2.1. Embedding Model

Based on the patchwork theory [2], De Vleeschouwer et al. [3] developed a HR-based embedding model. Fig. 1 shows the main idea of this model when the watermarks, i.e., 0 and 1, are embedded.



(a) Bit ‘0’ embedding (b) Bit ‘1’ embedding

Figure 1. Embedding Model of HR-Based Embedding Method

Consider an N -bit image with size of $H \times W$, we divide it into the non-overlapping blocks, which are associated with the watermarks one by one. For each block, two pseudo-random sets of pixels, i.e., zones A and B are generated. Then the grayscale histogram of each zone is mapped into a circle wherein the positions on the circle are indexed by the grayscale values and the weight of each position is represented by the number of pixels corresponding to the grayscale value. In Fig. 1, we use different colors, i.e., light gray, gray, and black, to indicate different weights. According to the distribution of the weights, the centers of the mass of zones A and B , denoted as M_{a} and M_{b} , are computed, as shown in Fig. 1. Let V_a and V_b represent the vectors pointing from the center of the circle to M_{a} and M_{b} , the watermarks embedding is thus achieved by slightly rotating them. To be specific, when the watermark bit is 0, V_a and V_b are rotated anti-clockwise and clockwise by the same angle, respectively; otherwise, they are rotated clockwise and anti-clockwise to embed bit “1”. Since the pixels in a given region, e.g., a block, are highly correlated, and the zones A and B are generated randomly, M_{a} and M_{b} are close to each other. In other words, the slight rotation can allow M_{a} and M_{b} to form a specified difference between them. In the receiver side, the sign of the difference is utilized to extract the watermark bits while the magnitude of it indicates the rotated angles, that is, the changed amount of the grayscale values of the block. In the rest of this paper, we refer to the changed amount as the embedding level. After the hidden messages are extracted, the stego centers of the mass, denoted as M_{a}' and M_{b}' in Fig. 1, can be rotated back to their original ones. That is, the host images can be recovered without any distortion and thus the reversibility is achieved. Meanwhile, the modulo-256 operation is applied to prevent overflow and underflow, which may result in the wrapped around pixels and thus the salt-and-pepper noise in the stego images.

In summary, the HR-based embedding as well as the modulo-256 operation are the key factors guaranteeing the reversibility and robustness in [3]. In the next subsection, we will make an analysis on the performance of this kind of method in terms of capacity, invisibility and robustness.

2.2. Analysis

In practical images, the centers of the mass of zones A and B in some blocks may differ greatly, which will lead to the failure of the aforementioned embedding model. To tackle this problem, De Vleeschouwer et al. defined the problematic blocks which were not used to embed watermarks. Based on this, the pure capacity can be denoted as

$$C \approx \left\lfloor \frac{H}{h} \right\rfloor \times \left\lfloor \frac{W}{w} \right\rfloor - \gamma , \tag{1}$$

where $h \times w$ is the block size and γ represents the number of the problematic blocks. Since γ is often small, the pure capacity is mainly determined by the number of the regular blocks. The bigger the block is, the lower the capacity is, and vice versa. Noted that $\approx$ instead of $=$ in Eq. (1) is used because the information of some specified problematic blocks are needed to be embedded into the images as well as the watermarks, which will decrease the pure capacity a little.

The invisibility, which is used to evaluate the visual quality of the stego image, depends on the rotated angle of the vectors, i.e., the embedding level in [3]. Given a fixed embedding level, the invisibilities are approximately same, even if different block sizes are employed. This point is sufficiently demonstrated by the experimental results in Section 4.

The invariant features, which serve as a key role in the robust LDH methods, are employed to ensure the robustness. It can be observed that the good stability of the difference between M_{a} and M_{b} allows this method to be robust against the high quality JPEG compression. Experimental comparisons with other robust LDH methods will further verify this superiority.

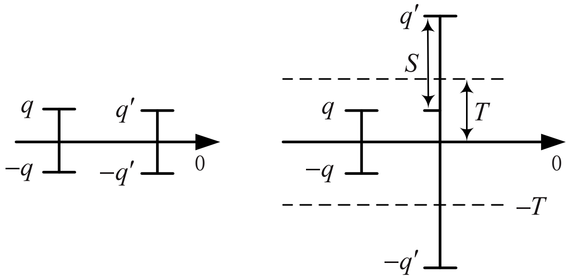
3. HDC EMBEDDING METHOD

In this section, the second kind of the robust LDH method, i.e., the HDC embedding methods, is summarized and analyzed. In our analysis, the works proposed by Zou et al. [9][10], Ni et al. [7][8] and Gao et al. [6] fall in this category. By studying the ideas of them, we propose a comprehensive embedding model, which will be discussed in the next subsection followed by the theoretical analysis.

3.1. Embedding Model

As aforementioned, the method in [3] suffers from serious salt-and-pepper noise, which degrades the visual quality of the stego images greatly. To overcome the disadvantage, a novel embedding mechanism, called as HDC embedding, was developed in [7][8][9][10]. This mechanism utilizes the statistical characteristics of the images, e.g., arithmetic average of differences between grayscale values within a block in [7][8] and the mean value of wavelet coefficients in [9][10], to embed the messages. In this paper, we refer to these objects as the embedding carrier, denoted as q . Based on this, the comprehensive embedding model can be illustrated in Fig. 2.

According to this model, the embedding can be achieved by modifying the embedding carrier q or keeping it intact. Specifically, q will be increased or decreased by the embedding level, S , to embed bit "1" while it is kept intact to embed bit "0". In the extraction process, a threshold, T , is applied to distinguish the different watermark bits by comparing it with the stego embedding carrier, q' . As shown in Fig. 2, S is equal to or greater than T (e.g., $S = 2*T$ in [8]), it is possible, therefore, to apply this condition to extract the watermark bits correctly. Following this step, the original q can be obtained by modifying q' with the same amount as the embedding process in opposite direction. That is, the host images can be recovered losslessly.



(a) Bit '0' embedding (b) Bit '1' embedding

Figure 2. Embedding Model of HDC Embedding Method

It is noticed that the histogram distributions of different blocks differ from each other, the above embedding model, therefore, may result in the overflow and underflow of pixels, which is absolutely not permitted in the LDH methods. To target the problem, the histogram distributions of the blocks as well as the to-be-embedded bits are considered at the same time. Taking the method in [10] as an example, if $q < 0$ and the grayscale values under consideration are close to the lower boundary of the grayscale histogram i.e., 0, the underflow may occur. In this case, q is kept intact no matter whether the to-be-embedded bit is 0 or 1. This is the reason why we refer to this kind of method as the histogram distribution constrained (HDC) embedding method. Obviously, this strategy handle the overflow and underflow successfully, the errors, however, are resulted. As a sequence, the error correction coding (ECC) is needed to correct them, whereas it limits the embedding capacity to a specific range. In the next subsection, we will discuss this issue.

3.2. Analysis

The usage of ECC corrects the errors caused in the embedding process and thus ensures the lossless recovery of the hidden messages. Meanwhile, the permutation technique is employed to deal with the bursts of errors. Except for their positive contributions, these mechanisms decrease the capacity greatly. Suppose $\mathrm{BCH}(n,k,t)$ and the Arnold permutation are adopted here, the maximum of the pure capacity can be formulated as

$$C^* = \left\lfloor \frac{(\lfloor \sqrt{\nu} \rfloor)^2}{n} \right\rfloor \times k, \tag{2}$$

and

$$\nu = \left\lfloor \frac{H^*}{h} \right\rfloor \times \left\lfloor \frac{W^*}{w} \right\rfloor, \tag{3}$$

in which $H^* \times W^*$ represents the size of the embedding region, i.e. the size of the image in [8], and the size of the HL or LH subband in [10]. For instance, the block size is 8×8 and BCH$(31,6,7)$, C^* is only 792 bits for the 512×512 images in [8], which is too small compared with the one in [3].

As for the invisibility, it also depends on the embedding level, S, as same as the HR-based method. However, only half of the pixels in a block are modified in the embedding process in [8], the visual quality of the stego image is higher than that in [3]. Similar cases also occur in [10], which has been demonstrated by the experimental results.

In our analysis, the robustness of this kind of method lies in the fact that the statistical quantity used as the embedding carrier, e.g., the mean value of the wavelet coefficients [9][10], is robust against JPEG compression.

4. EXPERIMENTAL RESULTS

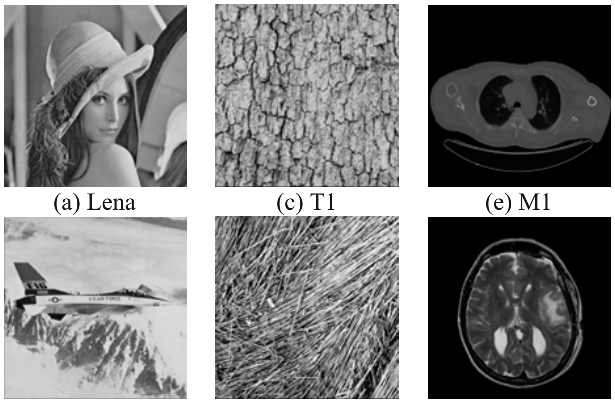

(a) Lena
(c) T1
(e) M1
(b) Airplane
(d) T2
(f) M2

Figure 3. Test Images

In this section, we use two commonly used images, *Lena*, *Airplane*, two texture images and two medical images, as shown in Fig. 3, to build a small data set for comparing the performances of the aforementioned two kinds of the robust LDH methods from two aspects: robustness and the capacity versus the image quality. In order to facilitate empirical comparison, all test images have a fixed size of $512 \times 512 \times 8$, the watermark is the binary sequence generated randomly, and BCH(15,11,1) is utilized in [8][10].

In the robust comparison, JPEG compression is considered here. Similar to that in [3] and [8], we use the major voting to obtain the hidden message of 100 bits, which is embedded into the image repeatedly. Fig. 4 shows the robust comparisons between the aforementioned methods for the six test images, where the block size is 8×8, the embedding levels are 8 in [3] and 16 in [8], and the bit error rate (BER) is utilized to evaluate the robustness. According to that in Fig. 4, it can be seen that the method in [3] outperforms that in [8] and [10] in terms of robustness. When the JPEG quality factor is 50, the hidden messages can be extracted correctly for the images except *T2*. Also observe that Ni's method achieves the similar robustness for the commonly used images and medical images with [3].

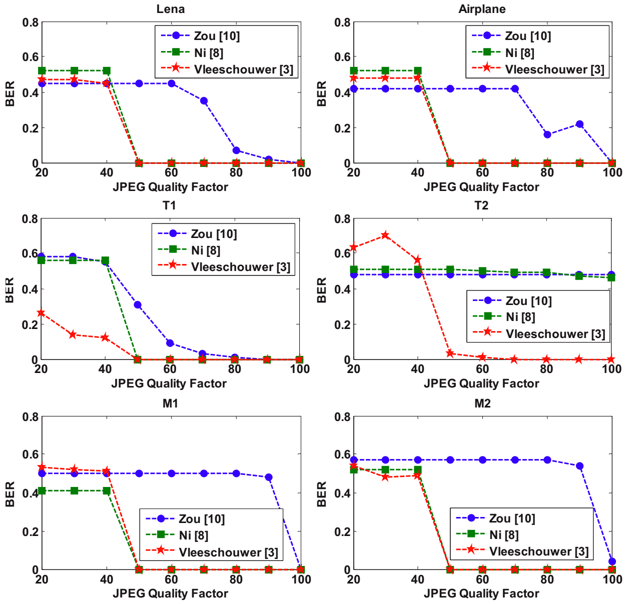


Figure 4. Robustness Comparison

Apart from the robustness, the experiments are carried out for the pure capacity and invisibility evaluation, and the block size ranges from 4×4 to 16×16 with step 4. The experimental results are illustrated in Fig. 5 wherein the peak signal-to-noise ratio (PSNR) is adopted to evaluate the invisibility. On one hand, the pure capacities of three methods can be controlled by adjusting the block size, which is observed in the direction vertical to the *y*-coordinate in Fig. 5. However, the usage of the ECC in the HDC embedding methods reduces their capacities a lot. On the other hand, the method in [3] has the lowest PSNR because of the usage of the modulo-256 operation, especially for the medical images, lower than 10dB.

Meanwhile, the embedding level in [10] is dependent on the image content, that is, the bigger the block is, the smaller the embedding level is. As a sequence, the PSNR of Zou's method is increased with the increase of the block size. Also observe that the PSNRs in [3] and [8] are not varied basically from the block size, this is because the embedding level is 8 at different block sizes.

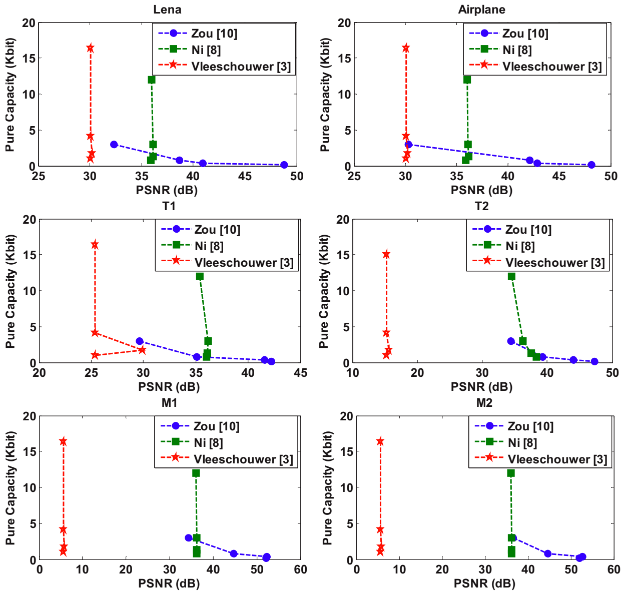


Figure 5. Comparison of Capacity versus Image Quality

In summary, the HR-based methods achieve the best robustness against the JPEG compression and the highest capacities while the lowest visual quality of the stego image. By contrast, the better visual quality is achieved in the HDC embedding methods but the capacity is limited. As for the robustness of the HDC embedding methods, different results are obtained: the better in [8] and the worse in [10].

5. CONCLUSION

In this paper, a comprehensive analysis and evaluation of the robust lossless data hiding (LDH) methods is presented. Based on the embedding mechanism, the existing methods are categorized, and their advantages and disadvantages are summarized by analyzing and comparing them theoretically and experimentally. In the future, it is worthwhile to design the novel robust LDH methods by introducing the more effective and efficient invariant features, e.g., feature points [4][5].

ACKNOWLEDGEMENTS

This research was partially supported by the National Natural Science Foundation of China (60771068, 60702061, 60832005, U0835004), the Ph.D. Programs Foundation of Ministry of Education of China (No. 20090203110002), the Natural Science Basic Research Plan in Shaanxi Province of China (2009JM8004), the Open Project Program of the National Laboratory of Pattern Recognition (NLPR) in China and the National Laboratory of Automatic Target Recognition (ATR), Shenzhen University, China.

REFERENCES

[1] L. An, X. Gao, C. Deng, and F. Ji, "Reversible watermarking based on statistical quantity histogram," *Lecture Notes in Computer Science*, vol. 5879, pp. 1300-1305, 2009.

[2] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3-4, pp. 313-336, 1996.

[3] C. De Vleeschouwer, J. F. Delaigle, and B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 97-105, 2003.

[4] C. Deng, X. Gao, D. Tao, and X. Li, "Geometrically invariant watermarking using affine covariant regions," IEEE Int. Conf. on Image Processing, San Diego, CA, USA, pp. 413-416, 2008.

[5] C. Deng, X. Gao, X. Li, and D. Tao. "A local Tchebichef moments-based robust image watermarking," *Signal Processing*, vol. 89, no. 8, pp. 1531-1539, 2009.

[6] X. Gao, L. An, X. Li, and D. Tao, "Reversibility improved lossless data hiding," *Signal Processing*, vol. 89, no. 10, pp. 2053-2065, 2009.

[7] Z. Ni, Y. Shi, N. Ansari, W. Su, Q. Sun, and X. Lin, "Robust lossless image data hiding," IEEE Int. Conf. on Multimedia Expo, Taipei, Taiwan, pp. 2199-2202, 2004.

[8] Z. Ni, Y. Shi, N. Ansari, W. Su, Q. Sun, and X. Lin, "Robust lossless image data hiding designed for semi-fragile image authentication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 4, pp. 497-509, 2008.

[9] D. Zou, Y. Shi, and Z. Ni, "A semi-fragile lossless digital watermarking scheme based on integer wavelet transform," IEEE 6th Workshop on Multimedia Signal Processing, Siena, Italy, pp. 195-198, 2004.

[10] D. Zou, Y. Shi, Z. Ni, and W. Su, "A semi-fragile lossless digital watermarking scheme based on integer wavelet transform," *IEEE Trans. Circuits Syst. Video Technol.* vol. 16, no. 10, pp. 1294-1300, 2006.