LOSSLESS DATA HIDING: FUNDAMENTALS, ALGORITHMS AND APPLICATIONS

Yun Q. Shi[1]*, Zhicheng Ni*[1]*, Dekun Zou*[1]*, Changyin Liang*[2] *and Guorong Xuan*[3]

[1]New Jersey Institute of Technology, Newark, NJ, USA, shi@njit.edu
[2] Shenzhen Polytechnic, Shenzhen, China
[3]Tongji University, Shanghai, China

ABSTRACT

Recently, among various data hiding techniques, a new subset called lossless data hiding has drawn tremendous interests. By lossless data hiding, it is meant that the marked media can be reversed to the original cover media without any distortion after the hidden data are retrieved. After a careful study of all lossless data hiding algorithms published up to today, we classify the existing algorithms into three categories: 1) Those developed for fragile authentication; 2) Those developed aiming at large embedding capacity; 3) Those developed for semi-fragile authentication. The mechanisms, merits, drawbacks and applications of these algorithms are analyzed, and some future research issues are addressed in this paper.

Key words—lossless data hiding, reversible data hiding, robust lossless data hiding, watermarking.

1. FUNDAMENTALS

Data hiding has recently been proposed as a promising technique for the purpose of information assurance, authentication, fingerprint, security, data mining, and copyright protection, etc. By data hiding, pieces of information represented by some data are hidden in a cover media. Many image data hiding algorithms have been proposed in the past several years. In most cases, the cover media will experience some permanent distortion due to data hiding and cannot be inverted back to the original media.

From the literature, it is observed that most of the current data hiding algorithms are not lossless. For instance, with the most popularly utilized spread-spectrum watermarking techniques, either in DCT domain [1] or block 8x8 DCT domain [2], *round-off* error and/or *truncation* error may take place during data embedding. As a result, there is no way to invert the marked media back to the original without distortion. With the popularly used least significant bit-plane (LSB) embedding method, the bits in the LSB are replaced according to the data to be embedded and the bit-replacement is not *memorized.* Consequently, the LSB method is not invertible. With another group of frequently used watermarking techniques, called quantization index modulation (QIM) [3], *quantization* error makes lossless data hiding impossible.

In some applications, such as in the fields of law enforcement, medical and military image systems, in addition to perceptual transparency it is desired to reverse the marked media back to the original cover media after the hidden data are retrieved for some legal or other considerations. The marking techniques satisfying this requirement are referred to as *reversible*, *lossless*, *distortion-free*, or *invertible* data hiding techniques.

Some lossless marking techniques have been reported in the literature over the last a few years. A classification of all existing algorithms into three categories is presented in Section 2. Some future research issues are discussed in Section 3. Conclusions are made in Section 4.

2. THREE CATEGORTIES OF EXISTING LOSSLESS DATA HIDING ALGORTIHMS

In this section, all existing lossless data hiding algorithms published in the literature up to today are classified into the following three categories.

.

2.1. Those for fragile authentication

The first several lossless data hiding algorithms belong to this category. Since fragile authentication does not need much data to be embedded in a cover media, the embedding capacity in this category is not large. Normally from 1k to 2k bits.

Barton's patent in 2000 [4] may be the earliest one. His algorithm was developed for authentication of digital media, including JPEG and MPEG coded image and videos. The main idea is to losslessly compress the bits to be overlayed and leave space for authentication bit-string. No specific performance result has been reported.

Honsinger et al.'s patent in 2001 [5] is the second lossless data hiding technique used for fragile authentication. Their method is carried out in the image spatial domain by using modulo 256 addition. In the embedding, $Iw = (I + W)$ mod 256, where Iw denotes marked image, I original image, W is the payload comes from the hash function of the original image. In the authentication side, the payload W can be reconstructed from the marked image, then subtract the payload from the marked image to losslessly recover the original image. By using modulo 256, over/underflow is avoided. However, the marked image may suffer from the salt-and-pepper noise during possible grayscale flipping over between 0 and 255 in either direction due to modulo 256 addition. This issue will be addressed later in this paper.

Fridrich's group explored a deep research on lossless data hiding techniques and developed some algorithms. Their first algorithm [6] is in the spatial domain, which losslessly compresses the bit-planes to leave room for data embedding. In order to leave sufficient room for data embedding, it needs to compress a relatively high level bit-plane, which usually leads to visual quality degradation. They also describe two reversible data hiding techniques [7] for lossy compressed JPEG image. The first technique is based on lossless compression of biased bit-streams derived from the quantized JPEG coefficients. The second technique modifies the quantization matrix to enable lossless embedding of one bit per DCT coefficient. In addition, Fridrich's group extended the idea of lossless authentication to MPEG-2 video [8].

2.2 Those for high embedding capacity

All the above mentioned techniques aim at fragile authentication, instead of data hiding. As a result, the amount of hidden data is rather limited. Hence, Goljan et al. [9] presented a first lossless marking technique that is suitable for data embedding. The details are as follows. The pixels in an image are grouped into non-overlapped blocks, each consisting of a number of adjacent pixels. For instance, it could be a horizontal block having four consecutive pixels. A discrimination function is established to classify the blocks into three different categories, Regular, Singular and Unusable. (The authors use the discrimination function to capture the smoothness of the groups.) An invertible operation F can be applied to groups. That is, it can map between a pair of gray level values, resulting in $F(R)=S$, $F(S)=R$, and $F(U)=U$. It is reversible since applying it to a gray level value twice produces the original gray level value. This invertible operation is hence called *flipping F*. The main idea for lossless embedding is that they scan the image group-by-group and losslessly compress the status of the image − the bit-stream of R and S groups (the RS-vector) with the U groups simply skipped − as overhead to leave room for data embedding. By assigning a 1 to R and a 0 to S they embed one message bit in each R or S group. If the message bit and the group type do not match, the flipping operation F is applied to the group to obtain a match. The data to be embedded consist of the overhead and the watermark signal. While it is novel and successful in reversible data hiding, the amount of data that can be hidden by this technique is still not large enough. From what is reported in [9], the estimated capacity ranges from 0.011 to 0.092 bits per pixel (bpp). This may not be high enough for some applications. Another problem with the method is that when the capacity increases, the visual quality will drop severely.

Xuan et al. proposed a high capacity lossless data hiding technique based on the integer wavelet transform (IWT) [10]. IWT is used in the algorithm to ensure the lossless forward transform and inverse transform. After IWT, the bias between '1' and '0' in the middle and high bit-planes of IWT coefficients becomes much larger than that in spatial domain. Hence, those coefficient bit-planes can be losslessly compressed to leave a large space for data embedding. Histogram modification is used in this algorithm to prevent the over/underflow problem. The experiments show that the capacity can reach 0.057 to 0.36 bpp, quite larger than the previous algorithms. While the PSNR of marked images are not very high due to histogram modification, there is no any annoying artifact and the visual quality is satisfying. Further improvement has been made, resulting in both higher embedding capacity and visual quality.

Ni et al. [11] proposed a new lossless data hiding technique based on the histogram modification. This algorithm utilizes the zero or minimum points of the image histogram and modifies the pixel value to embed the data. In the image histogram, they first find a *zero point* (no pixel assumes that gray value) and a *peak point* (a maximum number of pixels assume that gray value). Then they move the histogram between zero point and peak point toward zero point by one unit and leave the histogram near the peak point empty. Finally the histogram in peak point is moved to its neighbor or kept intact to finish the embedding of '1' or '0'. This algorithm has a quite large capacity (0.019 to 0.31 bpp) while keeping a very high visual quality for all images (the PSNR of marked images versus original images is guaranteed to be higher than 48 dB). This PSNR performance is much higher than any other algorithms at the same data embedding rate. .

Celik et al. [12] presented a high capacity, low distortion reversible data hiding technique. In the embedding phase, the host signal is quantized and the residual is obtained. Then they adopt the CALIC lossless image compression algorithm, with the quantized values as side information, to efficiently compress the quantization residuals to create high capacity for the payload data. The compressed residual and the payload data are concatenated and embedded into the host signal via generalized-LSB modification method. The experimental results show that the PSNR and capacity are satisfying.

Tian recently presented a new high capacity reversible data embedding algorithm in [13]. This algorithm employs two techniques, difference expansion and generalized least significant bit embedding, to achieve a high embedding capacity, while keep the distortion low. The reported results of 'Lena' image is shown in Table 1. It seems until now this algorithm reaches the highest capacity.

Table 1. Payload size vs PSNR of embedded Lena image.

Payload size (bits)	39k	82k	118k	172k	254k	369k
Bit rate(bpp)	0.15	0.32	0.46	0.67	0.99	1.44
PSNR(dB)	44.2	41.6	37.7	34.8	29.4	24.0

The main idea is described as follows. For a pair of pixel values x and y , they first compute the integer average l and difference h of x and y , where $h = x - y$. Then they shift h to the left one unit and append information bit b in LSB. That equals $h' = 2 \times h + b$. This method is called Difference Expansion. Finally they compute the new x and y , based on the new difference values h' and the original integer average value l . In this way, the marked image is obtained. They only embed information bits into the pixel pairs that cannot lead to overflow problem to overcome this overflow issue.

In summary, several lossless data hiding algorithms having large embedding capacity have been developed. Method in [13] may achieve the largest capacity, while method in [11] may keeps the highest PSNR. If the embedding capacity of the improved method in [10] developed in IWT is comparable to that in [13] is currently under investigation.

2.3. Those for semi-fragile authentication

For multimedia, content-based authentication makes more sense than representation-based authentication. The former is often called semi-fragile, while the latter fragile authentication. It is because semi-fragile authentication allows some incidental modification, say, compression within a reasonable extent, to be authentic. For this purpose, we need lossless data hiding algorithms robust to compression. That is, if the image has not been changed at all, it will be authentic. In addition, the original image will be able to recover. If the image has been compressed, the original image cannot be recovered, but the hidden data can still be recovered without error.

De Vleeschouwer et al. [14] proposed a lossless data hiding algorithm based on patchwork theory, which has certain robustness against JPEG lossy compression. This is the only existing robust lossless data hiding algorithm against JPEG compression. That is, each bit of the message is associated with a group of pixels, e.g. a block in an image. Each group is equally divided into two pseudo-random sets of pixels, i.e. zones A and B. The histogram of each zone is mapped to a circle (positions on the circle are indexed by the corresponding luminance). It is observed that in most cases the vectors pointing to the mass center of zones A and B are close to each other as shown in Figure 1 (a) and (b).. Hence slight rotation of these vectors in two opposite directions allows for embedding a bit. Embedding a binary 1 is shown in Figure 1 (c) and (d). As to the pixel values, rotations of the vectors correspond to luminance shifts. From Figure 1, it is obvious, modulo 256 addition is used. Therefore this algorithm suffers from the salt-and-pepper noise. More investigation in this regard is presented in Section 3.

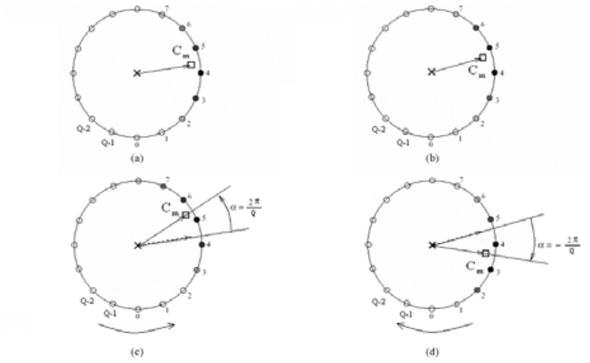

Figure 1: Data embedding diagram

3. SOME FUTURE RESEARCH ISSUES

As pointed in Section 2, De Vleeschouwer et al's method is the only existing lossless data hiding technique, which is robust to high quality JPEG compression. Hence it can be used for semi-fragile authentication. However, our extensive investigation reveals that it has some drawbacks that prevented it from practical usages. .

First of all, some marked images suffer from salt-and-pepper noise because the algorithm utilizes modulo 256 addition. That is, in doing modulo 256 addition, a very bright pixel with a large gray value close to 256 will be possibly changed to a very dark pixel with a small gray value close to 0, and vise versa. The salt-pepper noise becomes severe for some images, where more dark and bright pixels are contained. A typical example is medical images. We tried eight medical images and found that five images suffer from severe salt-pepper noise, and other three result in some salt-pepper noise. Figure 2 gives such an example of severe noise. Note that the slat-pepper noise becomes so "dense" that the original name of salt-pepper appears not to be appropriate.

Not only for medical images, the salt-pepper noise may be severe for color images as well. We have applied De Vleeschouwer et al's method to eight JPEG2000 test color images. Four images result in severe salt-pepper noise, while other four less severe salt-pepper noise. Figure 3 presents an example of severe case. This is a JPEG2000 test image. The data for authentication is embedded into the red color component. We can observe severe salt-pepper noise, which manifests itself as severe color distortion, i.e., about half of hair area becomes red and most of palm area of right hand turns out to be green.

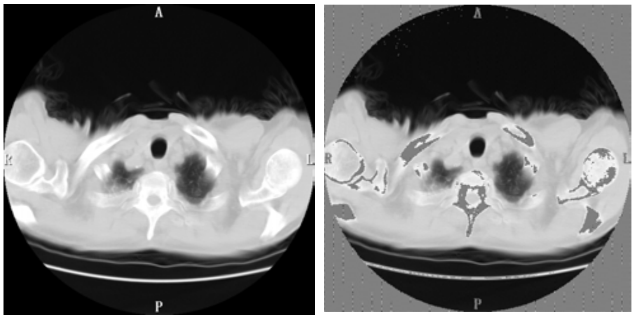

Figure2. (a) Original Mpic1 image , (b) Marked image.

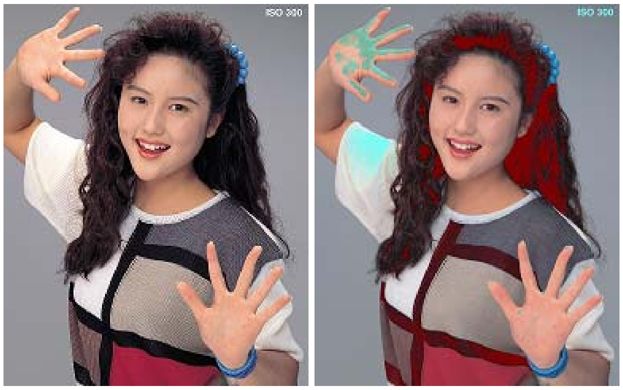

Figure 3. (a) Original Woman image (b) Marked image.

Secondly, the marked image does not have high enough PSNR. Table 2 contains test results for eight medical images. The PSNR of marked image is as low as 26 dB (as 476 information bits are embedded in image of 512x512x8). Note that the salt-pepper noise exists in each of eight marked medical images. When severe salt-and-pepper noise exists, the PSNR may drop to below 10 dB.

Table 2. Test results of eight medical images.

	PSNR (dB)	Robustness (bpp)	Salt-pepper noise
Mpic 1	9.28	1.0	severe
Mpic 2	4.73	2.0	severe
Mpic 3	26.38	0.8	some
Mpic 4	26.49	0.6	some
Mpic 5	26.49	0.6	some
Mpic 6	5.60	1.6	severe
Mpic 7	9.64	0.8	severe
Mpic 8	5.93	2.8	severe

Table 3 contains test results for eight JPEG2000 test images. The PSNR of marked image versus the original image is less than 30 dB (as 1412 information bits are embedded into a color image of 1536x1920x24). Note that the salt-pepper noise again exists in each of the eight test images. When severe salt-pepper noise takes place, the PSNR can be as low as less than 20 dB.

Table 3. Test results with eight JP2000 color images.

	PSNR (dB)	Robustness (bpp)	Salt-pepper noise
N1A	17.73	0.8	severe
N2A	17.73	2.2	severe
N3A	23.73	0.6	some
N4A	19.67	1.2	some
N5A	17.28	1.2	severe
N6A	23.99	0.6	some
N7A	20.66	1.4	some
N8A	14.32	1.4	severe

Thirdly, it is noted that in both Tables 2 and 3 there is an item called Robustness in the unit of bpp. This means that the hidden data can be error-freely retrieved when an image compression ratio is above this quantity. We noticed, for instance, in Table 3, the JPEG2000 test image called N6A corresponds to 0.6 bpp, meaning that when N6A goes through a compression, which is above 0.6 bpp then the hidden data by using De Vleeschouwer et al.'s method will be able to recover without error. However, this robustness is not strong enough for JPEG2000 compressed images. The reason is JPEG2000, due to its superiority, can compress an image to 0.0xx bpp while the compressed image is still in good quality for some applications. Therefore, a robustness of 0.6 bpp is not enough for semi-authentication of JPEG2000 compressed images.

It is noted that a method has been presented to eliminate salt-pepper noise in [14]. As the authors claimed [14], however, this method is not robust to JPEG compression.

Therefore, it is necessary to develop new robust lossless data hiding technologies that 1) do not use modulo 256 addition; 2) consequently, they will not generate the salt-pepper noises and the marked images have high PSNR versus the original image; 3) the robust lossless data hiding techniques are robust enough so that they can work for semi-fragile authentication of JPEG2000 compressed images.

4. CONCLUSION

This paper presents a thorough investigation on the development of all existing lossless data hiding techniques. These techniques are classified into three different categories. The mechanism, merits and drawbacks of each category are discussed. The future research issues on robust lossless data hiding algorithms, which may find wide applications in semi-fragile authentication of JPEG2000 compressed images, are addressed. It is expected that lossless data hiding may open a new door to link two groups of data: cover media data and embedded data. They will be applied to information assurance such as authentication, secure medical data system, intellectual property protection.

ACKNOWLEDGMENT

The work on this paper was partially supported by the Digital Data Embedding Technologies group of the Air Force Research Laboratory, Rome Research Site, Information Directorate, Rome NY, under contract F30602-03-1-0264. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Air Force Research Laboratory, or the U.S. Government.

REFERENCES

[1] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," in *IEEE Trans. on Image Processing*, vol. 6. No. 12, pp. 1673-1687, Dec. 1997.

[2] J. Huang and Y. Q. Shi, "An adaptive image watermarking scheme based on visual masking," *Electronics Letters*, 34 (8), pp. 748-750, 1998.

[3] B. Chen, G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transaction on Information Theory,* vol. 47, no. 4, pp. 1423-1443, May 2001.

[4] J. M. Barton, "Method and apparatus for embedding authentication information within digital data," U.S. Patent 5,646,997, 1997.

[5] C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel, "Lossless recovery of an original image containing embedded data," US Patent: 6,278,791, 2001.

[6] J. Fridrich, M. Goljan and R. Du, "Invertible authentication," *Proc. SPIE, Security and Watermarking of Multimedia Contents*, pp. 197-208, San Jose, CA, January 2001.

[7] J. Fridrich, M. Goljan and R. Du, "Invertible Authentication Watermark for JPEG Images," *ITCC 2001,* Las Vegas, Nevada, pp. 223-27, April 2001.

[8] J. Fridrich, Rui Du, Lossless "Authentication of MPEG-2 Video," *Proc. ICIP 2002*, Rochester, NY

[9] M. Goljan, J. Fridrich, and R. Du, "Distortion-free data embedding," *Proceedings of 4th Information Hiding Workshop*, pp. 27-41, Pittsburgh, PA, April 2001.

[10] G. Xuan, J. Zhu, J. Chen, Y. Q. Shi, Z. Ni, W. Su "Distortionless Data Hiding Based on Integer Wavelet Transform," *IEE journal, ELECTRONICS LETTERS,* Volume 38, No 25, pp.1646-1648, Dec.2002

[11] Z. Ni, Y. Q. Shi, N. Ansari and W. Su, "Reversible Data Hiding," *IEEE International Symposium on Circuits and Systems,* Bangkok, Thailand, May 2003.

[12] M. Celik, G. Sharma, A.M. Tekalp, E. Saber, "Reversible data hiding," in *Proceedings of the International Conference on Image Processing 2002*, Rochester, NY, September 2002.

[13] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 13, No. 8, August 2003.

[14] C. De Vleeschouwer, J. F. Delaigle and B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management," *IEEE Tran. Multimedia,* vol. 5, pp. 97-105, March 2003.