International Journal of Research in Signal Processing, Computing & Communication System Design
ISSN: 2395 – 3187
Vol: 01, Issue: 01, Jan-June, 2015

A Novel Digital Image Tamper Detection and Recovery by grouping Most Significant Bits

Dr. B. Vijay Kumar
Professor, CSE,
Malla Reddy College of Engineering & Technology, Secunderabad

Abstract - **Chan proposed an image authentication method by producing the parity check bits from pixels whose bits have been rearranged (in 2011) [6]. Due to this rearrangement, the value of the most-significant bit of each tampered pixel can be determined according to its parity check bits. With the help of the most-significant bit of the pixel, the pixel can be recovered by selecting two possible (7, 4) Hamming code words. However, if the distance between two Hamming code words is within a certain range, incorrect selection may occur. Chan's method added one additional bit to indicate the correct one. It is trivial that this may degrade the quality of the authenticated image. This paper groups four most-significant bits into different groups to form a mapping codebook and the mapping codebook is used to produce authentication data instead of the (7, 4) Hamming code book. The experimental results show that the proposed method has a greater ability to recover tampered areas**.

1. INTRODUCTION

The most important property of digital media is that it is easily tampered with. Consequently, media security and copyright protection mechanisms have developed quickly, and many image authentication schemes have been developed. Among these various image authentication schemes, some have been proposed to embed recovered data into the digital images [1]–[4] so that the tampered areas can be detected and recovered. It is trivial that the fineness of the recovered areas is related to the recovery unit of the different methods. However, the recovery unit of Yang and Shen's method [4] is just a block with 4×4 pixels, and that of Lee and Lin's method [2] is just a block with 2×2 pixels.

In order to improve the fineness of the recovered areas, Chan and Chang's method [1] produced parity check bits from each pixel through a hamming code technique [5]. The recovery unit of Chan and Chang's method becomes one pixel. Although Chan and Chang's method can reduce the recovery unit to one pixel, there are some drawbacks in their method. Chan and Chang's method must predict the most-significant bit of each tampered pixel first. The tampered pixel then can be recovered by referring to the predicted bit and the extracted parity check bits. However, once an incorrect prediction is made in the recovery procedure, the tampered pixel cannot be recovered successfully. Furthermore, a pixel with an incorrect prediction may affect the prediction accuracy of the following pixels. On the contrary, in Chan's method [6], the parity check bits are produced from pixels whose bits have been rearranged. The value of the most-significant bit of each tampered pixel can be determined according to its parity check bits. Therefore, the recovery procedure of the proposed method does not need to predict the most-significant bit of the pixel. Although the value of the most-significant bit can be known through the parity check bits, the method still has to predict pixels by selecting one from two candidates of (7, 4) Hamming code words. However, if the distance between two Hamming code words is within a certain range, there is still the possibility of making incorrect predictions. In the new version of Chan's method in [6], one additional bit was added to record the correct one from two candidates.

In this paper, we group four most-significant bits into different groups to form a mapping codebook. The mapping codebook has an important property in that the distancebetween two candidates in each group is always out of the mentioned range. The proposed method will produce authentication data through a mapping codebook instead of a (7, 4) hamming codebook. The experimental results show that the proposed method has better ability to recover the tampered areas with good quality.

The rest of this paper is organized as follows. Chan's method is reviewed in Section 2. The method proposed in this paper will then be presented in Section 3. In Section 4, the experimental results are offered to demonstrate the effectiveness of the proposed method. Finally, the conclusions will be made in Section 5.

2. RELATED WORK

In this section, Chan's method [6] is reviewed. Chan's method contains three procedures: the embedding procedure, the

detection procedure, and the recovery procedure. The embedding procedure produces parity check bits from the rearranged bits of the pixel, and the parity heck bits are embedded into another pixel by using a modulus function [7]. In the detection procedure, the extracted parity check bits are used to check whether the pixel has been tampered with. According to the indication of the tampered pixels, the tampered areas are also located in this procedure. The final procedure, the recovery procedure, recovers the tampered areas. The details of Chan's method are described as follows.

2.1 Embedding Procedure

The embedding procedure contains three steps: Hamming code production, bit rotation, and Torus automorphism [8]. In the first step, Hamming codeproduction, parity check bits are produced from therearranged four most-significant bits of each pixel. Therelation between rearranged bits and parity check bits isshown in Fig. 1. The values of the parity check bits (*P*1, *P*2,*P*3) can be decided by achieving the goal that the number of"1" in each circle should be even. According to the examplein Fig. 1, the values of the parity check bits *P*1, *P*2, and *P*3 are 1, 0, and 0, respectively.

It should be noticed that the parity check bits aregenerated from the reversed four most-significant bits. Thereason why we reverse the bits is that by referring to Fig. 2,the most significant bits of two original data bits are thesame if the values of their parity check bits are the same.

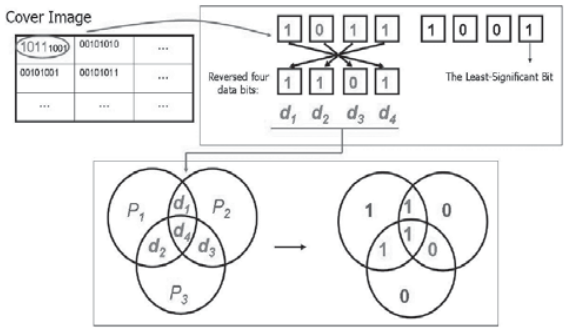


Fig.1. Way to generate the parity check bits from the reversed data bits

That means whether the value of the original data bits islarger than 128 or not can be decided according the value ofthe parity check bits.In the second step the produced parity check bits arerotated. The purpose of bit rotation is to rotate the order ofthe parity check bits to increase security. This step firstselects one secret key *k*1 as a seed to generate a sequence ofrandom numbers, *R*1, *R*2, …,*RN*×*N*, where *N* is the pixelnumber of the height and width of the cover image. For the*i*th pixel, its three parity check bits must be rotatedaccording to the random number *Ri*as follows.

$$J' = (J + R_i) \bmod 3 \qquad (1)$$

where*J* represents the original order of the parity check bits,and *J*denotes the new order of the parity check bits. Thevariables *J* and *J*2 are both in the range from 0 to 2. Aftergoing through (1), the bit at the *J*th position will be rotated to the *J*th position. In the third step, the rotated parity check bits areembedded into the three least-significant bits of anotherpixel indicated by the Torus auto-morphism. The formula of the Torus auto-morphism is shown as follows.

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ k_2 & k_2+1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \bmod N \qquad (2)$$

Where the variable *k*2 represents the second key. The symbol (*xi*, *yi*) denotes the position at where the *i*th pixel P_i *is* located. On the other hand, the position (*xi*2, *yi*2)represents the new position in which the parity check bits of*Pi* are going to be embedded and the symbol *N* is the pixelnumber of the height and the width of the cover image.Once the embedded position for each *Pi* is known, theparity check bits can be embedded by using a modulus

Function [7]. Finally, the authenticated image can beobtained.

2.2 Detection and Recovery Procedures

The purpose of the detection procedure is to locate thetampered areas so that the recovery procedure has targetareas to recover. In the detection procedure, the paritycheck bits for the pixel *Pi* at (*xi*, *yi*) can be extracted fromthe pixel at (*xi*', *yi*') according to (2). Meantime, the paritycheck bits of the pixel *Pi* can be produced through the fourmost-significant bits of the pixel *Pi*. Once the extractedparity check bits are not the same as the produced bits, both positions (*xi*, *yi*) and (*xi*', *yi*') are marked as the tampered pixels. After checking all pixels, morphological operations[9] are performed to eliminate the isolated faulty judgments. Finally, the tampered areas can be located through the detection procedure.

In the recovery procedure, if the parity check bits of the tampered pixel are not modified, these bits can be used topredict the value of the tampered pixel. According to the value of the parity check bits, the value of the most-significant bit, *d*4, can be known by referring to Fig. 2. Moreover, the two candidates of the four most-significantbits can also be known according to Fig. 2. The value of the most-significant bit of the pixel can help us to select thecorrect one from two candidates so as to recover the tampered pixel.

More precisely, in the detection procedure, the tampered area has been located. Chan's method uses an indicated matrix **M** to indicate the number of the surrounding un tampered pixels for each tampered pixels. It is obvious that the tampered pixels with a larger number at the corresponding positions of indicated matrix **M** have a better opportunity to select a correct one from two candidates because they have more un tampered pixels for references. Therefore, only the pixels with the largest number of un tampered pixels will be processed in each round. Because the most-significant bit of the predicted pixel can be known by referring to the parity check bits, the un tampered pixels whose most-significant bit is the same as the predicted pixel are gathered to calculate their average value.

In this section, the experimental results are demonstrated. To begin with, the test image, Lena, with 512×512 pixels, is shown in Fig. 5. We first show the authenticated images and their PSNR values in Fig. 6. Because both Chan's method and the proposed method use a modulus function to embed data instead of least

The candidate that has the minimal distance with the average value is selected to recover this tampered pixel. After that, the recovered pixels are marked as un tampered pixels, and the same procedures are performed to recover tampered pixels until all tampered pixels are marked as untampered pixels. The recovery procedure is shown in Fig. 3.

Generate Parity Check Bits

Original Data Bits	Reversed Data Bits	Parity Check Bits	The Decimal Values	Distance Between Two Original Data Bits
0 0 0 0	0 0 0 0	0 0 0	$(0)_{10}$	| 0 – 7 | = 7
0 1 1 1	1 1 1 0	0 0 0	$(0)_{10}$	
1 1 1 0	0 1 1 1	0 0 1	$(1)_{10}$	| 14 – 9 | = 5
1 0 0 1	1 0 0 1	0 0 1	$(1)_{10}$	
1 0 1 0	0 1 0 1	0 1 0	$(2)_{10}$	| 10 – 13 | = 3
1 1 0 1	1 0 1 1	0 1 0	$(2)_{10}$	
0 1 0 0	0 0 1 0	0 1 1	$(3)_{10}$	| 4 – 3 | = 1
0 0 1 1	1 1 0 0	0 1 1	$(3)_{10}$	
1 1 0 0	0 0 1 1	1 0 0	$(4)_{10}$	| 12 – 11 | = 1
1 0 1 1	1 1 0 1	1 0 0	$(4)_{10}$	
0 0 1 0	0 1 0 0	1 0 1	$(5)_{10}$	| 2 – 5 | = 3
0 1 0 1	1 0 1 0	1 0 1	$(5)_{10}$	
0 1 1 0	0 1 1 0	1 1 0	$(6)_{10}$	| 6 – 1 | = 5
0 0 0 1	1 0 0 0	1 1 0	$(6)_{10}$	
1 0 0 0	0 0 0 1	1 1 1	$(7)_{10}$	| 8 – 15 | = 7
1 1 1 1	1 1 1 1	1 1 1	$(7)_{10}$	

Fig. 2. Original data bits and the parity check bits

3. PROPOSED METHOD

In Chan's method[6], although the value of the most-significant bit can be known through the parity check bits, the method still has to predict pixels by selecting one from two candidates. Most pixels can be predicted correctly by using the method described in Section 2. However, there is still the possibility of making incorrect predictions while two candidates are too close. In fact, if the distance between two candidates is small enough, selecting any one of two candidates to recover the pixel is acceptable. The quality of the recovered pixel is still good. However, by referring to the last column in Fig. 2, it can be seen that the distance between two candidates becomes larger when the values of the parity check bits are 2, 5, 1 and 6. Their distances are 3 and 5. In the new version of Chan's method in [6], one additional bit is added to record the correct one from two candidates whose parity check bits are 2, 5, 1, and 6. It is trivial that this may degrade the quality of the authenticated image.

Original Data Bits	The Decimal Values	The Mapping Values	Distance Between two Original Data Bits
0 0 0 0	0	0	| 8 – 0 | = 8
1 0 0 0	8	0	
0 0 0 1	1	1	| 9 – 1 | = 8
1 0 0 1	9	1	
0 0 1 0	2	2	| 10 – 2 | = 8
1 0 1 0	10	2	
0 0 1 1	3	3	| 11 – 3 | = 8
1 0 1 1	11	3	
0 1 0 0	4	4	| 12 – 4 | = 8
1 1 0 0	12	4	
0 1 0 1	5	5	| 14 – 5 | = 8
1 1 0 1	13	5	
0 1 1 0	6	6	| 15 – 6 | = 8
1 1 1 0	14	6	
0 1 1 1	7	7	| 16 – 7 | = 8
1 1 1 1	15	7	

Fig. 3. Recovery procedure

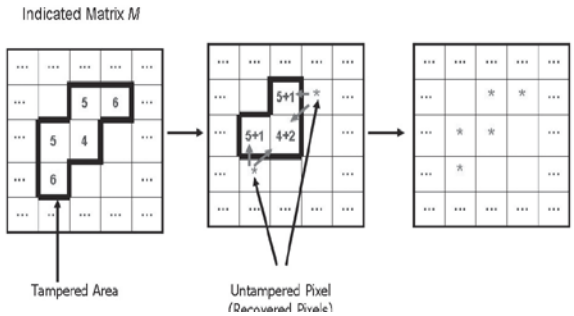


Fig. 4. Mapping Codebook.

In order to improve the quality, we propose grouping four most-significant bits into different groups to form a mapping codebook. There are two characteristics of the mapping codebooks. The data bits with same MSB are grouped together. Thus the value of difference between the bits will be 8. This will be useful while retrieving the tampered pixels as the most significant bit can be easily found out. This is shown in figure4.

In the embedding procedure, the proposed method is almost the same as Chan's method except for the Hammingcode production step. Chan's method obtained authentication data from four most-significant bits byproducing parity check bits. On the contrary, our proposed method obtains authentication data from four most significant bits through the Mapping Codebook. The mapping value can be decided by referring to the fourmost-significant bits and the Mapping Codebook. The role of the mapping value is treated as the parity bits in Chan's method. This means that the mapping value will go through the steps of bit rotation and the rotated bits are embedded to the three least-significant bits of another pixel indicated by Torus automorphism.

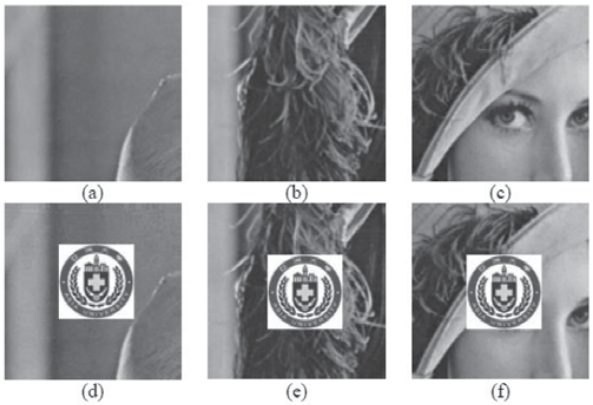
4. EXPERIMENTAL RESULTS

Significant bit (LSB) replacement, the peak signal-to-noise ratio (PSNR) values of Chan's method and the proposed method are higher than that of Chan and Chang's method.

Next, we show the recovery ability for different methods in different areas. There are three areas, including a non-detailed area, a detailed area and a complex area (which includes a non-detailed area and a detailed area), used in our experimental procedures. These areas are shown in Figs. 7 (a), (b) and (c). The tampered areas are shown in Figs. 7 (d), (e), and (f). The size of the tampered area is 64×64 pixels.

Fig. 5. Test image.

Fig. 6. Authenticated images: (a) Chan and Chang's method (PSNR: 38.57), (b) Chan's method (PSNR: 40.02), and (c) proposed method (PSNR: 40.07).



Fig, 7. Experimental areas: (a) to (c): original areas, and (d) to (f): tampered areas.

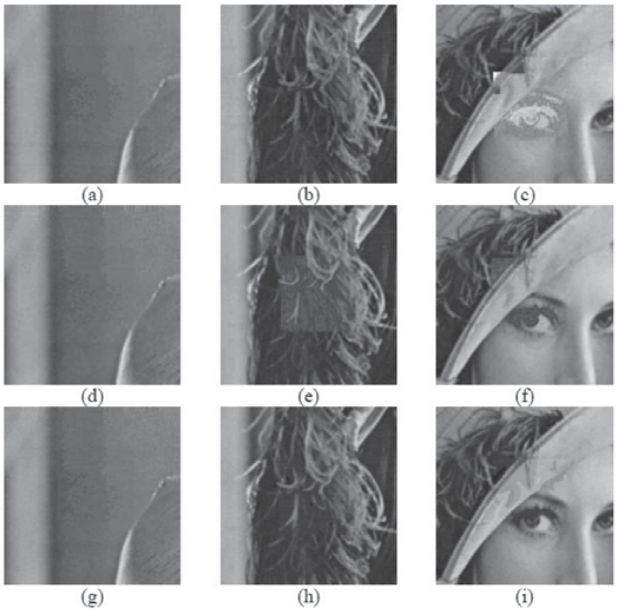


Fig. 8. Results of the recovered areas: (a) to (c): Chan and Chang's method, (d) to (f): Chan's method, and (g) to (i): proposed method.

In Fig. 8 we demonstrate the experimental results of the related methods and our proposed method. Although the recovered areas of Chan and Chang's method shown in Fig. 8 (a) and (b) have good results, the pixels in the recovered area in Fig. 8 (c) are different to the original area. On the other hand, Chan's method [6] and the proposed method can recover the tampered areas successfully. Comparing the experimental results of the proposed method with those of Chan's method, we can see that the recovered areas, by using our proposed method, have better image quality. Because the distance of two candidates for some Hamming code words is in the mentioned range, it is hardfor Chan's method to judge which one is correct. Therefore, the recovered area in Fig. 8 (e) is not very good. On thecontrary, because the distance of the two candidates of each group is out of the mentioned range, the recovered area in Fig. 8 (h) is better than that in Fig. 8 (e). It is trivial that the recovered areas of the proposed method are better than those of Chan's method.

5. CONCLUSIONS

In this paper, we group four most-significant bits into different groups to form a mapping codebook. The mappingcodebook is used to replace the role of (7, 4) Hamming code book in Chan's method. The mapping codebook hastwo important properties. First, only the candidates with the same value of the most significant bit can be gathered in the same group. This means the value of the most-significant bit of the predicted pixel can be determined according to its mapping values. Second, the distance between two candidates of each group is out of the range from 3 to 5. This means one additional bit for some pixels is not necessary in our method. According to the experimental results, the proposed method has a better ability to recover the tampered areas with good quality.

REFERENCES

[1] C.-S. Chan and C.-C. Chang, "An efficient image authentication method based on hamming code," *PatternRecognition*, vol. 40, no. 2, pp. 681"690, 2007.

[2] T.-Y. Lee and S.-D.Lin, "Dual watermark for image tamper detection and recovery," *Pattern Recognition*, vol. 41, no. 11, pp. 3497"3506, 2008.

[3] S.-S. Wang and S.-L.Tsai, "Automatic image authentication and recovery using fractal code embedding and image in painting," *Pattern Recognition*, vol. 41, no. 2, pp. 701"712, 2008.

[4] C.-W. Yang and J.-J.Shen, "Recover the tampered image based on VQ indexing," *Signal Processing*, vol. 90, no. 2, pp. 331"343, 2010.

[5] R. W. Hamming, "Error detecting and error correcting codes," *The Bell System Technical Journal*, vol. 26, no. 2, pp. 147"160, 1950.

[6] C.-S. Chan, "An image authentication method by applying hamming code on rearranged bits," *Pattern RecognitionLetter*, vol. 32, no. 14, pp. 1679"1690, 2011.

[7] C.-C. Thien and J.-C. Lin, "A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function," *Pattern Recognition*, vol. 36, no. 12, pp. 2875"2881, 2003.

[8] G. Voyatzis and I. Pitas, "Chaotic mixing of digital images and applications to watermarking," in *Proc. of EuropeanConf. on Multimedia Applications*, Louvain-la-Neuve, Belgium, 1996, pp. 687"695.

[9] J. Serra, *Image Analysis and Mathematical Morphology*, New York: Academic Press, 1982.