ISSN 2277-3061

Reversible data hiding based on histogram modification of difference images by s-type and Hilbert curve scanning

Jin Chen[1], Ruisong Ye[2]*
[1]School of Software Engineering, South China University of Technology, Guangzhou, 510006, China
[2]Department of Mathematics, Shantou University, Shantou, Guangdong, 515063, China
*Correspongding author, Email：rsye@stu.edu.cn

ABSTRACT

Reversible data hiding technology has currently attracted great attention recently. Using reversible data hiding methods, original cover images can be recovered losslessly from the cover image embedded information. In this paper, spatial reversible data hiding schemes based on histogram modification of difference images are discussed, in which the pixel gray values of difference image are altered slightly to perform the embedding and extraction. Specifically, we design two reversible data hiding schemes based on S-type scanning and Hilbert curve scanning, and compare them with the reversible data hiding scheme by difference between adjacent columns of the cover image. Embedding capacity and PSNR are calculated to figure out how block size, difference manner and embedding level influence the effect of hiding.

Keywords

Reversible data hiding; Difference; Histogram modification

Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol. 15, No. 2

www.ijctonline.com, editorijctonline@gmail.com

INTRODUCTION

With the rapid development of computer network technology and multimedia processing technology, image, video and other digital products are largely distributed, communicated and shared online, which brings great convenience for people's life. However, information disclosure, content tampering, infringement problems become increasingly serious consequently. In order to solve these problems, encryption technology, information hiding technology provides two effective approaches for protecting confidential information, works copyright certification, information authenticity identification, etc. In image processing research field, data hiding is refered to a process to embed the secret data into a cover image. According to the relationship between embedded data and cover image, data hiding is roughly divided into two categories. One kind is called steganography which embedded data has nothing to do with the cover image; another kind is called watermarking technology whose embedded data has a close relationship with cover image. In the digital watermarking technology, the embedded data can provide information for cover image authentication, author signature, etc. Embedded information will affect the cover image quality more or less, but after the watermark extraction, one hopes to be able to fully reconstruct the original cover image, i.e., realizing reversible data hiding. Reversible data hiding has been widely used in sensitive areas, such as medical, military and judicial field [1-4].

Reversible data hiding has become a hot spot in the field of digital watermarking research. Many scholars put forward many kinds of reversible data hiding algorithms recently [5-11]. Reversible watermarking algorithms based on compressing secret data are designed to increase the embedding capacity [5, 6]. Reversible watermarking algorithms based on difference expansion perform the reversible watermark data embedding by extending the difference of adjacent pixels [7]. Reversible data embedding based on histogram translation embeds watermark data in the vacancy of grayscale using statistical characteristic of the image histogram [8,9]. In recent years, combining histogram modification and reversible data hiding algorithms has become one of the important methods in the field of reversible data hiding technology research [10,11]. In 2004, Lee et al. used difference histogram modification to embed secret information data, resulting in higher image quality as well as large capacity [10]. In 2008, Lin et al. proposed a multi-level reversible watermarking algorithm based on difference histogram modification and further improve the embedded information capacity [11]. Because the peak value frequency of the difference image among the columns of the original cover image is greatly improved compared with the cover image, Lee et al. designed a multi-level reversible data embedding algorithm, in which block and embedding level are introduced to improve the embedding capacity and reduce distortion. How to generate the difference image, improve the peak value frequency of difference image, the difference method adopted is a worth considering issue. This paper considers the design of two different difference schemes, s-shaped scanning difference and Hilbert curve scanning difference. As the s-type curve and Hilbert curve can traverse the whole cover image, one can determine the calculation rules of the corresponding difference of adjacent pixels according to the pixel scanned successively. S-scanning and Hilbert curve scanning are used respectively to design reversible hiding algorithms based on difference image histogram modification. The simulation results are compared with those using the column difference method. The influence of block size, differential method and embedding level to embedded capacity and peak signal to noise ratio are analyzed as well.

MULTI-LEVEL SCHEME BASED ON HISTOGRAM MODIFICATION OF DIFFERENCE IMAGE USING CURVE SCANNING

It is common sense that the adjacent pixels's gray values are relatively close for a natural image. There is a large probability that adjacent pixels in a natural image have similar gray values. It follows from this property that the difference between two adjacent pixels in a natural image is expected to be near zero. If the difference is set to be one value in the difference image, then in the obtained difference image, the grayscale value with the maximum number of pixel values tends to be around 0. Furthermore, we can observe that the peak value frequence of the difference image becomes larger that that of that of the original cover image. How to choose adjacent pixels difference methods, however, there is little discussion in the literature. One can choose simple difference operation of adjacent columns or rows [11] , or one can determine the calculation rules of the corresponding difference of adjacent pixels with the help of a specific curve tranversing the whole image. In this paper, we adopt S-scanning and Hilbert curve sanning to calculate the difference image and then design the histogram modification based reversible data hiding scheme. Experimental results are then compared with those of column difference method. As for reversible data hiding scheme based on column difference, please refer to [11] for details.

S-type Scanning Based Reversible Hiding Scheme

Difference method based on S-type scanning converts the two-dimensional image data into one-dimensional data by S-type curve. The hiding phase and the extracting-recovery phase of the proposed scheme are explained in details in the sequel. We assume that the proposed cover image is sized $M \times N$ and let $T = (M \times N)/(A \times B)$. The hiding stage based on S-type scaning is described as follows.

Step 1. The original cover image is divided into samll blocks H_b, each of size $A \times B$. For S-type scanning, the image block pixel values are shown in Figure 1 (the origin is referred to the scanning starting point and the arrow represents the scanning end point). The two-dimensional matrix block H_b is scanned and converted to one vector H'_{bs} with length $A \times B$, and then get the difference image D_{bs} with length $A \times B - 1$ by neighboring pixels gray value subtraction, i.e.,

$$D_{bs}(i)=\left|H'_{bs}(i)-H'_{bs}(i+1)\right|,\ 0\le i\le A\times B-2,\ 0\le b\le T-1. \tag{1}$$

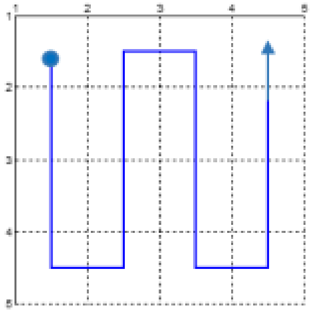


Fig. 1. S-type scanning.

Step 2. Generate the histogram of the difference image D_{bs} and record its peak value P_{bs} for each block.

Step 3. If the pixel value $D_{bs}(i)$ of block *b* is larger than its peak value P_{bs}, change the pixel value $D_{bs}(i)$ to $D_{bs}(i)+1$. Otherwise, the pixel value $D_{bs}(i)$ remains unchanged, which is represented as follows:

$$D'_{bs}(i)=\begin{cases}D_{bs}(i)+1 & \text{if } D_{bs}(i)>P_{bs},\\ D_{bs}(i) & \text{otherwise,}\end{cases}$$

for $0\le i\le A\times B-2$, $0\le b\le T-1$.

Step 4. The pixels in the modified difference image D'_{bs} having grayscale values the same as peak value P_{bs} can be modified to hide embedded message bit $m\in\{0,1\}$ as follows:

$$D''_{bs}(i)=\begin{cases}D'_{bs}(i)+m & \text{if } D'_{bs}(i)=P_{bs},\\ D'_{bs}(i) & \text{otherwise,}\end{cases}$$

for $0\le i\le A\times B-2$, $0\le b\le T-1$.

Step 5. The original scanned cover image H'_{bs} and its hidden difference image D'_{bs} are used to reconstruct the cover image hidden embedded message by performing the following transformation. For the first two pixels, the inverse operation is represented as:

$$S_{bs}(0)=\begin{cases}H'_{bs}(0) & \text{if } H'_{bs}(0)\le H'_{bs}(1),\\ H'_{bs}(1)+D''_{bs}(0) & \text{otherwise,}\end{cases}$$

$$S_{bs}(1)=\begin{cases}H'_{bs}(0)+D''_{bs}(0) & \text{if } H'_{bs}(0)\le H'_{bs}(1),\\ H'_{bs}(1) & \text{otherwise.}\end{cases}$$

For the remaining pixels, the inverse operation is expressed as:

$$S_{bs}(i)=\begin{cases}S_{bs}(i-1)+D''_{bs}(i-1) & \text{if } H'_{bs}(i-1)\le H'_{bs}(i),\\ S_{bs}(i-1)-D''_{bs}(i-1) & \text{otherwise,}\end{cases}$$

for $2\le i\le A\times B-2$, $0\le b\le T-1$.

The generated vector S_{bs} with length $A\times B$ is then reconstructed to be one two-diemnsional matrix block with size $A\times B$. All the blocks are accumulated to be the resulted cover image SH with secret embedded message.

We can extract the embedded message and reconstruct the original cover image losslessly. The process is the reverse of the above-mentioned embedding phase. The detailed extraction and reversing process are stated as follows.

Step 1. The marked cover image SH is divided into samll blocks SH_b, each of size $A\times B$. The two-dimensional matrix block SH_b is scanned according to S-type scanning and converted to one vector S_{bs} with length $A\times B$, and then get the difference image SD_{bs} with length $A\times B-1$ by neighboring pixels gray value subtraction by (2).

$$SD_{bs}(i) = |S_{bs}(i) - S_{bs}(i+1)|,\ 0 \le i \le A \times B - 2,\ 0 \le b \le T - 1. \quad (2)$$

Step 2. Extract the embedded message m on the difference image SD_{bs} of block *b* by

$$m = \begin{cases} 0 & \text{if } SD_{bs}(i) = P_{bs}, \\ 1 & \text{if } SD_{bs}(i) = P_{bs} + 1, \end{cases}$$

for $0 \le i \le A \times B - 2$, $0 \le b \le T - 1$.

Step 3. Remove the embedded message from the difference image SD_{bs} for block *b* using the following formula:

$$SD'_{bs}(i) = \begin{cases} SD_{bs}(i) - 1 & \text{if } SD_{bs}(i) = P_{bs} + 1, \\ SD_{bs}(i) & \text{otherwise,} \end{cases}$$

for $0 \le i \le A \times B - 2$, $0 \le b \le T - 1$.

Step 4. Modify the difference image $SD'_{bs}(i)$ for block *b* to obtain the difference image RD_{bs} to reconstructing the original cover image:

$$RD_{bs}(i) = \begin{cases} SD'_{bs}(i) - 1 & \text{if } SD_{bs}(i) > P_{bs} + 1, \\ SD'_{bs}(i) & \text{otherwise,} \end{cases}$$

for $0 \le i \le A \times B - 2$, $0 \le b \le T - 1$.

Step 5. The inverse difference transformation is applied to get the original scanned cover image RH_{bs} by performing the following formulae. Similarly, for the first two pixels, the inverse operation is represented as:

$$RH_{bs}(0) = \begin{cases} S_{bs}(0) & \text{if } S_{bs}(0) \le S_{bs}(1), \\ S_{bs}(1) + RD_{bs}(0) & \text{otherwise,} \end{cases}$$

$$RH_{bs}(1) = \begin{cases} S_{bs}(0) + RD_{bs}(0) & \text{if } S_{bs}(0) \le S_{bs}(1). \\ S_{bs}(1) & \text{otherwise.} \end{cases}$$

For the remaining pixels, the corresponding inverse operation is expressed as:

$$RH_{bs}(i) = \begin{cases} RH_{bs}(i-1) + RD_{bs}(i-1) & \text{if } S_{bs}(i-1) \le S_{bs}(i), \\ RH_{bs}(i-1) - RD_{bs}(i-1) & \text{otherwise,} \end{cases}$$

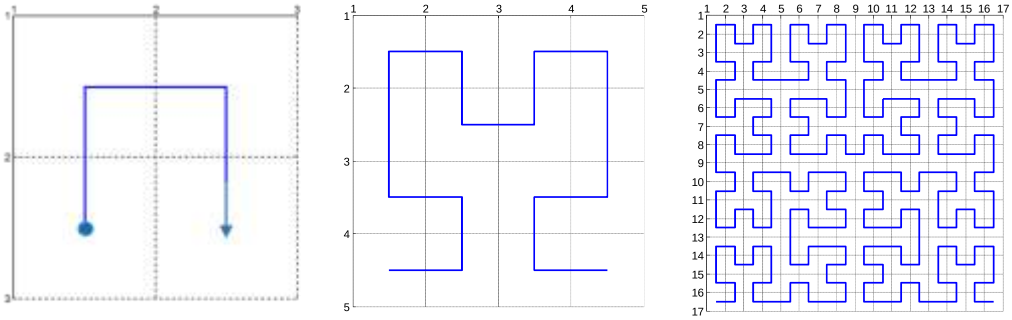
for $2 \le i \le A \times B - 2$, $0 \le b \le T - 1$.

The above process is one round of embedding and extraction process. The algorithm can be implemented several rounds and it is referred to one multi-level algorithm. The marked cover image can be regarded as a new cover image where other secret data can be embedded. As a result, more secret message can be hidden to improve the embedding capacity. For the sake of convenience, we will denote the embedded level as n if the algorithm is performed n rounds.

Hilbert Curve Scanning Based Reversible Hiding Scheme

Thanks to the good filling nature, space-filling curves have been widely applied in many fields, especially in digital image processing, such as image clustering, image encryption, image encoding, image storing, image retrieving, and pattern recognition [12]. The Hilbert curve, one famous space-filling curve presented by Hilbert to traverse points on a two-dimensional square grid, preserves point neighborhoods as much as possible [13]. This paper will use the Hilbert curve to scan square image pixels, and realize the difference histogram modification based reversible data hiding. Based on the Hilbert curve generation algorithm [14], the Hilbert curves with first, second and fourth order are dipicted in Figure 2. The curve starts from buttom left, then scans the pixels in its unique way and ends at bottom right.

Compared with the difference between ajacent columns and S-type scanning based difference, difference method based on Hilbert curve scanning is more complex to describe the image information. The process is a little more complicated than that of S-type scanning based method as well. The image data information is first read according to the Hilbert curve and rearranged to be one-dimensional vector data orderly. Difference operation is then performed for the adjacent pixels to get the difference image. The histogram translation and difference image's modification are next carried out and the inverse operation is finally implemented to reconstruct the marked image. Furthermore, we note that the classical Hilbert curve can traverse the square lattice, and therefore hiding algorithm based on Hilbert curve scanning is only suitable to square images.



order 1 order 2 order 4

Fig. 2. Hilbert curves with orders 1, 2 and 4

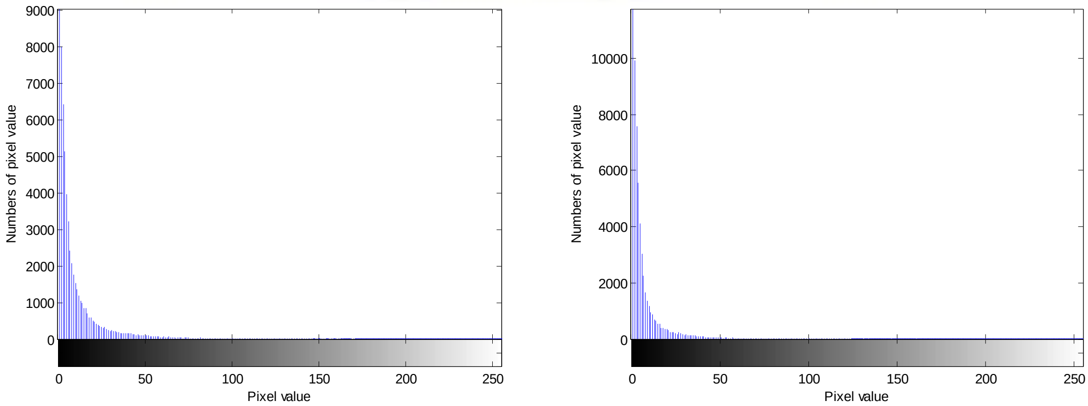
ANALYSIS OF HIDING EFFECT

Comparison of Embedding Capacity Expansion

The peak value frequency of the difference image is worthy of attention which is related to the cover image embedding capacity. Different difference methods will generate different peak value frequencies for the same processed cover image. Four test cover images with size 256×256, Lena.bmp, Rice.png, Baboon.png, and Cameraman.tif (see Figure 3), all 256×256 in size, are used to perform the difference operation respectively. We generate the difference image histogram, and record the corresponding pixel peak value and its frequency for each image. For example, the histograms of the three differnce images for Lena.bmp using column difference, S-type scanning difference and Hilbert curve scanning difference, are shown in Figure 4. The peak value frequency for the four test images and their difference images are demonstrated in Table 1. One can draw the following conclusions from Table 1. 1) No matter what kind of difference method is adopted, the peak value frequency of the difference image is larger than that of the original cover image. Therefore ajacent pixels difference operation is helpful for improving image embedding capacity. 2) The S-type scanning based difference is the most effective method to improve the embedding capacity among the three difference methods. 3) When the peak value frequency of the original cover image is large, the column difference method seems better. 4) Hilbert curve scanning based difference method can usually improve the embedding capacity compared with the column difference method.

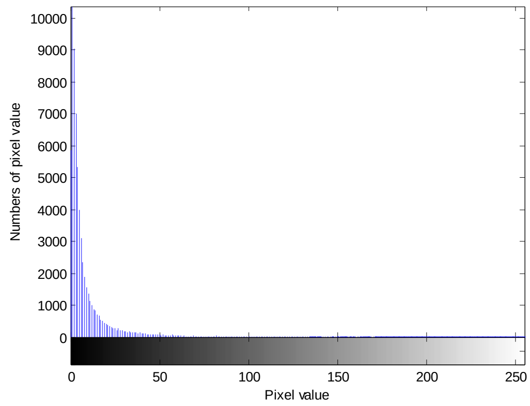
Lana.bmp Rice.png Baboon.png Cameraman.tif

Fig. 3. Origional four test images usesd for simulation



(a) Column difference (b) S-type scanning

ISSN 2277-3061



(c) Hilbert curve scanning

Fig. 4. Histograms of difference image generated by "Lena.bmp" using methods of adjacent column, S-type scanning and Hilbert curve scanning respetively

Table 1. The histogram peak value numbers for four test images and their difference images (Pb, Pb_D, Pb_Ds and Pb_Dh stand for the frequencies of original test images, adjacent column, s-type scanning and Hilbert curve scanning methods respectively)

Test image	Pb	Pb_D	Pb_Ds	Pb_Dh
Lena.bmp	580	9032	11758	10368
Rice.png	1418	6519	8063	7366
Baboon.png	1029	4074	4336	4267
Cameraman.tif	1684	14561	14272	14356

Block Size

In the three reversible data hiding algorithms, the cover image is seperated into small blocks and embedded secret message in the blocks, so as to improve the quality of image after embedded information. To the best of our knowledge, there is less detailed discussion on how the block size affects the marked cover image quality in the literature. To this end, we analyze the specific influence of block size on the embedding result. The original cover image is separated into small pieces with size $A \times B$. Considering the main purpose of histogram modification of the difference image, we should allow more adjacent pixels with similar gray values to be in the same block to create more positions embedded secret message. Note that if the block size is too small, then too many adjacent pixels with similar intensity will be separately divided, and hence it is hard to maintain high embedding capacity. If the block size is too large, the marked image has to afford the visual distortion. Therefore, we only discuss the cases $A, B \in (2, 32]$. In addition, considering the Lena image' size is 256×256, the block size is limited in the those cases shown in Table 2. Assuming that the embedding level is at round 1, we apply the algorithm based on column difference to get the results.

From Table 2, we have the following observations: 1) Under normal circumstances, as the block size varies large, the PSNR value of the marked image declines, as well as embedding capacity. 4×4 is the best block size to get the best embedding effect. 2) Under the condition of equal block area, a square block can well guarantee the marked image quality.

Moreover, similar results like above can be obtained when applying methods of S-type scanning and Hilbert curve scanning based difference.

Embedding Level

Furthermore, we turn to discuss the influence of embedding level. Given the block size 4×4, the results of the image quality and embedding capacity with respect to embedding level by different difference methods are demonstrated in Figures 5-6. It follows from the experimental results that the quality of image degrades as the embedded level increases. The three algorithms have almost the same degree of decline because the image quality is caused by embedded information. The more secret message is embedded, the worse is the image quality. However, the image embedding capacity rate increases as the embedding level increases. The embedding capacities for both S-type scanning and Hilbert curve scanning algorithms are consistent, while that for clolumn difference algorithm is different in some ways. Within the same embedding level, the embedding capacities obtained by S-type scanning and Hilbert curve scanning algorithms are a little larger than that generated by the algorithm based on column difference. The conclusion regarding PSNR is just the reverse.

Table 2. The index values for different block sizes yielded by adjacent column difference method

Block size	PSNR	Embedding capacity	Embedding rate
4×4	48.3599	12784	0.19507
4×8	43.44834	11472	0.175049
8×8	43.3484	10071	0.15367
4×16	39.33462	10615	0.161972
8×16	39.17705	9786	0.149323
4×32	35.57131	10027	0.153
16×16	39.07	9188	0.1402
8×32	35.42179	9523	0.145309
16×32	35.40629	9240	0.140991
32×32	35.36775	9054	0.138153

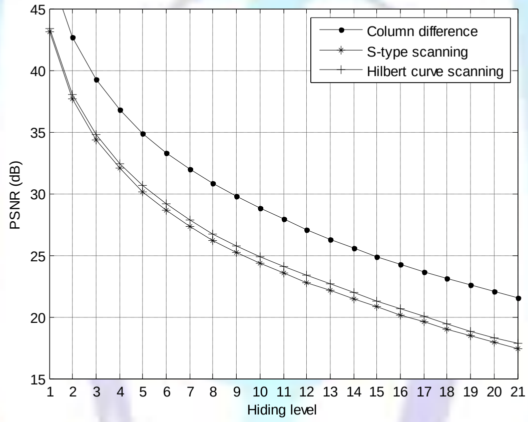


Fig. 5. The PSNR value curves for three methods as block size 4×4

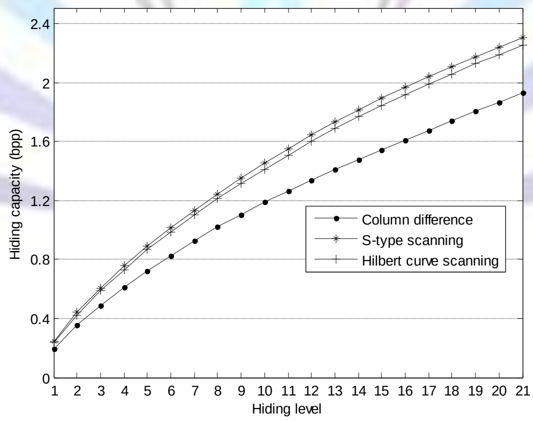


Fig. 6. The embedding rate curves for three methods as block size 4×4

Difference Method

From the analysis above, we can see that when the embedding level is more than 5, the PSNR values for the marked cover image will be lower than 35 usually, and hence the marked cover image will have large distortion. Here we only delve into cases of embedding levels ranging from 1 to 5. The results are shown in Table 3 and Table 4. We can draw the following conclusions: 1) Algorithm based on column difference under high embedding level can still maintain a good image quality, while S-type scanning and Hilbert curve scanning algorithm can't achieve it. 2) Considering the same image quality and block size, S-type scanning algorithm can obtain larger embedding capacity if embedded with the level 1 to avoid attack, otherwise, column difference algorithm with high embedding level is adopted to embed more secret data.

CONCLUSION

In this paper, two multi-level reversible data hiding algorithms based on S-type scanning and Hilbert curve scanning and difference image histogram modification are proposed. Specifically, we provide concrete steps of the algorithms, and then present the implemented results and make comparision with algorithm based on column difference in detial. Difference image generated by adjacent pixels difference operation can effectively improve the original cover image's embedding capacity. The column difference based algorithm is simple but effective. The S-type scanning difference based algorithm is the most effective in improving the embedding capacity among these three algorithms. In practical applications, the blocking and multi-level embedding can effectively guarantee the embedding capacity as well as the image quality to avoid excessive distortion. Nevertheless, along with the increase of embedding levels, the PSNR value of the marked cover image will be gradually reduced. Therefore, the embedding level should not be too large to ensure the quality of the marked cover image.

Table 3. The PSNR values for three methods with different block sizes and embedding levels

Method	Block size	Embedding level				
		1	2	3	4	5
Column difference	4×4	48.35986	42.672	39.241	36.789	34.87728
	8×8	43.34841	37.69197	34.36294	32.00799	30.17034
	16×16	39.07	33.46919	30.20253	27.92186	26.13643
S-type scanning difference	4×4	43.1114	37.63708	34.36733	32.02735	30.16527
	8×8	37.53802	32.40906	29.24712	26.99794	25.27972
	16×16	32.93542	27.29597	24.00697	22.11375	20.31838
Hilbert curve scanning difference	4×4	43.4134	38.0018	34.79347	32.45182	30.62282
	8×8	37.69001	32.67724	29.66672	27.35201	25.65839
	16×16	33.17623	27.71339	24.52619	22.35632	20.84716

Table 4. The embedding rates for three methods with different block sizes and embedding levels

Method	Block size	Embedding level				
		1	2	3	4	5
Column difference	4×4	0.195068	0.35269	0.488724	0.610718	0.72287
	8×8	0.153671	0.276672	0.382355	0.476227	0.562317
	16×16	0.140198	0.257309	0.356537	0.444351	0.524155
S-type scanning difference	4×4	0.24585	0.440079	0.607315	0.757126	0.892838
	8×8	0.200226	0.358292	0.492325	0.611816	0.720139
	16×16	0.184067	0.334579	0.458145	0.56813	0.669006
Hilbert curve scanning difference	4×4	0.23676	0.42482	0.587662	0.733246	0.866028
	8×8	0.186462	0.335129	0.460953	0.572556	0.674057
	16×16	0.167603	0.305664	0.420868	0.522888	0.615555

REFERENCES

[1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Process., 1997, 6:1673-1687.

[2] M. Wu, B. Lin, Data hiding in image and video: part I—fundamental issues and solutions, IEEE Trans. Image Process. 2003, 12 (6): 685–695.

[3] M. Wu, H. Yu, B. Liu, Data hiding in image and video: part II—designs and applications, IEEE Trans. Image Proc-ess. 2003, 12 (6): 696–705.

[4] Ruisong Ye, Huiqing Huang, Ruifeng Chen, A Color Image Watermarking Scheme Based on 2D Tent Map and Chebyshev Map, Applied Mechanics and Materials, 2014, Vols. 543-547: 2304-2307.

[5] J. Fridrich, M. Goljan, R. Du, Lossless data embedding for all image formats, in: Proceedings of SPIE Photonic West, Electronic Imaging 2002, Security and Watermarking of Multimedia Contents, vol. 4675, San Jose, California, January 2002, pp. 572–583.

[6] M. U. Celik, G. Sharma, A. M. Tekalp, E. Saber, Lossless generalized-LSB data embedding. IEEE Transactions on Image Processing (S1057-7149), 2005, 14(2): 253-266.

[7] Jun Tian. Reversible data embedding using a difference expansion. IEEE Transaction on Circuits and Systems for Video Technology (S1051-8215), 2003, 13(8): 890-896.

[8] Z. Ni, Y. Q. Shi, N. Ansari, et al. Reversible Data Hiding. IEEE Transaction on Circuits and Systems for Video Technology (S1051-8215), 2006, 16(3): 354-362.

[9] J. Hwang, J. W. Kim, J. U. Choi. A reversible watermarking based on histogram shifting, Int. Workshop on Digital Watermarking, Lecture Notes in Computer Science, Jeju Island, Korea (S0302-9743), 2006. 4283: 348-361.

[10] S. K. Lee, Y. H. Suh, Y. S. Ho, et al. Lossless data hiding based on histogram modification of difference images, Advances in Multimedia Information Processing-PCM 2004. Germany: Springer Berlin Heidelberg (S0302-9743), 2005: 340-347.

[11] Chia-Chen Lin, Wei-Liang Tai, Chin-Chen Chang. Multilevel reversible data hiding based on histogram modification of difference images. Pattern Recognition, 2008, 41: 3582-3591.

[12] Zhexuan Song, Nick Roussopoulos, Using Hilbert curve in image storing and retrieving, Information Systems, 2002, 27: 523-536.

[13] H. Sagan, Space-Filling Curve, Springe- Verlag, 1994.

[14] Li Liu, Ruisong Ye, A Matrix Iterative Approach to Systematically Generate Hilbert-type Space-filling Curves, International Journal of Computers and Technology, 2015, 14(1): 6281-6294

AUTHOR BIOGRAPHY

Jin Chen, a master degree candidate at School of Software Engineering, South China University of Technology.

Ruisong Ye, born in 1968 and received the B.S. degree in Computational Mathematics in 1990 from Shanghai University of Science and Technology, Shanghai, China and the Ph. D. degree in Computational Mathematics in 1995 from Shanghai University, Shanghai, China. He has been a professor at Department of Mathematics in Shantou University, Shantou, Guangdong, China since 2003. His research interest includes bifurcation theory and its numerical computation, fractal geometry and its application in computer science, chaotic dynamical system and its application in computer science, specifically the applications of fractal chaotic dynamical systems in information security, such as, digital image encryption, digital image hiding, digital image watermarking, digital image sharing. So far, he has published more than one hundred papers.