Reversible data hiding based on histogram modification using S-type and Hilbert curve scanning

Rui-Song Ye[1], Jin Chen[2] and Xing Zhang[1]

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

Reversible data hiding technology has 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 is 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 difference ways influence the effect of hiding.

Keywords: Reversible Data Hiding; Difference; Histogram Modification.

1. Introduction

With the rapid development of network technology and multimedia processing technology, images, videos and other digital products are largely distributed, communicated and shared online, which brings great convenience for people's life. However, information disclosure, content tampering and infringement problems become increasingly serious consequently. In order to solve these problems, encryption and information hiding provide two effective approaches. In image processing research field, data hiding is referred to a process to embed the secret data into a cover image. 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, 2].

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, see for example [3-6]. Reversible watermarking algorithms based on difference expansion perform the reversible watermark data embedding by extending the difference of adjacent pixels [3]. Reversible data embedding based on histogram translation embeds watermark data in the vacancy of grayscale using statistical characteristic of the image histogram [4]. 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. In 2004, Lee et al. used difference histogram modification to embed secret information data, resulting in higher image quality as well as large capacity [5]. In 2008, Lin et al. proposed a multi-level reversible watermarking algorithm based on difference histogram modification and further improve the embedded information capacity [6]. How to generate the difference image, improve the peak value frequency of difference image, the adopted difference way is a worth considering issue. This paper considers the design of two different difference schemes, S-type 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 scanned adjacent pixels. S-type 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 proposed in [6]. The influence of differential method and embedding level to embedded capacity and peak signal to noise ratio are analyzed.

2. Multi-level Scheme Based on Histogram Modification of Difference Image Using Curve Scanning

2.1. *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 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 scanning is described by Step 1-5 and $0 \le i \le A \times B - 2$, $0 \le b \le T - 1$ in Step 1-4.

Step 1. The original cover image is divided into small blocks H_b, each of size $A \times B$. For S-type scanning, the image block pixel values are shown in Figure 1(a) (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 neighbouring pixels gray value subtraction: $D_{bs}(i) = |H'_{bs}(i) - H'_{bs}(i+1)|$.

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

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

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

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

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

For the remaining pixels, the inverse operation is expressed by

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

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-dimensional matrix block with size $A \times B$. All the blocks are accumulated to be the resulted cover image SH embedded secret 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. We note that 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 information 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.

2.2. *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, etc. 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. 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 [7], the Hilbert curves with first, second and fourth order are depicted in Figure 1(b)-(d). The curve starts from bottom left, then scans the pixels in its unique way and ends at bottom right. The image data information is first read according to the Hilbert curve and rearranged to be 1D 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.
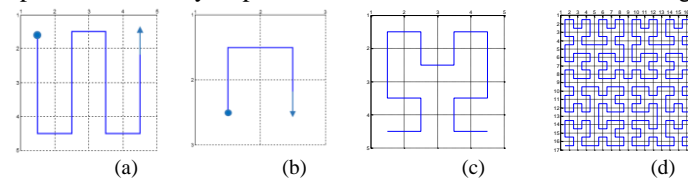

| (a) | (b) | (c) | (d) |

Fig. 1 Scanning type: (a) S-type; (b)-(d) Hilbert curves with orders 1, 2 and 4

3. Analysis of Hiding Effect

(i) **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 ways will generate different peak value frequencies for the same processed cover image. Three test cover images with size 256×256, Lena.bmp, Rice.png, and Baboon.png (see Figure 2(a)-(c)), 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. The peak value frequencies for the test images and their difference images are demonstrated in Table 1. One can draw the following conclusions from Tab. 1. 1) No matter what kind of difference way is adopted, the peak value frequency of the difference image is larger than that of the original cover image. Therefore adjacent pixels difference

operation is helpful for improving image embedding capacity. 2) The S-type scanning based difference is the most effective way to improve the embedding capacity among the three difference ways. 3) When the peak value frequency of the original cover image is large, the column difference way seems better. 4) Hilbert curve scanning based difference method can usually improve the embedding capacity compared with the column difference method.

(ii) **Embedding level**. 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 ways are demonstrated in Figure 2(d)-(e). 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 column 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.
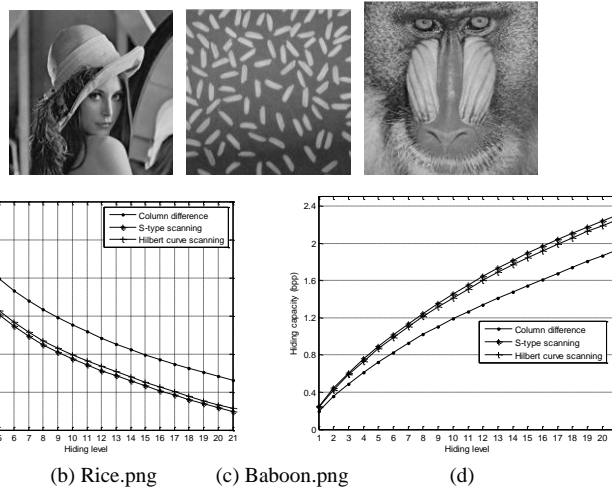
(a) Lena.bmp   (b) Rice.png   (c) Baboon.png   (d)   (e)

Fig. 2 Original test images (a)-(c), the PSNR values curve (d) and the embedding rates curve for three methods with block size 4×4

Tab. 1 The histogram peak values for three 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 ways 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

Tab. 2 The PSNR values for three ways with different block sizes and embedding levels

Method	Block size	Embedding level				
		1	2	3	4	5
Column difference	4×4	48.36	42.67	39.24	36.79	34.88
	8×8	43.35	37.69	34.36	32.01	30.17
	16×16	39.07	33.47	30.20	27.92	26.14
S-type scanning difference	4×4	43.11	37.64	34.37	32.027	30.17
	8×8	37.54	32.41	29.25	27.00	25.28
	16×16	32.94	27.30	24.01	22.11	20.32
Hilbert curve scanning difference	4×4	43.41	38.00	34.79	32.45	30.62
	8×8	37.694	32.68	29.67	27.35	25.66
	16×16	33.18	27.71	24.53	22.36	20.85

Tab. 3 The embedding rates for three ways with different block sizes and embedding levels

Method	Block size	Embedding level				
		1	2	3	4	5
Column difference	4×4	0.20	0.35	0.49	0.61	0.72
	8×8	0.15	0.28	0.38	0.48	0.56
	16×16	0.14	0.26	0.36	0.44	0.52
S-type scanning difference	4×4	0.25	0.44	0.61	0.76	0.89
	8×8	0.20	0.36	0.49	0.61	0.72
	16×16	0.18	0.33	0.46	0.57	0.67
Hilbert curve scanning difference	4×4	0.24	0.42	0.59	0.73	0.87
	8×8	0.19	0.34	0.46	0.57	0.67
	16×16	0.17	0.31	0.42	0.52	0.62

(iii) **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 2 and Table 3. 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 cannot 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.

Acknowledgments

References

1. I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Process., 6, 1673-1687(1997).
2. R. Ye, H. Huang, and R. Chen, A Color Image Watermarking Scheme Based on 2D Tent Map and Chebyshev Map, Applied Mechanics and Materials, Vols. 543-547, 2304-2307(2014).
3. J. Tian, Reversible data embedding using a difference expansion. IEEE Transaction on Circuits and Systems for Video Technology (S1051-8215), 13(8), 890-896(2003).
4. Z. Ni, Y. Q. Shi, N. Ansari, et al. Reversible Data Hiding. IEEE Transaction on Circuits and Systems for Video Technology (S1051-8215), 16(3), 354-362(2006).
5. 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), 340-347(2005).
6. Chia-Chen Lin and Wei-Liang Tai, and Chin-Chen Chang. Multilevel reversible data hiding based on histogram modification of difference images. Pattern Recognition, 41, 3582-3591(2008).
7. R. Ye and L. Liu, An iterated function system based method to generate Hilbert-type space-filling curves, I. J. Information Technology and Computer Science, 12, 12-22(2015).