The LSB-based High Payload Information Steganography

Wang Yu
bengbu navy petty officer academy

Abstract—The capacity and the distortion are the two important indexes in the information steganography field. The research tries adopting the famous Median Edge Detection to judge the flat area and the complex area, combining the concept of the Human Vision System to embed a fewer confidential information. A more confidential information should be embedded in the complex area for increasing the capacity. The research adopts the steganography method with the Least Significant Bit and the Optimal Pixel Adjustment Process method to embed the confidential information for reducing the distortion. The experimental results show that the research can obtain the efficiency of the high capacity and the low distortion.

Keywords- The center point; Matlab; Tilt; Bending; Distortion

I. INTRODUCTION

The capacity and the distortion are the two important indexes in the information steganography field. In general, the Human Vision System is sensitive to the changing of the flat area, that is, the surrounding pixel values are so close that a small-wide change of the pixel points can be easily observed. If the differences of the surrounding pixel values are so large that a large-wide change the pixel points cannot be easily observed in the complex area. The research adopts the Median Edge Detection to judge the flat area and the complex area, combining the concept of the HVS to embed the confidential information.If it is the flat area, a fewer information can be embedded. If is the complex area, a more information should be embedded. The experimental results show that the research can obtain the efficiency of the high capacity and the low distortion, avoiding the attack of the pseudo written analysis aiming to the LSB.

II. RELATED RESEARCHES

The LSB Steganography, the OPAP and the MED are respectively introduced in the followings.

A. The LSB Steganography

The LSB Steganography introduced by Chan and other scholars [2] is a famous irreversible Steganography, and its advantage is the high capacity and the low distortion. The method hides the confidential information in the last n bits of the image pixel values. The confidential information and the pixel values are represented by the binary system in the process of the Steganography.

For example, if the pixel value is x=153=(10011001)2, the confidential information with the hidden 3 bits is (110)2, and the newly obtained pixel value is y=(10011110)2=158 after hiding the confidential information.

B. The OPAP

The disadvantage of the LSB Steganography is that the longer the hidden confidential information is, the worse the quality of the obtained new image is. In addition, Chan and other scholars (2004)[2]adopts the OPAP to improve the distortion of disguising the images.

For example, x is a certain pixel value in the image, y is the new pixel value obtained from the LSB Steganography passing by the x, z is the new image value obtained from the OPAP passing by the y. If d equal to y-x, the y can be obtained from the confidential information whose length is n bits hidden in the x in terms of the LSB Steganography.

$$d \in \left[2^{-n}, 2^{n} \right] \tag{1}$$

According to the range of the *d* value, the different solving methods are as follows:

$$\begin{aligned}
&\text{Methods 1:}\quad d \in \left(2^{n-1}, 2^{n}\right) \\
&\textit{If} \quad y \geq 2^{n}, \textit{then} \quad z = y - 2^{n}, \textit{otherwise} \quad z = y \\
&\text{Methods 2:}\quad d \in \left[-2^{n-1}, 2^{n-1}\right] \\
&z = y \\
&\text{Methods 3:}\quad d \in \left(-2^{n}, -2^{n-1}\right) \\
&\textit{If} \quad y \geq 256 - 2^{n}, \textit{then} \quad z = y + 2^{n}, \textit{otherwise} \quad z = y
\end{aligned} \tag{2}$$

Assumption 1: If a certain pixel value in the image is x=25=(00011001)2, the confidential information with the hidden 3 bits（n=3）is (111)2, and the newly obtained pixel value is y=31=(00011111)2 after hiding the confidential information. The differential value of the two images can be calculated, d=y-x=31-25=6. The formula (2) is suitable for the method 1 according to the d value. As for equation y=31≥23, the obtained new pixel value is z=y-23=31-8=23=(00010111)2, and the error is reduced from 6 to 2.

Assumption 2: If a certain pixel value in the image is x=26=(00011010)2, the confidential information with the hidden 3 bits（n=3）is (101)2, and the newly obtained pixel value is y=29=(00011101)2 after hiding the confidential information. The differential value of the two images can be calculated, d=y-x=31-25=6. The formula (2) is suitable for the method 2 according to the d value. The obtained new pixel value is z=y=29.

Assumption 3: If a certain pixel value in the image is x=7=(00000111)2, the confidential information with the hidden 3 bits（n=3）is (001)2, and the newly obtained

pixel value is y=29=(00011101)2 after hiding the confidential information. The differential value of the two images can be calculated, d=y-x=1-7=-6. The formula (2) is suitable for the method 3 according to the d value. The obtained new pixel value is z=y+23=1+8=9=(00001001)2, and the error is reduced from 6 to 2.

The above each assumption can make the absolute value of the error is less than 2n-1（in which n is the length of the embedded-willing confidential information）through the OPAP.

C. *The MED*

The MED introduced by Weinberger and other scholars (2000)[3] is a famous MED. It does not only predict whether there is an edge existing the exterior, also predicting the predicted values of the target pixels. The research discovers that the changing ranges of the predicted errors are very large. In other words, if the general differential values are regarded as the basis of the length of the hidden confidential information, the high distortion can be caused.

The MED adopts the three pixels a, b and c next to the target pixel x to predict the pixel value of the x.

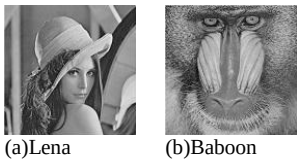
$$x\begin{cases}\min\{a,b\} & if \quad c \geq \max\{a,b\} \\ \max\{a,b\} & if \quad c \geq \min\{a,b\} \\ a+b-c & otherwise\end{cases} \qquad (3)$$

c	b
a	x

Figure 1. The diagram between the target pixel x and its neighboring pixel

The research just adopts it to judge whether there are the features of the edge. If so, the edge exists and it is called the complex area. If not, the edge does not exist and it is called the flat area.

The following explains the reason that the predicted values without using the MED can be used as the main basis of the information Steganography. The Lena figure and the Baboon figure in the SIPI image database are the examples. (as shown in the Fig .2)



(a)Lena (b)Baboon

Figure 2. The Lena figure and the Baboon figure in the SIPI image database

The Fig .3 and the table I is the error statistical figure and the error cumulative percentage in the Lena figure, and the percent of the n≤16 is 95.84%. It represents that the predicted values in the MED is accurate. (the red line equals to the n value in the LSB+OPAP)

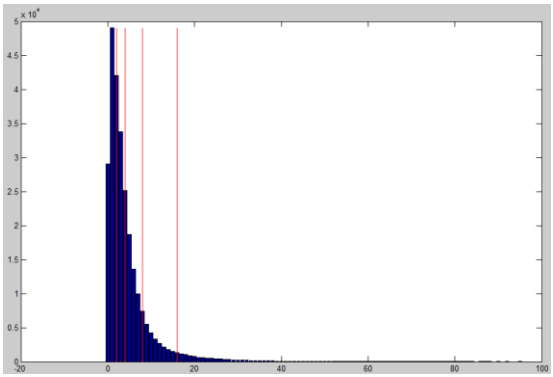


Figure 3. The error statistical figure in the Lena figure

TABLE I. THE ERROR CUMULATIVE PERCENTAGE IN THE LENA FIGURE

n	*n*≤2	*n*≤4	*n*≤8	*n*≤16
the cumulative percentage	45.85%	68.31%	87.25%	95.84%

The figure 43 and the table II is the error statistical figure and the error cumulative percentage in the Baboon figure, and the percent is just 69.78% for the Baboon figure tens to the complex figures. It represents that the predicted values in the MED is inaccurate. (the red line equals to the n value in the LSB+OPAP)

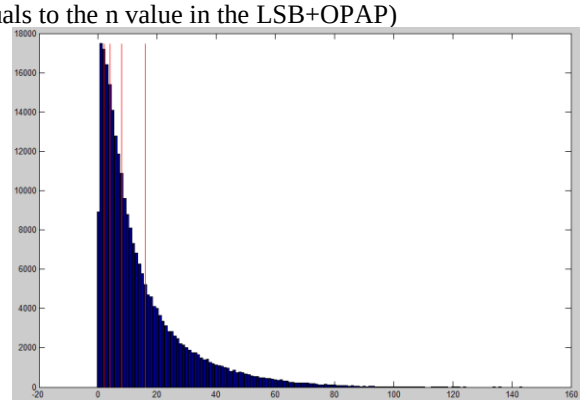


Figure 4. The error statistical figure in the Baboon figure

TABLE II. THE ERROR CUMULATIVE PERCENTAGE IN THE BABOON FIGURE

n	*n*≤2	*n*≤4	*n*≤8	*n*≤16
the cumulative percentage	16.64%	28.77%	47.69%	69.78%

Therefore, the research does not select the predicted values of the MED as the basis of the hidden information, just adopting the LSB+OPAP. When n bits are embedded, the absolute values of the error are less than 2n-1.

III. RESEARCHING METHODS

The hidden method and the removal method are respectively explained in the followings.

A. *The hidden method*

The hidden procedure is divided into two phases. At first, aiming to each pixel point in the first row and the first line, the confidential information of the (n-1) bits hidden in

the LSB+OPAP is adopted, as shown in the grey part of the Fig .5.

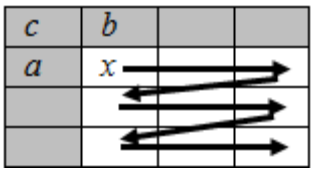


Figure 5. The embedding in the first phase

Next, other white parts of the Fig .5 are adopted to judge whether there are the edges in terms of the zig-zag method with the use of the formula (3). If so, n confidential information is hidden. If not, (n-1) confidential information should be hidden for complying with the HVS.

For example, if n is 3, the target pixel equals to the equation x=55=(110111)2, its neighboring pixel points a=28, b=47 and c=20, the formula (3) should be adopted to predict it. As for the inequation c≤min{a,b}, the edges are existed. If the hidden confidential information is (000)2, the y=48=(110000)2 can be obtained through the LSB Steganography. After the adjustment of the OPAP, d=-7 can be obtained with the formula (2), and the situation is applied in the method 1 so that the new pixel point z=56=(111000) can be obtained after the Steganography. The whole image can be dealt with through the Steganography and the disguised image hidden with the confidential information can be obtained.

B. The removal method

At first, the confidential information in the grey part of the Fig .5 should be fetched out, and the pixel x is represented by the binary system so that the final (n-1) bits are fetched out as the confidential information. Then the confidential information in the white parts of the Fig .5 should be fetched out, and the pixel x should be MED predicted for judging whether there are the edges in terms of the zig-zag method. If so, the n confidential information should be etched after the pixel value x is represented by the binary system. If not, (n-1) confidential information should be fetched out.

For example, if n is 3, the target pixel x=56 is the pixel point of the white area in the Fig .5, its neighboring pixel points a=28, b=47 and c=20, the MED prediction should be conducted. As for the inequation c≤min{a,b}, the edges are existed. If the pixel x=56 represents the binary system (111000)2, the final three bits (000)2 can be fetched out, that is, the confidential information hidden in the target pixel can also be fetched out.

IV. EXPERIMENTAL RESULTS

The research adopts the six images, Lena,Tiffany, Baboon, F16, Scene and Peppers in the SIPI image database as the experimental objects, and the method in the paper is compared with the LSB Steganography and the OPAP. (as shown in the Fig .6)

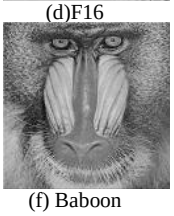
(a) Lena

(b) Peppers

(c) Tiffany

(d)F16

(e) Scene



(f) Baboon

Figure 6. The six experimental images in the SIPI image database

As to the LSB and the OPAP, the n value from the table III to the table Ⅷ is the length of the embedded confidential information, and there are two different significance to the method researched in the paper. If the target pixel is judged that there are the edges, the pixels are located in the complex area and n is embedded. If not, the pixels are located in the flat area and (n-1) is embedded.

Generally speaking, the flat areas in the natural images are more than the complex areas so that the comparisons from the table III to the table Ⅷ adopt the n value in the flat areas as the contrast. Take the table III as an example, when n is 4, the Payload in the OPAP is 4bpp and the PSNR value is 34.80dB. Compared with the method adopted by the research, the n is 5, the Payload in the OPAP is 4.68bpp and the PSNR value is 30.02dB. It shows that the research can have the high capacity complying with the condition of the HVS.

In addition, the research is based on the edge features of the image itself. Different target images can hide different lengths of the confidential information so that the attack to the pseudo written analysis of the LSB.

TABLE III. THE EXPERIMENTAL RESULTS IN THE LENA

Lena		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.70
	PSNR	44.15	46.37	47.36
n=3	Payload	3.00	3.00	2.69
	PSNR	37.92	40.72	41.83
n=4	Payload	4.00	4.00	3.68
	PSNR	31.78	34.80	35.97
n=5	Payload	5.00	5.00	4.68
	PSNR	25.86	28.81	30.02

TABLE IV. THE EXPERIMENTAL RESULTS IN THE TIFFANY

Tiffany		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.71
	PSNR	44.16	46.34	47.27
n=3	Payload	3.00	3.00	2.70
	PSNR	37.91	40.65	41.73
n=4	Payload	4.00	4.00	3.69
	PSNR	31.82	34.67	35.82
n=5	Payload	5.00	5.00	4.68
	PSNR	26.10	28.36	29.57

TABLE V. THE EXPERIMENTAL RESULTS IN THE BABOON

Baboon		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.63
	PSNR	44.15	46.37	47.61
n=3	Payload	3.00	3.00	2.63
	PSNR	37.92	40.74	42.08
n=4	Payload	4.00	4.00	3.63
	PSNR	31.86	34.81	36.19
n=5	Payload	5.00	5.00	4.64
	PSNR	25.81	28.81	30.19

TABLE VI. THE EXPERIMENTAL RESULTS IN THE F16

F16		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.69
	PSNR	44.16	46.37	47.40
n=3	Payload	3.00	3.00	2.68
	PSNR	37.98	40.73	41.89
n=4	Payload	4.00	4.00	3.67
	PSNR	31.84	34.82	36.04
n=5	Payload	5.00	5.00	4.67
	PSNR	26.07	28.82	30.07

TABLE VII. THE EXPERIMENTAL RESULTS IN THE SCENE

Scene		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.68
	PSNR	44.16	46.37	47.40
n=3	Payload	3.00	3.00	2.68
	PSNR	37.91	40.73	41.89
n=4	Payload	4.00	4.00	3.66
	PSNR	31.86	34.82	36.05
n=5	Payload	5.00	5.00	4.66
	PSNR	25.78	28.80	30.10

TABLE VIII. THE EXPERIMENTAL RESULTS IN THE PEPPERS

Peppers		LSB	OPAP	The research method
n=2	Payload	2.00	2.00	1.73
	PSNR	44.16	46.38	47.26
n=3	Payload	3.00	3.00	2.72
	PSNR	37.92	40.72	41.70
n=4	Payload	4.00	4.00	3.70
	PSNR	31.81	34.81	35.90
n=5	Payload	5.00	5.00	4.68
	PSNR	25.75	28.82	29.98

V. CONCLUSION

The research adopts the famous MED to judge the flat areas and the complex areas on the basis of the edge features of the image itself. The (n-1) bits are embedded in the flat areas and the n bits are embedded in the complex areas. The experimental results show that the method can obtain the efficiency of high capacity and the low distortion. In addition, the research can hid different lengths of the confidential information for avoiding the attack to the pseudo written analysis of the LSB.

REFERENCES

[1] W. J. Chen, C. C. Chang, T. Le. “High payload steganography mechanism using hybrid edge detector,” Expert Systems with Applications, 2010, pp. 3292-3301.

[2] C. K. Chan, L. M. Cheng. “Hiding data in images by simple LSB substitution, ” Pattern recognition , 2004, pp. 469-474.

[3] M. L. Weinberger, Seroussi G, “The LOCO-I lossless image compression algorithm:principles and standardization into JPEG-LS, ” IEEE Trans. Image Process, 2000, pp. 1309-1324.

[4] Kalaivani B, Padmaa M, “Authentication of encrypted secret information in image steganography,” International Conference on Communication and Signal Processing, ICCSP 2014 - Proceedings, pp. 803-807, 2014.

[5] Fouroozesh Zohreh, Al Ja'am Jihad,” Image steganography based on LSBMR using Sobel edge detection,” 2014 3rd International Conference on e-Technologies and Networks for Development, ICeND 2014, pp. 141-145, December 17, 2014.

[6] Bagade Anant M, Talbar Sanjay N,” Secure transmission of morphed stego keys over internet using IP steganography,” International Journal of Information and Computer Security, vol. 6, no. 2, pp. 133-142, 2014.

[7] Zou Ming-Guang, Li Zhi-Tang,” Wav-audio steganography algorithm based on amplitude modifying,” Tongxin Xuebao/Journal on Communications, vol. 35, pp. 36-40, October 1, 2014.

[8] Vijay M, Vigneshkumar V,” Image steganography algorithm based on Huffman encoding and transform domain method,” 2013 5th International Conference on Advanced Computing, ICoAC 2013, pp. 517-522, October 12, 2014.

[9] Zou Mingguang, Li Zhitang,” A wav-audio steganography algorithm based on amplitude modifying,” Proceedings - 2014 10th International Conference on Computational Intelligence and Security, pp. 489-493, January 20, 2015.

[10] Satir Esra, Isik Hakan,” A Huffman compression based text steganography method,” Multimedia Tools and Applications, vol. 70, no. 3, pp. 2085-2110, June 2014.

[11] Ou Duanhao, Sun Wei,” High payload image steganography with minimum distortion based on absolute moment block truncation coding,” Multimedia Tools and Applications, May 16, 2014.

[12] Ye Tian-Yu,” Quantum steganography in cavity QED based on the entanglement swapping of any two Bell states,” Guangdianzi Jiguang/Journal of Optoelectronics Laser, vol. 25, no. 8, pp. 1571-1577, August 15, 2014.

[13] Vimal Jithu, Alex Ann Mary,” Audio steganography using dual randomness LSB method,” 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014, pp. 941-944, December 18, 2014.

[14] Yang Ren-Er, Zheng Zi-Wei, Jin Wei,” Cover selection for image steganography based on image characteristics,” Guangdianzi Jiguang/Journal of Optoelectronics Laser, vol. 25, no. 4, pp. 764-768, April 2014.

[15] Iranpour Mehran, Safabakhsh Reza,” Reducing the embedding impact in steganography using Hamiltonian paths and writing on wet paper,” Multimedia Tools and Applications, March 8, 2014.