http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (5): 895-903, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.895.903

Information Hiding Algorithm with Strong Security Based on Run Length

[1]Jianquan Xie, [1]Xiaoping Fan, [1]Hui Peng, [2]Qing Xie and [2]Guang Sun
[1]Research Institute of Information Security, Hunan University of Finance and Economics, Changsha, China
[2]Department of Information Management, Hunan University of Finance and Economics, Changsha, China

Abstract: Imperceptibility, robustness, security and hiding capacity are important indexes in evaluating information hiding algorithms. Recently, in most case, only imperceptibility and robustness were considered rather than the security ability. The later one is highly required to guarantee secret communication. With fully consideration of this requirement, one kind of hiding information algorithm is proposed based on run length, the main idea of which is to separate the image into several binary images first. Then according to the parities of the run length of the binary images, one bit of information can be successfully embedded while no more than one pixel that located along black-white boundary of the binary image is modified. The blind extraction of the hiding information can be achieved. The algorithm neither obviously changes of distribution property of 0 and 1 in image's low plane, nor reduces the number of long run length. Hence it can successfully defend various detections that against Least Significant Bit (LSB) algorithm and improved LSB algorithm and it can also be applied in secret communication and other situations which require high security and large capacity.

Key words: Information hiding, run length, hiding capacity, imperceptibility, security

INTRODUCTION

Information hiding is to embed important information into other public carriers and then secretly deliver the information with almost no external characteristic and internal use value alteration of the carriers. As an important way to ensure safe information transmission, it can be effectively and widely used in several fields like digital watermark and secret communication. Recent carriers include text, image, voice or video and various kinds of document. These carriers have the same essences on application, which is utilizing the perception limitation of human vision or audition to hide information. Image is widely used in Internet as an information carrier and is the most widely used information hiding carrier as well since it has large redundant space. Recently most information hiding algorithms tend to study at imperceptibility and robustness (Jiang *et al*., 2010; Lou and Hu, 2012; Nezhadarya *et al*., 2011; Yang *et al*., 2012; Zeng *et al*., 2012; Zhou *et al*., 2010) of hidden information rather than undetected ability. In order to improve security and achieve safe information secret transmission, the hiding algorithm must realize vision imperceptibility of embedded information without detectable statistic abnormality.

Information hiding technology based on image has two categories: transformation domain and spatial domain. Generally, the spatial domain method is simple and convenient; it also has large hiding quantity and fast speed when embedding and extracting information. However, many spatial domain methods have poor performances on resisting attack. Some methods even change statistic characteristic resulting the hiding been easily detected (Bandyopadhyay *et al*., 2010; Chen and Chen, 2011; Fillatre, 2012; Ker, 2007; Lee *et al*., 2012; Lei *et al*., 2010; Zhang *et al*., 2010). The low security level constrains their application.

On the other side, lots of transformation domain algorithms share idea of spatial domain algorithm. For example, algorithm alters the lowest effective bit of quantified Discrete Cosine Transform (DCT) coefficient. Those algorithms also have security issue while the hiding capacity is even much smaller than the spatial domain algorithm. A kind of hiding algorithm based on run length that is proposed in this study has stronger security and larger hiding capacity, which can satisfy hiding quantity requirement of secret communication. The main idea is to transform grey image or color image into multiple binary images by bit plane and then use the parity of the longer (or shorter) run length to represent 0 (or 1). This algorithm will not cause aberrant change in low plane or run characteristic. It can also defend steganalysis that against LSB, adaptive LSB and run characteristic.

Corresponding Author: Jianquan Xie, Research Institute of Information Security, Hunan University of Finance and Economics, Changsha, China

MATERIAL AND METHODS

Security analysis of classical algorithm: Human eyes eventually perceive image in spatial domain, so the imperceptibility of all kinds of image hiding algorithms can be analyzed in spatial domain. For the purpose to guarantee the imperceptibility of hidden information, many images embed secret information into those locations that human eye can not perceive very well, where the Least Significant Bit (LSB) algorithm is a typical case. The LSB algorithm is simple and efficacious with very good imperceptibility, so majority of recent spatial watermarking scheme are based on image pixel's LSB or improved LSB.

LSB method even effect algorithms in transformation domain, classical Jsteg algorithm for instance. The JSteg algorithm transforms gray degree in spatial domain into quantized DCT coefficient, which makes it different from LSB method in spatial domain. With the purpose to enhance security of hided information, secret information usually take encryption process before insertion and information can then be seen as bit string of 0 and 1 that distributes stochastically. The security based on LSB method depends on stochastic characteristic of the LSB plane. However, many images' LSB planes do not appear to have stochastic characteristic. For example, the 256×256×8 bird image in Fig. 1a, b to i are its decomposed bit planes from the highest to the lowest one.

It is easy to see that every bit plane does not have stochastic characteristic. In practical information hiding application, generally the information that embedded in LSB is enciphered, like sequence with fake randomness or scrambled image. This method can avoid that secret information be detected by executing bit plane decomposition to intercepted image (Amirtharajan and Rayappan, 2012; Lu *et al*., 2012). If one uses enciphered

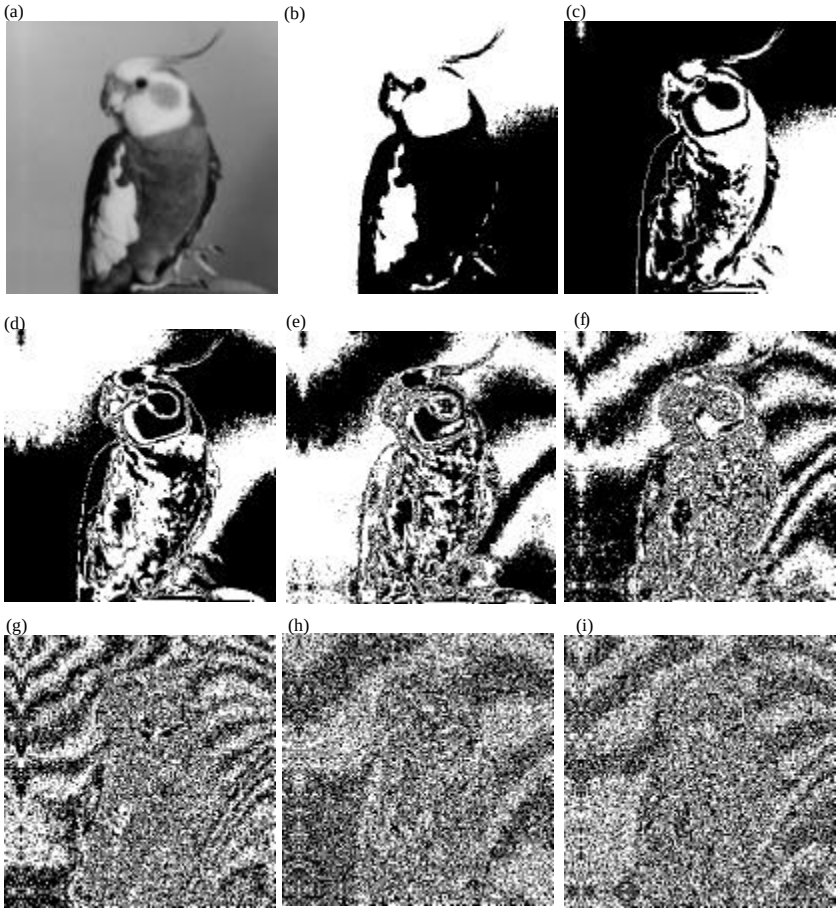

Fig. 1(a-i): Bit plane decomposition of the bird image, (a) Original gray image (b) Eighth bit plane, (c) Seventh bit plane, (d) Sixth bit plane, (e) Fifth bit plane (f) Fourth bit plane, (g) Bit plane (h) Second bit plane and (i) First bit plane

information to substitute the LSB bit, then this LSB bit plane will has stronger noise characteristic. Nevertheless, the noise difference between this plane and nature image bit plane actually gives a hint to detector and consequently brings threat to the security of information that hidden in LSB bit plane. Many LSB algorithms do not notice this problem that sometimes detector can find obvious information hidden trace even by simply applying visual detect to bit plane. LSB algorithm or other information hiding algorithms that hide data in low bit plane break nature attribute so their anti-detect abilities are not good. Chi-square (χ^2) analysis method applies steganalysis according to whether LSB plane of image has stochastic characteristic. Therefore, although hiding algorithm based on LSB has relative good imperceptibility, its security is not strong enough consequently makes it not suitable for secret communication.

In steganalysis, the similar ratio between original image and carrier image is not the key point of imperceptibility for information hiding. Instead, the identification ratio of hidden information in image is the key point, where hidden information is considered as the noise. In fact, steganalysis is applied without original image.

The insertion of encrypted information will enhance the stochastic characteristic of corresponding bit plane, especially the lowest bit plane. The number of short run will increase, that is to say the number of original image's long run will decrease along with the increasing of the number of short run (Zhou and Yang, 2005). Hence exclude using Chi-square (χ^2) analysis, checking statistical characteristic of run length can also tell whether the image contains secret information. This problem needs to be figured out in order to enhance the security of the algorithm.

The main function of human vision system is to abstract structure information of vision field (Li *et al*., 2009; Luo *et al*., 2007; Xie *et al*., 2011). Human vision characteristic is mainly used in information hiding technology from three aspects (Liu *et al*., 2007): Sensitive characteristic to vision space frequency, covering characteristic to contrast ratio and illumination. At present, majority of algorithms only utilize the illumination covering characteristic.

Actually, utilizing other characteristics can hide information as well. Even when the illumination characteristic of one single pixel does not satisfy certain demand, taking information hiding in binary image as a typical case (Ma and Lin, 2011). Binary image only has two colors-black and white with less redundant space, so for the purpose to guarantee imperceptibility, usually embed data into the common boundary of black and white. In a binary image, if there emerges continuous multiple pixels with the same value, then adding or reducing one same value pixel will not cause obvious perception change to human vision system. Moreover, along with the increase of the length of the continuous same value pixels string, the perception to such kind of change is weaker. In other words, if the run length is larger than some certain value, adding 1 or reducing 1 to the length will not affect vision perception, so it has better imperceptibility. Still, in order to solve security problem, the decrease of number of long run must be prevented.

Information hiding algorithm based on run length: Each bit plane of gray image (or color components of color image) can be considered as one binary image. Then by decomposing gray image into multiple binary images with appropriate embedding method, information can be embedded into every bit plane with good imperceptibility. Furthermore, if statistic characteristic of the image can keep steady, the algorithm will have stronger security. The basic idea of the algorithm proposed in this study is to first change at most one pixel value at common border of black and white. Then use parity of the longer (or shorter) run length to respectively represents 0 or 1 which needs to be hidden, while blind extraction can be achieved.

Embedding algorithm:

Step 1: Apply bit plane decomposition to a gray image (or a color image) and obtain multiple binary images

Step 2: Progressively scan the binary image. The traversal algorithm is to scan adjacent one black segment and one white segment (adjacent run couple of one black segment and one white segment) B and W. Scanning starts from a white run or a black run depends on the key k (i). k (i) determines original scanning run length in the row, which can enhance security. Assume the length of those two run is a and b, respectively

Step 3: Compare scale of a and b:

$$c1 = \max(a, b) \quad (1)$$

$$c2 = \min(a, b) \quad (2)$$

Without loss of generality, suppose c1 = a, c2 = b

Step 4: Judge if the run couple can hide information. According to former analysis, information can be embedded when the shortest run length exceed a certain value. It indicates that information can be

inserted when c2 surpasses the designated value x. (Due to different perceptibility of different bit plane, those bit planes could have different x values, the value can be designed small at low bit plane or large at high bit plane). Execute the following embedding process to one obtained bit that needs to be embedded in. Otherwise the run couple will be considered as cannot embed information. Then the process will jump to step 5 to choose another run and then retry the embedding manipulation again

Step 5: If the parity of c1-the length of long run-is the same with the information value that need to be embedded, then the insertion is accomplished. If there are different, then modify one pixel at common boundary of the two run segment to achieve the information hiding. The modify principle is to make the parity of the longer run has the same value with the information (when the run length is odd, embed 1; when the run length is even, embed 0). There are three cases in specific procedures:

- Modify one pixel at the common boundary of two run to add 1 to c1 and minus 1 to c2. If min (c_1, c_2)>x is satisfied, then embedding is accomplished. Turn to Step 6
- On the basis of (1), minus 2 to the length of c1 (equals to minus 1 to c1 and add 1 to *c*2 on the basis of the former one). If min (c1, c2)>x is satisfied, then the embedding is accomplished and tur to Step 6. If the length of the short run is still longer than x and the parity of long run is the same with the information, then embedding is completed and then turn to step 6. x is designated as the shortest length of run that can hidden information. During this procedure, the length comparison between two run might change (only happens when the length of longer run is only 1 larger than the short run). If the parity of the changed long run is the same with information that need to be hidden, then embedding is completed and then turn to step 6. Otherwise turn to the next step
- After 1 and 2, if embedding requirement still not satisfied, the run will be considered cannot embed information. This case only occurs when the runs in one couple have the same length and the shortest run length is smaller than the designated value after insertion. For avoiding misinterpretation, an embedding procedure is required (add 1 to length of one run and minus 1 to another run) but this embedding is considered invalid. In Step 6, information will be embedded again instead of pick new information to embed and no extraction will be applied to this run

Step 6: Scan next run couple and apply embedding manipulation to the next bit that is waiting for embedding until the entire information been inserted. During the scanning, if one line is finished, then scan the next line under the control of key

Extraction algorithm:

Step 1: Decompose gray image (or color image) in bit planes like the embedding procedure and obtain multiple binary images

Step 2: Scan run couple (adjacent white segment and black segment) B and W in lines according to key k (i) and record the length of those two run respectively as a and b

Step 3: Abstract embedded information. If the length value a and b both larger than designated value x, it represents that1 bit information has been embedded (otherwise the run couple has no information embedded in). The parity of the longer run is the hidden information (if the length of two run is the same, then choose the former one), that is:

$$w(j) = \max(a, b) \bmod 2 \quad (3)$$

Step 4: Scan next run couple with respect to embedding algorithm and then abstract information until the entire information has been abstracted

RESULTS AND DISCUSSION

For the need to test security, hiding capacity and imperceptibility, use four 256×256×8 images, “Bird”, “Lena”, “Bead”, “House”, “Mandrill” and “Cavas” as shown from Fig. 2a to f to execute full insertion and extraction experiments.

Images with information embedded in are shown in figures from Fig. 3a to f.

The hiding capacities of six images are listed as follows: Figure 3a is 2961, b is 2283, c is 1950, d is 2703, Fig. 3e is 3282, f is 1778, respectively. The extraction accuracy is 100%. From results, although hiding capacity of this algorithm is lower than LSB algorithm, but is much higher than the hiding capacity of most transformation domain algorithms and algorithms in Fillatre (2012),

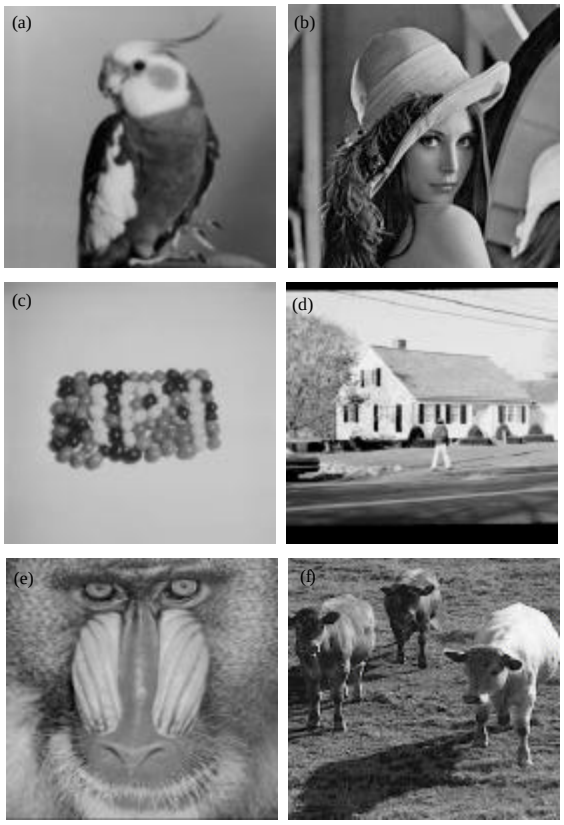

Fig. 2(a-f): Standard gray images for embedding, (a) Bird (b) Lena, (c) Bead (d) House, (e) Mandrill and (f) Cavas

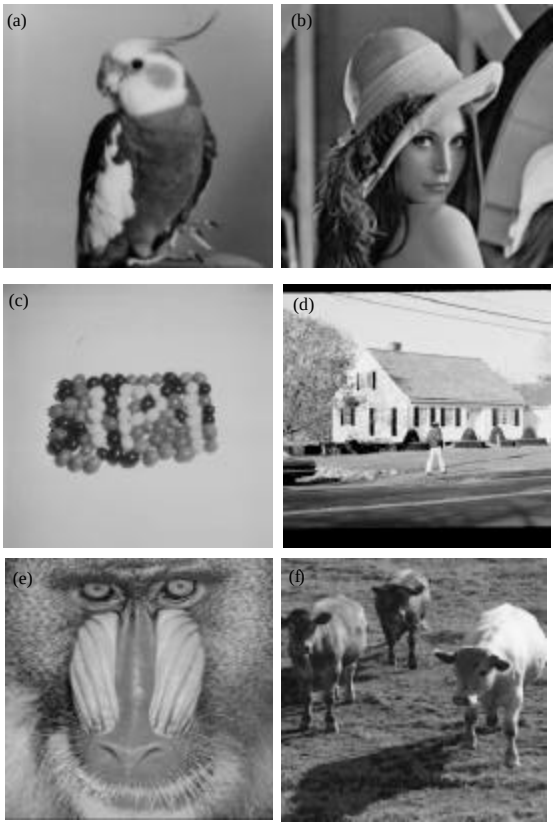

Fig. 3(a-f): Images with information embedded, (a) Bird (b) Lena, (c) Bead, (d) House, (e) Mandrill and (f) Cavas

Ker (2007), Liu *et al*. (2007) and Zhang *et al*. (2010). Moreover, this algorithm is more secure than other algorithms like those in Fillatre (2012), Ker (2007) Liu *et al*. (2007), Luo *et al*. (2007) and Zhang *et al*. (2010). For example, the algorithm in Luo *et al*. (2007) can only defend SPA attack. With only taking enlarging hiding capacity into consideration, one can adapt LSB algorithm on low bit plane and adapt the algorithm in this study on high bit plane. Then the hiding capacity is larger than LSB algorithm. However, the security of hidden information in low plane cannot be guaranteed as well as executing LSB. From human vision the change caused by insertion is imperceptible. Then objective evaluation is executed to evaluate imperceptibility after the image took insertion. The Peak Signal to Noise Rate (PSNR), which is widely used in gray image distortion evaluation, is not suitable to evaluate the imperceptibility of information hiding. Hence the improved PSNR evaluation index Weighted Peak Signal to Noise Rate (WPSNR) and objective measurement index Contrast Sensitivity Function (CSF), which is proposed in Bandyopadhyay *et al*. (2010), are used to inspect the imperceptibility instead. Evaluation results are shown in Table 1, from which one can see that imperceptibility index of all of the four images much exceed vision perceptive threshold value. This implies that the algorithm proposed in this study has good imperceptibility.

Table 1: Imperceptibility test results of images after information insertion

Image	CSF	WPSNR
Bird	51.39	60.63
Lena	46.22	54.46
Bead	51.34	56.44
House	50.35	51.72
Mandrill	42.32	57.94
Cavas	41.83	67.25

CSF: Contrast sensitivity function, WPSNR: Weighted peak signal to noise rate

Apply bit plane decomposition to images as shown in Fig. 3 and four low bit plane images are illustrated in

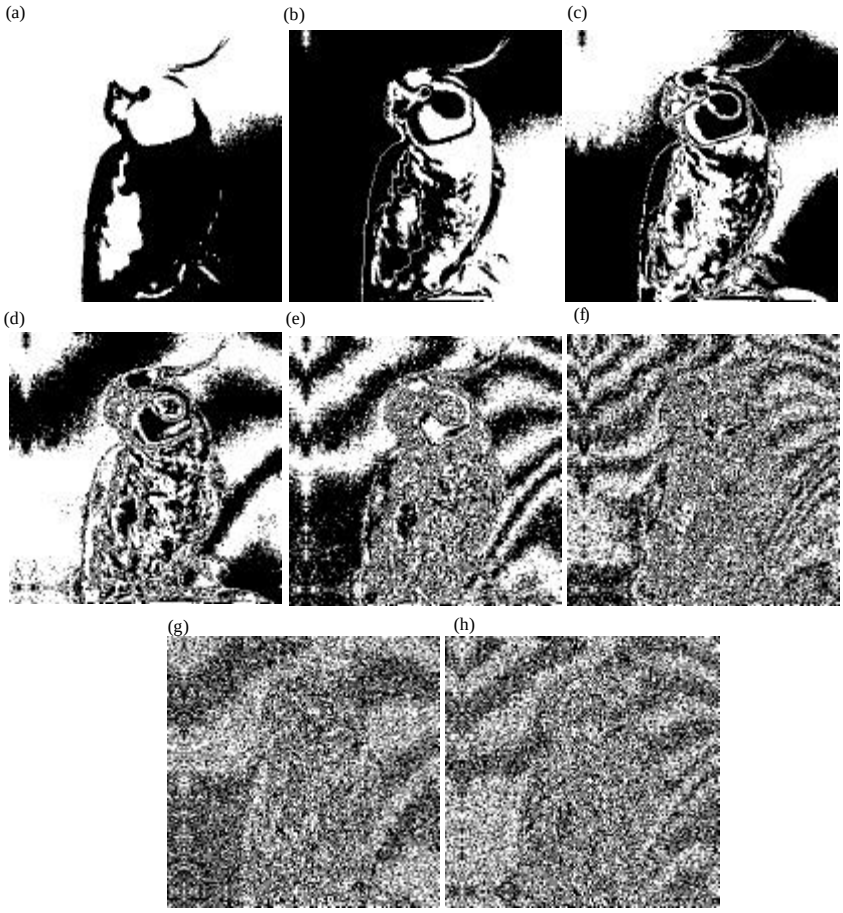

Fig. 4(a-h): Bit plane decomposition of carrying secret image, (a) Eighth bit plane, (b) Seventh bit plane, (c) Sixth bit plane, (d) Fifth bit plane, (e) Fourth bit plane (f) Third bit plane, (g) Second bit plane and (h) First bit plane

Fig. 4, respectively. From those figures, the LSB plane and low bit plane can save the nun-stochastic feature of the image and they are basically remaining the same with corresponding bit planes of Fig. 1. Hence it can defend steganalysis detection that based on whether the LSB plane and other low bit plane has stochastic characteristic.

Other images that have obvious texture characteristic on low bit planes can save the texture feature after information has been hidden. For example, four low bit planes of Fig. 3c are shown in Fig. 5a-d and non-stochastic characteristic maintained very well.

Run length of bit plane will change after information is hidden. This change between the four images in Fig. 3 and four images in Fig. 2 are illustrated from Fig. 6a to f. After information is embedded, neither significant increase nor decrease of the number of run, among different length, is observed from Fig. 6. No secret information is detected by using run length detection algorithm that proposed in Zhou and Yang (2005). Therefore it can defend several of steganography detection method that is based on run length statistic characteristic.

Then use 256×256×1 standard binary test images "Circle" and "Soil" as shown in Fig. 7a and b to realize full insertion and extraction experiments. Images with information inserted are shown in Fig. 8a and b. The hiding capacities of image "Circle" is 2961 and image "Soil" is 2283. The extraction accuracy is 100%. The imperceptibility evaluation results of those two images are as follow: for image "Circle", the CSF is 55.41 and the WPSNR is 43.98; for image "Soil", the CSF is 41.75 and the WPSNR is 41.12.

Use Chi-square (χ^2) analysis method, Regular and Singular groups method (RS) analysis method, Sample Pair Analysis (SPA) analysis method and Gray-level Plane

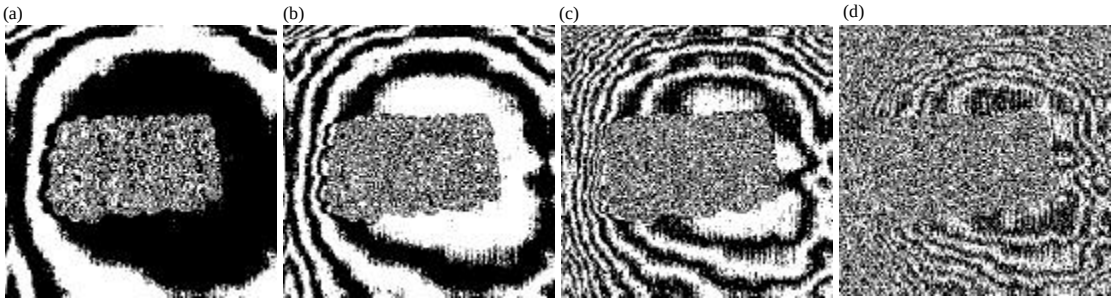

Fig. 5(a-d): Four low bit planes of secret carrying image Bead, (a) Fourth bit plane (b) Third bit plane, (c) Second bit plane and (d) First bit plane

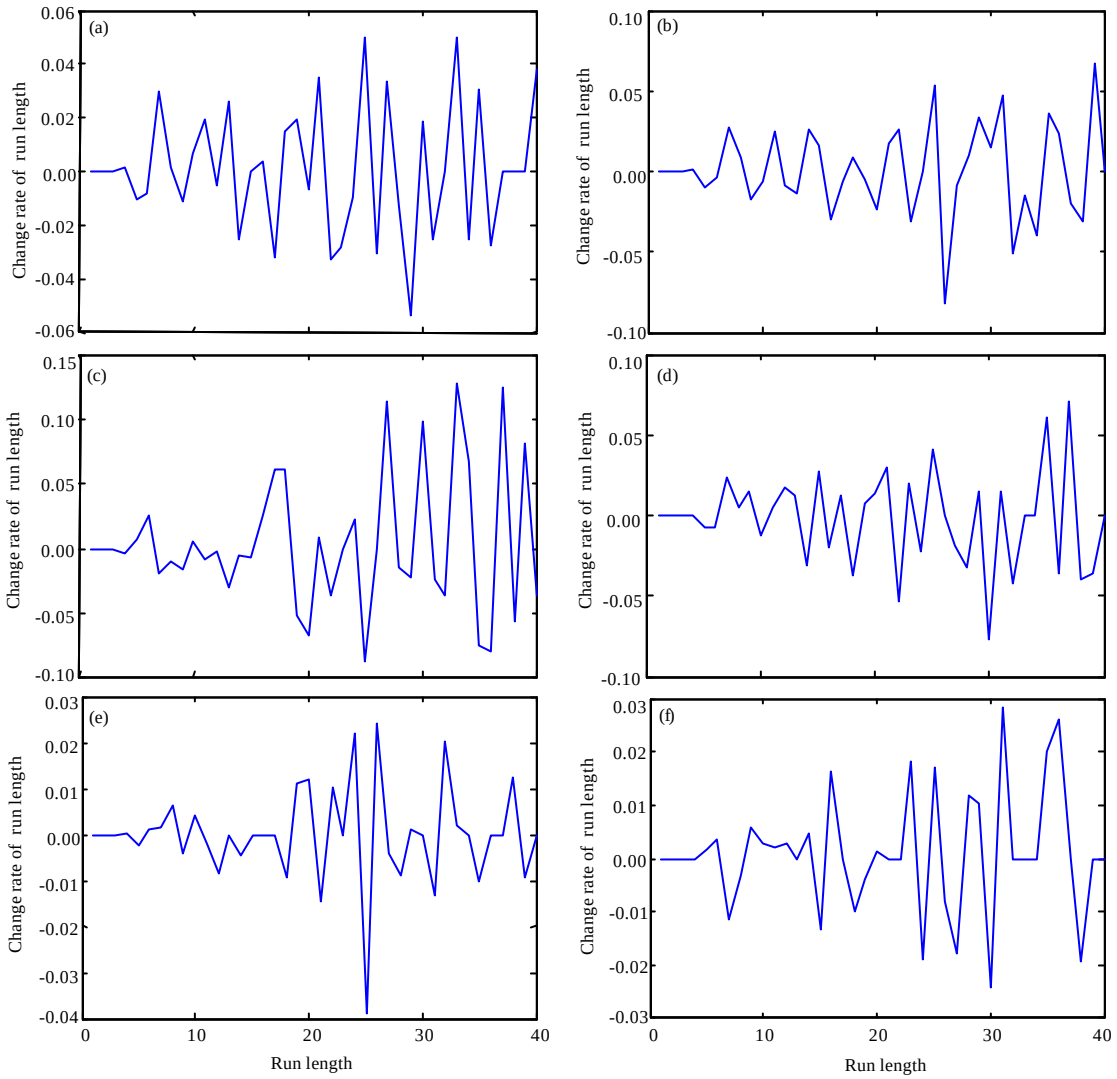

Fig. 6(a-f): Rate of change of partial run length with information embedded, (a) Bird , (b) Lena, (c) Bead, (d) House, (e) Mandrill and (f) Cavas

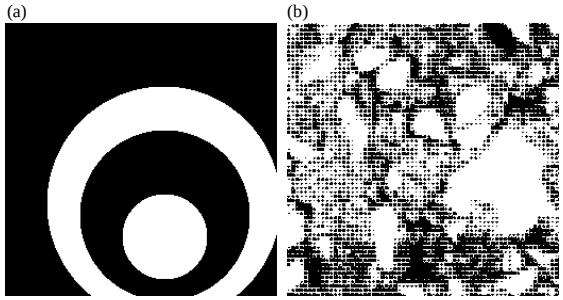

Fig. 7(a-b): Two standard binary images for test, (a) Circle and (b) Soil

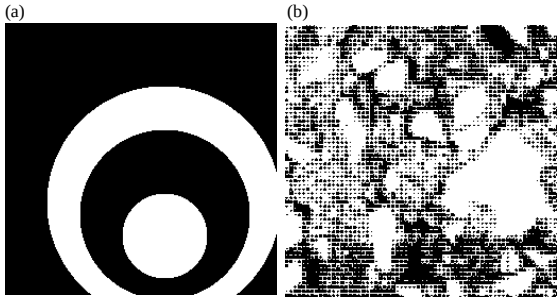

Fig. 8(a-b): Two stego-images after information has been embedded in the standard binary images, (a) Circle and (b) Soil

Crossing Analysis (GPC) analysis method to detect steganography, no secret information can be detected either. Consequently, the algorithm proposed in this study is adaptive to binary information hiding as well and it has better imperceptibility, ability to against detection and can defend steganalysis effectively.

Hiding capacity of image showed in Fig. 8 is equivalent to the 3×3 partition embedding algorithm's hiding capacity, which is relatively larger then other algorithms. Meanwhile, the imperceptibility of this algorithm is better than the partition algorithm and it will not arouse decrease of number of long run, hence it is safer than partition algorithm. But if one binary image has relatively thinner text line, then hiding capacity is smaller and saw tooth phenomenon may occur along vertical fringe of some smooth vertical line. This algorithm needs to be improved while be utilized on binary images.

The distribution of run of the binary image in Fig. 7 will change after information hiding and the most severe change is shown in Fig. 9. The figure shows that there is no significant increase or decrease among the total number of all kinds of run length. Hence the algorithm that

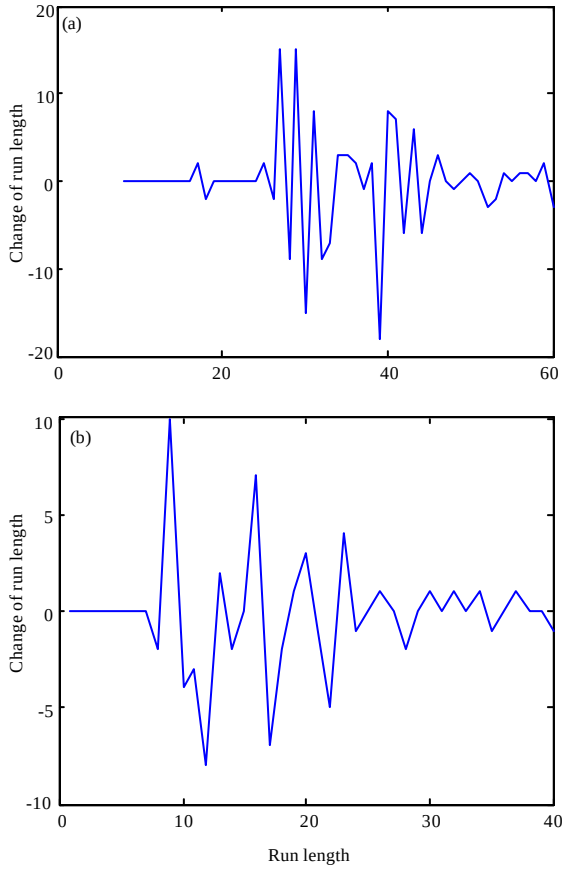

Fig. 9(a-b): Change of run length of the binary image after information hiding, (a) Circle and (b) Soil

is proposed in this study can be applied to binary images and can also defend many steganalysises that based on statistic feature of run length.

CONCLUSION

By using human vision characteristic, a kind of hiding algorithm is proposed in this study based on run length. Decomposing gray or color image into multiple binary images first, one can judge if every white-black run couple is able to embed information. In the positive case, the algorithm can insert information at the end of one long run, with no more than one pixel changed. The algorithm will not cause noise property on every bit plane and reducing of number of long run. It can also effectively defend steganalysis method like Chi-square (χ^2) analysis method, RS analysis method and GPC analysis method that makes the algorithm has better security. Meanwhile, since the algorithm can embed information into every bit plane including the highest bit plane, the algorithm has

larger hiding capacity, which makes the algorithm satisfies larger capacity commanded by some application. Simulation results show that the algorithm proposed in this study has stronger security, better imperceptibility and larger hiding capacity. The algorithm can be used in situation that requires larger hiding capacity and stronger security like secret communication.

ACKNOWLEDGMENTS

This study is supported by Educational Science Subject of Hunan 12-th Five-Year Plan under Grant No., XJK011BXJ008. Educational Science Subject of Hunan, 12C0560. Science and Technology Programs of Hunan Province, 2012GK3064, 2012GK4006. Supported by the Construct Program of the Key Discipline in Hunan Province.

REFERENCES

Amirtharajan, R. and J.B.B. Rayappan, 2012. An intelligent chaotic embedding approach to enhance stego-image quality. Inform. Sci., 193: 115-124.

Bandyopadhyay, S.K., A. Raychoudhury and T.U. Paul, 2010. A palette based approach for invisible digital watermarking using the concept of run-length. Proceedings of the International Conference on Computational Intelligence and Communication Networks, November 26-28, 2010, Bhopal, India, pp: 83-87.

Chen, G.X. and J.J. Chen, 2011. Research security for batch steganography. J. Chin. Comput. Syst., 32: 644-646.

Fillatre, L., 2012. Adaptive steganalysis of least significant bit replacement in grayscale natural images. IEEE Trans. Signal Process., 60: 556-569.

Jiang, C.X., X.W. Chen and Z. Li, 2010. Robust text watermarking based on significant components. Acta Automatica Sinica, 36: 1250-1256.

Ker, A.D., 2007. Steganalysis of embedding in two least-significant bits. IEEE Trans. Inform. Forensics Secur., 2: 46-54.

Lee, Y.P., J.C. Lee, W.K. Chen, K.C. Chang, I.J. Su and C.P. Chang, 2012. High-payload image hiding with quality recovery using tri-way pixel-value differencing. Inform. Sci., 191: 214-225.

Lei, Y., X.Y. Yang, X.Z. Pan and D.T. Guo, 2010. YASS steganalysis based on local randomness. Chin. J. Comput., 33: 1997-2002.

Li, H., Y. Lu, H. Cui and K. Tang, 2009. Image quality assessment based on frequency domain of structural similarities. J. Tsinghua Univ. Sci. Technol., 49: 559-562.

Liu, C., G. Liang and S. Wang, 2007. Steganography in color images using a binary image data hiding scheme. J. Applied Sci., 25: 342-347.

Lou, D.C. and C.H. Hu, 2012. LSB steganographic method based on reversible histogram transformation function for resisting statistical steganalysis. Inform. Sci., 188: 346-358.

Lu, X., Y. Cao, P. Lu and A. Zhai, 2012. Digital audio information hiding based on Arnold transformation and double random-phase encoding technique. Optik-Int. J. Light Electron Opt., 123: 697-702.

Luo, X.Y., P.Z. Lu and F.L. Liu, 2007. A dynamic compensation LSB steganography method defeating SPA. Chin. J. Comput., 30: 463-473.

Ma, X. and J. Lin, 2011. Performance evaluation of information hiding. J. Image Graphics, 16: 209-214.

Nezhadarya, E., Z.J. Wang and R.K. Ward, 2011. Robust image watermarking based on multiscale gradient direction quantization. IEEE Trans. Inform. Forensics Secur., 6: 200-1213.

Xie, J., Q. Xie and D. Huang, 2011. Study on imperceptibility index of information hidden based on image. J. Chin. Comput. Syst., 32: 953-957.

Yang, C.N., J.F. Ouyang and L. Harn, 2012. Steganography and authentication in image sharing without parity bits. Optics Commun., 285: 1725-1735.

Zeng, X.T., Z. Li and L.D. Ping, 2012. Reversible data hiding scheme using reference pixel and multi-layer embedding. AEU Int. J. Electron. Commun., 66: 532-539.

Zhang, Z., G. Liu, Y. Dai and Z. Wang, 2010. A novel second-order distribution maintained steganographic algorithm based on Markov chain security. J. Image Graphics, 15: 1175-1181.

Zhou, J. and Y. Yang, 2005. A detecting algorithm based run -length to LSB embedding in images. J. Xi'an Univ. Post Telecommun., 10: 1-5.

Zhou, X., S. Wang, S. Xiong and J. Yu, 2010. Attack model and performance evaluation of text digital watermarking. J. Comput., 5: 1933-1941.