A NEW STEGANOGRAPHY TECHNIQUE USING KNIGHT'S TOUR ALGORITHM, AFFINE CIPHER AND HUFFMAN CODING

ALLA TALIB MOHSIN

A dissertation submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

JULY 2013

This dissertation is dedicated to my family especially my father, my mother, my wife, my son and my friends for their endless support and encouragement.

ACKNOWLEDGEMENTS

All praise is to Allah and my peace and blessing of Allah be upon our prophet, Muhammad and upon all his family and companions. In particular, I wish to express my sincere appreciation to my thesis supervisor, Prof. Dr. Ghazali Bin Sulong, and to all my friends for encouragement, guidance, critics, advices and supports to complete this research.

In addition, I am extremely grateful to my father for unlimited support and encouragement during this research. I would like to thanks my beloved family: mother, fiancée, brother, and sisters who I am always beholden to them for their everlasting patience, support, encouragement, sacrifice, and love which have been devoted to me sincerely, so. I could endure for being away and eventually my master program came to fruition. For that, I ask Allah to bless all of them.

Besides, I would like to thank the authority of Universiti Teknologi Malaysia (UTM) and faculty of computing (FK) for providing me with a good environment and facilities such as Computer laboratory to complete this project with software which I need during process.

ABSTRACT

Steganography is a way to conceal secret message under a cover image, and is one of the most widespread approaches of securing information. The challenge of steganographic methods is to find a balanced trade-off between visual quality and capacity. In addition, robustness and security are also important that need to be considered. The Least Significant Bit (LSB) insertion approach is one of the most popular steganographic techniques, which can provide both high visual quality and large capacity. However, the LSB approach suffers from robustness and security. Thus, this study proposes a new steganography approach that involves five phases: (i) RGB colour cover image is converted to YCbCr colour space. Then, Cb component is selected to hide the secret data. (ii) The secret message is encrypted by using the Affine Cipher. (iii) The encrypted message is then compressed using the Huffman Coding. (iv) The Knight Tour algorithm is applied to embed the secret code onto the Cb component using LSB approach. (v) Once completed, the Cb is merged with the Y and Cr components and is subsegueatiy converted to RGB image to obtain a stego-image. Experimental results showed that not only has the proposed method improved the security and capacity as compared to the last LSB method, it has also increased the visual quality of the stego-image.

ABSTRAK

Steganografi merupakan satu cara untuk menyembunyikan pesanan rahsia dengan menggunakan imej sebagai samaran, dan ianya digunakan secara meluas untuk keselamatan maklumat. Cabaran steganografi adalah untuk mendapatkan keseimbangan diantara kualiti visual dan kapasiti. Selain itu, kelasakan dan keselamatan juga mustahak yang perlu dipertimbangkan. Pendekatan kemasukan Bit Signifikan Terkecil (BST) merupakan salah satu teknik yang popular yang mampu menghasilkan kualiti visual yang tinggi dan kapasiti yang besar. Walaubagaimanapun, pendekatan BST tersebut kecundany dari segi kelasakan dan keselamatan. Oleh itu, kajian ini mencadangkan satu pendekatan baru steganografi yang melibatkan lima fasa: (i) Imej warna *RGB* ditukar kepada ruang warna *YCbCr*, kemudian komponen *Cb* dipilih untuk menyembunyikan pesanan rahsia. (ii) pesanan rahsia dienkripkan menggunakan *Affin Cipher*. (iii) pesanan yang telah dienkrip kemudian dimampatkan menggunakan Pengkodan *Huffman*. (iv) Algoritma *Knight Tour* digunakan untuk membenamkan kod rahsia tersebut kedalam komponen *Cb* menggunakan pendekatan BST. (v) Setelah ianya selesai, komponen *Cb* digabungkan semula dengan komponen-komponen *Y* dan *Cr*, dan seterusnya ditukar kepada imej warna *RGB* bagi mendapatkan imej stego. Hasil eksperimen menunjukkan bahawa pendekatan yang dicadangkan itu bukan sahajc dapat menambahbaik keselamatan dan kapasiti, malah ianya juga dapat mempertingkatkan kualiti visual berbanding dengan pendekatan BST klasik.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

LSB	Least significant Bit
ISB	Intermediate Significant Bit
MSB	Most significant Bit
EMD	Exploiting Modification Direction
DWT	Discrete Wavelet Transform
DCT	Discrete Cosine Transform
PSNR	Peak Signal to Noise Ratio
MSE	Mean Squared Error
NCC	Normalized Cross Correlation
HVS	Human Visual System
LZW	Lempel Ziv Welch
GIF	Graphic Interchange Format
PNG	Portable Network Graphics
JPEG	Joint Photographic Expert Group
TIFF	Tagged Image File Format

CHAPTER 1

INTRODUCTION

1.1 Overview

The need for methods that provide efficient work on the protection of data and the private property of individuals has become necessary due to the huge growth of multimedia applications on networks. It is therefore important to create methods that provide security for the media to protect them from thieves and hackers and to prevent them from tampering and misrepresenting the data. There is therefore the need to devise techniques of data security.

Data protection consists of two techniques: cryptography and data hiding. Cryptography means the provision of protection for data storage and using a secret key data transfer. Encryption is still a successful way to protect data stored and transmitted over a network. But, with the growing use of networks to send and receive data on the global information network, it has become very difficult to maintain this data.

Data hiding has two main approaches: steganography and digital watermarking. These two approaches have many techniques (Zaidane, *et al*., 2010), one of them LSB. Steganography is defined as the art and science of communicating in a way which hides the existence of the communication. Also,

watermarking is defined as the practice of imperceptibly altering a work to embed a message about that work (Cox, *et al.* 2006).

There are two procedures that are used in steganography. The first one is embedder, for this procedure there are two inputs: payload and cover image (host image). Payload means the quantity of the secret message that is going to embed. Whereas, the cover image is used as a cover contains the message inside it. After the embedding process is completed the resulting image is now called stego-image, and will be transmitted to the receiver. The second procedure is detector, the input for this procedure is stego-image, and the detector can recognize the secret message by using extraction process (Cox, *et al*. 2006). As result, the stenography is considered one of the protection methods of the secret information that used the host media as a cover for instance: text, images, audio or video.

1.2 Background of the problem

The basics of steganography are imperceptibility, robustness and data rate (payload). The required trade-offs between the image quality and the payload. The steganography algorithm must be at an acceptable level.

In steganography, it is very hard to embed large amount of data and preserve the high image quality at the same time. Therefore, if it is required to have more payload steganography algorithm, its image quality will be low and conversely. Steganography algorithms are usually not efficient with high payload embedding (Cvejic, 2004).

The steganography technique is mostly divided in two groups: However others divide it into three types: spatial domain (Chan and Cheng 2004; Bender, *et al*. 1996; Chang, *et al.*, 2003), frequency domain (Inoue *et al*. 1999; Du and Hsu 2003; In, *et al*., 1999; Nasrabadi and King, 1988) and the compression domain (Lu,

et al., 2000). The former embeds messages directly into the image pixels. One of the most popular methods is called Least Significant Bit (LSB). Many researchers used it to embed data in the cover image because this method can embed high payload and at the same time can preserve a good quality of the image. But, LSB method is vulnerable to attacks such as statistical attack (histogram and Chi-square).

Other researchers used Intermediate Significant Bit (ISB) planes [Zeki, and Manaf, (2009); Yan, et al., (2009); Perumal, and Kumar, (2011), Mehemed, et al., (2009); Emami and Sulong (2011)] to overcome the LSB drawbacks.

The Exploiting Modification Direction (EMD) is proposed by (Xinpeng and Shuozhong, 2006). In this method, the cover image is segmented into many groups and each group contains *n* of pixels to embed secret message. During the embedding stage, it is required to increase or decrease one from the value of a particular pixel within the group. The limitation of this method is that the image still has a low quality because the size of the group consists of two pixels.

(Kai Yung, *et al*., 2010), proposed the relationship between the value of *n* and the amount of payload that minimizes cover image distortion. Thus, the imperceptibility is high whereas it is affected on the mount of the payload.

The second technique is the frequency domain. This technique embeds messages in the frequency coefficients of images by using a frequency-oriented mechanism such as Discrete Cosine Transform (DCT) (Chen, *et al.* 1999) Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) (Munteanu, *et al.* 1999).

Transformation domain methods hide messages in a significant area of the cover image that makes them more robust to attacks (such as compression, and some image processing operations) and imperceptibility than the LSB approach.

The other study (Ullah, 2010) introduced a new way in steganography that is different from the previous methods of steganography. Confidentiality and compression of large text using steganography is presented. The first process in this approach is to compress reduce the size of the secret message. And then, encode the compressed text in the cover image using a shared key between sender and receiver.

Generally, embedding in the low frequency can withstand many attacks but provide clear effect on the media. In contrast, embedding on a high frequency provides less impact on the quality of images but creates low robustness to different attacks (Shjul and Kulkarni, 2011).

Proposed using Discrete Wavelet Transform (DWT) and Huffman coding with 3 bits LSB in steganography. At first, compress the secret message using Huffman coding. And then convert the image to the frequency domain by using DWT. The human eye is very sensitive to the small changes in the edges but cannot detect the modifiers in the smooth parts. Therefore, this technique embedded the data in a high frequency sub-bands to increase the robustness of the embedding process (Nag, *et al*., 2009).

The last one is the compression domain. Here the secret information is embedded into the image compression such as JPEG, Block Truncation Coding (BTC) (Chen, 1999), and other methods which are used to embed data for extending the types of cover images. The drawbacks of their schema appear when using a large amount of secret data; for example, 64Kb. In this case, the embedding process requires a long time (Shie, *et al.*, 2006).

Now the question is: How can we increase the amount of the imperceptibility and robustness while maintaining the high quality of image?

The need to devise a new technique that has the ability to preserve high image quality and to embed a large payload is very important. Therefore, in this research, the LSB method and Affine cipher and Huffman coding are applied to embed a large amount of the secret data with privacy of information and at the same time give a high image quality as compared with previous methods.

1.3 Problem Statement

As the computer and communication technology grow up rapidly, the digital contents like images, audio, and video are distributed easily to the internet. However, this also causes substantial financial damage and becomes an imperative concern of copyright protection.

With regard to high payload, previous studies proposed various techniques including LSB, DCT, DWT and EMD. However, the results of the methods revealed that when the payload is increased the quality is: decreased and vice versa. Now, the question is how to obtain high capacity and robustness without sacrificing the stego-image quality.

- Huffman Coding is a popular data compression technique to reduce the size of storage. That means the compression can be used to increase the capacity. The question is: how to integrate this compression technique with steganography.
- With regard to robustness, previous studies proved that most of the steganography techniques are vulnerable to various attacks especially Chi-square attack. The question here is: How to design a robustness method that can withstand such a severe attack.

1.4 Aim of the study

The aim of this research is to propose a technique to hide information in color image using an improved LSB technique, affine cipher and Huffman coding algorithms. The aims of the proposed hiding data technique are achieving, improving the imperceptibility and reducing distortion on image. Furthermore, the system is producing stego-image similar to original image in term of human visual system (HVS) measured by Signal-to-Noise Ratio (PSNR).

1.5 Objectives of the study

This research intends to perform these objectives:

1. To improved steganography technique on color image based on LSB method.
2. To apply affine cipher on secret message to increase the security of the proposed technique.
3. To integrate the Huffman coding with the proposed method to increase the capacity.

1.6 Scope of the study

- Cover images are taken from standard dataset of USC-SIPI Data include 24-bit color in spatial domain with size (512*512).
- Secret messages including lower or upper cases letters of the alphabet are arbitrarily generated.
- Affine cipher technique is used for crypto the secret message.
- Compression: Huffman coding is used to compress the secret message.

- Evaluation: Peak Signal-to-Noise Ratio (PSNR) and NCC is used to evaluate the imperceptibility and robustness.

1.7 Significant of the study

The needing for security system in transferring the data between the sender and receiver has much of value. Therefore, implementing some secure techniques is very important. This thesis studied two techniques affine cipher coding algorithm and Huffman method for compression to increase the security and the amount of embedding data. This technique has good results in two terms imperceptibility and robustness. In addition, the system is considering complex. As a result, it is very difficult to reveal the message inside the image.

1.8 Research Framework

This study starts with the collection of requirements such as images and prepares the secret messages. The second step involves the design and implementation of the proposed method which include embedding process and extraction process. The next step is the evaluation of the proposed method by measuring the imperceptibility and robustness. Lastly, the other step is the result of the discussion to explain and compare the results of the proposed method with the previous methods. Figure 1.1, depicts the framework of this study.

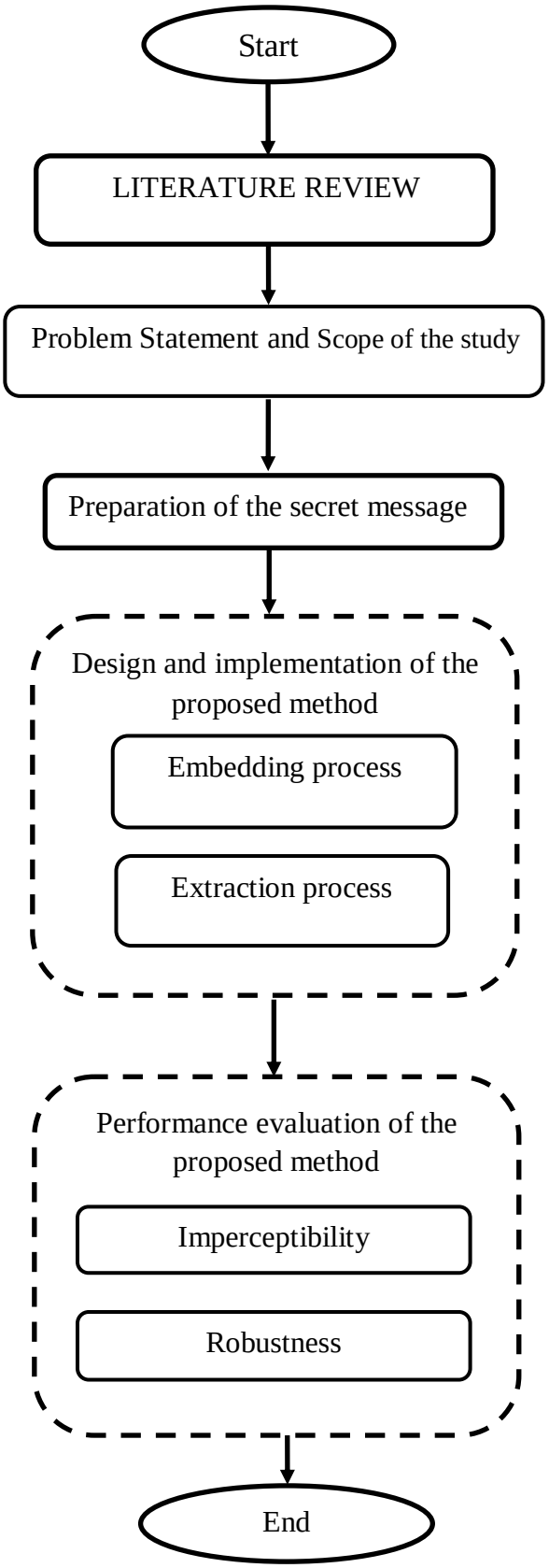

Figure 1.1 Framework of the study

REFERENCES

Chang, C.C., Hsiao, J.Y. and Chan, C.S. (2003).Finding optimal least-significant-bit substitution in image hiding by dynamic programming strategy. Pattern Recognition, 36(7), 1583-1595. doi: 10.1016/s0031-3203(02)00289-3.

Chen, B., Latifi, S. and Kanai, J. (1999).Edge enhancement of remote sensing image data in the DCT domain. Image and Vision Computing, 17(12), 913-921. doi: 10.1016/s0262-8856(98)00165-6.

Connell, J. B. (1973). A Huffman-Shannon-Fano code. Proceedings of the IEEE, 61(7), 1046-1047.

Cox, I. J., Miller, M. L. and Bloom, J. A(2002).Digital Watermarking. Morgan Kaufmann, 50-55.

Cox, I. J., Miller, M.L., Bloom, J. A., Fridrich, J. and Kalker, T. (2006). Digital and watermarking (2nd ed.). USA: ELSEVIER.

CVEJIC, N. (2004). Algorithm for Audio Watermarking and Steganography. University of Oulu, Oulu, Finland.

Emami, M. andSulong, G. B. (2011). A Statistical Method based on L2Norm Technique for EISB Information Watermarking Scheme. Proc. of International Conference on Future Information Technology IPCSIT, 13, 139-143.

Inoue, H., Miyazaki, A. and Katsura, T. (1999). An image watermarking method based on the wavelet transform. Proceedings of the 1999 Image Processing, 1999.ICIP 99.Proceedings.1999 International Conference on. 1999. 296-300 vol.291.

Kai Yung, L., Wien, H., Chen, J., Tung Shou, C. and Wen Chin, C. (2010).Data hiding by Exploiting Modification Direction technique using optimal pixel grouping.Proceedings of the 2010 Education Techno+logy and Computer (ICETC), 2010 2nd International Conference on. 22-24 June 2010. V3-121-V123-123.

Katzenbeisser, S. andPetitcolas, F. (2000). Information Hiding Techniques for Steganography and Digital Watermarking: Artech House, Inc.

Li, B., He, J., Huang, J. and Shi, Y. Q. (2011).A survey on image steganography and steganalysis. Journal of Information Hiding and Multimedia Signal Processing, 2(2), 142-172.

Mehemed, B. A., El-Tobely, T. E. A., Fahmy, M.M., Said Nasr, M. E. L. and El-Aziz, M. H. A. (2009). Robust digital watermarking based falling-off boundary in corners board-MSB-6 gray scale images. International Journal of Computer Science and Network Security, 9(8), 227-240.

Munteanu, A., Cornelis, J., Van Der Auwera, G. and Cristea, P. (1999).Wavelet image compression - the quadtree coding approach. Information Technology in Biomedicine, IEEE Transactions on, 3(3), 176-185.

Nag, A., Biswas, A., Sarkar, D., and Sarkar, P. P. (2009). A Novel Technique for Image Steganography Based on DWT and Huffman Encoding. International Journal of Computer Science and Security (IJCSS), 4(6), 561-570.

Perumal, S. M. and Kumar, V. V. (2011). A Wavelet based Digital Watermarking Method using Thresholds on Intermediate Bit Values. International Journal of Computer Applications, 15(3), 29-36.

Stallings, W. (1999). Cryptography and Network Security: Principles and Practice (edition 2nd ed.): Prentice-Hall, Inc.

Ullah, F., Naveed, M.,Inayatullah, B. M. andIqbal, F. (2010).Novel Use of Steganography for Both Confidentiality and Compression. IACSIT International Journal of Engineering and Technology, 2(4), 361-366.

Xinpeng, Z. and Shuozhong, W. (2006).Efficient Steganographic Embedding by Exploiting Modification Direction. Communications Letters, IEEE, 10(11), 781-783.

Yan, D., Yang, R., Yu, Y. and Xin, H. (2009). Blind Digital Image Watermarking Technique Based On Intermediate Significant Bit and Discrete Wavelet

Transform. Proceedings of the 2009 Computational Intelligence and Software Engineering, 2009.CiSE 2009.International Conference on. 11-13 Dec. 2009. 1-4.

Zeki, A. M. andManaf, A. A. (2009). A Novel Digital Watermarking Technique Based on ISB (Intermediate Significant Bit). International Journal of Information Technology, 5, 989-996.

Lin-Yu, T., Yung-Kuan, C., Yu-An, H. and Yen-Ping, C. (2008).Image Hiding with an Improved Genetic Algorithm and an Optimal Pixel Adjustment Process.Proceedings of the 2008 Intelligent Systems Design and Applications, 2008.ISDA '08.Eighth International Conference on. 26-28 Nov. 2008. 320-325.

Mathkour, H., Al-Sadoon, B. and Touir, A. (2008).A New Image Steganography Technique.Proceedings of the 2008 Wireless Communications, Networking and Mobile Computing, 2008.WiCOM '08.4th International Conference on. 12-14 Oct. 2008. 1-4.

Michael, B. andCachin, C. (2004). Public-Key Steganography with Active Attacks. IBM Research, 1-16.

Morkel, T., Eloff, J. H. P. and Olivier, M. S. (2005). AN OVERVIEW OF IMAGE STEGANOGRAPHY.Proceedings of the 2005 Information Security South Africa (ISSA2005). June-July 2005. Sandton, South Africa, 1-12.

Rabah, K. (2004). Steganography -- The Art of Hiding Data. Information Technology Journal, 3(3), 245-269.

Shjul, A. A. andKulkarni, U. L. (2011). A secure skin tone based steganography Using wavelet transform. International Journal of computer theory and Engineering, 3(1), 16-22.

Wu, H. C., Wu, N. I., Tsai, C. S. and Hwang, M. S. (2005). Image steganographic scheme based on pixel-value differencing and LSB replacement methods. Vision, Image and Signal Processing, IEE Proceedings -, 152(5), 611-615.

Yang, H., Sun, X. and Sun, G. (2009).A High-Capacity Image Data Hiding Scheme Using Adaptive LSB Substitution. RADIOENGINEERING, 18(4), 509-516.

Zaher, M. A. (2011). Modified Least Significant Bit (MLSB). Computer and Information Science, 4(11), 60-67.

Anand, V. J. and Dharaneetharan, G.D. (2011).New Approach in Steganography by Integrating Different LSB Algorithms and Applying Randomization Concept to Enhance Security Rourkela. Paper presented at the ICCCS'11, International Conference on Communication, Computing & Security Odisha, India.

Chan, C.K. and Cheng, L.M. (2004).Hiding data in images by simple LSB substitution. Pattern Recognition, 37(3), 469–474.

Chang, C.C., Chou, Y.C. and Lu, T.C. (2007). A semi-blind watermarking based on discrete wavelet transform. Paper presented at the Proceedings of the 9th international conference on Information and communications security, Zhengzhou, China.

Manjunatha, P. R. and Koliwad, S. (2009). A Comprehensive Survey of Contemporary Researches in Watermarking for Copyright Protection of Digital Images. *IJCSNS International Journal of Computer Science and Network Security*. 9 (4). P 91-107.

Almohammad, A. & Ghinea, G. Year. Stego image quality and the reliability of PSNR. In: Image Processing Theory Tools and Applications(IPTA),2010 2nd International Conference on, 7-10 July 2010 2010. 215-220.

Anand, J. V. & Dharaneetharan, G. D. 2011. New approach in steganography byvintegrating different LSB algorithms and applying randomization concept to enhance security. Proceedings of the 2011 International Conference on Communication, Computing & Security. Rourkela, Odisha, India: ACM.

Dumitrescu, S., Xiaolin, W. & Zhe, W. 2003. Detection of LSB steganography via sample pair analysis. Signal Processing, IEEE Transactions on, 51, 355-372.

Parberry, I. 1997. An efficient algorithm for the Knight's tour problem. Discrete Applied Mathematics, 73, 251-260.

Xiangyang, L., Bin, L. & Fenlin, L. Year. Detecting LSB steganography based on dynamic masks. In: Intelligent Systems Design and Applications, 2005. ISDA '05. Proceedings. 5th International Conference on, 8-10 Sept. 2005 2005. 251-255.

Ali, D. B.(2004).Digital Image Watermarking Techniques for Copyright Protection. Ph. D., University of Mosul, Mousl.

Al-laham, M. and El Emary, I. M. M. (2007). Comparative Study Between Various Algorithm of Data Compression Techniques. Paper presented at the WCECS, San Francisco, USA.

Kharrazi, M., Husrev, T. S. andMemon, N. (2004).Image Steganography and Steganalysis. Concepts and Practice, 2939, 35-49.

Lee, Y.K., Bell, G., Huang, S.Y., Wang, R.Z. and Shyu, S.J. (2009).An AdvancedLeast-Significant-Bit Embedding Scheme for Steganographic Encoding.Advances in Image and Video Technology. Berlin / Heidelberg: Springer, 349-360.

Westfeld, A. and Pfitzmann, A. (2000).Attacks on Steganographic Systems - Breaking the Steganographic Utilities EzStego.Jsteg, Steganos, and S-Tools -and Some Lessons Learned. [Lecture Notes in Computer Science]. Springer-Verlag, 61-75.

\ Zaidan, A. A., Zaidan, B. B.,Taqa Y. A., Sami, M. K., Alam, G. M. and Jalab, A. H. (2010). Novel multi-cover steganography using remote sensing image and general recursion neural cryptosystem. International Journal of Physical Sciences, 5(11), 1776-1786.

Wu, H. C., Wu, N. I., Tsai, C. S. & Hwang, M. S. 2005. Image steganographic scheme based on pixel-value differencing and LSB replacement methods. Vision, Image and Signal Processing, IEE Proceedings -, 152, 611-615.

Tariq Al, H., Mahmoud Al, Q. & Hassan, B. Year. A testbed for evaluating security and robustness of steganography techniques. In: Circuits and Systems, 2003 IEEE 46th Midwest Symposium on, 27-30 Dec. 2003 2003. 1583-1586 Vol.

Chai, P., Liu, J., Pei, D. & Yang, Z. Year. LPC Prediction Error Combined with LSB Steganography for Blind Speech Authentication. In: Multimedia Signal Processing, 2005 IEEE 7th Workshop on, Oct. 30 2005-Nov. 2 2005 2005. 1-4.

Roque. J.J. , M. J. M. Year. SLSB: Improving the Steganographic Algorithm LSB. In: 7th International Workshop on Security in Information Systems, WOSIS, 2009 Milan, Italy. INSTICC Press, 57-66.

Gillman, D. W., Mohtashemi, M. & Rivest, R. L. 1996. On breaking a Huffman code. Information Theory, IEEE Transactions on, 42, 972-976.