Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 5, May 2014, pg.821 – 826

SURVEY ARTICLE

Reversible Data Hiding and its Methods: A Survey

Sukhdeep Kaur[#1], Manshi Shukla[*2]

#M. Tech, Research Scholar, Department of Computer Science and Engineering, RIMT-IET, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India

[1] sukhdeepkaur1991@gmail.com

*Assistant Professor, Department of Computer Science and Engineering, RIMT-IET, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India

[2] manshishukla07@gmail.com

Abstract — Hiding information destroys the host image even though the distortion introduced by hiding is imperceptible to the human visual system. Reversible data hiding techniques are designed to solve the problem of lossless embedding of large messages in digital images so that after the embedded message is extracted, the image can be restored completely to its original state before embedding occurred. This paper presents a survey on various reversible data hiding methods.

Keywords — Reversible data hiding (RDH); LSB; Difference expansion; Histogram modification; Cover image

I. INTRODUCTION

Today, in the digital era, any sort of data such as images, text, audio, can be digitized and stored indefinitely, and can be transmitted at high speeds. Therefore there is a need to hide secret identification inside certain types of digital data. This information can be used to identify attempts to tamper with sensitive data, to embed annotations and to prove copyright ownership. Storing, hiding, or embedding secret information in all types of digital data is one of the tasks of the field of steganography. Secret data can be embedded in various types of cover. If the data are embedded in an image (cover image), the result is a stego-image (or stegoimage) object. The data can also be embedded in text file, audio, video etc. Embedding data in a cover is a technological challenge. The size of the embedded data should not increase the size of cover as it becomes noticeable to an attacker who is familiar to the original cover. Therefore secret data should be embedded in "holes" in the cover (places where the cover data have redundancies).

II. REVERSIBLE DATA HIDING

Data hiding is a term encompassing a wide range of applications for embedding messages in content. Usually, hiding information destroys the host image even though the distortion introduced by hiding is imperceptible to the human visual

system. However, there are some sensitive images for which any embedding distortion of the image is intolerable, such as medical images, military images or artwork preservation. For images like in medical field, even slight changes are unacceptable because of the potential risk of a physician misinterpreting the image. In other applications, such as remote sensing it is also desired that the original cover media can be recovered because of the required high-precision nature. In these cases a special kind of data hiding method called reversible data hiding or lossless data hiding is used. Reversible data hiding (RDH) techniques are designed to solve the problem of lossless embedding of large messages in digital images so that after the embedded message is extracted, the image can be restored completely to its original state before embedding occurred.

Steps of RDH

- Data Embedding
- Data Extraction

To embed the data in the image we need these inputs:

- The data to be embedded i.e. secret data.
- The cover data (cover image or host image)
- The key.

By combining these a suitable algorithm is generated which produce a stego image (stego cover) that can be stored or transmitted. To the other end the decoder or extractor receives the stego image and the stego key (optional) and extracts the data. In some algorithms the decoder work is only to check that data is actually embedded in the file or not. It is in the case where the hidden data are a watermark originally placed in the cover to prove ownership. The block diagram of reversible data hiding is shown in figure 1[1].

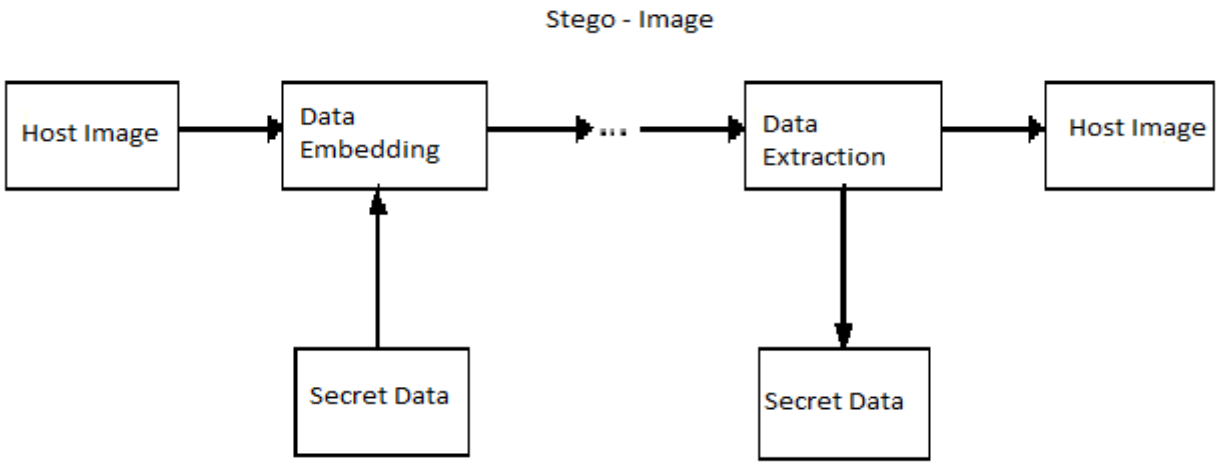

Figure 1. Reversible Data Hiding

An important feature of reversible data hiding is the reversibility, that is, one can remove the embedded data from the stego image to restore it to the original image. From the viewpoint of information hiding, reversible data embedding hides some information in a digital image such a way that only an authorized party could decode the hidden data/information and also could restore the image to its pristine, original state. The important metrics to determine the performance of reversible data-embedding algorithm are[2]

1) *Payload capacity limit:* determines the maximal amount of information that can be embedded
2) *Visual quality:* determines how is the visual quality on the embedded image
3) *Complexity:* determines the algorithm complexity

The motivation of reversible data embedding is distortion-free data embedding. Reasons behind data hiding is personal data, private data, sensitive data, confidential data ,to avoid misuse of data, unintensional damage to data , human error, accidental deletion, blackmail purposes, hide traces of a crime, for fun, trade secrets and many more.

III. METHODS IN RDH

A considerable amount of research on reversible data hiding has been done over the years. Three important methods are discussed here.

A. LSB substitution

The least-significant-bit (LSB) is the most widely used spatial domain data hiding technique. It generally embeds the same amount of data as the LSB pixels[3]. Least significant bit (LSB) is one of the simplest techniques that hides a secret message in the LSBs of pixel values without introducing many perceptible distortions. Changes in the value of the LSB are imperceptible for human eyes. In this substitution method of steganography the right most bit in a binary notation is replaced with a bit from the embedded message. LSB substitution replaces the least significant bit with a secret bit stream.

Least significant bit (LSB) insertion is a simple approach for embedding information in a cover image. The least significant bit (i.e. the 8th bit) of some or all of the bytes inside an image is changed to a bit of the secret message. In 24-bit image, a bit of each of the green, red and blue color components can be used, since each of them are represented by a byte. For example, a grid for 3 pixels consisting of a 24-bit image can be as follows:

00100101 00011100 11101001
00100110 00101111 11010111
01101100 00100010 11001011

If the number is 400, the binary representation is 110010000, is to be embedded into the least significant bits of this part of the cover image, then the resulting grid is(where underlined bits have been changed):

00100101 0001110**1** 1110100**0**
00100110 00101111 1101011**0**
01101100 00100010 1100101**0**

So if the number was embedded into the first 9 bytes of the grid only the four underlined bits needed to be changed according to the embedded message. Taking the average, only half of the bits in an image will need to be modified to hide a secret message using maximum cover size[4].

The advantages are [5]:

1. Chances for degradation of the original image are less.
2. More data/information can be stored in an image.

The disadvantages are [5]:

1. It is less robust as the hidden data can be lost with image manipulation.
2. The hidden data can be destroyed easily by simple attacks.

B. Difference Expansion

It is an outstanding reversible data hiding scheme in terms of low distortion in image quality and high embedding capacity. The method divides the image into pairs of pixels, then embeds one-bit of information into the difference of the pixels of each pair from those pairs that are not expected to cause an overflow or underflow. A pair generally consists of two neighboring pixels or two with a small difference value. The location map indicates that the modified pairs are compressed and included in the payload[7].

Jun Tian [6] calculate the differences of neighboring pixel values, and from that values some difference values are selected for the difference expansion (DE). The message authentication code, original content restoration information and additional data will all be embedded into the difference values.

Data embedding DE algorithm consists of following steps[6][8]

- Calculate the differences of neighboring pixel values,
- determine the changeable bits in that differences,
- some differences are chosen to be expandable by 1-bit so that increases the changeable bits,
- creating a location map which contains the location information of all selected expandable difference values,
- collecting original LSB values,
- Data embedding by replacement i.e. embed the location map, the original LSBs , and a payload and finally an inverse integer transform.

Data extraction DE algorithm consist of following steps [6][8]

- Calculate the differences of neighboring pixel values,
- determine the changeable bits in that differences,
- collect least significant bits of difference values,
- separate the compressed original changeable bit-stream , the hash of original image (payload) from extracted bit-stream and decode the location map,
- decompress the compressed separated bit-streams and reconstruct the original image replacing the changeable bits,
- Apply the inverse integer transform to reconstruct a restored image. To authenticate the content, we compare the authentication hash with the hash of the restored image.

The advantages are [8]:

1. No loss of data due to compression and decompression.
2. Applicable to audio and video data.
3. Encryption of compressed location map and changeable bit-stream of different numbers increases the security.

The disadvantages include [8]:

1. There may be some round off errors as division by 2 is though very little.
2. Due to bit-replacements of gray scale pixels there is significant degradation of visual quality .

C. Histogram Modification

The histogram modification technique involves generating histogram and finding the peak point and the zero point and shifting histogram bins to embed message bits[9].

Ni et al. [12] proposed a reversible data hiding method , their method uses the histogram of an original image to embed secret messages. In the histogram, they found multiple pairs of peak and zero points, where a peak point corresponds to the pixel value with a maximum number of pixels in the cover image assume and a zero point corresponds to the pixel value with no pixel in the cover image assumes. It uses a pair of peak and zero points to embed the secret messages[11][13][14].

Rajkumar et al. [10] proposes the differences between adjacent pixels instead of simple pixel value is considered, since image neighbor pixels are strongly correlated the difference is expected to be very close to zero, at the sending side, the image is scanned in an inverse s-order as shown in figure 2 and then calculate the pixel difference d_i between pixels x_{i-1} and x_i and peak points of histogram are determined

$$d_i = \begin{cases} x_i & \text{if i} = 0 \\ |x_{i-1} - x_i| & ,otherwise \end{cases}$$

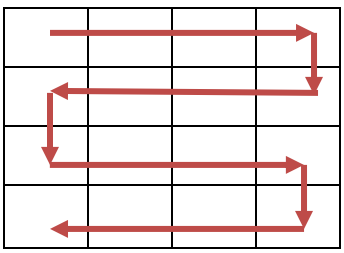

Figure 2. Inverse s ordering scanning

P.H.Pawar et al. [2] uses histogram based RDH method. In this approach the cover image is divided into several equal blocks/tiles and then the histogram is generated for each of these blocks. Maximum and minimum points are computed for these histograms so that the embedding space can be generated to hide the data at the same time increasing the embedding capacity of the image. A one bit change is used to record the change of the minimum points. This improve the level of hiding places. This technique of block division successfully enhances the data hiding capacity because the total data that can be hidden in multiple blocks is generally larger than that can be hidden in a single cover image[15].

The advantages are[8][2]:

1. High payload
2. Distortions are quite invisible

The disadvantages are[8] :

1. Capacity is limited by the frequency of peak pixel value in the histogram
2. It searches the image several times, that makes the algorithm time consuming.

IV. CONCLUSION

Reversible data hiding methods are getting popular because of the reversibility of carrier medium in the receiving end after extraction of secret data. The motivation of reversible data embedding is distortion-free data embedding .In this paper different methods of reversible data hiding for digital images are presented: Least significant Bit substitution (LSB), Difference expansion , Histogram modification each with its advantages and disadvantages. The focus of all methods is on high payload with less degradation of data. The performance can be evaluated by determining the visual quality of the image and by determining the complexity of an algorithm.

ACKNOWLEDGMENT

The author would like to thank the RIMT Institutes, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India. Author would also wish to thank editors and reviewers for their valuable suggestions and constructive comments that help in bringing out the useful information and improve the content of paper.

REFERENCES

[1] D.R.Denslin Brabin, and Dr.J.Jebamalar Tamilselvi "Reversible Data Hiding: A Survey" IJIRCCE: International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1, Issue 3, May 2013.

[2] P.H.Pawar and K.C.Jondhale "Histogram Based Reversible Data Hiding Using Block Division" International Conference on Advanced Communication Control and Computing Technologies ICACCCT,2012,pp-295-299.

[3] Rhythm Katira, Prof. V. Thanikaiselvan "Random Traversing Based Reversible Data Hiding Technique Using PE and LSB" IJET: International Journal of Engineering and Technology, Vol 5, No 2, Apr-May 2013, pp-1579-1583.

[4] Shilpa Gupta, Geeta Gujral and Neha Aggarwal "Enhanced Least Significant Bit Algorithm For Image Stegnography" IJCEM International Journal of Computational Engineering and Management, Vol. 15 Issue 4, July 2012, pp-40-42.

[5] Mehdi Hussain and Mureed Hussain " A Survey of Image Steganography Techniques" International Journal of Advanced Science and Technology Vol. 54, May, 2013, pp-113-124.

[6] Jun Tian. "Reversible data embedding using a difference expansion." IEEE Transaction on Circuits and Systems for Video Technology, Vol. 13. No. 8. August 2003, pp. 890-896.

[7] Hyoung Joong Kim, Vasiliy Sachnev, and Dong Hoi Kim "New reversible data hiding algorithm based on difference expansion method" Feburary 2005, pp-112-119.

[8] Masoud Nosrati, Ronak Karimi, Mehdi Hariri "Reversible Data Hiding: Principles, Techniques, and Recent Studies" World Applied Programming, Vol (2), Issue (5), May 2012. Pp-349-353.

[9] Ya-Fen Chang and Wei-Liang Tai "Histogram-based Reversible Data Hiding Based on Pixel Differences with Prediction and Sorting" KSII Transactions on Internet and Information Systems Vol. 6, No. 12, Dec 2012.

[10] Rajkumar Ramaswamy and Vasuki Arumugam "Lossless Data hiding Based on Histogram Modification " The International Arab Journal of Information Technology, Vol 9, No. 5, Sept 2012, pp-445-451.

[11] Der-Chyuan Lou, Chen-Hao Hu and Chung-Ching Chiu "Steganalysis of Histogram Modification Reversible Data Hiding Scheme by Histogram Feature Coding" International Journal of Innovative computing Information and Control, Vol 7, No. 11, Nov 2011.

[12] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," IEEE Transactions on Circuits Systems and Video Technology, Vol. 16, No.3, 2006, pp. 354–362.

[13] Che-Wei Lee and Wen-Hsiang Tsai "A Lossless Large-volume Data Hiding Method Based on Histogram Shifting Using an Optimal Hierarchical Block Division Scheme**"** Journal of Information Science and Engineering.

[14] C.-H. Yang and M.-H. Tsai "Improving histogram-based reversible data hiding by interleaving predictions" IET Image Processing , Dec 2009.

[15] Harshavardhan Kayarkar and Sugata Sanyal "A Survey on Various Data Hiding Techniques and their Comparative Analysis".

[16] Sona Ignacious "Literature Survey on Performance of Reversible Data Hiding Algorithm" International Journal of Scientific Research Engineering & Technology (IJSRET), Vol 2, Issue 10,pp-665-670, January 2014.

