IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 43-48
www.iosrjournals.org

Reversible Data Hiding Technique and its Type, a survey

[1]Manisha G. Gedam, [2]Shruti M. Rakhunde, [3]Usha P. Kosarkar
[1](manisha.gedam@raisoni.net, Assit. Prof., G. H. Raisoni Institute of Information Technology, RTM Nagpur University, Nagpur (MS), India)
[2](shruti.rakhunde@raisoni.net, Assit. Prof., G. H. Raisoni Institute of Information Technology, RTM Nagpur University, Nagpur (MS), India)
[3](usha.kosarkar@raisoni.net, Assit. Prof., G. H. Raisoni Institute of Information Technology, RTM Nagpur University, Nagpur (MS), India)

***ABSTRACT**: This is a survey paper describes different types of algorithms for Reversible Data Hiding (RDH). Reversible data hiding can be defined as an approach where the data is hidden in the host media such as image, audio and video files. Reversible Data Hiding (RDH) or lossless data hiding, is a method by which the original cover can be lossless restored after the embedded information is extracted. Many RDH techniques have been developed. This paper summarizes and reviews these techniques. Previous literature has shown that difference expansion, interpolation technique, prediction and sorting, histogram modification are the most common methods for data hiding, but previously these methods are implemented in plain images. Recently these methods are used in encrypted images to improve security. Different RDH algorithms have their own merits and no single approach is optimal and applicable to all cases. RDH is still an active topic. This paper provide comprehensive analysis of all the major reversible data hiding approaches implemented as found in the literature. Also paper presents a new method RDH.*

***Keywords -** Reversible Data Hiding (RDH), Difference Expansion (DE), sorting and prediction, histogram modification*

I. INTRODUCTION

Data hiding is the art and science of communicating secret data in an appropriate multimedia carrier, e.g., image, audio, and video files. Digital Steganography and watermarking are the two kinds of data hiding. Reversible data hiding can be defined as an approach where the data is hidden in the host media that may be a cover image. A reversible data hiding is an algorithm, which can recover the original image losslessly after the data have been extracted. Reversible data embedding, which is also called lossless data embedding, embeds invisible data (which is called a payload) into a digital image in a reversible fashion. As a basic requirement, the quality degradation on the image after data embedding should be low. An intriguing feature of reversible data embedding is the reversibility, that is, one can remove the embedded data to restore the original image.

Data hiding is a technique for embedding information into covers such as audio, image and video files, which can be used for copyright protection, media notation, integrity authentication, covert communication, etc. Most data hiding methods embed messages into the cover media like image or video to generate the marked media by only modifying the least significant part of the cover and, thus, ensure perceptual transparency. The embedding process will usually introduce permanent distortion to the cover, that is, the original cover can never be reconstructed from the marked cover. However, in some applications, such as medical imagery, military, and law forensics, no degradation of the original cover is allowed. We need a special kind of data hiding method, for such cases which is referred to as reversible data hiding (RDH) or lossless data hiding, by which the original cover can be lossless restored after the embedded message is extracted. The block diagram of RDH is shown in figure 1.1. Reversible Steganography or watermarking can restore the original carrier without any distortion or with ignorable distortion after the extraction of hidden data. So reversible data hiding is now getting popular. Fig 1.1 Reversible data hiding as a basic requirement, the quality degradation on the image after data embedding should be low. An intriguing feature of reversible data embedding is the reversibility, that is, one can remove the embedded data to restore the original image. From the information hiding point of view, reversible data embedding hides some information in a digital image in such a way that an authorized party could decode the hidden information and also restore the image to its original, pristine state. An information-hiding system is characterized using four different aspects: capacity, security, perceptibility and robustness [2] shown in Fig. 1

♣**Capacity** refers to the amount of information that can be hidden in the cover medium.
♣ **Security** refers the inability of the hacker to extract hidden information.
♣**Perceptibility** means the inability to detect the hidden information.
♣**Robustness** is the amount of modification the stego-medium can withstand before an adversary can destroy the hidden information.

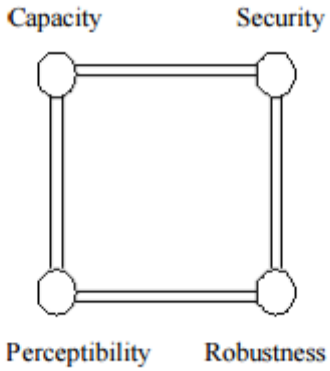


Fig. 1. Characteristics of Data Hiding System

II. REVERSIBLE DATA HIDING

Reversible data hiding can be defined as an approach where the data is hidden in the cover media that may be an image. A reversible data hiding is an approach, which can recover the original image losslessly after the data have been extracted from the cover image. Reversible data embedding, which can be called lossless data embedding, embeds confidential data (which is called a payload) into a digital image in a reversible manner. As a basic requirement, the quality degradation on the cover image after data embedding should be low. An interesting feature of reversible data embedding is its reversibility, that is, one can remove the embedded data to restore the original image.

The data embedding process will usually introduce permanent loss to the cover medium. However in some applications such as military, medical, and law forensics where degradation of cover is not allowed. In these cases, a special kind of data hiding method called reversible data hiding or lossless data hiding is used. Reversible Data Hiding (RDH) in digital images is a technique that embeds data in digital images by altering the pixel values of image for secret communication and the cover image can be recovered to its original form after the extraction of the secret data from it. The block diagram of RDH is shown in Fig.2. Watermarking & Reversible Steganography can restore the original carrier without any or with ignorable distortion after the extraction of hidden data. Thus reversible data hiding method are now getting popular. In this paper some important reversible data hiding techniques for digital images are explained and the results are analyzed.

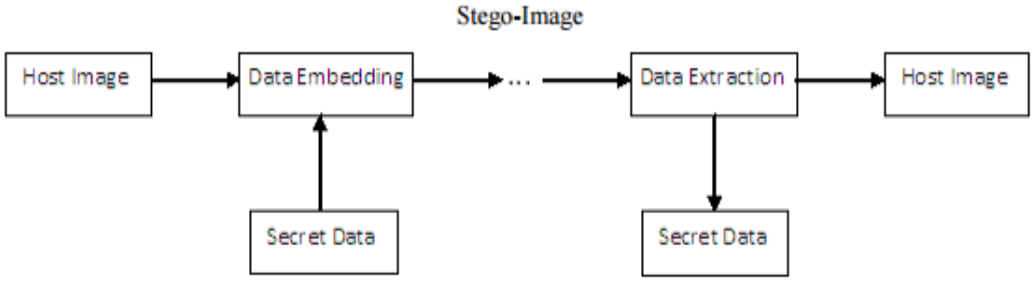


Fig.2. Reversible Data Hiding

III. REVERSIBLE DATA HIDING TECHNIQUES

Reversible data hiding has been there from several years. Researchers have come up with various different methods for reversible data hiding. Following are different techniques which have been proposed over years. Circular visual cryptography was introduces in 2005 [4]. In this scheme, circular shadow image can hide two or more confidential data sets into circular images and display them at both the inner and outer region of the circular images. However, it can only produce a circular shadow image without the central part causing a low

resolution on the images at the inner portion as seen in figure 4. It encrypts data into two ringed shadow images allowing hiding two confidential data sets simultaneously [9].

A. LSB Modification Based Technique.

One of the earliest methods is the LSB (Least Significant Bit) modification. In this well known method, the LSB of each signal sample is replaced (over written) by a secret data bit. During extraction, these bits are read in the same scanning order, and secret data is reconstructed.

B. Difference Expansion (DE) Based Technique.

Difference expansion (DE) based technique for Reversible Data Hiding is proposed by Tian [2]. In DE technique extra storage space is discovered by exploring the redundancy in the image content. The DE technique is used to reversibly embed a payload into digital images. Both the payload capacity limit and the visual quality of embedded images of the DE method are among the best in the literature, along with a low computational complexity. The method of embedding is as follows. The two neighbor pixels *(x, y)* are considered the mean value and the difference is calculated first.

$$l = \lfloor (x + y)/2 \rfloor, \; h = x - y \text{----------(1)}$$

Where $\lfloor . \rfloor$ represents the floor operation which rounds the elements to the nearest integers towards minus infinity. To embed a binary data bit $b\big(b \in (0,1)\big)$ into a difference, the expanded difference is calculated as:

$$h' = 2 \times h + b \text{----------(2)}$$

Finally, the new pixels (x', y') are computed as follows

$$x' = l + \lfloor (h' + 1)/2 \rfloor, y' = l - \lfloor h'/2 \rfloor \text{----------(3)}$$

In extraction phase, the average and the difference of the pixels $(x', y')$are also calculated first:

$$l = \lfloor (x' + y')/2 \rfloor, h' = x' - y' \text{---------(4)}$$

The embedded data is least significant bit ofh', and the original difference h is calculated by:

$$b = LSB(h'), \; h = \lfloor h'/2 \rfloor \text{----------(5)}$$

And the original pixels can be restored by:

$$x = l' + \left\lfloor \frac{(h+1)}{2} \right\rfloor, y = l' - \lfloor h/2 \rfloor \text{----------(6)}$$

In [2] Jui Tian has introduced a difference expansion technique which discovers extra storage space by exploring the redundancy in the image content. Both the secret data holding capacity limit and the visual quality of embedded images of the DE method are among the best in the literature, along with a low computational complexity. Alattar [13] used the vectors instead of pairs to increase the hiding ability. From equation (6) to prevent the overflow and underflow problems i.e. to restrict *x, y* in the range of [0,255], it is equivalent to have $0 \le l + \left\lfloor \frac{h+1}{2} \right\rfloor \le 255, \text{and } 0 \le l - \left\lfloor \frac{h}{2} \right\rfloor \le 255$ since both l and h are integers, on can derive that the above inequalities are equivalent to $|h| \le 2(255 - \le l), \text{and} |h| \le 2l + 1$

To extract the embedded data and restore the original values, the decoder needs to know which difference values have been selected for the DE. To facilitate it, we need to embed such location information, such that the decoder could access and employ it for decoding. For this purpose, we will create and embed a location map, which contains the location information of all selected expandable difference values.

From the detail procedure, we can get the conclusion that, DE-based algorithm didn't take the content into account; it brings big distortion to images when the texture of the image is complex. So, many new algorithms based on DE are proposed. In scheme proposed in [5] authors have divided the image logically into smooth and complex region thus used only the smooth region for applying RDH using DE this improves the performance of RDH algorithm and also the distortion caused to the complex region of image will be less as compared to other schemes. Using these methods the capacity of hiding data is also increased to a level. Thus visual quality of encoded image is improved in scheme. Similar approach is in [3] where there is Hsiao's algorithm, the image is divided into blocks, and the variances of each block are computed to classify the block into different class with different texture, then different amount data are embedded into different block, the quality has improved and the payload has increased.

C. Histogram Shift Based Technique.

The histogram-shifting based reversible data hiding schemes embed data by shifting the histogram into a fix direction. And there are two points which are important in these schemes, which are peak point and zero point. The peak point corresponds to the grayscale value, which corresponds to the maximum number of pixels in the histogram of the giver image. And the zero point is usually the point that the number is histogram is zero. And the minimum number of pixels is selected as the zero point to increase the embedded capacity. In histogram-shifting based algorithms, the pixels between the peak and zero pairs were modified in the embedding processing, the pixel in the peak point was used to carry a bit of the secret message, and the others were

modified and no secret data were embedded The hiding capacity of the histogram shifting based data hiding equals the number of pixels in the peak points, the larger the number of pixels in the peak point, the higher the hiding capacity. To increase the hiding capacity, more of the peak points and zero pairs can be used. Sometimes, it is difficult to find out more pairs of peak and zero points because the zero points are not searched. The basic procedure of histogram–shift algorithm is as follows:

Step 1: Create the histogram of image.

Step 2: Find the peak points and the zero points.

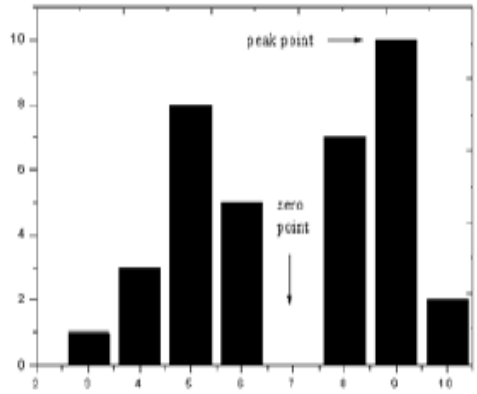
Step 3: We assume the peak point is *a* and the zero point is *b*. *(a>b);* Shift the points between *b+1* and a-1 by reducing 1.

Step 4: If the embedded bit is 1, the peak point is reserved; otherwise, change the peak point value by reducing 1.

Step 5: To achieve the reversibility requirements, the location of the pixels in the minimum point must be recorded and embedded. Then record the peak point, the zero points and some other auxiliary information.

Figure 4 gives an example to demonstrate the basic procedure of histogram-shifting based reversible data hiding algorithm. We assume the cover image is Figure 3-(a). Figure 3-(b) is the histogram of the cover image, we can see that the peak point is 9 and the zero point is 7, and the number of 9 is 10. That means we can embed 10 bits into host image. We assume the embedded data is 2 (1101110001). Figure 4-(a) and 4-(b) are stego-image and recovered image, respectively.

9	8	8	8	9	5
9	5	9	9	6	5
8	6	9	8	5	3
6	5	10	9	6	4
5	5	10	9	9	5
8	4	9	6	8	4



(a) Cover image (b) Histogram of image

Figure 3: Demonstration of basic procedure of Histogram-shifting based RDH

9	7	7	7	9	5
8	5	9	9	6	5
7	6	9	7	5	3
6	5	10	8	6	4
5	5	10	8	8	5
7	4	9	6	7	4

9	8	8	8	9	5
9	5	9	9	6	5
8	6	9	8	5	3
6	5	10	9	6	4
5	5	10	9	9	5
8	4	9	6	8	4

(a) Stegno Image (b) Recovered Image

Figure: 4. Example of the histogram-based algorithm

D. Prediction Error Based Technique

In order to enhance the embedding capacity, Yen proposed an efficient data hiding scheme based on predict error method. Reversible data hidings based on prediction error use predicted system to embed data, there are many predictors which have been proposed. They are horizontal predictor, vertical predictor, Causal weighted average, Causal and SVF. One well known predictor is the median edge detection (MED) predictor. His embedding procedure as following:

Step 1. Divide the cover image by using 3 * 3 blocks.
Step 2. Take the block's center point as the base point and obtain the prediction error value between it and the surrounding pixel.
Step 3. Do they predict error value histogram and find out the peak point.
Step 4. Embed the secret data into two-side region of peak point.
Step 5. Get the stego-image.
Recover the secret data by using the following steps:
Step 1. Divide the stego-image by using 3 * 3 blocks.
Step 2. Take the block's center point as the base point and obtain the prediction error value between it and the surrounding pixel.
Step 3. Find the peak point from the predict error value histogram and recover the secret data.
Step 4. Restore error histogram and recover the original image.
The major characteristic of prediction error method is the little difference between neighboring pixels. Therefore, it will be enhanced the embedding capacity and keep better image quality because the statistic data concentrates on zero point.

E. Vector quantization Based Technique

Different from the above three scheme, there is a scheme in compression domain, which named reversible data hiding based on Vector quantization (VQ). VQ is one of efficient compression technique, and it has widely used for its good character of easy implementation and high efficiency. In Use the upper and left blocks to predict the current block to embed data. During the encoding phase, adjacent blocks are used to encode the current block, but the additional flag bits are required.

Vector Quantization is a method which is lossy compression. For fewer stores in images, videos and transports, obtains the lower data rate and rebuilds the signal that has some loss. Vector Quantization is proposed first time by Y. Linde, A. Buzo, and M. Gray in 1980. This method produces codebook that combining with each representative vectors which call code word symbolically by data training. The size and domain of codebook decide the rate that compress. The generating, optimization, encoding, and decoding are included in codebook of VQ encoding. First, select original codebook that results in the problem of local optimization after optimization is important in generating. Next, Linde-Buzo-Gray (LBG) was used to maximize codebook often. Last, divide image to equal size in non-overlapping block. The size of each blocks are the same with coding dimension and regard as vectors.

The codebook is combined with code words that are a set of representative sample. The flow chart of coding book is show as Figure.5. Figure.6 is the flow chart of decoding. When decoding, it only uses the result which is produced in coding to search the sets of index. Then find the corresponding code words that according to original sort in codebook and recover image data. After all of indexes process, the decoding is finished.

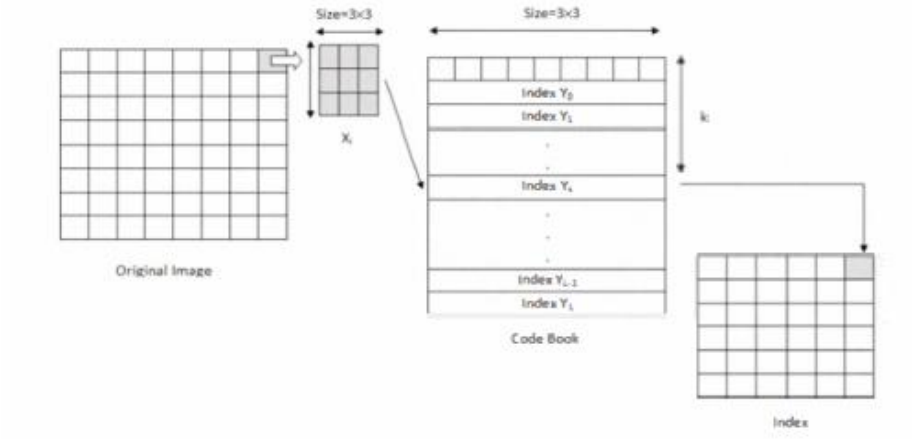


Figure: 5. The flow chart of coding (Code words 3 x3, codebook Size=L)

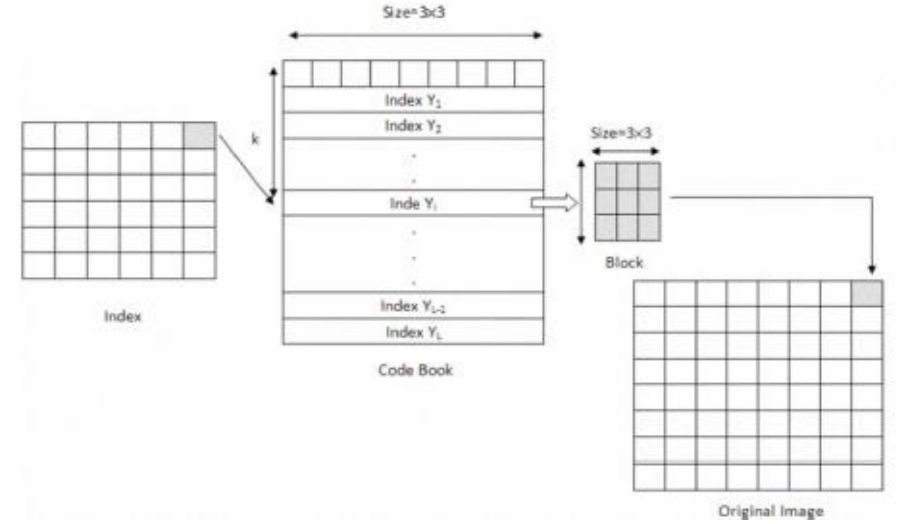


Figure: 6.The flow of decoding (Code words 3 X3, codebook Size=L)

IV. CONCLUSION

Reversible data hiding techniques getting popular because of the reversibility of carrier medium in the receiving end after extraction of secret data. In this paper five different types of reversible data hiding techniques for digital images: LSB Modification Based Technique**,** Difference expansion technique, Histogram modification technique and Interpolation technique are studied, analyzed and compared. The survey results show each technique has its own advantage and disadvantages.

REFERENCES

[1]. Abbas Cheddad, Joan Condell, KevinCurran and Paul McKevitt, (2010), "Digital Image Steganography: Survey and Analysis of Current Methods", Elsevier signal processing, Vol. 90, pp. 727-752.

[2]. Jun Tian, "Reversible Data Embedding Using a Difference Expansion", IEEE transactions on circuits and systems for video technology, vol. 13, no. 8, august 2003.

[3]. Ali Al-Ataby and Fawzi Al-Naima,(2010), " A Modified High Capacity Image Steganography Technique Based on Wavelet Transform", The International Arab Journal of Information Technology, Vol. 7, No. 4, pp 357-363.

[4]. Bret Dunba r, (2002), "A detailed look at stegonagraphic Techniques and their use in an Open – Systems Environment", SANS Institute.

[5]. Shruti M. Rakhunde, Archana A. Nikose, "New Approach for Reversible Data Hiding Using Visual Cryptography", IEEE International Conference on Computational Intelligence and Communication Networks 2014, Print ISBN: 978-1-4799-6928-9, Pages 846 - 855

[6]. Chang C.C, Lin C.C and Chen Y.H,(2008), " Reversible Data Embedding Scheme using Differences Between Original and Predicted Pixel values" , IET Information Security, Vol. 2,

[7]. Chin-Chen Chang, Wei-Liang Tai, and Chia-Chen Lin, (2006), "A Reversible Data Hiding Scheme Based on Side Match Vector Quantization", IEEE Transaction on Circuits and Systems for Video Technology, Vol.16, No. 10, pp 1301-1308.

[8]. Coltuc D and Tremeau A,(2005), " Simple Reversible Watermarking Schemes ", Proc. of SPIE, Security, Steganography, Watermarking of Multimedia Contents, Vol. 5681, pp. 561–568.

[9]. Dalel Bouslimi, Gouenou Coatrieux, Michel Cozic and Christian Roux, (2012), "A Joint Encryption/Watermarking System for Verifying the Reliability of Medical Images", IEEE Transactions on Information Technology in Biomedicine, Vol. 16, No.5, pp. 891- 899.

[10]. Shruti M. Rakhunde, Archana A. Nikose, "A Novel and Improved Technique for Reversible Data Hiding using Visual Cryptography", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 6, June 2014, ISSN (Online) : 2278-1021, ISSN (Print) : 2319-5940

[11]. J. Fridrich, D. Soukal, (2006), "Matrix Embedding for Large Payloads", IEEE Transactions on Information Forensic and Security, Vol. 1, No.3, pp. 390-395

[12]. Jennifer L. Wong, Gang Qu and Miodrag Potkonjak,(2004), "Optimization-Intensive Watermarking Techniques for Decision Problems" , IEEE Transaction on Computer- Aided Design of Integrated Circuits and Systems, Vol. 23, No. 1, pp. 119-127.

[13]. Lixin Luo, Shenyang Chen, Ming Chen, Xiao Zeng and Zhang Xiong, (2010), "Reversible Image Watermarking Using Interpolation Technique", IEEE Transaction on Information Forensics and Security, Vol. 5, No. 1, pp 187 – 193.

[14]. A. Nag, S. Biswas, D. Sarkar, P.P. Sarkar ,(2010), "A Novel Technique for Image Steganography Based on Block-DCT and Huffman Encoding", International Journal of Computer Science and Information Technology, Vol. 2, N0. 3, pp. 103-112.

[15]. Shruti M. Rakhunde, Archana A. Nikose, "Reversible Data Hiding using Color Visual Cryptography", International Journal of Advance Foundation and Research in Computer (IJAFRC), Volume 1, Issue 2, Feb 2014. ISSN 2348 – 485

[16]. Neminath Hubballi and Kanya kumari D P,(2009), "Novel DCT based watermarking scheme for digital images", International Journal of Recent Trends in Engineering, Vol. 1, No. 1, pp. 430-43..

[17]. Niels Provos and Peter Honeyman, (2003), "Hide and Seek: An Introduction to Steganography", IEEE Security & Privacy, pp.32-44.

[18]. Shaowei Weng, Yao Zhao, Jeng-Shyang Pan and Rongrong Ni,(2008), "Reversible Watermarking Based on Invariability and Adjustment on Pixel Pairs", IEEE signal processing letters, Vol. 15, pp. 721-724