ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

A Survey on Lossless and Reversible Data Hiding

Uma Choudhari[1], Mrunal Jadhav[2], Priyanka Kadam[3], Vaishanavi Mane[4]

Student, Dept.of CS, Pad.Dr.D.Y.Patil College of Engineering and Technology, Pimpri, Savitribai Phule Pune University,Pune,India[1,2,3,4]

ABSTRACT: Data hiding is a way by which we can hide the important or secret data into some other data, other data can be text, image etc. It is means for providing privacy protection. This process of hiding data into another form of data is known as steganography. During data hiding data can be lossed because when we encrypt and embed the data into image pixel values may get changed when we try to recover the original image some distortions may occur so lossless technique can be used. In lossless data loss can be avoided. Using reversible data hiding original image can be recovered with minimum distortion. Reversible data hiding consists of three phases image encryption, data embedding and data-extraction image-recovery phases. This technique is widely used in medical field , military field and law forensics. In such fields distortion is not allowed.

In this paper we done survey about Reversible Data Hiding Techniques using difference expansion, using optimal value transfer. Reversible Data Hiding Technique used in Media Encryption and Media Watermarking , in Encrypted JPEG Bitstream.

KEYWORDS: Encryption, Data hiding, Lossless, Reversible.

I. INTRODUCTION

Data hiding and lossless:

Data hiding is method to hide or to embed data in digital image to provide protection to secrete data, providing authentication. There are various techniques for data hiding. First techniques included invisible ink, secret writing using chemicals, templates overcomes text messages, microdots, changing letter/word/line/paragraph spacing, changing fonts. PDF files and HTML can be used for data hiding but it can be used to a limited extend. Data hiding can be used for various purposes such as for authentication of image, copyright, data integrity, fraud detection, ownership of images.

Lossless technique is used to avoid data loss due to encryption and embedding of the data into image pixel values may get changed. It is based on image blocking. By using blocking good performance can be achieved, but always not possible to restore full content.

II. RELATED WORK

1)Least Significant Bit (LSB) modification[1]:

a) GENERALIZED-LSB (G-LSB) EMBEDDING.: LSB modification method is used for data-embedding. In this, the Least Significant Bit of each signal sample is over written by a payload data bit. This data bit is embedded by one bit of data per input sample. two or more LSBs may be over written allowing for a corresponding bits per sample, if additional capacity is required . While extraction process, these bits should be read in the same scanning order, and payload data should be reconstructed. LSB modification is a simple, no robust embedding technique by using small bounded embedding distortion and high embedding capacity. G-LSB method is employed here. G-LSB is generalized method of LSB. If the host signal is represented by a vector , the G-LSB embedding and extraction processes can be represented as

$$S_w = Q_L(S) + w$$
$$W = S_w - Q_L(S_w) = S_w - Q_L(S)$$

DOI: 10.15680/IJIRCCE.2016. 0402230

Where, S_w represents the signal containing the embedded information, w represents the embedded payload vector of L–ary symbols, i.e., w_i € {0,1,….,L-1} , and

$$Q_L(x) = L\left\lfloor\frac{x}{L}\right\rfloor$$

A. Binary to L-ary (L-ary to Binary) Conversion.

B. Embedding Capacity and Distortion

b) LOSSLESS GENERALIZED-LSB DATA EMBEDDING

The G-LSB embedding algorithm can be directly used for data embedding with low distortion. But the method is irreversible, i.e., When its lowest levels containing the residual signal are replaced with the watermark signal the host signal is permanently distorted. This drawback can be overcome by including information for construction of the residual signal again along with the embedded data in the payload.

2) Reversible data hiding[2]-

There are two types of reversible data hiding:

a) Non separable reversible data hiding- A owner of data encrypts original image which is uncompressed for this encryption key is used. Then data hider compress the LSB of encrypted image for this data hiding key is used. After this process sparse space is created to which additional data can be accommodated.

owner of data use encryption key and encrypt the original image then data hider use data hiding key to embed additional data into encrypted image, it is not necessary that data hider know the original content .in this technique data extraction is not separable means if data hiding key is available but encryption key is not available then it is not possible to extract any information from the encrypted image containing additional data.

b)Separable reversible data hiding- In this scheme if only data hiding key is available then additional data can be extracted even if receiver does not know the original content . if only encryption key is available then receiver can decrypt the received data to obtain original image but additional data cannot be extracted .this scheme has three phases:-

1. Image encryption
2. Data embedding
3. Data extraction /image recovery phases

For this technique in future comprehensive combination of image encryption and data hiding which is compatible with lossy compression can be used. If lossless compression method is used for encrypted image which contain embedded data and still additional data can be extracted and original content can be recovered because in lossless compression content of the image which contain embedded data does not change.

3) Reversible data embedding using difference expansion[3] :

In this technique payload is embedded into digital image, this supports reversibility which means original image can be restored by removing the embedded data. In difference expansion differences of neighboring pixel values are calculated and then some difference values are selected for difference expansion. In reversible data embedding, one approach is to select embedding area which can be list significant bit of some pixels. In this area payload and original values can be embedded. Difference expansion discovers extra storage space by exploring the redundancy in the image.

a)Reversible Integer Transform:

In Reversible Integer Transform for an 8bit grayscale, pair of value(x, y),x, y € Z,0 ≤ x, y ≤ 255,then define integer average I and difference h as

$$l := \left\lfloor\frac{x+y}{2}\right\rfloor, h := x - y$$

The reversible integer transforms set up a one-to-one correspondence between (x, y) and (I, h),

b) Expandable and Changeable Difference Values

DOI: 10.15680/IJIRCCE.2016. 0402230

Embed a bit into the difference value by the DE, the new, expanded difference value will be $h^{'}$

$$h^{'}=2*h+b.$$

prevent overflow and underflow, should be $h^{'}$satisfy

$$|h| \leq \min(2(255-I),2I+1).$$

c)Data-Embedding Algorithm

Data embedding algorithm consists of six steps:

1. First the difference values are calculated.
2. Then difference values are Partitioned into four sets
3. Creating a location map
4. Collecting original LSB values
5. Data embedding by replacement
6. Inverse integer transform

The performance of data embedding algorithm can be measured by using following factors:

1. Payload capacity limit - The maximum amount of information that can be embedded is Payload capacity limit .
2. Visual quality - This factors checks how is the visual quality.
3. Complexity – Complexity is the algorithm complexity.

Advantage of this technique is that redundancy in the digital content is explored and because of this reversibility is achieved.

4) Optimal value transfer[4]:

In reversible data hiding values of sender's data is modified by using some rules and then at receiver end original content can be restored perfectly after hidden data is extracted. Optimal value transfer differences between original pixel values are found from neighbors . and optimal value transfer rule is used and estimation errors are modified .

Under a payload distortion criterion the optimal rule of value modification is used. Target function is maximized for that iterative algorithm is used and optimal value transfer matrix can be obtained. First denote the histogram of the available data as:-

$$H = \{\ldots,h-2,h-1,h0,h1,h2,\ldots\}$$

Where, h_k is the amount of available data with a value k.

Then denote the number of available data possessing an original value *i* and a new value *j* caused by data hiding as $t_{i,j}$, and transfer matrix is made up of $t_{i,j.}$

$$T=\begin{bmatrix} t_{M1,M1} & \cdots & t_{M1,M2} \\ \vdots & \ddots & \vdots \\ t_{M2,M1} & \cdots & t_{M2,M1} \end{bmatrix}$$

Where M_1 and M_2 are the minimum and maximum of the available cover data.

Steps to calculate the optimal transfer matrix:-

1. Initialize

   And D = 0,P = 0.
2. Calculate the new histogram H'
3. For each $t_{i,j} > 0$ and each $k(k \neq j, h'_i > h'_k)$

$$\leftthreetimes (i,j,k) = \frac{\log\left(\frac{h'_j}{h'_k}\right)}{d(i,k)-d(i,j)}$$

4. Calculate the distortion level D. Then goto step2

5) Media Encryption and Media Watermarking[5]:

a) Media Encryption:

In Media encryption media data is encrypted into unintelligible ones with ciphers, which protects confidentiality of media contents. The encrypted videos are difficult to understand.

In order to meet real time applications video encryption often requires the scheme be time efficient and format complaint Different from text or binary data encryption . It is not possible to practically encrypt video data completely with traditional ciphers, such as standard used for data encryption or standard used for advanced encryption , because of high computational cost. Alternatively,partial encryption encrypts only a fraction of video data and improves the efficiency.

b) Media Watermarking:

Media Watermarking embeds some information into media data perceptibly or imperceptibly, and it protects media data's identification. For invisible video watermarking imperceptibility and robustness are often required. In robustness algorithms against H.264/AVC compression which has three types: raw video watermarking, compression domain watermarking and compressed data watermarking.

Raw video watermarking embeds watermarks into videos before video compression. Compression domain watermarking embeds watermarks into DCT coefficients during H.264/AVC encoding. Compressed data watermarking embeds watermarks into the compressed data stream.

Process of Media Encryption and Media Watermarking:

media data X is encrypted and watermarked partially. Set X be composed of two independent parameters,i.e. Y, and Z . Among them, Y is encrypted, and Z is watermarked.

$$\begin{Bmatrix} Y' = E(Y, K_e) \\ Z' = W(Z, B, K_w) \end{Bmatrix}$$

where,$Y^{'}$,K_e ,E() ,$Z^{'}$, B, K_w,K() are the encrypted copy of Y, encryption key, encryption algorithm, watermarked copy of Z, watermark, watermark key and watermark algorithm, respectively.

Encryption Process:

It is necessary that both motion information (MVD) and texture information (IPM and residue data)should be encrypted, because it is not secure enough to encrypt only IPM. MVD, IPM, and residue data are all encrypted partially with a cipher to decrease the time cost .

MVD Encryption: For each macroblock, the signs ("0"—positive, "1"—negative) of MVD [x,y] are encrypted with a cipher. That is, MVD is encrypted from [x,y] to [$x^{'}$,$y^{'}$] with the following condition being satisfied:

$$\begin{Bmatrix} |x'| = |x| \\ |y'| = |y| \end{Bmatrix}$$

where "|x|" denotes the absolute value of x.

IPM Encryption: Encoding is done on IPM with Exp-Golomb code. In Exp-Golomb code, each 2M+1-length codeword is composed Of M zeros, "1" and M bits of information.

Residue Encryption: DCT coefficients are encrypted partially for each non-zero residue macroblock. That is, only the first 8 coefficients' signs in each 4 4 DCT block are encrypted with a cipher.

6) Reversible Data Hiding[6]:

In this a reversible data hiding algorithm is used which can recover original image without any distortion.This algorithm uses points of the histogram of an image which can be the zero or minimum and the pixel grayscale values are slightly modified and then embed to data into the image.The algorithm has been successfully applied to a wide range of images. In data hiding process two sets of data, a set of the embedded data and another set of the cover media data are linked. The relationship between these two sets of data characterizes different applications. It is shown experimentally that the peak signal-to-noise ratio (PSNR) of the marked image generated by this method versus the original image is assures to be above 48 dB.

This reversible data hiding technique will be deployed for a wide range of applications like secure medical image data systems, and authentication of image in the medical field and law enforcement, and the other fields in which the rendering of the original images is needed or desired.

Embedding Algorithm:

1) We first find a zero point, and then a peak point in the histogram. A zero point corresponds to the grayscale value. A peak point corresponds to the grayscale value .

2)The whole image is scanned in a sequential order, i,e row-by-row, from top to bottom, or, column-by-column, from left to right.

3)The whole image is scanned once again in the same sequential
order. In this step we check the to-be-embedded data sequence.

If the corresponding to-be-embedded bit in the sequence is binary “1,” the pixel value is increased by 1. Otherwise, the pixel value remains intact.

Extraction Algorithm:

1) In first step Scan the marked image in the same sequential order which is used in the embedding procedure.
2) Scan the image again, for any pixel whose grayscale value x€(a,b), the pixel value is subtracted by 1.
3) If there is overhead found in the extracted data, set the pixel grayscale value (whose coordinate (i,j) is saved in the overhead) as b.

7) Reversible Data Hiding in Encrypted JPEG Bitstream[7]:

In this technique JPEG bitstream is encrypted into properly organized structure and by performing slight modifications in JPEG stream secret message can be embedded into encrypted bitstream . the secret message bits are encoded with error correcting codes so that perfect data extraction and image recovery is achieved. The encryption and embedding is performed by using encryption and embedding keys. If both keys are available the secret bits can be extracted by analyzing the blocking artifacts of the neighboring blocks, to perfectly recover the original bitstream even if only encryption key available it is possible to decode the bitstream .

a)BITSTREAM PARSING AND ENCRYPTION

A JPEG Bitstream Parsing an image is decomposed to a set of quantized DCT coefficients in non-overlapped blocks, and then coded into a bitstream by using entropy encoding. In entropy encoding, the DC coefficients and the AC coefficients are handled separately. The DC coefficients are coded with the Huffman codes after using a one-dimensional predictor. For AC coefficients, since there are many zeros, with the run lengthcoding (RLC) coefficient are encoded and this process is efficient . The quantization tables and Huffman/VLC coding tables are defined and stored in the JPEG file header, which are important for entropy encoding and decoding.The entropy encoded bits are structured by the Huffman codes in which the Huffman code identifies the range of the coefficient magnitude and the length of appended bits. Bitstream parsing is a part of the entropy decoding, which analyzes the compressed bits according to the JPEG structure and the Huffman tables extracted from the JPEG file header.

*b)*Bitstream Encryption

Encrypting a JPEG bitstream into the one that can be decoded into an unrecognizable image directly by a JPEG decoder.JPEG data has strict structure because of this if any alteration is made on individual bit it may cause the failure of decoding. So that the care should be taken. So the encryption can be done according to the encoding structure by selecting and modifying the changeable bits.and this procedure consists of two steps:-

I. Encryption of the appended bits
II. Encryption of the quantization table.

III. CONCLUSION

The We survey all these papers, after doing survey it is observed that in different papers different techniques are used for specific purpose ,such as lossless technique is used to avoid data loss and reversible technique is used in fields like medical, military and law forensics where distortion of image is not allowed. These techniques are separately used , in future these techniques can be combined to improve efficiency and performance.

DOI: 10.15680/IJIRCCE.2016. 0402230

REFERENCES

[1] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, “Lossless Generalized-LSB Data Embedding,” IEEE Trans. on Image Processing, 14(2), pp. 253–266, 2005.

[2] X. Zhang, “Separable Reversible Data Hiding in Encrypted Image,” *IEEE Trans. Information Forensics & Security*, 7(2), pp. 526−532, 2012.

[3] J. Tian, “Reversible Data Embedding Using a Difference Expansion,” IEEE Trans. on Circuits and Systems for Video Technology, 13(8), pp. 890−896, 2003.

[4] X. Zhang, “Reversible Data Hiding with Optimal Value Transfer,” *IEEE Trans. on Multimedia*, 15(2), 316−325, 2013.

[5] S. Lian, Z. Liu, Z. Ren, and H. Wang, “Commutative Encryption and Watermarking in Video Compression,” *IEEE Trans. on Circuits and Systems for Video Technology*, 17(6), pp. 774−778, 2007.

[6] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, “Reversible Data Hiding,” IEEE Trans. on Circuits and Systems for Video Technology, 16(3), pp. 354−362, 2006.

[7] Z. Qian, X. Zhang, and S. Wang, “Reversible Data Hiding in Encrypted JPEG Bitstream,” *IEEE Trans. on Multimedia*, 16(5), pp. 1486−1491, 2014.