International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Reversible Data Hiding on Encrypted Digital Images

Kranti U. Patil[1], J. V. Shinde[2]

[1]Late G. N. Sapkal College of Engineering, Anjaneri, Nasik- 422212

[2]Professor, Late G. N. Sapkal College of Engineering, Anjaneri, Nasik- 42221

Abstract: ***Recently, Reversible data hiding is a point of attention. Reversible data hiding in encrypted images is mostly used since it maintains excellent property that the original cover image can be recovered losslessly after the embedded data is extracted. By using this technique the confidentiality of image contents is also maintained. This paper proposes a novel reversible data hiding technique for digital color image. This technique guarantees effective retrieval of hidden data in color image without degradation in the quality of color image. The cryptographic methods are used to encrypt & decrypt the data to increase the security. The proposed method is enhanced from text data to any data type like text, video, audio etc. In this system only digital image is considered as cover for data hiding purpose.***

Keywords: Reversible data hiding, Steganography, Cryptography, Encryption, Decryption

1. Introduction

In recent years, the technologies are developed so much that most of the media's use internet as the primary medium to transfer the data from one end to other end across the world. Now-a-days, the data transmission is very simple and fast using the internet. However, security threats are one of the main problems with data transition over the internet. On internet the personal data can be stolen or misused in many ways. Therefore, data security is one of the most essential factors taken into consideration during the process of data transferring.

Reversible data hiding (RDH) in images is a recently used technique, by which the original cover can be recovered losslessly after the embedded data is extracted. Here the image used for hiding the data is called cover and the image generated after embedding data is called stego image. This technique is mostly used in military, medical and law forensics, where no distortion of the original cover is allowed.

In current trends, information security domain is a point of interest. To overcome this, a technique of cryptography was developed for keeping the communication secret. Many different methods for encryption and decryption of data have been developed for keeping the data secret. Unfortunately, keeping only the contents secret could not work. There is a requirement of keeping the existence of message secret. The technique developed for this purpose is called steganography. Steganography is an information hiding technique. Using steganography we can embed secret data into digital cover media like images, video, audio, etc. file formats. Steganography has several characteristics like storage capacity, invisibility and resistance against attack. Due to these characteristics it is an efficient data hiding technology.

Data hiding techniques generally falls into two groups i.e. spatial and frequency domain. In the first group, least significant bits (LSBs) of image pixels are used for embedding the data. In the second group, frequency coefficients of image are used for embedding purpose. In RDH, the lossless recovery of cover after the data extraction is an important aspect. The application of cryptography with steganography can increase the capacity of information. In this paper, reversible data hiding on digital cover image is proposed. In this paper data files are embedded after first 8 bits of image. Here multiple data files are hidden in the cover image. There are no limitations on embedded data size. With respect to cover image we can embed larger data files into that. The data used for embedding can be of any type like image, audio, video etc. There is no restriction of data type.

2. Literature Survey

In W. Zhang, B. Chen, and N. Yu [2] method, a decompression algorithm for embedding the data is used. They proved that using this construction they can achieve the rate-distortion bound as long as the compression algorithm reaches the entropy. In this they have improved three RDH schemes that are using binary features sequence as covers. Using this system, embedding distortion can be reduced. It also improves reversible data hiding schemes for binary JPEG images. This system did not work on gray scale covers for designing recursive codes.

J. Fridrich, M. Goljan, and D. Rui's [3] system, a general framework for RDH is proposed. Extracted compressible features of cover images are firstly introduced by them. In this system, they have reserved a space to hide data by compressing the proper bit-planes having minimum redundancy. The lowest bit-planes which offer lossless compression are used if the image is not noisy. In completely noisy image some bit-planes are having strong correlations. These bit-planes are used to vacate room space to store hash. This system provides high capacity and security levels and can be used for authentication purpose of JPEG, audio file, digitized holograms etc. But this system forces noisy images to embed information in the higher bit-planes. Small images having single bit-plane cannot offer enough space to hide

hash. This system has not enough capacity to embed large payload.

J. Tian has proposed a system which uses difference expansion method for embedding data. This system uses the features which are compressed by expansion i.e. the differences between two neighboring pixels. Some differences are selected for expansion by one bit i.e. the difference is multiplied by 2. Thus, LSB's of the differences are all zero and this LSB's can be used for embedding messages. The advantages of the system are: 1. Use of compression and decompression causes no loss of data. 2. This system is also applicable to audio and video data. 3. The compressed location map and changeable bit streams of different numbers are encrypted which increases the security. The disadvantages of the system are: 1. as there is division by 2 there may be some round off errors. 2. Depends largely on the smoothness of natural image that's why can't be applied to images who's capacity is zero or very low. 3. Degradation of visual quality due to bit replacements [4].

Z. Ni, Y. Shi, N. Ansari, and S. Wei have proposed a system which uses histogram shift strategy for RDH. In this system, the space is saved for embedding the data by shifting the bins of histogram of gray values. The authors make use of zero point and a peak point of given image histogram to embed messages. In this system, the embedding capacity is the number of pixels with peak point. For embedding, the whole image is searched for peak point. The advantages of the system are: 1. Simple to implement. 2. Constant PSNR of 48.0dB is obtained. 3. Distortions are quite invisible. 4. Capacity is high. The disadvantages of the system are: 1. Capacity is limited by the frequency of peak-pixel values in histogram. 2. Time consuming as image is searched several times [5].

X. L. Li, B. Yang, and T. Y. Zeng have used a hybrid algorithm which makes the combination of three techniques of PEE (i.e. Prediction-error Expansion), adaptive embedding and pixel selection. In the proposed system, depending on the threshold values the image pixels is divided into two parts. Then the pixels are selected depending on their capacity-parameter and threshold. The smooth pixels are selected from two parts. Finally, data is embedded by modifying the histograms that are derived from selected pixels. The advantages of the system are: 1. By decreasing the modifications to pixel values, the system reduces the embedding impact. 2. More sharply distributed prediction-error histogram can be obtained. 3. The visual quality of watermarked image is greatly improved [6].

L. Luo et al. have used an interpolation technique for developing their reversible image watermarking system. This system can embed a large amount of converted data into images with imperceptible modification. The interpolation errors which are residuals of this technique have greater decorrelation ability. The highly efficient reversible watermarking scheme is developed by applying additive expansion to these interpolation-errors. The advantages of the system are: 1. High image quality. 2. Greater embedding capacity. 3. Less Computational cost [7].

G. Xuan, J. Chen, J. Zhu, Y.Q. Shi, Z. Ni, and W. Su have proposed a integer wavelet transform based lossless data hiding technique. This system hides the authentication information. For preventing gray scale overflowing during data embedding, the histogram modification or integer modulo addition techniques are used. This method uses second-generation wavelet transform IWT. The information is hided into middle bit-plane and in the high frequency sub-bands respectively. This makes the watermarked image greatly as same as the original image. Also the PSNR value is increased. The advantages of the system are: 1. High embedding capacity. 2. Security level is raised due to the use of secret key during embedding of data. The disadvantages of the system are: 1. only gray scale mapping is done. 2. Often multiple bit planes are needed to have enough space [8].

In this paper, V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi have proposed a system which gives reversible or lossless watermarking for image without using a location map. In this system data embedding depends on the prediction errors. Prediction errors based on magnitude of its local variance can be recorded using sorting technique. Using the sorted prediction errors and reduced size location map whenever needed improves the data embedding capacity by decreasing the distortion. The histogram shift significantly reduces the size of location map. The double embedding scheme in this system allows using each pixel for hiding data. The advantages of the system are: 1. Capacity can be significantly increased. 2. Double embedding scheme is used. 3. Less distortion [9].

M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran have tried to first encrypt the data and then compressing it, so that the compressor don't know the encryption key. The encrypted data is compressed using distributed source coding principles. The key will be available only to the decoder. They have shown that encrypted data can be compressed to same rate as that of original unencrypted data could have been. The perfect secrecy and original cover recovery is obtain in this system [10].

In the above methods, the data hiding is done but the size of the data hided is limited. Also, some of the systems have considered gray-scale image as the cover image. So, we are trying to overcome these problems through our system.

3. Existing System

In the existing system, histogram shifting technique is used. The image is firstly divided into two planes A and B using image partition technique. Then the LSBs of A are reversibly embedded into B using standard RDH technique. The estimating error is calculated so that data can be embedded into estimating error sequence using histogram shift technique. By using bidirectional histogram shift, some messages can be embedded on each error sequence. The rearranged image is then encrypted using stream cipher to generate its encrypted version. The information for data hider is embedded into LSBs of first 10 pixels in encrypted image. After image encryption, the data hider or any third person cannot access the content of the original image without using encryption key.

Once the image is passed on to data hider, he can embed data into image without getting access to original image. The data hider can easily read the information in the LSBs of first 10 encrypted pixels. After getting the information about how many bit-planes and rows of pixels he can modify, the data hider hides additional data into available bit-planes using LSB substitution method. At the end of data embedding process, data hider encrypts the additional data using data hiding key to generate marked encrypted image. After that, no one except data hiding key holder can extract the data.

Next, the marked encrypted image is passed on to database manager. The database manager can extract data from either encrypted or decrypted image.

Case 1: Data extraction from Encrypted image
In this case, the data manager can decrypt the LSB planes and extract the data by directly reading the encrypted version using data hiding key. Here both data embedding and data retrieval are done on encrypted image so the privacy of original contents is maintained.

Case 2: Data extraction from Decrypted image
In this case, the data extraction and image decryption is a sequential process. The database manager firstly decrypts the image using encryption key then after he can retrieve the data using data hiding key.

Disadvantages of Existing system:
a) It works only on text messages.
b) There is limitation on data size to be embedded on the cover image.
c) The cover image considered here is gray-scale only.
d) Multiple data files cannot be embedded using this system.

4. Implementation Details

In this paper, we propose a reversible data hiding for digital image. Here, we are going to propose the system for embedding multiple data files in the digital cover media.

A reversible data hiding is a technique which gives lossless or distortion-free data hiding. Using this technique, not only the security of data is obtained but also the exact data and original image after extraction is obtained. In this system the digital image used for embedding the data files is called the cover image and the image generated after embedding process is called stego-image. The following diagram shows the data hiding process:

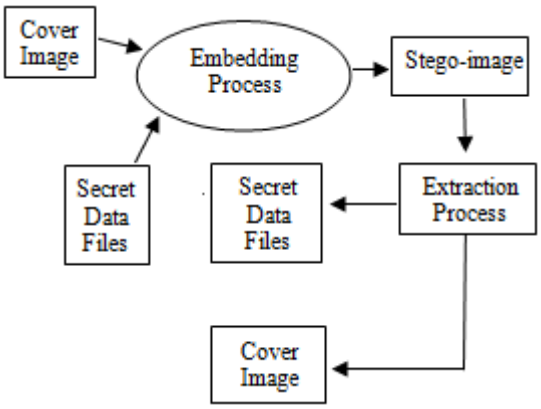

Figure 4.1: Reversible Data Hiding process

1) Platform: Microsoft Visual Studio 2010
Microsoft provides an integrated development environment called Microsoft visual studio. Microsoft visual studio is used for developing the graphical user interface, web application, web sites etc. Visual studio also has code editor which also provides syntax highlighting.

2) Proposed System
The proposed system architecture is as shown in the figure below; it consists of image encryption, data embedding, image decryption, data extraction and image recovery block. The cover image here is the original image used for hiding data. This system uses two keys one for encryption and one for data hiding. The encryption key is required for decryption process, if the key doesn't matched with encryption key image decryption is not possible. Same for data hiding key, data retrieval is possible only with data hiding key.

3) Advantages
a) The cover used in this system is color image.
b) The data to be hidden on image can be of any format.
c) Multiple data files can be added on cover image.
d) No size limitations for data files to be embedded.

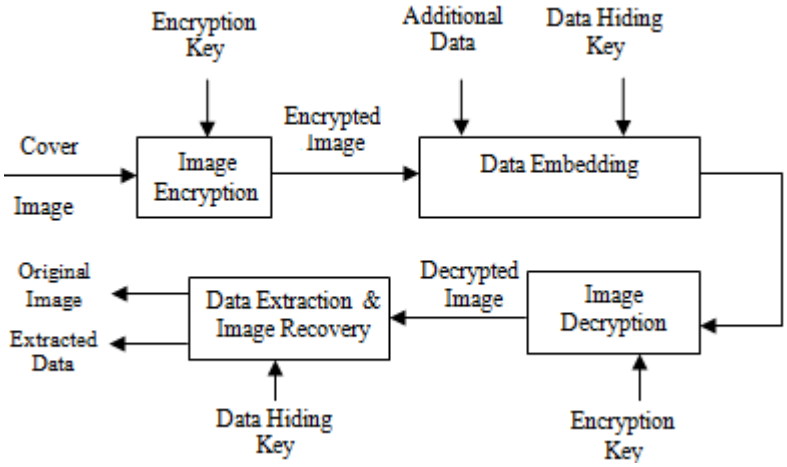

Figure 4.2: Propose system architecture

Modules Description

1. Encrypted Image Generation:

Input: In this the image passed by first module, encryption key are the inputs.

Algorithm: Advanced Encryption standard Algorithm (Rijndael Algorithm)
AES is a symmetric block cipher algorithm. It uses same key for encryption and decryption process. The stages of AES algorithm are as follows:
1. Substitute bytes
2. Shift rows
3. Mix Columns
4. Add Round Key

Output: This module outputs the marked encrypted image.

2. Data Hiding In Encrypted Image:

Input: In this the marked encrypted image, the data hiding key, the data files to be embedded are taken as inputs.

Algorithm- The steps of algorithm are as follows:
1. Transform the image into a temp jpeg first.
2. Read all the bytes from input image and store in byte array byInputImage.

```
byInputImage-> File.ReadAllBytes(InputImage)
```

3. Calculate byte array length of byInputImage.
4. Copy all the selected files into a single folder.
5. Compress all the files in a zip file.

```
ZipFile zpFiles = ZipUnzip.ZipFiles(arrFiles)
```

6. Convert the zip file to byte array byZipFiles.

```
byte[] byZipFiles = smt.ToArray()
```

7. Now, concatenate input image byte array byInputImage with zip file byte array byZipFiles.

```
byInputImage.CopyTo(byOutput, 0);
byZipFiles.CopyTo(byOutput,byInputImage.Length)
```

8. Save the bytes array byoutput on the image file.

Output: This module gives the data containing encrypted image as the output.

3. Data Extraction and Image Recovery:

Input: In this the data embedded image, data hiding key and encryption key is taken as input.

Algorithm: - To Extract data files and image

The steps for data extraction and image recovery are as follows:
1. Separate the byte array byInputImage and byZipfiles from byoutput byte array with respect to byInputImage array length.
2. Save the first part of byte array as bitmap image.
3. Save the second part as zip files array.
4. Extract the zip files array to get the attached files.

Algorithm- A.E.S Decryption Algorithm

The stages of DES are same as that of AES algorithm with the exception that each step is inverse of its counterpart in the encryption algorithm.
1. Inverse Shift rows
2. Inverse Substitute bytes
3. Inverse Add Round Key
4. Inverse Mix Columns

Output: this module gives the original data files and image as output.

Table 5.1: Comparison of PSNR values in marked decrypted grey-scale and marked decrypted RGB images

Embedding Rate(bpp)	0.005		0.01		0.05		0.1	
	ES gray	PS RGB	ES gray	PS RGB	ES gray	PS RGB	ES gray	PS RGB
Lena	65.151	67.80	62.1	65.54	55.28	57.31	51.809	54.54
Baboon	65.10	65.75	62.083	64.89	55.219	60.17	51.825	56.20
Peppers	65.120	65.87	62.128	63.80	55.214	60.79	51.831	56.16
Embedding Rate(bpp)	0.2		0.3		0.4		0.5	
	ES gray	PS RGB	ES gray	PS RGB	ES gray	PS RGB	ES gray	PS RGB
Lena	49.31	51.72	47.42	50.16	46.373	49.51	46.36	48.76
Baboon	49.209	53.16	47.402	51.71	46.363	50.84	46.365	48.55
Peppers	49.205	52.404	47.404	50.36	46.36	48.28	46.37	46.87

5. Performance Evaluation and Results

5.1. Data Used

In this paper digital color image of JPEG format and other formats like bmp, png are used as cover image. The data files used for embedding are of txt, docx, pdf etc. format. The size of data is independent of the cover size.

5.2. Results

We take standard images Lena, Baboon and Peppers to demonstrate the feasibility of proposed method. Here we analysis PSNR values of previous method in marked decrypted grey-scale images with PNSR values of our proposed method in corresponding marked decrypted color images. Table.5.1 shows comparisons result under various embedding rates. Fig.5.1 plots the PSNR results of marked decrypted image of Lena under given embedding rates. From the Fig.5.2, it can be observed that overall range of embedding rate, for all cases, our approach outperforms state-of- the-art RDH algorithms in encrypted color images. The original image and stego-image are as below:

Figure 5.1: Lena image a. original image b. recovered image after data extraction

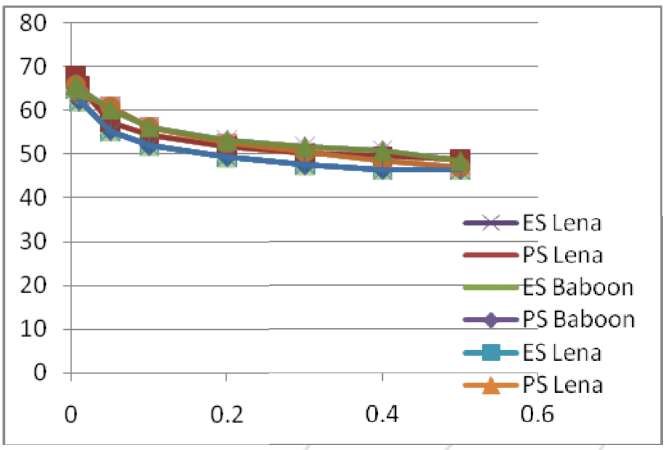

Figure 5.2: Graphical representation for PSNR comparison in test images

6. Conclusion and Future Scope

In this paper, we proposed a reversible data hiding technique to embed the secret data in color image with high security, imperceptibility and robustness. Due to proposed embedding algorithm the data embedding capacity will be significantly increased in this system. Multiple data files will be embedded in cover image. The security of the data is much more increased due to the use of two keys: encryption key and data hiding key. No one can recognize the stego image as the appearance looks same as cover image. All the data files are compressed and then embedded with cover, so more data can be embedded.

In future, we will extend this system considering audio or video files as the cover. In this paper only digital image is considered as cover.

7. Acknowledgment

First and foremost, I would like to thank my guide, Prof. Ms. J. V. Shinde, for his guidance and support. I will forever remain grateful for the constant support and guidance extended by guide, in making this report. Through our many discussions, he helped me to form and solidify ideas. The invaluable discussions I had with him, the penetrating questions he has put to me and the constant motivation, has all led to the development of this project.

I wish to express my sincere thanks to the Head of department, Prof. N. R. Wankhede, also grateful thanks to the M.E Co-coordinator Prof. Ms. J. V. Shinde and the departmental staff members for their support.

References

[1] Kede Ma, Weiming Zhang, Xianfeng Zhao, Nenghai Yu, and Fenghua Li, "Reversible data hiding in encrypted images by reserving room before encryption," IEEE Transactions on Information Forensics And Security, vol. 8, no. 3, pp. 553-562, March 2013

[2] W. Zhang, B. Chen, and N. Yu, "Improving various reversible data hiding schemes via optimal codes for binary covers," IEEE Trans.Image Process., vol. 21, no. 6, pp. 2991–3003, June. 2012.

[3] J. Fridrich, M. Goljan, and D. Rui, "'Invertible Authentication", In Proc. of SPIE Photonics West, Security and Watermarking of Multimedia Contents III, San Jose, California, USA, Vol. 3971, pp. 197-208, January 2001.

[4] J. Tian, "Reversible data embedding using a difference expansion,"IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 8, pp. 890–896,August. 2003.

[5] . Ni, Y. Shi, N. Ansari, and S. Wei, "Reversible data hiding," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 3, pp. 354–362, March.2006.

[6] X. L. Li, B. Yang, and T. Y. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," IEEE Trans. Image Process., vol. 20, no. 12, pp. 3524–3533, December.2011.

[7] L. Luo et al., "Reversible image watermarking using interpolation technique," IEEE Trans. Inf. Forensics Security, vol. 5, no. 1, pp. 187–193,March. 2010.

[8] G. Xuan, J. Chen, J. Zhu, Y.Q. Shi, Z. Ni, and W. Su, "Lossless Data hiding based on Integer Wavelet Transform", In Proc. of IEEE International Workshop on Multimedia Signal Processing. Marriott Beach Resort St. Thomas, US Virgin Islands, 9-11 December 2002.

[9] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi, "Reversible watermarking algorithm using sorting and prediction," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 7, pp. 989–999, July. 2009.

[10] M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran, "On compressing encrypted data," IEEE Trans. Signal Process., vol. 52, no. 10, pp. 2992–3006, October. 2004.

[11] K. Hwang and D. Li, "Trusted cloud computing with secure resources and data coloring," IEEE Internet Comput., vol. 14, no.5, pp.14–22, Sep./October.2010.

[12] X. Zhang, "Separable reversible data hiding in encrypted image," IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, pp. 826–832, April. 2012.

[13] Y. Chakrapani and K. SoundaraRajan," Genetic algorithm applied to fractal image compression," in ARPN Journal of Engineering and Applied Sciences , vol. 4, no. 1, February.2009

[14] MedisettyNagendra Kumar and S. Srividya,"Genetic Algorithm Based Color Image Steganogaphy Using Integer Wavelet Transform And Optimal Pixel Adjustment Process," in International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-3, Issue-5, October 2013.

[15] X. Zhang, "Reversible data hiding in encrypted images," IEEE Signal Process. Lett., vol. 18, no. 4, pp. 255–258, April. 2011.

[16] W. Hong, T. Chen, and H.Wu, "An improved reversible data hiding in encrypted images using side match," IEEE Signal Process. Lett., vol. 19, no. 4, pp. 199–202, April. 2012.