

IMPROVING AND ANALYSIS EMBEDDING SECURITY IN DIGITAL IMAGES USING RDH APPROACHES

E.JEEVAJOTHI¹, DR.K.KUPPUSAMY²

^{*1}Department of Computer Science, Alagappa University, Karaikudi, Tamilnadu, India. jeevathiru2@gmail.com¹

^{*2}Professor and Head, Department of Computational Logistics, Alagappa University, Karaikudi, Tamilnadu, India. kkdiksamy@yahoo.com²

Abstract—Reversible Data Hiding technique (RDH) is an approach to embed secret data into the original cover image in reversed way. This work proposes a new framework for hiding data into the JPEG Bit stream which integrates RSA public key cryptography, Discrete Cosine Transformation (DCT) Watermarking and the blocking artifact function. It achieves a perfect data and quality image extraction, high embedding capacity, robustness and high security in the network.

We proposes a lossless, a reversible, and a combined data hiding schemes for ciphertext images encrypted by public key cryptosystems with probabilistic and homomorphic properties. In the lossless scheme, the ciphertext pixels are replaced with new values to embed the additional data into several LSB-planes of ciphertext pixels by multi-layer wet paper coding. Then, the embedded data can be directly extracted from the encrypted domain, and the data embedding operation does not affect the decryption of original plaintext image. In the reversible scheme, a preprocessing is employed to shrink the image histogram before image encryption, so that the modification on encrypted images for data embedding will not cause any pixel oversaturation in plaintext domain. Although a slight distortion is introduced, the embedded data can be extracted and the original image can be recovered from the directly decrypted image. Due to the compatibility between the lossless and reversible schemes, the data embedding operations in the two manners can be simultaneously performed in an encrypted image. We analyze the embedding rate and PSNR with encrypted embedded image with original image .

confidential data. The major two areas steganography and cryptography provides secure data transmission over internet. Reversible Data Hiding (RDH) is based on the steganography. The data is to be hide is embedded in an encrypted image. At first the image is encrypted using any encryption algorithm then the data to be hide is embedded in the encrypted image. If the receiver has the data-hiding key he can extract the hidden data from the encrypted image even though he does not know the contents of the image. If the receiver has the key for encryption , then he can decrypt the received data to recover an image similar to the original image, but not able to extract the hidden data. If the receiver has both the keys, then he can extract the hidden data and also he can recover the original content which is errorless. We can say that a data hiding method is reversible if the original image content can be perfectly recovered from the image containing embedded data

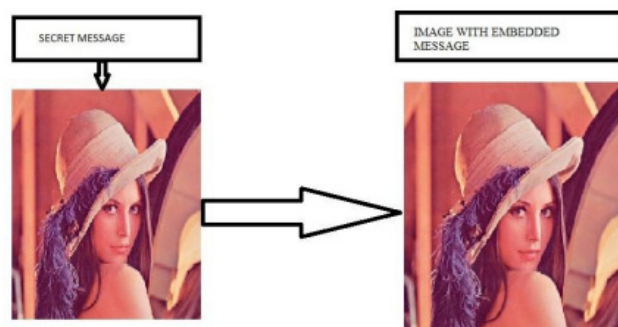


Figure 1.1 Image Based Data Hiding

Keywords – Facet, weightage, utility mining

I. INTRODUCTION

In Health care Sector during the secure transmission of health care reports over network, the sensitive data of a patient can be transmitted by embedding data in medical images. This technique improves the security of the data. Reversible Data Hiding is a kind of technique that is mainly used in case of embedding data in encrypted images. Therefore the security of the cover image can be ensured. We can use this technique where situation in which both the transmitted data and the cover image is confidential. Encryption provides security to

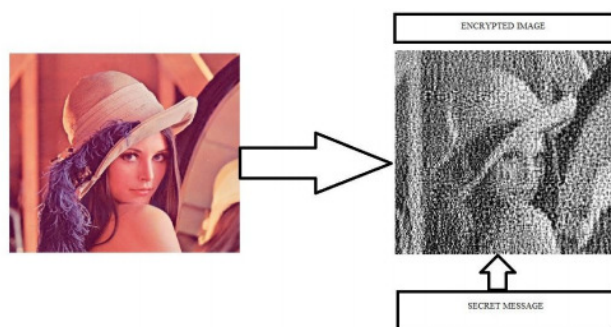


Figure 1.2 Encrypted Image Based Data Hiding

The amount of digital images has increased rapidly on the Internet. Image security becomes increasingly important for many applications, e.g., confidential transmission, video surveillance, military and medical applications. For example, the necessity of fast and secure diagnosis is vital in the medical world. Nowadays, the transmission of images is a daily routine and it is necessary to find an efficient way to transmit them over networks. To decrease the transmission time, the data compression is necessary. The protection of this multimedia data can be done with encryption or data hiding algorithms. Since few years, a problem is to try to combine compression, encryption and data hiding in a single step. For example, some solutions were proposed in to combine image encryption and compression. Two main groups of technologies have been developed for this purpose. The first one is based on content protection through encryption. There are several methods to encrypt binary images or gray level images. The second group bases the protection on data hiding, aimed at secretly embedding a message into the data. Nowadays, a new challenge consists to embed data in encrypted images. Previous work proposed to embed data in an encrypted image by using an irreversible approach of data hiding or data hiding, aimed at secretly embedding a message into the data. A new idea is to apply reversible data hiding algorithms on encrypted images by wishing to remove the embedded data before the image decryption. Recent reversible data hiding methods have been proposed with high capacity, but these methods are not applicable on encrypted images. Data security basically means protection of data from unauthorized users or hackers and providing high security to prevent data medication. This area of data security has gained more attention over the recent period of time due to the massive increase in data transfer rate over the internet. In order to improve the security features in data transfers over the internet, many techniques have been developed like: Cryptography, Steganography.

While Cryptography is a method to conceal information by encrypting it to cipher texts and transmitting it to the intended receiver using an unknown key, Steganography provides further security by hiding the cipher text into a seemingly invisible image or other formats. In recent years, signal processing in the encrypted domain has attracted

considerable research interest. As an effective and popular means for privacy protection, encryption converts the ordinary signal into unintelligible data, so that the traditional signal processing usually takes place before encryption or after decryption. However, in some scenarios that a content owner does not trust the processing service provider, the ability to manipulate the encrypted data when keeping the plain content unrevealed is desired. For instance, when the secret data to be transmitted are encrypted, a channel provider without any knowledge of the cryptographic key may tend to compress the encrypted data due to the limited channel resource. While an encrypted binary image can be compressed with a lossless manner by finding the syndromes of lowdensity parity-check codes. With the lossy compression method an encrypted gray image can be efficiently compressed by discarding the excessively rough and fine information of coefficients generated from orthogonal transform. When having the compressed data, a receiver may reconstruct the principal content of original image by retrieving the values of coefficients. Data hiding is a technique that is used to hide information in digital media such as images, audio, video etc. The information that is hidden depends upon the purpose of application. Owing to data hiding, some distortion may occur in the original cover medium and cannot be inverted back to the original medium. Such a data hiding is called lossy data hiding. But in applications such as medical image system, law enforcement, remote sensing, military imaging etc it is desired to recover the original image content with greater accuracy for legal considerations. The data hiding scheme that satisfies this requirement is called reversible or lossless data hiding. Reversible data hiding was first proposed for authentication and its important feature is reversibility. It hides the secret data in the digital image in such a way that only the authorized person could decode the secret information and restore the original image. Several data hiding methods have been proposed. The performance of a reversible data embedding algorithm is measured by its payload capacity, complexity, visual quality and security. Earlier methods have lower embedding capacity and poor image quality. As the embedding capacity and image quality improved, this method became a covert communication channel. Not only should the data hiding algorithm be given importance. The image on which the data is hidden should also be highly secured.

II. RELATED WORK

Lots of research has been done in the area of reversible data hiding. In last few years various efficient methods have been proposed for reversible data hiding. Some noticeable work in area of reversible data hiding is as follows: In Bhaskara Reddy,et.al suggested an Effective Algorithm of Encryption and Decryption of Images Using Random Number Generation Technique and Huffman coding .In this paper, they implemented security for image. They considered an image, read its pixels and convert it into pixels matrix of order as height and width of the image. Replace that pixels into some

fixed numbers, generate the key using random generation technique. Encrypting the image using this key, performing random transposition on encrypted image, converting it into one dimensional encrypted array and finally applied Huffman coding on that array, due to this size of the encrypted image is reduced and image is encrypted again. The decryption is reverse process of encryption. Hence the proposed method provides a high security for an image with minimum memory usage. The main steps in the encryption algorithm is

Step 1. Replace each pixel by fixed number values.

Step 2. Generate the secret key by using random generation technique

Step 3. Huffman Coding.

The steps in image decryption is reverse of encryption algorithm

In Subhanya R.J., Anjani Dayanandh N presented the paper "Difference Expansion Reversible Image Watermarking Schemes Using Integer Wavelet Transform Based Approach". In this project, they present a new scheme of image watermarking to guard intellectual properties and to secure the content of digital images. It is an effective way to protect the copyright by image watermarking. The work concerns with the watermarking algorithm that embeds image/text data invisibly into a video based on Integer Wavelet Transform and to minimize the mean square distortion between the original and watermarked image and also to increase Peak signal to noise ratio. Here the message bits (image) are hidden into gray/color images. The size of secret data/image is smaller than cover image. To transfer the secret image/text confidentiality, the secret image/text itself is not hidden, keys are generated for each gray/color component and the IWT is used to hide the keys in the corresponding gray/color component of the cover image. The watermarks are invisible and robust against noise and commonly image processing methods.

Zhang suggests a novel method for separable reversible data hiding. Here content owner first encrypts the original uncompressed image using an encryption key to produce an encrypted image. Then, the data-hider compresses the least significant bits (LSB) of the encrypted image using a data-hiding key to create a sparse space to accommodate the additional data. At the receiver side, the data embedded in the created space can be easily retrieved from the encrypted image containing additional data according to the data-hiding key. Since the data embedding only affects the LSB, a decryption with the encryption key can result in an image similar to the original version. When using both of the encryption and data-hiding keys, the embedded additional data can be successfully extracted and the original image can be perfectly recovered by exploiting the spatial correlation in natural image. In C. Anuradha and S. Lavanya proposed a secure and authenticated discrete reversible Data hiding in cipher images deals with security and authentication. In the first phase, a content owner encrypts the original uncompressed image using an encryption key. Then, a data hider may compress the least significant bits

of the encrypted image using a data hiding key to create a sparse encrypted image containing additional data, if a receiver has the data hiding key, receiver can extract the additional data though receiver does not know the image content. If the receiver has the encryption key, can decrypt the received data to obtain an image similar to the original one, but cannot extract the additional data. If the receiver has both the data hiding key and the encryption key, can extract the additional data and recover the original content without any error by exploiting the spatial correlation in natural image when the amount of additional data is not too large. It is also a drawback because if the receiver has any one key as known, and then he can take any one information from the encrypted data. In order to achieve authentication SHA-1 algorithm is being used.

Che-Wei Lee and Wen-Hsiang Tsai proposed a lossless data hiding method based on histogram shifting, which employs a scheme of adaptive division of cover images into blocks to yield large data hiding capacities as well as high stego-image qualities. The method is shown to break a bottleneck of data-hiding-rate increasing at the image block size of 8×8 , which is found in existing histogram-shifting methods. Four ways of block divisions are designed, and the one which provides the largest data hiding capacity is selected adaptively.

Yang Yang et al. proposed a method [1] prior to embed message into the texture area of the medical images for improving the quality of the details information and helping in accurate diagnosis. Again in order to decrease the embedding distortion while enhancing the contrast of the texture area, this paper also proposed a message sparse representation method. Various experiments implemented on medical images showed that the proposed method enhances the contrast of texture area when compared with previous methods. They proposed a RDH method in medical images with texture area enhancement based on the idea of histogram stretching. The proposed method consists of four parts: 1) rhombus prediction and texture-based sorting; 2) embedding scheme and enhancing contrast of texture area; 3) message sparse representation; 4) message extraction and cover image recovery. Smita Agrawal and Manoj Kumar proposed, a novel reversible data hiding technique, based on integer-to-integer wavelet transform and histogram-bin-shifting for medical images [2]. In that proposed system images are divided into blocks and entropy of each block is calculated to evaluate the smoothness of the blocks. Integer-to-integer wavelet transform was applied over smooth blocks and watermark was embedded in all sub bands of detail part. Histogram-bin-shifting technique was used to embed the watermark. The proposed scheme was applied on various medical images and also compared with one of the recent existing reversible data hiding techniques. Higher PSNR values demonstrate the effectiveness of the proposed scheme. Yun-Qing Shi et al. [3] discussed the various RDH algorithms into the following six categories: 1) RDH into image spatial domain; 2) RDH

into image compressed domain (e.g., JPEG); 3) RDH suitable for image semifragile authentication; 4) RDH with image contrast enhancement; 5) RDH into encrypted images, which is expected to have wide application in the cloud computation; and 6) RDH into video and into audio. For each of these six categories, the history of technical developments, the current state of the arts, and the possible future researches are presented and discussed.

Zhenxing Quian and Xinpeng Zhang proposed a Reversible Data Hiding Technique (RDH) using distributed source encoding[4]. It consists of three phases: image encryption, data embedding and data extraction/image recovery. In the first phase, the sender turns the original image into plain bits by decomposing each pixel into 8 bits. The owner then chooses an encryption key to generate pseudo-random bits using a stream cipher function and encrypts the bit stream of the original image. After the image encryption, the content owner sends the encrypted image to the data hider. To embed additional data into the image, the data hider first decomposes the encrypted image into four sub images of equal sizes. Bits of three MSB planes of the sub images are collected and using a selection key the data hider pseudo randomly selects L bits from them and shuffles the selected bits. It is controlled by a shuffle key. The shuffled bits are divided into groups. Then the data hider uses the stepian-wolf codes to compress the selected bits C. On the receiver end, with the marked encrypted image, the hidden data can be extracted using the embedded key, and the original image can be approximately reconstructed using the encryption key, or lossless recovery using both of the keys.

III. PROBLEM FORMULATION

To maintain the security and authentication, Reversible Data Hiding i.e. RDH techniques are related to steganography and cryptography function. Encryption and data hiding are two techniques of data protection. Data hiding techniques embeds original data which we don't want to disclose into cover media by introducing slight acceptable modifications, while encryption techniques converts plaintext data into unreadable form i.e. ciphertext. It is beneficial to embed the data into a digital media to communicate the secret. The message signal can be recovered with no loss but the original cover can be lost. So in general reversible data hiding techniques can be used now a days. Method of reversible data hiding are emerging and difficult to implement with DCT and other Encryption standards.

IV. PROPOSED WORK IMPLEMENTATION

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 METHODOLOGIES

- Data Extraction
- Data Embedding
- 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

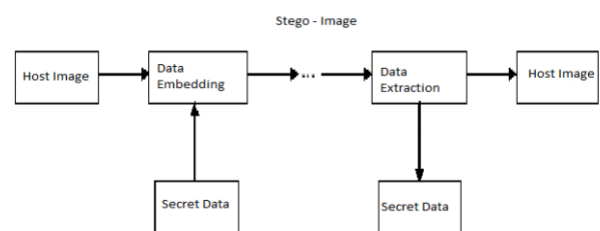


Fig: 3 Proposed work flow diagram

The first step of every image processing application is image acquisition or image capturing. The images of leaves are captured by using the camera and it will store it in some formats like .PNG, .JPG, .JPEG etc.

1. DCT Computation of Image

A **discrete cosine transform (DCT)** expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications in science and engineering, from lossy compression of audio (e.g. MP3) and images (e.g. JPEG) (where small high-frequency components can be discarded), to spectral methods for the numerical solution of partial differential equations. The use of cosine rather than sine functions is critical for compression, since it turns out (as described below) that fewer cosine functions are needed to approximate a typical signal, whereas for differential equations the cosines express a particular choice of boundary.

In particular, a DCT is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample. There are eight standard DCT variants, of which four are common.

In our work, using following command we compute the pixel values from DCT from the image pixels.

$$H = \text{dct}(I, 255)$$

ENCRYPTION PROCESS

Though the fields of steganography and cryptography are associated with one another, there is a distinction to be made. Cryptography is the art of jumbling a message so that a would-be eavesdropper cannot interpret the message. Steganography, on the other hand, is the art of hiding a message so that a would-be eavesdropper is unaware of the message's presence.

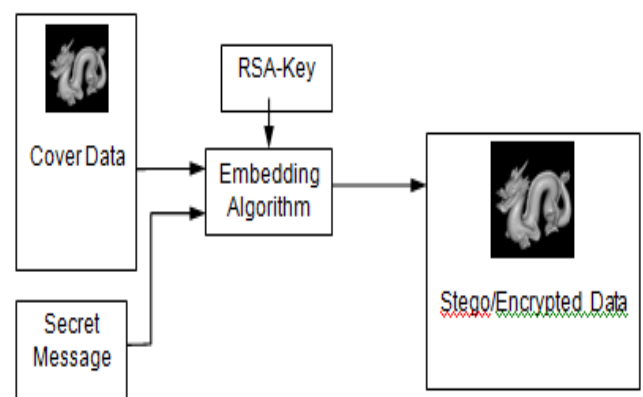
While steganography has been around for centuries, the Digital Revolution has sparked a renewed interest in the field. For instance, the mass media industry has shown increasing interest in steganography to fight piracy. It is even rumored that the terrorist organization, Al-Qaeda, has employed steganography to transmit orders to its operatives over the internet.

All digital file formats can be used to hide secret messages. Our propose work focuses specifically on the techniques employed in hiding information in digital image files.

I. A GENERIC STEGANOGRAPHIC/CRYPTOGRAPHY SYSTEM

As with any other science, steganography has its own set of terminology. The term *cover* is used to describe the original message in which we will hide our secret message. Once we embed our secret message into the cover, the new message is known as the *stego data*. The *stego data* is analogous the *cipher text* of cryptography.

A generic steganographic system, or *stegosystem*, works thusly. A secret message is embedded into the cover data using some sort of embedding algorithm. The cover data may be a single file, but that is not necessarily the case. The embedding algorithm then outputs the stego data. There is, however, a minor detail that needs to be added to the system. Recall Kerckhoff's principle, which states that the security of a system should not be based on the obscurity of the algorithm, but on the strength of its key. Therefore the embedding algorithm should require a key as an input. Additionally it is advisable to encrypt the secret message prior to embedding it.

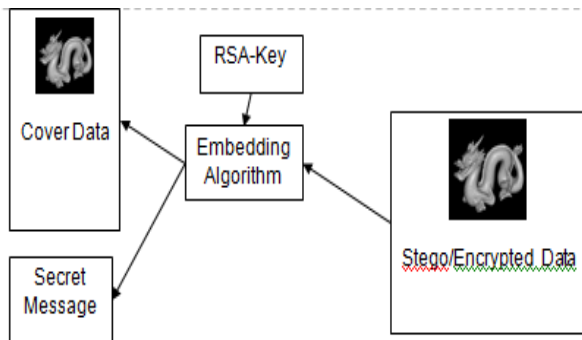


Though embedding algorithms may take many forms, there are some requirements that all embedding algorithms should meet. Firstly, the distortion of the cover data as a result of the embedding algorithm should be as imperceptible as possible. Secondly, no part of the secret message should be contained in the header of the stego data file. The message must become part of the cover data and should be immune to manipulation attacks such as re-sampling or filtering. Ideally, it would also be a good idea to include error correcting codes into the message so that if the stego data is damaged, the message can still be recovered. Finally, it is imperative that the original cover data never fall into the hands of an eavesdropper or be used twice. Since the embedding process is additive, the secret message can be recovered if an eavesdropper has different stego files which utilize the same cover data.

We will now explore some of the more popular techniques for embedding messages into cover.

DECRYPTION PROCESS:

The above Encryption process has applied reverse biasing we get image and hiding text separately with RSA algorithm basis.



EMBEDDING ALGORITHM

Input: Cover image, Secret data

Output: Encrypted image

Algorithm:

Step 1: Read both cover and secret image.

Step2: Divide the cover image in to 8x8 blocks of pixels.

Step3: Apply two dimensional DCT on each 8x8 blocks of pixels to get 64 DCT coefficients.

Step4: Quantize the 64 DCT coefficients in to the rounded value.

Step5: Encrypt the secret data using RSA algorithm.

Step6: Divide the encrypted image in to 8x8 blocks of pixels.

Step8: Decrypt the secret data and Image using RDH

Step 9: Performance Analysis of RDH

VI. EXPERIMENTAL RESULTS

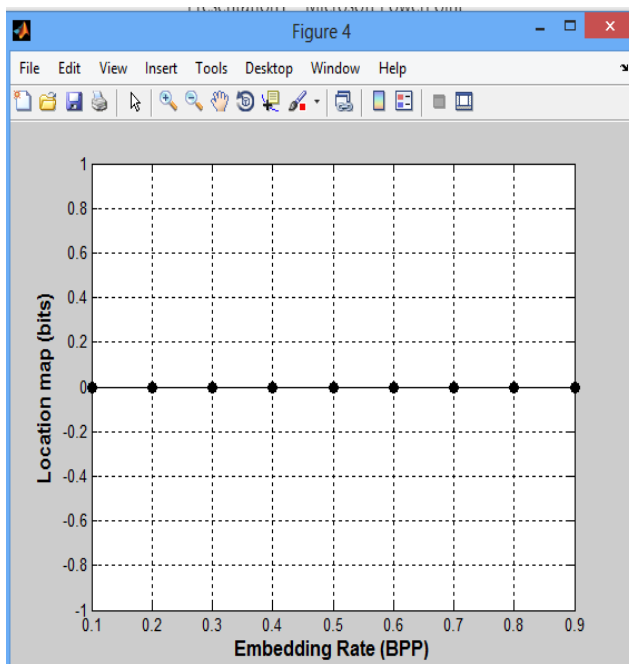


Fig: Embedded Rate Vs location Map of digital images

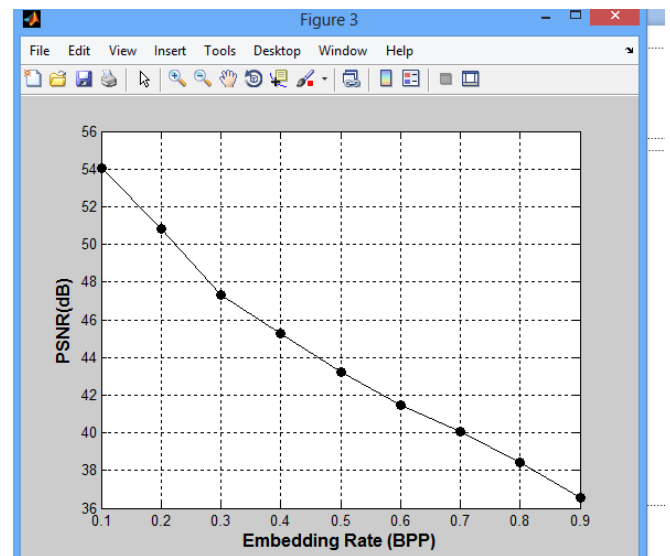


Fig: Embedded Rate vs PSNR

V CONCLUSION

In this correspondence, we propose an RDH framework for encrypted JPEG bitstream. The original JPEG bitstream is properly encrypted to hide the image content with the bitstream structure preserved. The secret message bits are encoded with image and embedded into the encrypted bitstream by modifying the appended bits. By using the encryption and embedding keys, the receiver can extract the embedded data capacity and perfectly restore the original image. When the embedding key is absent, the original image can be approximately recovered with satisfactory quality without extracting the hidden data. In future, we can implement the video RDH with RSA algorithms.

REFERENCES

- [1] Xiaochun Cao, Senior Member, IEEE, Ling Du, Xingxing Wei, Dan Meng, Member, IEEE, and XiaojieGuo, Member, IEEE, "High Capacity Reversible Data Hiding in Encrypted Images by Patch-Level Sparse Representation", IEEE Transactions on Cybernetics.
- [2] M. U. Celik, G. Sharma, and A. M. Tekalp, "Lossless generalized-LSB data embedding," IEEE Trans. Image Process., vol. 14, no. 2, pp. 253–266, Feb. 2005.
- [3] J. Fridrich, M. Goljan, and R. Du, "Invertible authentication watermark for JPEG images," in Proc. Inf. Technol. Coding Comput., Las Vegas, NV, USA, Apr. 2001, pp. 223–227.
- [4] H. J. Kim, V. Sachnev, Y. Q. Shi, J. Nam, and H. G. Choo, "A novel difference expansion transform for reversible data embedding," IEEE Trans. Inf. Forensics Security, vol. 3, no. 3, pp. 456–465, Sep. 2008.

[5] D. Coltuc, "Improved embedding for prediction based reversible watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 873–882, Sep. 2011

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

[7] Y. Hu, H. K. Lee, K. Chen, and J. Li, "Difference expansion based reversible data hiding using two embedding directions," *IEEE Trans. Multimedia*, vol. 10, no. 8, pp. 1500–1511, Dec. 2008

[8] B. Ou, X. Li, Y. Zhao, R. Ni, and Y.-Q. Shi, "Pairwise prediction error expansion for efficient reversible data hiding," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5010–5021, Dec. 2013.

[9] W. L. Tai, C. M. Yeh, and C. C. Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 6, pp. 906–910, Jun. 2009.

[10] C. C. Lin, W. L. Tai, and C. C. Chang, "Multilevel reversible data hiding based on histogram modification of difference images," *Pattern Recognit.*, vol. 41, no. 12, pp. 3582–3591, Dec. 2008.

[11] P. Tsai, Y. C. Hu, and H. L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting," *Signal Process.*, vol. 89, no. 6, pp. 1129–1143, Jun. 2009.

[12] S. L. Lin, C. F. Huang, M. H. Liou, and C. Y. Chen, "Improving histogram-based reversible information hiding by an optimal weight-based prediction scheme," *J. Inf. Hiding Multimedia Signal Process.*, vol. 4, no. 1, pp. 19–33, Jan. 2013

and Information Security, Algorithms, Neural Networks, Software Engineering and Optimization Techniques.



E. Jeevajothi M.sc., B.Ed., is working as a computer instructor in Private higher secondary school from Paramakudi. She has 7 years of teaching experience. Her areas of Project UG level is VB Editor and PG level project area is Real time Audio 5.1 surround digital media.



Dr. K. KUPPUSAMY is working as Professor and Head, Department of Computational Logistics, Alagappa University, Karaikudi, Tamilnadu, India. He received his Ph.D in Computer Science and Engineering from Alagappa University, Karaikudi, Tamilnadu in the year 2007. He has 30 years of teaching experience at PG level in the field of Computer Science. He has published many research papers in the reputed International and National Journals and presented in the National and International conferences. His areas of research interests include Network