

A Spatial Domain Image Steganography Technique Based on Plane Bit Substitution Method

S.M.Riyazoddin¹

¹ CMR Institute of Technology, Hyderabad, Andhra Pradesh.

Received: 9 February 2012 Accepted: 5 March 2012 Published: 15 March 2012

Abstract

Steganography is the art and science of hiding information by embedding data into cover media. In this paper we propose a new method of information hiding in digital image in spatial domain. In this method we use Plane Bit Substitution Method (PBSM) technique in which message bits are embedded into the pixel value(s) of an image. We first, proposed a Steganography transformation machine (STM) for solving Binary operation for manipulation of original image with help to least significant bit (LSB) operator based matching. Second, we use pixel encryption and decryption techniques under theoretical and experimental evolution. Our experimental, techniques are sufficient to discriminate analysis of stego and cover image as each pixel based PBSM, and operand with LSB.

Index terms— spatial domain, pbsm, stm, lsb.

1 Introduction

The word Steganography is of Greek origin and means "concealed writing" from the Greek words steganos meaning "covered or protected", and graphein meaning "to write". The first recorded use of the term was in 1499 by Johannes Trithemius in his *Steganographia*, a treatise on cryptography and steganography disguised as a book on magic. Generally, messages will appear to be something else: images, articles, shopping lists, or some other cover text and, classically, the hidden message may be in invisible ink between the visible lines of a private letter. "Steganography niche in security is to supplement cryptography, not replace it. If a hidden message is encrypted, it must also be decrypted if discovered, which provides another layer of protection." There are several approaches in classifying steganographic systems. One could categorize them according to the type of covers used for secret communication. A classification according to the cover modifications applied in the embedding process is another possibility. Although in some cases an exact classification is not possible, the group steganographic methods are of six categories: Substitution systems substitute redundant parts of a cover with a secret message; Transform domain techniques embed secret information in a transform space of the signal (e.g., in the frequency domain); Spread spectrum techniques adopt ideas from spread spectrum. Author : CMR Institute of Technology, Hyderabad, Andhra Pradesh.

But as we know steganography deals with hiding of information in some cover source. On the other hand, Steganalysis is the art and science of detecting messages hidden using steganography; this is analogous to cryptanalysis applied to cryptography. The goal of steganalysis is to identify suspected packages, determine whether or not they have a payload encoded into them, and, if possible, recover that payload. Hence, the major challenges of effective Steganography are:-1. Security of Hidden Communication: In order to avoid raising the suspicions of eavesdroppers, while evading the meticulous screening of algorithmic detection, the hidden contents must be invisible both perceptually and statistically. 2. Size of Payload: Unlike watermarking, which needs to embed only a small amount of copyright information, steganography aims at hidden communication and therefore usually requires sufficient embedding capacity. Requirements for higher payload and secure communication are often contradictory. Depending on the specific application scenarios, a tradeoff has to be sought.

One of the possible ways of categorizing the present steganalytic attacks is on the following two categories 1. Visual Attacks: These methods try to detect the presence of information by visual inspection either by the naked eye or by a computer. The attack is based on guessing the embedding layer of an image (say a bit plane) and then visually inspecting that layer to look for any unusual modifications in that layer. 2. Statistical Attacks: These methods use first or higher order statistics of the image to reveal tiny alterations in the statistical behavior caused by steganographic embedding and hence can successfully detect even small amounts of embedding with very high accuracy. This system can be explained using the 'prisoners problem' (Figure 1.1) where Alice and Bob are two inmates who wish to communicate in order to hatch an escape plan. However communication between them is examined by the warden, Wendy. To send the secret message to Bob, Alice embeds the secret message 'm' into the cover object 'c', to obtain the stego object 's'. The stego object is then sent through the public channel. In a pure steganographic framework, the technique for embedding the message is unknown to Wendy and shared as a secret between Alice and Bob. In private key steganography Alice and Bob share a secret key which is used to embed the message. The secret key, for example, can be a password used to seed a pseudorandom number generator to select pixel locations in an image cover-object for embedding the secret message. Wendy has no knowledge about the secret key that Alice and Bob share, although she is aware of the algorithm that they could be employing for embedding messages. In public key steganography, Alice and Bob have private/public key pairs and know each other's public key. In this thesis we confine ourselves to private key steganography only.

2 b) Different Kinds of Steganography

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display. The redundant bits of an object are those bits that can be altered without the alteration being detected easily. Image and audio files especially comply with this requirement, while research has also uncovered other file formats that can be used for information hiding. Hiding information in text is historically the most important method of steganography. An obvious method was to hide a secret message in every n th letter of every word of a text message. It is only since the beginning of the internet and all the different digital file formats that this has decreased in importance.

This Paper is organized as follows: In section II, "Literature Survey", we give a background of the existing state of the steganographic research. We cover briefly the main categories of steganographic algorithms covered till date although the survey is not exhaustive and we may have missed out some of the algorithms. In section III we discuss the method how the steganography transformation takes place, in section IV we discuss the proposed PBSM an adaptive method for secured image steganography is discussed and in last section we tried to show the experimental results.

3 II.

4 Literature survey

Here we discuss the necessary background required for this work. In section 2.1 we discuss briefly some of the existing steganographic techniques. In section 2.2 we present some of the steganalytic attacks proposed till date as a counter measure to the steganographic algorithms.

5 a) Existing Steganographic Techniques

The steganographic algorithms proposed in literature can broadly be classified into two categories.

6 Spatial Domain Techniques 2. Transform Domain Techniques

Each of these techniques is covered in detail in the next two subsections.

7 i. Spatial Domain

These techniques use the pixel gray levels and their color values directly for encoding the message bits. These techniques are some of the simplest schemes in terms of embedding and extraction complexity. The major drawback of these methods is amount of additive noise that creeps in the image which directly affects the Peak Signal to Noise Ratio and the statistical properties of the image. Moreover these embedding algorithms are applicable mainly to lossless image-compression schemes like TIFF images. For lossy compression The most common algorithm belonging to this class of techniques is the Least Significant Bit (LSB) replacement technique in which the least significant bit of the binary representation of the pixel gray levels is used to represent the message bit. This kind of embedding leads to an addition of a noise of $0.5p$ on average in the pixels of the image where p is the embedding rate in bits/pixel. This kind of embedding also leads to an asymmetry and a grouping in the pixel gray values $(0,1);(2,3); \dots (254, 255)$. this asymmetry is exploited in the attacks developed for this technique as explained further in section 2.2. To overcome this undesirable asymmetry, the decision of changing the least significant bit is randomized i.e. if the message bit does not match the pixel bit, then pixel bit is either increased or decreased by 1. This technique is popularly known as LSB Matching. It can be observed

that even this kind of embedding adds a noise of 0:5p on average. To further reduce the noise, [19] have suggested the use of a binary function of two cover pixels to embed the data bits. The embedding is performed using a pair of pixels as a unit, where the LSB of the first pixel carries one bit of information, and a function of the two pixel values carries another bit of information. It has been shown that embedding in this fashion reduces the embedding noise introduced in the cover signal. In [21], a multiple base number system has been employed for embedding data bits. While embedding, the human vision sensitivity has been taken care of. The variance value for a block of pixels is used to compute the number base to be used for embedding. A similar kind of algorithm based on human vision sensitivity has been proposed by [22] by the name of Pixel Value Differencing. This approach is based on adding more amounts of data bits in the high variance regions of the image for example near "the edges" by considering the difference values of two neighboring pixels. This approach has been improved further by clubbing it with least significant bit embedding in [17]. According to [17], "For a given medium, the steganographic algorithm which makes fewer embedding changes or adds less additive noise will be less detectable as compared to an algorithm which makes relatively more changes or adds higher additive noise." Following the same line of thought Crandall [17] have introduced the use of an Error Control Coding technique called "Matrix Encoding". In Matrix Encoding, q message bits are embedded in a group of $2^{1/q}$ cover pixels while adding a noise of $1/2^{1/q}$ per group on average. The maximum embedding capacity that can be achieved is $1/2^{1/q}$. For example, 2 bits of secret message can be embedded in a group of 3 pixels while adding a noise of 0:75 per group on average. The maximum embedding capacity achievable is $2/3 = 0:67$ bits/pixel. F5 algorithm [17] is probably the most popular implementation of Matrix Encoding. LSB replacement technique has been extended to multiple bit planes as well. Recently [20] has claimed that LSB replacement involving more than one least significant bit planes is less detectable than single bit plane LSB replacement. Hence the use of multiple bit planes for embedding has been encouraged. But the direct use of 3 or more bit planes leads to addition of considerable amount of noise in the cover image. [17] Have given a detailed analysis of the noise added by the LSB embedding in 3 bit planes. Also, a new algorithm which uses a combination of Single Digit Sum Function and Matrix Encoding has been proposed. It has been shown analytically that the noise added by the proposed algorithm in a pixel of the image is 0:75p as compared to 0:875p added by 3 plane LSB embedding where p is the embedding rate. One point to be observed here is that most of the approaches proposed so far are based on minimization of the noise embedded in the cover by the algorithm. Another direction of steganographic algorithm is preserving the statistics of the image which get changed due to embedding. Chapter 2 of this thesis proposes two algorithms based on this approach itself. In the next section we cover some of the transform domain steganographic algorithms.

8 Global

ii. Transform Domain These techniques try to encode message bits in the transform domain coefficients of the image. Data embedding performed in the transform domain is widely used for robust watermarking.

Similar techniques can also realize large capacity embedding for steganography. Candidate transforms include discrete cosine Transform (DCT), discrete wavelet transform (DWT), and discrete Fourier transform (DFT). By being embedded in the transform domain, the hidden data resides in more robust areas, spread across the entire image, and provides better resistance against signal processing. For example, we can perform a block DCT and, depending on payload and robustness requirements, choose one or more components in each block to form a new data group that, in turn, is pseudo randomly scrambled and undergoes a second-layer transformation.

Modification is then carried out on the double transform domain coefficients using various schemes. These techniques have high embedding and extraction complexity. Because of the robustness properties of transform domain embedding, these techniques are generally more applicable to the "Watermarking" aspect of data hiding. Many steganographic techniques in these domain have been inspired from their watermarking counterparts. F5 [17] quantized coefficients to the nearest data bit. It also uses Matrix Encoding for reducing the embedded noise in the signal. F5 is one of the most popular embedding schemes in DCT domain steganography, though it has been successfully broken in [17].

The transform domain embedding does not necessarily mean generating the transform coefficients on blocks of size 8 x 8

as done in JPEG compression techniques. It is possible to design techniques which take the transforms on the whole image [17]. Other block based JPEG domain and wavelet based embedding algorithms have been proposed in [17] with respectively.

9 b) Existing Attacks

The steganalytic attacks developed till date can be classified into visual and statistical attacks.

The statistical attacks can further be classified as 1. Targeted Attacks 2. Blind Attacks Each of these classes of attack is covered in detail in the next two subsections along with several examples of each category.

10 i. Targeted Attacks

These attacks are designed keeping a particular steganographic algorithm in mind. These attacks are based on the image features which get modified by a particular kind of steganographic embedding. A particular steganographic

algorithm imposes a specific kind of behavior on the image features. This specific kind of behavior of the image statistics is exploited by the targeted attacks. Some of the targeted attacks are as follows:

1. Histogram Analysis: The histogram analysis method exploits the asymmetry introduced by LSB replacement. The main idea is to look for statistical artifacts of embedding in the histogram of a given image. It has been observed statistically that in natural images (cover images), the number of odd pixels and the number of even pixels are not equal. For higher embedding rates of LSB Replacement these quantities tend to become equal. So, based on this artifact a statistical attack based on the Chi-Square Hypothesis Testing is developed to probabilistically suggest one of the following two hypothesis: Null Hypothesis H0: The given image contains steganographic embedding Alternative Hypothesis H1: The given image does not contain steganographic embedding the decision to accept or reject the Null Hypothesis H0 is made on basis of the observed confidence value p . A more detailed discussion on Histogram Analysis can be found in [24]. another LSB steganalysis technique that can detect the existence of hidden messages that are randomly embedded in the least significant bits of natural continuous-tone images. It can precisely measure the length of the embedded message, even when the hidden message is very short relative to the image size. The key to this methods success is the formation of 4 subsets of pixels (X, Y, U, and V) whose cardinalities change with LSB embedding (as shown in Figure 2.1), and such changes can be precisely quantified under the assumption that the embedded bits are randomly scattered. A detailed analysis on Sample Pair technique can be found in [23].

Another attack called RS Steganalysis based on the same concept has been independently proposed by [25].

ii. Blind Attacks

The blind approach to steganalysis is similar to the pattern classification problem. The pattern classifier, in our case a Binary Classifier, is trained on a set of training data. The training data comprises of some high order statistics of the transform domain of a set of cover and stego images and on the basis of this trained dataset the classifier is presented with images for classification as a non-embedded or an embedded image. Many of the blind steganalytic techniques often try to estimate the cover image statistics from stego image by trying to minimize the effect of embedding in the stego image. This estimation is sometimes referred to as "Cover Image Prediction". Some of the most popular blind attacks are defined next. 1. Wavelet Moment Analysis: Wavelet Moment Analyzer (WAM) is the most popular Blind Steganalyzer for Spatial Domain Embedding. It has been proposed by [40]. WAM uses a denoising filter to remove Gaussian noise from images under the assumption that the stego image is an additive mixture of a nonstationary Gaussian signal (the cover image) and a stationary Gaussian signal with a known variance (the noise). As the filtering is performed in the wavelet domain, all the features (statistical moments) are calculated as higher order moments of the noise The detailed procedure for calculating the WAM features in a gray scale image can be found in [17]. WAM is based on a 27 dimension feature space. It then uses a Fisher Linear Discriminant (FLD) as a classifier. It must be noted that WAM is a state of the art steganalyzer for Spatial Domain Embedding and no other blind attack has been reported which performs better than WAM. The source code of this attack is available at [17]. After extraction of features, a Support Vector Machine (SVM) is used for classification. We would like to mention that although in [17] a SVM has been used for classification we have used the Linear Discriminant Analysis for classification. Some other blind attacks have also been proposed in literature. [17] Have modeled the difference between absolute value of neighboring DCT coefficients as a Markov process to extract 324 features for classifying images as cover or stego. [27] Have extended the features of [26] to 193 and clubbed them with 72 features derived by reducing the 324 extracted by [17].

III.

Steganography transformation a) LSB Replacement

LSB replacement is a well-known steganographic method. In this embedding scheme, only the LSB plane of the cover image is overwritten with the secret bit stream according to a pseudorandom number generator (PRNG). As a result, and thus it is very easy to detect the existence of hidden message even at a low embedding rate using some reported steganalytic algorithms b) LSB Matching LSB matching (LSBM) employs a minor modification to LSB replacement. If the secret bit does not match the LSB of the cover image, then +1 or -1 is randomly added to the corresponding pixel value. Statistically, the probability of increasing or decreasing for each modified pixel value is the same and so the obvious asymmetry artifacts introduced by LSB replacement can be easily avoided.

Adaptive Method for Steganography

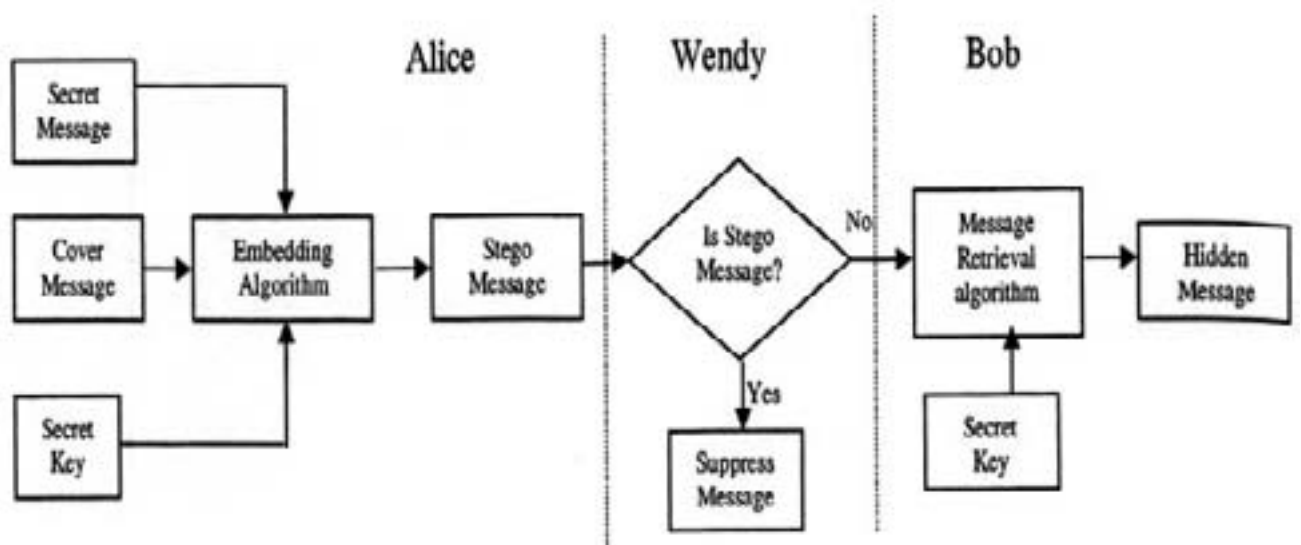
To increase the security and the size of stored data, a new adaptive LSB technique is used. Instead of storing the data in every least significant bit of the pixels, this technique tries to use more than one bit in a pixel in such a way that this change will not affect the visual appearance of the host image. It uses the side information of neighboring pixels to estimate the number of bit which can be carried in the pixels of the host-image to hide the secret data called PBSM.

¹© 2012 Global Journals Inc. (US)

²© 2012 Global Journals Inc. (US) Global Journal of Computer Science and Technology



Figure 1:



1

Figure 2: Fig. 1 .

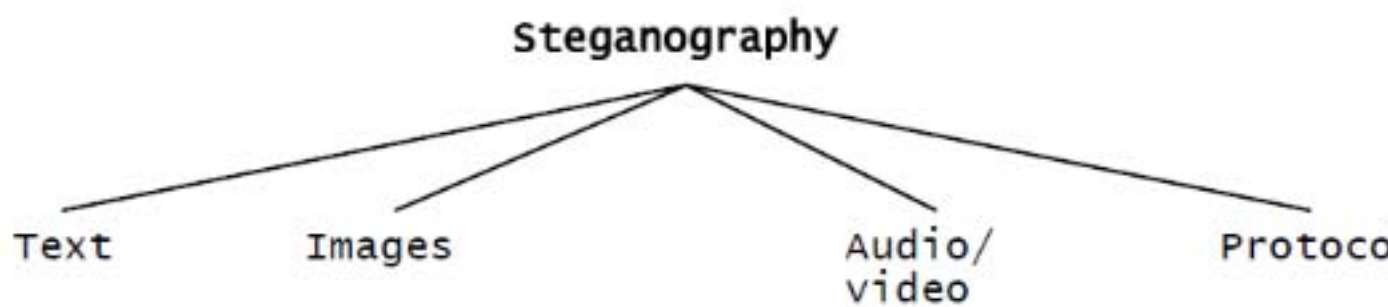


Figure 3:

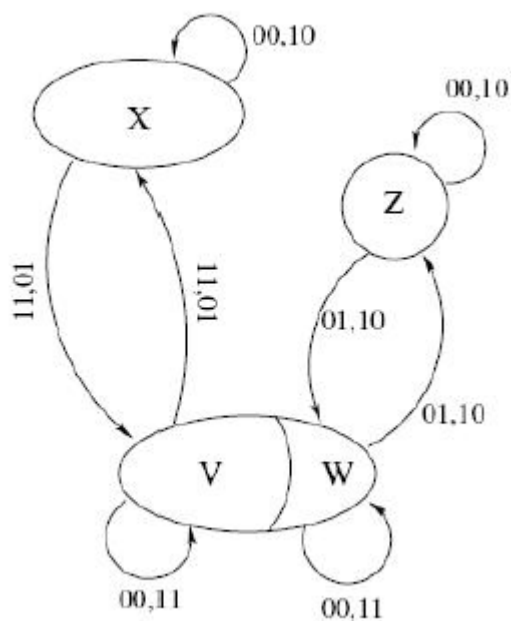
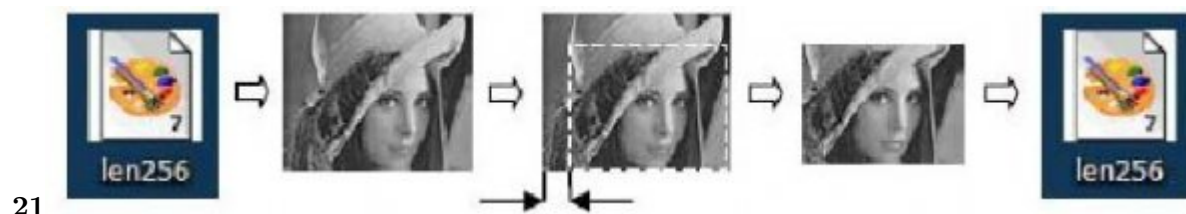


Figure 4:



21

Figure 5: Figure 2 . 1 :

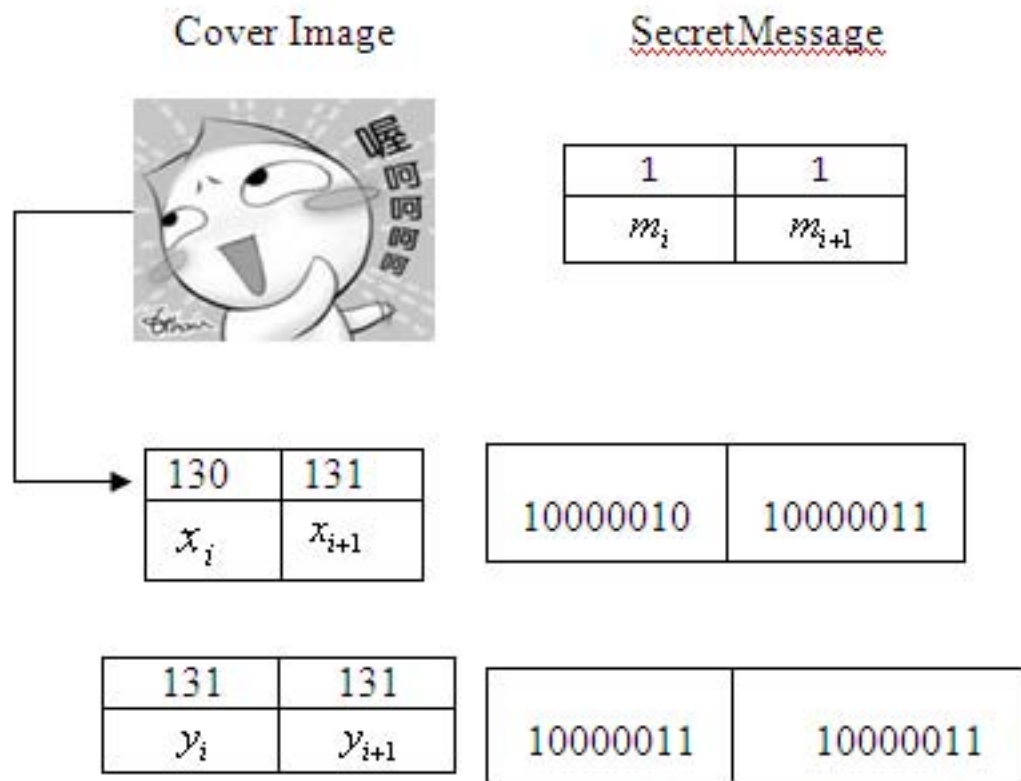


Figure 6:

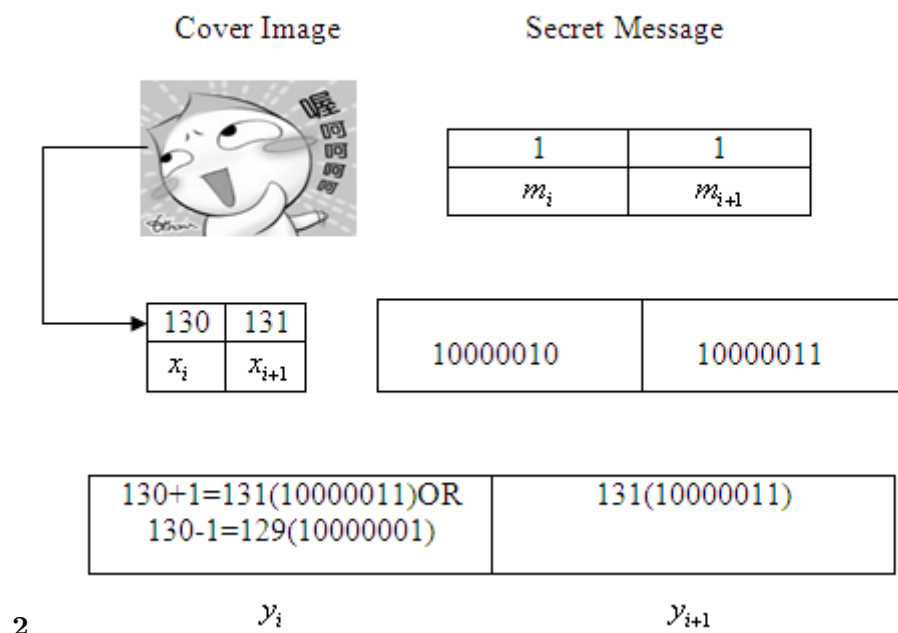


Figure 7: Fig. 2

Step1: Binary function

$$f(y_i, y_{i+1}) = LSB([y_i / 2] + y_{i+1})$$

$$\text{Primary1: } f(L-1, R) \neq f(L+1, R)$$

$$\text{Primary2: } f(L, R) \neq f(L, R+1) \\ \neq f(L, R-1)$$

$$f(129, 140) = 1$$

$$f(131, 140) = 0$$

$$f(130, 140) = 0$$

$$f(130, 140) = 0$$

$$f(130, 141) = 1$$

$$f(130, 139) = 1$$

$$f(y_i, y_{i+1}) = LSB([y_i / 2] + y_{i+1})$$

$$f(130, 140) = 0$$

$$f(130, 140) = 0 \quad f(130, 141) = 1$$

$$f(130, 139) = 1$$

1

Figure 8: a) Sending Algorithm 1 .

Step 2:

$$m_i = LSB(y_i)$$

$$m_{i+1} = f(y_i, y_{i+1})$$

130	140
x_i	x_{i+1}

10000010	10001100
----------	----------

$$f(130, 140) = 0$$

0	0
m_i	m_{i+1}



130	140
-----	-----

Figure 9:

[Goel and Pdf] , Goel , Pdf .

[Speaker and Tai (2006)] , Jarnomielikainen Speaker , : Wei-Liang Tai . *Source: IEEE signal processing letters* May 2006. 13 (5) p. 285287. (LSB matching Revisited)

[Wu and Tsai (2003)] ‘A Steganographic method for images by pixel-value differencing’. D C Wu , W H Tsai . *Pattern Recognition Letters* Jan. 2003. 24 p. .

[Zhang and Zhang (2010)] ‘A Survey of Recent Advances in Face Detection’. C Zhang , Z Zhang . MSR-TR-2010-66. *Microsoft Technical Report* June 2010.

[Swain and Kumar (2006)] ‘A Technique for Secret Communication Using a New Block Cipher with Dynamic Steganography, Research Scholar-CSE’. Gandharba Swain , Saroj Kumar , Lenka . *IEEE Signal Processing Letters* May 2006. 13 (5) p. . SOA University ; MITS University (LSB Matching Revisited)

[D’enziz ()] ‘An Incremental Learning Algorithm for Face Recognition’. O D’enziz , M . *Biometric Authentication*, 2002. Springer. 2359. (Castrill’on and others)

[Goel (2008)] ‘Data Hiding in Digital Images: A Steganographic’. Piyush Goel . [http://cse.iitkgp.ac.in/~abhij/facad/03UG/Report/03CS3003_Piyush_Indian Institute of Technology Kharagpur](http://cse.iitkgp.ac.in/~abhij/facad/03UG/Report/03CS3003_Piyush_Indian%20Institute%20of%20Technology%20Kharagpur) May. 2008.

[Yang et al. (2002)] ‘Detecting faces in images: A survey’. M.-H Yang , D J Kriegman , N Ahuja . *IEEE Trans. Pattern Anal. Mach. Intell* Jan. 2002. 24 (1) p. .

[Xie et al. ()] ‘Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform’. X Xie , J Lai , Others . *Pattern Recognition* 2010. Elsevier. 43.

[Fridrich (2004)] ‘Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes’. J Fridrich . *Proc. 6th Int. Workshop on Information Hiding*, (6th Int. Workshop on Information Hiding Toronto, Canada) 23-25 May 2004. p. .

[S. Z. Li A. K. Jain (ed.) ()] *Handbook of Face Recognition*, S. Z. Li & A. K. Jain (ed.) 2011. Springer.

[Li and Chu (2007)] ‘Illumination Invariant Face Recognition Using Near-Infrared Images’. S-Z Li , R-F Chu . *IEEE Transactions on Pattern Analysis And Machine Intelligence* April 2007. 29 (4) .

[Chandramouli et al. (2003)] ‘Image Steganography and Steganalysis: Concepts and Practices’. R Chandramouli , M Kharrazi , N Memon . *Proc. 2nd Int. Workshop on Digital Watermarking*, (2nd Int. Workshop on Digital Watermarking Seoul, Korea) 20-22 Oct. 2003. p. .

[Kim et al. ()] ‘Incremental Linear Discriminant Analysis Using Sufficient Spanning Sets and its Applications’. T.-K Kim , ? B Stenger , ? J Kittler , R Cipolla . *International Journal Computer Vision* 2010. Springer.

[Daneshkhah et al. (2011)] ‘Incremental Training of a Detector Using Online Sparse Eigende composition’. Ali Daneshkhah , ; S Hassan Aghaeinia , C Paisitkriangkrai , J Shen , Zhang . *IEEE, Transactions on Image Processing*, 2011. January 2011. 20. Amirkabir University of Technology (AUT) Tehran, Iran IEEE (A More Secure Steganography Method in Spatial Domain)

[Pevny and Fridrich (2007)] ‘Merging Markov and DCT features for multi-class JPEG steganalysis’. T Pevny , J Fridrich . *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, (SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX San Jose, CA) Jan 2007. 6505 p. .

[Wu (2011)] *Offline and Online Adaboost for Detecting Anatomic Structures*, H Wu . August 2011. Arizona State University (MS Thesis)

[Dumitrescu et al. (2002)] ‘On steganalysis of random lsb embedding in continuous-tone images’. S Dumitrescu , X Wu , N Memon . *Proc. IEEE International Conference on Image Processing*, (IEEE International Conference on Image essing Rochester, New York) September 2002.

[Grabner and Bischof ()] ‘On-line Boosting and Vision’. H Grabner , H Bischof . *Austrian Joint Research Project Cognitive Vision under projects S9103-N04*, 2004.

[Fridrich et al. (2001)] ‘Reliable Detection of LSB steganography in Color and Grayscale Images’. J Fridrich , M Goljan , R Dui . *Proc. ACM Workshop on Multimedia and Security*, (ACM Workshop on Multimedia and Security Ottawa, CA) 5 th Oct. 2001. p. .

[Viola and Jones ()] ‘Robust real-time face detection’. P Viola , M J Jones . *Int. J. Comput. Vis* 2004. 57 (2) p. .

[Zhang and Wang (2005)] ‘Steganography using multiple-base notational system and human vision sensitivity’. X Zhang , S Wang . *IEEE Signal Processing Letters* Jan. 2005. 12 (1) p. .

[Xing and Li ()] *Study on Wavelet Transformation-based Low Illumination & High Dirt Face Detection Algorithm*, S Xing , H , X-G Li . 2011. IEEE.

[Wagner et al. (2012)] ‘Toward a Practical Face Recognition System: Robust Alignment and Illumination by Sparse Representation’. J Wagner , A Wright , Z Ganesh , H Zhou , Yi Mobahi , Ma . *IEEE Transactions on Pattern Analysis And Machine Intelligence* February, 2012. 34 (2) .

[Zhao and Chellappa ()] W Zhao , R Chellappa . *Face Processing: Advanced Modeling and Methods*, 2006. Elsevier.