An Analysis of LSB Image Steganography Techniques in Spatial Domain

Mamta Juneja, and Parvinder S. Sandhu

Abstract—(Markus Kahn, 1995) defines Steganography as art and science of communicating in a way which hides the existence of the communication. In contrast to cryptography, where the enemy is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a cryptosystem, the goal of steganography is to hide messages inside other innocent messages in a way that does not allow any enemy to even detect that there is a second message present. In this paper we would review various steganography techniques LSB substitution, LSB matching, Adaptive LSB and Pixel-Value Differencing (PVD), Edge detection filter based, Pixel Indicator techniques Component based LSB and Texture based techniques. The review concludes with the need of better steganography technique for color images meeting criteria’s of imperceptibility, robustness and capacity equally.

Keywords—Adaptive LSB, Component based LSB, Edge detection filter based Techniques, LSB Matching, LSB Substitution, Pixel Indicator Techniques, Pixel Value Differencing, Steganography.

I. REVIEW OF LSB IMAGE STEGANOGRAPHY TECHNIQUES IN SPATIAL DOMAIN

Neil F. Johnson and Sushil Jajodia, 2000. Discuss three popular methods for message concealment in digital images. These methods are LSB insertion, masking and filtering and algorithmic transformations.

(Kevin Curran and Karen Bailey, 2003) analyze seven different image based steganography methods. These methods are Stego1bit, Stego2bits, Stego3bits, Stego4bits, Stego Colour Cycle, StegoPRNG, and StegoFridrich.

Generally, Steganography techniques are classified into LSB substitution, LSB matching, Adaptive LSB and Pixel-Value Differencing (PVD), Edge detection filter based, Component based LSB and Texture based per the survey conducted for this paper.

(Chang et al., 2002)(Thien et al., 2003) proposes that LSB substitution is the most commonly used method directly replacing the LSBs of pixels in the cover image with secret bits to get the stego-image. LSB substitution algorithm is the simplest scheme to hide message in a host image. It replaces the least significant bit (LSB) of each pixel with the encrypted message bit stream. Authenticated receivers can extract the message by deciphering the LSB of every pixel of the host image with a pre-shared key. Since only the least significant bit of pixels is altered, it is visually imperceptible by human. The capacity of the algorithm is 1 bit per pixel.

(Wang et al., 2000 and 2001) had worked on to improve the perceptual quality of the stego-image and employed a genetic algorithm to generate a substitution table. According to this substitution table, the value of the secret data to be embedded into each host pixel is transformed to another value in advance which is closer to the original value of the host pixel; however, owing to the nature of a genetic algorithm, although the substitution table is good, it may not be the optimal solution. In order to obtain the optimal solution, (Chan et al., 2001 and 2004), genetic algorithm is not required. An Optimal Pixel Adjustment Process (OPAP) is used to improve efficiency and enhance the visual quality of the stego-image generated by simple LSB substitution.

The LSB matching scheme was introduced by (Ker et al., 2004). LSB matching also modifies the LSBS of the cover image for data hiding, but it does not simply replace the LSBS of the cover image as LSB replacement does. On the other hand, if one secret bit does not match the LSB of the cover image, then another one will be randomly added or subtracted from the cover pixel value. A revised version of LSB matching is proposed by (Mielikainen, 2006) which greatly improve it by lowering the expected number of modifications per pixel (ENMPP), from 0.5 to 0.375. Therefore, the histogram affected by the scheme is less significant. Generalization of LSB matching (G-LSB-M), is presented by (LI, 2009) in which sum and difference covering set of finite cyclic group were used to further reduce ENMPP and providing better security.

But still these techniques had various problems like these are limited mainly by artificial noises in the smooth regions of the image. Artificial noises badly damage the visual quality of the stego-image. These LSB techniques replace the same length bits of each original pixel with the embedding data. However, not all pixels in the image can tolerate equal amounts of changes without noticeable distortion. Therefore, the stego-image has low quality when equally changing LSBS of all pixels. Research found out that, if an image is processed with simple LSB substitution the histogram of the image will be showed in a “pair-wise” manner. These pair-wise blocks are known as Pairs of Values (PoV) which can be identified by Chi-square Test given by(Westfled et al.,1999),(Provos et al.,2002),(Stanley,2005). All LSB matching techniques were

Mamta Juneja is Assistant Professor in University Institute of Engineering and technology, Panjab Univesity, Chandigarh, e-mail: er_mamta@yahoo. Com.

Dr. Parvinder S. Sandhu is working as Professor with Rayat and Bahara Institute of Engineering and Bio-Technology, Sahauran, Punjab.
successfully attacked by best-known detector for LSB matching which is based on the center of mass (COM) of the histogram characteristic function (HCF) discussed by (Ker, 2005).

To solve this issue, LSB based methods employed Human Visual System (HVS) masking characteristics to embed the secret data into the variable sizes of LSBs of each pixel called Adaptive LSB. (Lie et al., 2000) created a piecewise mapping function according to the HVS contrast sensitivity to determine the adaptive numbers of LSBs for data hiding. (Lee et al., 2000) exploited the contrast and luminance property of HVS and achieved a variable-sized LSB insertion. In research carried by (Liu et al., 2004), each pixel of the original image is grouped according to its intensity, then the frequency of the original pixel in each group is counted, and a bit plane wise data hiding method is used to embed the secret message into the original image by the principle of the pixel with high frequency priority. Similarly, (Kekre et al., 2008) determined the embedded capacity of each pixel by considering the luminance from the highest bits residual image. But these improved LSB schemes do not fully exploit the HVS masking characteristics; especially the Edge masking effect and they cannot obtain good imperceptibility.

PVD (Pixel value differencing) method was introduced by (Wu and Tsai, 2003) to further improve the quality of the stego-image, which utilized the HVS sensitivity to intensity variations from smoothness to high contrast by the selection of the width of the range which the difference value of two neighbor pixels belongs to. (Chang and Tseng, 2004) proposed Side Match method and (Park et al., 2005) gave Neighborhood pixel information (NPI) method which checked more than two neighborhood pixels to determine the payload of each pixel; however, the embedding capacity of these methods is far less than that of PVD methods. By combining the LSB insertion and PVD methods, (Wu et al., 2005) proposed a data hiding scheme with a better image quality by using PVD methods alone. In their approach, two consecutive pixels are embedded by the LSB replacement method if their difference value falls into a lower level; similarly, the PVD method is used if the difference value falls into a higher level. In other words, the secret data is hidden into the smooth areas by LSB substitution and PVD methods in the edge areas. Furthermore, (Yang et al., 2006) gave Multi pixel differencing (MPD) and (Jung et al., 2008) combined MPD with LSB to estimate smoothness of each pixel. Generalization of PVD method was provided by (Liu et al., 2008). (Wang et al., 2008) exploited modulus function with PVD to rectify its falling off boundary problem and resistant to RS-analysis attack.

(Yang et al., 2008), proposed an adaptive LSB steganographic method using PVD and LSB replacement. In their scheme, the difference value of two consecutive pixels is used to estimate the hiding capacity into the two pixels. Pixels located in the edge areas are embedded by a k-bit LSB substitution method with a greater value of k than that of the pixels located in smooth areas. The scheme embeds more secret data into edged areas than smooth areas in the host image. But even this has various problems like Falling off the boundary problem for edge pixels, embedding data in the whole image even at low embedding rate, Poor in resistance to statistical attacks. These were easily attacked by (Zhang et al., 2004), by change in histogram. All of the evaluation criteria’s (quality, capacity, security and complexity of data) embedding were not met. Some more media rather than audio, video and media should also be tried. These methods follow the principle that the edge areas can tolerate more changes than smooth areas. However, this principle obeyed by some existing data hiding schemes does not discriminate texture features from edge ones; the edge areas used by these schemes contain both edges and textures.

(Luo et al., 2010) worked on the problem of uniform embedding at all parts of an image irrespective of size of secret message and proposed LSB matching revisited. This edge adaptive scheme can select the embedding regions according to the size of secret message and the difference between two consecutive pixels in the cover image. For lower embedding rates, only sharper edge regions are used while keeping the other smoother regions as they are. When the embedding rate increases, more edge regions can be released adaptively for data hiding by adjusting just a few parameters. The experimental results evaluated on 6000 natural images with three specific and four universal steganalysis algorithms show that the new scheme can enhance the security significantly.

(Maleki et al., 2011) worked on falling of boundary problem and security issues and provided an adaptive data hiding method based on four-pixel differencing combined with modulus function. The average differencing value of a four-pixel block via a threshold secret key determines whether current block is located in edge or smooth areas. Pixels in the edge areas are embedded by Q-bit of secret data with a larger value of Q than that of pixels placed in smooth areas and provides five secret keys to protect embedded secret data and problem of overflow or underflow does not occur.

(Joo et al., 2011) proposed Adaptive Steganographic Method Using the Floor Function and Modulus function with Practical Message to provide better resistance to attacks.

(Liao et al., 2011) proposed four-pixel differencing and modified LSB substitution to improve quality. Secret data are hidden into each pixel by the k-bit modified LSB substitution method, where k is decided by the average difference value of a four-pixel block. Readjustment has been executed to extract the secret data exactly and to minimize the perceptual distortion but it compromised resistance to attacks for achieving quality.

(Mandal et al., 2011) proposed DHPVD in which more number of secret bits are inserted to the edge areas than smooth areas to improve capacity. 2x2 non overlapping mask is chosen from the source image in row major order. The difference among two consecutive may fall into any one of four levels such as lower, middle1, middle2 and higher. Then depending upon the difference level, variable numbers of secret bits are embedded using a hash function in the consecutive two pixels in the non overlapping 2x2 mask. In addition to embedding the contents of the hidden image, dimension of the hidden image has also been embedded. A bit handling is used to minimize the difference between the source and embedded pixels but not resistant to attacks.

(Kumar et al., 2012) integrates Tri way pixel value differencing and LSB matching revisited for executable file as secret data as a means to show that the steganography can also work with other medias than that of image, audio and video.
The above stated techniques worked on PVD which lacks to discriminate texture features from edges, time consuming, complex and can be attacked. 

(Alwan et al., 2005) introduced Edge detection filtering based approach to overcome this problem. They used Sobel mask filter for embedding data in images using LSB, gray level connectivity using a fuzzy approach and the ASCII code. Further (Negi et al., 2006) proposed adaptive steganography based on filtering approach using both global and local image features and (Hempstalk, 2006) introduced two new techniques FilterFirst and BattleSteg against more traditional image steganography techniques, BlindHide and HideSeek. Results show that FilterFirst beat all the steganalysis techniques until embedding rates became greater than 7% and performed better than all other steganography algorithms tested. Also, Features of the cover, such as edges, are better way of hiding information. 

(Singh et al., 2007) encrypted messages in nonadjacent and random pixel locations in edges of images on gray images to avoid sequential attacks of all above mentioned techniques. 

(Chen et al., 2010) proposed a novel steganography scheme which is based on the LSB steganography mechanism and employs a hybrid edge detector which combines the fuzzy edge detector with the canny edge detector. The hybrid edge detector assists the new scheme in generating a better quality stego image with high embedding rate of 2.86bpp with some resistance to statistical attacks.

(Hussain, 2011) worked on data hiding method around the edge boundary of an object by varying threshold value of filter. The experimental results show very high rate of PSNR but proposed scheme is targeted for low rate of hidden data capacity. 

(Bassil et al., 2011) proposed a simulation tool GhostBit based on Parameterized Canny edge detection algorithm to provide high resistance to Steganalysis attack. The parameters used are size of the Gaussian filter, a low threshold value, and a high threshold value. These parameters can yield to different outputs for the same input image and secret data. As a result, discovering the inner-workings of the algorithm would be considerably ambiguous, misleading steganalysts from the exact location of the covert data.

The above techniques focused on gray images but the advancement in image technology to RGB leads to steganography application for color images. 

(Gutub et al., 2008, 2009) introduced Pixel Indicator Techniques. They merged random pixel manipulation method and the stegokey to propose a technique, which uses the least two significant bits of one of the channels to indicate existence of data in the other two channels. He further targets high capacity in RGB image based steganography by introducing the concept of storing variable number of bits in each channel (R, G or B) of pixel based on the actual color values of that pixel i.e. lower color component stores higher number of bits. He also proposed Triple-A concealment technique method to hide digital data inside image-based medium. The algorithm adds more randomization by using two different seeds generated from a user-chosen key in order to select the component(s) used to hide the secret bits as well as the number of the bits used inside the RGB image component. The randomization adds more security especially if an active encryption technique is used such as AES. The capacity ratio is increased above SCC and pixel indicator scheme. Triple-A has a capacity ratio of 14% and can be increased if more number of bit is used inside the component(s).

(Gandharba et al., 2012) proposed an approach to RGB channel based steganography technique which uses RSA algorithm for encryption and decryption. In an RGB image, each pixel (24 bits) is having R channel of 8 bits, G channel of 8 bits and B channel of 8 bits. The image is divided into 8 blocks and the cipher text is divided into 8 blocks. One cipher block is allocated to be embedded in only one image block by a user defined sub key. Out of the three channels in each pixel of the image one is used as the indicator channel. The indicator channel for the different blocks is not the same. The other two channels (called data channels) are used for hiding cipher text bits in 4 least significant bit (LSB) locations. In a data channel 4 bits of cipher text can be embedded if after embedding the change in pixel value is less than or equal to 7. The two LSBs of indicator will tell whether the cipher text is embedded in only one data channel or in both data channels, so that retrieving can be done accordingly at the receiver.

But pixel indicator techniques had a drawback that they treated all Red, Green, Blue components equally but in actual the contribution of all Red, Green, Blue components is not same for visual perception. So component based approaches were introduced.

(Imran et al., 2007) proposed Component based Steganography for Color images. They integrated NPI, modified Least Significant Bits technique for data embedding and uses the green component of the image as it is less sensitive to human eye and thus it is totally impossible for human eye to predict whether the image is encrypted or not.

(Chang et al., 2008) proposed a large payload data embedding method for color images. The proposed method modifies the blue value of the color pixel in order to imply the secret data because the blue value is an insensitive color to human eyes. Furthermore, the number of secret bits that can be embedded into a cover pixel is dynamic and can be applied to both RGB and YUV color systems.

(Roque et al., 2009), presents a novel Steganography algorithm based on the spatial domain: Selected Least Significant Bits (SLSB). It works with the least significant bits of one of the pixel color components in the image and changes them according to the message’s bits to hide. The rest of bits in the pixel color component selected are also changed in order get the nearest color to the original one in the scale of colors. This new method has been compared with others that work in the spatial domain and the great difference is the fact that the LSBs bits of every pixel color component are not used to embed the message, just those from pixel color component selected.

The pixel of an color image consists of red, green and blue component and each of the component ranges from 0 to 255, in case of 24-bit representation, (Mandal et al., 2012) proposed pixel value differenting (PVD) method for secret data embedding in each of the component of a pixel in a color image while eliminating overflow (exceed 0-255) problem. Further security is provided by using different number of bits in different pixel components.
(Moro et al, 2007) presented a steganographic algorithm called ConDith, ConDithSpread and Context that selects pixels to embedding information from non homogeneous texture regions.

(Hamid et al, 2009) proposed LSB steganography approach based on texture analysis i.e. dividing the image into simple and complex textures and hiding more data on complex textures than simple one. They proposed two approaches from LSB algorithm; the 3-3-2 approach without any limitations on the type of images being used and can reach up to 33.3% of size of hidden data, and the second one is the 4-4-4 limitations on the type of images, the new approach features will increase the data hidden in the image by merge the above approaches. The major Drawback of these texture based algorithms is that after the selection of pixels, they use LSB to insert the message and if we apply any filter in the stego-image the message is lost.

II. CONCLUSION ANALYSIS

Initial work on LSB steganography was on LSB Substitution and was explored by (Chang et al., 2002) (THIEN et al., 2003) (Wang et al., 2000, 2001) (Chang et al., 2003, 2006) (Chen et al., 2001, 2004).which replaces the same length bits of each original pixel with the embedding data so were easily attacked by Chi-square Test given by (Westfled et al., 1999), (Provost et al., 2002) and (Stanley, 2005).

LSB Matching introduced by (Ker et al., 2004) and researched by (Mielikainen, 2006) (LI, 2009) (Luo et al., 2010) (Kumar et al., 2012) was attacked by (Ker, 2005) based on the center of mass (COM) of the histogram characteristic function (HCF).

Adaptive LSB was worked on by (Liu et al., 2000) (Liu et al., 2004) (Kekre et al., 2008) is based on variable number bits substitution but doesn’t fully exploit the HVS masking characteristics; especially the Edge masking effect and they cannot obtain good imperceptibility.

PVD methods are among the most popular method which was explored by (Wu and Tsai, 2003) (Park et al., 2005) (Wu et al, 2005) (Yang et al, 2006) (Jung et al., 2008) (Liu et al., 2008). (Wang et al., 2008) (Yang et al, 2008) (Maleki et al., 2011) (Liao et al., 2011) (Mandal et al., 2011).It follows the principle that the edge areas can tolerate more changes than smooth areas. However, this principle obeyed by some existing data hiding schemes does not discriminate texture features from edge ones; the edge areas used by these schemes contain both edges and textures. Moreover were complex to work and were easily attacked by (Zhang et al., 2004).

Edge detection Filter based technique was utilized by (Alwan et al., 2005) (Negi et al, 2006) (Hempstalk, 2006) , (Singh et al, 2007) , (Chen et al, 2010), (Hussain, 2011) (Bassil et al., 2011) for steganography in Gray images. But the advancement in image technology to RGB leads to steganography application for color images Pixel indicator techniques introduced by (Gutub et al, 2008, 2009) (Gandharba et al, 2011) for color images had a major drawback of treating all color components (red, green, blue) equally contradicting Hecht principle, which reveals that the visual perception of intensely red objects is highest and than of intensely Green objects and is least for intensely blue objects i.e. red plays the most significant and Blue plays a least significant role in color formulation. So, we can integrate maximum changes in Blue component and average changes in green component and least change in red component without making much difference in color image.

Color component based techniques researched by (Imran et al., 2007) (Chang et al., 2008) (Roque et al, 2009) (Mandal et al., 2012) is not exploited fully for all types of attacks like targeted and universal and were focused on single component. They didn’t utilize cryptography and Random Sequence generator techniques to be better resistant to attacks.

(Chen et al., 2010) proposed steganography technique using hybrid filter but tested it for gray images moreover didn’t test it for targeted and universal attacks. Capacity of 2.8 bpp and highest PSNR value attained is 28.6 which is very low. Like wise (Mandal, 2011) achieved 49 % PSNR but didn’t even mention capacity factor and (Hussain et al., 2011) achieved highest PSNR for very small text messages. (Mandal et al., 2011) achieved 39% PSNR with good capacity but also explicitly mentioned this in their research paper that they targeted quality and capacity from tradeoff between capacity, quality and robustness. They compromised resistance to attacks for quality and capacity. (Kekre et al., 2009), (Husain et al, 2010) also worked on quality and capacity but successfully attacked by (Singh et al., 2012).

The three most required evaluation criteria’s for any good steganography techniques are Robustness, Imperceptibility and Capacity. But there is no technique so far for color images which would target all these criteria’s fully. So there is urgent requirement of technique which would provide good capacity and high PSNR value and resistant to all targeted as well as universal steganalysis attacks.

REFERENCES


