International Journal of Electronics Communication and Computer Engineering
Volume 10, Issue 4, ISSN (Online): 2249–071X

A High-Capacity Data Hiding Algorithm Based on Switching and Matrix Encoding Techniques

Sunita Waykole[1*], Deepa Indrawal[2] and Dr. Archana Sharma[3]
[1]Research Scholar, Mewar University, India.
[2]Research Scholar, Mewar University, India.
[3]Prof. TIT Bhopal.
[*]Corresponding author email id: sunitawaykole@gmail.com

Date of publication (dd/mm/yyyy): 09/08/2019

Abstract **– In this paper, we present a novel high-capacity data hiding algorithm that would address the need of secured communication over un-encrypted communication channels for digital images. The proposed framework investigates the spatial redundancy within the images and embeds the information using matrix based adaptive switching data hiding. The proposed system could be easily extended to other digital media (such as audio & video). In brief the proposed system incorporates localized and inherent image features based on redundancy of range to determine the capacity of each pixel within the neighborhood. Matrix encoding is employed over the pixels in a specific manner that embedding capacity of each pixel is attained. In addition, the proposed algorithm is blind data hiding technique i.e. the receiver retrieves the stego information without any reference and information regarding the original cover image. The simulation results would reveal that the visual artifacts of stego image are conserved with respect to the cover image. In addition, the proposed system offers significant immunity to common statistical attacks as statistics of original and stego image are nearly similar.**

Keywords **– Switching Technique, Matrix Encoding, Steganographical, Data Hiding.**

I. INTRODUCTION

In the current digital era and evolving awareness of digital empowerment of individuals enhanced the use and transmission of digital information significantly. "Digital India" program wherein significant portion of government records, personnel information, and financial details of Indians would available online [1]. Its main aim was to enhance the electronic literacy among the people so that they could acknowledge the concern government benefits which in turn reduce corruption. Unfortunately, every person's financial, medical, social and criminal history would be online and digitally available for individuals of both nature (i.e. good and bad). Henceforth, the demand for the data security has been growing exponentially now-a-days. The researchers view that data security is a hot topic and ever growing topic with numerous fields. It is a well known fact that the right information about an organization at the right time could help a group (or individual) to support or destroy a designated organization in consideration.

Data hiding has been incorporated into several major applications that stores and/or transmits secure information [1] for all the core sectors. Data hiding is a science that deals with the embedding of the information within the digital media while conserving the integrity of the cover. A perfect data hiding system is not only flexible in the security mechanism, but also has high overall embedding efficiency. With several possible channels and wide range of applications associated with data hiding can be classified into two major areas i.e.

1. Digital Steganographical.
2. Digital Watermarking.

Both these approaches have been evolving into active research areas within the well-defined constraints and differences between them. The purpose and objective of each of the technique differ drastically from one another but concept (i.e. hiding information) remains the same. The purpose of employing digital watermarking techniques is generally related to authentication and annotation based application. Therefore, these algorithms focus on hiding low amount of data that could be exploited for ownership and intellectual rights of the cover media. It is essential that these techniques are robust to noise, compression and other image processing attacks.

On the other hand, steganography is employed as a distraction technique to foil any attempt to detect the communication mechanism. It is imperative that these techniques require high capacity while integrity of the digital media is not considered vital. But immunity against stego-detection is essential for recognizing the algorithm suitable for covert communication between authorized parties. In this paper, we focus on developing an adaptive and high capacity steganographic algorithm based on the matrix encoding. The simulation results of the proposed technique would reveal that the visual artifacts of stego image are conserved with respect to the cover image. In addition, the proposed system offers significant immunity to common statistical attacks as statistics of original and stego image are nearly similar and suitable for secure communication.

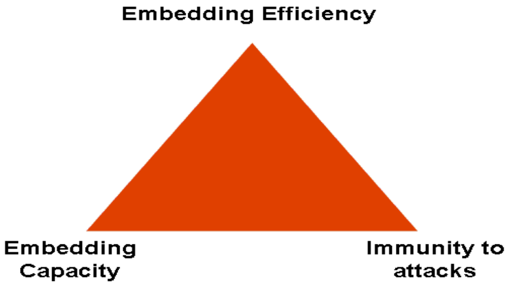

Fig. 1. Various issues on steganographic algorithm.

At the initial phases of digital steganography most of the research is limited to non-blind techniques wherein the cover media is essential for recovering embedded information. In the last decade, blind-steganography techniques have evolved into a major research area. It is a commonly accepted fact that a blind-steganographic technique should address the following issues that are presented in figure 1. i.e.

1. Embedding Capacity: the capacity of the cover image should be utilized to maximum extent without significant distortions.
2. Immunity to attacks: the technique should be robust against various stego-detection algorithms as statistical and visual distortions should be minimal.

It is commonly exploited technique by the several algorithms that they focus on the least significant bit-plane (LSB) for hiding information. Since it is commonly accepted fact that this plane has little impact on the cover images statistical and visual features thus enhancing security to embedded information. These algorithms assume that all the regions of the image are homogeneous but the regions are in fact heterogeneous. Thus, even though they offer significant embedding capacity their immunity against simple statistical is minimal. Hence, we need an approach that could exploit these heterogeneous areas and design an algorithm that employ inherent image features as a basis of region selection framework to overcome the above discussed problems.

Henceforth, it is important topic to understand the ever evolving topic of pixel/region selection in the scientific field of adaptive data hiding. R.C. Cherukuri et.al [9] presented an algorithm wherein the embeddable regions are classified into smooth and noise regions based on PN-sequences. This technique exploits the noise regions for hiding the information in a secured manner thus making the algorithm immune to simple statistical attacks. N. Noda et.al [10] presented a steganographic technique that classifies the least significant bit-plane into "noise" or "informative" sub-blocks based on the complexity measure. The secure information bits are hidden within the noisy blocks of LSB by manipulating the values based on a embedding function while the rest of the blocks are employed for adjusting the complexity measure so as to ensure successful recovery of hidden information bits. Based on the simulation results, it is evident these algorithms offer improved immunity while the embedding capacity of cover image has been reduced in comparison with primitive steganographic approaches i.e. Stools, WbStego and, Secur Engine [6-8].

Which evolving stego-detection algorithm the complexity associated with steganographical algorithms has also been increasing. Even though, the driving force of steganography was to act as an alternative covert communication channel wherein the means of communication is invisible while in encryption means of communication is unreadable. In fact, the presence of communication could be detected but cannot be decoded in case of steganography making it viable solution in the digital era.

It is well known fact that the embedding capacity is directly proportional to the number of bit-planes employed for embedding the information. Hence, to design a high capacity embedding systems it is vital to embed in each pixel to its full capacity while preserving the visual and statistical artifacts of the cover image and enhancing the immunity of the embedding algorithm. E. Franz [11] presented a complex approach for embedding secure data by concentrating on maintaining the visual and order statistics of the cover image with stego data. This algorithm exploits the histogram equalization and manipulation for modeling the embedding function as a Markov source and altering the cover image bits with sensitive data so that it resembles similar statistical distribution that of the cover image. D.C Wu et.al [12] presented a novel approach wherein the difference between the pixel-values of corresponding non-overlapping blocks is exploited for hiding the information in a secured manner. S. Agaian et.al. [13] presented a complexity measure that could classify the image into various regions designated as embeddable and non-embeddable. It also evaluated the embedded capacity of each embeddable region based on the complexity measure. It is evident that this algorithm provided a new concept to determine the embedding capacity of a particular image sub-block and exploit that information to hide maximum information.

It was clear that the focus of steganographic research has been shifted from hiding information to selection of regions. This adaptive selection of regions has evolved in the recent years wherein different concepts were employed to select the noise regions within the image to hide information. This process has two limitations namely.

1. Since the selected regions for embedding were limited that capacity has been reduced drastically.
2. With a prime focus on preserving the statistics that existing methods were immune to some statistical but were vulnerable to other attacks.

These issues added to a new variant within adaptive steganography that focused on selecting the pixels rather than the regions. C. Cherukuri et.al. [14] presented a novel adaptive pixel based data hiding approach that selects the embeddable pixels based t-order statistics. The proposed approach exploited the inherent features of images such as edges, noisy regions, textures and others to suggest a pixel is embeddable or not. This algorithm also incorporated the complex features to preserve the order statistics which enhanced the immunity against various attacks. The pixels selected by the proposed algorithm were able to withstand changes of greater magnitude.

The prime motive of this paper was to evaluate a process that could select the pixels whose embedding capacity can be reached. Since most of the adaptive techniques failed to attain the maximum embeddable capacity of the cover image. Another important aspect that a researcher does not focus on the embedding efficiency of the approach as it can boast capacity by incorporating the non-embeddable pixels/regions to embed multiple bits. These are the challenge's that were addressed and became instrumental in developing this novel concept. The proposed system makes steganography a viable covert communication channel that improves immunity while enhancing the capacity with a new variant of embedding efficiency.

II. Pixel Variation Measure (PVM)

I. The basic step in developing any new steganographic algorithm lies in the manner a pixel is selected for embedding sensitive information. Based on the analysis is was found that the non-embeddable (smooth) regions often contain pixels that could be employed for hiding similarly not all pixels of embeddable region are suitable. This concept laid the foundation for pixel based adaptive steganographic techniques. These algorithms would perform a pre-processing step that conducts a detail image analysis over a 3-by-3 over lapping image sub-blocks to determine the suitable pixels. Many image processing based techniques are employed such as filtering based approaches, pattern recognition concepts and other t-order statistics [14]. The fundamental principle employed by the researchers to foil the steganographic attacks is to alter the natural noise of cover image to hide the data. There are several reasons, why natural noise exists within the original cover images it could due to anyone of the conditions i.e. hardware issues, lighting conditions, improper handling of the instrument and/or changes in the background.

Considering the above fact, we need to utilize this principle while hiding the information in noisy regions in a secured manner. It is essential that even after the embedding the change in magnitude should be within the scope of the noisy regions. Furthermore, the pixels in these regions have high variation range within the given neighborhood. In addition, a close inspection reveals that the embedding along the vertical and horizontal edges could induce visible distortions and blurring effect while diagonal edges preserve the artifacts from visible distortion. Active research was concentrated to address this issue in systematic process that offers improved capacity and limits statistical changes in the cover image.

In this paper, we determine the PVM of each pixel in an image sub-block as it plays a critical role for hiding the data. The pixel within the image are classified into various classes based on the neighborhood variation range a) high variation pixel, b) vertical medium variation pixel, c) horizontal medium variation pixel and d) low variation pixel. The distance measure that is commonly employed in selection of threshold pixel is used as basis which is presented below:

$$\left[dist(F)\right]_{i,j} = \min_{\forall x,y}\left\{\sqrt{|i-x|^2+|j-y|^2}\,[F]_{i,j} \neq [F]_{x,y}\right\}$$

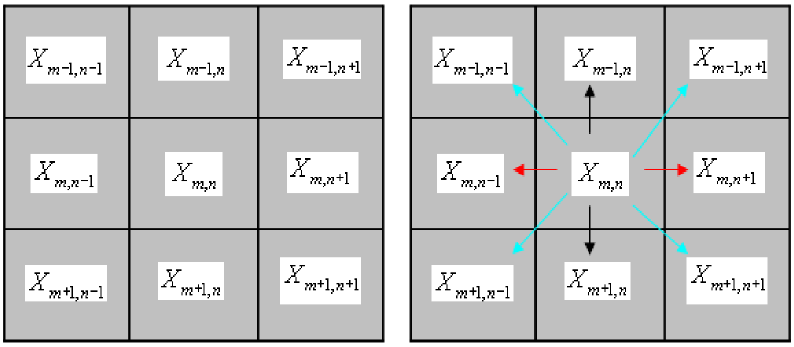

Fig. 2. Distance measure of center pixel using PVM

The Pixel Variation Measure (PVM) is utilized for selecting the pixels that would be suitable for embedding sensitive data bits in a secured manner. Furthermore, the proposed measure also determines the capacity of each selected pixel. We evaluate the PVM for each pixels as presented in figure 2 and are classified into three groups based on soft thresholding namely:

- Non-embeddable pixels with an embedding capacity “0”.
- Low-priority pixels with an embedding capacity “1”.
- High-priority pixels with an embedding capacity “2”.

$\hat{X}_{mn} < \propto_1 \; Non - Embeddable\ pixel$

$\propto_1 < \hat{X}_{mn} < \propto_2 \; Low\ priority\ Embeddable\ pixel$

$\propto_2 < \hat{X}_{mn} \; High\ priority\ Embeddable\ pixel$

Where, α_1, α_2 are the variance limiting values that signify the magnitude change in the pixel with each bit hidden. In addition, we introduce the concept of altering the hiding process from one image sub-block to another based on distribution of embeddable pixels. Switching theory is an advantageous process for the applications that make rapid decisions about routing information. In case of signal processing, this theory is devoted to the processing of a particular class of signals that could be modelled by thresholds. In this paper, the threshold logic is estimated based on the PVM of each pixel.

The blocks within the image which are suitable for embedding are determined during the first phase of switching detection using the corresponding threshold function. The basic architecture of the switching detection technique is presented in Figure 3. This framework offers an ability to exploit inherent image analysis such as, localized information in time and variation within the image sub-block. In recent years, the demand for dual measures based transforms defined over the finite field has increased exponentially. Switching concept that is commonly used in telecommunication wherein the routing of the calls would be carried out based on caller input. We employ a similar routing for each sub-block based on the soft-thresholding PMV as the logic controller to improve the embedding efficiency of the entire algorithm. The distribution of the pixel variation measure varies from one pixel to other but the embedding algorithm would treat each pixel on similar lines without acknowledging its individual capacity that several issues arise. Hence, we exploited the switching concept in digital perspective for improving the embedding efficiency of the proposed algorithm while maintaining other constraints.

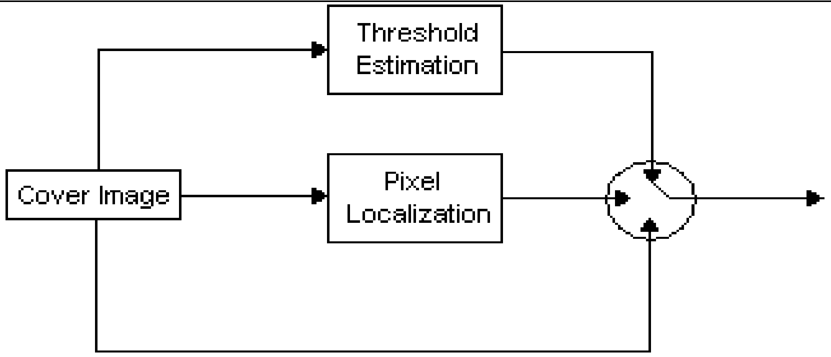

Fig. 3. General structure of the switching detection technique.

III. Proposed Algorithm

In this section, we present the proposed algorithm based on switching theory and matrix encoding that could overcome the limitations of capacity barriers while maintaining first order statistics of the image. The embedding technique is an obscure procedure with various defined constraints based on the PVM and the level of security needed along with the amount of information embedded. It is a straight forward algorithm that offers flexibility over the three realms of steganography i.e. capacity; immunity; and efficiency. Figure 4 provides the general structure of the encoding process.

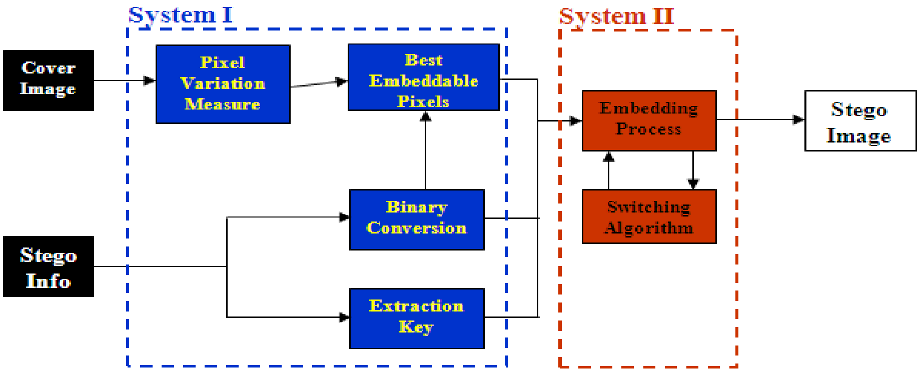

Fig 4. General structure of the embedding technique

The purpose of the threshold derivation process is to ensure an even distribution of stego information throughout the image. This process selects the threshold in a manner in which the safest pixels will be used first in relevance to the size of the stego file to be embedded. To facilitate extraction, information is sequentially embedded from left to right, top to bottom. If the threshold is randomly selected, one may be unsure of the resultant distribution of steganographic data. A very high threshold may not accommodate the amount of space required.

IV. Simulation

For computer simulation, we test the performance of the algorithm over 50 test images of varying sizes and features with different patterns of stego information. The experimental results are presented for some of the image where the cover image is of size 512 by 512 and the embedded information is of size 128 by 128.

Simulation I: Analytical Analysis

Capacity of an Embedding system ($\hat{C}$):

The capacity of an embedding system is defined as the ratio of number bits embedded to the total number embeddable pixels available in a cover image.

$$\hat{C} = \frac{Number\ of\ embeddable\ Pixels}{Number\ of\ Pixels\ in\ the\ cover} * 100$$

Embedding Efficiency of an Embedding system ($\widehat{E}$):

The embedding efficiency of an embedding system is defined as the ratio of number bits embedded to the total number pixels that have been altering for hiding the bits.

$$\hat{E} = \frac{Number\ of\ embeded\ bits}{Number\ of\ Pixels\ altered\ in\ the\ cover}$$

Change Density rate of an Embedding system (CDR):

The change density rate of an embedding system is defined as the ratio of number pixels altered during the embedding process to the total number of pixels in a cover image.

$$\widehat{CDR} = \frac{Number\ of\ altered\ pixels}{Number\ of\ pixels\ in\ the\ cover} * 100$$

In addition, the change density rate (CDR) can also be defined as follows.

$$CDR = \frac{\widehat{C}}{\widehat{E}}$$

In this phase, we test the security barriers associated with the new proposed algorithm. Table 2 gives the analytical analysis results over the embedded information length in terms of bytes. 10% pertains to a message length of 9747. 33% pertains to a message length of 32448 and 50% pertains to a message length of 49152.

Table 2. Illustrates the variations in analytical analysis for various images with varying embedded information.

		$\hat{C}$	$\hat{E}$	$\widehat{CDR}$
Fish	10%	48.352	5.127	9.431
Fish	33%	48.352	4.873	9.922
Fish	50%	48.352	4.321	11.190
Building	10%	38.518	5.023	7.668
Building	33%	38.518	4.734	8.136
Building	50%	38.518	3.913	9.844
Rock	10%	44.806	5.209	8.602
Rock	33%	44.806	4.675	9.584
Rock	50%	44.806	3.935	11.386

From table 2, it is evident that if the possible capacity is higher than the embeddable data then the proposed algorithm ensures the changes are minimal. Based on the proposed algorithm, possible embeddable pixels are limited to noisy, texture and edge pixels of the image hence the capacity available "$\hat{C}$" of the proposed system is fixed. While "$\hat{E}$" is altered with increase in changes with minimum being value 2 that indicates 50% change in the pixels with the increase in value the possible changes per pixel embedded information decreases. Based on the analysis, in this phase it is clear that the proposed algorithm offers enhanced capacity while limiting the changes in pixel to lower number than existing algorithms.

V. Conclusion

In this paper, we introduced a novel high capacity data hiding algorithm that addressed the need of secured communication over un-encrypted communication channels for digital images in an effective manner. The proposed framework incorporated the spatial redundancy within the images and embeds the information using matrix based adaptive switching data hiding. Matrix encoding is employed over the pixels in a specific manner that embedding capacity of each pixel is attained while the embedded information was retrieved lossless from stego media without prior knowledge of original media. In addition, the visible and first order statistics of the image are preserved in a significant manner.

The simulation results revealed that the visual artifacts of image are preserved even after hiding significant amount of the sensitive data. In addition, the first order statistics of the image are also preserved while the system is immune to commonly used steganographic attacks.

References

[1] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing", Prentice Hall Inc., 2002.

[2] Van Schyndel, R.G.; Tirkel, A.Z.; Osborne, C.F.; "**A digital watermark"** Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference Volume 2, 13-16 Nov. 1994 Page(s):86 - 90 vol.2

[3] Tirkel, A.Z.; Osborne, C.F.; Van Schyndel, R.G.; "**Image watermarking a spread spectrum application"** Spread Spectrum Techniques and Applications Proceedings, 1996., IEEE 4th International Symposium, 22-25 Sept. 1996 Page(s):785 - 789 vol.2

[4] Wolfgang, R.B.; Delp, E.J.; "**A watermark for digital images"** Image Processing, 1996 Proceedings, International Conference on Volume 3, 16-19 Sept. 1996 Page(s): 219 - 222.

[5] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Systems. J.,* vol. 35, 1996.

[6] wbStego, 11 March 2003,

[7] Secure Engine, 11 March 2003, URL: http://securengine.isecurelabs.com/

[8] S_Tools, March 2003

[9] Sos. S. Agaian, I. Gurevich, R.C. Cherukuri, E. Metlitski ; "Two New M-sequence Based Data Hiding Algorithms" 7th International Conference on Pattern Recognition and Image Analysis: (MAIK "Nauka/Interperiodica" Publishing, Moscow), 2005, Vol. 15, No. 2.

[10] Michiharu Niimi, Hideki Noda, Eiji Kawaguchi, Richard O. Eason: Luminance Quasi-Preserving Color Quantization for Digital Steganography to Palette-Based Images. ICPR (1) 2002: 252-254

[11] Elke Franz. "Steganography Preserving Statistical Properties." Information Hiding: 5th International Workshop, IH. pp. 287-294. July 2003.

[12] Wu D. C. and Tsai W. H., "Spatial-domain image hiding using image differencing," IEE Proceedings on Vision Image and Signal Processing. Vol. 147, No. 1, PP. 29-37, Feb 2000.

[13] S.S. Agaian, R.C. Cherukuri, S.Ronnie, "A New Secure Adaptive Steganographic Algorithm using Fibonacci Numbers", 2006 IEEE Region 5 Technology and Science conference, San Antonio, USA, April 7-8 2006.

[14] S.S. Agaian, R.R. Sifuentes, R.C. Cherukuri, "T-Order Statistics and Secure Adaptive Steganography", *SPIE Optics & Photonics advance technical program, San Diego, USA, 31July-4August 2005*.