A Novel Algorithm for Hiding Information in Video using Spatial Domain

Hassan H. Soliman, Hossam E. Mostafa and Eman A.E. Ahmed

Department of Electronics and Communications engineering, Faculty of Engineering,
El-Mansoura University, Mansoura, Egypt.

, hossam_moustafa@mans.edu.eg, engineer_em@hotmail.comhhsoliman@hotmail.com

Abstract

Video Steganography is a technique that is used to transmit information by modifying video frames in an imperceptible manner and depending on the weakness of the Human Visual System (HVS) in distinguishing the simple differences between coloured images. This paper is about embedding encrypted information (text and image) which have been encrypted using RSA (Rivest-Shamir-Adleman) technique at sender side using the public key. The video frames will be used as a cover for the hidden data. This embedding has been applied using two methods. The first is considered a parity Least significant Bit (LSB) in the R - image (Red-image) of the video frame and the second method was XORing of LSB for each pixel in the R - image (Red-image) of video frame and the bit next to it. The stego video file (the video that contains the hidden data) will be sent over the data network to the receiver who can extract the secret encrypted message then decrypt it using the private key to get the secret message. The proposed system has been tested using Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) for different noise (salt and pepper) densities have been added as a sort of effective attack that may change the secret message. A filtering stage has been used to treat this attack; the results of each method were compared as well.

Keywords: *Information Hiding, Encryption (RSA), Video Steganography, Human Visual System (HVS)*

1. Introduction

Information security gained a significant importance by the development of the computer and the expansion of its use in different areas of life and work, steganography is the art of concealing data within other digital media such as an image, an audio or video for providing higher security. The security includes both imperceptibility and undetectability. Encryption is introduced for the data security, the commonly used encryption schemes include DES (Data Encryption Standard) and RSA. The Equation in (1) provides a very generic description of the pieces of the steganographic process:

$$\boldsymbol{cover\ medium + hidden\ data + stego\ key = stego\ medium} \quad (1)$$

In this context, the cover medium is the file in which we will hide the data and the stego key must be available to increase the security level, the resultant file is the stego medium. Any steganography technique has to satisfy two basic requirements, the first requirement is perceptual transparency, i.e. the cover object (the object which does not contain any additional data) and the stego object (the object which contains secret message) must be perceptually indiscernible, and

the second constraint is high data rate of the embedded data. Among the methods of Steganography, the most common one is the method, which uses images for applying steganography. Image steganography has been explored extensively with various steganographic schemes. Since nowadays, video files are available everywhere and; today's technology allows the copying and redistribution of video files over the Internet at a very low or almost no cost. Therefore, it is necessary to have methods that confine access to these video files and for its security. Video Steganography is one of the solutions. In Video Steganography, the weakness of the Human Visual System (HVS) is used to hide information in the video. That is, while using digital video frames as cover files the difficulty of the human eye to distinguish colours is taken advantage of. Video Steganography has a wide range of applications such as covert communication, digital watermarking, access control, digital rights management, etc. An effective video steganographic scheme should possess the following three characteristics: Perceptual Transparency, Data Rate (Capacity) and Robustness. These characteristics (requirements) called the magic triangle for data hiding and are contradictory. Many watermarking methods have been proposed for image and video authentication **[1, 2, 3]**.The rest of the paper is organized as follows: **Section 2** explains in brief the Literature survey of video Steganography, **Section 3** explains the proposed methods, **Section 4** gives experimental results and its discussion and **Section 5** concludes the paper and introduces the future vision.

2. Related Work

The existing video steganography techniques can be classified on several criteria given below:

2.1 According to the domain of steganography insertion

2.1.1 Spatial Domain Steganography [2, 4, 5]

Spatial domain techniques embed messages in the intensity of the pixels directly. Least Significant Bit (LSB) is the first most widely used in a spatial domain steganography technique. It embeds the bits of a hidden message in the LSB of the image pixels . But the problem with this technique is that if the image is compressed then the embedded data may be lost. Thus, there is a fear of loss of data that may have sensitive information. LSB has been improved by using a Pseudo Random Number Generator (PRNG) and a secret key in order to have private access to the embedded information .The embedding process starts with deriving a seed for a PRNG from the user password and generating a random walk through the cover frame that makes the steganalysis hard. Another recent improvement based on a random distribution of the message was introduced by M. BaniYounes and A. Jantan . In this method they utilize an encryption key to hide information about horizontal and vertical blocks where the secret message bits are randomly concealed.

2.1.2 Frequency Domain Steganography [6, 7, 8, 9]

In frequency domain, frames are first transformed and then the message is embedded in the transformed image. When the data is embedded in frequency domain, the hidden data reside in more robust areas, spread across the entire image, and provides better resistance against statistical attacks. There are many techniques used to transform frame from the spatial domain to frequency domain. The most common frequency domain methods usually used are Fourier Transform (FT), Short Time Fourier Transform (STFT), and Continuous Wavelet Transform

(CWT), Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT) or a Combination of DCT and DWT. Also Steganograohy was done by mixing more than one transformation **[10]** that was applied by embedding the secret data in the first and second level DWT sub-bands of the host image, followed by the application of DCT on the selected DWT sub-bands. The combination of the two transforms improved the steganography performance considerably when compared to the DWT-Only steganography approach.

2.2 According to the video steganography methods

2.2.1 Considering video as separate images [5, 7, 8].

In this method, each video frame is considered as a separate image, in which secret message is hidden. The main advantage of this method is the possibility of using the algorithms used in image steganography for video, but it requires a large amount of computations. The algorithms supposed in this paper use this model.

2.2.2 Finding new dimensions in video [2, 11].

Videos have potential characteristics and dimensions, in which dimensions, if identified, one can use the special characteristics of the human visual system to hide information in the video. The human eye cannot identify the changes made in the video or one can consider the video as a one-dimensional signal – as the one received by the TV-to hide information in the signal. This method provides powerful watermarks in video, but it requires a great deal of computations.

2.2.3 Using the special characteristics of video saving formats. [12, 13]

Each of these video saving formats, like MPEG, AVI, etc, has their own specific characteristics. For example, some formats benefits from special conversions, where information can be hidden. The main advantage of this method is having simple algorithms that can be used real-time. Unfortunately, the steganography in this method totally depends on the video saving format.

3. Proposed Methods

The first stage in the proposed system has been executed at sender side where the secret message (text or image) will be encrypted using the RSA technique. The encrypted message will be embedded in MJPEG video frames after converting it into binary format. MJPEG video must be divided into frames, each frame is a JPEG (256*256) image. At the receiver side the stego video will be converted into frames and extract the encrypted hidden data, then decrypt it to get the original secret message. **Figure 1** shows the block diagram of a secure steganographic system, input messages can be image or text.

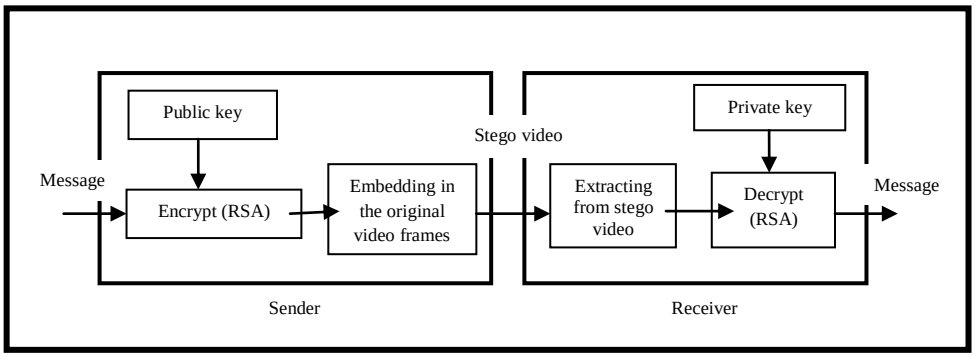

Figure 1. The proposed steganography system.

3.1 Preparing data for embedding

To prepare text for embedding it must be converted to binary pattern. The conversion process of text into a binary pattern is done using ASCII code. In ASCII code every letter, digit, and miscellaneous symbols are represented to unique seven-bit binary codes. After the hidden text is converted to a set of 0's and 1's, the steganography algorithm can be implemented. To prepare the gray image (uint8) that will be embedded in the video frames, each pixel in it must be converted into 8 binary bits then the steganography algorithm can be implemented.

3.2 Steps at sender using Parity LSB Algorithm

a) Encrypt the message (text or image) using the public key
b) Convert the original video into frames, and get the (Red) matrix of each frame.
c) Depending upon the value of the message bit to be embedded (0/1), the LSB of the red value of the current pixel will be modified or unchanged.
d) If the encrypted message bit to be embedded is 0, then the LSB of the red value of the current pixel will be modified or unchanged such that the parity of the red value of the current pixel after embedding of this message bit is even.
e) If the encrypted message bit to be embedded is 1, then the LSB of the red value of the current pixel will be modified or unchanged such that the parity of the red value of the current pixel after embedding of this message bit is odd.
f) The modified R-value is then will be written to the frame
g) Assemble the frames into the stego video.
h) Send via email to the receiver over data network.

3.3 Steps at the receiver using Parity LSB Algorithm:

a) Divide the stego video into frames and get the (Red) matrix for each frame.
b) Convert the pixel to binary format, the result will be 8 bits.
c) Extract the message bit by testing the parity of each value for each pixel in red matrix, if it is even, then the encrypted message bit that will be retrieved is 0.
d) If the parity is odd, then the encrypted message bit will be retrieved is 1.
e) After all encrypted message bits are retrieved; they will be converted to the original format.
f) Decrypt the data using the private key and get the secret message.

3.4 Steps at sender using XORing of LSB's:

a) Encrypt the message (text or image) using public key
b) Convert the original video into frames then get the (Red) matrix of each frame.
c) Convert each pixel into binary format, the result will be 8 bits.
d) Every encrypted secret message bit will be embedded into the LSBs of the cover frame (pixels) after processing.
e) Processing is done as follows: If the message bit to be embedded is 0, then adjust or flip the LSB such that the XORing of LSB and next to LSB will be 0. If the message bit to be embedded is 1, then adjust or flip the LSB such that the XORing of LSB and next to LSB will be 1, **Table 1** illustrates the action for all value probabilities of LSB and bit next to LSB. After embedding all bits of encrypted message, the (R) matrix of the frame will be recovered, and the modified frames that will be got will be called stego frames.
f) Reconstruct the frames into stego video.
g) Send via email to the receiver over data network.

Table 1. Procedure for data embedding Using Xoring Algorithm

LSB	Bit next to LSB	XOR	Action if message bit is 0	Action if message bit is 1
0	0	0	No Change	Flip LSB
0	1	1	Flip LSB	No Change
1	0	1	Flip LSB	No Change
1	1	0	No Change	Flip LSB

3.5 Steps for Data extraction using XORing of LSB's:

a) Convert stego video into frames then read each frame, and get the (Red) matrix of each frame.
b) Convert each pixel in (R) matrix to binary format (8 bit).
c) Retrieve the message bit by XORing the LSB and the bit next to LSB. If the result of XORing is 0, then the message bit is 0. If the result of XORing is 1, then the message bit is 1.
d) After all encrypted message bits are retrieved.
e) Decrypt the encrypted message using the private key to get the secret message.

4. Results and Discussions

4.1 Database description

The Proposed methods were implemented using Matlab2012b. They have been tested on 50 different video files in MJPEG format, the duration of each video was 50 seconds and the frame rate was 10 frames/sec. Three videos have been used to represent the results in **Tables 2** and **3** (SampleVideo1, SampleVideo2 and SampleVideo3). The resolution of each video was 256*256.

4.2 Hidden Data

The secret messages for embedding were text or image files. The text files used for embedding were Text1, Text2 and Text3. Text1 was (hidden data) the size of it was 12 bytes and Text2 is (data hiding is a very interesting field) the size of this text file was 37 bytes and Text3 was (data hiding is a very interesting field we can hide any data) the size of this text file was 58 bytes. The image files used for embedding were gray images (Image1.jpg (1.26Kbyte) 35*37, Image2.jpg (1.9 Kbyte) 45*45 and Image3.jpg (2.44Kbyte) 50*50).

4.3 Tested values

The correctness of the proposed algorithms was measured by calculating two values, which were the mean Peak Signal to Noise Ratio (PSNR) and the mean square error (MSE). These values were compared with the existing values of the related work **[3, 4, 5]**. The proposed algorithms achieved an excellent values comparing with the existing values of the related work.

The PSNR for a video with a number of C frames is calculated as described in (2) **[5].**

$$\mathrm{PSNR}_{\mathrm{video}} = \frac{\sum_{i=1}^{C} \mathrm{PSNR}(i)}{C} \quad \mathrm{dB\ (Decibel)} \qquad (2)$$

Where the PSNR for a frame can be defined as described in (3) **[5]**:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{L^2}{\mathrm{MSE}}\right) \qquad (3)$$

Where, Mean Square Error (MSE) of the stego frame can be calculated as in (4) **[5]**

$$MSE = \frac{1}{N * M}\sum_{i=1}^{N-1}\sum_{j=1}^{M-1}(X(i,j) - Y(i,j))^2 \qquad (4)$$

Where $X(i,j)$ is the cover image that contains N*M pixels and $Y(i,j)$ is the stego image and L is the maximum pixel value of the image, in other words, $L=2^b-1$ where b is the bit depth of the original frame so, in this work L=255 in the case of 8 bits depth.
Matlab program was used to add salt and pepper noise on each frame to measure the robustness of the proposed methods, the function that was used to add salt&pepper noise in Matlab is described as following:
Noisyframe=IMNOISE(I,'salt&pepper',D) this function adds "salt and pepper" noise to the frame I, where D is the noise density. The default for D is 0.05 and the range for this value is between 0 and 1 **[14]**.

In **Tables 2** and **3**, the first three entries in the first column are the cover video clips followed by the secret messages. In the second column and the third column display the PSNR and MSE for different salt&pepper noise densities (0, 0.1 and 0.2) which have been added to the stego video as an intruder effect or active attacker. This type of attacking tries to change the content of the hidden message. To treat this attack or reduce the influence of it, the Median filter will be applied at the receiver side on each frame before applying the extraction processing.

Figure 2 illustrates the Simulink model which has been used to apply this treatment, the results show that, the extracted images were enhanced a lot using the median filter stage but text messages were not enhance. To handle this, an agreement of letter case and preventing from using special character (just use capital character and numbers) must be applied between sender and receiver and the text message embedded more than one times as can as possible this handled the salt &the pepper attack with high percentages. **Table 4** shows sample video1 frames for different noise densities after Image3 has been embedded. **Table 5** shows the extracted messages (Image3 and Text1) for different noise densities before using the filtering stage. **Table 6** shows the same extracted messages for different noise densities after using the filtering stage.

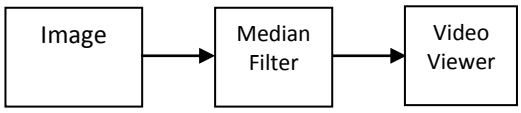

Figure 2. The Simulink model used [7]

Table 2. The PSNR results for LSB parity method with different noise densities

Cover	Secret message	Salt& pepper noise densities					
		0		0.1		0.2	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
Sample video1	Text1 12 byte	68.8	0.01	65.6	0.06	63.2	0.12
	Text2 37 byte	68.4	0.01	65.3	0.07	63	0.13
	Text3 58 byte	68	0.02	65.2	0.07	62.6	0.13
	Image1 35*37	67.2	0.03	65	0.08	62.1	0.14
	Image2 45*45	66.5	0.05	64.2	0.09	61.6	0.17
	Image3 50*50	66	0.06	63.7	0.1	61	0.18
Sample video2	Text1 12 byte	68.4	0.01	65.4	0.07	63	0.13
	Text2 37 byte	68	0.02	65.2	0.07	62.8	0.13
	Text3 58 byte	67.5	0.03	64.8	0.08	62.6	0.13
	Image1 35*37	67	0.04	64.4	0.09	62.2	0.15
	Image2 45*45	66.4	0.05	64	0.1	61.7	0.17
	Image3 50*50	66	0.06	63.6	0.1	61.2	0.18
Sample video3	Text1 12 byte	67.2	0.03	65.8	0.06	63.1	0.12
	Text2 37 byte	67	0.04	65.2	0.07	62.7	0.13
	Text3 58 byte	66.5	0.05	65	0.08	62.2	0.15
	Image1 35*37	65.6	0.06	64.4	0.09	61.8	0.16
	Image2 45*45	65	0.08	64.1	0.09	61.3	0.17
	Image3 50*50	64.2	0.09	63.8	0.1	61	0.18
Average		66.87	0.040	64.70	0.081	62.17	0.148

Table 3. The PSNR results for XORing of LSB's method with different noise densities

Cover	Secret message	Salt& pepper noise densities					
		0		0.1		0.2	
		PSNR	MSE	PSNR	MSE	PSNR	MSE
Sample video1	Text1 12 byte	68.5	0.01	65.4	0.07	63.2	0.12
	Text2 37 byte	68.2	0.01	65.3	0.07	62.9	0.13
	Text3 58 byte	67.8	0.02	65	0.08	62.7	0.13
	Image1 35*37	67.3	0.03	64.6	0.08	62.2	0.15
	Image2 45*45	66.9	0.04	64.2	0.09	61.7	0.16
	Image3 50*50	66.4	0.05	63.7	0.1	61.4	0.17
Sample video2	Text1 12 byte	68.2	0.01	65.4	0.07	63.1	0.12
	Text2 37 byte	67.9	0.02	65.2	0.07	62.9	0.13
	Text3 58 byte	67.5	0.03	65	0.08	62.6	0.13
	Image1 35*37	67	0.04	64.7	0.08	62.1	0.14
	Image2 45*45	66.6	0.04	64.2	0.09	61.7	0.16
	Image3 50*50	66.3	0.05	63.7	0.1	61.3	0.17
Sample video3	Text1 12 byte	67.8	0.02	65.6	0.06	63.2	0.12
	Text2 37 byte	67.5	0.03	65.3	0.07	63.1	0.12
	Text3 58 byte	67.1	0.03	65.1	0.07	62.8	0.13
	Image1 35*37	66.7	0.04	64.8	0.08	62.4	0.15
	Image2 45*45	66.4	0.05	64.3	0.09	62	0.16
	Image3 50*50	66	0.06	64	0.1	61.5	0.17
Average		67.22	0.032	64.75	0.080	62.37	0.14

Table 4. sample video frames for different noise densities

Noise density	Frame30 (sample video1)	Frame35 (sample video1)	Frame45 (sample video1)
0			
0.1			
0.5			

Table 5. The extracted secret messages for different noise densities before using the filtering stage

Sample video1	Noise density			
Secret message	0	0.1	0.15	0.2
 Image1				
Text1	Hidden data	Hidden daua	Hidteb fata	Hrdden Sada

Table 6. The same extracted secret messages for different noise densities after using the filtering stage

Sample video1	Noise density			
Secret message	0	0.1	0.15	0.2
 Image1				
Text1	Hidden data	Hidden data	Hideebcata	Hodden Wata

4.4 Results discussion

For images and videos, PSNR between 30dB-50dB is acceptable **[15]**. The larger PSNR indicates the best image quality. **Table 2** shows that, the resulting PSNR has values between **61** and **68.8** dB and the steganography videos appear visually identical to the original ones and these values decreases as noise increases. **Table 3** shows that, the resulting PSNR has values between **61.5** and **68.5 dB** and the steganography videos appear visually identical to the original ones and these values decreases as noise increases. The results in **Tables 2 and 3** indicate that there are no big differences between the LSB parity method and the XORing of LSB's method, they both introduced excellent results even with noises. If the video frames exposed to salt and pepper noise with different noise densities, the filtering stage was used to treat this attack. **Table 4** shows that embedding Image3 with no noises does not affect in noticeable manner on the video frames and they appear identical to the original video for the human eyes. The results in **Table 5 and 6** show that, the extracted messages for different noise densities after using the filtering stage were better than before using the filtering stage especially with images.

5. Conclusion & Future work

The paper proposed two methods for hiding information (image and text) in digital video frames using the spatial domain, one was considered a parity Least significant Bit (LSB) and the second method was the XORing of LSB for each pixel in the R - image of video frame and the bit next to it. Different noise densities were applied to simulate the influence of intruders over the network who will try changing the content of the secret message. Using encryption along with steganography, these methods provide an additional level of security. From the experimental results, it was seen that the proposed methods were effective comparing with the other spatial domain methods. From seeing tests, no big difference was found between the original video file and the stego video file. The hidden information is recovered without any error for no noise added and the two methods introduced very similar and acceptable PSNR and MSE to each other. Using median filtering stage at the receiver side enhanced the retrieval of the secret message, especially with images and reduced the effect of adding noise.

The future work will focus on embedding secret messages using the same proposed methods and the same videos in the frequency domain using combined DCT-DWT and make a comparison for the results of both domains. Another challenge in the future work will be the network problems which may face the stego video during broadcasting over data network and how these problems will affect the hidden data retrieval.

References

[1] Y. Shi, Miao Qi, Y. Lu, J. Kong and D. Li "Object based self-embedding watermarking for video authentication" , Transportation, Mechanical, and Electrical Engineering (TMEE), 2011 IEEE International Conference on, Communication, Networking & Broadcasting ; Components, Circuits, Devices & Systems 2011.

[2] X. Zhang, S. Wang, Z. Qian and G. Feng, " Reference Sharing Mechanism for Watermark Self-Embedding, IEEE Transactions on Image Processing, vol. 20, pp. 485-495, 2011.

[3] D. Xua, b, R. Wangc, J. Wang, "A novel watermarking scheme for H.264/AVC video authentication," Signal processing: Image Communication, ELSEVIER, Volume 26, Issue 6, July 2011, Pages 267–279

[4] V. Sathya; K , Balasubramaniyam,; N. Murali,;M. Rajakumaran; Vigneswari, "DATA HIDING IN AUDIO SIGNAL, VIDEO SIGNAL TEXT AND JPEG IMAGES", IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012) March 30, 31, 2012

[5] Y. WANG, Q. HE, H. WANG, B. YIN and S. DING," Steganographic Method Based on Keyword Shift", Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on ,Communication, Networking & Broadcasting 2010 , Page(s): 454 – 456

[6]B. L. GunJal and R.R Manthalker ," An overview of transform domain robust digital image watermarking algorithms", Journal of Emerging Trends in Computing and Information Sciences volume 2 No. 1. ©2010

[7] Matlab Simulink R2012b

[8] G . Prabakaran,and R. Bhavani, "A Modified Secure Digital Image Steganography Based on Discrete Wavelet Transform", 2012 International Conference on Computing, Electronics and Electrical Technologies [ICCEET]

[9] S. Singh and G. Agarwal,"Hiding image to video: A new approach of LSB replacement", International Journal of Engineering Science and Technology Vol. 2, 2010, pp. 485-495.

[10] A. Al-Haj,"Combined DWT-DCT Digital Image Watermarking", Journal of Computer Science 3 (9): 740-746, 2007 - ISSN 1549-3636 © 2007 Science Publications

[11] W. Jue, Z. Min-qing, S. Juan-li, " Video Steganography Using Motion Vector Components", Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference , Page(s): 500 – 503

[12] J. Paul Cruz, N. Joseph Libatique, and G. Tangona, "Steganography and Data Hiding in Flash Video", TENCON 2012 - 2012 IEEE Region 10 Conference Date of Conference: 19-22 Nov. 2012 Page(s): 1 – 6.

[13] K. Kancherla and S. Mukkamala,"Block Level Video Steganalysis Scheme", 2012 11th International Conference on Machine Learning and Applications, 2012 IEEE

[14] http://www.mathworks.com/help/images/ref/imnoise.html

[15] "Peak Noise to Signal Ratio". [online]. Available: http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio