International Conference on Computational Intelligence: Modeling, Techniques and Applications (CIMTA) 2013

Optimized Video Steganography using Genetic Algorithm (GA)

Kousik Dasgupta\textsuperscript{a,}\textsuperscript{*}, Jyotsna Kumar Mondal\textsuperscript{b}, Paramartha Dutta\textsuperscript{c}

\textsuperscript{a}Department of Computer Science and Engineering, Kalyani Government Engineering College, Kalyani-741 235, India

\textsuperscript{b}Deptt. of CSE, Kalyani University, Kalyani-741 235, India

\textsuperscript{c}Deptt. of CSS, Visva-Bharati University, Santiniketan-731 235, India

Abstract

This paper proposes a novel video steganography scheme for efficient and effective information hiding. In this era of Internet communication video is considered to be an effective and important tool for communication. Video steganography uses video as cover media for embedding secret data. A 3-3-2 LSB based scheme has been used as a base technique for video steganography. Imperceptibility and video quality are supposed to be two key parameters for deciding goodness of any steganographic scheme. Thus the base technique is enhanced using Genetic Algorithm (GA) which thrives to get an optimal imperceptibility of hidden data. An anti-steganalysis test is used to check for the innocence of the frame with respect to original frame. Experimental results show a substantial improvement in the Peak Signal Noise Ratio (PSNR) and Image Fidelity (IF) values after optimization over the base technique. Complexity analysis of the proposed method is also reported in the paper.

© 2013 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords: Information hiding, Video steganography, LSB embedding, Genetic Algorithm

1. Introduction

The revolution brought by the digital information to the modern era is immense \cite{1}. This has also generated new opportunities of innovation and challenges. Several modern day devices like digital camera, powerful camcorder, digital voice recorder, multimedia personal digital assistant (PDA) has resulted into rich multimedia contents. Video is one such popular medium for communication and enjoyment, in today’s day to day life. The ease of editing and perfect reproduction in digital domain has brought forward two important domains, watermarking and data hiding (steganography). Of which the former is mainly concerned for maintaining the authentication and the latter used here, is for sending secret information to the intended users. In this paper Video is used as a cover media for embedding secret message, where videos can be said as a collection of frames and audio, either in compressed domain or in uncompressed domain.

The advantage of using video files in hiding information is primarily because video is more secure against hacker attacks due to the relative complexity of video compared to image files and audio files. Video based steganography
techniques are mainly classified into spatial domain and frequency domain based methods. Frequency domain techniques are mainly based on discrete cosine transforms (DCT) and wavelet transforms. S. Suma et. al. [2] proposed an integer wavelet transformation in cover video so as to get the stego-video. Where as Li. et. al. in [3] proposed a DCT method for hiding the secret message. In spatial domain the most widely used method is LSB substitution [4] where as MSB substitution can also be used. Daniel Socek et. al. [5] proposed a novel video encryption with steganography in digital videos. Tamer Shanableh [6] proposes two data hiding approaches using compressed MPEG video. Some other methods exist in literature [7] and [8] for video stegaography or data hiding. Subtlety, video-based steganography techniques generally takes such analysis into account and tries to maintain the statistics of the carrier before and after message hiding.

In recent years, it has been observed that Imperceptibility is the most important requirement in steganographic schemes. So researchers aim to achieve a stego version of the media where modification is slight and also transparent to human eye. Modern day Steganalysis however can detect even slight modifications. This has motivated researchers to design steganographic schemes [9] and [10] that are capable of resisting steganalysis. The capability of these “anti-steganalysis” techniques are enhanced by two approaches parametric and non parametric. “Parametric” means achieving anti-steganalysis by directly tuning the embedding parameter such as magnitude of the inserted noise per message embedding, embedding rate, embedding positions, etc [11]. In “Nonparametric” approach anti-steganalysis is achieved by modulating pixel values individually.

The architecture of the system is described in section 2. A 3-3-2 based LSB substitution scheme [12] used as a base technique is described in section 3. After embedding GA has been used as an optimizer to modify embedded pixels coefficients, so that some target performance are optimized as described in section 4. A performance analysis of the proposed optimizer with the base technique is described in section 5. Section 6 concludes the work.

2. System Architecture

The proposed system architecture for Video Steganography (Encoding) is given in Fig. 1(a). In the closed loop system, the carrier video is first converted to frames by the module Splitter. The Splitter module breaks the video into audio and frames. Though both audio and frames can be used to embed secret data, one or multiple frames have been used as a carrier in the paper. The carrier frame(s) is given as input to the Embedder. The embedding is done using the 3-3-2 LSB base embedding technique (as described in 3). The output of the embedder is stego frame(s). Now the stego frame(s) goes through an Optimizer, which optimizes the stego frame such that it is indistinguishable from the original version. The optimizer can use any optimization methods, e.g. GA, SA (Simulated Annealing) etc. In this paper we have proposed GA as the optimization technique. The optimizer optimizes the stego signal using the objective function as given in Equation 1. Next the optimized value goes through a Anti-steganalysis test module. In this a steganalytic subsystem as described in [13] has been used. The module analyses the gradient energy as statistical features. However it is difficult to achieve anti-stegalysis and optimization at the same time. Hence, an iterative procedure is used which works in closed loop. The stego frame(s) are then passed through a Merger module. It merges the stego frame(s) and all the remaining non stego frames and audio obtained from splitter module to make a Stego Video.

The system architecture for Decoding is depicted in Fig. 1(b). The Stego Video goes through the module Splitter, which breaks the video into audio and frames. The Stego frame(s) in particular is passed through a Decoder. It extracts the secret data from the stego frame. The output is the Secret data which was embedded inside the carrier video.

The proposed system architecture is implemented as a VStego Engine using Visual C++ 2012 as IDE (Integrated Development Environment) and Opencv 1.0 as the graphics library.

3. Base Technique: A 3-3-2 based LSB Video Steganography

A generic steganographic scheme can be described as [14] is a system where inside a host media or cover media ($M_0$), the embedding module inserts a set of secondary data ($e$), which is referred to as secret data, to obtain the stego media ($S_1$). A generic LSB based steganographic scheme can be described as choosing a subset ($j_i$) of cover elements ($C$) and performing the substitution operation $LSB_i(C_{m_i})$. Where $m_i$ can be either 1 or 0 and substitution can be in
multiple bits of LSB also.
In the base technique eight bits of secret data are considered for embedding at a time in the LSB of RGB (Red, Green and Blue) pixel value of the carrier frames in 3, 3, 2 order respectively [12]. Thus first three bits of the secret message are concealed inside three (03) bits of LSB of Red pixel, next three bits in the three (03) bits of LSB of Green pixel. The remaining two bits of secret message are concealed in two (02) bits of LSB of Blue pixel. The detailed technique has been depicted in Fig. 2. The particular distribution pattern is taken considering that the chromatic influence of blue to the human eye is more than that of red and green pixels. Hence without sacrificing the quality of the video an optimum payload can be achieved. Also this small variation in colors inside the large number of video frames would be very difficult for the human eye to detect. Thus producing stego frames.

The Base technique of Video Steganography, **encoding algorithm** is enumerated below:

1: Find 4 LSB bits of each RGB pixels of the cover frame.
2: Embed the eight bits of the secret image into 4 bits of LSB of RGB pixels of the cover frame in the order of 3, 3, 2 respectively.

3: Regenerate stego video frames.

Whereas the Decoding algorithm is explained below:

1: Find 4 LSB bits of each RGB pixels of the stego frame.

2: Retrieve the bits of secret data from LSB of RGB pixel of the stego frame in the order of 3, 3, 2 respectively.

3: Reconstruct the secret information.

4: Regenerate video.

4. Genetic Algorithm as an Optimizer over Base Technique

The stego frames obtained from the base technique has resulted in changes of RGB pixel of the original frames but imperceptibility of the video needs to be taken care for successful steganography. For design of any steganographic schemes [1] several factors should be considered like imperceptibility, embedding capacity, statistical undetectability (antisteganalysis), Bit error rate (BER) after data extraction and robustness to attacks. However some of the factors conflict with one another, such as, increasing embedding capacity might reduce the imperceptibility, etc.. Hence any steganographic problem can be viewed as an Optimization problem where a steganography scheme maps a secret data (or stego signal) to a host media (or undetected region) [14]. Thus an objective function that minimizes all the mentioned parameters and giving a completely optimal solution is not possible. Hence, in this paper an objective function as in Equation 1 has been proposed where preferred parameters are optimized and letting all others be inequality constraints. The proposed objective function $E$ has Mean square error (MSE) ($f_1$) and Human vision system (HVS) deviation ($f_2$) as preferred parameters,

$$E = w_1 \times f_1 + w_2 \times f_2$$  \hspace{1cm} (1)

where $w_1$ and $w_2$ are predefined weights. It is very difficult to decide/optimize the weights, one criterion can be that more general the factor larger is the weight. Another logic is user’s preference or importance given to a particular factor over the other. Here the later approach has been used and the optimization is then performed on the given set of weights. The weights are considered as $w_1 = 0.8$ and $w_2 = 0.2$. The most widely adopted statistical image quality feature for accessing image quality is $MSE$, given by Equation 4. It measures the distortion between pixels of stego frame and original frame. The other preferred parameter in objective function is $SSIM$ (structural similarity) [15] accounts for HVS characteristics. It takes care of substantial point-by-point distortions that are not perceptible, such as spatial and intensity shifts, as well as contrast and scale changes. $SSIM$ is a function of luminance comparison $l(x,y)$, contrast comparison $C(x,y)$ and structure comparison $s(x,y)$ as given in Equation 2:

$$SSIM = f(l(x,y), C(x,y), s(x,y))$$  \hspace{1cm} (2)

This optimization problem is solved by Genetic Algorithm using the Optimizer module of system architecture explained in section 2. A little research has been done in application of GA to video steganographic problems, though some work exits in literature on image steganography [16].

Genetic Algorithm [17] has been used by researchers as an optimization tool in varied set of problems. This paper uses a basic GA approach for optimization. The proposed algorithm for GA as an optimizer of the base video steganography technique:

**Input:** Stego frame(s) with secret data embedded in 3-3-2 target layers of LSB of each RGB pixels.

**Output:** Optimized Stego frame(s).

**Initialization of population:** Objective of this step is to get different chromosomal representation of the pixel value of the stego frame. A random selection of data points are made as initial population. Where each of the data points have same target layers.

**Mutation:** This step selects most of the times the best fitted pair of individuals for crossover. The fitness value of each individual chromosomes are calculated using the fitness function as given in 1. The best fitted value chromosome is selected twice and the least fitted value is discarded for mutation. A very small value (5%) is chosen as mutation
probability. Depending upon the mutation value the bits of the chromosomes, except the target layers, are changed from ‘1’ to ‘0’ or ‘0’ to ‘1’. The output of this is a new mating pool ready for crossover.

**Crossover:** Objective of this step is to perform crossover between the Mating pool selected in the previous step. A random single point crossover is chosen and portion lying on one side of crossover site is exchanged with the other side. Thus it generates a new pair of individuals.

The steps **Mutation** and **Crossover** are repeated iteratively till, either maximum number of iterations are exceeded or we get a chromosome having pixel value closest to the original value.

The optimized stego frame(s) are then merged with non stego frames and audio in the merger module as explained in Fig 1(a). The final output is an optimized stego video.

5. **Performance evaluation**

For performance evaluation of the proposed technique three video are considered, details of each are given in Table 1. The details of the secret image is also given in Table 1. Any Steganographic technique is evaluated on basis of payload and imperceptibility. Where the former describes the capacity of secret data embedded in the carrier media and the later gives the measure of embedded data imperceptible to the observer (perceptual invisibility) and computer analysis (statistical invisibility). The measure of capacity for the different carriers are listed in Table 2 in terms of payload (bits per byte or $bpB$). Increase or maintaining the payload and maintaining an acceptable level of Stego quality is considered as a good contribution.

Two types of perceptibility measure are listed in Table 2 namely fidelity and quality. Fidelity means the perceptual similarity between signals before and after processing. However, quality is an absolute measure of the goodness of a signal to avoid any suspension and therefore detection. The quality measure is given by $PSNR$ [18] as given in Equation 3.

$$PSNR = 10 \log_{10} \frac{L^2}{MSE}$$  \hspace{1cm} (3)

where, $L$ is peak signal level for a grey scale image it is taken as 255. The value of MSE is calculated by Equation 4.

$$MSE = \frac{1}{H \times W} \sum_{i=1}^{H} (P_{(i,j)} - S_{(i,j)})$$  \hspace{1cm} (4)

where, $H$ and $W$ are height and width and $P_{(i,j)}$ represents original frame and $S_{(i,j)}$ represents corresponding stego frame. Whereas the fidelity measure is measured by Image Fidelity ($IF$) [18] as given in Equation 5.

$$IF = 1 - \frac{\sum_{i,j} (I_{i,j} - \bar{I}_{i,j})^2}{\sum_{i,j} \bar{I}_{i,j}^2}$$  \hspace{1cm} (5)

where, $i$ and $j$ are coordinates of the pixel, $I_{i,j}$ is pixel value of carrier frame and $\bar{I}_{i,j}$ is pixel value of stego frame. The results are also compared with the corresponding base technique. As eight (08) bits are embedded per three (03) bytes, so payload is 2.66 $bpB$. Comparing the results it can be observed that, though the payload is same as that of the base method, the $PSNR$ and $IF$ values show improvement.

5.1. **Complexity analysis**

The complexity analysis of any algorithm includes computation complexity (time complexity) analysis and space complexity analysis. The proposed 3-3-2 based LSB substitution replaces $k_1$ bits of LSB of each pixel during encoding. Where $k_1$ represents total number of bits replaced per pixel at a time constant, $c_1$. Thus the time complexity of encoding is at most $(c_1 \times k_1)$. In the same lines for decoding time complexity will be at most $(c_1 \times k_1)$. In addition, for accessing concerned pixel of carrier frame there will be time complexity of at most $n^2$ during encoding and another $O(n^2)$ during decoding. So total time complexity $S$ is as given below,

$$S = O((c_1 \times k_1) + (c_1 \times k_1) + n^2 + n^2)$$  \hspace{1cm} (6)
Table 1: Cover Video File details

| S.No. | Cover video file information | Secret message Resolution W*H |
|-------|------------------------------|-------------------------------|
|       | Name of video                | Resolution W*H                | Frame sec | No. of frames | |
| 01    | tree. avi                    | 320*240                       | 30        | 450           | 150*150 |
| 02    | globe. avi                   | 320*240                       | 30        | 107           |         |
| 03    | computer. avi                | 320*240                       | 30        | 510           |         |

Table 2: Performance evaluation of GA as an optimizer over Base Video Steganography Technique 3-3-2 LSB

| Name of video | Results obtained using GA as an optimizer over Base technique | Results obtained using Base Video Steganography Technique 3-3-2 LSB |
|---------------|-------------------------------------------------------------|---------------------------------------------------------------|
|               | PSNR            | IF (bpB)           | Payload (bpB)       | PSNR            | IF (bpB)           | Payload (bpB)       |
| tree. avi     | 39.374          | 0.99               | 2.66                | 38.03           | 0.87               | 2.66                |
| globe. avi    | 34.372          | 0.99               | 2.66                | 32.67           | 0.89               | 2.66                |
| computer. avi | 41.613          | 0.99               | 2.66                | 39.21           | 0.86               | 2.66                |

Thus the time required by 3-3-2 based LSB substitution is \( O(n^2) \). The next part of the proposed technique uses GA for optimization where initialization of population is considered as GA preprocessing step so it can be ignored for complexity analysis. For encoding into binary string for chromosomal representation time complexity is at most \( n \). Next for evaluation of fitness function 1 a time complexity of at most \((c_2 \times k_2)\), where \( c_2 \) is time constant and \( k_2 \) represents number of chromosomes. Next for mutation a complexity of at most \( m \), where \( m \) is the chromosome length. For single point crossover another \( m \), so the total complexity of GA process \( G \) is given by,

\[
G = O((c_2 \times k_2) + m + m)
\]

Which comes out to be at most \( O(m) \). So the total time complexity of the proposed technique is approximately \( O(n^2) \). The space complexity of the technique is due to storage of eight bits of secret data and twenty four (RGB) bits of carrier frame for the 3-3-2 based LSB substitution. This can be considered of the order of at most \( (k_3) \), where \( k_3 \) is thirty two bits of memory. For the genetic algorithm it is due to storage of chromosomes in memory. Let the length of chromosome is \( N \) and the size of total population is \( M \) making the space complexity of the proposed GA is of at most \( (M \times N) \). So the space required for proposed technique is approximately \( O(M \times N) \).

6. Conclusion

A GA based Optimized video steganographic scheme has been proposed. The optimizer optimizes the values over basic video steganography done using a 3-3-2 LSB technique. The optimizer uses a cost function consisting of two factors, however other factors can also be included for further detailed study. A Performance evaluation has been done of the proposed technique with the base techniques on the basis of perceptibility and fidelity. The \( PSNR \) values lies between 20 and 40 dB, which is considered as standard. Further it becomes difficult for the human visual to recognize...
any difference between a cover and Stego file if the PSNR value exceeds 36dB. An anti-steganalysis test has been done on the stego frame in a closed loop system, so separate video steganalysis studies is avoided. The techniques are applied in uncompressed domain it can be extended to compressed domain. Though GA has been quite appreciable used as an optimizer other optimizing technique are on the anvil.

Acknowledgements

The authors would like to thank their institutions for supporting the research. The work is supported by AICTE as a part of RPS project entitled, "Design and Development of Framework for Multimedia security and Forensic to Provide Digital Rights Management of Video".

References

[1] Min Wu, Bede Li, *Multimedia Data Hiding*, Springer, 1st edition, 2003.
[2] S. Suma, “Improved Protection in Video Steganography using compressed Video Bitsterams,” in *International Journal on Computer Science and Engineering*, Vol. 02, No. 03, pp. 764–766, 2010.
[3] Y. Li, H.-X. Chen, and Y. Zhao, “A new method of data hiding based on H.264 encoded video sequences,” in Proc. IEEE Int. Conf. Signal Processing, ICSP, pp. 1833–1836, 2010.
[4] M. Ramalingam, “Stego Machine – Video Steganography using Modified LSB Algorithm,” in Proc. World Academy of Science, Engineering and Technology 74 2011, pp. 502–505, 2011.
[5] D. Socek, H. Kalva, Spyros S. Magliveras, O. Marques, D. Culibik and B. Furrh, “New approaches to encryption and steganography for digital videos,” in Proc. Multimedia Systems, Springer-Verlag 2007.
[6] Tamer Shanableh, “Data Hiding in MPEG Video Files Using Multivariate Regression and Flexible Macroblock Ordering,” in Proc. IEEE Transactions on Information Forensics and Security, VOL. 7, NO. 2, pp. 455-464, 2012.
[7] V. Sampat, K. Dave, J. Madia and P. Toprani, “A Novel Video Steganography Technique using Dynamic Cover Generation,” in Proc. National Conference on Advancement of Technologies-Information Systems and Computer Networks, ISCON-2012, Proceedings published in International Journal of Computer Applications, 2012.
[8] Wang Jue, and Zhang Min-qing, “Video Steganography Using Motion Vector Components,” in Proc. of IEEE 978-1-61284-486-2/11, pp. 500-503, 2011.
[9] Y.-T. Wu and F. Y. Shih, “Genetic algorithm based methodology for breaking the steganalytic systems,” in Proc. of IEEE Trans. Syst., Man, Cybern. B, vol. 36, no. 1, pp. 25–31, 2006.
[10] M. Kharrazi, H. T. Senchar, and N. Memon, “Cover selection for steganographic embedding,” in Proc. of Int. Conf. Image Processing, pp. 117–120, 2006.
[11] Guo-Shiang Lin, Yi-Ting Chang, and Wen-Nung Lie, “A Framework of Enhancing Image Steganography with Picture Quality Optimization and Anti-Steganalysis Based on Simulated Annealing Algorithm,” in Proc. of IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 12, NO. 5, pp. 345–357, 2010.
[12] Kousik Dasgupta, J.K.Mandal, Paramartha Dutta, “Hash Based Least Significant Bit technique for Video Steganography(HLSB),” inInternational Journal of Security, Privacy and Trust Management (USPTM), pp 1-11, April 2012 (DOI: 10.5151/ijsptm.2012.2201).
[13] W.-N. Lie, G.-S. Lin, and S.-L. Cheng, “Dual protection of JPEG images based on informed embedding and two-stage watermark extraction,” in Proc. of IEEE Trans. Inf. Forensics Security, vol. 1, no. 3, pp. 330-341, 2006.
[14] M. Wu and B. Liu, “Data hiding in image and video: Part IFundamental issues and solutions,” in Proc. of IEEE Trans. Image Process., vol. 12, no. 6, pp. 685-695, 2003.
[15] Z. Wang, A.C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error measurement to structural similarity,” IEEE Trans. Image Processing, vol. 13, no.4, pp.600-612, Apr. 2004.
[16] S. P. Maity and M. K. Kundu,“Genteic ALgorithm for optimality of data hiding in digital images”, in SOft Comput(2009)13, Springer-Verlag, pp. 361-373, 2009.
[17] Goldberg D, “Genetic Algorithm in Search, Optimization and Machine Learning,” Addison-Wesley, 1989.
[18] M. Kutter and F. A. P. Petikolas, “A fair benchmark for image watermarking systems”, in Proc. of Electronic Imaging ’99. Security and Watermarking of Multimedia Contents, vol. 3657, The International Society for Optical Engineering, pp. 1-14, 1999.