URSI GASS 2020, Rome, Italy, 29 August - 5 September 2020

A New Embedding Technique Based On Psychovisual Threshold for Robust and Secure Compressed Video Steganography

Ferda Ernawan*, Muhammad Fuad Abdullah
Faculty of Computing, Universiti Malaysia Pahang

Abstract

Videos are often compressed to reduce storage and transmission payload at the expense of lower quality due to bandwidth-related issues. Most video steganography techniques do not provide robustness against compression technique. Thus, it is vital to develop a steganography technique that can be resistant against compression. This research proposed a new embedding technique in video steganography based on object motion and modified entropy. The object motions in the video frame were determined by horizontal and vertical motion vectors. The video frames that had object motion were computed by modified entropy. The proposed scheme embedded data along with the object motion by modifying Discrete Cosine Transform (DCT) coefficients in the video frames. Six DCT coefficients were selected in the middle frequency using DCT-psychovisual effects of hiding messages. The experimental results showed that the scheme achieved good robustness of message recovery in terms of Bit Error Rate (BER) and Normalised Cross-Correlation (NC). The recovered message of the proposed steganography scheme can survive video compression

1 Introduction

Video file provides extensive data and it requires a large processing transmission. Due to the massive amount of data, videos are often compressed to reduce the storage and the transmission payload at the expense of lower quality due to bandwidth-related issues. Due to limited bandwidth, videos are usually compressed to minimise the size of the file before transferring the video data. Videos provide additional hidden space in a scene change; those hidden spaces are highly imperceptible to human visual systems. The existing video steganography schemes are not resistant under video compression. The hidden message is destroyed by the quantisation process in video compression [1]. Therefore, it is quite challenging to develop a video steganography technique that is resistant under video compression and maintains the quality of the video close to the original one. This has motivated researchers to reveal the hidden space of the video data and design a new hiding technique that is robust against video compression.

Many steganography techniques present a data hiding scheme by randomly selecting frames of videos. They do not consider the bit error of hidden data. Concealed messages by randomly selecting frames may produce some distortions in the stego-videos. In order to manage the video quality, there exist a number of data hiding techniques that conceal messages among scene changes. A video steganography technique based on scene-change provides better security and less distortion in the quality of the video [2]. The scheme implements Least Significant of Bits (LSB) for concealing the messages, which are easy to be removed against compression methods.

Furthermore, this scheme does not consider the optimal bits of the hidden data. Scene-change redundancies have not been revealed like the redundancy model that was investigated in audio steganography [3]. The field of audio steganography has made significant progress in identifying the redundant part of the audio to conceal a secret message while reducing the distortion of the stego-audio. Unlike audio steganography, the scene-change process in videos has not been fully investigated in the video steganography system.

This research proposed DCT psychovisual effect and object motion for concealing a message in video data. The proposed video steganography used Discrete Cosine Transform (DCT) for its compact transform and ease of implementation. The message concealed in the selected DCT coefficients did not give a significant effect on the quality of the video. The proposed scheme provided the right level of video quality and it can be resistant against the compression technique.

2 Existing of Video Steganography

Video steganography is a technique of communication by hiding data through a cover medium. There exist a number of data hiding techniques that conceal messages using the LSB method and transform domain. A video steganography technique based on LSB provides less distortion in the quality of the video.

The scheme by [2] presented a video steganography scheme using scene change detection, DCT-DWT, and Least Significant of Bits (LSB) for concealing the messages. Furthermore, this scheme did not sufficiently consider the optimal bits of the hidden data. The proposed steganography technique provided rendering payload to increase the absolute visual quality. However, the steganography technique may produce some distortions on the extracted secret message when the stego-video is compressed by compression techniques. The quality of the extracted hidden data needs to be improved when the cover-medium is compressed.

Kar et al. [4] presented a video steganography using DNA alphabets to secure communication and privacy. Their scheme hid the message in the video by using LSB substitution method. Their experimental results showed that the proposed algorithm was able to maintain the steganography video quality. Their scheme produced lower distortion for various payloads and claimed to be secured using the random video frame selection. However, this approach is highly vulnerable to stego-analysis under compression. The hidden data may be lost because of the quantisation process in the compression technique.

Moreover, Nguyen et al. [5] proposed embedding data after quantising DCT coefficient in the video H.264/AVC sequence. The authors also claimed that the proposed technique was able to produce high visual quality of the stego-video and improve the embedding capacity. Their experimental results showed that the proposed method outperformed the existing technique and maintained high quality stego-videos.

Furthermore, BanuPriya et al. [6] presented adaptive LSB embedding in video steganography for security application. The secret data was encrypted by using RC7 encryption. It was embedded in the randomly selected video frames. Their experimental results showed the hidden secret message was able to be recovered without any loss. Ningsih et al. [7] also presented video steganography using LSB and the advanced encryption standard (AES). The authors used an encrypted image as a secret data. Video steganography using LSB produced less degradation in the video quality. Meanwhile, the hidden message was not resistant under video compression and was destroyed by the quantisation process in video compression. Therefore, it is quite challenging to develop a video steganography technique that is resistant under video compression and maintains the quality of video close to the original video.

Kumar and Singh [8] presented a video steganography scheme based on Discrete Wavelet Transform (DWT). The authors considered the human skin region as the Region of Interest (ROI) for embedding the secret data. The scheme achieved high imperceptibility and robustness against MPEG-4 compression. However, the authors did not consider the video frame selection for hiding data. The hidden data might be easy to recognise and removed by unauthorised users. The scheme also required a large computational complexity due to wavelet transform.

3 Proposed Hiding Message

The Firstly, the video was split frame-by-frame. Each frame was computed by the motion detection method to find the motion vector. Motion detection was performed with non-overlapping blocks of 8×8 pixels from left to right and top to bottom. In order to select the embedded frame, the selection process started to find the object motion in the video. The motion vector can be calculated by Cumulative Absolute Difference (CAD) between current and previous frames. All the coordinates of object motion were saved and the modified entropy was applied to all of the video frames. Each video frame was sorted according to the modified entropy value.

The blocks that had motion vector can be utilised to conceal the message. The magnitude of motion vector can be defined by:

$$m = \sqrt{x^2(i) + y^2(i)} \tag{1}$$

where m denotes the magnitude value, $x[i]$ represents the horizontal motion vector, and $y[i]$ denotes vertical motion vector in the i-th macro-block. The macro-block with significant magnitude values were selected for the hiding location.

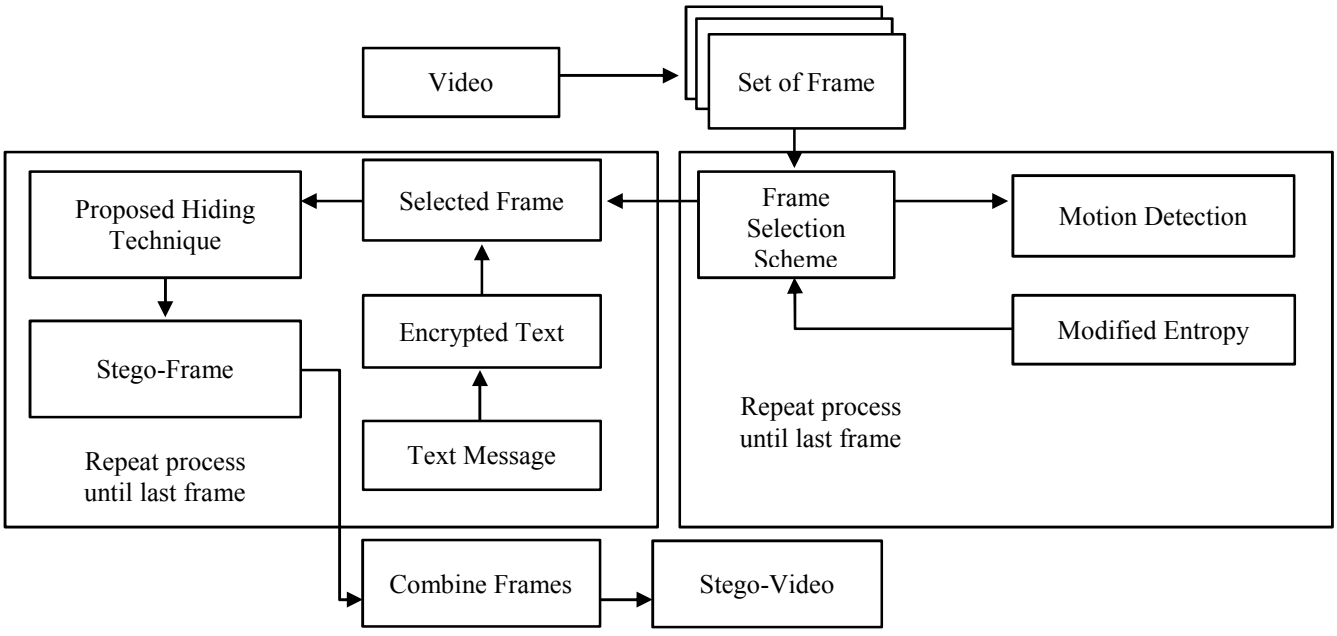

Figure 1. Block diagram of hiding message

After we obtained all motion vectors, each frame computed by the modified entropy. The modified entropy defined by

$$E = -\sum_{i=1}^{n} p_i \log 2(p_i) + p_i \exp 1 - p_i \qquad (2)$$

pi indicates the probability of *i*, $0 \le pi \le 1$. The entropy of image frames was sorted by ascending order. The frame with the lowest modified value was selected for the first hidden message. The hiding location was determined by the significant magnitude of the motion vector. Before the message was inserted, it was encrypted using Advance Encryption Standard (AES) 128 bits. The scrambled message bits by AES provided additional security in the video steganography from attackers.

Each selected frame was transformed by DCT. The message was hidden by examining the selected DCT coefficients of ((3,2) (2,3)), ((1,3) (0,4)), and ((1,4) (2,3)) coordinates. The selected coefficients were grouped into three pairs. If the hidden message is equal to 1, the first coefficient is designed less than the second coefficient. The value was swapped and the threshold value was added to the first coefficient. Or else, the value of the first coefficient will be replaced with the second coefficient and added to the threshold value. The hiding technique for message bit equal to 1 is given as follows:

If (|SC(2x)|<|SC (2x+1)|) *then*
C= SC (2x);
SC (2x)= SC (2x+1)+ s;
SC (2x+1)=C;
else
SC (2x)= SC (2x)+ f;
SC (2x+1)=A(2x+1);

The hidden message will consider the length of message bits. If the hidden message is equal to 0, then the first coefficient is less than the second coefficient, and the value of the second coefficient is added with a threshold coefficient. Or else, the value of the first coefficient will be swapped with the second coefficient and the second coefficient will be added with a threshold. The hiding technique for message bit equal to 0 is given as follows:

If (|SC (2x)|<| SC (2x+1)|) then
SC (2x)= SC (2x)+ s;
SC (2x+1)= SC (2x+1);
else
C= SC (2x);
SC (2x)= SC (2x+1);
SC (2x+1)=C+ f;
end (if)

where x = 0, 1, and 2. SC (2x) represents SC (0), SC (2), and SC (4) and SC(2x+1) denotes SC (1) and SC (3). f and s are the proposed scaling factor for hiding data. Then all frames were merged into a video data. The proposed hiding scheme is summarised in Algorithm 2 as follows:

Algorithm 2: Hiding Technique
Input: Selected video frame; thresholds (*f* and *s*)
Step 1: Select blocks that have object motion.
Step 2: Selected blocks are transformed by 8×8 DCT.
Step 3: Select DCT coefficients based on psychovisual effect and re-arrange it into a vector *D*.
Step 4: Hiding data algorithm is defined by:

S=1;
for u=0 to 2
if S<=length(frame)
if frame(S)=1 then if(|D(2u)|<|D (2u+1)|)then
C= D (2u);
D (2u)= D (2u+1)+ s;
D (2u+1)=C;
else
D (2u)= D (2u)+ f;
D (2u+1)=A(2u+1);
end (if)
else
if(|D (2u)|<| D (2u+1)|)then
D (2u)= D (2u)+ s;
D(2u+1)= D (2u+1);
else
C= D (2u);
D (2u)= D (2u+1);
D (2u+1)=C+ f;
end (if)
end (if)
S=S+1;
end (for)

for u = 0,1 and 2. D (2u) represents $D(0)$, $D(2)$ and $D(4)$ and D(2u+1) denotes $D(1)$ and $D(3)$. f and s are the proposed scaling factor for hiding data as described in Algorithm 2.
Step 9: The modified coefficients are arranged into two-dimensional matrix.
Step 10: Apply inverse DCT on each selected block.
Step 11: Merge all blocks into image frame and then sequence of images
Output: Stego-video

4 Proposed Extracting Message

A stego-video was divided into the image frames. Referring to the motion analysis from the database, x and y block coordinates of the hidden message were identified for each selected frame. Each selected block was transformed by 8x8 DCT. The message can be extracted with certain rules. If the first coefficient pair is less than the second coefficient, the message bit is equal to 1. If the first coefficient pair is greater than the second coefficient, then the message bit is equal to 0. The extracted message was decrypted by AES 128 bits. Next, the message bits were rearranged to reconstruct the message. The proposed step-by-step process of extracting message is discussed in Algorithm 3.

Algorithm 3: Extracting Technique
Input: Stego-video; x and y coordinates of the selected blocks based on object motion;
Step 1: x and y coordinates of object motion are used to select concealed message region.
Step 2: Each selected block is transformed by 8×8 DCT.

Step 3:Select DCT coefficients based on psychovisual effect and re-arrange it into a vector *D*.
Step 4: Selected DCT coefficients *D* are computed by the following rule:

if $D(k) < D(k+1)$ for $k=0,2,4$ *then*
    message_bit =1,
else
    message_bit =0
end(*if*)

Step 5: Each message bit is arranged to recover the message.
Output: Message recovery

5 Experimental Results

The experimental results of the proposed hiding technique in video steganography demonstrated the statistical robustness of stego-video. This experiment used 20 videos to test the proposed algorithm. All the video sequences were in the uncompressed format. The robustness performance of the proposed stego-videos was tested by implementing MPEG-4 compression attack. The hidden message was extracted from the compressed stego-videos. The experimental results of the BER values from the proposed scheme are listed in Table I.

TABLE I. BIT ERROR RATE OF THE RECOVERED MESSAGE UNDER MPEG-4 COMPRESSION

Video	LSB	Without Frame Selection	Proposed scheme
Akiyo	0.4816	0.0333	0.0499
Foreman	0.5039	0.0621	0.3317
Xylophone	0.4934	0.0220	0.0122
Soccer	0.4705	0.0300	0.4049
Football	0.5018	0.0377	0.0315
Coastguard	0.4971	0.1261	0.1295
Mobile	0.4882	0.1094	0.1218
Waterfall	0.4830	0.0290	0.00071
Flower	0.4823	0.4174	0.4034
Bus	0.4941	0.1264	0.0959

Referring to Tables I, the proposed scheme was tested under various amounts of hidden data, such as 700 characters or 5,600 bits. The experimental results showed the comparison between video steganography using LSB method, the proposed scheme without frame selection and with frame selection. The proposed scheme with frame selection provided additional security, where hiding the message was performed based on modified entropy and object motion for selecting video frame. The hidden data might not be easy to detect by unauthorised people. The proposed scheme without frame selection concealed the message from the first video frame. The number of selected video frame considered the amount of message bits.

These experiments evaluated the robustness of hidden data with various amounts of message bits against video compression. The experiments have been tested by MPEG-4 compression. The proposed scheme outperformed the LSB technique in terms of BER and NC values for recovered message under MPEG compression. The proposed scheme hid a message into the selected frame based on modified entropy frames. It can avoid message removal when the stego-video was cropped at the first video frame. The frame selection was also able to improve the security from unauthorised users. The technique is shown to be potentially resistant of the hidden data against MPEG compression. The statistical Mean Peak Signal to Noise Ratio (MPSNR) values for hidden data of 5,600 bits of the "Akiyo", "Foreman", "Coastguard", and "Bus" is shown in Figure 2.

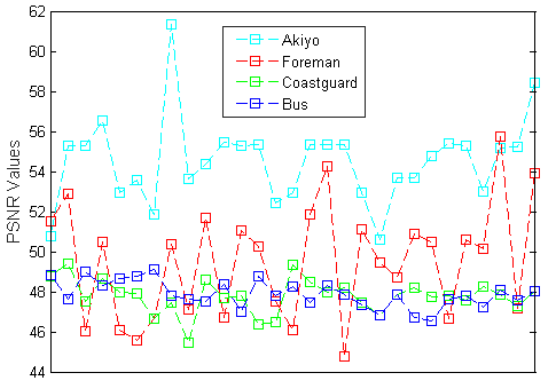

Figure 2. MPSNR value obtained from proposed scheme for Akiyo, Foreman, Coastguard, Bus with hidden message 5600 bits.

Figure 2 shows the sample of 30 concealed video frames for different amounts of hidden data. The proposed video steganography achieved average PSNR value of stego-video around 52 dB. The embedding capacity of hidden data depends on the number of detecting motion in the video data. Hence, the proposed scheme is not suitable for videos which have minimum object motion. The histogram of the stego-video is shown in Figure 3.

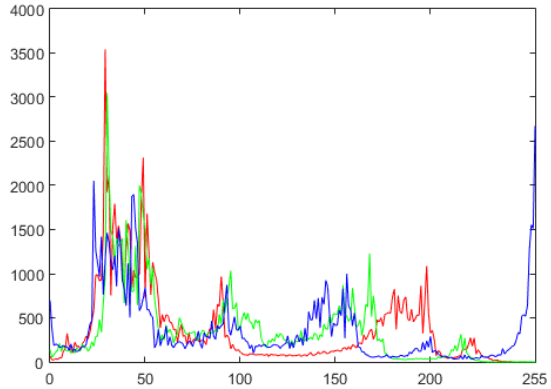

Figure 3. Histogram of hidden data in RGB channel

Referring to Figure 3, the histogram of stego-video frame is closer to the histogram of the original video frame. In order to show the significant change of pixel

distribution for both histograms, the RGB colour was converted into YCbCr colour.

6 Conclusion

This research presented video steganography based on DCT psychovisual and object motion. The object motion in video data was used to determine the selected video frame and the embedded block locations. The psychovisual threshold was used to determine the hiding location on the selected blocks. The concealing message was designed based on certain rules by modifying the selected DCT coefficients in the middle frequencies. The experimental results showed that the proposed scheme produced a good visual quality of video and undetectable to the human visual system. The proposed scheme also achieved high robustness of the concealed message under MPEG-4 compression.

6 Acknowledgements

This work was supported by Fundamental Research Grant Scheme (FRGS) No. RDU190117 from Ministry of Higher Education, Malaysia

7 References

1. M. Liśkiewicz, R. Reischuk and U. Wölfel, "Security levels in steganography – Insecurity does not imply detectability," *Theoretical Computer Science*, ***692***, 2017, pp. 25-45, https://doi.org/10.1016/j.tcs.2017.06.007
2. M. Ramalingam, and N.A.M. Isa, "A data-hiding technique using scene-change detection for video steganography," *Computers and Electrical Engineering*, ***54***, 2016, pp. 423–434, https://doi.org/10.1016/j.compeleceng.2015.10.005
3. H. Ghasemzadeh, M. Tajik Khass and M. Khalil Arjmandi, "Audio steganalysis based on reversed psychoacoustic model of human hearing," *Digital Signal Processing: A Review Journal*, ***51,*** 2016, pp. 133-141, https://doi.org/10.1016/j.dsp.2015.12.015
4. N. Kar, K. Mandal and B. Bhattacharya, "Improved chaos-based video steganography using DNA alphabets," *ICT Express*, ***4***(1), 2018, pp. 6–13. https://doi.org/10.1016/j.icte.2018.01.003
5. D. C. Nguyen, T.S. Nguyen, F.R. Hsu and H.Y. Hsien, "A novel steganography scheme for video H.264/AVC without distortion drift," *Multimedia Tools and Applications*, ***78***(12), 2019, pp. 16033–16052. https://doi.org/10.1007/s11042-018-6976-3
6. R. BanuPriya, J. Deepa and S. Suganthi, "Video steganography using LSB algorithm for security application," *International Journal of Mechanical Engineering and Technology*, **1**, 2019, pp. 203–211.
7. P.A.S.L.E. Ningsih, G.M.A. Sasmita and N.M.I.M. Mandenni, "MP4 video steganography using least significant bit (LSB) substitution and advanced encryption standard (AES)," *Journal of Theoretical and Applied Information Technology*, ***95***(21), 2017, pp. 5805–5814.
8. P. Kumar and K. Singh, "An improved data-hiding approach using skin-tone detection for video steganography," *Multimedia Tools and Applications*, *77*(18), 2018, pp. 24247–24268. https://doi.org/10.1007/s11042-018-5709-y