Video Steganography – A Novel Approach

[1]Prof. Samir Kumar Bandyopadhyay

[1]Advisor to Chancellor
JIS University, Kolkata, India

***Abstract*: The basic model of Digital Steganography is composed of three types of objects: Cover or Carrier or Vessel object, Secret object and Stego object, which is nothing but combination of Cover and Secret. Depending on the types of the cover object, different types of Digital Steganography have been defined, like Image Steganography, Audio Steganography, Video Steganography and Text Steganography. In this work, an initiative has been taken to identify the open areas Video Steganography.**

***Index Terms*: VOIP, Text Steganography, Video Steganography, GOP, Predicted Frame.**

I. INTRODUCTION

A video is a collection of still images called frame. Therefore, to hide any secret within video, first it needs to be broken into collection of frames. Then frames can be selected to hide data in similar approach of image steganography.

A video can be composed of three types of frames:

- Intra-coded Frame or I-Frame: This frame is popularly known as key frame. This frame is self-referential and doesn't use information of any other frame. These are largest frame amongst the three types of frames and possess high quality but less efficient in compression perspective.
- Predicted Frame or P-Frame: This frame is forward predicted from last P-frame or previous I-Frame. P-Frames are more efficient than I-Frame but not than B-Frame. This is inter-coded frame.
- among three; it looks forward and backward for redundant picture information. This is also inter-coded frame.

There are two types of compression techniques: inter-frame and intra-frame. The inter-frame or temporal compression defines that a video is composed of group of frames that references each other. It requires more processing power to look back neighbour frames to display one frame. In case of intra-frame compression each frame is self-contained doesn't require looking back its neighbours. Group of pictures (GOP) specifies how inter-frame and intra-frames are arranged in a video [1].

In this paper, first construction of video and how a frame has been chosen to hide secret data by scene change detection have been explained. In the second part a novel technique of video steganography has been discussed which key based and robust technique of data is hiding. To validate the imperceptibility, Structural Content (SC), Peak Signal to Noise Ratio (PSNR in dB), Normalized Cross-Correlation (NCC), Average Difference (AD), Maximum Difference (MD), and Normalized Absolute Error (NAE) have been calculated. To ensure the effectiveness, embedding capacity has been calculated and to validate the robustness, intentional Steganalysis attacks have been applied.

II. LITERATURE REVIEW

To process video easily each video frame is evenly partitioned into smaller chunks which are called Macroblocks [2-4]. Authors have proposed that Macroblock composed of a 16x16 block of luma and chroma [5-6]. Chroma or Chrominance (C) is the colour information of the picture. Luma or Luminance (Y) is the brightness of the light. Chrominance and Luminance collectively define the colour of interest. In High Efficiency Video Coding (HEVC, one of the video compression standards; e.g. H.265, MPEG-H Part 2 etc.), a concept of Coding Tree Block has been introduced, which splits the Macroblocks into multiple Coding Blocks. Author shows simultaneously CB can be partitioned into transform blocks which can be used as input to linear block transform DCT in Video Codec, using which video compression is done [7-9]. Figure 1 reflects how different types of frames can compose a Group of Pictures (GOP).

In video steganography a specific frame can be chosen to hide data. Now this frame can be chosen depending on scene change. It triggers a topic known as scene change detection and shot detection [3-6]. A scene is collection of shots. In other words, a scene is unit of video which takes place at specific location in same time whereas a shot results from one continuous recording by a single camera. The gap between two shots is called shot boundaries. There are broadly two types of shot boundary detection method:

- Abrupt Transition: In this transition sudden transition from one shot to another happens. This is sometimes called hard cut or simply cut. Figure 2 demonstrates an example for abrupt scene change.
- Gradual Transition: This transition is correlated by spatial-chromatic effect which means one shot gradually replaces another one. This soft transition can be of different types like:

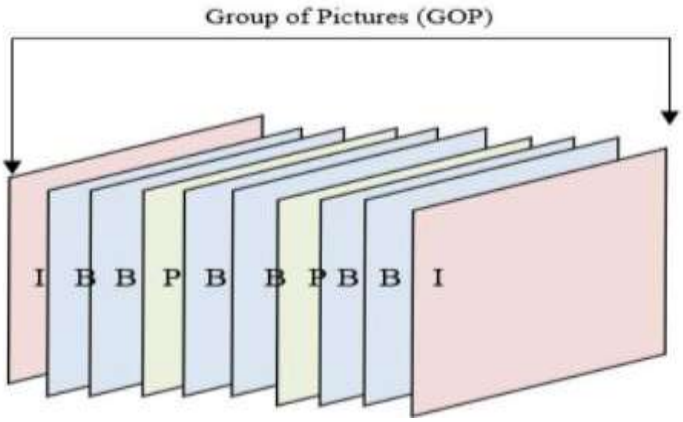

Figure 1 MPEG display order

- Fade: This can be fade-in or fade-out. The fade-in happens when the image is displayed from a black image and fade-out when an image fades to a black screen.
- Dissolve: When fade-in and fade-out occurs one after another.
- Wipe: When one frame wipes another frame, means during transition a virtual line going across the frame clearing the old one and displaying the new screen.
- old one and displaying the new screen.

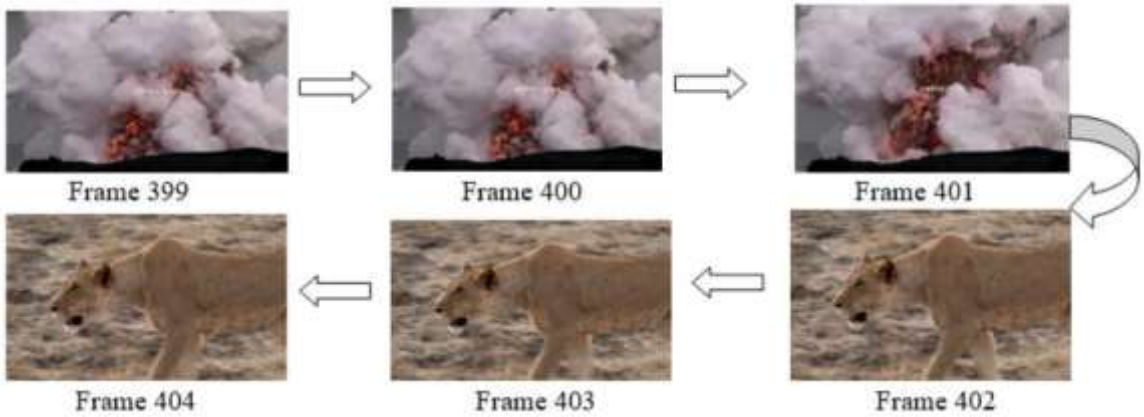

Figure 2 Example of Abrupt Scene Change

One of the important parts of video processing is to find the place where scene change happens. This is called scene change detection. There are four algorithms are available for scene change detection. Author has tested the frames by mean absolute frame differences (MAFD) [10]. Author has computed correlation between histogram of reference frame and all other frames [11]. If change is sharp, then the correlation factor is high and for no scene change the factor is zero.

III. Proposed Method

In this paper a novel technique of video steganography has been proposed which is validated and verified by thorough testing. The block diagram for Video Steganography process is shown in figure 3.

This testing needs a range of video which are not copyright protected. Since Videvo video database is a stock of royalty free science and technology video footage, hence in this research, Videvo has been used as video database. For testing, HD videos have been chosen having certain values in video features like resolution, aspect ratio and frame rate. Resolution of frame is width x height, means number of pixel in width and height. It signifies high quality of video, higher the resolution better the quality. Aspect Ratio is the ratio of width to height. It is an important feature which affect the feel of the video. The frame rate is the number of frames that are displayed per second, often measured in fps (Frame per second). The higher the fps, the smoother the video appears to the user. For this experiment, Full HD video footages have been used as cover which have following features:

- Resolution: 1920x1080
- Aspect Ratio: 16:9
- Frame Rate: ~30 fps
- Duration: Less than a Minute

A video can be treated as sequence of frames. Each frame is nothing but a still image. Rapid succession of collection of frames creates impression of movement. Hence in video steganography, at first cover video has divided into number of frames. Then one of the frames is selected for data hiding, which would be known as Stego frame once embedding is completed. Now the question arises how that Stego frame is chosen to hide data. There is a concept called Shot transition detection (simply, shot detection) or cut detection or scene change detection or scene boundary detection in video processing. A shot or scene or cut is series of frames shot

at a stretch with one camera. There are several existing methods for scene change detection. Although all these algorithms generally follow two steps:

- Scoring: A score which is basically a number of excellence is calculated between a pair of frames (framei and framei+1). These scores can be evaluated through Histogram Differences, Sum of Absolute Differences and Edge Change Ratio generally [10].
- Decision: The frame which scored highest in scene change detection has been chosen for further processing.
- In [9,11]], authors have demonstrated a new technique of shot transition detection using DCT coefficient of two consecutive frames.

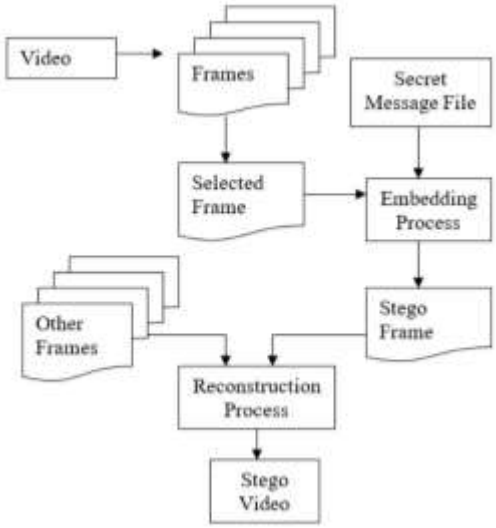

Figure 3 Block diagram technique for video steganography

Here the secret image is processed through Arnold Transform which creates scrambled image. The advantage of using Arnold Transform in steganography is that it rearranges the pixel in such a way that random cropping Steganalysis attack on Stego object doesn't destroy the secret image. Figure 4 and Figure 5 demonstrates the approach of current proposed technique of video steganography.

Once shot transition frame is detected embedding algorithm is applied to hide the scrambled image. In this proposed method colour video is used as cover, where every frame is a true-colour image. A true-colour image stores its Red, Green and Blue pixel values in three different matrices. Hence a true colour image has the dimension of mxnx3 where m is the size of length and n is the size of breadth. A true-colour image often called RGB image. RGB space is shown in Figure 6.

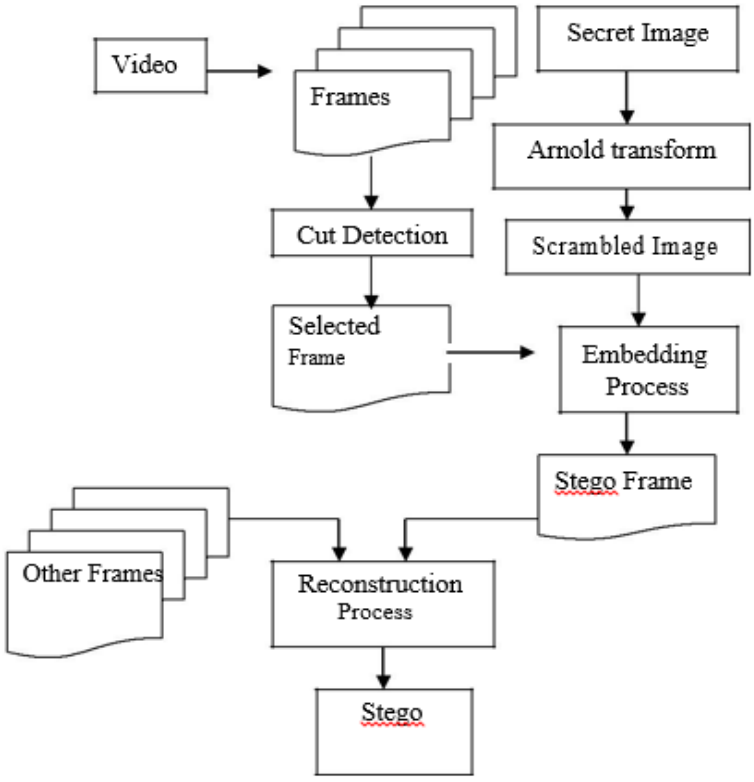

Figure 4 General Block Diagram of embedding algorithm of proposed system

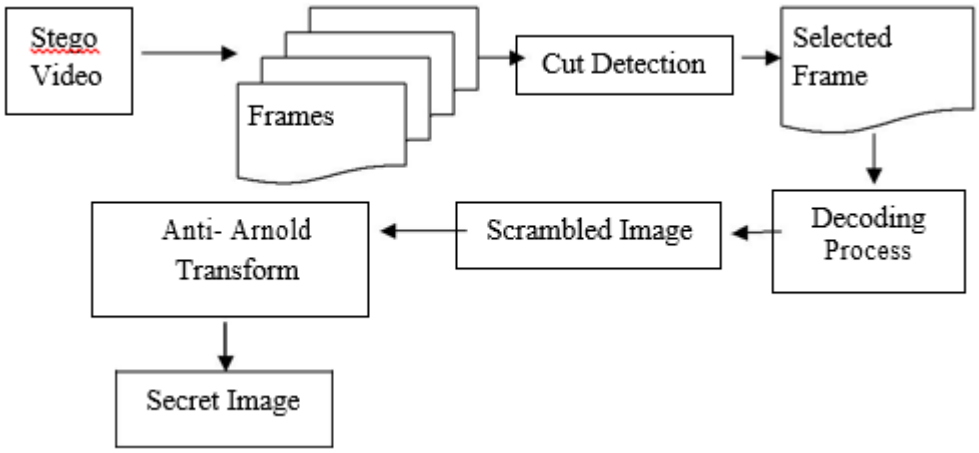

Figure 5 General Block Diagram of decoding algorithm of proposed system

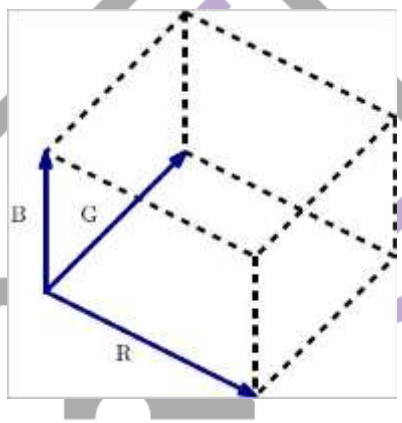

Figure 6 RGB space representation

In this proposed method, chosen frame which scored highest in scene change detection process, is itself a true colour image where secret data embedding takes place; that RGB image can be divided into Red, Green and Blue plane. Among these three image planes, blue light has short wavelengths. Hence, this is undetectable by human eyes and that's why blue colour plane has been selected for data hiding.

Initially two-dimensional DWT is applied on the blue plane matrix which divides frequency component into four sub-bands: LL, LH, HL and HH. In the proposed method except LL any other sub-bands can be chosen for data hiding. It has been seen that HL sub-band is more suitable for data hiding. Since, HL includes information regarding mid-frequency range used by low-pass and high-pass filters. Therefore, data hiding at HL sub-band assures that the proposed method will be resistant to the filtering Steganalysis attack.

Then, HL matrix is divided into 4x4 non-overlapping blocks onto which DCT is performed. The advantage of applying block DCT on image against standard DCT is - it is computationally efficient and less costly in terms of memory and hardware.

Now it needs to be find out the mid-band frequency region in 4x4 DCT block to embed secret data. In this proposed method, secret image is scrambled by Arnold transform followed by that scrambled image has been embedded into mid-band frequency region of 4x4 DCT block of cover using pseudo-random number. During secret extraction process, similar scene change detection algorithm based on differences in DCT coefficient has been performed on the Stego video to identify the Stego true-colour image frame. Then that RGB image frame is divided into Red, Green, and Blue plane. Next, two-dimensional DWT is applied on the Blue plane to choose HL sub-band which would be further divided into 4x4 blocks and DCT is applied on each block. Thereafter correlation coefficient is calculated between DCT coefficients

IV. Analysis of Results

This algorithm has been tested over different sets of HD video. However here only four sets of test results have been shown Table 1 and Table 2; Table 3 shows the set of original and extracted image. The quality of the proposed method is analyzed through Peak Signal to Noise Ratio, Normalized Cross-Correlation, Average Difference, Structural Content, Maximum Difference, Normalized Absolute Error and SSIM.

Table 1 Quality Analysis of Results for Cover vs. Stego HD Video

Full HD Cover Video	Video #1	Video #2	Video #3	Video #4
Mean Square Error	0.5342	0.6604	0.4468	0.2129
Peak Signal to Noise Ratio	50.8535	58.1707	53.2739	50.5931
Normalized Cross-Correlation	1.0000	1.0027	1.0007	1.0002
Average Difference	0.0037	-0.5179	-0.1259	-0.3446
Structural Content	1.0000	0.9940	0.9978	0.9988
Maximum Difference	4	13	12	15
Normalized Absolute Error	0.0024	0.0146	0.0218	0.0244
Structural Similarity Index	0.9184	0.8989	0.5083	0.6512

Table 2 Quality Analysis of Results for Original vs. Extracted Secret

Secret Binary Image	Image #1	Image #2	Image #3	Image #4
Mean Square Error	0	0	0	1.2346e-04
Peak Signal to Noise Ratio	99	99	99	87.2157
Normalized Cross-Correlation	1	1	1	1
Average Difference	0	0	0	-1.2346e-04
Structural Content	1	1	1	0.9998
Maximum Difference	0	0	0	0
Normalized Absolute Error	0	0	0	1.5444e-04
Structural Similarity Index	1	1	1	1.0000

Table 3 Images for Original Secret vs. Extracted Secret

Original Secret				
Extracted Secret				

V. CONCLUSIONS

In this paper a blind, key based, novel technique of video steganography has been discussed, by utilising DWT and DCT together to hide secret image data. The probable all perceptual metrics has been analysed to validate the outcome of the method with respect to imperceptibility. Moreover, all relevant Steganalysis attacks has been performed to show the robustness of the proposed method. The capacity of the method also can be improved with the size of video resolution.

REFERENCES

[1] Dr. M. U. Umamaheshwari, "Analysis of different steganographic algorithm for secured data hiding" , International journal of computer science and network security , 2010.

[2] Mrudul Dixit, Nikita bhide, Sanika Khankhoje, "Video Steganography", International conference on pervasive computing, 2015.

[3] Tarik Faraj idbea, salina abdul samad, "An adaptive compressed video steganography based on pixel value differencing schemes", international conference on adaptive technologies for communication, 2015.K. Elissa, "Title of paper if known," unpublished.

[4] Pooja yadav, Nischol Mishra , Sajneev sharma, " A secure video steganography with encryption based on LSB Technique", IEEE international conference on computational intelligence and computing research, 2013.

[5] Vladimir Hajduk, Martin Broda , "Image steganography with using QR Code and cryptography", 26th conference on Radioelektronica, 2016.

[6] Rajesh G.R, A. Shajin Narguman , "Steganography algorithm based on discrete cosine transform for data embedding into raw video streams", Chennai fourth international conference on sustainable energy and intelligent system, 2013.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] Ramadhan mastafa, Khaled M. Elleithy, "A DCT-based robust video steganography method using BCH error correcting codes", IEEE, 2016.

[8] Essam H. Houssein , Mona A. S. Ali, Aboul Ella Hassanien , "An image steganography algorithm using Haar discrete wavelet transform with Advanced encryption algorithm", Federated conference on computer science and information system, 2016.

[9] B. Bing, "In Video Coding Fundamentals, Next-Generation Video Coding and Streaming", John Wiley & Sons, 2015.

[10] Xiaoquan Yi, & Nam Ling, " Fast Pixel-Based Video Scene Change Detection", IEEE, 2005.

[11] Radwan, N. I., Salem, N. M., & El Adawy, M. I. , " Histogram Correlation for Video Scene Change Detection. In D. C. Wyld, J. Zizka, & D. Nagamalai (Eds.)", Advances in Computer Science, Engineering & Applications, 2012.