Video Watermarking using DWT and Elgamal for Authentication and Security

Shaik Hedayath Basha[1], U.Karthik[2] D.Mahendra[3], M.Sai Harsha[4] E.Nirmal Raj[5]
[1]*Assistant Prof.,* [2,3,4,5] *UG Students, ECE Department,*
RMK College of Engineering and Technology
Chennai, India.

Abstract

In the present scenario there is a requirement to merge watermarking with cryptography for better authentication and security. The proposed work is carried out based on the above requirement and it is done in two folds. The host video is initially divided into frames then divided into two folds as static frames and dynamic frames. In static frame watermarking process is carried out the watermark logo is embedded in the chosen static frames using Haar discrete wavelet transform and in dynamic frame Elgamal encryption process is used for security purpose. The frames were joined to obtain the embedded video, the frames were subjected to various attacks the PSNR and Correlation coefficient factor is obtained for analyzing the quality of the video.

Keywords

video watermarking; elgamal encryption; PSNR; attacks; security

I. INTRODUCTION

Watermarking process is one of the techniques to protect the authentication of the multimedia content. The multimedia content can be audio signal, image file, video file or 3D video file. The generic watermark embedding and recovery schemes are shown in the below figure 1 and figure 2. [1]

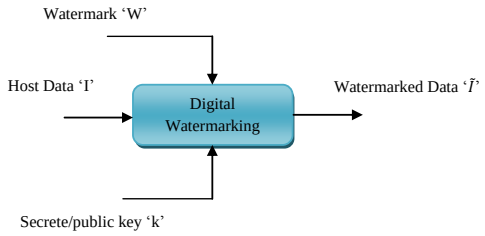

Fig 1 Generic digital watermark embedding scheme

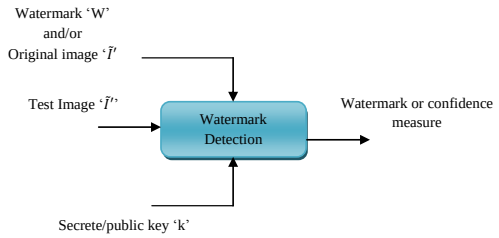

Fig 2 Generic digital watermark recovery scheme

There are four different types of watermarking systems as discussed in [1] the first type is private watermarking, second type is semi-private watermarking, third type is public watermarking and fourth type is asymmetric watermarking. In general the digital content of the above formats can easily be modified and it can misuse by any person with good knowledge in computers. The watermarking system can provide better authentication and copy right protection with certain limitations. The vulnerability of the system can be decreased with the help of good security. The basic concept behind watermarking is that the watermark logo which is to be embedded in the multimedia signal need to have one property that the watermark cannot disturb the quality of the multimedia signal. The property is watermark size need to small compared with the host signal (multimedia signal), here size is referred as payload of the watermark, therefore the payload of the watermark needed to be less. Next point to concentrate is the visibility of the watermark, the watermark can be visible or it can be invisible it depends up on the requirement of the user. But it is good to have imperceptible watermarking so that the attacker feels that the watermark is not present. Last important point to concentrate is that the logo should be robust to recover for the adversary. The probability of detecting the logo by the adversary need to very low that is the logo need to be secured. There is a trade-off between data payload, imperceptibility and robustness. If data payload is increased the logo will be less imperceptible and the adversary attacks will increase. In any watermarking system the trade-off between the above said parameters should be compromised to attain proper security and authentication.

In this paper second type of watermarking system that is semi-private watermarking system is chosen for implementation in MATLAB and in future thinking to implement in FPGA using VHDL.

Here the paper is organized with literature survey which gives different ideas for implementing the work, next followed with proposed work, then Elgamal encryption process is discussed, at last the results were discussed.

II. LITERATURE SURVEY

An algorithm to find video clips with dissimilar temporal durations and some spatial

variation, a longest common sub-sequence (LCS) matching method for measuring the temporal similarity between video clips based on the measure propose three techniques are present to improve the retrieval efficiency. First, a small amount of coefficients in the low frequency region of DCT block is as the basis to represent spatial features is presented. Second to determine a suitable quantization step-size for visual features to obtain better tolerate spatial variations of similar video clips and propose a paired quantized method. Third, the compactness and/or continuity of matched common sub-sequences in the LCS measure to better reflect temporal characteristics of video. The performance of the proposed algorithm shows an improvement of 63.5% in terms of MAP (mean average precision) as compared to previous algorithm. The results show that this approach is effective for news video retrieval [2]. An Improved DCT domain block video watermarking scheme based on the video sequence characteristics fully based on human visual system and moving object detection technology, choosing the movement and complex field as embedding regions. The watermarking embedding process uses the sub pixel blocks to reduce the block effects of pixel classification [3]. In video surveillance scenario with real-time moving object detection and tracking. The detection of moving object is important in many tasks, such as video surveillance and moving object tracking. The design of a video surveillance system is intended for automatic identification of events of interest, especially on tracking and classification of moving objects. Normally a video surveillance system combines three phases of data processing.

1. Moving object extraction.
2. Moving object recognition and tracking and decisions about actions.
3. The extraction of moving objects, followed by object tracking and recognition, can often be defined in very general terms.

This survey reviews briefly research works on object detection and tracking in videos. The definition and tasks of object detection and tracking are first described, and the important applications are mentioned [4]. The improvement of information security, the conventional image encryption algorithm has been far from to ensuring the security of images in the transmission process. This presents a new image encryption algorithm, which can improve the security of image during transmission more effectively. The traditional scrambling algorithm based on Arnold transformation only applies to the square area, which is a big limitation. Focus on this, a multi-region algorithm for image scrambling encryption model is proposed, which splits the non-square image to multiple square regions, and scrambles each region. Experimental consequences show that the new algorithm improves the image security effectively to avoid deciphering, and it also can restore the image as same as the original image, which reaches to the purposes of image safe and reliable transmission [5].

Motion analysis is based on the extraction of the human target. The purpose of motion target detection is extract the human target for the background from a video sequence, requiring extract the complete human outline and reflect the morphological characteristics of the human target. Effective extraction of motion targets is very important for the next steps of action recognition. Here are several motion object detection methods: background subtraction method, temporal difference method, optical flow method. Background subtraction method is the most common method in moving object segmentation of image. This method is more suitable for a fixed camera position. The basic idea is to use the current video frame, and the background model which had been constructed to differential and threshold, detect the human region, the advantages of this algorithm is complete human region. However, this method has large amount of calculation and a high computational complexity. The background subtraction method generally includes four steps: background modeling, background updating, background difference and post-processing. The database of this article is in the case of a fixed camera position the light would not change much, so the background subtraction for target detection is choosing. Optical flow method is to use time domain variation and correlation of pixel intensity data in the image sequence to determine the movement of the respective pixel position. The advantage of this method is that under the premise of movement of the camera, it can detect the independent movement of the target. The drawback is that the calculation is quite complex, and anti-noise performance is poor. Temporal difference method is extraction the difference between the two or three adjacent frame of a video sequence, to extract the information of motion object through difference and threshold [6]. Moving object detection in a video sequence is one of the most challenging research topics due to its various applications and complexity. Sometimes, it is also required to identify a moving object without any previous information about its shape or motion. Here a global mesh generation technique which can construct a hierarchical mesh on moving object(s) inside a video frame. This technique is efficient enough to solve the problem to a great extent and can identify a moving object in a video frame without any previous knowledge. The generated mesh is triangular and probabilistic in nature. The granularity of the mesh depends upon the level of hierarchy, neighborhood triangle decomposition, gradient and temporal information of the video sequence. Experiments show that it can identify the moving object(s) with good accuracy without any pre or post-processing steps [7].

In video surveillance, there are many interference factors like target changes, complex scenes, and target deformation in the moving object tracking so

to resolve this issue based on the comparative analysis of several common moving object detection methods, a moving object detection and recognition algorithm combined frame difference with background subtraction is used where calculated the average values of the gray image of the continuous multi-frame image present in the dynamic image, and then obtained background image by the statistical average of the continuous image sequence i.e., the uninterrupted interception of the N-frame images are added and the average is calculated. In this case, weight of object information has been increasing, and also restrains the static background. Eventually the motion detection image contains both the target contour and more target information of the target contour point from the background image, so as to achieve separating the moving target from the image. The simulation outcome shows the effectiveness of the planned algorithm [8].

III. Proposed Work

A. *Discerte Haar Wavelet Transform*

In general wavelets are special functions which exhibit oscillatory behavior for a short period of time and then die out. It uses a single function, its dilations and translations to generate a set of orthonormal basis function to represent a signal. Numbers of such functions are infinite and it is user responsibility to choose the required functions with the prior knowledge of application undergoing more iteration, so fortunately there are two special functions called Haar Wavelet and Scaling function which have explicit expression, so Haar Wavelet has been chosen in this thesis for simplicity. Haar scaling function defined as in equation 1 and it is represented in figure 3.

$$\phi(t) = \begin{cases} 1 & 0 \le t \le 1 \\ 0 & elsewhere \end{cases} \quad (1)$$

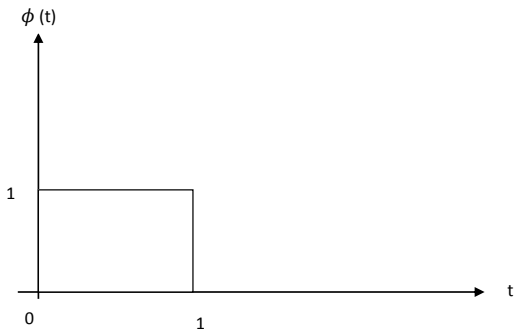

Fig 3 Haar Scaling function

Haar Wavelet function is defined as in equation 2 and it is represented in figure 4.

$$\psi(t) = \begin{cases} 1 & 0 \le t \le 1/2 \\ -1 & -1 \le t \le 1 \\ 0 & elsewhere \end{cases} \quad (2)$$

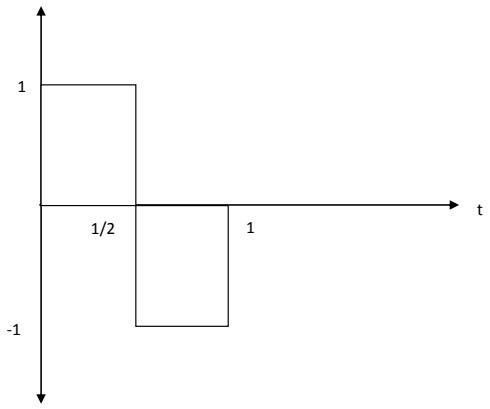

Fig 4 Haar Wavelet function

The Haar wavelet transformation is self-possessed a sequence of Low-Pass and High-Pass filters, known as a Filter Bank. These filter sequence can be applied in the same way as a discrete FIR (Finite Impulse Response) filter in DSP (Digital Signal Processing), as multiple successive FIR filters. The Low Pass filter performs an Averaging/Blurring operation, and is expressed in equation 3 and the High-Pass filter performs a differencing operation and can be expressed in equation 4.

$$H = \frac{1}{\sqrt{2}}(1,1) \quad (3)$$

$$G = \frac{1}{\sqrt{2}}(1,-1) \quad (4)$$

The complete wavelet transform can be represented in matrix format and it is expressed in equation 5.

$$W_N = \frac{H}{G} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (5)$$

The above equation 6 is used of transforming a 4x4 pixel image. The equivalent matrix can be expanded for larger images.

B. Video Watermarking Process

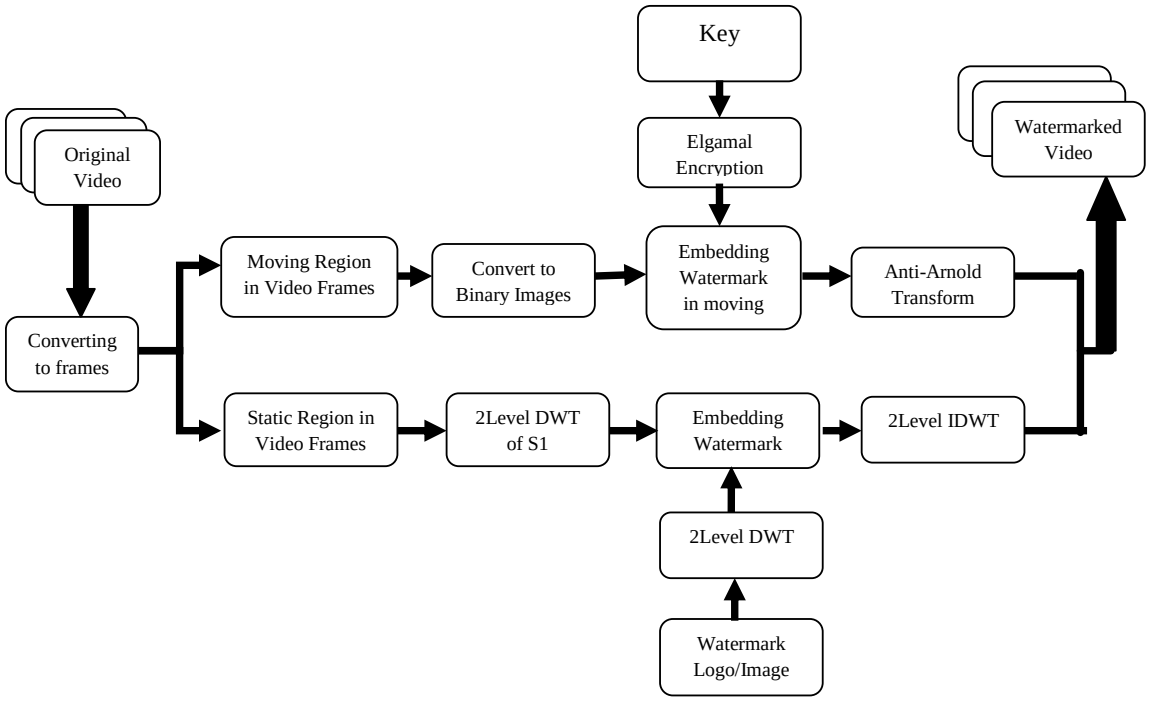

Fig 5 Video Watermarking Embedding Block diagram

In this proposed work the video is divided into static frames and dynamic frames using inter frame difference method. The block diagram is shown in the figure 5. The logo size is 80 X 90 and is shown in figure 6.

Fig 6 Logo image 80 X 90 size

The one level DWT is performed on the blue frame of the logo is shown in the figure 7.

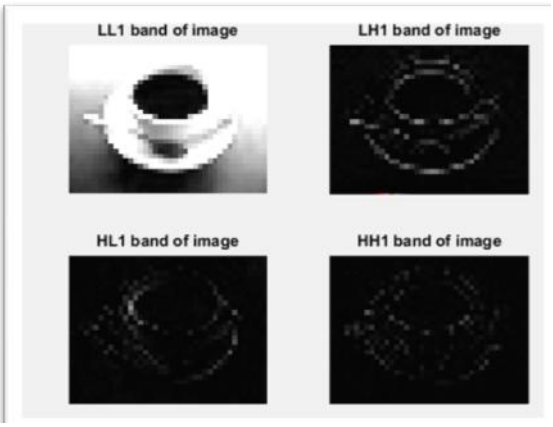

Fig 7 One Level of blue frame of Logo

The video "Rhino.avi" is taken for embedding logo and performing the elgamal encryption in the dynamic frames. The first frame is shown in the figure 8.

Fig 8 First frame of the Rhino.avi video

The embedding process is as follows the video is divided into frames, then the successive frame difference is applied to obtain static and dynamic frame. In static frames the logo is embedded in only blue frame because according to HVS red and green color frames are sensitive to any small changes compared to blue frame. Selective frame are undergone for embedding logo. Two level DWT is performed on the selective frames similarly the logo after one level DWT is embedded in the LHL band of selective frame. The elgamal encryption is performed and added in the dynamic frames. The advantage is that the embedded video will work only with the key.

IV. RESULTS

The video chosen for embedding is converted into frames using MATLAB and obtained 114 frames from the 9 seconds video with 14fps; frames of the video are shown in figure 9.

Fig 9 Frames from the chosen video

The frames are then divided into two folds as static and dynamic frames the static frames are shown in figure 10.

Fig 10 Static Frames of the chosen video

The dynamic frames are shown in the figure 11.

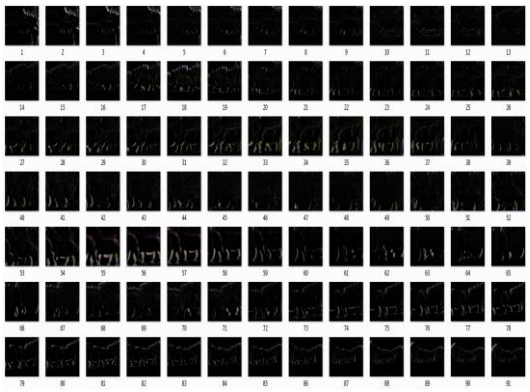

Fig 11 Dynamic frames of the chosen video

The watermark is embedded in every 8th frame, that is first frame is taken for embedding and then 9th frame is chosen, then 15th frame and so on. So, totally 15 frames are selected for embedding the logo.

The chosen frames are shown in the below figure 12.

Fig 12 Selected 15 frames for embedding logo

The peak signal to noise ratio (PSNR) values of all 15 frames after embedding logo is shown in the below table I. Correlation coefficient of original and embedded logo is shown in table II.

TABLE I. PSNR OF SELECTED FRAMES

Frame No	PSNR in dB	Frame No	PSNR in dB	Frame No	PSNR in dB
1	61.9982	39	62.9767	79	61.9424
9	61.9618	47	62.8508	85	61.4253
15	62.3361	55	62.6689	93	61.6211
23	62.9275	63	62.0152	101	61.7182
31	62.9340	71	61.7462	109	61.8475

TABLE II. CORRELATION COEFFICIENT OF LOGO

Frame No	PSNR in dB	Frame No	PSNR in dB	Frame No	PSNR in dB
1	0.95	39	0.99	79	0.965
9	0.94	47	0.976	85	0.954
15	0.97	55	0.967	93	0.963
23	0.977	63	0.965	101	0.959
31	0.98	71	0.954	109	0.968

Attacks like resizing, rotation, salt and pepper noise addition, cropping, horizontal shear and image adjustment is performed on the video and the logo is obtained after the attacks and it shown in the Table III.

TABLE III. MSE AND PSNR OF LOGO WITH ATTACKS

S.No	Attack	MSE	PSNR
1	Resizing	0.2915	53.4852
2	Rotation with 10 degrees	1.2342	47.2169
3	Salt and Pepper Noise	0.2915	53.4852
4	Cropping	0.2915	53.4852
5	**Horizontal shear**	**0.6813**	**49.7974**
6	Adjust	0.000074	79.5778

The Horizontal shear attack decreases the quality of the logo and even the frame in the other hand image adjustment attack has less effect on the quality. The remaining attacks have moderately better PSNR values.

V. CONCLUSION AND FUTURE SCOPE

The embedding is properly implemented and the blue frame is taken for embedding logo which increases the PSNR of the embedded frames. If red or green frames are taken then the quality of embedded video would be decreased. In future the video embedding can be better implemented in the hardware using FPGA and VHDL. The security can be increased by better encryption technique like elliptical curve cryptography.

REFERENCES

[1] M. Kutter and F. A. P. Petitcolas A Fair benchmarking for Image Watermarking systems, Elecronic Imaging'99, Security and Watermarking of multimeda contents, vol. 3657, Sans Jose, CA, USA, 25-27 January 1999. The International Society for optical Engineering.

[2] Young-tae, Kim and Tat-Seng, Chua -"Retrieval of News Video using Video Sequence Matching", Proceedings of the 11th International Multimedia Modeling Conference, IEEE Computer society 2005lerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] X. Jing et al., - "The Domain Block Video Watermarking scheme based on Video sequence' Characteristics and DCT", 7th International Conference on Computer science & Education, pp. 448-452, July 14-17, 2012.

[4] "Study of Moving Object Detection and Tracking for Video Surveillance, International Journal of Advanced Research in Computer Science And Software Engineering, Volume 3, Issue 4, April-2013.

[5] Min Li, Ting Liang and Yujie He, - "Arnold Transform based Image Scrambling method", 3rd International Conference on Multimedia Technology, 2013

[6] Jing Cao, Dong Zhang, Dong Fang Wang, - "Background Subtraction based Human Region Extraction Method", 3rd International Conference on Multimedia Technology, pp. 430-437, 2013.

[7] Kalyan Goswami et al., - "A Novel Mesh-Based Moving Object Detection Technique in Video Sequence", Journal of Convergence, Volume 4, No. 3, September 2013.

[8] Kuihe Yang, Zhiming Cai and Lingling Zhao, - "Algorithm Research on Moving Object Detection of Surveillance Video Sequence", Optics and Photonics Journal, 308-312, 2013.