Privacy Preservation of Video using Sample Image and SVM Classification on Video frames

[1]Sayyada Fahmeeda Sultana, [2]Dr. Shubhangi D C
[1]Asst. Professor, [2]Professor
[1]PDA College of Engineering,
[2]PG-Center,VTU RO, Kalaburagi, India
[1]https://orcid.org/0000-0001-8265-2237

Abstract : Privacy Preservation of video provides confidentiality and prevents unauthorized access of the content. Real time constraints, large amount, and unique characteristics of video data inhibits the use of traditional cryptographic algorithms over video data. Frames which require more security need to be encrypted, to search such frames content based video frame retrieval is performed in the proposed system using sample image and support vector machine (SVM). The frames which require more security can be searched using this method then on the retrieved frames fully homomorphic encryption is performed through the proposed Trigonometry based video frame encryption and decryption algorithm, which can encrypt multiple number of pixels at a time to reduce the encryption time and reduce the number of keys required to ¼ of other state of art encryption techniques like AES, DES, RSA and produce the same result.

***IndexTerms* - Trigonometry Based Encryption,SVM,Content Based Retrieval,Video Frame Encryption,Video.**

I. Introduction

Recent development in Cloud Computing technologies has resulted easy storage and transmission of huge amount of multimedia data towards the cloud. Multimedia data like videos require privacy, confidentiality or authorization constraints on their content and therefore need protection. Conventional cryptographic algorithms (RSA, DES, AES, etc.) previously designed on text data may not be suitable for multimedia applications due to their large data size, data redundancy and real time constraint [1] with limited processing power, memory and bandwidth. Video cryptosystems for real-time applications depends on performance parameters such as security, computational efficiency, compression efficiency and format compliance [2]. Selective frames to be encrypted may be better solution. To meet this requirement the proposed system perform content based retrieval on video frames based on sample image using SVM such that the frames which have critical confidentiality requirement are retrieved by giving sample image, then on the retrieved video frames the proposed trigonometry based video frame encryption algorithm is applied. The proposed trigonometry based video frame encryption algorithm is a light-weight technique especially for cloud storage.

The Content Based Image Retrieval uses image content to retrieve the digital images from huge database of images. Content based image retrieval is a technique for retrieving semantically-relevant images from an image database based on automatically-derived image features. In CBIR system, image features are categorized in three main classes: color, texture and shape. Color is the most common visual feature used in CBIR as it is simple to extract the color information from images. To extract information about shape and texture feature is complex usually performed after the initial filtering provided by color features.

The objective of support vector machine algorithm is to find a hyperplane in an N-dimensional space(N—the number of features) that distinctly classifies the data points to separate the two classes of data points.

Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line. If the number of input features is 3, then the hyperplane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, Maximize the margin of the classifier. Deleting the support vectors will change the position of the hyperplane. These are the points that lead to the use of SVM in our proposed system.

Rest of the paper is organized as Related work is presented in Section II, Section III gives the detail of methodology used in the proposed system, Section IV presents the experimental results and there analysis, concluding remarks are given in Section V.

II. Related Work

Privacy preservation of videos on cloud can be performed using different encryption methodologies. Fully layered encryption scheme, the whole content is first compressed, then, the compressed bit-stream is entirely encrypted using a standard cipher DES[3] or AES [4]. Naïve Technique is a straight forward method to encrypt every byte in the whole Moving Picture Expert Group (MPEG) stream using standard encryption schemes such as DES or AES. The idea of naïve algorithm is to treat the MPEG bit-stream as text data and does not use any of the special structure [5]. Permutation based Encryption works mainly on achieving visual degradation on permutation principle. Pure Permutation: Pure permutation algorithm scrambles bytes within a frame by permutation. It is extremely useful in situation where the hardware decodes the video, but decryption must be done in software

[6]. Zig-Zag Permutation: Zig-Zag permutation [7],instead of mapping the 8X8 block to 1X64 vector in “Zig-Zag” order, it maps individual 8x8 block to 1x64 vector by using a random permutation list (secret key). Key based Permutation: block-based video encryption algorithms using Faro IN OUT Shuffle and rotation techniques for real-time video application. They proposed two algorithms in which each video frame is passed through for key generation and scrambling. The algorithms rotate the first video frame by an angle then key is generated based on the block size using Faro IN OUT perfect. When compared with random permutation, their proposed method provides more scrambling of video frame[8]. Huffman Codeword Permutation It is a lightweight mpeg video encryption which incorporates encryption with MPEG compression in one step [9]. Compression Logic based Random Permutation The proposed algorithm is Compression logic based video encryption algorithm [10].

III. To retrieve videos from database, effective video analyzing, indexing and retrieval techniques are required. Video retrieval using query-by-image is not successful as it gives result of videos with less relevancy and accuracy. A method is proposed where input is a video clip, to achieve the high quality of content based video retrieval by discovering the temporal patterns in the video contents[11]. video structure parsing, content parsing, content representation, and content based video retrieval and browsing technology[12].

IV. Methodology

The proposed method for Privacy Preservation of video using sample image system is shown in Figure 1. Video to be encrypted is converted into frames stored as database, the feature vectors are extracted from the frames in the database and described by multidimensional feature vectors, which form a feature database. To find which frame to be encrypted from the database, the feature vectors are extracted from the given sample image. Similarities between the feature vectors of the Sample image and the feature vectors of the database frames are then calculated. The retrieval is performed with the aid of Sample Image similarity to database video frames and SVM method. And, the last stage of the proposed method is to encrypt the retrieved video frame using the proposed Trigonometry based encryption Algorithm.

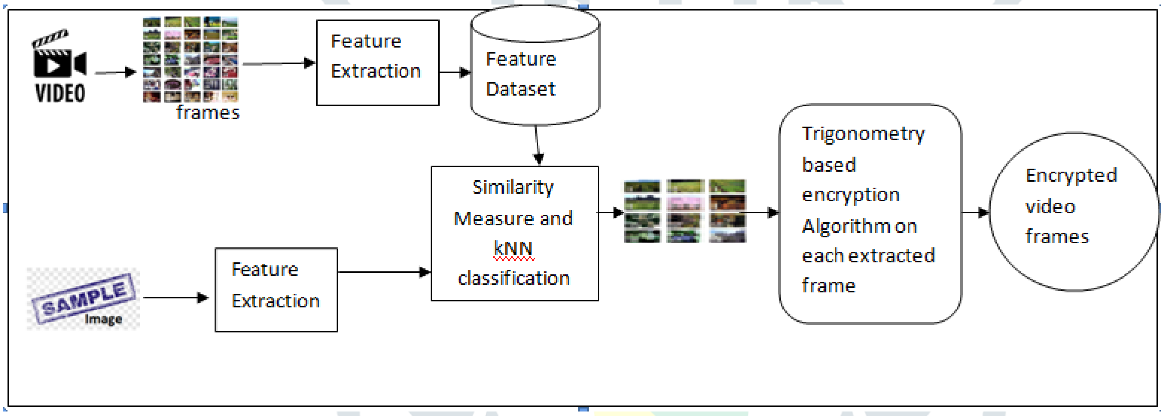

VIDEO
frames
Feature Extraction
Feature Dataset
Trigonometry based encryption Algorithm on each extracted frame
Encrypted video frames
Similarity Measure and kNN classification
SAMPLE Image
Feature Extraction

Figure 1. Block Diagram of Proposed Methodology

The frame work for the proposed system Privacy Preservation of video using sample images includes the following steps

A. Feature Extraction of video frame and SVM

Feature Extraction of video frame is the process by which key features of the sample are selected. Typically, the process of feature extraction relies on the set of following features for feature dataset creation and sample image feature extraction. The feature set of each frame is created using Color Histogram, Color Auto-Correlogram, Color Moments, Gabor Wavelet, and Wavelet Moments. All the extracted features are concatenated to form 190 dimensional feature vector.

Color histograms [13] are defined as a set of bins where each bin denotes the probability of pixels in the image being of a particular color. A color histogram for a given image is defined as a vector:

$$CH = \{ch[1], ch[2], \ldots ch[i] \ldots, ch[N]\} \qquad (1)$$

Where i represents a color in the color histogram and corresponds to a sub cube in the HSV color space, H[i] is the number of pixels in color i in that image, N is the number of bins in the HSV histogram and each H, S,V component is uniformly quantized into 8,2,2 bins for HSV respectively, making 32 dimensions out 190 dimension feature vector.

A color correlogram expresses how the spatial correlation of pairs of colors changes with distance. A color histogram captures only the color distribution in an image and does not include any spatial correlation information.

The correlogram (γ) and auto-correlogram (α) of the frame I is defined for (g_i, g_j) at a distance l using equ (2)[14].

$$\alpha^{l}_{g_{i,g_j}}(I) = \gamma^{l}_{g_{i,g_j}}(I) = Prob\left|P_2 \in I_{g_i}\right| \quad where\ |P_1 - P_2| = l \quad P_1, P_2 \in I \qquad (2)$$

Where probability that given any pixel P_1 of level g_i, a pixel P_2 at a distance l in certain direction from the given pixel P_1 is of level gi. Using equ(2) image is quantized into 4X4X4 =64 colors in RGB space, making 64 dimensions out of 190 dimensions.

Color moments are measures that can be used differentiate images based on their features of color. The first two moments from the RGB plans are extracted using Mean and Standard deviation given by equ(3),(4), respectively.

$$Mean_i = \sum_{n}^{j=1} \frac{1}{n} P_{ij} \qquad (3)$$

$$Std(\sigma_i) = \sqrt{\frac{1}{n} \sum_{n}^{j=1} (P_{ij} - Mean_i)^2} \qquad (4)$$

Mean is the average color value in the image, and Standard deviation is given by σ is the standard deviation. Generating 6 dimensions out of 190 dimensions.

Gabor wavelets in image processing algorithms for interest point detection, Gabor wavelet filters spanning four scales: 0.05, 0.1, 0.2, 0.4 and six orientations$\theta = 0, \theta_{n+1} = \theta_n + \frac{\pi}{n}$ are applied to frames. Mean and standard deviation of Gabor wavelet coefficients are calculated to from feature vector add in 48 dimensions.

Applying the wavelet transform to the image with a 3-level decomposition, the mean and the standard deviation of the transform coefficients are used to form the feature vector. Zernike, PseudoZernike and Fourier-Mellin ones are defined in a continuous form equ(5), one can define the wavelet moments by replacing the function $T_n(r)$ with a wavelet basis functions.

$$W_{nm} = \iint T_n(r) e^{-jm\theta} f(r,\theta) r dr d\theta \qquad (5)$$

Since the dataset is constructed from a combination of features using similarity metrics to take advantage of features Relative L_1 defined in equ (6).

$$l_1 = \sum\nolimits_{i=1}^{n} \frac{|x_i - y_i|}{(1 + x_i + y_i)} \qquad (6)$$

Where x_i,y_i represents features from feature dataset and sample image features.

Support Vector Machine (SVM)

Support Vector Machine is a mathematically rigorous, machine learning technique to build a linear classifier. It creates a hyperplane in a high dimensional space that can accurately slice into two segments according to the desire objective.

The SVM algorithm consists of two phases:

a. Training phase:

The training dataset of n frames is given by

$$(X_1, y_1) \dots (X_i, y_i) \qquad (7)$$

Where X_i represents the feature vector for point I and y_i is its binary class value of 1 or -1. Thus there are two classes represented as 1 and -1.

Decision boundary should classify the points correctly, thus

$$l_{1_i} \geq 1 \; for \; all \; i \qquad (8)$$

b. Testing Phase

The resulting classifier is applied to unlabelled images to decide whether they belong to the positive or the negative category. The label of x is simply obtained by computing l_1 from equ(6)

B. Proposed Trigonometry Based Algorithm for Encryption video frame

The Frames retrieved from the above steps based on sample image under goes the execution of proposed Trigonometry based algorithm for encryption of selected video frames. The proposed algorithm is based on trigonometric rotation matrix based on trigonometric equ(9)

$$tan\theta = 1 - \sec\theta \qquad (9)$$

$$R = \begin{bmatrix} tan\theta & -sec\theta \\ sec\theta & -tan\theta \end{bmatrix} \qquad (10)$$

$$R^2 = \begin{bmatrix} tan\theta & -sec\theta \\ sec\theta & -tan\theta \end{bmatrix} \times \begin{bmatrix} tan\theta & -sec\theta \\ sec\theta & -tan\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (11)$$

Given an 2X2 matrix A=$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ multiply with R matrix in equ(10)

$$R' = A \times R \qquad (11)$$

R' represents encrypted A. To gain the original A use equ(12)

$$A = R' \times R \qquad (12)$$

Algorithm 1 shows the steps of encryption and Algorithm 2 shows the steps of decryption

```
    Algorithm 1: Trigonometry Based Encryption algorithm
    Input: Video frame $I$, Key $k_1, k_2, \dots, k_{M/4}$ where $M$ is number of row in video frame
    Output: Ciphered video frame $CI$
    Step 1: [m,n]=size(I)
    Step 2:
      $\mathbf{for\ k = 1:3}$
        $\mathbf{for\ i = 1:2:m-2}$
    Step 2.1: Find $R$ such that $R_i(1,1) = tan(k_i), R_i(1,2) = -sec(k_i),$
          $R_i(2,1) = sec(k_i), R_i(2,2) = -tan(k_i)$
          $\mathbf{for\ j = 1:2:n-2}$
     Step2.2  Select $A_{i,j,k}$ from $I$, where $A_{i,j,k}$ is 2X2 matrix of pixels from video frame
    Step 2.3: $CI_{i,j,k} = A_{i,j,k}XR$
          end
          end
          end
    Step 3: Return $CI$
    Step 4: Stop

    Algorithm 2: Trigonometry Based Decryption algorithm
    Input: Video frame $CI$, Key $k_1, k_2, \dots, k_{M/4}$ where $M$ is number of row in video frame
    Output: Ciphered video frame $DI$
    Step 1: [m,n]=size(CI)
    Step 2: for k=1:3
          $\mathbf{for\ i = 1:2:m-2}$
    Step 2.1: Find $R$ such that $R_i(1,1) = tan(k_i), R_i(1,2) = -sec(k_i),$
          $R_i(2,1) = sec(k_i), R_i(2,2) = -tan(k_i)$
          $\mathbf{for\ j = 1:2:n-2}$
     Step2.2  Select $A_{i,j,k}$ from $CI$, where $A_{i,j,k}$ is 2X2 matrix of pixels from video frame
    Step 2.3: $DI_{i,j,k} = CI_{i,j,k}XR$
          end
          end
          end
    Step 3: Return $DI$
Step 4: Stop
```

V. Experimental Result and Analysis

The result of proposed method is evaluated in two fold, firstly, Experimental results are presented and then evaluated on Content based retrieval using Sample image on Video frames, Secondly, the proposed Trigonometry Based Encryption and Decryption algorithm on extracted video frame.

Results and Analysis of Content based retrieval using Sample image on Video frames

This section shows, the proposed work is quantitatively validated using two performance metrics, namely, precision and recall. Precision is the ratio of number of relevant images retrieved to the total number of images retrieved, whereas recall is the ratio of number of relevant images retrieved to the total number of relevant images present in the database. Precision and recall can be calculated using equ (13) and equ (14) respectively.

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \qquad (13)$$

$$Recall = \frac{True\ Positive}{(True\ Positive + False\ Negative)} \qquad (14)$$

Where True Positive is number of correct frames retrieved, False positive is number of wrong frames retrieved, False Negative number of wrong frames not retrieved. The comparison of precision and recall of proposed Content based retrieval system with combination of features and SVM with other existing system is shown in table 1.

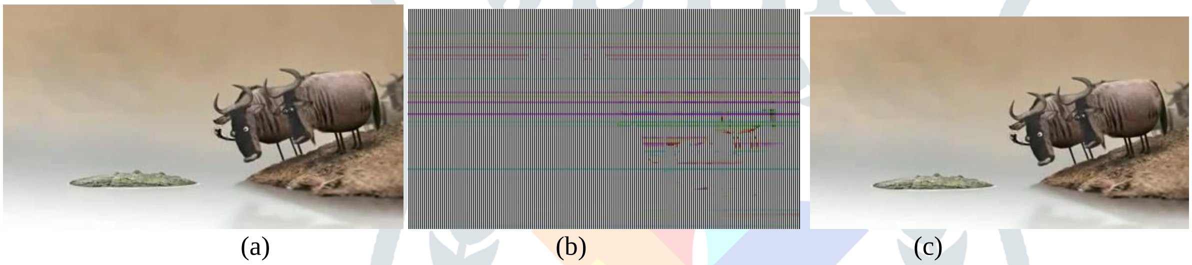
Table 1. Comparison of Existing and Proposed Technique Content Retrieval System

Performance Matrix	Existing System[15]	Proposed System
Recall	0.15	0.1
Precision	0.75	0.9

The proposed system of content based video frame retrieval on multiple dynamic feature along with SVM provides more accuracy then the other exiting system.

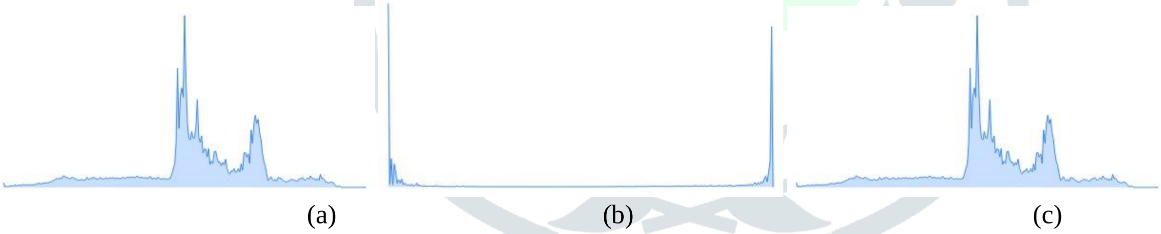
Result and Analysis of Proposed Trigonometry based video frame Encryption and Decryption Algorithm

The proposed Trigonometry based video frame encryption and decryption algorithm is a fully homomorphic encryption algorithm like other state of art AES, DES, RSA algorithm which provide more efficiency text data and are less efficient on multimedia data like video frames as the state of are method needs to encrypt single pixel at time with different key for each pixel, the proposed method provides the efficiency of fully homomorphic encryption by encrypting four pixels at a time and number of keys required is reduced to $\frac{number\ of\ row}{4}$ in image. Figure 2 shows the result of encryption on one of the video frame, figure 2a original video frame, figure 2b and 2c shows the encrypted and decrypted video frames respectively. Figure 3a,3b,3c shows the histograms of images in figure 2, as shown in figure 3a and 3b, the histograms of original and encrypted video frames are different indicating the difference in figure 3a and figure 3b.



(a) (b) (c)

Figure 2 (a) Original Video Frame (b) Encrypted Video Frame (c) Decrypted Video Frame



(a) (b) (c)

Figure 3 Histogram of (a) Original Video Frame (b) Encrypted Video Frame (c) Decrypted Video Frame

The proposed method of content based video retrieval based on sample image using SVM and proposed Trigonometry based video frame encryption and decryption algorithms have been tested on video clips taken from social media, youtube, and test video dataset. The Proposed algorithms found to be more efficient in terms of time, memory utilization and accuracy of Results to provide Privacy to videos.

VI. CONCLUSION

The paper proposes content based retrieval from video frames based on sample image and Trigonometry based video frame encryption and decryption algorithm. The Trigonometry based video encryption and decryption algorithm is fast and require only few steps for encryption, it provides ciphered image which will be totally different from original video frame. Proposed content based retrieval of frames based on sample image using SVM and then encryption of retrieved video frame allows to encryption only the required frames reducing the effort of encrypting all video frames.

REFERENCES

[1]. B Bhargava, C Shi, S Wang. (2004). MPEG Video Encryption Algorithms. Multimedia Tools and Applications. Kluwer Academic Publishers. https://doi.org/10.1023/B:MTAP.0000033983.62130.00

[2]. A. Kulkarni, S. Kulkarni , K. Haridas, A. More. (2013). Proposed Video Encryption Algorithm v/s Other Existing Algorithms: A Comparative Study. International Journal of Computer Applications. https://arxiv.org/ftp/arxiv/papers/1303/1303.3485.pdf

[3]. NIST: Data Encryption Standard , FIPS 46-3, 1999.

[4]. NIST: Advance Encryption Standard, FIPS 197, 2001.
[5]. Shiguo lian. (2008). Multimedia Content Encryption: Algorithms and Application. CRC Press. https://www.crcpress.com/Multimedia-Content-Encryption-Techniques-and-Applications/Lian/p/book/9781420065275
[6]. Adam J. Slaggel. (2004). Known-Plaintext Attack Against a Permutation Based Video Encryption. Algorithm. https://slagell.info/Adam_J._Slagell/Publications_files/slagell04a.pdf
[7]. L.Tang. (1996). Methods For Encrypting and Decrypting MPEG Video Data Efficiently. Proceedings of the Forth ACM International Multimedia Conference. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.8471
[8]. S. Sultana, Shubhangi D C. (2017). Video Encryption Algorithm and Key Management using Perfect Shuffle. Int. Journal of Engineering Research and Application. DOI: 10.9790/9622-0707030105
[9]. C. Shi and B. Bhargava. (1998). Light Weight MPEG video Encryption Algorithm. Proceeding of the International Conference on Multimedia. https://ieeexplore.ieee.org/abstract/document/740527
[10]. Hao Wang and Chong-wei Xu. (2007). A New Lightweight and Scalable Encryption Algorithm for Streaming Video over Wireless Networks. International Conference on Wireless Network.
https://pdfs.semanticscholar.org/c163/59d793dcf019b4f7a0b68e419bc48490cb50.pdf
[11]. Zhang H. (2003). Content-Based Video Analysis, Retrieval and Browsing. In: Feng D.D., Siu WC., Zhang HJ. (eds) Multimedia Information Retrieval and Management. Signals and Communication Technology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-05300-3_2
[12]. P. Kulkarni, B.Patil, B.Joglekar. (2015). An effective content based video analysis and retrieval using pattern indexing techniques. International Conference on Industrial Instrumentation and Control (ICIC).
https://ieeexplore.ieee.org/document/7150717
[13]. D. Srivastava, R. Wadhvani and M. Gyanchandani. (2015). A Review: Color Feature Extraction Methods for Content Based Image Retrieval. International Journal of Computational Engineering & Management.
https://www.ijcem.org/papers052015/ijcem_052015_02.pdf
[14]. J.Huang, S R. Kumary, M. Mitraz, W. Zhux, R.Zabih. (1997). Image Indexing Using Color Correlograms.Proceeding Conference on Computer Vision and Pattern Recognition. http://www.cs.cornell.edu/~rdz/Papers/Huang-CVPR97.pdf
[15]. S. Kaur, A.Gupta. (2015). Support Vector Machine Based Approach to Content Based Image Retrieval Using Combination of Color and Texture Features. International Journal of Advanced Computer Engineering and Communication Technology.http://www.irdindia.in/journal_ijacect/pdf/vol4_iss3/2.pdf