A NEW TECHNIQUE FOR DETERMINING REGION OF INTEREST IN SELECTIVE VIDEO PROTECTION APPROACH

Rose M. Alhasany [1], **Lahieb M. Jawad** [2]
[1,2] College of Information Engineering, Al-Nahrain University, Baghdad, Iraq
rosemohammed246108@gmail.com [1] , lahieb.moh@coie-nahrain.edu.iq [2]

Received:6/11/2020, Accepted:25/2/2021

***Abstract*- Today, the use of video communication in real-time applications is rising at a rapid rate and most of these videos require secure transmissions like surveillance, video conferencing, video on demand, medical and military imaging systems. The trade-off between data security and real-time performance is the main challenge in this field. This paper presents a protection system that provides a balancing between the security level and coding efficiency. A smart selection for the motion information of the video is done based on canny edge detection scaling which reduces the amount of encrypted data and strong and fast stream encryption is done using key generation model of chaos and RC4 algorithm. The experimental results of the performance metrics confirmed the high perceptual and cryptographic security of the proposed system against attacks and showed that all the requirements of compression efficiency are satisfied.**

keywords: **Compression efficiency, Smart selection, Canny edge detection, Chaos, The RC4 algorithm.**

I. Introduction

Digital medium data (images, sounds, videos, etc.) is rapidly entering people's lives with the continuous growth and advancement of computer networks and wireless mobile communications technologies [1]. As a result, video networking systems, such as the associated video conferencing system, on-demand videos, and on-the-spot video surveillance system, have become more diversified [2]. In the case of attacks in open network communications, such as interception of data and theft of personal data, unauthorized copying and piracy, video data without encryption will become very vulnerable. Video data protection is of great importance for multimedia use and is currently a challenge in developing video technology [3]. H.264 / AVC is the commonly used IP- and wireless network video encoding format that supports a variety of applications such as video storage, television transmission, real-time video communications, and on-demand video. H.264/ AVC has unique coding structure characteristics, huge data volumes, and high real-time demand. It is, therefore, a complex and challenging task to attempt to encrypt video data without affecting the coding quality. Efficient features, stability, and compression in a video encryption device must be taken into account. Encryption has proven to be a viable method of protecting video data content [4]. The original data of your video is still unknown even though the encrypted data is intercepted by foreign spies because they have not the right decryption key and this allows for the protection of video data. The conventional encryption algorithms can be highly protected the compressed video data but not format compliant and the computational complexity is high [5], therefore; selective encryption technique is widely used in this field as it provides another efficient way in which the compressed video can be secure on the multimedia social network with satisfying the compression efficiency requirements where only the important or sensitive video data are chosen for encryption. In the proposed work, a smart selection technique based on edge detection is used to select the significant and effective parameters of the H.264 compression standard in an intelligent way. As the working nature of the compression standard is streaming processing, RC4 stream cipher is used with a chaotic system to encrypt the selected parameters of H.264. As it is known that the security of RC4 depends significantly on it's key and a lot of researches was done in this field to solve this problem [6], [7], [8], [9]. Thus in this work, RC4 is strengthened in this part by using a combination of two nonlinear chaotic methods, Henon-map and Sine-map in a key generation module. This paper aims to make a balance between two paws, satisfying a good level of security that protects the compressed video against various attacks and on the other hand, preserving the compression efficiency by avoiding the increase in the video bitrate, don't increasing the encoding time and the computational complexity and ensuring the video displaying in the decoder without problems which means maintaining the format compliance. The proposed strategy provided these requirements by determining the best significant and little amount of video data based on a smart selection technique, and design a fast stream and strong encryption algorithm.

II. Related Studies

[4] In this scheme, analyzation was done to the impact of the quantization parameter (QP) on the encryption of the sign of $T_1\,s$ and the inter macroblock non-zero coefficients and based on their results, the syntax elements chosen for encryption are the sign of intra-macroblock non-zero DCTcoefficients, the sign of trailing ones $(T_1\,s)$, the intra prediction modes (IPMs) and the sign of motion vector difference (MVD) to keep the H.264/AVC video secure in multimedia social networks. This scheme used an RC4 encryption algorithm with keyspace (2^{2048}) which made it very efficient against brute-force attacks. The simulation results ensured that this scheme is efficient in achieving perceptual security, preserving bitrate, and having low computational cost. [10] presented a partial encryption method for H.264 of video conference applications based on chaos. The scheme used two piecewise linear chaotic maps (PWLCMs) in building pseudo-random bit generators that generated two renewable key streams, one used for the encryption operation and the other used for the encryption decision whether to encrypt or not. Three syntax elements of H.264 bitstream were selected for encryption, non-zero quantified coefficients for both (I) and (P) frames, the signs of motion vectors (MVs) differences, and the Intra-Prediction Mode (IPMs). The results of the performance analysis indicated that this scheme is secure and very efficient in terms of computational complexity and coding efficiency. [11] designed a real-time video protection system based on the CABAC module of the H.264 entropy coding stage and the chaotic systems built from Logistic map. The encrypted bin-strings of CABAC were Intra-Prediction Mode (IPM), Motion Vector Difference (MVD), and residue coefficients, the encryption process was xoring the bin-strings with the chaotic sequence. The simulation results ensured that this encryption technique is very suitable for real-time application due to its low computational complexity. [12] in this work, they designed a video protection system suitable for the constrained devices in an Internet of multimedia things environment where they adopted the EXPer (extended permutation with exclusive OR) innovative to encrypt the H.264 bitstream selectively. The selected parameters include the signs of motion vector difference (MVD) and the absolute values of delta QP of both CAVLC and CABAC and the textural syntax elements of CAVLC (signs of $T_1\,s$, suffix and sign of NZ levels) and CABAC (UEG0 suffixes, signs of NZ-TC levels). The proposed cipher includes confusion and diffusion processes represented by permutation and xoring with three dynamic keys generated per video sequence. The simulation analysis confirmed that the EXPer system provided significant confidentiality with a small computational cost and a negligible bitrate overhead that made the proposed system very suitable for real-time applications.

III. Theoretical Background

A. A General Review of H.264

H.264/MPEG-4 Advanced Video Coding standard (H.264/AVC) is an efficient video coding standard jointly developed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) [13]. H.264/AVC has achieved a significant improvement in compression performance compared to prior standards, and it provides a network-friendly representation of the video that addresses both conversational and non-conversational applications [14]. Since the H.264 compression standard is based on YUV color space, the first step is converting the input video of RGB color format into YUV color format [15] . In H.264, the video passes through many divisions where in the first stage the source video sequence is divided into cycles that consist of a fixed number of pictures called the group of picture (GOP), then each picture is subdivided into one or more slices and each slice consist of an integral number of Macroblocks (MB), MB is the smallest coding unit in a frame that comprises information belonging to a region of (16×16) Y luma samples along with the related (U) and (V) component samples [25]. The cycle starts with Intra-prediction (I) frame and followed by many predicted (P) or bi-predicted (B) frames. I-frame is predicted using intra prediction for the current frame without using previously encoded pictures and the produced information includes the IPM and the residual data, there are different prediction modes for each block size where the luma components have two block sizes are (16×16) and (4×4), and one of $(8\times)$ block size for each chroma components, (16×16) has 4 prediction modes, (4×4) has 9 prediction modes and (8×8) chroma has 4 prediction modes [16]. P-frame and B-frame are predicted using inter prediction where each macroblock is predicted from an area of the same size in a reference picture to give the MV using motion estimation methods. Thus, the produced data from the prediction stage includes the prediction information (IPMD, MVD) and the residual error. The residual data will then pass through the next compression stages as shown in Fig. 1 that include transformation, quantization, and entropy coding which is CAVLC for the H.264 and the prediction information will be encoded using the fixed-length coding method (Exp-Golomb) where the encoder converts the syntax element into an index code-num.

$$M = \log 2(code_{num} + 1) \tag{1}$$

$$INFO = code_{num} + 1 - 2M \tag{2}$$

the codeword constructs from three parts: Mzeros, 1, INFO where INFO is M-bit field carrying information, the length of each codeword is equal to $(2M + 1)$ bits [17].

B. Canny Edge Detection

Edges are considered the most important aspect which provides valuable information for image analysis [18]. Edges are essentially the outline that distinguishes the object from its context. Edge detection is very complicated and is influenced by a deteriorating mechanism due to varying fluctuating noise levels [18]. The edge detector Canny is an operator for edge detection. Multi-stage algorithms are used to identify a wide range of borders [19]. The canny edge algorithm for edge detection is efficient. It takes a gray image as an input, processes it and generates the result that displays the intensity discontinuities. The canny edge detector performs five phases: input smoothing by Gauze filter, picture gradient detection, non-maximum detection, dual thresholding, and edge tracking through hysteresis. Several critical edge detection conditions can be satisfied by canny edge detection [20]:

1) Canny has better identification (criteria for identification). The canny method will illuminate all the current edges that conform to the threshold of the user-specified parameter.
2) Canny has the best way to locate. Canny is capable of generating minimum distance between the observed edge and the actual image edge.
3) It provides a clear response. It gives a single answer for every tip. This makes the edge detection for the next picture less confusing. Canny Edge Detection's identification of parameters would affect both detection outcomes and edge. Parameters are:
    - Gaussian default value.
    - Value of threshold.

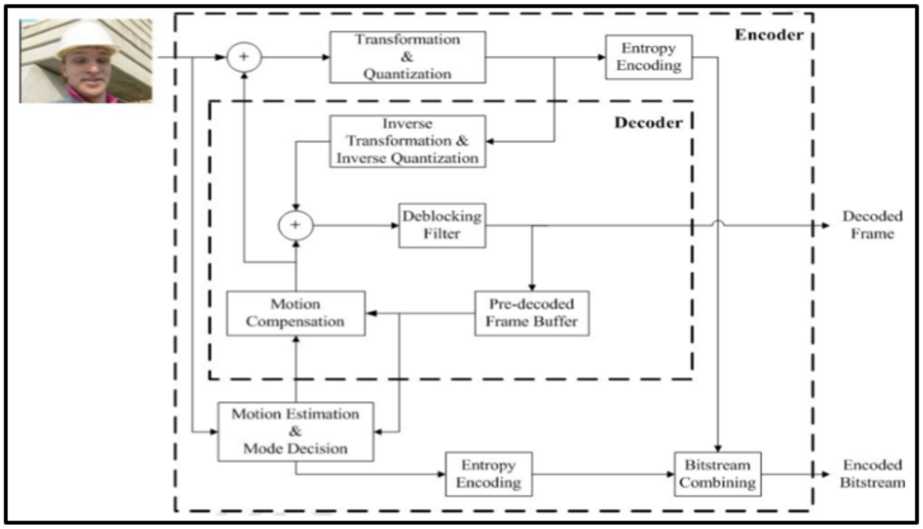


Figure 1: H.264 encoder and decoder

The implementation steps of canny edge detection are [20] :

1) Allow a Gaussian filter to remove all noise from an image. The resulting image will be less blurred after applying a Gaussian filter. The filter is used to obtain the real edge of the image. When a Gaussian filter is not used, the noise itself is observed sometimes as an edge.
2) Detect the Sobel operator's edge of the value "4.75" , so that you can see the edges and distortions. The results from both operators are combined to get the cumulative result by the equation given below:

$$[G] = [Gx] + [Gy] \tag{3}$$

3) Use the following formula to determine the direction:

$$G = \sqrt{G2x} + G2y \tag{4}$$

The detection uses two thresholds (maximum and minimum threshold values). The pixel is negated as the image background when the pixel gradient is greater than the maximum threshold. As pixel gradient from maximum to minimum threshold and if linked to another edge pixel higher than the maximum threshold, it is adopted as the edge.

4) Reduce the emerging edge line by using non-maximum suppression. this process produces a slimmer edge line.
5) The last step is to set up two threshold values to determine the binary value of image pixels. An example of the result of applying the canny edge detection method to an image is described in Fig. 2.

C. Chaotic Map Methods

The nature of the chaotic map methods in it's fundamental concepts of high sensitivity to the initial parameters and conditions has led many researchers to produce new chaos-based encryption algorithms [21],[22], [6], [8], [9], [23], [24], [25]. The chaotic map schemes have a good combination of randomness, velocity, high security and good encryption efficiency.

- Henon Map Henon-map is a two-dimension discrete dynamical system, defined by the equations [7]:

$$Xn + 1 = 1 - a\,xn^2 + Yn \tag{5}$$

$$Yn + 1 = b\,xn \tag{6}$$

It generates two random and unpredictable sequences (X, Y) based on initial values (x_0, y_0) and control parameters (a, b) where $a \in (1.54, 2)$ and $b \in [0, 1]$, the values(a=1.4, b=0.3) give the higher chaotic behavior of henon-map.

- Sine Map Sine-map is a simple type of discrete chaotic system. It is obtained from the sine function and its mathematical equation is [24]:

$$Zn + 1 = r\,sin(\pi Zn) \tag{7}$$

It is based on the initial value (z_0) and the control parameter (r) which its value in the range of $r \in [0, 1]$ and the value $(r = 0.94)$ is the best for the chaotic behavior of sine-map.

D. Traditional RC4 Algorithm

The RC4 is a stream cipher, symmetric key, designed in 1987 by Ron Rivest for RSA Security and characterized by its speed and simplicity. It is a variable key length (from 1 to 256 bytes) stream cipher with byte-oriented operations and based on the use of random permutation[26]. RC4 algorithm procedure mainly consists of four stages which are: Initialization, Key Scheduling Algorithm (KSA), Pseudo-Random Generation Algorithm (PRGA) and Xoring stage. The input to the algorithm is the key, the first stage is initializing two string arrays, (S) array with elements from (0) to (255) and (T) array which is filled with the entered key by the user. the second stage includes Randomizing (S) array depending on (T) array. The following stage is generating the final random keystream from the produced array of the KSA stage, with a length equal to the plaintext length. The last stage is xoring the plaintext with the keystream and obtaining the ciphertext.

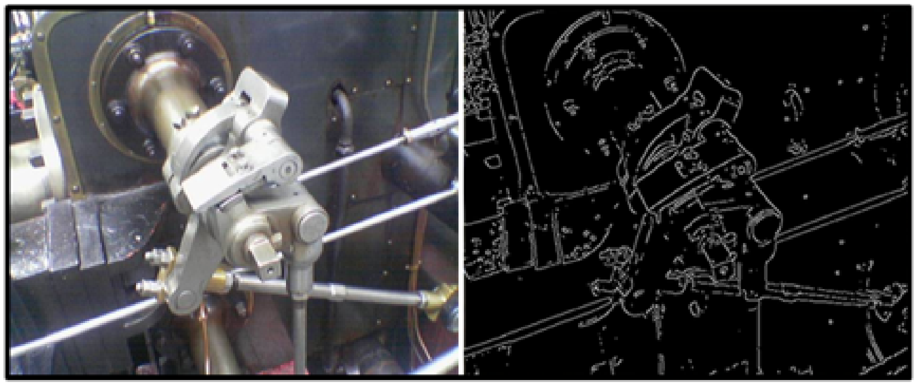


Figure 2: Result of applying canny technique to an image

IV. Proposed System

The proposed system mainly consists of two phases: the selective video encryption phase which is performed on the sender side, and the selective video decryption phase which is executed on the receiver side, and each phase has its stages as shown in Fig. 3.

A. *Selective Video Encryption Phase*

The selective video encryption phase includes three parts: key generation based on a chaotic system, selection technique, and encryption procedure for the selected parameters using the RC4 algorithm.

- **Key generation:** The cryptographic key plays an essential role in any cryptosystem and this is why the key generation process should be accurate as it has a very large impact on the strength of the encryption algorithm. The proposed scheme based on generating a One-Time Pad (OTP) and random sequence using two chaotic maps, Henon-map and Sine-map. As shown in Fig. 4, X and Y are the paths of Henon-map and Z is the path of Sine-map, they are given by the equations 5, 6, 7. The control parameters of them are set to the values (a= 1.4, b= 0.3, r= 0.95) to obtain a highly chaotic behavior. To increase the security and the randomness of the generated key, several operations are done between the chaotic paths (X, Y, Z), the first operation is sorting X sequence and the new positions of its sequence bytes is taken to order Y sequence bytes with it. The following operation is Xoring the rearranged Y with Z sequence, then the produced sequence is 1, kollpXored with the sorted X to generate the secret key sequence which is tested with NIST and proved that it is unpredictable and has high randomness. The length of the generated sequence from the chaos is equal to the key length of RC4 which is (256 bytes) and it is the number of iterations for the two chaotic maps.
- **Selection Technique:** The selection technique of data that would be encrypted should consider important performance metrics which are compression efficiency, time cost, and security. To satisfy these requirements, the selection must be for a small amount of encoding data to not cause delay to the encoding time and don’t change the compression ratio. At the same time, these data must be significant, and sensitive to achieve a suitable level of security.

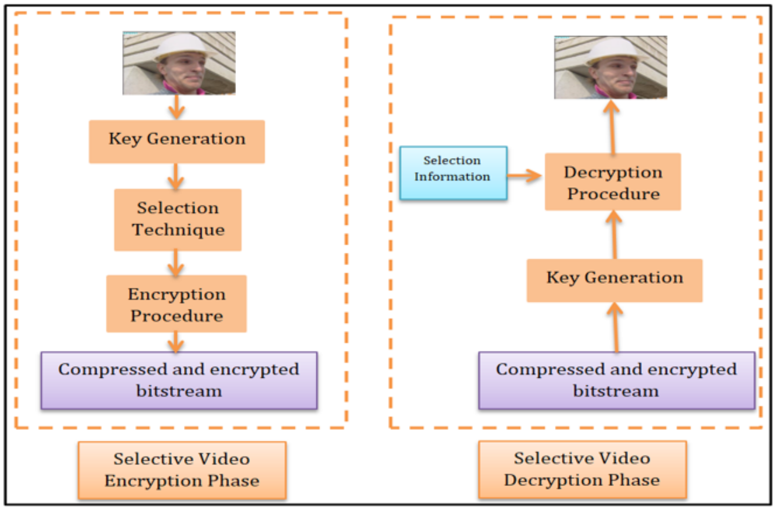

Key Generation
Selection Technique
Encryption Procedure
Compressed and encrypted bitstream
Selective Video Encryption Phase
Selection Information
Decryption Procedure
Key Generation
Compressed and encrypted bitstream
Selective Video Decryption Phase

Figure 3: Selective video encryption and decryption phases

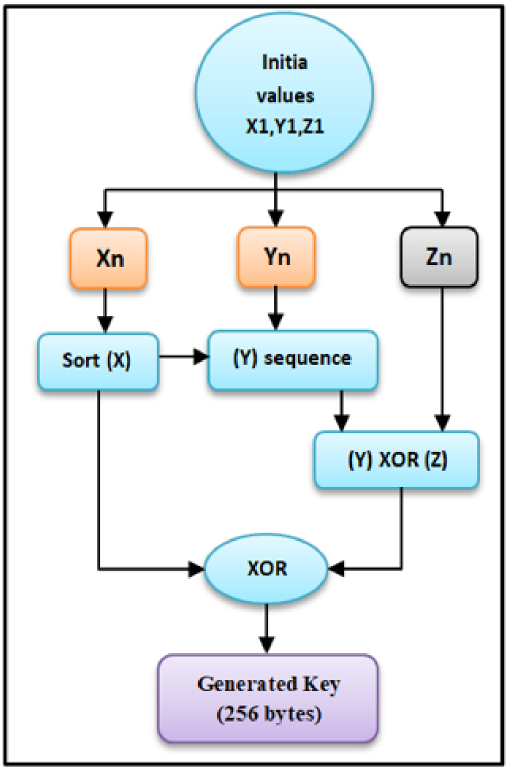


Figure 4: Key generation module

The selection technique is performed in two levels and each level is applied in a different stage where the first level will be applied to the original (RGB) video before the compression and the second level will be applied after entering the raw video to the H.264 compression standard. The following points explain the selection levels in detail:

1) **Cycle-level:** The selection levels are built based on the H.264 processing procedure where the video is encoded after dividing it into multiple groups of pictures called cycles. To encrypt the most important and sensitive compression data accurately, each cycle of the video will be specified as a significant or insignificant cycle. This specification is applied to the raw video and using canny edge detection which measures the amount of information, this measurement will be done for the frames of each cycle. The selection of this level is shown in Fig. 5, and is executed in the following steps:

a) Divide the raw video into cycles.

b) Convert the raw colored video into the grayscale video.

c) Apply the canny edge detection to the frames of all cycles.

d) Determine the number of edges pixels for the frames of each cycle.

e) Calculate the average of the determined edges pixels for each cycle (X).

f) Determine the threshold value (T) which is equal to the average of (X), as the following equation:

$$T = \frac{summation\,of\,(X)}{No.\,of\,cycles} \tag{8}$$

g) Based on the threshold, the cycles are specified as significant or insignificant where if (X) of the cycle is equal or larger than (T) then the cycle is significant, else the cycle is insignificant.

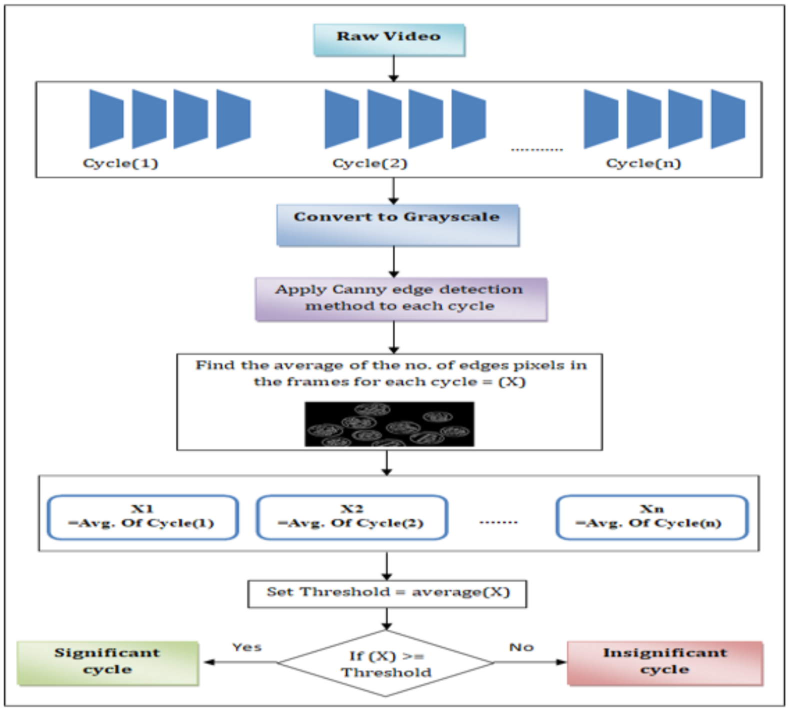


Figure 5: Flowchart of the cycle level

2) **Frame-level:** The second level of the selection is the frames inside each cycle, these frames are classified into three types, I-frame, P-frame, and B-frame. Since the prediction process to P and B frames is depending on I-frame as the reference frame, the encryption of the I-frame will be affected on them greatly and make them unintelligible. Thus I-frame is chosen to be encrypted in all cycles. As the cycle level classifies the cycles into significant and insignificant based on their information scale where the significant cycle has more information, subsequently it has more motion information, therefore; the movement information is selected to be encrypted for all P-frames in the significant cycle as P-frame is the reference for the adjacent P and B frames and as a result, its encryption will be effected on them, while only I-frame will be encrypted for the insignificant cycle because it has less motion information. This level of selection is shown in Fig. 6.

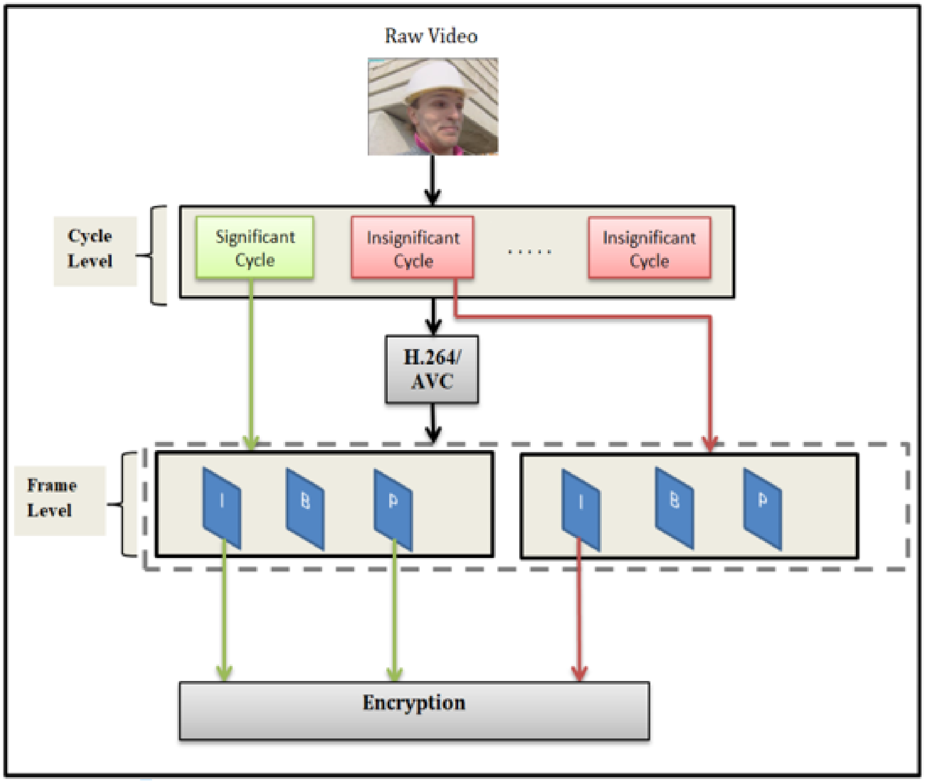


Figure 6: Selection technique levels

- **Encryption procedure:** this part clarifies the encryption process for the selected parameters of the compression standard that are selected depending on the selection technique levels. These parameters are: intra-prediction mode (IPM) of I-frame in all cycles (significant and insignificant) is selected for encryption which is one of the H.264/AVC main features and it is the directional spatial prediction for Intra coding areas, and motion vector difference (MVD) of all P-frames in the significant cycles is chosen to be encrypted to hide the movement of objects in videos, especially the smart motion videos. To keep the length of the bitstream produced from the compression standard, the encryption of these syntax elements (IPM, MVD) would be in the entropy coding stage. In H.264, IPM and MVD are encoded using Exp-Golomb coding which converts the symbols intocodewords with fixed lengths and in regular construction. The codeword consists of three parts: a prefix of M zero bits, a '1' and a suffix of M bits called INFO, the selected part for encryption is INFO. After the selection, the encryption algorithm that used is the RC4 algorithm which uses the generated key explained in the above part. After executing the steps of the RC4 algorithm, the last step (PRGA) that generating a keystream with a length equal to the plaintext length, this keystream is XORed with the plaintext. In the proposed scheme, the keystream bytes are yield to a modification stage to avoid the increase in the sensitive parameters data size when xoring them with the keystream. As shown in Fig. 7, the modification stage to the keystream bytes is based on the length of the INFO part which is (M bits), first the length of the keystream byte (KBlen) is checked where if (KBlen) is greater than (M), the bits of keystream byte would be shortened to be equal to the number of M bits and if (KBlen) is shorter than or equal to (M), the keystream byte would remain without change. After the modification process to the keystream bytes, it would Xored with the selected bits of the INFO part to produce the encrypted bits that return to take their place in the codeword and give the encrypted bitstream.

B. *Selective Video Decryption Phase*

The selective video decryption phase is the inverse of the encryption phase where the compressed and encrypted bitstream, selection information and seed are the input. This phase consists of two parts to reconstruct the original video. The first part is the key generation where the same steps of the key generation that executed in the encryption side will be performed with needing the initial values and control parameters of the chaos. The second part is the decryption Procedure which starts with the H.264 decoder where the first stage is the entropy decoding that converts the compressed bitstream into a series of symbols representing syntax elements of the video sequence. The encrypted bits of the syntax elements (IPM, MVD) are extracted when passing to the Exp-Golomb method using the information of the selection technique. As shown in Fig. 8, after the selection of the encrypted INFO part from the codeword bits, it would be Xored with the keys of RC4 keystream bytes after undergoing the modification process. Thus, the encrypted INFO bits are returned to their original and the H.264 completes its stages to reconstruct the video and gave the original video.

V. Simulation Analysis

The experimental analysis for the proposed video encryption system is done on a personal computer configured with a processor of 2.9 GHz core i7, and 16 GB RAM and using the programming language MATLAB R2016b. The experiments are evaluated using the raw video tasting sequences (Carphone, News, Foreman, Mobile) of Quarter Common Intermediate Format (QCIF) (176×144 pixels/frame) and (YUV) color format which is the most suitable color format for H.264 compression standard, the first (50) frames of the four test videos are selected and coded at the H.264 with (QP) value is taken (18) and (GOP) is (7).

1) **Visual Analysis**

Visual analysis ensures that no information about the original video can be extracted from the encrypted video. As shown in Table I, the encrypted videos are very noisy and unclear which means that the proposed selective encryption guarantees visual perception security.

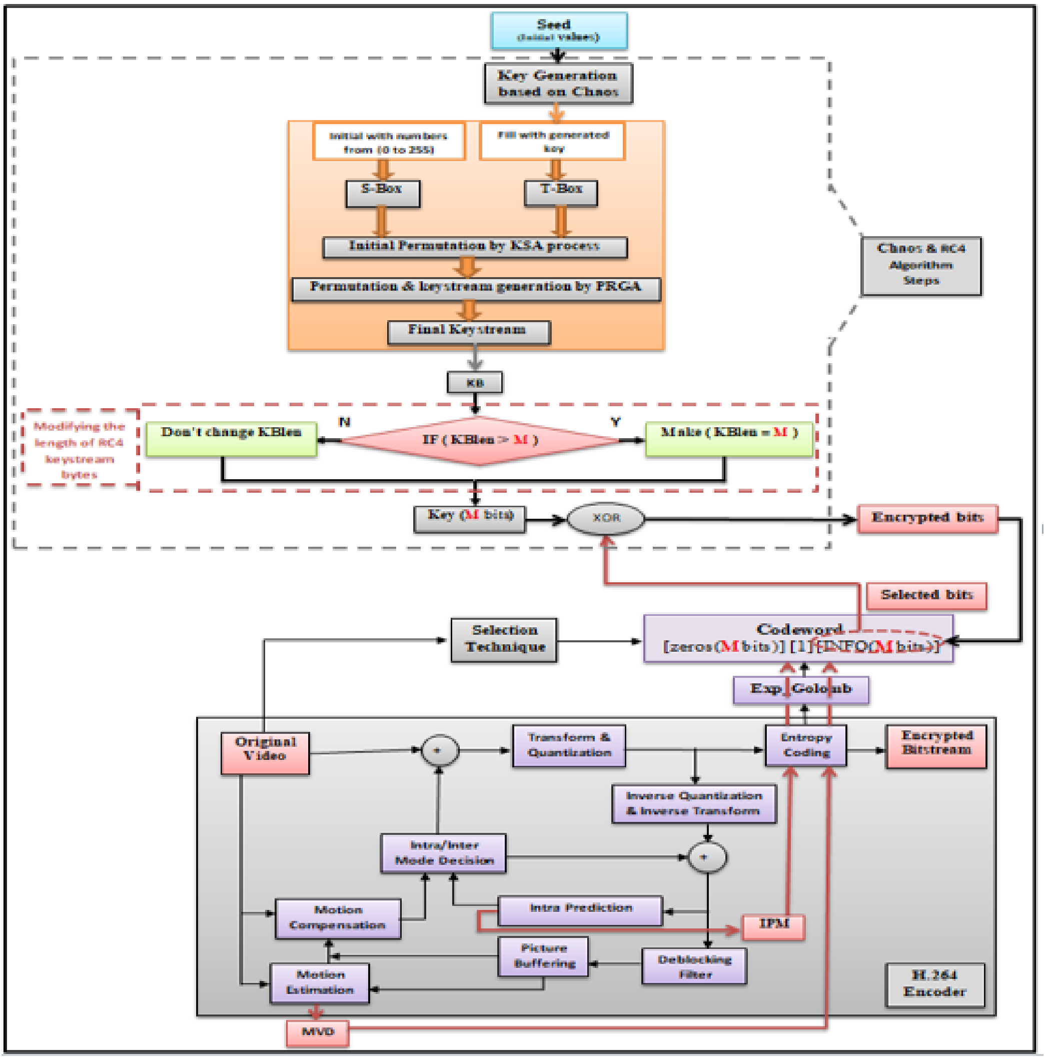


Figure 7: The encryption phase

2) **Histogram Analysis**

A histogram is a representation of the frequency distribution of image pixels. The results of histogram analysis for the source and encrypted videos frame number (10) is shown in Table II and as it is clear that the distribution of the pixel in the histogram of the original frame is wholly different from it in the encrypted frame which indicates to the strength of the proposed system against the statistical attacks.

TABLE I
Visual Video Quality

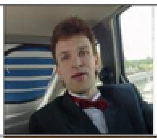
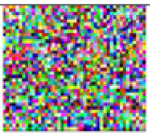
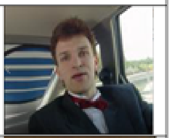
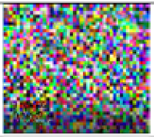
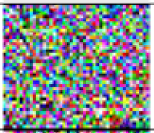
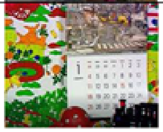
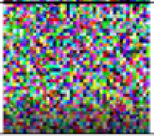
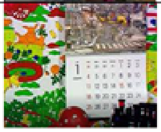
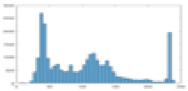
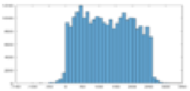
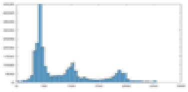
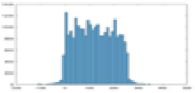
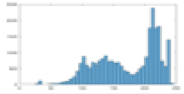
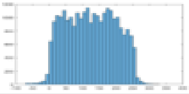
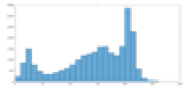
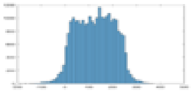
Sequence	Original Video	Encrypted Video	Decrypted Video
Carphone			
News			
Foreman			
Mobile			

TABLE II
Histogram Results

Sequence	Original Video	Encrypted Video
Carphone		
News		
Foreman		
Mobile		

3) **Encrypted Area (EA)**

The encrypted area is the ratio of encrypted bits to the whole bitstream length in terms of percentage [27]. The EA for videos calculated for the selected syntax parameters (IPM, MVD), when the video contains a lot of objects, this indicates that there is more movement, therefore; the more details the video contains, the more syntax elements are selected, and as a result, the EA would be higher. As it is obvious in Table III that each of the testing video sequences has a different EA because each video represents different combinations of the scene such as fast motion, complicated texture, related to camera motion, still background and active foreground.

$$EA = \frac{No.of\ Encrypted\ Bits}{Total\ Bitstream\ Length} \tag{9}$$

TABLE III
Percentage of Encrypted Area

Sequence	Encrypted Area (%)
Carphone	2.6701
News	3.2166
Foreman	3.3028
Mobile	3.7369

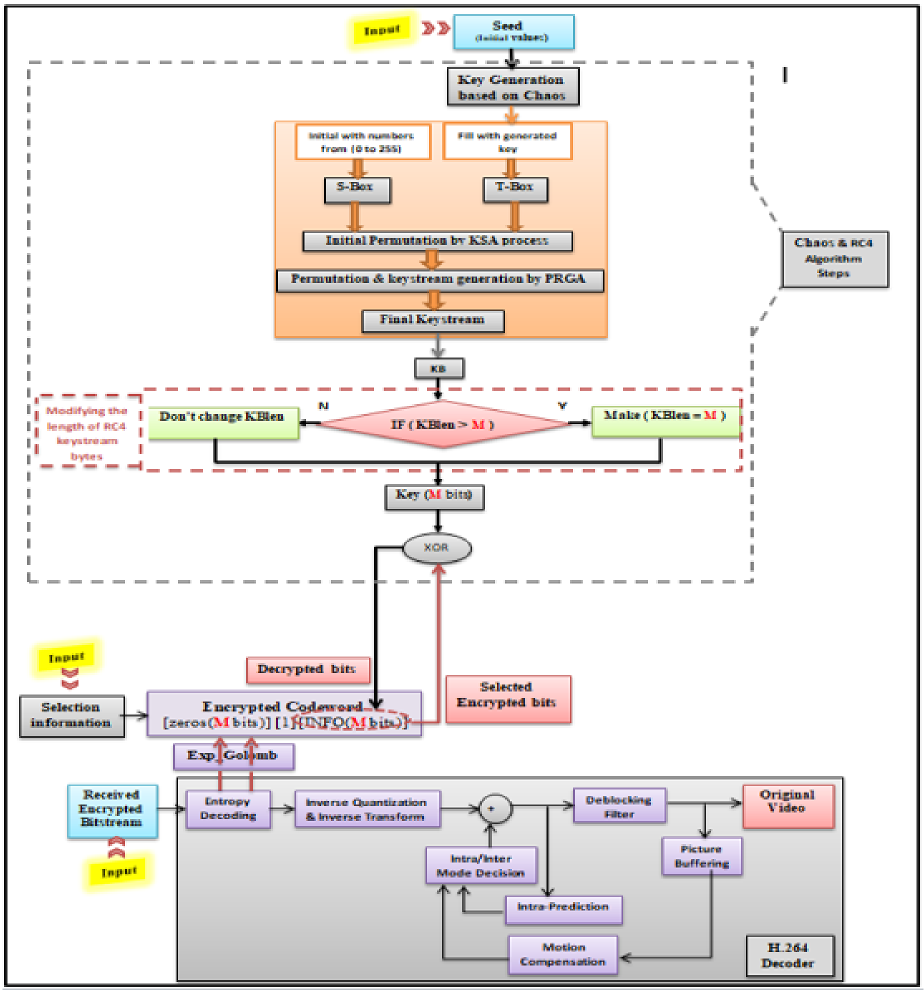


Figure 8: The decryption phase

4) **File Size Overhead**

In the proposed system, the selected parameters (IPM, MVD) are encrypted during the entropy coding stage with preserving it's the length of encoding, therefore; the bitstream length is kept without any change or overhead as it is obvious in Table IV were the experiments that done to (4) video sequences resulted from a negligible difference that can't change the size of the encoded video bitstream. Thus, the proposed encryption system satisfying the compression efficiency and is suitable for massive videos. The equation for calculating this metric is [27]:

$$Bit\,Overhead = \frac{Difference\,Between\,Encrypted\,and\,Encoded\,Bits}{Encoded\,Bits\,Alone} \tag{10}$$

TABLE IV
Percentage of File Size Overhead

Sequence	File size overhead(%)
Carphone	4.1355e-05
News	4.2802e-05
Foreman	4.3236e-05
Mobile	5.3831e-05

5) **Time Cost**

Time Cost is the additional processing delay caused by the encryption process to the encoding time [27]. Table V gives the rate of the changing to the encoding time for the testing video sequences and as it is obvious that the impact of encryption is a very small and not noticeable delay (null) which means that the proposed system satisfying time efficiency. The percentage of time cost is calculated using the following equation:

$$TC = \frac{TE + TD}{TCE + TCD} \tag{11}$$

where TCE, TCD, TE, and TD are the time of compression, decompression, encryption and decryption respectively.

TABLE V
Percentage of Time Cost

Sequence	Time cost(%)
Carphone	0.2491
News	0.2525
Foreman	0.3941
Mobile	0.7204

6) **Format Compliant**

The selected syntax parameters for encryption are not format or control information, therefore; the general bitstream format is preserved and the encrypted bitstream is decoded without any problem. Thus, the proposed system is compatible with the compression format.

7) **Peak signal-to-noise ratio (PSNR)**

In this evaluation, the original video is considered as a signal and the encrypted video is considered as noise. The calculation equations of PSNR are [6]:

$$PSNR = 20 * \log 10(\frac{255}{\sqrt{MSE}}) \tag{12}$$

Where the MSE equation is,

$$MSE = \frac{1}{M\,N}\sum_{i=1}^{M}\sum_{j=1}^{N}(|I(,) - I'(,)|)^2 \tag{13}$$

I(i, j) is the pixel value of the original video and $I'(i, j)$ is the pixel value of the encrypted video at the location (i, j). The PSNR evaluation for the testing videos is done for the three color planes (Y U V) of the video. In Table VI, the PSNR values for different reference videos in their original and the encrypted status are mentioned. As it is obvious that the values of all the encrypted sequences suffer from severe decent by comparing with the values of the original, this indicates the high level of security and protection provided by the proposed encryption scheme.

TABLE VI
PSNR Results for The Original, Encrypted and Decrypted Sequences

Sequence	PSNR					
	Original Video			Encrypted Video		
	Y	U	V	Y	U	V
Carphone	37.2327	37.7873	37.7141	7.481	8.6758	8.8202
News	35.6509	35.5307	35.3127	7.6648	8.3979	8.4231
Foreman	40.4286	40.5863	40.9782	7.1657	8.4057	8.8729
Mobile	41.082	41.4672	41.7402	7.6845	7.8559	8.0041

8) **Key Space Analysis**

As the key of RC4 is generated using a chaotic system, the keys of the proposed scheme are the chaotic initial values and parameters. Thus, the key set is (X0, Y0, Z0, a, b, r). The key length of the proposed scheme is (228 bits) by considering (52 bits) for the first three parameters and (24 bits) for the last three parameters. The key-space is (2^{288}), it is sufficient to resist the brute-force attacks.

9) **NIST Test**

NIST (National Institute of Standards And Technology) is a statistical test to measure the randomness of the output pseudo-random number generator sequence in binary form. NIST test includes 15 tests that focus on a variety of different randomness types that could exist in a sequence. All fifteen tests decision rules at the %1 PaletteTicks Level is based on if the computed test value is smaller than (0.01), then this sequence is nonrandom. Otherwise, conclude that the sequence is random. Table VII shows the randomness test of the generated key by the chaotic system [25].

TABLE VII
Results of NIST Tests

Test No.	Test Type	P-value	State
1	FrequencyTest	0.91589	SUCCESS
2	BlockFrequency	0.91043	SUCCESS
3	Runs	0.57456	SUCCESS
4	LongestRun	0.15075	SUCCESS
5	Rank	0.35637	SUCCESS
6	FFT	0.93468	SUCCESS
7	NonOverlappingTemplate	0.83235	SUCCESS
8	OverlappingTemplate	0.56524	SUCCESS
9	Universal	0.20486	SUCCESS
10	LinearComplexity	0.34955	SUCCESS
11	Serial	0.98937	SUCCESS
12	ApproximateEntropy	0.85752	SUCCESS
13	CumulativeSums	0.99041	SUCCESS
14	RandomExcursions	0.07285	SUCCESS
15	RandomExcursionsVariant	0.02234	SUCCESS

10) **Entropy Analysis**

Information entropy is a measure of the randomness amount in information content, it is calculated by the following equation [6]:

$$H(S) = \sum_{i=1}^{n-1} P(S_i) \log 2(\frac{1}{PS_i}) \tag{14}$$

Where (S_i) is the pixel value, the probability of the symbol (S_i) is $P(S_i)$ and (n) is the total number of symbols which is (256). In Table VIII, the entropy results assessed the randomness in the content of the video. as it is illustrated that the results values of all the encrypted sequences are closer to (8) which refers to the high randomness in the encrypted videos and the high degradation in its quality. If it is focused on the entropy values of the encrypted videos, it will be noticed that the videos with the highest values and the closest to 8 are the videos with the highest amount of encrypted data (EA) where the proportion between the entropy and the EA is directly proportional.

TABLE VIII
Entropy Results of Original and Encrypted Sequences

Sequence	Entropy	
	Original Video	Encrypted Video
Carphone	7.2267	7.964
News	7.5684	7.8915
Foreman	7.013	7.8569
Mobile	6.5791	7.9573

11) **Comparative Evaluation**

The performance of the proposed system is compared with prior works proposed recently, each one of these researches used various encryption techniques and protected different syntax elements of the H.264 standard. The comparison between them is done based on the essential performance metrics (PSNR, Encrypted Area, Time-Cost, Encrypted Data, Encryption Algorithm, Key Space), these criteria's measure the security level and the effect of the encryption on the compression. As it's described in Table IX, the proposed approach has less PSNR value than the rest references which means it has a higher security level. On the other hand, the rate of the encrypted area of the proposed approach is less than others while the security level is higher than others. That means the proposed approach balance the security level and the rate of encrypted area. Also, the rate of time cost of the proposed approach is fewer than the rest except [1] still has a few time cost because this research used only chaos for encryption. Therefore; the security level of [1] is lower because of its focus on speed only without taking care of the other requirements of the efficient performance like security.

TABLE IX
Comparison Between Proposed Work and Others

Results Metrics	[1]	[10]	[12]	Proposed
Average PSNR (db)	9.89	9.83	10.26	8.1
Encrypted Area(%)	4.7	41	11.92	3.3
Time-Cost (%)	0.14	3.12	3.1	0.39
Encrypted Data	IPM, signs of T1s, signs of the NZ coefficients, signs of MVD	IPM, sign of MVD, T1, suffix of levels	MVD, residue coefficients , delta QP	IPM, MVD
Key Space (bits)	626	147	128	228
Encryption Algorithm	Chaos	Permutation and XOR	Permutation and XOR	Chaos and RC4

VI. Conclusion

This paper proposed a selective encryption system for the H.264 based on reducing the encryption area using an intelligent selection technique that controls the amount of encrypted data depending on the scaling of the information by the canny edge detection method. This system reduced the encrypted area which enhanced the encryption efficiency in terms of time cost and bit rate overhead which makes it very suitable for real-time applications. The results of the security metrics like visual, histogram, PSNR, SSIM, keyspace, and NIST proved the high level of protection provided by this system to the encrypted video. Besides, the comparative evaluation between the proposed method and the other recent schemes clarified the improvement that was satisfied by balancing between security and compression efficiency.

References

[1] H. Xu, X. J. Tong, M. Zhang, Z. Wang, and L. H. Li, "Dynamic Video Encryption Algorithm for H264/AVC Based on A Spatiotemporal Chaos System" , J. Opt. Soc. Am. A Vol. 33, No. 6, p. 1166, 2016, doi: 10.1364/josaa.33.001166.
[2] L. Wang, W. Wang, and J. Ma, "Perceptual Video Encryption Scheme for Mobile Application Based on H.264" , J. China Univ. Posts Telecommun. 15, 73-78 (2008) .
[3] G. Van Wallendael, A. Boho, and J. De Cock, "Encryption for High Efficiency Video Coding with Video Adaptation Capabilities" , in Proceedings of IEEE Conference on Consumer Electronics (IEEE, 2013), pp. 31-32.
[4] F. Peng, X. Ging Gong, M. Long, and X. Ming Sun, "A Selective Encryption Scheme for Protecting H.264/AVC Video in Multimedia Social Network" , Multimed. Tools Appl. , Vol. 76, No. 3, pp. 3235-3253, 2017, doi: 10.1007/s11042-016-3710-x.
[5] Goldreich O (2009) , "Foundations of Cryptography Volume II Basic Applications" , Cambridge University Press, New York, pp. 373-481.
[6] G. Hanchinamani and L. Kulkarni, "An Efficient Image Encryption Scheme Based on a Peter De Jong Chaotic Map and a RC4 Stream Cipher" , 3D Res. , Vol. 6, No. 3, 2015.
[7] D. S. Alani and S. A. Al Iesawi, "Image Encryption Algorithm Based on RC4 and Henon Map" , J. Theor. Appl. Inf. Technol. , Vol. 96, No. 21, pp. 7065-7076, 2018.
[8] M. Kumari and S. Gupta, "A Novel Image Encryption Scheme Based on Intertwining Chaotic Maps and RC4 Stream Cipher" , 3D Res. , Vol. 9, No. 1, 2018.
[9] Methaq Talib Gaata and Fadya Fouad Hantoosh, "An Efficient Image Encryption Technique using Chaotic Logistic Map and RC4 Stream Cipher" , Vol.3, pp.213-218, 2016.
[10] N. Khlif, A. Masmoudi, F. Kammoun, and N. Masmoudi, "Secure Chaotic Dual Encryption Scheme for H.264/AVC Video Conferencing Protection" , IET Image Process. , Vol. 12, No. 1, pp. 42-52, 2018, doi: 10.1049/iet-ipr.2017.0022.
[11] F. K. Tabash and M. Izharuddin, "Efficient Encryption Technique for H.264/AVC Videos Based on CABAC and Logistic Map" , Multimed. Tools Appl. , Vol. 78, No. 6, pp. 7365-7379, 2019, doi: 10.1007/s11042-018-6494-3.
[12] A. Shifa, M. N. Asghar, S. Noor, N. Gohar, and M. Fleury, "Lightweight Cipher for H.264 Videos in The Internet of Multimedia Things with Encryption Space Ratio Diagnostics" , Sensors (Switzerland), Vol. 19, No. 5, 2019, doi: 10.3390/s19051228.
[13] M. R. and K. N. , "Comparative Study of Video Compression Techniques H264/AVC" , Int. J. Adv. Res. Comput. Sci. Softw. Eng. Res. , Vol. 4, No. 11, pp. 874-877, 2014.
[14] T. Wiegand and G. J. Sullivan, "The H.264/AVC Video Coding Standard" , IEEE Signal Process. Mag. , Vol. 24, No. 2, pp. 148-153, 2007.
[15] H. Riiser, "Adaptive Bitrate Video Streaming over HTTP in Mobile Wireless Networks" , 2013.
[16] F. K. Tabash, M. Izharuddin, and M. I. Tabash, "Encryption Techniques for H.264/AVC Videos: A Literature Review" , J. Inf. Secur. Appl. , Vol. 45, pp. 20-34, 2019, doi: 10.1016/j.jisa.2019.01.001.
[17] I. E. G. Richardson, "H.264 and MPEG-4 Video Compression" , 2003.
[18] Mahajan, S.R. , et al. ," Review of An Enhanced Fracture Detection Algorithm Design Using X-Ray Image Processing" , IJIRSET J. 1(2) (2012).
[19] Fazal-E-Malik, " Mean and Standard Deviation Features of Color Histogram Using Laplacian Filter for Content-Based Image Retrieval" , J. Theor. Appl. Inf. Technol. 34(1) (2011) .
[20] Kurniawan, S.F. , "Bone Fracture Detection Using OpenCV" , J. Theor. Appl. Inf. Technol. 64 (2015) .
[21] D. A. W. Sami Ibrahim and M. I. Sameer, "A New Approach of Color Image Encryption Based on RC4 Algorithm and Chaotic Map" , Int. J. Comput. Appl. Technol. Res. , Vol. 6, No. 9, pp. 422-429, 2017.
[22] I. Hussain, T. Shah, and M. A. Gondal, "Application of S-Box and Chaotic Map for Image Encryption" , Math. Comput. Model. , Vol. 57, No. 9-10, pp. 2576-2579, 2013.
[23] A. T. Sadiq, A. K. Farhan, and S. A. Hassan, "A Proposal to Improve RC4 Algorithm Based on Hybird Chaotic Maps" , Vol. 6, No. 4, pp. 74-81, 2016.
[24] R. E. de Carvalho and E. D. Leonel, "Squared Sine Logistic Map" , Phys. A Stat. Mech. its Appl. , Vol. 463, pp. 37-44, 2016.
[25] R. U. Ginting and R. Y. Dillak, "Digital Color Image Encryption Using RC4 Stream Cipher and Chaotic Logistic Map" , Proc. , 2013 Int. Conf. Inf. Technol. Electr. Eng. Intelligent Green Technol. Sustain. Dev. ICITEE 2013, pp. 101-105, 2013.
[26] William Stallings, W. , "Cryptography and Network Security: Principles and Practice" , Upper Saddle River, NJ: Prentice-Hall, 2013, 752p.
[27] S. Lian, "Multimedia Content Encryption Techniques and Applications" , 1st ed. New York: Taylor & Francis Group, 2008.