symmetry

MDPI

Article

A Selective Video Encryption Scheme Based on Coding Characteristics

Shuli Cheng [1], Liejun Wang [1,*], Naixiang Ao [2] and Qingqing Han [1]

[1] College of Information Science and Engineering, Xinjiang University, Urumqi 830046, China; cslxjuedu@126.com (S.C.); hq_564992125@163.com (Q.H.)
[2] Xinjiang Lianhai INA-INT Information Technology Ltd., Urumqi 830000, China; aonaixiang@sina.cn
* Correspondence: wljxju@xju.edu.cn; Tel.: +86-139-9981-6618

Received: 16 January 2020; Accepted: 21 February 2020; Published: 26 February 2020

Abstract: The protection of video data has become a hot topic of research. Researchers have proposed a series of coding algorithms to ensure the safe and efficient transmission of video information. We propose an encryption scheme that can protect video information with higher security by combining the video coding algorithm with encryption algorithm. The H.264/AVC encoding algorithm encodes the video into multiple slices, and the slices are independent of each other. With this feature, we encrypt each slice while using the cipher feedback (CFB) mode of the advanced encryption standard (AES) with the dynamic key. The key is generated by the pseudo-random number generator (PRNG) and updated in real time. The encryption scheme goes through three phases: constructing plaintext, encrypting plaintext, and replacing the original bitstream. In our scheme, we encrypt the code stream after encoding, so it does not affect the coding efficiency. The purpose of the CFB mode while using the AES encryption algorithm is to maintain the exact same bit rate and produce a format compatible bitstream. This paper proposes a new four-dimensional (4-D) hyperchaotic algorithm to protect data privacy in order to further improve the security of video encryption. Symmetric encryption requires that the same key is used for encryption and decoding. In this paper, the symmetry method is used to protect the privacy of video data due to the large amount of video encrypted data. In the experiment, we evaluated the proposed algorithm while using different reference video sequences containing motion, texture, and objects.

Keywords: coding characteristics; semantic element; selective encryption; hyperchaos

1. Introduction

Sharing video in multimedia applications has become an indispensable part of people's lives, especially in the medical industry [1] and military applications, due to the continuous development of multimedia applications and network technologies. With the continuous maturity and in-depth application of office automation in the medical industry, as the core information assets of medical care, these data assets are intentionally or unintentionally leaked, which will cause economic and reputation loss to the continuous operation of the enterprise, and even face more strict regulatory penalties. Therefore, some researchers have invested a lot of effort to further improve the video encryption scheme to ensure the security of private video data.

We all know that video encryption is generally divided into full encryption and selective encryption [2,3], both of which have advantages and disadvantages. Full encryption is usually suitable for small amounts of encrypted data and it requires high security. However, selective encryption is suitable for large amounts of encrypted data and strong real-time performance. Video encryption is a very time-consuming process due to the large amount of video data. Many video encryption algorithms are proposed according to the different semantic elements of encryption. Some schemes are

proposed in those articles [4,5] to encrypt the video by scrambling the intra prediction mode (IPM) of the intra-coded macro block. Exclusive OR (XOR) operation is usually used in intra prediction mode (IPM) to protect data privacy. In reference [6], Khlif et al. proposed encrypting the symbol of the current motion vector (MV) of each macroblocks (MB). Scheme [7] encrypts not only the intra prediction mode but also the motion vector difference (MVD). Lian et al. [8] proposed encrypting intra prediction modes, motion vector differences, and residual coefficients. Shi et al. proposed other encryption parameters in scheme [9], such as the encryption sequence parameter set and image parameter set. Jiang et al. [10] analyzed the perception performance, plaintext scrambling space, and key security of existing encryption algorithms. Next, while considering the key distribution and synchronization, the researchers propose a new encryption algorithm. This encryption algorithm chooses all IPM as plaintext, uses key-controlled cyclic sequence to encrypt intra prediction mode (IPM) of Intra 4×4 twice, and chooses chaotic random sequence as key, which improves the security of the algorithm and reduces the calculation cost. Sbiaa et al. proposed an encryption system in reference [11]. The system encrypts the alternating current (AC) and direct current (DC) coefficients generated after the quantization step, and then encodes them to encrypt and compress the video. Khlif et al. proposed a chaos-based encryption and compression scheme in reference [12]. The scheme uses chaotic functions with two different keys to generate two renewable key streams. The first is used for encryption and the other is used to determine whether to encrypt. The scheme inserts the encryption and decryption processes into the compression and decompression processes, respectively. The semantic elements that are encrypted in the encryption process are: non-zero quantized coefficients of intra (I) and inter (P) predicted frames, symbols of motion vector differences of inter prediction frames, and intra prediction modes of intra prediction blocks. Asghar et al. [13] proposed a new cryptanalysis method. The experiments show that video blurring can be achieved by only encrypting a few motion vector difference symbols (MVDs) in each chip. The Advanced Encryption Standard (AES) algorithm [14] based on cipher feedback mode (CFB) mode is a symmetric encryption method with strong security. This algorithm can effectively protect data privacy. Radanliev et al. [15–17] proposed that future developments be applied to IoT, and the main goal is to explore the development of IoT. These studies are very important for video encryption and privacy protection.

Traditional encryption algorithms are not suitable for image encryption [18]. Chaotic systems are known for their sensitivity to initial conditions and parameters, pseudo-randomness, ergodicity, and reproducibility [19]. It is very suitable for video image encryption. Researchers have proposed many chaotic encryption schemes based on this consideration. Cheng et al. [20] proposed a privacy-preserving image retrieval scheme based secure kNN, DNA coding and deep hashing. In addition, the researcher proposed histopathological image retrieval that was based on asymmetric residual hash and DNA coding [21]. In the current research, researchers have proposed some classic selective video encryption methods. Hamidouche et al. [22] applied the hyperchaotic system to selective video encryption. Altaf et al. [23] proposed efficient computation of selective video encryption that was based on chaotic block ciphers. Therefore, hyperchaotic encryption is also the main technology of video image encryption, and this technology has become a research focus of selective video encryption in recent years.

In this paper, a novel selective video encryption scheme is proposed, which extracts IPM, MVDs, residual coefficient (RC), and Delta quantization parameters (QP) parameters in each slice by using the slice characteristics of H.264 video coding. Then, the dynamic key generated by PRNG and the CFB of Advanced Encryption Algorithms (AES) are used to encrypt the parameters extracted from each chip. This paper proposes a new four-dimensional (4-D) hyperchaotic algorithm to protect data privacy in order to further improve the security of video encryption. Finally, the encrypted data is put back into the compressed stream. The scheme has the characteristics of real-time key updating and independent decryption between the slices, which greatly improves the security of encrypted video. The experiments confirm the effectiveness of the proposed method.

The organization of the article: In this section, we briefly introduce some application backgrounds of video encryption and some common methods of video encryption. In the next section, we will

focus on H.264/AVC video encryption systems. In the third section, we will introduce our proposed selective video encryption scheme. In the fourth section, the experimental results and experimental analysis of our proposed selective encryption scheme are given. Finally, we completed the study by giving conclusions.

This paper is organized, as follows: In Section 1, we briefly introduce some application backgrounds of video encryption and some common methods of video encryption. In Section 2, we will focus on H.264/AVC video encryption system. In Section 3, we will introduce our selective video encryption scheme. We present the experimental results and analyses in Section 4. In Section 5, we elaborate the relevant conclusions of the paper.

2. Description of Video Encryption Scheme Based on H.264/AVC

We all know that H.264/AVC [24] is a highly compressed digital video codec standard jointly proposed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Motion Picture Experts Group (MPEG). We introduce the relevant knowledge in this section. These related works are closely related to our proposed selective video encryption scheme that is based on encoding characteristics.

2.1. H.264/AVC Bit-Stream Syntax Structure

We will focus on analyzing the syntax structure of the H.264/AVC bitstream to propose a format compatible and secure video encryption scheme. By reading the paper, we know that H.264/AVC has the characteristics of bit-stream hierarchy [11], and the video stream is composed of individual syntax elements. Figure 1 shows the data hierarchy structure of H. 264/AVC bitstream.

In the H.264 structure, the encoded video is composed of group of picture (GOP), and the image group is composed of frame images. A frame image is composed of one or more pieces, one piece is composed of one or more macroblocks (MB) and one macroblock is composed of 16 × 16 YUV data.

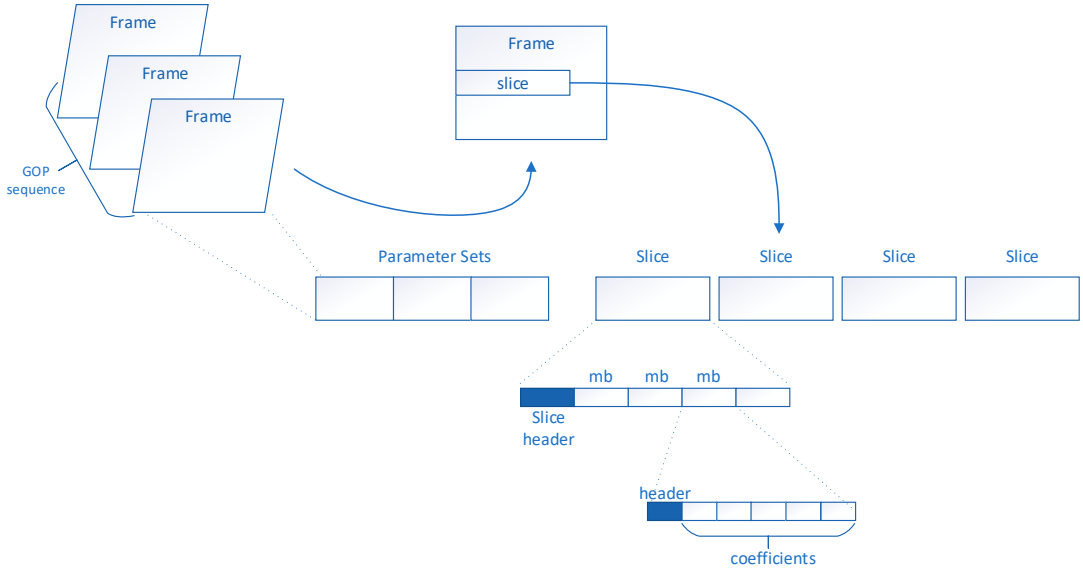


Figure 1. Data hierarchy in a video stream. Figure 1 contains three different types of frames (I-frame, P-frame and B-frame). Among them, I frames are encoded independently of other images. P-frames are encoded using the predictions of previous I or P-frames. B frames use double prediction of the previous and next I or P frames. The GOP sequence takes I frame as a start frame, and it is a periodic sequence. Each frame is sliced to limit error transmission.

2.2. Pseudo Random Sequence Generator

For known encryption algorithms, the encryption algorithm is more secure if the key is more difficult to obtain. Most researchers believe that pseudo-random sequence generator is a good choice

in order to achieve this goal. These sequences generated by a pseudo-random sequence generator can be used as a dynamic key for a cryptographic system, which confirms the security of the encryption algorithm. The chaotic function is very sensitive to the initial value and it is close to the noise characteristics, so using the chaotic function to make a pseudo-random sequence generator is a good choice. However, the key length of the chaotic function is short, and the chaotic value distribution is uneven. Therefore, Xu et al. [25] proposed an effective pseudo-random sequence generator. Figure 2 shows the working mode of PSNG.

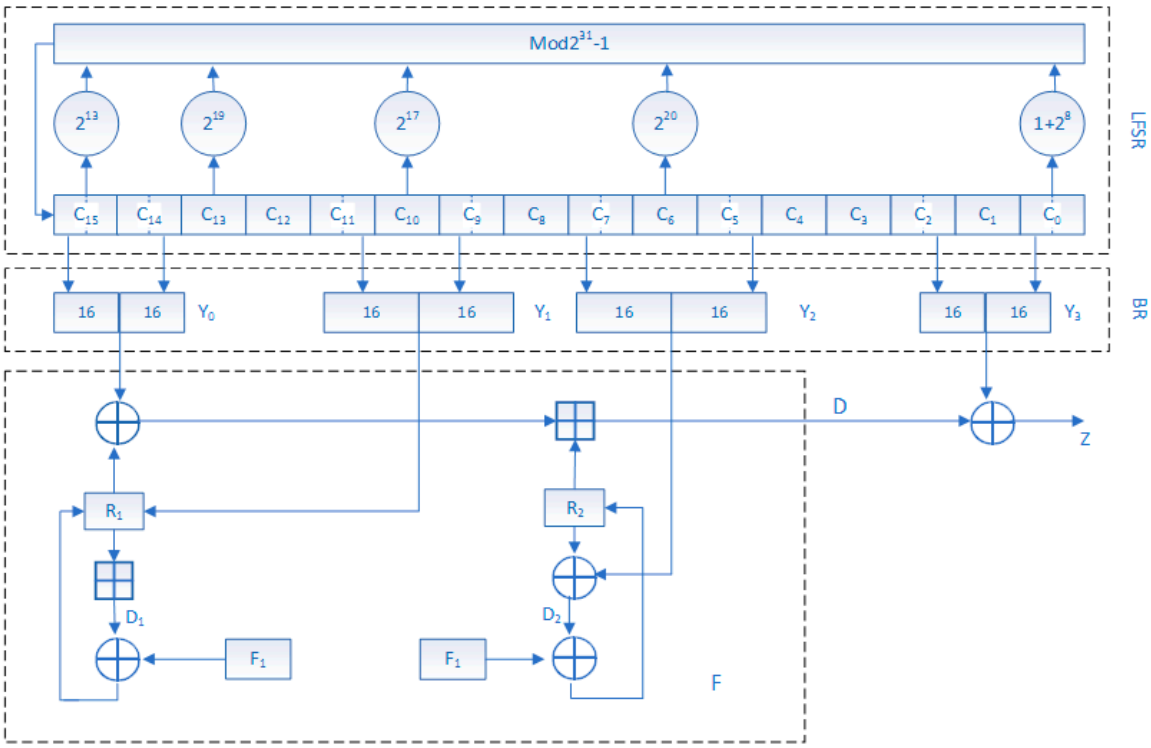
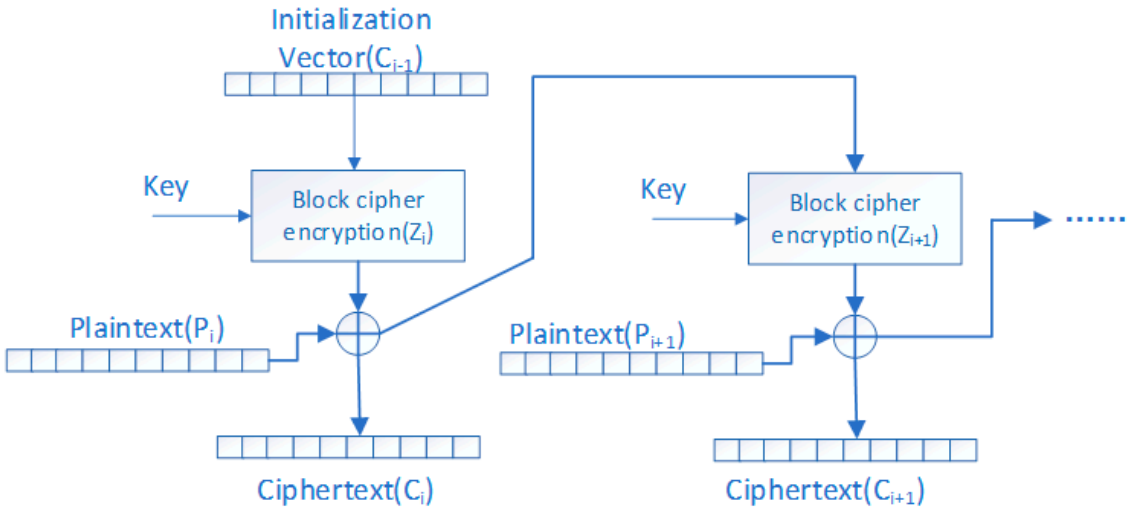


Figure 2. Working mode of Pseudo-random number generator (PRNG). PRNG contains three logical layers (top, middle, and bottom). The top layer consists of 16 linear feedback shift registers, the middle layer is a bit recombination layer, and the bottom layer is a non-linear mapping layer.

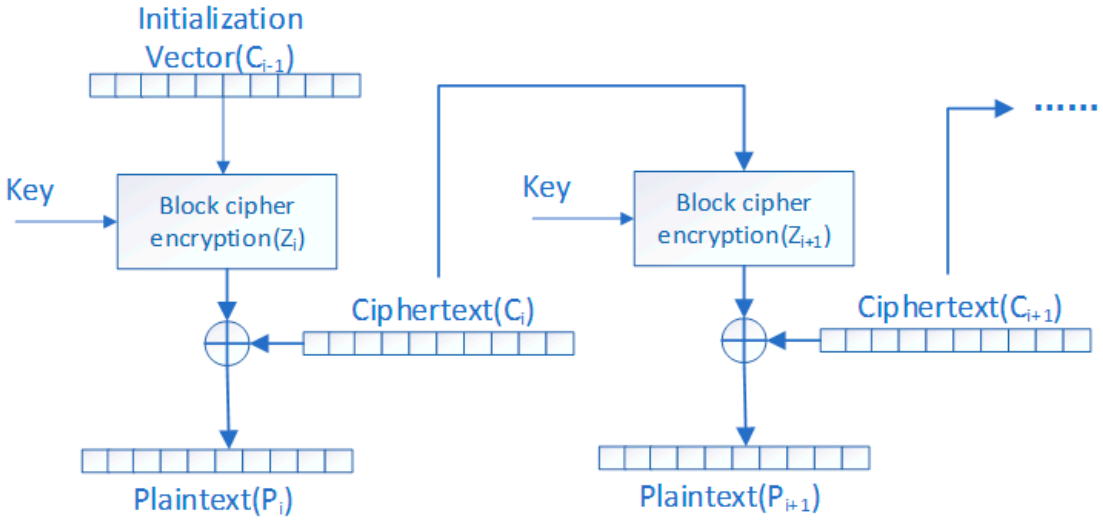
As shown in Figure 2, the Pseudo-random number generator (PRNG) is composed of a linear feedback shift register (LFSR), a bit recombination (BR), and nonlinear function F. Among them, R_1 and R_2 are two 32-bit memory cells, and F_1 and F_2 are two chaotic functions, respectively. PRNG execution produces a 32-bit Z at a time, which forms the output of the key stream by combining 32-bit Z. Literature [25] suggests that PRNG discards the first 100 values and uses post-order values, given the security of the password.

2.3. AES Encryption Algorithms

As a block encryption algorithm, AES supports multiple encryption modes to meet different security requirements and transmission requirements. For example, Electronic codebook (ECB), Cipher-block chaining (CBC), Cipher feedback (CFB), Output feedback (OFB), and so on. In fact, the ECB mode is a common encryption mode of the AES. In CFB mode, CFB converts block ciphers to stream ciphers. Figure 3 shows the working mode of CFB.



(**a**) cipher feedback mode (CFB) mode encryption.



(**b**) CFB mode decryption.

Figure 3. The working mode of CFB. Figure 3a CFB mode encryption. The encryption process requires a shift register that is the same size as the block to initialize the register. Then, the contents of the register are encrypted using a block cipher. Figure 3b CFB mode decryption. Based on Exclusive OR (XOR) operation, CFB mode decryption can be implemented.

The generation of the key stream Z_i and the ciphertext block C_i is as follows:

$$\begin{cases} Z_i = E_{key}(C_{i-1}) \quad , for \ \ i \geq 1 \\ C_i = Z_i \oplus P_i \end{cases} \tag{1}$$

There, $\oplus$ is an XOR operation.

2.4. The Four-Dimensional Hyperchaotic System

The four-dimensional hyperchaotic system can be expressed, as follows:

$$\begin{cases} y_1 = \mathrm{a}(y - x) + w \\ y_2 = \mathrm{d}x - xz + cy \\ y_3 = \mathrm{xy} - \mathrm{bz} \\ y_4 = \mathrm{yz} + \mathrm{ew} \end{cases} \tag{2}$$

where a, b, c, and d are system control parameters. If the parameters are set to a = 35, b = 3, c = 12, d = 7, and e = 0.2, the four-dimensional chaotic system is hyperchaotic, and the system can generate four hyperchaotic sequences (y_1, y_2, y_3, y_4).

3. Video Selective Encryption Scheme

We know that researchers encrypt semantic elements through different combinations to make the video content as obscure as possible, according to the previous introduction. Video encryption usually needs to consider compatibility, security, and timeliness, so it is important to propose a selective encryption algorithm to protect video privacy.

3.1. Selection of Important Syntax Elements

Selective encryption is a technology that saves time by encrypting a small amount of bitstream data, but it still has sufficient security [26]. Analyzing the research work of the predecessors, we found that important semantic elements in the video stream were encrypted in most video encryption schemes. For example, IPM, MVD, residual coefficient, and so on.

In reference [13], Asghar et al. proposed a new cryptanalysis method. Asghar et al. believe that the probability of each syntax element obeys Poisson distribution in the video test sequence. Based on two assumptions: (1) the values of the symbols in the file are evenly distributed; and, (2) For syntax elements with a range of continuous values, it is very difficult to correctly guess the cryptographic elements on a sufficient number of elements. Asghar et al. analyzed that there are 119,904 MVD symbols in a 21.7 M video file. It is necessary to correctly guess that the MVD symbol has 2119904 possibilities, so the probability that the MVD symbol can be fully guessed is 1/2119904. Suppose that the attacker needs to correctly guess 80% of the MVD symbols to make the video visible, and guess that the combination of 80% MVD symbols has C_{119904}^{95923}. It is very difficult to correctly guess 80% of the MVD symbols. Moreover, the analysis given [8] is reasonable in encrypting spatial information and motion information during H.264/AVC encoding. However, the scheme of encrypting the syntax elements only encrypts the equal-length codes to maintain the same bit rate as the original code stream [14]. Therefore, a selective video encryption scheme with high security, high efficiency, and format compatibility is proposed in this paper. When compared with reference [8], our proposed encryption scheme directly encrypts the semantic elements in the H.264 bitstream, thereby directly changing the code stream. From previous research work, we know that a video is encoded into multiple frames, and a frame consists of multiple fragments, and these fragments are independent of each other. According to this idea, we use the independence between video slices in the proposed encryption scheme. The AES with CFB mode is used to encrypt the IPM, MVDs, residual coefficient, and Delta QP [27] semantic elements in video. More importantly, the key between the slices is updated in due course. This keeps the encrypted video format compatible and real-time secure. The following video semantic elements need to be encrypted.

Intra Prediction Mode (IPM): IPM supports four prediction modes of Intra_4 × 4, Intra_16 × 16, Intra chroma and I_PCM 24. The IPM in the Intra_4 × 4 and Intra_16 × 16 blocks is selected for encryption in our encryption scheme.

Motion Vector Difference symbol (MVDs): We need to encrypt motion vectors in order to destroy the movement information. In H.264/AVC, the motion vector mv is further predicted to obtain mv', and

the motion vector difference $mv_d = mv - mv'$. In the H.264/AVC baseline profile, the MVD is encoded while using Exp-Golomb entropy coding. The value of MVD and the corresponding Exp-Golomb codeword is indicated in the paper [28], and the last bit of the Exp-Golomb codeword is given to affect the MVD symbol. Therefore, we only need to encrypt the last bit of the Exp-Golomb codeword in the motion vector difference.

Residual coefficient: We need to encrypt some other sensitive data, that is, residual data, in order to further promote the security of the encryption video. In the H.264/AVC baseline profile, the format of quantization coefficient [29] used to encode residual blocks is as follows: {coeff_token, sign_of_trailing ones, level, total_zeros, run_before}. The encryption of the residual coefficients is done by modifying the trailing ones sign and the level codeword [28,30]. The encryption of the residual coefficients is implemented in our scheme by replacing the trailing ones sign, level, and totalzeros codewords.

Delta QP: Differences in quantization parameters (QP) can affect video quality. We need to encrypt the Delta QP parameters in order to make the texture information more distorted.

3.2. Selective Encryption Process

We select the IPM, MVDs, and the residual coefficients affecting the texture information, Delta QP [27] to encrypt the video, by analyzing the structure of H.264/AVC video code stream. Figure 4 shows the flow chart of selective video encryption.

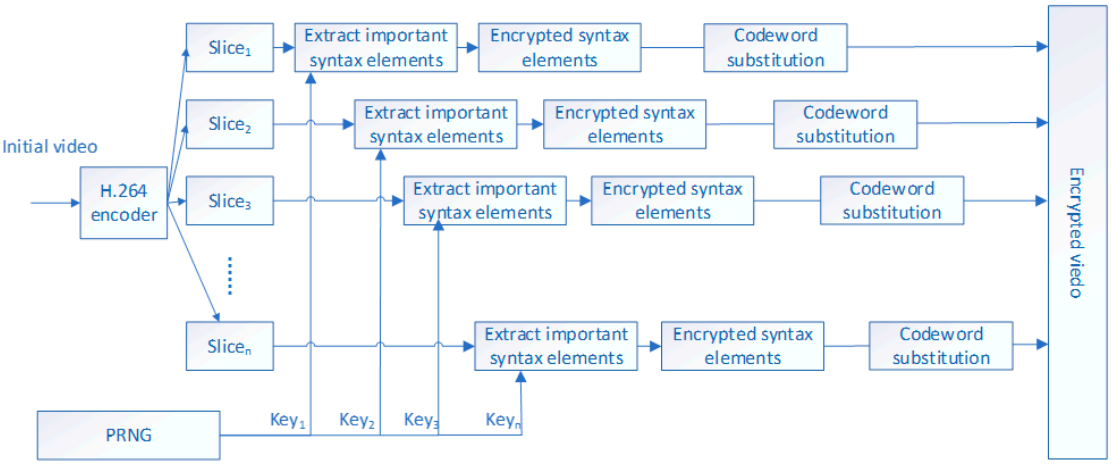


Figure 4. Encryption flow chart. Figure 4 is selective video encryption, which mainly encrypts important elements in the video slice. At the same time, there is independence between video slices.

The H.264 encoder encodes a video into multiple slices, and the slices are independent of each other. We use this feature of video coding to propose an efficient and secure video encryption scheme. We regard $C_i = E_{key}(C_{i-1}) \oplus P_i$ as the mark of the CFB mode while using key k to encrypt the n-bit P_i. Figure 3 shows the implementation scheme. We chose to use CFB mode to maintain the compression ratio. As can be seen from Figure 3, plaintext block P_i is the same size as the ciphertext block C_i. Furthermore, ciphertext from the previous block is directly used to encrypt the current block. The algorithm proposed in this paper has three stages: constructing the plaintext block P_i, encrypting the plaintext block P_i, and replacing the encrypted information with the original codeword.

Plaintext construction: Since the slices are independent from each other, we encrypt the important semantic elements of each slice in units of slices while using the PRNG generation key in Section 2.2. In our scheme, we can copy the important semantic elements of each slice in the H.264 code stream directly to the vector P_i to create the plaintext, and until the vector P_i is completely filled or reaches the edge of the slice. We think that L is the size of the vector P_i, and $L(P_i)$ is the length at which the vector P_i is filled. Where $L = 128$. If $L(P_i) < L$, we fill P_i with $P(i) = 0$, where $i \in \{L(P_i) + 1, \ldots, L\}$. From the perspective of encryption algorithms, padding is to improve the security of the algorithm.

Encrypting plaintext while using CFB mode of AES. Figure 3a shows the encryption process for CFB mode. Z_i is created by the previous ciphertext block C_{i-1}. According to Formula (1), we perform xor operation between Z and plaintext block P to achieve the encryption of the current plaintext block. In the key generation phase, we used the PRNG that was proposed by Xu et al. [25] in Section 2.2 to generate the key. For security reasons, we will consider the sequence after the 100th 32-bit as the key stream.

Replace the original bit: this is the last stage of encryption. In this phase, we replace the P_i vector in the original stream with the encrypted ciphertext vector C_i. The H.264 codestream is sequentially accessed in the plaintext P_i construction phase, given the length of each important semantic element (l_n, l_{n-1}, ... , l_1). Subsequently, we replace the corresponding part of the plaintext P_i in the H.264 stream with the ciphertext C_i. At this point, our proposed selective encryption scheme ends.

3.3. Decryption Process

Figure 3b shows the decryption process for the CFB mode. The ciphertext C_{i-1} of the previous block is used as an input of the AES encryption algorithm to generate Z_i (here, the same encryption function E_{key} and encryption key k as the encryption phase are used.), and the current ciphertext block C_i and Z_i are XORed to generate a plaintext block P_i. Repeat the above operation until the plaintext vector P_i is generated. Afterwards, the plaintext vector P_i is divided into segments of length l_n, l_{n-1}, ... ,l_1 to replace the corresponding stream in ciphertext to generate the original stream.

3.4. Introduce 4-D Hyperchaotic Selective Video Encryption Process

Figure 5 shows the encryption and decryption process of the proposed 4-D hyperchaotic system. Figure 4 is selective video encryption, which mainly encrypts important elements in the video slice. At the same time, there is independence between video slices. It can be seen from Figure 5 that we use the 4-D hyperchaotic system to encrypt selective video encryption results. Therefore, the selective video encryption result is the input of the 4-D hyperchaotic system for Figure 5. At this time, the ciphertext image in Figure 5 is a result that is based on video encoding and 4-D hyperchaotic hybrid encryption.

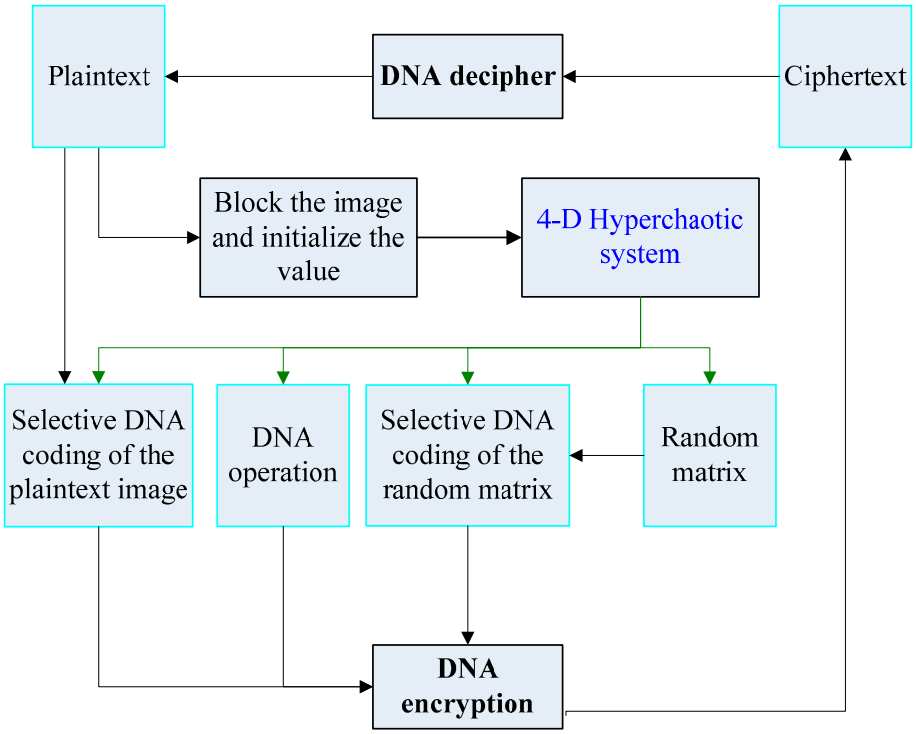


Figure 5. The encryption and decryption process of four-dimensional (4-D) hyperchaotic system.

4. Experimental Results and Analysis

In this section, we will present the results of selective video encryption. We used the JM8.6 reference software in the video coding phase. In the experiment, we analyzed the proposed selective encryption scheme using seven reference sequences with a frame rate of 30 frames/s and a format of QCIF (176 × 144). These seven reference video sequences are combined by different motions, colors, contrasts, and objects. Each of the seven reference sequences has 300 frames for experimental analysis, and the structure of the GOP is "IPPP".

4.1. Security Analysis of Selective Video Encryption Scheme

Perceptual security means that the video sequence after encryption cannot be correctly identified, depending on the attributes of the encryption. In this section, we will discuss the perceived security of our proposed encrypted video images from two aspects: subjective video quality and objective video quality.

4.1.1. Subjective Video Quality Analysis

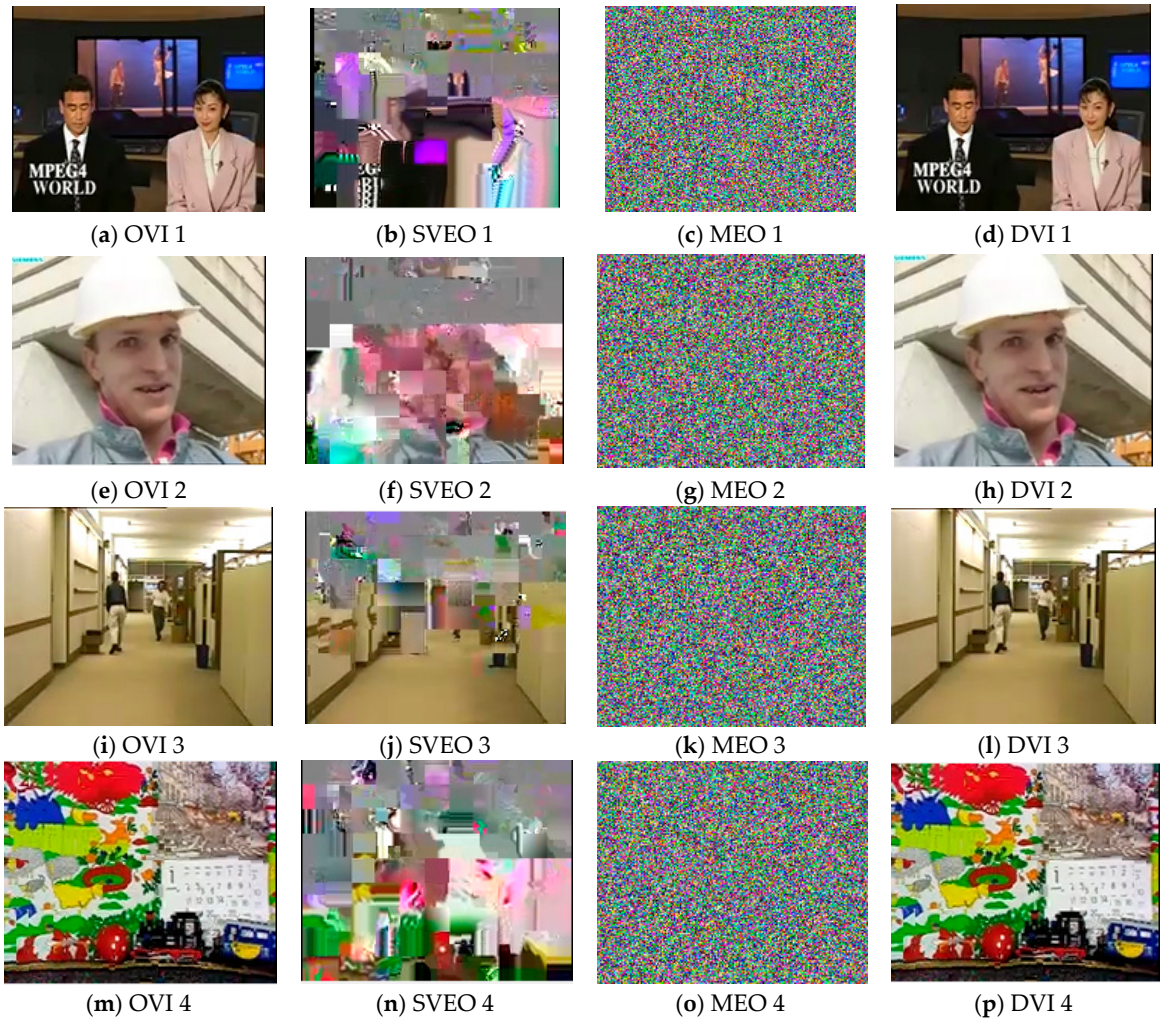
In our proposed scheme, we encrypt the IPM, RC, Delta QP, and MVDs semantic elements in the video sequence by means of fragmentation, which guarantees the visual perception security of video encryption. This paper proposes a new five-dimensional (5-D) hyperchaotic algorithm to protect data privacy in order to further improve the security of video encryption. Only four benchmark sequence renderings are shown here due to space limitations (Figure 6). This includes the original video image (OVI) (OVI 1, OVI 2, OVI 3, OVI 4), Selective video encryption output (SVEO) (SVEO 1, SVEO 2, SVEO 3, SVEO 4), Mixed encrypted output (MEO) (MEO 1, MEO 2, MEO 3, MEO 4), and the decrypted video image (DVI) (DVI 1, DVI 2, DVI 3, DVI 4). Where, the first column in Figure 6 is the original video image. The second column in Figure 6 is selective video encryption output. The third column in Figure 6 is the mixed encrypted output. The last column in Figure 6 is the decrypted video image. The mixed encryption output is the mixed encryption result. The specific step is to perform selective encryption first and then perform 5-D hyperchaotic encryption.

In Figure 6, if the user needs to ensure sufficient real-time and large amount of data, we recommend that the user choose a selective video encryption algorithm. This method can encrypt video data in real time to a certain extent. We recommend that users choose a hybrid encryption method, which mainly focuses on video encryption security, if the user needs sufficient security and the amount of data is small. Figure 6 shows the subjective evaluation of the algorithm. In addition, the encrypted video sequence maintains the same size as the original video sequence (as shown in Table 1), and the bitstream format compatibility is well accomplished.

Table 1. Video Sequence Size Comparison.

Video Sequence	Video Size	Encoded Video Size	Encrypted Video Size	Decrypted Video Size
coastguard	10.8 KM	444 K	444 K	444 K
containe	10.8 KM	251 K	251 K	251 K
foreman	10.8 KM	328 K	328 K	328 K
hall	10.8 KM	257 K	257 K	257 K
mobile	10.8 KM	927 K	927 K	927 K
mother-daughter	10.8 KM	180 K	180 K	180 K
news	10.8 KM	301 K	301 K	301 K

Table 1 shows that the size of the reference video sequence is 10 KM in our encryption scheme. The video sequence size has significantly dropped after H.264 encoding. Additionally, the encrypted video bit stream is the same size as the encoding video bit stream, which retains the original bit characteristics very well.



(**a**) OVI 1 (**b**) SVEO 1 (**c**) MEO 1 (**d**) DVI 1

(**e**) OVI 2 (**f**) SVEO 2 (**g**) MEO 2 (**h**) DVI 2

(**i**) OVI 3 (**j**) SVEO 3 (**k**) MEO 3 (**l**) DVI 3

(**m**) OVI 4 (**n**) SVEO 4 (**o**) MEO 4 (**p**) DVI 4

Figure 6. Subjective evaluation of the algorithm. The first column in Figure 6 is the original video image. The second column in Figure 6 is selective video encryption output. The third column in Figure 6 is the mixed encrypted output. The last column in Figure 6 is the decrypted video image.

4.1.2. Objective Video Quality Analysis

In this section, we also used Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) to measure the perceived effect of the video. PSNR is widely used to objectively evaluate video quality. A larger value of PSNR indicates a better video quality, and vice versa. The calculation formula of PSNR is obtained according to the mean square error (MSE) of the error of two images. The calculation formula of MSE is expressed by Equation (3), and the calculation formula of PSNR is represented by Equation (4). However, PSNR does not evaluate video quality very well because of the nonlinearity of the human visual nervous system. Therefore, it has been proposed to use SSIM to evaluate the similarity of two images. The value of SSIM ranges from 0 to 1. A value that is close to 1 indicates that the reference image is more similar to the target image.

$$\mathrm{MSE} = \frac{1}{H \times \mathrm{W}} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} \left(P_{org}(x,y) - P_{enc}(x,y) \right)^2 \tag{3}$$

Therefore, the formula for calculating PSNR is described, as follows:

$$\mathrm{PSNR} = 10 \ln\left(\frac{\left(2^k - 1\right)^2}{MSE}\right) \tag{4}$$

where, H and W indicate the height and width of the video image, respectively. $P_{org}(x, y)$ show the pixel value of the original video image, and $P_{enc}(x, y)$ indicate the pixel value of the encrypted video image. k represents the number of bits per pixel, which is generally 8.

Table 2 gives the average PSNR values of the video after encoding and after encryption. Through experiments, we know that the benchmark sequence outputs 264 streams after passing through the H.264 encoder. However, our proposed selective encryption scheme encrypts the important semantic elements in the 264 code stream, and the encrypted video file is still 264 streams. In the Table 2, we compare the code stream after encoding with the code stream after encryption to calculate the PSNR value. Similarly, Table 3 provides the average SSIM value of the encrypted video. We observe Table 3 to find that the video image after encryption is quite different from the video image after encoding.

Table 2. Average Peak Signal to Noise Ratio (PSNR) comparison of video sequences.

Sequence	PSNR (dB)					
	Encoded Video			Encrypted Video		
	Y	U	V	Y	U	V
coastguard	34.82	42.72	44.61	14.28	31.34	40.05
containe	37.06	41.36	41.22	14.62	30.10	29.02
foreman	36.71	40.62	41.59	12.37	28.08	28.41
hall	38.12	39.71	41.46	14.27	21.32	28.05
mobile	33.93	35.72	35.39	12.62	19.03	19.02
mother-daughter	38.22	41.56	42.53	16.54	22.05	25.11
news	37.62	40.61	41.06	11.25	22.46	26.24

Table 3. Average Structural Similarity Index (SSIM) comparison of encrypted video sequences.

Video Sequence SSIM Performance Analysis						
Coastguard	**Containe**	**Foreman**	**Hall**	**Mobile**	**Mother-Daughter**	**News**
0.31	0.53	0.36	0.42	0.10	0.48	0.29

4.2. Evaluation Comparison

In this section, we will further elaborate the objective evaluation indicators of the proposed algorithm. It includes original image performance indicators, selective video encryption performance, and hybrid encryption performance. We choose information entropy, PSNR, and SSIM to evaluate the performance of the algorithm. Table 4 shows the objective performance evaluation of the algorithm.

Table 4. Algorithm objective performance evaluation.

Video Image Type	Video Image 1			Video Image 2		
	Entropy	PSNR	SSIM	Entropy	PSNR	SSIM
original image	7.2191	–	1	7.5146	–	1
Selective video encryption	7.5944	19.983	0.29	7.2495	21.233	0.36
Introducing 4-D hyperchaotic hybrid encryption	7.9899	9.0644	0.0231	7.9897	9.6759	0.0016

We randomly selected two video images to evaluate the performance of the algorithm. The hybrid encryption algorithm is superior to the original selective video encryption from a security perspective, as can be seen from Figure 4. The closer the value of the information entropy is to 8, the better the

algorithm's encryption performance. The smaller the PSNR value, the better the algorithm's encryption performance. The smaller the value of SSIM, the better the encryption performance of the algorithm.

We compare the encryption scheme that is presented in this article with the other seven newest video encryption schemes. These encryption schemes use different algorithms to encrypt different parameters in the video. Therefore, we mainly compare encryption parameters, encryption algorithms, format compatibility, and bit addition. Table 5 gives the results of the comparison.

Table 5. Comparison of existing video encryption schemes with proposed encryption schemes.

Existing Schemes	Encrypted Semantic Element	Format Compliant	Bit Increase	Encryption Algorithm
Xu [25]	IPM, MVDs, T1s, signs of the NZ coefficients	yes	no	Chaos
Abomhara [31]	I frame	no	no	AES
Shahid [32]	T1s, NZ level	yes	no	AES
Fei [33]	IPM, MVD, Signs of residual	yes	yes	Chaos
Sung [34]	Motion vector	yes	yes	RC4
Wei [35]	NALUs	yes	yes	RC4
Wang [36]	IPM, MVD, Quantization coefficients	yes	yes	Hash and AES
Ours	IPM, MVDs, Signs of residual, delta QP	yes	yes	Chaos and AES

As can be seen from Table 5, the format compatibility of the encrypted video data is a research hotspot of video encryption. We all know that the target of video encryption is to damage the economic value of video. In general, our proposed video selective encryption scheme is completely compliant with the decoder, and the compression ratio has not changed, and it has a sufficiently long key length to protect the video information.

4.3. Security Analysis

In the video selective encryption scheme, we use a pseudo-random sequence generator (PSNG) that is composed of double chaotic maps to generate an encryption key. The semantic features of the H.264 slice are used to selectively encrypt important semantic elements in the video sequence. After reading a large amount of literature, we know that the video sequence after H.264 encoding is divided into multiple slices, and the slices are independent of each other. This scheme uses this feature to achieve the selective video encryption. We will analyze the security of selective video encryption schemes from two other aspects.

4.3.1. Key Volume

The key volume is another factor in measuring password security. In this scheme, we did not use AES's original key generation scheme, but used the pseudo-random key generator that was proposed by Xu et al. [25] to generate the key. The generator's key space is $2^{(2*32+16*31)} * \left(2 * 2^{32}\right)\left(2 * 2^{32}\right) = 2^{626}$. It can be seen that the key volume is very large enough to withstand exhaustive attack. The 4-D hyperchaotic system is introduced to further protect the privacy of video images, which can meet the requirements of small amount of video data and strong security video encryption. The size of the key space of the 4-D hyperchaotic system is: $10^{15} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16} = 10^{95} \approx 2^{315}$. Therefore, the proposed algorithm has strong security.

4.3.2. Security Analysis of Selective Video Encryption Scheme

We use the fragmentation feature of H.264 video coding to encrypt important semantic elements in the video. In our experiments, we encode the video into 300 independent slices and encrypt the important semantic elements in each slice. The encrypted video size is the same as the encoded video size (Table 1 has been given). The security of this scheme not only depends on the security of the encryption algorithm, but also on the security of each slice in the video. Fortunately, the slices after

H.264 encoding are independent of each other. Therefore, even if one slice in the video is correctly decrypted, it will not affect the security of other video information.

For the convenience of experiment, we encrypt each video frame as a slice. The benchmark video sequence we use is 300 frames, so each video has 300 slices, so each video is encrypted with 300 keys. We know that it is very difficult to correctly guess a key, and then the probability that 300 keys are correctly guessed is almost zero. Furthermore, we know that each video frame can be encoded into a plurality of mutually independent slices. If each video frame is encoded as n (n > 1) slices, then n * 300 keys are required to encrypt a video. In this case, it is even more difficult to obtain a clear video.

In the encryption process, we choose the CFB mode of AES to ensure that the video is encrypted and it still meets the requirements of real-time and format compatibility. Furthermore, the AES algorithm for 128-bit encryption keys is a good choice compared to most encryption algorithms, because it is assumed that all possible keys are used to crack 128-bit keys at 50 billion keys per second. The required time is 5×1021 years. Therefore, the AES algorithm can not only guarantee the security of the encryption mean, but also ensure the real-time and format compatibility requirements after the video is encrypted. Therefore, theoretically speaking, our selective video encryption scheme is superior to other video encryption schemes.

Figure 5 depicts a subjective evaluation of the algorithm. The first column in Figure 5 is the original video image. The middle column in Figure 5 is the encrypted video image. The last column in Figure 5 is the decrypted video image. Table 2 illustrates the peak signal-to-noise ratio of the algorithm in each color space. Table 3 shows the structural similarity index of the encryption algorithm. The smaller the peak signal-to-noise ratio and the lower the structural similarity index, the better the performance of the selected encryption algorithm.

In Table 4, we will further elaborate the objective evaluation indicators of the proposed algorithm. It includes original image performance indicators, selective video encryption performance, and hybrid encryption performance. We choose information entropy, PSNR, and SSIM to evaluate the performance of the algorithm.

Table 5 compares the existing video encryption schemes. The algorithm performance comparison analysis is completed from the perspectives of encryption semantic elements, format complexity, bit number impact, and encryption unit. The experimental results show that the proposed selective encryption algorithm has better performance. Therefore, the security analysis of selective video encryption algorithm is given in this section.

In summary, this section mainly simulates and analyzes the proposed algorithm. From the experimental results, the performance of the proposed algorithm is better than that of most current selective video encryption algorithms.

5. Conclusions

The selective video encryption scheme that is proposed in this paper cleverly utilizes the segmentation characteristics of H.264 encoding and encrypts the key semantic elements (IPM, MVD, residual coefficient, delta_QP) in each slice. The main contribution of this paper is to propose a new selective video encryption method that is based on video coding technology and 4-D hyperchaotic system. This method provides users with two alternatives: (1) When the amount of encrypted data is small and sufficient security is required, the user can choose to complete selective video encryption based on video encoding and 4-D hyperchaotic system; (2) When the amount of encrypted video data is large and sufficient real-time is required, users can choose a method that is based on video characteristic encoding to implement selective video encryption. In this paper, we use different reference video sequences containing motion, texture, and objects to analyze the perceived quality of encrypted video. Its evaluation indicators include information entropy, PSNR, and SSIM. The experiments show that our proposed video encryption scheme has better encryption effect and less encryption time. In addition, we compare the selective video encryption scheme proposed in this paper with other existing video

encryption schemes. Theoretical analysis confirms that our proposed selective encryption scheme is superior to other video encryption schemes.

Author Contributions: Conceptualization, L.W. and S.C.; methodology, Q.H.; software, Q.H.; validation, S.C., Q.H. and L.W.; formal analysis, N.A.; writing—original draft preparation, Q.H.; writing—review and editing, Q.H.; visualization, S.C.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Xinjiang Uygur Autonomous Region, grant number 2019D01C033, in part by the National Science Foundation of China under Grant 61771416 and U1903213, in part by the CERNET Innovation Project under Grant NGII20180201, and in part by the Creative Research Groups of Higher Education of Xinjiang Uygur Autonomous Region under Grant XJEDU2017T002.

Conflicts of Interest: The authors have declared that no competing interests exist.

References

1. Guang, M.A.; Zhao, P.J.; Cui, M.X. Research and Implementation of Medical Video Encryption and Decryption Player. *China Med. Devices* **2016**, *6*, 87–89.
2. Schwarz, H.; Marpe, D.; Wiegand, T. Overview of the scalable video coding extension of the H. 264/AVC standard. *IEEE Trans. Circuits Syst. Video Technol.* **2007**, *17*, 1103–1120. [CrossRef]
3. Lian, S. *Multimedia Content Encryption: Techniques and Applications*; Auerbach Publications: New York, NY, USA, 2008.
4. Ahn, J.; Shim, H.J.; Jeon, B.; Choi, I. Digital video scrambling method using intra prediction mode. In Proceedings of the 5th Pacific Rim Conference on Multimedia, Tokyo Waterfront City, Japan, 30 November–3 December 2004; pp. 386–393.
5. Khlif, N.; Damak, T.; Kammoun, F.; Masmoudi, N. A very efficient encryption scheme for the H.264/AVC CODEC adopted in Intra prediction mode. In Proceedings of the International Image Processing, Applications and Systems Conference, Sfax, Tunisia, 5–7 November 2014; pp. 1–7.
6. Khlif, N.; Damak, T.; Kammoun, F.; Masmoudi, N. Motion vectors signs encryption for H.264/AVC. In Proceedings of the 2014 1st International Conference on Advanced Technologies for Signal and Image Processing, Sousse, Tunisia, 17–19 March 2014.
7. Lian, S.; Liu, Z.; Ren, Z.; Wang, Z. Selective video encryption based on advanced video coding. In Proceedings of the 6th Pacific Rim Conference on Multimedia, Jeju Island, Korea, 13–16 November 2005; pp. 281–290.
8. Lian, S.; Liu, Z.; Ren, Z.; Wang, H. Secure advanced video coding based on selective encryption algorithms. *IEEE Trans. Consum. Electron.* **2006**, *52*, 621–629. [CrossRef]
9. Shi, T.; King, B.; Salama, P. Selective encryption for H.264/AVC video coding. In Proceedings of the SPIE Security, Steganography, and Watermarking of Multimedia Contents VIII, San Jose, CA, USA, 15 January 2006; p. 607217.
10. Jiang, J.; Liu, Y.; Su, Z.; Zhang, S.; Xing, S. An Improved Selective Encryption for H.264 Video based on Intra Prediction Mode Scrambling. *J. Multimed.* **2010**, *5*, 464–472. [CrossRef]
11. Sbiaa, F.; Kotel, S.; Zeghid, M.; Tourki, R.; Machhout, M.; Baganne, A. A Selective Encryption Scheme with Multiple Security Levels for the H.264/AVC Video Coding Standard. In Proceedings of the 2016 IEEE International Conference on Computer and Information Technology (CIT), Nadi, Fiji, 8–10 December 2016.
12. Khlif, N.; Masmoudi, A.; Kammoun, F.; Masmoudi, N. Secure chaotic dual encryption scheme for H.264/AVC video conferencing protection. *IET Image Process.* **2018**, *12*, 42–52. [CrossRef]
13. Asghar, M.N.; Ghanbari, M.; Fleury, M.; Reed, M.J. Confidentiality of a selectively encrypted H.264 coded video bit-stream. *J. Vis. Commun. Image Represent.* **2014**, *25*, 487–498. [CrossRef]
14. Shahid, Z.; Chaumont, M.; Puech, W. Fast protection of H.264/AVC by selective encryption of CABAC. In Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, New York, NY, USA, 28 June–3 July 2009; pp. 1038–1041.
15. Radanliev, P.; De Roure, D.C.; Nicolescu, R.; Huth, M.; Mantilla, M.; Canaday, S.; Burnap, P. Future developments in cyber risk assessment for the internet of things. *Comput. Ind.* **2018**, *102*, 14–22. [CrossRef]
16. Radanliev, P.; De Roure, D.C.; Nurse, J.R.C.; Montalvo, R.M.; Canady, S.; Santos, O.; Maddox, A.; Burnap, P.; Maple, C. Future developments in standardisation of cyber risk in the Internet of Things (IoT). *SN Appl. Sci.* **2020**, *2*, 169. [CrossRef]

17. Nicolescu, R.; Huth, M.; Radanliev, P.; Roure, D. Mapping the values of IoT. *J. Inf. Technol.* **2018**, *33*, 345–360. [CrossRef]
18. Zhu, Z.; Zhang, W.; Wong, K.W.; Yu, H. A chaos-based symmetric image encryption scheme using a bit-level permutation. *Inf. Sci.* **2011**, *181*, 1171–1186. [CrossRef]
19. Ravichandran, D.; Praveenkumar, P.; Balaguru Rayappan, J.B.; Amirtharaja, R. Chaos based crossover and mutation for securing DICOM image. *Comput. Biol. Med.* **2016**, *72*, 170–184. [CrossRef] [PubMed]
20. Cheng, S.L.; Wang, L.J.; Huang, G.; Du, A.Y. A privacy-preserving image retrieval scheme based secure kNN, DNA coding and deep hashing. *Multimed. Tools Appl.* **2019**, 1–23. [CrossRef]
21. Cheng, S.; Wang, L.; Du, A. Histopathological image retrieval based on asymmetric residual hash and DNA coding. *IEEE Access* **2019**, *7*, 101388–101400. [CrossRef]
22. Hamidouche, W.; Farajallah, M.; Raulet, M.; Deforges, O.; Assad, S.E. Selective video encryption using chaotic system in the SHVC extension. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 19–24 April 2015.
23. Altaf, M.; Ahmad, A.; Khan, F.A.; Uddin, Z.; Yang, X. Computationally efficient selective video encryption with chaos based block cipher. *Multimed. Tools Appl.* **2018**, *77*, 27981–27995. [CrossRef]
24. Wiegand, T.; Sullivan, G.J.; Bjontegaard, G.; Luthra, A. Overview of the H.264/AVC video coding standard. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, *13*, 560–576. [CrossRef]
25. Xu, H.; Tong, X.; Meng, X. An efficient chaos pseudo-random number generator applied to video encryption. *Optik* **2016**, *127*, 9305–9319. [CrossRef]
26. Shukla, P.K.; Khare, A.; Rizvi, M.A.; Stalin, S.; Kumar, S. Applied cryptography using chaos function for fast digital logic-based systems in ubiquitous computing. *Entropy* **2015**, *17*, 1387–1410. [CrossRef]
27. Van Wallendael, G.; Boho, A.; De Cock, J.; Munteanu, A.; Van De Walle, R. Encryption for high efficiency video coding with video adaptation capabilitie. *IEEE Trans. Consum. Electron.* **2013**, *59*, 634–642. [CrossRef]
28. Xu, D.; Wang, R.; Shi, Y.Q. Data Hiding in Encrypted H.264/AVC Video Streams by Codeword Substitution. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 596–606. [CrossRef]
29. Richardson, I.E.G. *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*; Wiley: Hoboken, NJ, USA, 2003.
30. Khlif, N.; Damak, T.; Kammoun, F.; Masmoudi, N. Selective encryption of CAVLC for H.264/AVC. In Proceedings of the 14th International Conference on Sciences and Techniques of Automatic Control & Computer Engineering—STA'2013, Sousse, Tunisia, 20–22 December 2013.
31. Abomhara, M.; Zakaria, O.; Khalifa, O.O.; Zaidan, A.A.; Zaidan, B.B. Enhancing Selective Encryption for H. 264/AVC Using Advanced Encryption Standard. *Int. J. Comput. Electr. Eng.* **2010**, *2*, 223–229. [CrossRef]
32. Shahid, Z.; Chaumont, M.; Puech, W. Fast protection of H.264/AVC by selective encryption of CAVLC and CABAC for I and P frames. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *21*, 565–576. [CrossRef]
33. Peng, F.; Zhu, X.; Long, M. An ROI Privacy Protection Scheme for H.264 Video Based on FMO and Chaos. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 1688–1699. [CrossRef]
34. Sung-Sam, H.; Myung-Mook, H. The study of selective encryption of motion vector based on the S-Box for the security improvement in the process of video. *Multimed. Tools Appl.* **2014**, *71*, 1577–1597.
35. Wei, Z.; Wu, Y.; Ding, X.; Deng, R.H. A scalable and format-compliant encryption scheme for H.264/SVC bitstreams. *Signal Process. Image Commun.* **2012**, *27*, 1011–1024. [CrossRef]
36. Wang, X.; Zheng, N.; Tian, L. Hash key-based video encryption scheme for H: 264/AVC. *Signal Process. Image Commun.* **2010**, *25*, 427–437. [CrossRef]