electronics

MDPI

Article

A Novel Separable Scheme for Encryption and Reversible Data Hiding

Pei Chen [1,2], Yang Lei [1,2], Ke Niu [1,2,*] and Xiaoyuan Yang [1,2]

[1] Key Laboratory of Network and Information Security Under the Chinese People Armed Police Force (PAP), Xi'an 710086, China
[2] College of Cryptography Engineering, Engineering University of PAP, Xi'an 710086, China
* Correspondence: niuke@163.com; Tel.: +86-029-138-0859-3399

Abstract: With the increasing emphasis on security and privacy, video in the cloud sometimes needs to be stored and processed in an encrypted format. To facilitate the indexing and tampering detection of encrypted videos, data hiding is performed in encrypted videos. This paper proposes a novel separable scheme for encryption and reversible data hiding. In terms of encryption method, intra-prediction mode and motion vector difference are encrypted by XOR encryption, and quantized discrete cosine transform block is permutated based on logistic chaotic mapping. In terms of the reversible data hiding algorithm, difference expansion is applied in encrypted video for the first time in this paper. The encryption method and the data hiding algorithm are separable, and the embedded information can be accurately extracted in both encrypted video bitstream and decrypted video bitstream. The experimental results show that the proposed encryption method can resist sketch attack and has higher security than other schemes, keeping the bit rate unchanged. The embedding algorithm used in the proposed scheme can provide higher capacity in the video with lower quantization parameter and good visual quality of the labeled decrypted video, maintaining low bit rate variation. The video encryption and the reversible data hiding are separable and the scheme can be applied in more scenarios.

Keywords: encrypted video; reversible data hiding; sketch attack; difference expansion; QDCT; permutation; embedding capacity

Citation: Chen, P.; Lei, Y.; Niu, K.; Yang, X. A Novel Separable Scheme for Encryption and Reversible Data Hiding. *Electronics* **2022**, *11*, 3505. https://doi.org/10.3390/electronics11213505

Academic Editor: Silvia Liberata Ullo

Received: 1 October 2022
Accepted: 26 October 2022
Published: 28 October 2022

1. Introduction

With the rapid development of the internet, a large number of videos are stored in the cloud [1]. To protect the video content, these videos are stored in encrypted format [2]. To achieve the goals of fast retrieval of encrypted video and protection of video after decryption, data hiding is combined with encryption to embed label information into encrypted video [3]. Furthermore, in order to realize the lossless recovery of video, reversible data hiding in encrypted videos emerges as the times require [4].

Recently, reversible data hiding in encrypted images has been widely studied [5–7]. However, due to the different structures of video coding, most schemes of reversible data hiding in encrypted images are difficult to apply to video. Therefore, the research on reversible data hiding in encrypted videos (RDH-EV) develops slowly and the framework used is singular. In [4], the first RDH-EV scheme is presented. The intra-prediction mode (IPM), the sign of motion vector difference (MVD), and the sign of quantized discrete cosine transform (QDCT) coefficient are encrypted by RC4. Histogram shifting (HS) is implemented to embed information in the encrypted QDCT coefficient. The scheme achieves the separation of encryption and reversible data hiding, which means that decryption and data extraction can be performed in any order. The most recent studies [8–11] apply the separable framework and focus on the improvement of data hiding. The scheme presented in [8] focuses on the embedding capacity of data hiding and adds a scale factor for embedding zone selection to expand according to different capacity requirements. In [9], the authors estimate the distortions caused by embedding information into the QDCT coefficient and set different priorities of the QDCT coefficient for embedding to decrease the interframe distortion drift. This work focuses on imperceptibility of data hiding. In [10], every two adjacent coefficients are grouped into pairs and information is embedded in the pairs by two-dimensional histogram shifting. The scheme utilizes the correlation of adjacent QDCT coefficients and provides high capacity. In [11], the authors translate the framework to H.265. In the scheme, the amplitude of MVD is encrypted by displacing vertical and horizontal components and the signs of MVD and QDCT coefficients are encrypted by RC4. The conventional HS is used for embedding information into QDCT coefficients. In [12], a robust framework based on multidomain embedding is proposed. In the scheme, partial QDCT coefficients of I-frame are permutated by logistic chaotic scrambling and then used for robust embedding. The QDCT coefficients of P-frame are encrypted by the ZUC algorithm and then used for reversible embedding. The scheme focuses on the robustness of data hiding.

To the best of our knowledge, above are all papers in this field in recent years. Although some literature [13–15] combining video encryption and data hiding techniques has been published in recent years, the kind of their data hiding algorithms is traditional steganography [16] and not reversible, so those schemes do not belong to the scope of this paper. It can be seen that in current schemes, there are not many encryption and reversible data hiding techniques used in combination. In terms of encryption technique, IPM and the signs of MVD and QDCT coefficient are usually encrypted by stream cipher in the schemes. In 2017, Minemura et al. [17] presented a novel sketch attack for encrypted video, and attackers can obtain the rough outline of the original video directly from its encrypted counterpart. Unfortunately, the encryption methods used in the current schemes cannot resist the attack. This means that the schemes will not be applicable to application scenarios with high security requirements, such as military. Therefore, RDH-EV needs to adopt more secure encryption methods. In terms of reversible data hiding, only HS is adopted to combine the encryption methods. Conventional reversible data hiding techniques include HS [18–20], difference expansion (DE) [21–23], and integer transform [24,25]. At this stage, how to achieve more diversified combination of reversible data hiding and encryption is also the focus of the field. Based on the above two considerations, a novel scheme combining logistic chaotic scrambling and DE is proposed for RDH-EV in this paper. QDCT blocks (4 × 4) in each frame are permutated by logistic chaotic scrambling. In order to further enhance the visual confusion effect, IPM and MVD signs are also encrypted by XOR encryption. Then information is embedded in fixed regions of each 4 × 4 QDCT block of P frame by DE. The highlights of this paper can be summarized as follows:

- In terms of encryption, the proposed scheme can resist the sketch attack and has better visual security than the existing ones while maintaining format compliance and the bit rate.
- In terms of reversible data hiding, DE is applied to RDH-EV for the first time and can provide labeled video with good quality.
- Encryption and reversible data hiding in the proposed scheme are separable and can be performed in any order, which can be applied in more application scenarios.

The rest of the paper is organized as follows. In Section 2, related work about the separable framework is summarized. The proposed scheme is described in Section 3. The experimental results are shown in Section 4. Section 5 concludes the paper.

2. Related Work

In past decades, many works on video encryption [26–28] and reversible data hiding [18–25] have been done. However, few RDH-EV schemes were put forward and the framework used is singular. In this section, we summarize the framework used in the current schemes.

In most schemes, IPM and MVD sign are encrypted only for content distortion. The separable framework in the schemes is mainly realized by encryption and embedding in QDCT domain. In this framework, the signs of QDCT coefficients are flipped through stream cipher encryption and the amplitudes of QDCT coefficients are used for data embedding. An origin-symmetric histogram shifting algorithm is designed and combined with sign encryption algorithm to achieve separability. The strategy of this shifting algorithm is shown in Figure 1. Its prominent feature is symmetry about the origin. In the embedding phase, the bit embedded by shifting "1" to "2" is consistent with the bit embedded by shifting "−1" to "−2." The bit embedded by shifting "1" to "1" is consistent with the bit embedded by shifting "−1" to "−1." In this way, in the information extraction phase, the information extracted from the symmetrical points will be consistent, as shown in Equation (1). This makes the information extraction independent of the sign encryption algorithm, which means the extraction operation can be performed in both encrypted video and decrypted video. For example, in the encryption phase, the sign of the coefficient "1" is flipped by XOR encryption and "1" is encrypted to "−1." In the embedding phase, the information will be embedded in the encrypted coefficient "−1." According to the different bit to be embedded, the encrypted coefficient "−1" will be shifted to different values. If the bit to be embedded is "1"/"0," the encrypted coefficient "−1" will be shifted to "−2"/"−1." If the information extraction phase is performed in the encrypted video, the information bit "1"/"0" can be extracted from the encrypted shifted coefficient "−2"/"−1." If the information extraction phase is performed in the decrypted video, the encrypted shifted coefficient "−2"/"−1" will be decrypted to be the decrypted shifted coefficient "2"/"1" and then the information bit extracted from the coefficient "2"/"1" is still "1"/"0." Similarly, the strategy of 2DHS satisfies Equation (2).

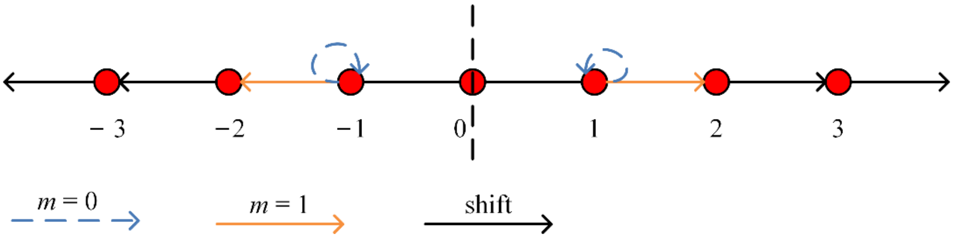


Figure 1. Origin-symmetric histogram shifting.

$$\text{Extract}\left(\pm x'\right) = m \tag{1}$$

where Extract( ) denotes data extraction algorithm; x″ denotes a modified QDCT coefficient; m denotes bit "0" or "1."

$$\text{Extract}\left(\pm x', \pm y'\right) = m_1 m_2 \tag{2}$$

where x″ and y″ denote a pair of modified QDCT coefficients.

It can be seen that the framework perfectly realizes the separation of encryption and data hiding, and is easy to implement. However, because the encryption method used by the framework cannot resist sketch attacks, there is an urgent need for a new separable framework that combines higher visual security encryption methods.

3. Proposed Scheme

The proposed scheme is shown in Figure 2. The proposed scheme includes three parts: video encryption, reversible data hiding in encrypted video, and data extraction and video recovery, which will be elaborated in this section. In video encryption stage, the video owner parses original bitstream and encrypts IPM, QDCT block and MVD to obtain encrypted bitstream. In the data embedding stage, the data hider embeds label information in the encrypted bitstream by DE. Finally, the receiver can extract the label information from encrypted bitstream or decrypted bitstream and restore the original bitstream. As shown in Scenario II, the labeled encrypted bitstream will be decrypted to labeled bitstream. The labeled bitstream can be normally decoded to play and the video quality is good.

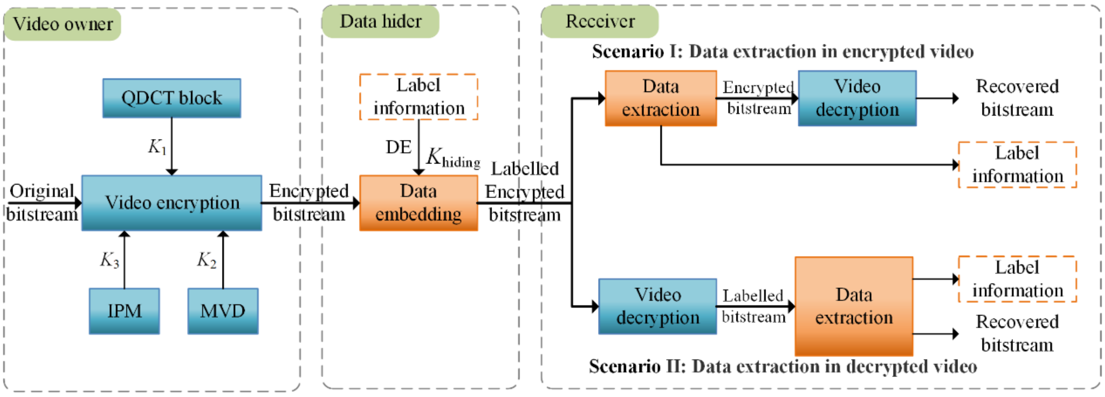


Figure 2. Scheme flow.

3.1. Video Encryption

As mentioned in Section 2, it is difficult to resist the sketch attack by using the encryption methods that IPM and the signs of QDCT and MVD are encrypted by XOR operation (I-Q-M encryption). In this section, the sketch attack presented in [17] is analyzed, and then an effective encryption method is proposed.

3.1.1. Sketch Attack Using Macroblock Bitstream Size (MBS)

As described in [17], the complexity of a macroblock can be inferred from the bitstream size required for encoding the macroblock. Specifically, the more complex the macroblock is, the larger the bitstream size needed to encode the macroblock is, and vice versa. Based on this feature, the following formula can be used to attack encrypted video to obtain the outline of original video:

$$\phi(i,j) = \text{round}\left(255 \times \frac{b(i,j)}{\max\{b(i,j)\}}\right) \tag{3}$$

where $b(i, j)$ denotes the number of bits spent on encoding the (i, j)-th macroblock, max{ } is the maximum function, round( ) is the rounding function, $\phi(i, j)$ denotes the (i, j)-th pixel value of the sketch image.

It can be seen from Equation (3) that MBS attacks mainly obtain plaintext information by calculating the encoding bitstream size of each macroblock. The attack effect is shown in Figure 3. The partial content of original video can be obviously seen from the sketch image obtained by MBS attack. The I-Q-M encryption does not change the bitstream size, so it cannot resist MBS attacks. Obviously, the encryption algorithm can resist MBS attacks through two operations: One is to change the code bitstream size of each element by using a specific replacement operation. In the most ideal case, the bitstream size of each macroblock will tend to be consistent, so the bitstream size of each macroblock cannot reflect its own complexity. Another method is to scramble the positions of each macroblock, so that the sketch image is scrambled. Due to the complexity of video coding structure, it is difficult to adjust the bitstream distribution to an ideal state while maintaining format compatibility. Therefore, the scrambling encryption method is used to improve the security in this paper, while keeping the bit rate unchanged.

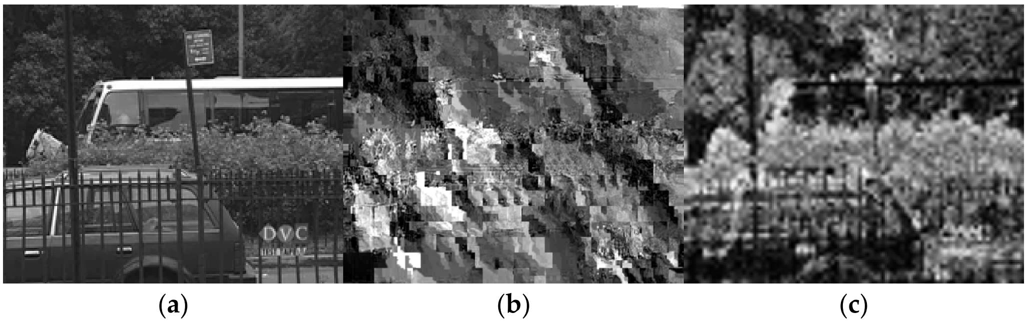


Figure 3. MBS attack effect. (a) Original frame, (b) I-Q-M encrypted frame, (c) sketch image obtained by MBS attack.

3.1.2. Permutation of QDCT Block

Through analysis, the bitstream size spent on encoding a macroblock mainly depends on that spent on encoding the QDCT block. In this paper, the QDCT blocks are permutated instead of the macroblocks that are permutated in [26]. In this way, the sketch image will be scrambled and decoding error of IPM will not occur.

Chaotic mapping is generally used in scrambling algorithms to generate random sequences. Chaotic systems have very complex dynamic behaviors and are widely used in the field of secure communications [29,30]. The one-dimensional logistic mapping formula is as follows:

$$x_{n+1} = \lambda\, x_n(1-x_n) \tag{4}$$

where $\lambda \in (0,4]$ is logistic parameter $x_n \in (0,1)$. When $3.57 < \lambda \leq 4$, the map is in a chaotic state. The closer λ is to 4, the more uniformly the x range is distributed at (0,1). Therefore, λ is equal to 4 in this paper.

When x_0 is given, a random sequence can be generated according to Equation (4), from which the random sequence $S = (s_1, \ldots, s_{l-1}, s_l)$ with length l can be intercepted. Then, the index sequence $K = (k_1, \ldots, k_{l-1}, k_l)$ can be obtained according to Equation (5) and will be used for permutation.

$$[S', K] = Sort(S) \tag{5}$$

Let $\mathrm{B}_{i,j}$ denote the (i, j)-th QDCT block in a frame that has $M \times N$ QDCT blocks. The QDCT blocks in a frame are divided into two sequences Q_0 and Q_1 as follows:

$$Q_0 = \left(\mathrm{B}_{1,1}\right) \tag{6}$$

$$Q_1 = \left(\mathrm{B}_{1,2}, \ldots, \mathrm{B}_{1,N}, \mathrm{B}_{2,1}, \ldots, \mathrm{B}_{2,N}, \ldots, \mathrm{B}_{M,1}, \ldots, \mathrm{B}_{M,N}\right) \tag{7}$$

The permutation process is shown in Figure 4. The sequence Q_1 will be permutated to generate the permutated sequence Q_1' according to Equation (8) with the index sequence K_1. The QDCT block will be rearranged according to Q_1'.

$$Q_1'(n) = Q_1\left(K_1(n)\right), \quad n = 1, 2, \ldots, M \times N - 1 \tag{8}$$

where $K_1(n)$ denotes the n-th number in the permutation key K_1 and $Q_1(n)$ denotes the n-th $B_{i,j}$ in the sequence Q_1.

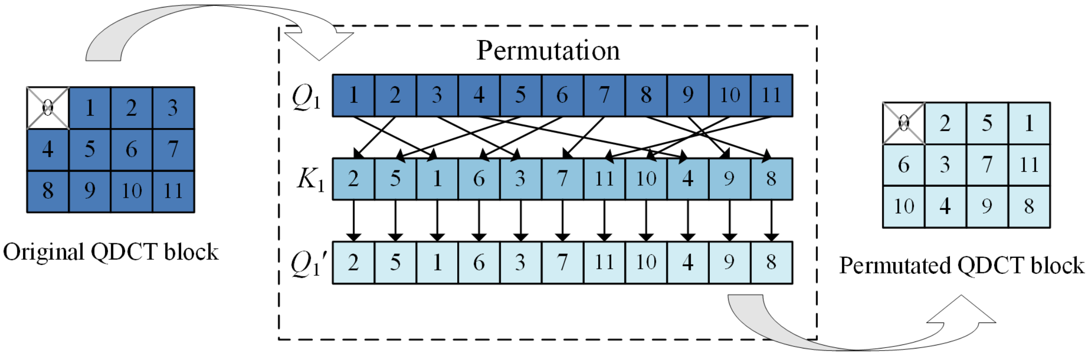


Figure 4. Permutation of QDCT block.

3.1.3. Encryption of MVD Sign

The MVD codeword structure is encoded as [*M* zero][1][*INFO*] using Exp-Golomb code in the video encoding process [31]. *INFO* is the information value of *M* bit. The *M* and *INFO* can be calculated as follows:

$$\begin{cases} M = \text{floor}\left(\log_2\left[codeNum+1\right]\right) \\ INFO = codeNum+1-2M \end{cases} \tag{9}$$

where *codeNum* is calculated according to the value *k* to be encoded, as follows:

$$codeNum = \begin{cases} 2|k|, & k \le 0 \\ 2|k|-1, & k > 0 \end{cases} \tag{10}$$

The characteristics of the codeword are shown in Table 1. It can be seen from the table that the codewords of the opposite numbers are different only in the last bit. Therefore, the last bit of the MVD codeword is encrypted by XOR using stream cipher to realize MVD sign encryption in this paper. It should be noted that an MVD contains horizontal and vertical components, and two components are encrypted independently, as follows:

$$\boldsymbol{mvd}_i\left(mvd\_h_i, mvd\_v_i\right) = \begin{cases} -mvd\_h_i, & \text{if } S_1\left(2i-1\right)=1 \\ mvd\_h_i, & \text{if } S_1\left(2i-1\right)=0 \\ -mvd\_v_i, & \text{if } S_1\left(2i\right)=0 \\ mvd\_v_i, & \text{if } S_1\left(2i\right)=0 \end{cases} \tag{11}$$

where mvd_i denotes the *i*-th MVD; mvd_h_i and mvd_v_i represent the horizontal and vertical components of MVD respectively; S_1 is generated by the ZUC algorithm [32] with the key K_2.

Table 1. Exp-Golomb codeword of MVD.

MVD	Exp-Golomb Codeword
0	1
1, −1	010, 011
2, −2	00100, 00101
3, −3	00110, 00111
4, −4	0001000, 0001001

3.1.4. Encryption of IPM Codeword

To enhance visual safety, the encryption of IPM is added. As in most schemes, XOR encryption is performed on the following three codewords when flag is 0, as shown in

Figure 5. The IPM of the first row or column will not be encrypted because it may cause decoding errors. The encryption method is as follows:

$$p_i' = p_i \oplus S_2(i) \tag{12}$$

where p_i denotes the i-th bit of the whole IPM codewords in a frame and S_2 is generated by ZUC algorithm with the key K_3.

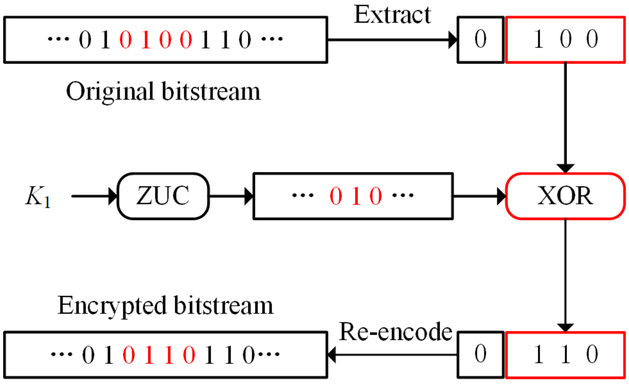


Figure 5. IPM encryption.

3.1.5. Adaptive Key Generation

In order to further strengthen the security of the encryption method and facilitate key management, each frame of the video will be encrypted with a different permutation key K_f, which is adaptively generated based on the data of each frame and an initial seed S_0 with 256 bits. The steps of adaptive generation are as follows:

Step 1: Arrange the DC coefficients of all QDCT blocks in a frame into a one-dimensional sequence from large to small $D = (dc_1, dc_2, \ldots, dc_{MN})$.

Step 2: Calculate the adaptive value x_a of the frame according to Equation (11).

$$x_a = \frac{\sum_{i=1}^{MN/2} dc_{2i-1}}{\sum_{i=1}^{MN} dc_i} \tag{13}$$

Step 3: Generate a chaotic sequence $(x_k)_{k=0}^{\infty}$ by logistic mapping using λ and x_a. Extract $(x_k)_{k=1000}^{k=1255}$ from $(x_k)_{k=0}^{\infty}$ and generate binary random sequence S_p according to Equation (12).

$$S_p(i)\begin{cases} 1, & x_k \le 0.5 \\ 0, & x_k > 0.5 \end{cases} \tag{14}$$

Step 4: XOR S_p and S_0 to generate a new sequence S_n.

Step 5: Divide S_n into 16 subsequences S_{ni} ($i = 1, \ldots, 16$), each of which has a length of 16 bits. Calculate x_0 of the frame according to Equation (13).

$$x_0 = \sum_{i=1}^{16} \mathrm{mod}\left(\sum_{j=1}^{16} S_{ni}(j), 10\right) 10^{-i} \tag{15}$$

Step 6: Calculate the permutation sequence K_f of the frame according to Equations (4) and (5) using λ and x_0.

3.2. Data Hiding in Encrypted Videos

3.2.1. Embedding Region Selection

As mentioned in Section 3.1.2, the complexity of macroblocks can be inferred according to the bitstream size spent on encoding its QDCT block. The human visual system is less sensitive to small changes in complex texture regions. Therefore, the QDCT block with large bitstream size is selected as the embeddable region to achieve better imperceptibility. In this paper, QDCT blocks, of which DC coefficients are not 0 and of which AC coefficients are not all 0 are selected for information embedding. In these blocks, AC_{10} and AC_{12} of each QDCT block form a pair of coefficients, as shown in Figure 6.

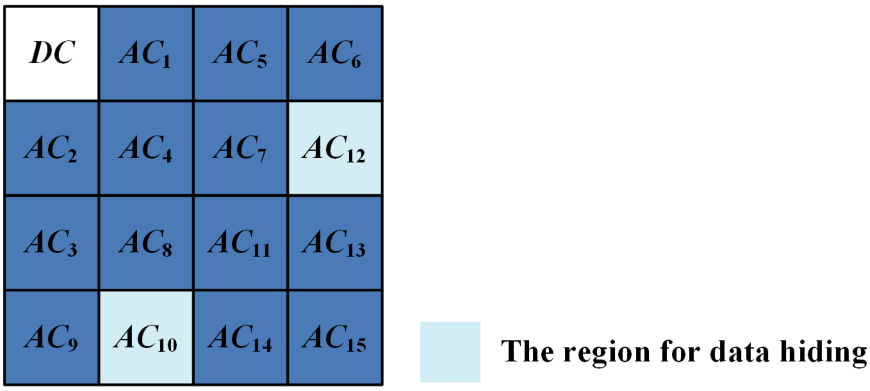


Figure 6. The region for data hiding.

3.2.2. Difference Expansion in Encrypted Videos

Let (x, y) denote the coefficient pair. The average value a and difference h of the coefficient pair can be calculated as follows:

$$\begin{aligned} a &= \left\lfloor \frac{x+y}{2} \right\rfloor, \\ h &= x - y, \end{aligned} \tag{16}$$

The inverse transformation of the above formula is as follows:

$$\begin{aligned} x &= a + \left\lfloor \frac{h+1}{2} \right\rfloor, \\ y &= a - \left\lfloor \frac{h}{2} \right\rfloor, \end{aligned} \tag{17}$$

The difference h will be expanded and one bit m can be embedded in the expanded difference h', as follows:

$$h' = 2h + m \tag{18}$$

Then, the original coefficient pair will be modified as follows:

$$\begin{aligned} x' &= a + \left\lfloor \frac{h'+1}{2} \right\rfloor, \\ y' &= a - \left\lfloor \frac{h'}{2} \right\rfloor, \end{aligned} \tag{19}$$

The whole embedding process as shown in Algorithm 1.

Algorithm 1. Embedding algorithm (Single frame)

Input: Encrypted bitstream *F*, label information *m*

Output: Encrypted bitstream containing label information *F′*

```
1. for $Q_1'$ in F do://Traverse the QDCT block in the encrypted video stream
2.     $DC$ ← The DC coefficient of the QDCT block
3.     $NZ$ ← Number of non-zero AC coefficients excluded AC10 and AC12 in QDCT block
4.     if $DC > 0$ and $NZ > 0$ then
5.         $(x_i, y_i)$ ← $AC_{10}$ and $AC_{12}$ of the i-th QDCT block in $Q_1'$
6.         $a_i$, $h_i$ ← the average value a and difference h of $(x_i, y_i)$ calculated by Equation (14)
7.         $h_i'$ ← the expanded difference calculated according to Equation (16) and m
8.         $(x_i', y_i')$ ← the modified coefficient pair calculated by Equation (17)
9.         $Q_1'$ ← Replace $(x_i, y_i)$ with $(x_i', y_i')$
10.    end if
11. end for
12. $F''$ ← Reencode $Q_1'$ into bitstream
13. return $F'$
```

3.3. Information Extraction and Video Recovery

Information extraction is the inverse process of data hiding. Information can be extracted as follows:

$$m = \mathrm{mod}(h', 2) \tag{20}$$

The original difference *h* can be obtained according to Equation (19). Then the original coefficient pair (*x*, *y*) can be recovered by Equation (15).

$$h = \left\lfloor \frac{h'}{2} \right\rfloor \tag{21}$$

The extraction operation can be implemented in encrypted video or decrypted video. The original QDCT block can be accurately restored to recover the original video.

3.4. Application Scenario

If an RDH-EV scheme is separable, it can be applied to more application scenarios. Assume the original bitstream, the label information to be embedded, the encryption operation, the encryption key, the data-embedding operation, the encrypted bitstream, the labeled bitstream, the decryption operation and the data-extraction operation as ***O***, ***M***, E, ***K***$_\mathrm{e}$, H, ***O***$_\mathrm{E}$, ***O***$_\mathrm{H}$, D and R. If video encryption and reversible data hiding are separable, the following equation must hold [13]:

$$\begin{aligned} \boldsymbol{O}_{\mathrm{EH}} &= \mathrm{H}\left(\mathrm{E}\left(\boldsymbol{O}, \boldsymbol{K}_{\mathrm{e}}\right), \boldsymbol{M}\right), \\ \boldsymbol{O}_{\mathrm{HE}} &= \mathrm{E}\left(\mathrm{H}\left(\boldsymbol{O}, \boldsymbol{M}\right), \boldsymbol{K}_{\mathrm{e}}\right) \end{aligned} \tag{22}$$

$$\mathrm{R}\left(\mathrm{D}\left(\boldsymbol{O}_{\mathrm{HE}}, \boldsymbol{K}_{\mathrm{e}}\right)\right) = \mathrm{D}\left(\mathrm{R}\left(\boldsymbol{O}_{\mathrm{HE}}\right), \boldsymbol{K}_{\mathrm{e}}\right) = \mathrm{R}\left(\mathrm{D}\left(\boldsymbol{O}_{\mathrm{EH}}, \boldsymbol{K}_{\mathrm{e}}\right)\right) = \mathrm{D}\left(\mathrm{R}\left(\boldsymbol{O}_{\mathrm{EH}}\right), \boldsymbol{K}_{\mathrm{e}}\right) = \left(\boldsymbol{O}, \boldsymbol{M}\right) \tag{23}$$

Specifically, the data embedding, data extraction and video recovery can be implemented in both plain and encrypted domains. As shown in Figure 2, the proposed scheme can realize these operations. It indicates that the scheme in this paper is separable. Figure 7 shows three application scenarios in cloud environment.

1. User 1 only owns the encryption key. User 1 can download the labeled encrypted bitstream from cloud and decrypted the bitstream to obtain labeled bitstream. The labeled bitstream can be normally decoded to play video with good quality.
2. Cloud provider only owns the hiding key. Cloud provider can directly extract the label information from the encrypted bitstream to complete indexing and authentication or other operations.
3. User 2 is authorized by user 1 and cloud provider, and owns the encryption key and the hiding key. User 2 can not only extract the label information for management, but also decrypt the bitstream to obtain the original video.

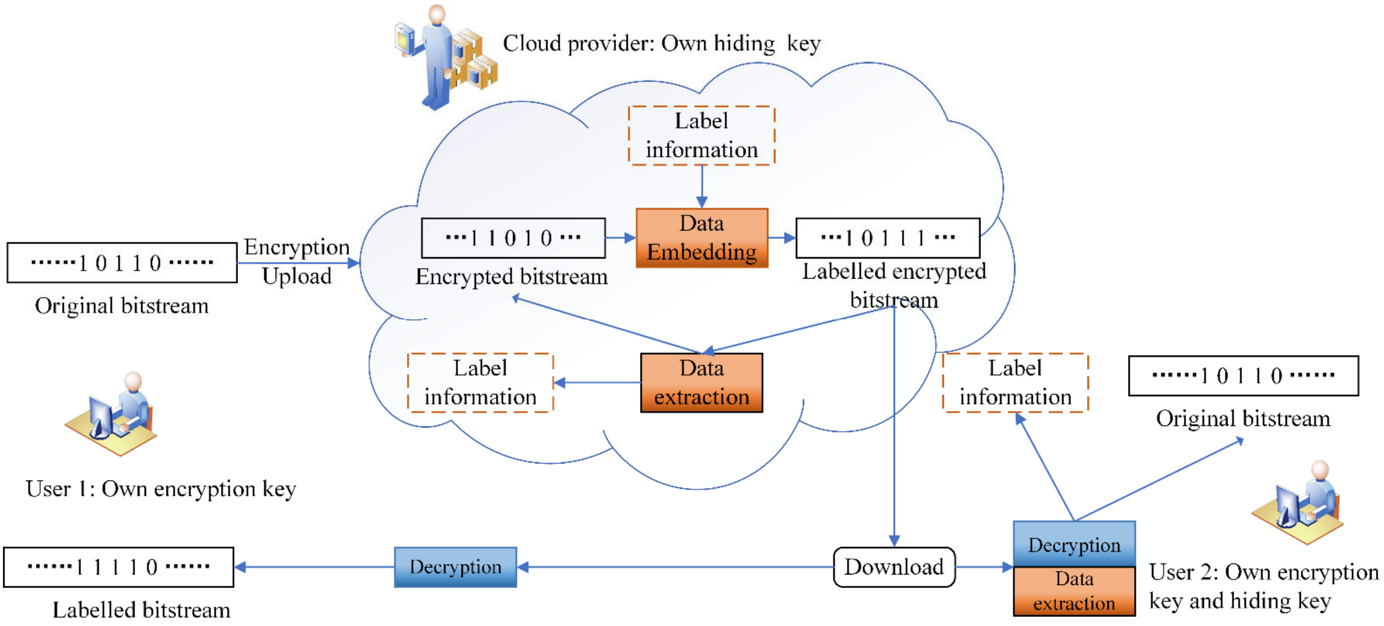


Figure 7. Application scenarios in cloud environment.

4. Experimental Results with Analysis

The effectiveness of the proposed scheme has been investigated through a series of simulation experiments. Section 4.1 introduces the video sequence used, the experimental runtime environment and the methods used for comparison. The experiments for the encryption method are analyzed in Sections 4.2–4.5. The visual effect of encryption method is analyzed in Section 4.2. The ability of resisting sketch attack is shown in Section 4.3. Section 4.4. analyzes computational complexity. The encryption space is analyzed in Section 4.5. Sections 4.6–4.9 are the experimental analysis of the embedding algorithm. The visual quality of labeled decrypted video is deeply analyzed in Section 4.6. The embedding capacity is analyzed in Section 4.7. Section 4.8 describes the robustness. The reversibility is analyzed in Section 4.9. Section 4.10 reports the bit rate variation caused by the encryption or embedding.

4.1. Experiment Setting

1. Video sequences

For the objectivity of the experimental results, nine standard common intermediate format (352 × 288) video sequences (container, stefan, coastguard, foreman, news, akiyo, bus, flower and bridge-close) are used for simulation. These video sequences can be obtained on the YUV Video Sequence website [32]. The nine video sequences selected are rich in content, including fast motion, slow motion, complex texture and simple texture scenes. In this paper, the luminance components of the first 100 frames of each video sequence are used for experiments.

2. Experimental operation environment and parameter setting

All experiments in this section were simulated on a computer equipped with an Intel i7-8550U 4 GHz CPU and 8 GB memory. The simulation experiment runs in MATLAB H.264 codec [33]. The group of pictures (GOP) is set to "IPPP" with length of 20, which means that the first frame is encoded as I frame and the rest 19 frames are encoded as P frame. The default quantization parameter (QP) is 28.

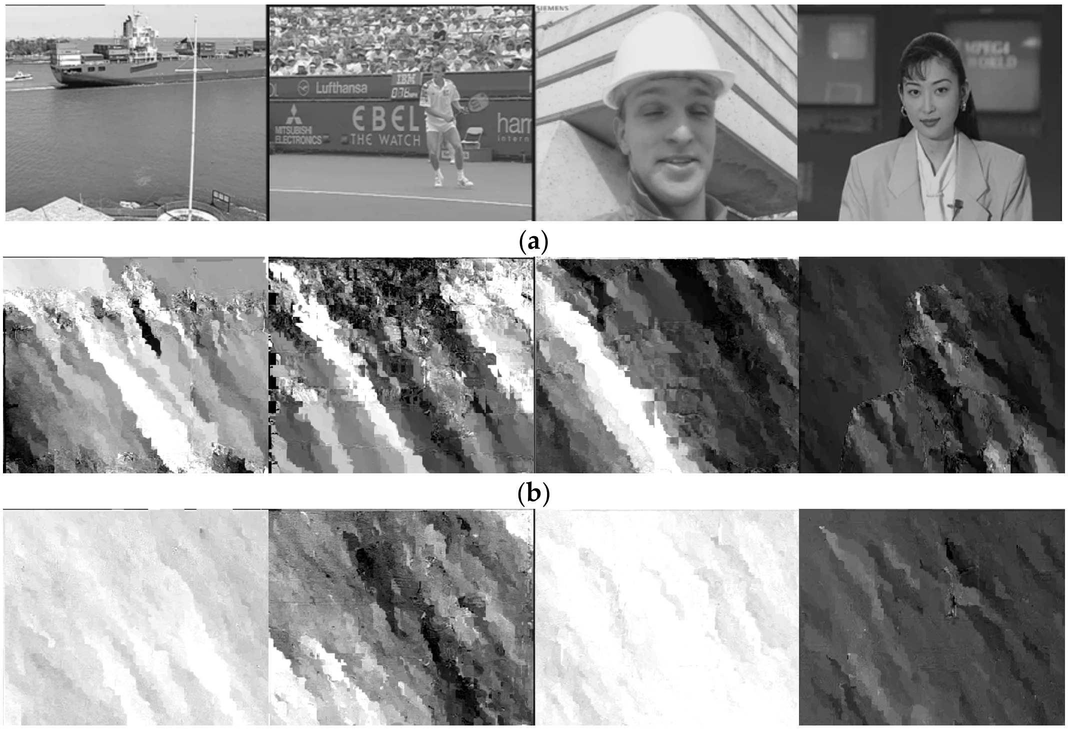
3. Contrast experiment setting

In terms of encryption algorithm, it is compared with the encryption algorithm commonly used in the current scheme [4,8–10] (IPM, QDCT coefficient sign and MVD sign encryption, recorded as I-Q-M encryption). As far as we know, the difference expansion algorithm is applied to the encrypted video for the first time in this paper, and no other expansion algorithm in encrypted video that can be used for comparison exists. Therefore, in terms of the embedding algorithm, this paper makes longitudinal comparative experiments as detailed as possible.

4.2. Visual Encryption Effect

The visual encryption effect of encryption algorithm is mainly evaluated by subjective video quality and objective evaluation metrics. The objective evaluation metrics include peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [34,35].

The subjective video quality experiment results are shown in Figure 8. It can be seen that the content of the original frame cannot be perceived from the encrypted frames. Compared with the I-Q-M encryption method, the visual effect of the proposed encryption method is smoother and shows less texture. This is because the proposed encryption method scrambles the QDCT blocks, thus further confusing the texture in the video.



(**a**)

(**b**)

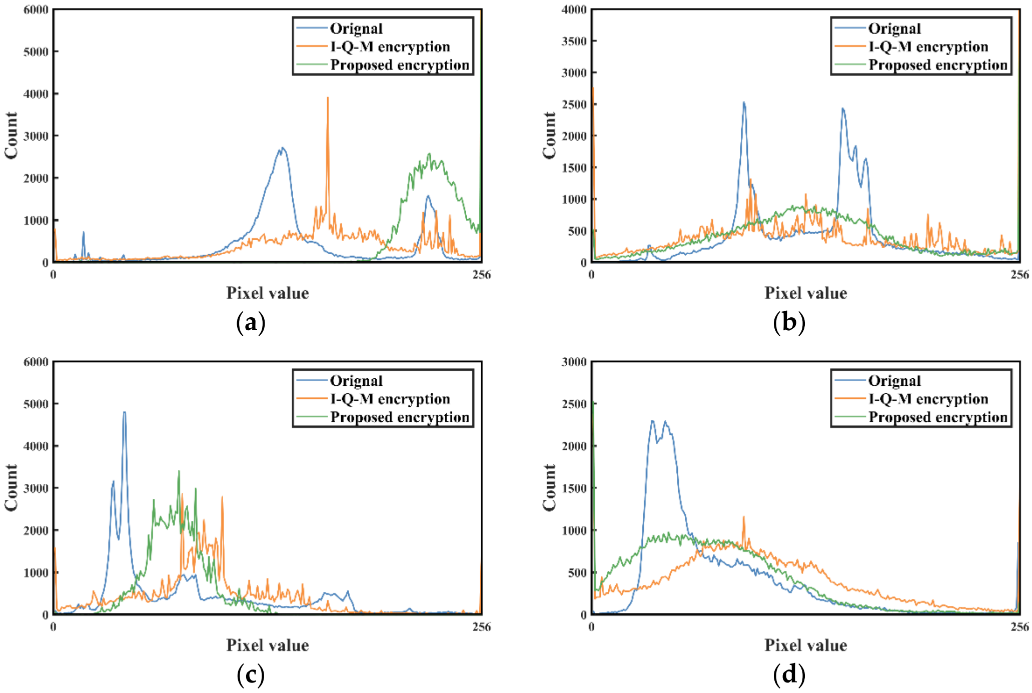
(**c**)

Figure 8. Encryption effect comparison. (**a**) Original frame, (**b**) I-Q-M encryption, (**c**) proposed encryption.

As shown in Table 2, PSNR and SSIM before and after video encryption are given. It can be seen that the PSNR and SSIM of the frame encrypted by proposed method shall not exceed 13.03 dB and 0.1015. Compared with I-Q-M encryption, the PSNR and SSIM of the frames encrypted by proposed method are similar. In addition, the histograms of the two kinds of encrypted videos are statistically analyzed. As shown in Figure 9, the histogram of the encrypted frame is different from that of the original frame, indicating that the proposed encryption algorithm has good performance. From the perspective of subjective video quality and objective standards, the encryption method used in this paper achieves a better visual confusion effect and meets the requirements of visual security.

Table 2. Comparison between PSNR and SSIM after encryption.

Video Sequence	PSNR/dB			SSIM		
	Original	I-Q-M Encryption	Proposed Encryption	Original	I-Q-M Encryption	Proposed Encryption
container	37.94	10.51	8.47	0.9977	0.0754	−0.0011
stefan	37.69	9.03	11.21	0.9971	−0.0022	0.0509
foreman	38.67	7.01	7.93	0.9985	−0.1358	0.0306
news	39.54	10.83	13.03	0.9986	−0.0578	−0.0726
akiyo	40.80	11.69	11.32	0.9988	0.1306	0.0103
bus	36.78	10.91	12.37	0.9967	0.1397	0.0913
flower	37.71	10.60	11.01	0.9984	0.2790	0.1015
bridge-close	36.98	8.61	8.38	0.9983	0.1405	0.0429



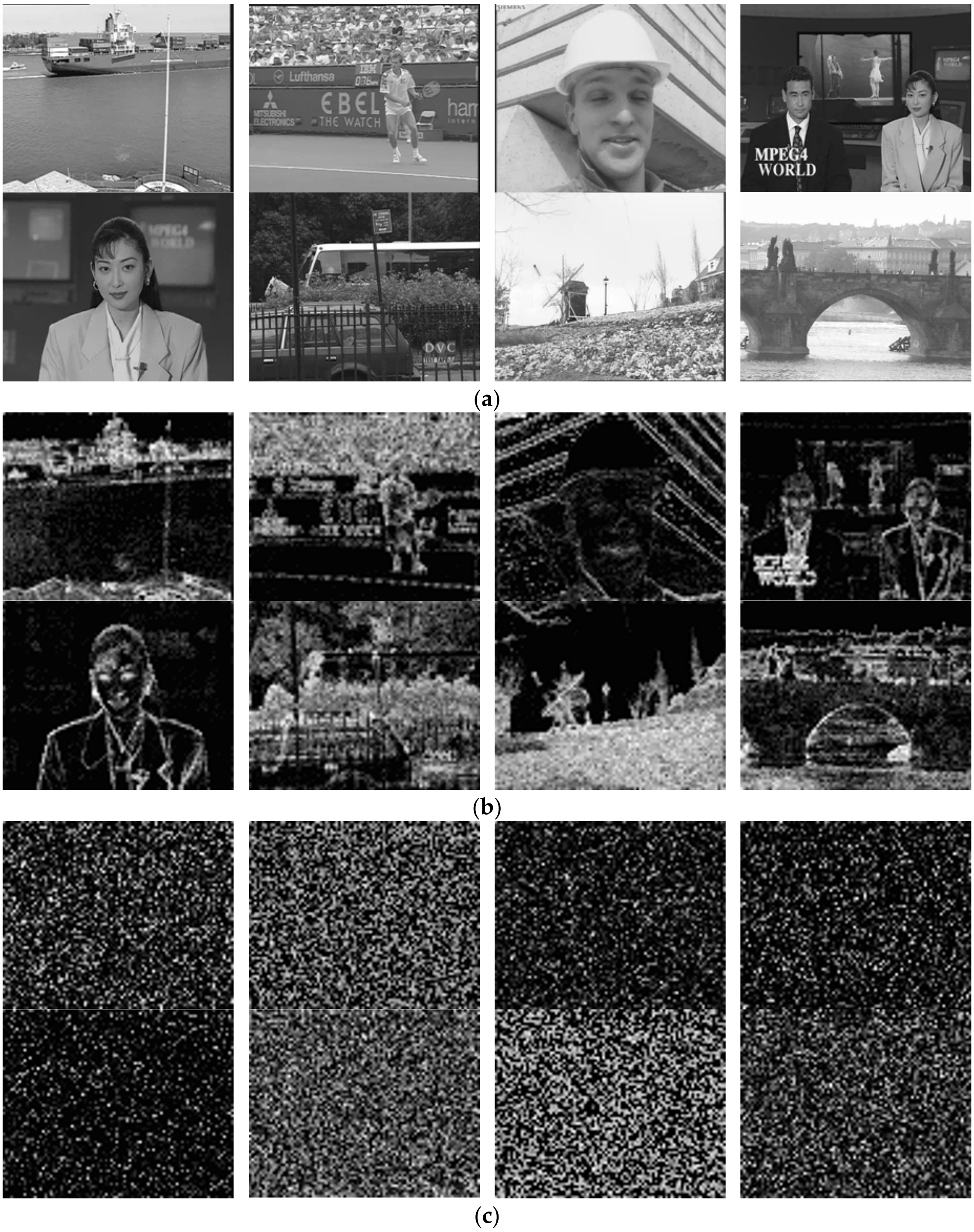
(**a**) (**b**) (**c**) (**d**)

Figure 9. Video frame histogram. (**a**) Container, (**b**) Stefan, (**c**) news, (**d**) bus.

4.3. Resistance to Sketch Attacks

Resisting sketch attacks is a remarkable characteristic of the proposed encryption method. At present, none of the encryption methods used in RDH-EV has this characteristic.

Figure 10 shows the result of sketch attack on videos encrypted by the I-Q-M method and the proposed method. It can be clearly seen that the I-Q-M encryption method cannot resist the sketch attack, and the outline of the original video content will still be stolen under the ciphertext state. Some information of video content can be easily obtained from these outline images, such as characters, scenes and vehicles. Under the sketch attack, the sketch image of the video encrypted by the proposed method is a noisy image without obvious information, from which any content of the original video cannot be distinguished.


(**a**)

(**b**)

(**c**)

Figure 10. Effect of resisting sketch attack. (**a**) Original frame, (**b**) I-Q-M encryption, (**c**) proposed encryption.

4.4. Computational Complexity Analysis

The proposed encryption method includes three parts: IPM encryption, MVD sign encryption and QDCT block permutation. IPM encryption and MVD sign encryption are XOR encryption of codewords, and their time complexity mainly depends on the number of corresponding codewords. In a 4 × 4 macroblock, the IPM codeword used for encryption is 3 bits. Therefore, in a frame containing $M \times N$ macroblocks, the time complexity of

IPM encryption is O (3(M−1)(N−1)). In terms of MVD, the maximum number of sign bits in P frame is 2 × M × N, so the time complexity of MVD sign encryption is O (2MN). Scrambling QDCT blocks requires sorting and scrambling operations, and the time complexity is O ($MN\log_2 MN+MN$). Since 3(M−1)(N−1) and 2MN are much smaller than ($MN\log_2 MN+MN$), the total time complexity of the proposed encryption operations is O ($MN\log_2 MN$).

The I-Q-M encryption method [4,8–10] includes IPM encryption, QDCT sign encryption and MVD sign encryption. The time complexity of the IPM encryption and the MVD encryption is consistent with that of our encryption methods. The time complexity of the IPM encryption is O (3(M−1)(N−1)) and of the MVD encryption is O (2MN). Signs of non-zero AC coefficients will be encrypted by XOR in the QDCT sign encryption and the time complexity of this encryption depends on the number of non-zero AC coefficients. Since the non-zero AC coefficients of a 4 × 4 macroblock cannot exceed 15, the time complexity of the QDCT sign encryption is O (15MN). To sum up, the total time complexity of the I-Q-M encryption is O (MN), which is lower than that of the methods proposed in this paper.

The encryption method proposed in this paper is based on the coding bitstream. First, some elements, IPM and QDCT coefficients in the bitstream, are decoded, and then the corresponding encryption is performed. Finally, the encrypted bitstream is obtained by re-encoding the encrypted elements. Although the complexity of the proposed encryption method is higher than that of the I-Q-M encryption method, the proposed encryption method only partially decodes and encodes in the whole process and the time cost in the actual operation process is not too large.

4.5. Encryption Space Analysis

Generally speaking, the encryption space should be large enough to resist brute force attacks. As described in Section 3.1, the encryption keys used in proposed encryption method include K_1, K_2 and K_3. The size of the keys is all 256 bits and they are used to generate random sequences by ZUC algorithm, which is difficult to attack [32]. However, the attacker may attempt to enumerate the encryption operations instead of the encryption key. Therefore, the encryption space of I frame and P frame is analyzed. Let m_s denote the number of MVD sign bits and m_Q denote the number of non-zero AC coefficients sign bits in a frame. For the proposed encryption method and the I-Q-M encryption, the encryption space of I frame is calculated by the following formula:

$$\begin{aligned} E_{\mathrm{I}}^{\mathrm{Proposed}} &= \left(MN-1\right)! \times 2^{3(M-1)(N-1)}, \\ E_{\mathrm{I}}^{\text{I-Q-M}} &= 2^{m_Q} \times 2^{3(M-1)(N-1)} \end{aligned} \tag{24}$$

The encryption space of P frame is calculated by the following formula:

$$\begin{aligned} E_{\mathrm{P}}^{\mathrm{Proposed}} &= \left(MN-1\right)! \times 2^{m_{\mathrm{S}}}, \\ E_{\mathrm{P}}^{\text{I-Q-M}} &= 2^{m_Q} \times 2^{m_{\mathrm{S}}} \end{aligned} \tag{25}$$

It can be seen from Equations (22) and (23) that, for the proposed encryption method, the size of encryption space mainly depends on the number of macroblocks and the number of MVD. For the I-Q-M encryption, the size of encryption space mainly depends on the number of macroblocks, the number of MVD and the number of non-zero AC coefficients. As shown in Table 3, although the encryption space of the P frame will be smaller than that of the I frame, the encryption space of both the proposed method and the I-Q-M encryption is large enough to resist brute force attacks. In addition, P frame content is decoded based on the I frame content. As long as I frame encryption is not cracked, even if P frame encryption is cracked, P frame content cannot be obtained. Accordingly, the subsequent decryption of the encrypted P frame requires the complete content of the forward P frame, otherwise the original content cannot be recovered. To sum up, for the proposed encryption, the encryption space of *i*-th frame in a GOP can be calculated as follows:

$$E_i = (MN-1)! \times 2^{3(M-1)(N-1)} \times (i-1)(MN-1)! \times 2^{m_S} \tag{26}$$

When $i = 1$, E_i takes the minimum value E_1, which is equal to E_I. In other words, the minimum encryption space to crack a frame in a GOP is E_I, which is large enough to resist the brute force attack.

Table 3. The encryption space of I frame and P frame.

Video Sequence	Proposed Method		I-Q-M Ecnryption	
	E_I	E_P	E_I	E_P
container	$6335! \times 2^{18531}$	$6335! \times 2^{1031}$	2^{36034}	2^{5042}
stefan	$6335! \times 2^{18531}$	$6335! \times 2^{1188}$	2^{45511}	2^{15327}
foreman	$6335! \times 2^{18531}$	$6335! \times 2^{1200}$	2^{30947}	2^{6298}
news	$6335! \times 2^{18531}$	$6335! \times 2^{1112}$	2^{31119}	2^{4540}
akiyo	$6335! \times 2^{18531}$	$6335! \times 2^{938}$	2^{25694}	2^{2227}
bus	$6335! \times 2^{18531}$	$6335! \times 2^{1115}$	2^{48383}	2^{17894}
flower	$6335! \times 2^{18531}$	$6335! \times 2^{1088}$	2^{53570}	2^{14067}
bridge-close	$6335! \times 2^{18531}$	$6335! \times 2^{1366}$	2^{39816}	2^{12370}

4.6. Visual Quality of Labeled Decrypted Video

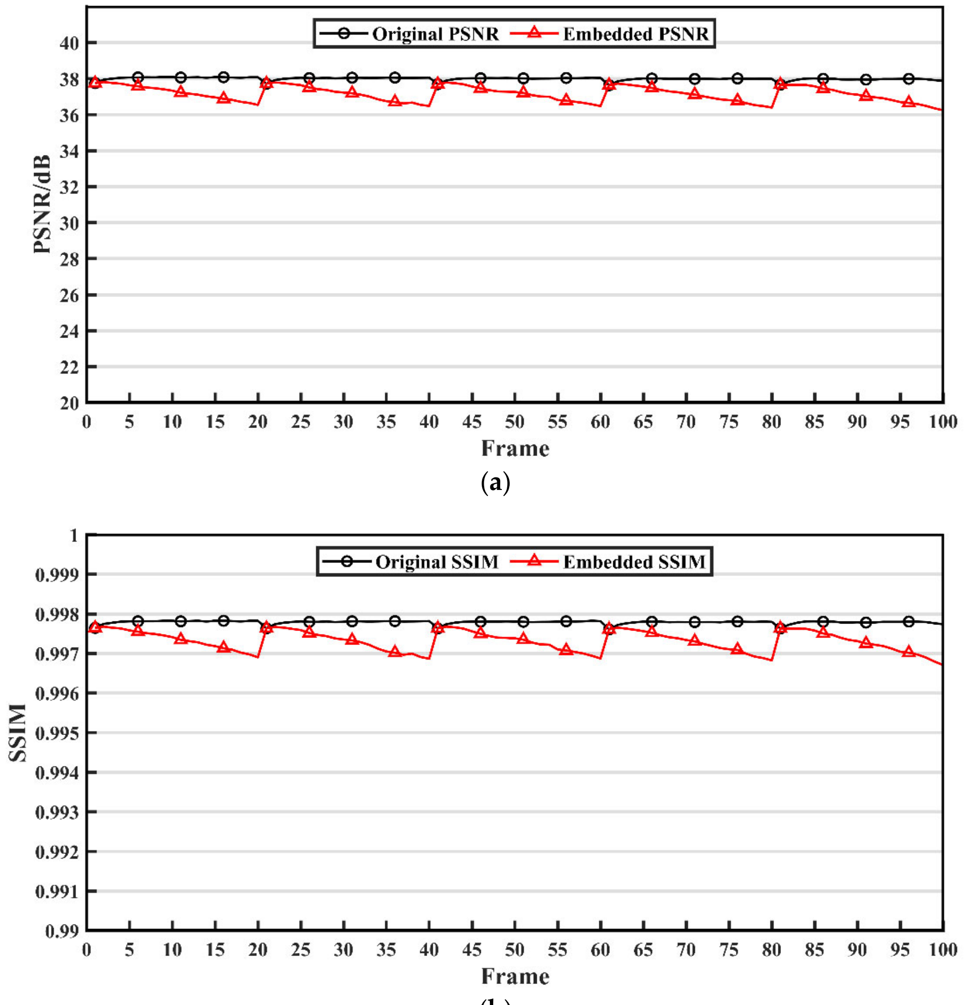
Because there is no difference expansion algorithm in the ciphertext domain to compare, this paper uses PSNR and SSIM to evaluate the visual quality of labeled decrypted video under different QP coding as vertically as possible.

Table 4 shows the average PSNR and SSIM of original video and labeled video with QP of 24, 28, and 32. The data in the table are obtained in full embedding. It can be seen from the table that the PSNR and SSIM of the embedded video have decreased compared with those of the original video, but the decrease is not large. The average PSNR value decreases by 3.57dB at most, and the average SSIM value decreases by 0.0027 at most, which can meet the requirements of good visual quality. Figure 11 shows the PSNR and SSIM changes of the container video sequence after embedding. It can be seen that, in a GOP, the PSNR and SSIM of labeled videos decrease slowly, reaching the lowest in the last frame. In other words, the visual quality of the last frame in a GOP is the worst. As shown in Figure 12, the 20th frame (GOP length in this paper is 20) of each labeled video sequence is given. As can be seen from the figure, even in the 20th frame, it is difficult to see the traces of information embedding. In the cloud environment, when authorized users directly decrypt videos for rough browsing, the embedding algorithm in this paper can provide good video quality.

Table 4. Average PSNR and SSIM before and after embedding.

Video Sequence	QP	PSNR		△PSNR	SSIM		△SSIM
		Original	Embedded		Original	Embedded	
container	24	40.81	39.36	−1.45	0.9988	0.9984	−0.0005
	28	38.05	37.31	−0.74	0.9978	0.9974	−0.0004
	32	35.18	34.69	−0.49	0.9957	0.9952	−0.0005
stefan	24	40.65	38.31	−2.34	0.9986	0.9975	−0.0011
	28	37.60	35.92	−1.68	0.9971	0.9957	−0.0014
	32	34.33	33.12	−1.21	0.9938	0.9918	−0.0020
foreman	24	40.97	39.54	−1.43	0.9991	0.9987	−0.0004

	28	38.53	37.75	−0.78	0.9984	0.9981	−0.0003
	32	35.91	35.43	−0.48	0.9971	0.9968	−0.0003
news	24	42.25	39.88	−2.37	0.9993	0.9987	−0.0006
	28	39.72	38.27	−1.45	0.9987	0.9981	−0.0005
	32	36.83	36.04	−0.79	0.9974	0.9969	−0.0005
akiyo	24	43.26	41.97	−1.29	0.9993	0.9991	−0.0002
	28	41.11	40.47	−0.64	0.9989	0.9987	−0.0002
	32	38.49	38.11	−0.38	0.9980	0.9978	−0.0002
bus	24	39.92	37.12	−2.79	0.9986	0.9973	−0.0013
	28	36.94	34.83	−2.10	0.9972	0.9955	−0.0017
	32	33.84	32.10	−1.74	0.9943	0.9916	−0.0027
flower	24	41.09	37.52	−3.57	0.9993	0.9983	−0.0010
	28	37.71	35.64	−2.07	0.9984	0.9974	−0.0010
	32	34.15	32.97	−1.17	0.9963	0.9952	−0.0011
bridge-close	24	40.09	37.22	−2.87	0.9992	0.9983	−0.0008
	28	37.23	35.18	−2.05	0.9984	0.9974	−0.0010
	32	34.42	33.21	−1.21	0.9969	0.9959	−0.0010



(**a**)

(**b**)

Figure 11. PSNR and SSIM before and after embedding. (**a**) PSNR before and after embedding, (**b**) SSIM before and after embedding.

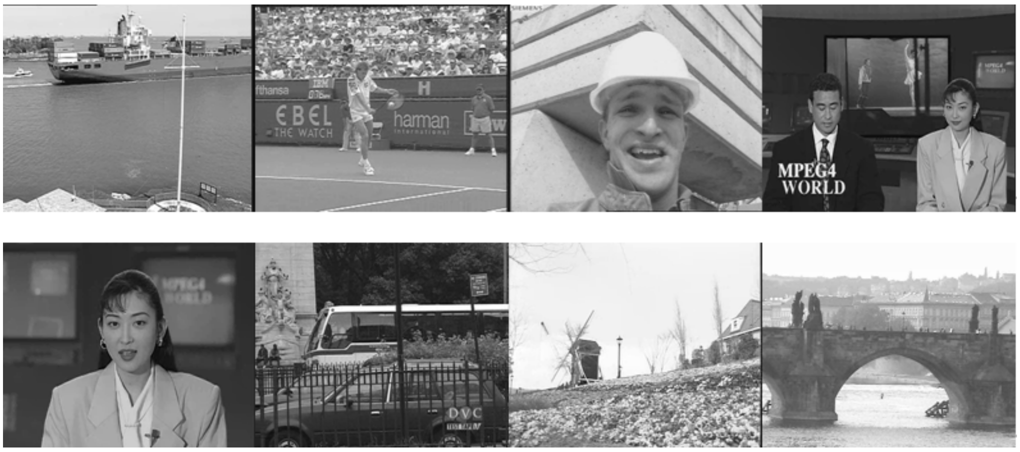


Figure 12. 20th frame of each labeled decrypted video.

4.7. Embedding Capacity

Embedding capacity is a basic evaluation metric of data hiding. The full embedded capacity of each video sequence is shown in Table 5. From different video sequences, video sequences with complex textures, such as stefan, bus and flower, have relatively large embedding capacity. The embedding capacity of video sequences with simple texture, such as akiyo and container, is relatively small. This is because the specific QDCT blocks are selected for embedding, of which DC coefficients are not 0 and AC coefficients are not all 0. The more complex the texture of the video sequence is, the more QDCT blocks can be used for embedding, and the larger the full embedding capacity. In video compression coding, the larger the QP value is, the more QDCT coefficients will be quantized to 0, resulting in fewer embeddable QDCT blocks and smaller full embedding capacity. According to the change of SSIM following QP in Table 4, the embedding capacity of this algorithm is large when the QP value is small, but the SSIM of embedded video will not decrease significantly. It shows that the embedding algorithm in this paper is more suitable for video encoded with low QP.

Table 5. The max embedding capacity of each video sequence.

Video Sequence	QP	Embedding Capacity	Video Sequence	QP	Embedding Capacity
container	24	44,015	akiyo	24	14,910
	28	17,615		28	6465
	32	7705		32	2700
stefan	24	194,965	bus	24	195,160
	28	136,535		28	129,165
	32	87,580		32	79,545
foreman	24	102,230	flower	24	133,255
	28	51,605		28	86,250
	32	23,800		32	49,105
news	24	32,570	Bridge-close	24	118,815
	28	17,340		28	59,450
	32	9285		32	27,610

4.8. Robustness Analysis

Robustness refers to the ability not to lose hidden information due to some changes in the carrier. Without any attack on the video containing information, the information embedded by the proposed embedding algorithm can be accurately extracted from the video. However, if the video is attacked by various attacks, such as recompression, the embedded information will not be extracted effectively. For example, after recompression, the coefficients of some QDCT blocks containing information may be fully quantized to 0, so these QDCT blocks containing information will be skipped during information extraction, resulting in information extraction errors. In addition, even if the QDCT block is not quantized into an all zero coefficient, the coefficients a and b will probably change, making it impossible to accurately extract information. In addition, even if the coefficients of the QDCT block are not all quantized to zero, the coefficients a and b will probably change, making it impossible to accurately extract information. Therefore, the embedding algorithm in this paper is not robust. This means that the scheme in this paper is not suitable for scenarios with high robustness requirements, and can be applied to content authentication, data indexing and other fields.

4.9. Reversibility Analysis

The reversibility of the embedding algorithm refers to the property that the carrier can be recovered without loss after information extraction. In this paper, the classical reversible algorithm, difference expansion, is used. The algorithm uses the difference between coefficients to extract and restore the carrier. Figure 13 shows the original PSNR and the PSNR after information embedding and extraction of news, container and bus video sequences. It can be seen from the figure that the two PSNR polylines corresponding to each video sequence coincide. It shows that the carrier can restore to the original value after extracting information and restore the original video quality, proving that the embedding algorithm in this paper is reversible.

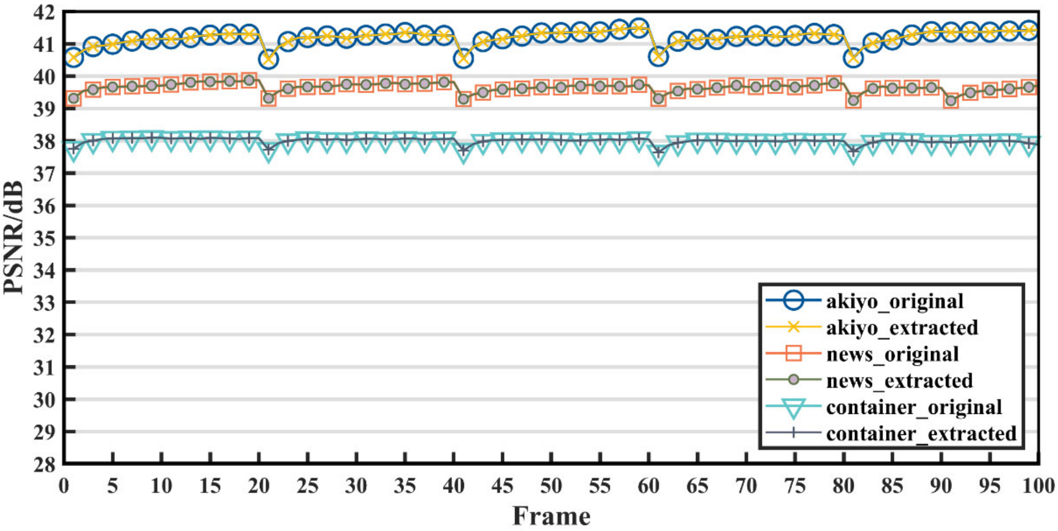


Figure 13. The original PSNR and the PSNR after information embedding and extraction.

4.10. Bit Rate Variation

In order to further evaluate the performance of this scheme, the bit rate variations caused by the encryption and information embedding are analyzed. Bit rate variation can be calculated as follows:

$$BR\_\text{var} = \frac{BR\_\text{e} - BR\_\text{ori}}{BR\_\text{ori}} \times 100\% \quad (27)$$

where *BR*_e denotes the bit rate after encryption or embedding; *BR*_ori denotes the original bit rate. The larger the *BR*_var is, the greater the bit rate variation caused by the operation, which affects the compression performance.

The encryption method in this paper includes three aspects—QDCT block permutation, IPM encryption, and MVD sign encryption—and it is implemented at the code stream. The QDCT block permutation only changes the location, not the coefficients in the QDCT block, so it will not cause bit changes. IPM encryption is to perform XOR encryption on three codewords in the code stream without changing the code length. The encryption of the MVD sign will only cause the reversal of its sign codeword, and will not increase the bit. Therefore, the encryption method proposed in this paper will not lead to changes in bit rate. The embedding algorithm used in this paper will change amplitude of some QDCT coefficients, so it will inevitably lead to the variation of bit rate. The more information is embedded, the greater the bit rate change.

Table 6 shows the bit rate variation caused by encryption and embedding of each video sequence under different QP conditions. It can be seen that the experimental results are consistent with the above analysis. In different video sequences encoded by different QP, the bit rate after encryption is always consistent with the original bit rate. As for the bit rate change caused by embedding, we can analyze it from two perspectives: video sequence and QP value. From the perspective of video sequence, the bit rate of video sequence with complex texture changes more. From the perspective of QP, the bit rate of video sequence with low QP changes more. Through analysis, the reason is that video sequences with complex texture or low QP contain more embeddable QDCT blocks, and more coefficients are modified, resulting in greater changes in bit rate.

Table 6. Bit rate variation caused by encryption and embedding.

Video Sequence	**QP**	*BR*_var/%	
		Encryption	**Embedding**
container	24	0	2.9142
	28	0	1.8050
	32	0	1.2280
stefan	24	0	4.2299
	28	0	4.2577
	32	0	4.1354
foreman	24	0	4.9393
	28	0	4.3054
	32	0	3.2288
news	24	0	3.0628
	28	0	2.2776
	32	0	1.7547
akiyo	24	0	2.2069
	28	0	1.3957
	32	0	0.8569
bus	24	0	4.0076
	28	0	4.0006
	32	0	4.0473
flower	24	0	3.3235
	28	0	3.1360
	32	0	2.6917
bridge-close	24	0	3.4948
	28	0	3.1446
	32	0	2.6155

5. Conclusions

In this paper, a novel separable scheme for encryption and reversible data hiding is proposed, which has improved security in encryption method. In the encryption stage, the QDCT blocks are permutated by logistic chaotic scrambling and the partial codewords of the IPM and the MVD are encrypted by XOR encryption, which will not lead to bit rate variation. In the reversible data hiding stage, difference expansion is applied for the first time and AC_{10} and AC_{12} in 4 × 4 QDCT block of the P frame are selected as coefficient pairs for embedding. The experiments show that compared with the encryption method used in other schemes, the proposed encryption method can resist sketch attack and has better security. The reversible data hiding algorithm can provide good visual quality of the labeled decrypted video, of which the average PSNR decreases by 3.57 dB at most and the average SSIM decreases by 0.0027 at most. The encryption and the data hiding in the proposed scheme are separable and the proposed scheme can be applied to more application scenarios, as described in Section 3.4. Based on the above conclusions, users concerned about data privacy can safely store video data in the cloud and can directly decrypt the video in the cloud to quickly browse the content. In addition, due to the high security of the proposed scheme, it may be applied to telemedicine or military fields. Of course, the proposed scheme can be further improved. Although the permutation encryption can resist contour attacks, the increased computational complexity makes the scheme less immediate. In future work, we focus on how to combine faster secure encryption techniques and reversible data hiding techniques.

Author Contributions: Conceptualization, P.C.; methodology, Y.L.; project administration, X.Y.; validation, K.N.; writing—original draft, P.C.; writing—review and editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the National Natural Science Foundation of China under grant No.61872384.

Data Availability Statement: Not applicable.

Conflicts of Interest: The funders had no role in the design of the study, the collection, analyses, or interpretation of data, the writing of the manuscript, or the decision to publish the results.

Abbreviations

RDH-EV	reversible data hiding in encrypted video
IPM	intra-prediction mode
MVD	motion vector difference
QDCT	quantized discrete cosine transform
HS	histogram shifting
DE	difference expansion
I-Q-M encryption	IPM encryption, QDCT sign encryption and MVD sing encryption
MBS	macroblock bitstream size
GOP	group of pictures
QP	quantization parameter
PSNR	peak signal-to-noise ratio
SSIM	structural similarity

References

1. Al-Abbasi, A.O.; Aggarwal, V. VidCloud: Joint Stall and Quality Optimization for Video Streaming over Cloud. *ACM Trans. Model. Perform. Eval. Comput. Syst.* **2021**, *5*, 1–32. https://doi.org/10.1145/3442187.
2. Singh, P.; Atrey, P.K. Recovering Tampered Regions in Encrypted Video Using POB Number System. *Signal Proc. Image Commun.* **2019**, *74*, 96–109. https://doi.org/10.1016/j.image.2019.01.009.
3. Lian, S.; Liu, Z.; Ren, Z.; Wang, H. Commutative Encryption and Watermarking in Video Compression. *IEEE Trans. Circuits Syst. Video Technol.* **2007**, *17*, 774–778. https://doi.org/10.1109/TCSVT.2007.896635.

4. Xu, D.; Wang, R.; Shi, Y.-Q. Reversible Data Hiding in Encrypted H.264/AVC Video Streams. In *Digital-Forensics and Watermarking, Proceedings of the 12th International Workshop, IWDW 2013, Auckland, New Zealand, 1–4 October 2013*; Shi, Y.-Q., Kim, H.-J., Pérez-González, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8389, pp. 141–152.
5. Puteaux, P.; Puech, W. A Recursive Reversible Data Hiding in Encrypted Images Method with a Very High Payload. *IEEE Trans. Multim.* **2021**, *23*, 636–650. https://doi.org/10.1109/TMM.2020.2985537.
6. Wu, F.; Zhou, X.; Chen, Z.; Yang, B. A Reversible Data Hiding Scheme for Encrypted Images with Pixel Difference Encoding. *Knowl. Based Syst.* **2021**, *234*, 107583. https://doi.org/10.1016/j.knosys.2021.107583.
7. Yu, M.; Yao, H.; Qin, C. Reversible Data Hiding in Encrypted Images without Additional Information Transmission. *Signal Proc. Image Commun.* **2022**, *105*, 116696. https://doi.org/10.1016/j.image.2022.116696.
8. Xu, D.; Wang, R. Efficient Reversible Data Hiding in Encrypted H.264/AVC Videos. *J. Electr. Imaging* **2014**, *23*, 053022. https://doi.org/10.1117/1.JEI.23.5.053022.
9. Yao, Y.; Zhang, W.; Yu, N. Inter-Frame Distortion Drift Analysis for Reversible Data Hiding in Encrypted H.264/AVC Video Bitstreams. *Signal Proc.* **2016**, *128*, 531–545. https://doi.org/10.1016/j.sigpro.2016.05.004.
10. Xu, D.; Zhu, Y.; Wang, R.; Fu, J.; Chen, K. Two-Dimensional Histogram Modification for Reversible Data Hiding in Partially Encrypted H.264/AVC Videos. In *Digital Forensics and Watermarking, Proceedings of the 15th International Workshop, IWDW 2016, Beijing, China, 17–19 September 2016*; Shi, Y.-Q., Kim, H.-J., Pérez-González, F., Liu, F., Eds.; Springer: Cham, Switzerland, 2016; Volume 10082, pp. 393–406.
11. Long, M.; Peng, F.; Li, H. Separable Reversible Data Hiding and Encryption for HEVC Video. *J. Real Time Image Proc.* **2018**, *14*, 171–182. https://doi.org/10.1007/s11554-017-0727-y.
12. Chen, P.; Zhang, Z.; Lei, Y.; Niu, K.; Yang, X. A Multi-Domain Embedding Framework for Robust Reversible Data Hiding Scheme in Encrypted Videos. *Electronics* **2022**, *11*, 2552. https://doi.org/10.3390/electronics11162552.
13. Xu, D.W. Commutative Encryption and Data Hiding in HEVC Video Compression. *IEEE Access* **2019**, *7*, 66028–66041. https://doi.org/10.1109/ACCESS.2019.2916484.
14. Guan, B.; Xu, D.W.; Li, Q. An Efficient Commutative Encryption and Data Hiding Scheme for HEVC Video. *IEEE Access* **2020**, *8*, 60232–60245. https://doi.org/10.1109/ACCESS.2020.2983330.
15. Xu, D.W.; Guan, B. An Improved Commutative Encryption and Data Hiding Scheme for HEVC Video. *Multim. Tools Appl.* **2022**, *81*, 18105–18127. https://doi.org/10.1007/s11042-022-12676-8.
16. Yang, C.; Luo, X.; Lu, J.; Liu, F. Extracting Hidden Messages of MLSB Steganography Based on Optimal Stego Subset. *Sci. China Inf. Sci.* **2018**, *61*, 119103. https://doi.org/10.1007/s11432-017-9328-2.
17. Minemura, K.; Wong, K.; Phan, R.C.-W.; Tanaka, K. A Novel Sketch Attack for H.264/AVC Format-Compliant Encrypted Video. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 2309–2321. https://doi.org/10.1109/TCSVT.2016.2589742.
18. Xu, D.; Liu, Y. Reversible Data Hiding in H.264/AVC Videos Based on Hybrid-Dimensional Histogram Modification. *Multim. Tools Appl.* **2022**, *81*, 29305–29319. https://doi.org/10.1007/s11042-022-12740-3.
19. Xu, Y.; He, J. Two-Dimensional Histogram Shifting-Based Reversible Data Hiding for H.264/AVC Video. *Appl. Sci.* **2020**, *10*, 3375. https://doi.org/10.3390/app10103375.
20. Kang, J.; Kim, H.; Kang, S. Genuine Reversible Data Hiding Technique for H.264 Bitstream Using Multi-Dimensional Histogram Shifting Technology on QDCT Coefficients. *Appl. Sci.* **2020**, *10*, 6410. https://doi.org/10.3390/app10186410.
21. Tian, J. Reversible Data Embedding Using a Difference Expansion. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, *13*, 890–896. https://doi.org/10.1109/TCSVT.2003.815962.
22. Kim, H.; Kang, S. Genuine Reversible Data Hiding Technology Using Compensation for H.264 Bitstreams. *Multim. Tools Appl.* **2018**, *77*, 8043–8060. https://doi.org/10.1007/s11042-017-4698-6.
23. Tang, X.; Wang, H.-X.; Chen, Y. Reversible Data Hiding Based on a Modified Difference Expansion for H.264/AVC Video Streams. *Multim. Tools Appl.* **2020**, *79*, 28661–28674. https://doi.org/10.1007/s11042-020-09315-5.
24. Ma, G.; Wang, J. Efficient Reversible Data Hiding in Encrypted Images Based on Multi-Stage Integer Wavelet Transform. *Signal Proc. Image Commun.* **2019**, *75*, 55–63. https://doi.org/10.1016/j.image.2019.03.013.
25. Weng, S.; Chen, Y.; Hong, W.; Pan, J.-S.; Chang, C.-C.; Liu, Y. An Improved Integer Transform Combining with an Irregular Block Partition. *Symmetry* **2019**, *11*, 49. https://doi.org/10.3390/sym11010049.
26. He, J.; Xu, Y.; Luo, W.; Tang, S.; Huang, J. A Novel Selective Encryption Scheme for H.264/AVC Video with Improved Visual Security. *Signal Proc. Image Commun.* **2020**, *89*, 115994. https://doi.org/10.1016/j.image.2020.115994.
27. Farajallah, M.; Gautier, G.; Hamidouche, W.; Déforges, O.; Assad, S.E. Selective Encryption of the Versatile Video Coding Standard. *IEEE Access* **2022**, *10*, 21821–21835. https://doi.org/10.1109/ACCESS.2022.3149599.
28. Li, X.; Yu, H.; Zhang, H.; Jin, X.; Sun, H.; Liu, J. Video Encryption Based on Hyperchaotic System. *Multim. Tools Appl.* **2020**, *79*, 23995–24011. https://doi.org/10.1007/s11042-020-09200-1.
29. Ayubi, P.; Barani, M.J.; Valandar, M.Y.; Irani, B.Y.; Sadigh, R.S.M. A New Chaotic Complex Map for Robust Video Watermarking. *Artif. Intell. Rev.* **2021**, *54*, 1237–1280. https://doi.org/10.1007/s10462-020-09877-8.
30. Farri, E.; Ayubi, P. A Robust Digital Video Watermarking Based on CT-SVD Domain and Chaotic DNA Sequences for Copyright Protection. *J. Ambient Intell. Humaniz. Comput.* **2022**, 1–25. https://doi.org/10.1007/s12652-022-03771-7.
31. Richardson, I.E.G. *H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia*; Wiley: Hoboken, NJ, USA, 2003; ISBN 978-0-470-84837-1.

32. Xue, S.; Qi, W.-F. Research on the Best Linear Approximation of Addition Modulo 2n. *J. Electron. Inf. Technol.* **2012**, *34*, 2156–2160. https://doi.org/10.3724/SP.J.1146.2012.00096.
33. YUV Video Sequences. Available online: http://trace.eas.asu.edu/yuv/index.html (accessed on 30 September 2022).
34. H.264 Baseline Codec. Available online: https://ww2.mathworks.cn/matlabcentral/fileexchange/39927-h-264-baseline-codec (accessed on 26 June 2022).
35. Horé, A.; Ziou, D. Image Quality Metrics: PSNR vs. SSIM. In Proceedings of the 20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.