A selective encryption scheme for protecting H.264/AVC video in multimedia social network

Fei Peng[1] · **Xiao-qing Gong**[1] · **Min Long**[2] · **Xing-ming Sun**[3]

Received: 27 July 2015 / Revised: 18 May 2016 / Accepted: 24 June 2016 /
Published online: 11 July 2016

Abstract Aiming to protect H.264/AVC video in multimedia social network, a selective encryption scheme is put forward. In the scheme, after analyzing the impact of quantization parameter (QP) on the encryption of the sign of T1 s and the impact of encrypting inter-macroblock non-zero coefficients, the sign of intra-macroblock non-zero DCT coefficients, the sign of trailing ones (T1 s), the intra prediction modes (IPMs) and the sign of motion vector difference (MVD) are encrypted to protect the texture and motion information of H.264/AVC. Experimental results and analysis show that the computation cost is low and the bitrate increment is negligible, and it can achieve good performance in resisting brute-force attacks, histogram-based attacks and replacement attacks. Based on the encryption scheme, a framework of its implementation in multimedia social network is put forward. It has great potential to be implemented for the video data protection in multimedia social network.

Keywords Multimedia security · Selective encryption · H.264/AVC, multimedia social network

1 Introduction

Nowadays, network technologies have been developed very fast. People can easily access multimedia social networks such as YouTube, Facebook, Youku and etc. via personal computer, mobile and tablet computer. Large scale of multimedia content can be watched at anywhere and

✉ Fei Peng
eepengf@gmail.com

[1] School of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan Province 410082, China

[2] College of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, Hunan Province 410014, China

[3] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, Jiangsu Province 210044, China

anytime. Meanwhile, the open multimedia social network may introduce problems on digital rights managements (DRM) [35, 36]. For example, when a multimedia content is accessed by a client user, it can be duplicated and abused by malicious persons when it is transmitted in the open network. Thus the rights of the content providers are harmed. As video is an important multimedia data in multimedia social network, its content protection is of great importance.

To countermeasure this situation, much effort has been put to improve the security of multimedia data [6, 23]. Encryption has been proved to be a feasible means to protect the content of video data. Even the encrypted video data is intercepted by eavesdroppers, the original information of video data is still unknown if they do not have the correct decryption key. In this way, the security of video data can be guaranteed. Video data has features of large volume, high real-time requirements and special coding structure. If traditional encryption methods are directly used for video data, it can achieve high security, but it is not format-compliant and the computational complexity is high [8]. Therefore, selective encryption provide another effective means for protecting video data in multimedia social network.

As a popular video coding standards in the existed multimedia social networks, H.264/AVC can obtain low bitrate and greatly reduce the cost of video transmission [4]. Meanwhile, it has good network adaptability and error robustness, and can efficiently prevent packet loss, bit error and blocking in the network transmission [32]. Thus, selective video encryption scheme for H.264 has broad application prospect [17].

The main contributions of the paper include:

(1) The impact of QP on the encryption of the sign of T1 s and the impact of encrypting inter-macroblock non-zero coefficients are analyzed in details, and it provides a guidance for the selection of syntax elements in H.264/AVC for encryption.
(2) Compared with the existed methods, the proposed selective encryption scheme can significantly reduce the data amount for encryption while obtain fairly good performance. It can strike a balance between security and efficiency.
(3) A framework of the implementation of the proposed selective encryption scheme in multimedia social network is put forward.

The remainder of the paper is organized as follows. The related works are introduced in Section 2. The proposed selective encryption scheme is presented in Section 3 and the experimental results and analysis are provided in Section 4. Finally, some conclusions are drawn in Section 5.

2 Related works

In the past decades, some encryption schemes for H.264/AVC have been put forward. These algorithms are designed for different purposes. Some for high security, some for real-time applications, some for compression ratio invariance and some for format-compliance. According to the data volume to be encrypted, the existed video encryption algorithms can be classified into complete video encryption and selective video encryption.

2.1 Complete video encryption

Complete video encryption encrypt video as a sequence of binary data. AES (Advanced encryption standard) and DES (Data encryption standard) are generally used to encrypt the entire video data [7]. Complete encryption can achieve high security and the length of bitstream can be kept. However, as the data volume of video is huge, the computational complexity is high. Meanwhile, it destroys the coding structure and the format of the encrypted video is not compatible. Therefore, the applications of complete video encryption is only confined to the circumstance with high security but no real-time requirements, such as video storage.

2.2 Selective video encryption

After that, the significant partial data for reconstruction of video is proposed to be selected for encryption, which is named as selective encryption [14]. Selective encryption generally has fast speed. At the same time, it can make use of the process of video coding and video data format, which is good for the compatibility and maneuverability. For H.264/AVC, IPMs [1, 5, 12], motion vector [11, 15, 16, 19, 20], DCT coefficients [9, 18, 22, 24, 25, 30, 31] are often selected for encryption.

2.2.1 Selective encryption based on IPMs

In the video decoding, IPMs are important for the reconstruction of intra-macroblock. J. Ahn et al. first proposed to scramble H.264/AVC using IPMs [1]. The properties of the intra-block are utilized to scramble the prediction modes of intra4 × 4 block and intra-16 × 16 block, respectively. It can keep the bitstream length and format-compliant. Nevertheless, it cannot resist Friedman attacks if the limited length of pseudo random sequence is adopted [5]. Aiming to expand the key space, two-dimensional chaotic pseudorandom sequence is implemented for the encryption of H.264 [12]. Intra, Inter prediction mode scrambling, transform results and motion vector encryption are combined in the encryption. Nevertheless, since only half of IPMs are encrypted, the security is still limited. Based on the works in [1, 12], S. Xing et al. proposed an improved encryption scheme based on IPMs [33]. During the scrambling of the most probable prediction modes, it is reset to 0 and inserts 3 bits chaotic key as the value of rem_intra4 × 4_pred_mode when the prev_intra4 × 4_pred_mode_flag equals to 1. It can achieve high security, low complexity and good real-time property, but it imposes impact on code length.

2.2.2 Selective encryption based on motion vector

Motion vector is a key data for inter-macroblock reconstruction in H.264/AVC decoder. A digital video scrambling scheme using motion vectors and slice relocation was proposed by S. Kwon et al. [11]. By arbitrarily relocating the differential motion vectors and macroblock starting positions in a slice, the video content is scrambled. However, it requires parsing bitstream twice. Meanwhile, it is not suitable for the situation when arithmetic coding is used for motion vector relocation. Different operators (addition, subtraction, multiplication, and division) for different frames and different operands for different macroblocks in the same frame were proposed to encrypt motion vector [19]. All of them are stored in an index table along with the frame number and macroblock number of motion vector. Nevertheless, the maintenance of the index table is complicated. S. Lian et al. proposed to encrypt advanced video coding by selectively encrypting intra-prediction mode, residue data and motion vector [15, 16]. The infra-prediction mode is encrypted based on Exp-Golomb entropy coding, the intra-macroblocks' direct currents are encrypted based on context-based adaptive variable length coding, and infra-macroblocks' ACs and the inter-macroblocks' MVDs are sign-encrypted with a stream cipher followed with variable-length coding. It is secure in perception, keeps format compliance, and obtains high time efficiency. To protect the ROI (Region of interest) privacy information in video, F. Peng et al. proposed to use chaotic systems to encrypt the amplitudes of MVDs [20]. Although it can achieve good scrambling for the ROI, the computational complexity is high and the compression ratio is marginally affected.

2.2.3 Selective encryption based on DCT coefficients

After transform and quantization, DCT coefficients are critical for video restoration. Thus, the encryption of DCT coefficients is significant for video. A selective video encryption scheme based on H.264 was proposed by Y. Wang et al. [30]. It combines the AES OFB (Output Feedback) mode with the sign encryption algorithm, and encrypts direct currents and parts of alternating currents, respectively. Since a low-resolution video frame can be restored by the absolute values of motion vector [18], the security is limited. A quality-controllable encryption method for H.264 coded video streams was proposed by G. Hong et al. [9]. According to the percentage of intra4 × 4 blocks, intra8 × 8 blocks and intra16 × 16 blocks in different video sequences, four levels of encryption policies including no encryption, encryption of the residuals of intra4 × 4 block, encryption of the residuals of intra8 × 8 block, and encryption of the residuals of intra16 × 16 block are implemented. The security is varied with different encryption levels. A video encryption scheme exploiting the distribution of the DCT coefficients was proposed by C. Raju et al. [22]. Direct current and alternating current coefficients are encrypted based on the distribution characteristics. Furthermore, the encrypted DCT coefficients within a frame are partitioned into groups, and they are scrambled within the group. It can obtain high security and has little impact on compression ratio. However, the computational complexity is high. Two fast encryption schemes for H.264/AVC were proposed by Z. Shahid et al. in [24, 25]. Sign bit and the suffix of non-zero coefficients are encrypted. They are format-compliant and fulfills real-time constraints. Nevertheless, only residual data is encrypted, the texture information and motion information are untouched. Thus, the security is still questionable. A tunable encryption scheme for H.264/AVC was proposed by Y. Wang et al. [31]. The sign of non-zero coefficients in the residual block is encrypted. Compared with the encryption methods in [24, 25], the scrambling effect is mainly achieved by encrypting the sign of non-zero coefficients. However, the encryption of the sign of T1 s cannot provide satisfactory scrambling effect when the quantization parameter (QP) is small. Thus, the encryption of the sign of inter-macroblock non-zero coefficients cannot provide good scrambling effect.

From the above analysis, although much advancement has been made on selective encryption for H.264/AVC, their still exist some shortcomings such as limited scrambling effect, high computation complexity, side influence on compression ratio, and the contradiction between security and efficiency.

To meet the needs for the protection of H.264/AVC in multimedia social network, a novel fast encryption scheme for H.264 is proposed based on quantization and CAVLC (Context-based Adaptive Binary Arithmetic Coding) entropy coding.

3 The proposed selective encryption scheme

In this Section, the impact of quantization parameter on the encryption of the sign of T1 s and the impact of encrypting inter-macroblock non-zero coefficients are first analyzed, respectively, and then the proposed selective scheme is presented. All results are obtained by encoding the first 30 frames of Foreman sequence with CIF (352 × 288) resolution and 4:2:0 sampling format under the baseline profile, IPP…P.

3.1 Impact of quantization parameter on the encryption of the sign of T1s

During CAVLC entropy encoding, the non-zero coefficients of 4 × 4 block are grouped into T1s and remaining non-zero coefficients (RLevel), and they are encoded separately. T1s refer to the coefficients whose values are +1/−1, and the maximum number is 3. If the number of the +1/−1 is exceeded 3, only the last 3 bits are considered as T1s. For each T1s, 1 bit is used to represent the sign. If T1s equals +1, the codeword is 0; otherwise, the codeword is 1. An example of the encrypted results of T1s are shown in Fig. 1.

As seen from Fig. 1, the scrambling effect of the encrypted results are varied with different QPs. By adjusting the quantization parameters, the average PSNR and SSIM (Structural Similarity) [28, 29] of the encryption results are shown in Fig. 2.

As seen from Fig. 2, It can be found that the change of the PNSRs and SSIMs of the encrypted video frame is not significant when the QP is small. Meanwhile, the PNSRs and SSIMs are significantly reduced when the QP is increased. Energy Packing Efficiency (EPE) [3] was proposed to calculate the energy distribution of T1s and RLevel, respectively. It is formulated as:

$$EPE = \sum_{i=M}^{N} E\{S_i^2\} / \sum_{j=0}^{15} E\left\{S_j^2\right\}, \tag{1}$$

where S_i represents the i^{th} $(0 \le i \le 15)$ DCT coefficient using zigzag scanning. With different M and N, the energy percentage of T1 s and RLevel in the whole block can be obtained, respectively.

As seen from Figs. 1 and 3, when QP is 12, the energy proportion of T1s is only 0.06 while the most energy of DCT coefficients is in RLevel. Under this circumstance, the encryption of the sign of T1 s cannot achieve satisfactory perceptual scrambling effect. When QP is 36, the energy proportion of T1s is about 0.57. The encryption of the sign of T1s cannot achieve much better perceptual scrambling effect than the fore-mentioned situation.

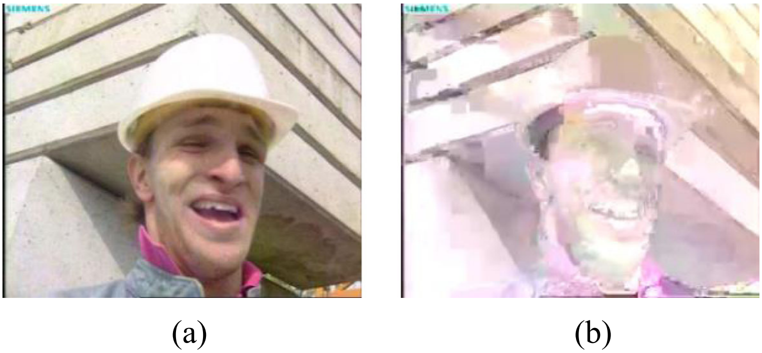


Fig. 1 Encryption results of the fifteen frame of Foreman. **a** QP = 12; **b** QP = 36

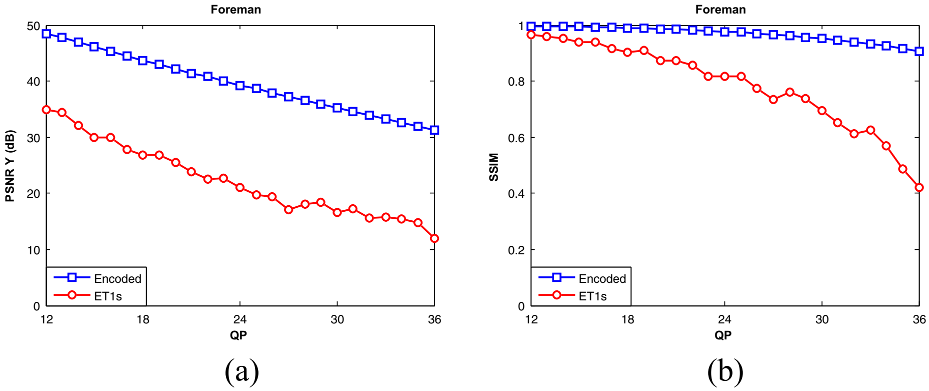


Fig. 2 PSNR and SSIM of the encrypted video frame. **a** PSNR; **b** SSIM

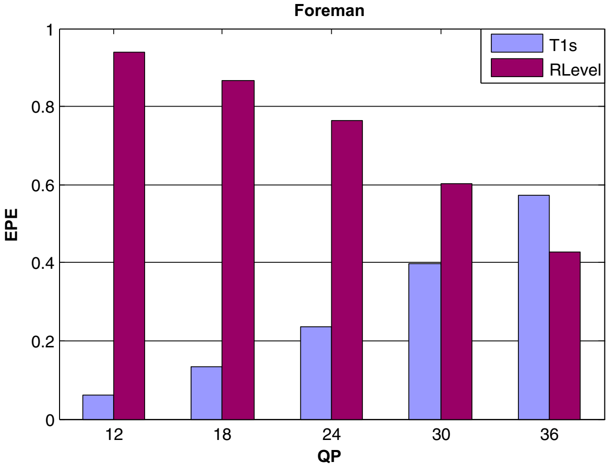
From the above analysis, QP has significant impact on the scrambling effect from the encryption of the sign of T1s. In order to achieve satisfactory encryption performance, a predefined threshold for QP should be selected by the PSNR and SSIM of the encrypted video.

3.2 Impact of encrypting the sign of inter-macroblocks NCs

For H.264/AVC, a frame can be partitioned into macroblocks and several macroblocks can be arranged as slice. For Intra-slice (I-slice), it contains only intra-macroblock (I-macroblock). While for inter-slice (P-slice or B-slice), it contains both intra-macroblock and inter-macroblock (P-macroblock for P-slice or B-macroblock for B-slice). During the H.264/AVC encoding process, P-macroblock refers to the previous reference picture, while B-macroblock is predicted from both the previous and next reference pictures. Thus, compared with P-macroblock, B-macroblock can achieve higher prediction precision and fewer NCs.

As the prediction block using intra-prediction is accomplished based on the current block and the decoded blocks of the same decoded slice, the energy of the I-macroblock is mainly concentrated on the residual data. In the decoding process, the encryption of the current block will have impact on the sample values of the right block and the below block in the video

Fig. 3 EPE of T1 s and RLevel in Foreman with different QP values



frame reconstruction, which results in influence on the decoding of the rest macroblocks. Besides, the subsequent P-macroblocks are also decoded based on I-macroblocks. Thus, the encryption of the residual data of I-macroblocks can achieve good perceptual scrambling effect.

For P-macroblock, each partition or sub-partition is predicted from a same size area of the reference picture according to inter-prediction. As the energy of P-macroblock is mainly concentrated on the corresponding reference picture, the residual data occupies only a small part of energy of the current block. Meanwhile, the decoding of the current block has no influence on the neighbored inter-prediction blocks. So the encryption of P-macroblock NCs only changes some detail parts of the video frame, it cannot achieve effective scrambling effect.

The encryption of the sign of different macroblock NCs of Foreman sequence are shown in Fig. 4.

As seen from Fig. 4, it can be found that the scrambling effect of the encryption is mainly come from I-macroblocks NCs, while the contribution from P-macroblock NCs is negligible. Meanwhile, the PSNR and SSIM of the encryption of the sign of P-macroblock NCs with different QPs are shown in Fig. 5.

As distortion can be found when the PSNR of luminance component is less than 23 [13], it can be seen from Fig. 5 that all PSNR of each frame are greater than 23 and the SSIM are also very large. It indicates that the encryption of the sign of P-macroblock NCs cannot achieve scrambling effect with different QPs.

From the above analysis, the encryption of the sign of P-macroblock NCs has negligible contribution to the scrambling of the whole video, while the encryption of the sign of I-macroblock NCs is desirable.

3.3 The proposed selective encryption scheme

According the analysis described above, the IPMs, sign of intra-macroblock NCs and he sign of MVD are selected for encryption, and the framework of the proposed encryption scheme is shown in Fig. 6.

The detailed procedures of the proposed scheme is described in the following:

Step 1. If the current macroblock is intra-macroblock and the *prev_intra4x4_pred_mode_flag* (abbreviated as *flag*) of the luminance 4 × 4 block equals 0, the *rem_intra4 × 4_pred_mode* (abbreviated as *IPM*) is encrypted as:

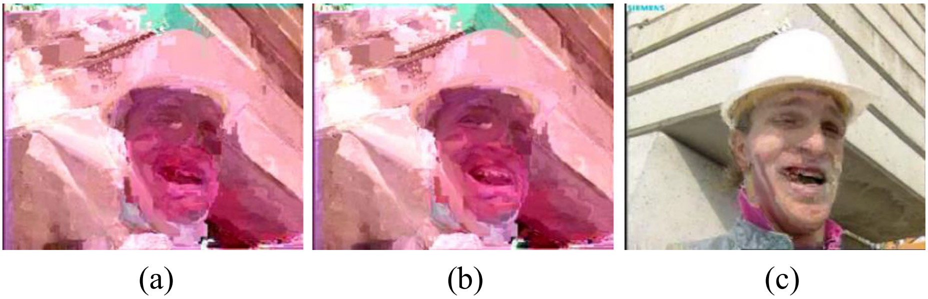


Fig. 4 Encryption of different macroblocks NCs in the fifteen frame of Foreman (QP = 18). **a** All macroblocks; **b** I-macroblock; **c** P-macroblock

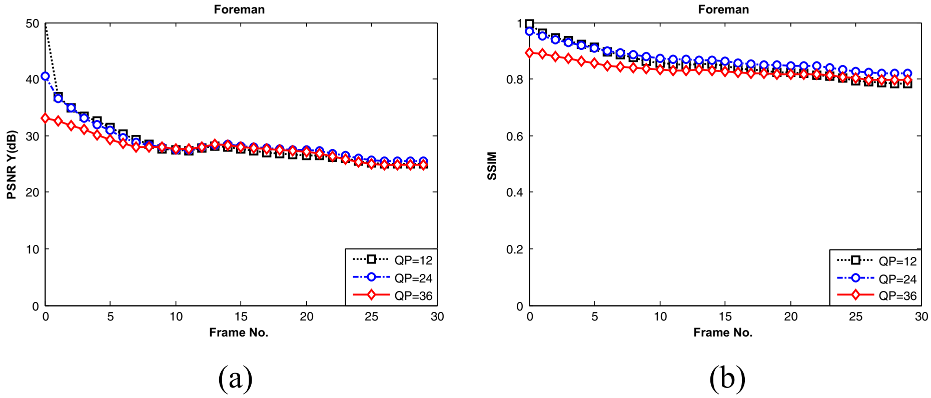


Fig. 5 PSNR and SSIM of the encrypted video frame by encrypting the sign of P-macroblocks NCs with different QPs. **a** PSNR; **b** SSIM

$$EIPM = \begin{cases} IPM \oplus key, & flag = 0 \\ IPM, & flag = 1 \end{cases}, \tag{2}$$

where *EIPM* represents the encrypted *IPM* and *key* is the key stream generated from RC4 [26].

Step 2. If the current macroblock is intra-macroblock, the sign of T1 s is encrypted when the QP is larger than or equals the predefined threshold *T*. Meanwhile the sign of RLevel is encrypted. The encryption processes are describe as:

$$ET1s = \begin{cases} T1s \oplus key, & QP \geq T \\ T1s, & otherwhis \end{cases}, \tag{3}$$

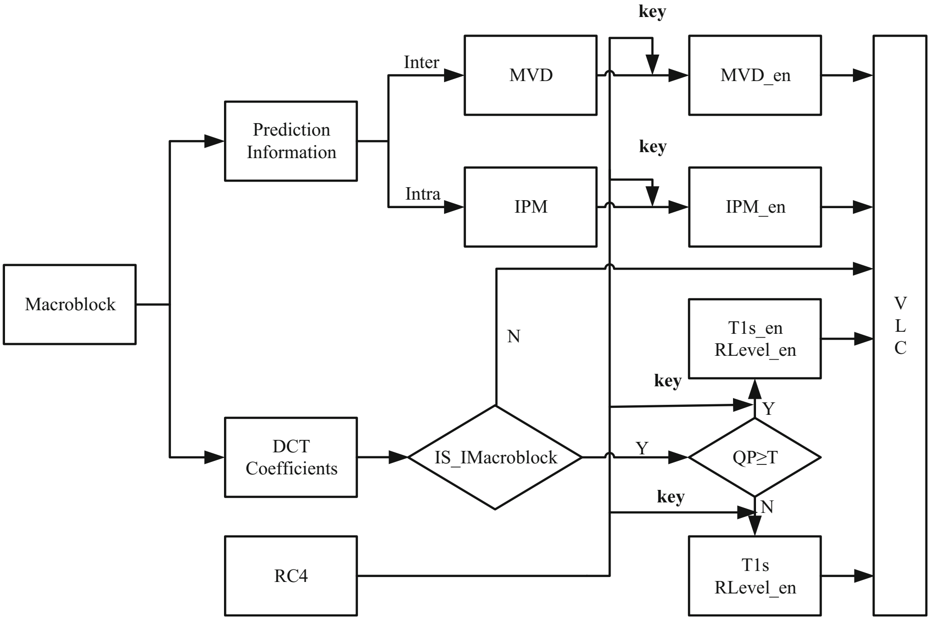


Fig. 6 The framework of the proposed encryption scheme

$$ERLevel = RLevel \oplus key, \tag{4}$$

$ET1\ s$ and $ERLevel$ represent the encrypted $T1\ s$ and $RLevel$, respectively.

Step 3. If the current macroblock is inter-macroblock, the sign of MVD is encrypted as:

$$EMVD = MVD \oplus key, \tag{5}$$

where $EMVD$ represents the encrypted MVD.

Since the encryption of the H.264/AVC video is accomplished by using XOR operations between syntax elements and key stream generated from RC4, the decryption is just the reverse of it.

4 Experimental results and analysis

All experiments are done with the H.264/AVC reference model JM10.2 [10], and they are performed on an IBM compatible PC with CPU G630 2.70GHz, 2GB RAM. Baseline profile is used in the coding, and the configuration parameters are: IPP...P sequence type encoding the first 30 frames with intra-period 30, QP is 18, the number of reference frames is 5, and the entropy coding mode is CAVLC. 9 video sequences with CIF (352 × 288) resolution and 4:2:0 sampling format [27, 34] have been used for the analysis. Each of them represents different combinations of objects (people, vehicle, buildings), color (bright/dull), contrast (high/low), and motion (fast/slow, pan/zoom/rotation). A stream cipher RC4 is used to generate binary key stream, and the threshold T is set as 26.

Experiments are done to the video sequences, and the some of the results are shown in Fig. 7. As seen from Fig. 7 (a) ~ (f), the content of the encrypted video frames are totally different from those of the original video frames, which indicate that the proposed scheme can achieve good scrambling effect.

4.1 Analysis of perceptual scrambling effect

Experiments are done 9 video benchmark sequences to evaluate the perceptual scrambling effect of the proposed method and the method in [31], respectively. The statistical analysis results of different video sequences are listed in Table 1.

As seen from Table 1, the average PNSR, SSIM of the encrypted video frames using the proposed scheme is 15.27 and 0.1197, while that for the method in [31] is 15.10 and 0.0948, respectively. Generally, human's eyes cannot distinguish the content of the video frame when the PNSR is about 15 or SSIM is about 0.10. Thus, the method in [31] and the proposed scheme both can achieve good perceptual scrambling effect. The scrambling effect of the proposed scheme is nearly the same as the method in [31], meanwhile, the amount of encrypted non-zero coefficients is significantly reduced compared with the method in [31].

4.2 Analysis of compression ratio

During the encryption, the IPMs with 4 × 4 luminance block are encoded with a fixed length and MVDs are encoded with signed Exp-Golomb, the encryption of them can keep the bitstream length and has no side effect on the compression ratio. As for NCs, only the

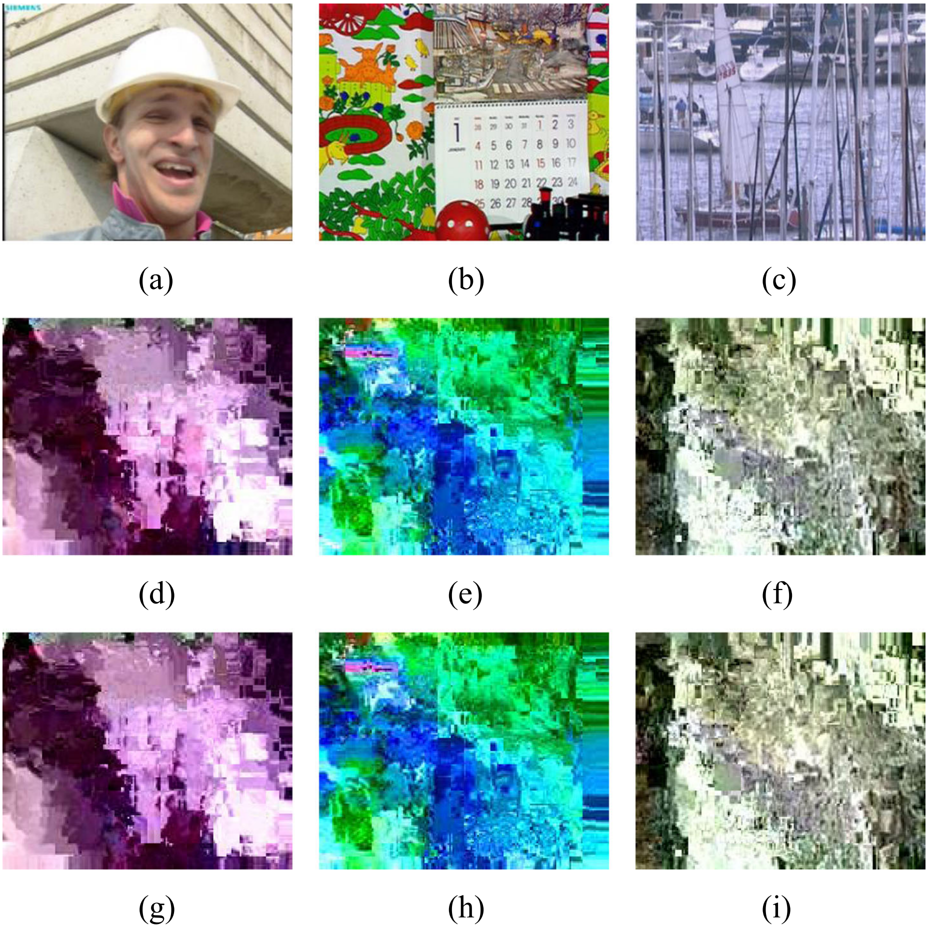


Fig. 7 Experimental results for the Fifteenth frame of Foreman, Mobile, and Harbour (QP = 18). **a~c** Original frame; **d~f** Encrypted frame using the proposed method (*key* = "ABCDEFGHIJKLMNOP"); **g~h** Encrypted frame using the method in [31]

encryption of the sign of the first RLevel has slight influence on compression ratio, while the encryption of the sign of other RLevel and T1 s have no influence on compression ratio. Thus, the total influence of the proposed scheme on the compression ratio is limited.

Experiments are done to 9 video sequences to find the difference of the bitstream length before and after encryption, the results are listed in Table 2.

As seen from Table 2, the increment of the bitstream length is varied from −0.007 % to 0.013 %, and the average increment is only 0.001 %. The results indicate that the influence of the encryption on compression ratio is negligible.

4.3 Analysis of computational cost

Generally, the computational cost of an encryption scheme mainly depends on the encrypted data volume and encryption operations. In the proposed scheme, the encrypted syntax elements include IPMs, MVDs and NCs. As for intra-macroblocks, IPMs, T1 s and RLevel

Table 1 PSNR and SSIM of different benchmark video sequences (QP = 18)

Sequence	PSNR			SSIM		
	ORIG	Method in [31]	Proposed	ORIG	Method in [31]	Proposed
Bus	43.46	15.86	16.14	0.9923	0.0380	0.0737
City	44.04	18.65	18.77	0.9911	0.0747	0.0899
Crew	44.59	16.30	16.65	0.9899	0.1740	0.2075
Football	43.18	15.35	15.59	0.9905	0.0922	0.1381
Foreman	44.59	13.49	13.64	0.9900	0.1171	0.1218
Harbour	43.45	15.87	16.21	0.9957	0.0322	0.0665
Ice	46.45	14.98	14.83	0.9914	0.1654	0.1745
Mobile	42.61	9.44	9.49	0.9966	0.0270	0.0449
Soccer	44.26	15.93	16.11	0.9841	0.1322	0.1604
Average	44.07	15.10	15.27	0.9913	0.0948	0.1197

are encrypted. Whereas for inter-macroblocks, MVDs are encrypted. Since all encryption are based on XOR operation, the encryption is fast and the realization is simple. Thus, the computational cost of the proposed scheme is very low.

Experiments are done to calculate the amount of the encrypted non-zero coefficients and the time consumption of the method in [31] and the proposed method. The results are listed in Table 3.

As seen from Table 3, compared with the method in [31], the amount of the encrypted non-zero coefficients of the proposed scheme is reduced 87.68 % and the time consumption is reduced 83.15 %. The results indicate that the computation cost of the proposed method is significantly reduced.

4.4 Security analysis

4.4.1 Analysis of key space

An ideal encryption algorithm should have sufficient key space to resist brute-force attacks. If the key space is small or the key is poorly chosen, the cryptosystem can be broken no matter

Table 2 Change of bitstream length before and after encryption (QP = 18)

Sequence	Before encryption (bit)	After encryption (bit)	Overhead (%)
Bus	4196016	4196096	0.002
City	2566456	2566456	0.000
Crew	2250296	2250143	−0.007
Football	5099520	5100198	0.013
Foreman	1877280	1877184	−0.005
Harbour	4977752	4977782	0.001
Ice	1308872	1308951	0.006
Mobile	5742808	5742957	0.003
Soccer	2754216	2754216	0.000
Average	–	–	0.001

Table 3 Comparison of computation cost for 9 video sequences in the residual block (QP = 18)

Sequence	Amount of encrypted non-zero coefficients (bit)			Time for encrypting non-zero coefficients (ms)		
	Method in [31]	Proposed	Reduction rate(%)	Method in [31]	Proposed	Reduction rate(%)
Bus	878925	64963	92.61	53.316	5.316	90.03
City	404749	40252	90.06	24.551	3.892	84.15
Crew	494387	47833	90.32	29.992	4.329	85.57
Football	1327409	502274	62.16	80.520	30.549	62.06
Foreman	253940	25973	89.77	15.404	3.042	80.25
Harbour	1022838	57938	94.34	62.041	4.910	92.09
Ice	296253	37152	87.46	17.970	3.713	79.34
Mobile	1051853	79001	92.49	63.804	6.127	90.40
Soccer	438389	44348	89.88	26.593	4.126	84.48
Average	–	–	87.68	–	–	83.15

how well and strong the encryption algorithm is designed. Generally, the size of key space equals the total number of encryption/decryption key pairs that can be used in the cryptosystem. In the proposed scheme, the stream cipher RC4 is used for the generation of key stream. As the length of the key size is varied from 1 to 256 byte, the maximum key space is $2^{256\times 8} = 2^{2048}$. In order to resist brute-force attacks [2], the length of the key size is recommended as 16 byte.

4.4.2 Analysis of key sensitivity

Key sensitivity is very important for a cryptosystem. Here, key sensitivity test is performed to the proposed scheme. For the encrypted video frame, the original key and the key with 1-bit difference are used to decrypt it, respectively. The results are shown in Fig. 8.

As seen from Fig. 8, app:addword:respectively it can be seen that 1-bit difference from the original key results in a completely wrong decrypted frame. Therefore, the proposed algorithm has good key sensitivity.

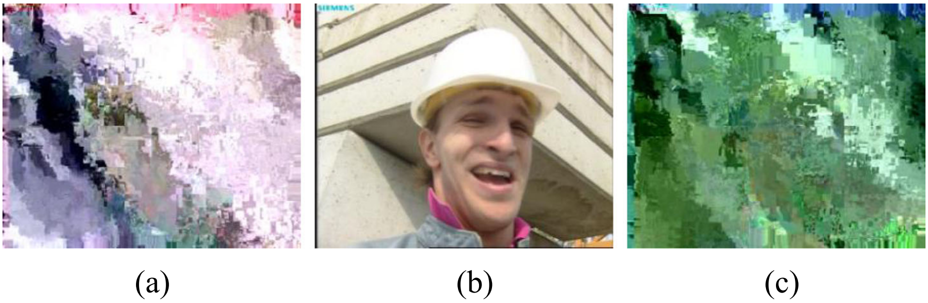


Fig. 8 Key sensitivity test for fifteenth frame in Foreman. **a** Decoded without key. PSNR =15.025 dB, SSIM = 0.209; **b** Decrypted with the original key (*key* = “ABCDEFGHIJKLMNOP”), PSNR =44.61 dB, SSIM = 0.983; **c** Decrypted with 1-bit different key from the original key (*key* = “ABCDEFGHIJKLMNOQ”). PSNR =14.34 dB, SSIM = 0.190

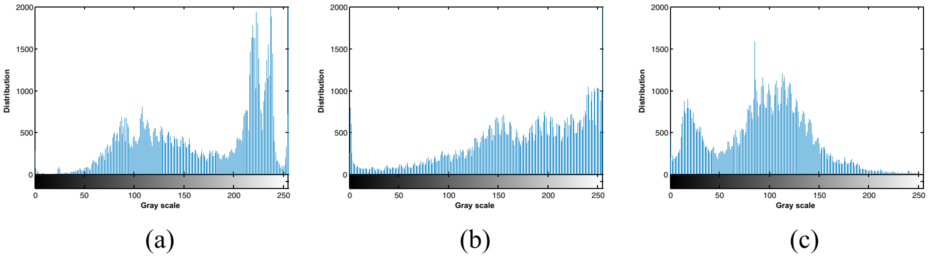


Fig. 9 Histogram analysis for fifteenth frame of Foreman. **a** Original frame; **b** Encrypted frame with the original key; **c** Encrypted frame with 1-bit different key from the original key

4.4.3 Analysis of histogram attack

Histogram reflects the distribution of the pixel values of an image. For a good encryption algorithm, the histogram of original and encrypted video frame should be different from each other in order to resist differential attacks and statistical attacks. The histograms of the original frame and encrypted frames with different keys are shown in Fig. 9.

As seen from Fig. 9, it can be found that the histograms of original and encrypted frames are completely different from each other, and the histograms cannot provide any clue to employ differential attacks and statistical attacks. The results indicate that the proposed method has good performance in resisting histogram attack.

4.4.4 Analysis of replacement attacks

Replacement attack is that the encrypted syntax elements are set to fixed values to make the encrypted video frame intelligible, and it is a kind of ciphertext-only attacks [21]. Taking Foreman sequence for example, the recovered fifteenth frames with different replacement attacks are shown in Fig. 10.

As seen from Fig. 10, they are still unintelligible with different syntax elements are set to fixed values, which prove that the proposed algorithm can resist different replacement attacks.

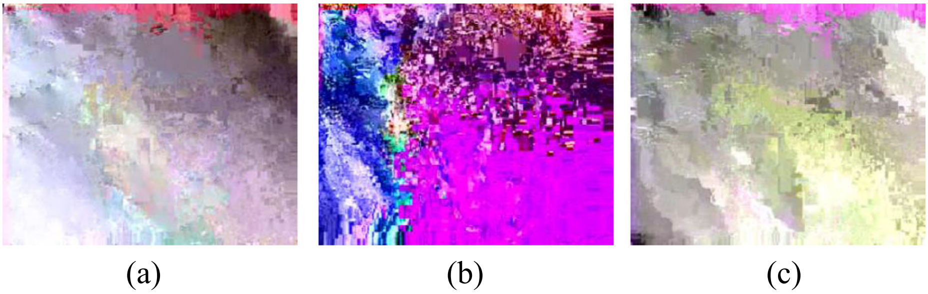


Fig. 10 The fifteenth frame of Foreman with replacement attacks. **a** Set all IPMs to be most probable prediction mode. PSNR =15.02 dB, SSIM = 0.317; **b** Set all the signs of MV to be positive values. PSNR =8.67 dB, SSIM = 0.067; **c** All the signs of NCs are set as positive values. PSNR =15.47 dB, SSIM = 0.238

Table 4 Comparison of some recent video selective encryption schemes

Video selective encryption schemes	Encrypted elements	Bitrate increase	Computational cost	Format compliance	Encryption algorithms	Security
[19]	MV using different operators and operands	Yes	Low	Yes	Index Table	Low
[16]	Intra-DC, sign of intra-AC, IPMs, suffix of MVD	No	High	Yes	AES(LKE and RDE)	High
[20]	Sign of NCs, IPMs, sign and amplitude of MVD in ROI region	Yes	High	Yes	Chaotic stream cipher	High
[22]	Encryption and scrambling DCT coefficients	Yes	High	Yes	Pseudo-Random Number	General
[25]	Sign of NCs, amplitude of RLevel	No	General	Yes	AES(CFB mode)	Low
[31]	Sign of NCs, IPMs, sign of MVD	No	High	Yes	Rabbit stream cipher	High
Proposed	Sign of RLevel and selected T1 s within intra-macroblock, IPMs, sign of MVD	No	Low	Yes	RC4	High

4.4.5 Analysis of data compatibility and operability

As there is no change to the format and the control information of H.264/AVC, the encrypted bitstream is still coincided with H.264/AVC standard. It means that the encrypted video can be normally decoded by all H.264/AVC compliant decoders without any undefined behavior. Meanwhile, the encrypted bitstream still have characteristics of data operability such as bitrate control, image block cut and paste, and addition/deletion. From above analysis, the data compatibility and operability is kept in the proposed encryption scheme.

4.4.6 Comparative evaluation

Here, a comparison is performed to different video selective encryption schemes [16, 19, 20, 22, 25, 31] proposed in the past few years. They are different from each other in several aspects, such as encrypted syntax elements, bitrate increase, computational cost, format-compliance, encryption algorithms and security. The comparison is listed in Table 4.

As seen from Table 4, it can be found that the proposed scheme is format-compliant to H.264/AVC decoder, and keeps the bitrate unchanged. Meanwhile, the computation cost is low and high security can be obtained. It can achieves good trade-off between encryption efficiency and security.

4.5 Implementation of the proposed scheme in multimedia social network

H.264/AVC is widely used in the popular multimedia social network. The framework of the implementation of the proposed selective encryption scheme in multimedia social network is illustrated in Fig. 11.

As seen from Fig. 11, for the digital video content providers in multimedia social network, they can encrypt the video before its distribution. The encryption key should be distributed to the authorized users via a secure channel. For the authorized users, they can obtain the correct video content because they can use the key to decrypt the encrypted video. While for unauthorized users, even they obtain the encrypted video content, they cannot normally watch the video content. Thus, the digital right management of H.264/AVC video can be guaranteed by using the proposed scheme.

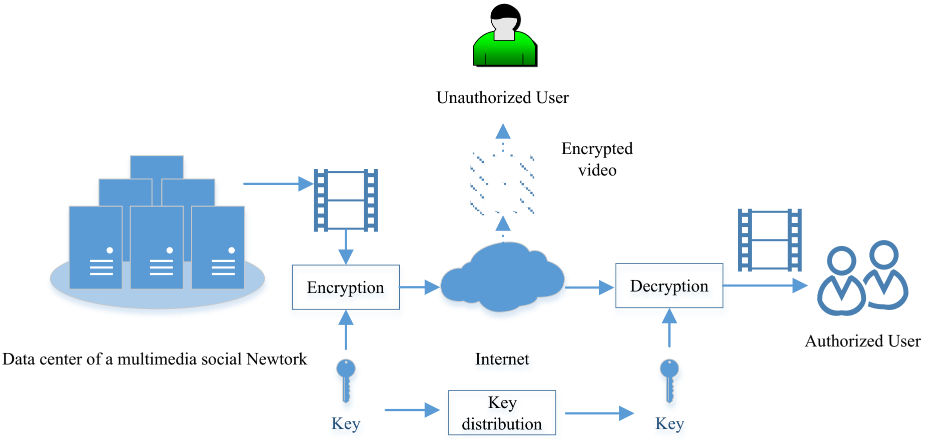


Fig. 11 The framework of the implementation of the proposed scheme in multimedia social network

5 Conclusions

By analyzing the impact of quantization parameter on the encryption of the sign of T1 s and the impact of encrypting inter-macroblock non-zero coefficients, a selective encryption scheme for H.264/AVC is proposed and a framework of its implementation in multimedia social network is presented. IPMs, sign of RLevel and selected T1 s within intra-macroblock, sign of MVD are used for encryption. In order to achieve tradeoff between security and efficiency, the encryption of the sign of T1 s is based on a predefined threshold according to QP value. Experimental results and analysis show that proposed scheme can achieve good performance in perceptual scrambling effect, bitrate and computational cost, and it can effectively resist brute-force attacks, differential attacks, histogram analysis, and replacement attacks. Besides, it is completely compatible to H.264/AVC standard. It has great potential to be implemented for the secure video content distribution in multimedia social network.

Acknowledgments This work was supported in part by project supported by National Natural Science Foundation of China (Grant No. 61370225, 61572182), project supported by Hunan Provincial Natural Science Foundation of China (Grant No.15JJ2007), supported by the Scientific Research Plan of Hunan Provincial Science and Technology Department of China (2014FJ4161), supported by Project of the Priority Academic Program Development of Jiangsu Higher Education Institutions, and supported by Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology.

Thank for the reviewers and the guest editor for their kind suggestions.

References

1. Ahn J, Shim H, Jeon B, Choi I (2004) Digital video scrambling method using intra prediction mode," in *Proc. PCM*, LNCS 3333. pp. 386–393
2. Alvarez G, Li S (2006) Some basic cryptographic requirements for chaos-based cryptosystems. Int J Bifurcation Chaos 16(8):2129–2151
3. Au-Yeung SK, Zhu S, Zeng B (2009) Partial video encryption based on alternating transforms. IEEE Signal Process Lett 16(10):893–896
4. *Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/ISO/IEC* 14496–10 *AVC)*, document JVT-G050, Joint Video Team (JVT), Mar. 2003
5. Friedman WF (1987) The index of coincidence and its applications in cryptanalysis. Aegean Park Press, California
6. Fu Z, Sun X, Liu Q, Zhou L, Shu J (2015) Achieving efficient cloud search services: Multi-keyword ranked search over encrypted cloud data supporting parallel computing. IEICE Trans Commun E98-B(1):190–200
7. Furht B, Socek D, Eskicioglu A (2004) Multimedia security handbook. CRC Press, Boca Raton, pp. 93–131
8. Goldreich O (2009) Foundations of cryptography volume II basic applications. Cambridge University Press, New York, pp. 373–481
9. Hong G, Yuan C, Wang Y, Zhong Y (2006) A quality-controllable encryption for H.264/AVC video coding, in *Proc. PCM*, LNCS 4261. pp. 510–517
10. JM Reference Software, ver. 10.2 (2011) [Online]. Available: http://iphome.hhi.de/suehring/tml/download/old_jm/
11. Kwon S, Choi W, Jeon B (2005) Digital video scrambling using motion vector and slice relocation, in *Proc. 2nd Int. Conf., ICIAR*, pp. 207–214
12. Li Y, Jiang J, Liang L (2007) Research and improvement of the video encryption algorithm for H.264. Acta Electron Sin 35(9):1724–1727
13. Li X, Feng Z, Hu Y (2009) Video encrypting scheme based on H.264 CAVLC entropy coding. Comput Eng Appl 45(34):114–117
14. Lian S, Sun J, Wang Z (2004) Quality analysis of several typical MPEG video encryption algorithms. J Image Graph 9(4):483–490
15. Lian S, Liu Z, Ren Z, Wang H (2006) Secure advanced video coding based on selective encryption algorithms. IEEE Trans Consum Electron 52(2):621–629
16. Lian S, Sun J, Liu G, Wang Z (2008) Efficient video encryption scheme based on advanced video coding. Multimed Tools Appl 38(1):75–89

17. Liu F, Koenig H (2010) A survey of video encryption algorithms. Comput Secur 29(1):3–15
18. Liu Z, Li X (2004) Motion vector encryption in multimedia streaming, in *Proc. 10th Int. Multimedia Modelling Conf.*, pp. 64–71
19. Liu Y, Yuan C, Zhong Y (2007) A new digital rights management system in mobile applications using H.264 encryption, in *Proc. 9th Int. Conf. Advanced Communication Technology*, pp. 583–586.
20. Peng F, Zhu X, Long M (2013) An ROI privacy protection scheme for H.264 video based on FMO and chaos. IEEE Trans. Inf Forensics Secur 8(10):1688–1699
21. Podesser M, Schmidt H, Uhl A (2002) Selective bitplane encryption for secure transmission of image data in mobile environments, in *Proc. 5th IEEE Nordic Signal Process. Symp.*, pp. 4–6
22. Raju C, Srinathan K, Jawahar C (2008) A real-time video encryption exploiting the distribution of the DCT coefficients, in *Proc. IEEE Region* 10 *Conf. (TENCON 2008)*, pp. 1–6
23. Ren Y, Shen J, Wang J, Han J, Lee S (2015) Mutual verifiable provable data auditing in public cloud storage. J Internet Technol 16(2):317–323
24. Shahid Z, Chaumont M, Puech W (2009) Fast protection of H.264/AVC by selective encryption, in *Proc. SinFra IPAL Symp.*, pp. 11–21
25. Shahid Z, Chaumont M, Puech W (2011) Fast protection of H.264/AVC by selective encryption of CAVLC and CABAC for I and P frames. IEEE Trans Circuits Syst Video Technol 21(5):565–576
26. Stallings W (2011) Cryptography and network security, principles and practice, Fifth edn. Pearson, Upper Saddle River
27. *Video Trace Library* (2011) [Online]. Available: http://trace.eas.asu.edu/yuv/index.html
28. Wang Z, Bovik A (2009) Mean squared error: love it or leave it? A new look at signal fidelity measures. IEEE Signal Process Mag 26(1):98–117
29. Wang Z, Bovik A, Sheikh H, Simoncelli E (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
30. Wang Y, Cai M, Tang F (2007) Design of a new selective video encryption scheme based on H.264, in *Proc. Int. Conf. Computational Intelligence and Security*, pp. 883–887
31. Wang Y, O'Neill M, Kurugollu F (2013) A tunable encryption scheme and analysis of fast selective encryption for CAVLC and CABAC in H.264/AVC. IEEE Trans Circuits Syst Video Technol 23(9):1476–1490
32. Wiegand T, Sullivan G, Bjontegaard G, Luthra A (2003) Overview of the H.264/AVC video coding standard. IEEE Trans Circuits Syst Video Technol 13(7):560–576
33. Xing S, Jiang J, Qi M (2009) An intra prediction mode-based video encryption algorithm in H.264, in *Proc. Int. Conf. Multimedia Information Networking and Security (MINES '09)*, pp. 478–482
34. *Xiph.org Video Test Media* (2013) [Online]. Available: http://media.xiph.org/video/derf/
35. Zhang Z (2011) Digital rights management ecosystem and its usage controls: a survey. Int J Digital Content Technol Appl 5(3):255–272
36. Zhang Z (2012) Frontier and methodologies on digital rights management for multimedia social networks. Int J Digital Content Technol Appl 6(9):245–249

Fei Peng received the Ph.D. degree in Circuits and Systems from the South China University of Science and Technology, Guangzhou, China, in 2006. He is a visiting fellow of the Department of Computer Science, University of Warwick, U.K. in 2009-2010. Currently, he is a professor in the School of Computer Science and Electronic Engineering, Hunan University, Changsha. He is also the Director of the Department of Information Engineering. His areas of interest include digital watermarking and digital forensics.

Xiaoqing Gong received the B.S. degree in Electronic and Information Engineering from Jiangxi Normal University, Nanchang, Jiangxi, China, in 2013. Currently, he is a master's candidate of the School of Computer Science and Electronic Engineering, Hunan University, Changsha. His areas of interest include multimedia security and digital watermark.

Min Long received the Ph.D. degree in Circuits and Systems from the South China University of Science and Technology, Guangzhou, China, in 2006. She is a visiting fellow of the Department of Computer Science, University of Warwick, U.K. in 2009-2010. Currently, she is a professor in the College of Computer and Communication. Her areas of interest include digital watermarking and chaos-based secure Communication.

Xingming Sun received the PhD degree in computing science from Fudan University, China, in 2001. Currently, he is a professor in the School of Computer & Software, Nanjing University of Information Science & Technology, China. His research interests include network and information security, digital watermarking, and data security in cloud. He is a senior member of the IEEE.