A Scalable Video Encryption Algorithm for H.264/SVC

Guo Jie, Qiu Weidong
School of Information Security Engineering
Shanghai Jiaotong University
Shanghai, 200240, China
guojie@sjtu.edu.cn, qiuwd@sjtu.edu.cn

Du Chao, Chen Kefei
Department of Computer Science and Engineering
Shanghai Jiaotong University
Shanghai, 200240, China
nssirius@126.com, kfchen@sjtu.edu.cn

***Abstract*— As the extension of H.264/AVC standard, Scalable Video Coding (SVC) has achieved significant improvements in coding efficiency and scalability. Most of the available video encryption algorithms are designed based on H.264/AVC standard. In this paper, the extensional technologies in SVC relative to H.264/AVC are analyzed. According to the new characteristics, a scalable video encryption algorithm is proposed for H.264/SVC. Some new features, such as motion prediction information of spatial scalability and key picture information of MGS (medium-grain quality scalability), are encrypted with three domains in hierarchical layers using different keys. Performance analysis of the new algorithm is reported and its strength against cryptanalysis attacks is discussed. Experimental results indicate that our scheme has adequate performance to allow for its use in real world applications.**

Keywords- H.264/SVC; video encryption; multimedia security

I. INTRODUCTION

With the rapid growth of video service in web applications, the requirements of video coding technology have changed from simply chasing high compression ratio to adapting to the various needs or preferences of end users as well as to varying terminal capabilities or network conditions. SVC (scalable video coding) is a highly attractive solution to the problems. The objective of the SVC standardization is to encode a high-quality video bit stream that also contains one or more subset bit streams. A subset video bit stream is derived by dropping packets from the larger video to reduce the bandwidth required for the subset bit stream. So the same bit stream can be adapted to different spatial resolution, different temporal resolution, or different quality resolution [1]. In July 2007, the SVC project is standardized as an amendment of the H.264/MPEG-4 AVC standard.

With continuous popularization of scalable video encoding, the necessity of practical security technologies for H.264/SVC is unquestionable. In the recent years, video encryption has been heavily researched. The encryption can be conducted before, in or after video compression process [2]. However, most of the available video encryption algorithms are designed based on H.264/AVC or MPEG2 standard. There is also some work in the literature on secure encryption for scalable bit streams. Because the base layer of SVC is encoded similar to AVC, the compression integrated encryption approaches for AVC can be basically employed for SVC [3]. The approach of [4] takes advantage of SVC to implement transparent encryption after video compression. The approaches of [5-8] have been proposed for SVC encryption, which all preserve the NALU (network abstraction layer units) structure and encrypt the NALU payload directly.

A series of new technologies, such as temporal scalability, spatial scalability and quality Scalability, are introduced in H.264/SVC standard. So there are more choices in selecting encryption information or designing encryption algorithm. In this paper, the new technologies in H.264/SVC standard are analyzed. Based on the new characteristics of SVC, a multi-layered scalable video encryption algorithm is proposed to provide different security levels according to different application situation. The proposed scheme offers the computational efficiency by encrypting different domains selectively and the security through the use of different keys. Meanwhile, the proposed scheme preserves SVC scalability and format-compliance. Experimental results indicate that our scheme has adequate performance to allow for its use in real world applications.

The rest of this paper is organized as follows. Section 2 presents some new characteristics of H.264/SVC. Section 3 elaborates our encryption scheme. Section 4 evaluates our scheme using real data. Section 5 concludes this paper.

II. NEW CHARACTERISTICS OF H.264/SVC

An SVC bit-stream consists of a base layer and enhancement layers. Each enhancement layer improves the video in one of three "scalability dimensions," namely resolution (to enable spatial scalability), quality (to enable SNR scalability), and time (to enable different frame rates) [2], as shown in Fig.1.

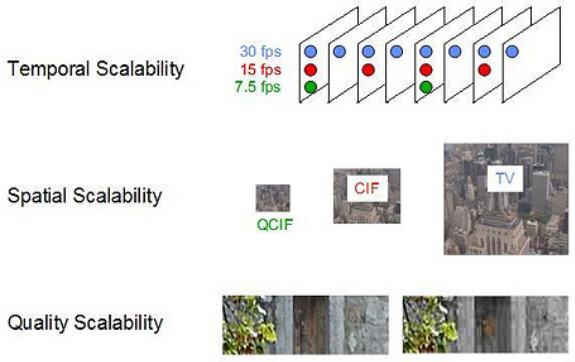


Figure 1. Example of SVC scalability.

Similar to H.264/AVC, the temporal scalability is supported with a reasonable number of temporal layers. The only related change in SVC is the signaling technology of temporal layers. For supporting spatial scalable coding, SVC follows the conventional approach of multilayer coding. In order to improve coding efficiency, additional inter-layer prediction mechanisms are incorporated. Quality scalability is supported by the general concept for spatial scalable coding, which is referred to CGS (coarse-grain quality scalable coding). For increasing the flexibility of bit stream adaptation and error robustness, a new approach MGS (medium-grain quality scalability) is included in the SVC design [1].

III. OUR SCHEME

According to the new characteristic of SVC, a new video encryption scheme is proposed. In our scheme, an encoder structure with two spatial layers is employed, as shown in Fig.2. The encryption scheme is designed in hierarchical levels with three keys, which can achieve different confidentiality at different level. Inter-layer prediction information between different layers is employed for the first level encryption. Key picture information of MGS is employed for the second level encryption. The intra DCs and the sign bits of nonzero ACs of I frames are employed for the third level encryption.

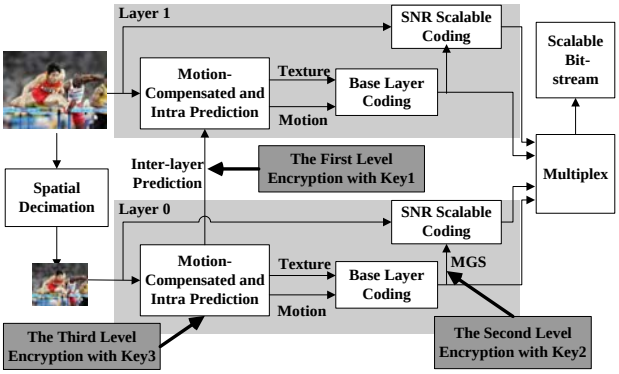


Figure 2. Multi-layered video encryption structure.

A. The First Level Encryption

Different from H.264/AVC, the main goal of inter-layer prediction in H.264/SVC is to enable the usage of the lower layer information to improve the rate-distortion efficiency of the enhancement layers. In order to improve the coding efficiency for spatial scalable coding, two additional inter-layer prediction concepts have been added in SVC: prediction of macroblock modes and associated motion parameters and prediction of the residual signal [1]. The partitioning data of the enhancement layer macroblock together with the associated reference indexes and motion vectors are derived from the corresponding data of the macroblock in the reference layer, as shown in Fig.3.

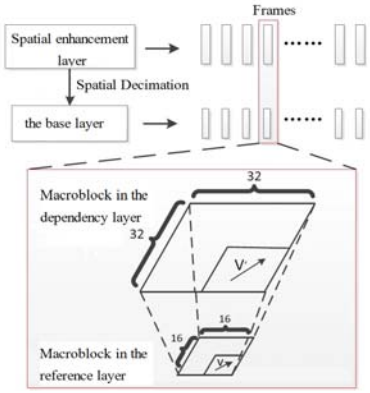


Figure 3. Inter-layer motion prediction schematic diagram.

Obviously, the modification of inter-layer motion prediction parameters will lead to the inevitable chaos of the enhancement layer data. So, the first level encryption is designed by encrypting the inter-layer prediction information, as shown in Fig.2. Denote a plaintext's payload as P_0, P_2, $\cdots$, P_m, and a stream key generator generates a key stream as k_0, k_2, $\cdots$, k_m, all in bytes. The cipher text C_m is:

$$C_m = P_m \oplus k_m \quad (1)$$

The stream key generator is constructed, as shown in Fig.4. An initial key is set to K_0. A series of hash transformation is applied to generate a secret key sequence (K_0, K_1, K_2, $\cdots$, K_n). According to the size of the plaintext, a key stream is chosen for the encryption.

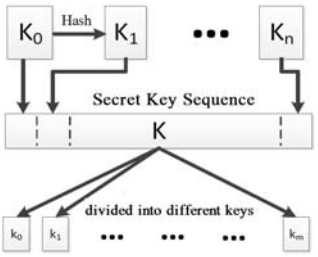


Figure 4. Generation of stream keys.

B. The Second Level Encryption

Different from H.264/AVC, a Key Picture concept for hierarchical prediction structures is introduced in the MGS design of H.264/SVC. The high-level signaling allows a switching between different MGS layers in any access unit. So, any enhancement layer NAL unit can be discarded from a bit stream to provide a quality scalable coding [1]. All pictures of the coarsest temporal layer are transmitted as key pictures, which are the reference pictures of the enhancement layer reconstruction.

In our scheme, the reconstruction information of the key picture is employed for the second level encryption, as shown in Fig.2. In H.264, every pixel in a macro block is represented by a luminance component and two chrominance components. In our scheme, AES encryption technology is applied to encrypt the luminance and chrominance blocks of

the reconstructed information. The encryption process is shown in Fig.5.

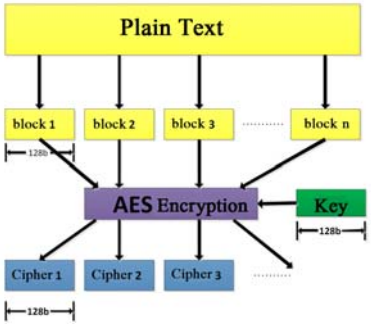


Figure 5. Encryption of luma and chroma block of reconstruction.

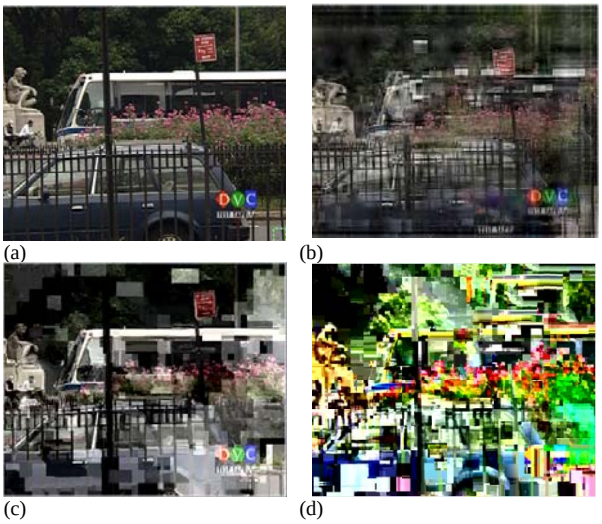
C. *The Third Level Encryption*

The first two schemes mainly target the security of the enhancement layer data. Multi-layered scalable bit-stream distribution requires the protection for its individual layers. So the security of the base layer data should be considered. Since the base layer of SVC is encoded similar to AVC, all encryption schemes for AVC can be basically employed in the base layer. Similar to [9], the intra DCs and the sign bits of nonzero ACs of I frames are employed for our third level encryption. The selected data is encrypted with a stream cipher, which is similar to the process of the first level encryption. The secret key is a randomly generated bit-stream of length m which can be represented as k_0, k_2, $\cdots$, k_m. The bit-stream can be represented as P_0, P_2, $\cdots$, P_m. The encryption function can be described as:

$$E_k(P) = P_m \oplus k_m \quad (2)$$

where $\oplus$ is the binary XOR operation.

IV. EXPERIMENTAL RESULTS



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

Figure 6. The fifth frame of original video and encrypted video. (a) Original video. (b) Encrypted video in the first scheme. (c) Encrypted video in the second scheme. (d) Encrypted video in the third scheme.

Some experiments have been conducted to test the real encryption performance of our scheme. Experiments are performed with the source code of JSVM (Joint Scalable Video Model) software, which is the official software for the SVC project of the Joint Video Team (JVT) of the ISO/IEC Moving Pictures Experts Group (MPEG) [10]. In our experiments, a video test sequence "Bus" are employed. The parameters are set to be 2 spatial layers, 4 temporal layers, and 2 quality layers. BUS_176_144 sequence with the frame size 176×144 is applied for the low space layer, and BUS_352_288 sequence with the frame size 352×288 is applied for the high space layer. The fifth frame of original sequences is shown in Fig.6 (a). Fig.6 (b) shows the effect of the first level encryption. The fifth frame is blurred, but still comprehensible. Fig.6 (c) shows the effect of the second level encryption. The fifth frame is obscured, but still comprehensible. Fig.6 (d) shows the effect of the third level encryption. The fifth frame is incomprehensible.

A. *Computational Complexity*

The computing complexity of the proposed encryption scheme depends on the amount of encrypted data, the cost of sub-key generation and the cost of the stream cipher. The cost of sub-key generation depends on the computing complexity of key generation and the number of sub-keys to be generated. In our scheme, some important features are selected for the encryption. Compared with that of original video, the amount of encrypted data is very low. In experiments, we test the time efficiency of the encryption/ decryption process, which is measured by the time ratio between encryption/decryption and compression/ decompression. Table I gives the experimental results. Seen from the table, the encryption/decryption operation makes up no more than 10% percent of the compression/ decompression operation. It is shown that the encryption/ decryption process does not affect the compression/ decompression process greatly.

TABLE I. TEST OF TIME EFFICIENCE AND STREAM FILE SIZE

	Run time (second)	***Time ration***		***Encoded file size***
		Encryption	***Decryption***	
No encryption	1470	-	-	477k
The first scheme	1512	2.85%	3.27%	477k
The second scheme	1557	5.92%	6.18%	477k
The third scheme	1493	1.56%	2.01%	477k

In our scheme, the encryption is operated during the process of video compression. Some bits of the encoding coefficients are modified, but the length of codewords is identical. Whether the video is encrypted or not, the file size of the encoded stream remains 477K, as seen in Table I. So, the proposed scheme does not increase the requirements of the network bandwidth.

B. *Perceptual Quality of Encrypted Video*

Till now, peak signal-to-noise ratio (PSNR) is often used to evaluate videos' perceptual quality. We test the PSNRs

(the average value of Y, U and V) of the unencrypted videos and the encrypted videos. Since the performance of video encryption scheme by employing DCT coefficients has been greatly researched in the published literature [2], we focus on the performance of the encryption schemes with the new characteristics of H.264/SVC. The experimental results are shown in Fig.7. The comparison of PSNR between the unencrypted video and the encrypted video obtained by the first level encryption is shown in Fig.7 (a). Fig.7 (b) shows the experimental results with the second level encryption. It is shown that the encrypted videos' PSNRs are all smaller than the ones of the unencrypted video. Thus, our encryption scheme can blur the videos, which keeps it secure in different perception. The customers can choose different level of our scheme to achieve the required perceptual protection.

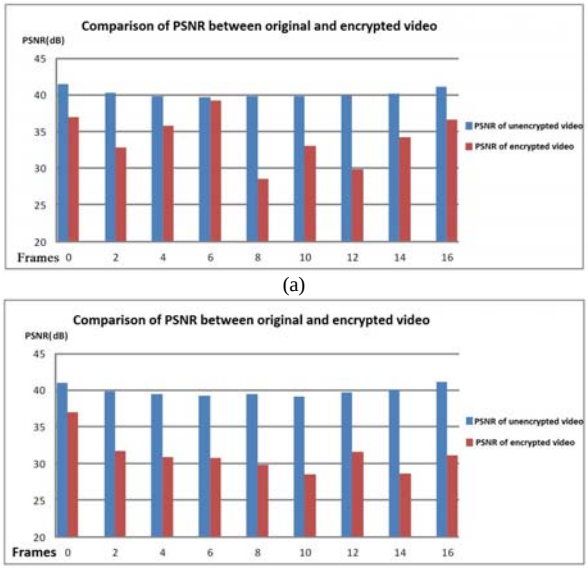


Figure 7. Comparison of PSNR between original and encrypted video.(a) Encrypting with the first scheme. (b) Encrypting with the second scheme.

C. *Security Analysis*

In our scheme, different key is applied for different level encryption. Each level encryption is controlled under a 128-bit sub-key. Thus, for each level encryption, the brute-force space is 2^{128}. This brute-force space is too large for attackers to break the cryptosystem. Meanwhile, the difficulty of brute-force attack is determined by the encryption scheme. Some typical encryption algorithms, such as AES and stream cipher, are employed in our scheme. We also can improve the security of the stream key generator with a keyed hash function. The proposed scheme can provide a good level of security against a brute-force attack.

V. CONCLUSION

This paper focuses on the problem of how to realize a scalable encryption of H.264/SVC videos. Based on an analysis of the extensional technologies in H.264/SVC, a new scalable video encryption algorithm is presented. Different from the previous video encryption algorithm, the new characteristics of H.264/SVC, such as inter-layer motion prediction of spatial scalability and key picture information of MGS, are employed for our design. A hierarchical encryption structure with different keys is constructed, which can satisfy the different requirements of different occasions. The proposed scheme is evaluated with JSVM software and some video test sequences. It is shown that our scheme can achieve satisfactory encryption results with small computations.

ACKNOWLEDGMENT

This work is supported by NSFC (61073157), NSFC (61070249) and NSFC (61133014).

REFERENCES

[1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Codingm Extension of the H.264/AVC Standard", IEEE Transactions On Circuits And Systems For Video Technology, vol. 17, no. 9, pp.1103-1120, SEPTEMBER 2007.

[2] T. St¨utz and A. Uhl, "A Survey of H.264 AVC/SVC Encryption", IEEE Transactions On Circuits And Systems For Video Technology, vol. 22, no. 3, pp. 325-339, march 2012.

[3] E. Magli, M. Grangetto, G. Olmo, "Transparent encryption techniques for H.264/AVC and H.264/SVC compressed video", Signal Processing,vol.91, no.5, pp.1103–1114, May 2011.

[4] T. St¨utz and A. Uhl, "Format-compliant encryption of H.264/AVC and SVC," in Proc. 10th IEEE ISM, Dec. 2008, pp. 446 - 451

[5] J. Apostolopoulos, "Architectural principles for secure streaming and secure adaptation in the developing scalable video coding (SVC) standard," in Proc. IEEE ICIP, Oct. 2006, pp. 729–732.

[6] H. Kodikara Arachchi, X. Perramon, S. Dogan, and A. M. Kondoz, "Adaptation-aware encryption of scalable H.264/AVC video for content security," Signal Process. Image Commun., vol. 24, no. 6, pp. 468–483,2009.

[7] H. Hellwagner, R. Kuschnig, T. St¨utz, and A. Uhl, "Efficient in-network adaptation of encrypted H.264/SVC content," Elsevier J. Signal Process. Image Commun., vol. 24, no. 9, pp. 740–758, Jul. 2009

[8] Z. Wei, Y.D Wu, X.H Ding, and R.H. Deng, "A Scalable and Format-Compliant Encryption Scheme for H.264/SVC Bitstreams", Signal Processing: Image Communication, vol.27, no.9, pp.1011–1024, October 2012

[9] S. Li, G. Chen, A. Cheung, B. Bhargava, and K. Lo, "On the Design of Perceptual MPEG-Video Encryption Algorithms," IEEE Transactions on Circuits and Systems for Video Technology, vol.17 , no.2, pp.214 - 223. 2007

[10] "Joint Scalable Video Model, JSVM", Available: http://ip.hhi.de/imagecom_G1/savce/downloads/SVC-Reference-software. html