Layered Encryption for Scalable Video Coding

Chunhua Li
Department of Computer Science and Technology, Tsinghua University,
Beijing 100084, China
lich07@mails.tsinghua.edu.cn

Chun Yuan
Division of Information Technology, Graduate School at Shenzhen, Tsinghua University,
Shenzhen 518055, China
yuanc@sz.tsinghua.edu.cn

Yuzhuo Zhong
Department of Computer Science and Technology, Tsinghua University,
Beijing 100084, China
zyz-dcs@mail.tsinghua.edu.cn

***Abstract*—A layered selective encryption scheme for Scalable Video Coding (SVC) is proposed in this paper. The main feature of this scheme is making use of the characteristics of SVC. This method fully meets the encryption requirements of SVC and the encryption procedures are carried out at the Network Abstractor Layer (NAL) level. Based on the different structure and importance of base tier and enhancement tiers, different domains are encrypted. For base tier, Intra-Prediction mode (IPM) and residual sign are selected. For enhancement tiers, temporal scalability and spatial/SNR scalability are distinguished. Furthermore, key generation and distribution schemes are presented. Stream cipher—Leak EXtraction (LEX) algorithm is adopted to reduce computational cost. Experiments were performed to verify the proposed method using the joint scalable video model (JSVM), and the experimental results show that the proposed method protects the SVC streams effectively and supports full scalability; meanwhile, it can guarantee the robustness to transmission errors.**

Keywords- Scalable Video Coding; layered encryption; LEX;Network Abstractor Layer

I. INTRODUCTION

Scalable Video Coding (SVC) is an extension of H.264/AVC with high efficiency and scalability. At present, SVC is deemed most promising video format for streaming applications over heterogeneous networks and devices. From a global SVC stream, a scalable bit stream representation with lower resolution and/or quality can be obtained by discarding selected data. Video encryption methods are needed for the copyright protection of the scalable video contents during transmission. The method should provide security, time efficiency and error robustness. Furthermore, it is very important that the proposed video encryption scheme keeps the scalability of SVC.

Currently, various encryption schemes for the non-scalable as MPEG-2 and H.264/AVC have been proposed, but the encryption schemes for the SVC are restrictive.

In scalable contents encryption, the enhancement structure for the scalability was considered in [5]. Base layer encryption could provide enough security of the whole bit-stream because scalable contents are enhanced from the base layer. However, base layer encryption causes loss of the scalability when the video is adapted.

Protection for scalable video coding has been researched in [6] [7] [8]. Y. G. Won et al. [6] proposed an encryption algorithm to protect SVC bit-stream and a method for conditional access control to consume the encrypted SVC bit-stream. An encryption method to protect the region of interest (ROI) of SVC was presented in [7]. Recently, a selective encryption scheme for SVC was designed by S. W. Park et al. [8], but didn't discuss the key's generation and distribution in further detail. So an integrity solution including both encryption method and key management for protection of scalable video coding is still necessary.

In this paper, we aim to propose a novel SVC video encryption scheme. By analyzing SVC codec's properties, a layered encryption scheme is presented, which encrypts both base layer and enhancement layer. Moreover, considering the security and time efficiency, a selective encryption method is proposed. In this scheme, different domains are encrypted in different layers.

The rest of the paper is organized as follows. Section 2 briefly introduces the structure of SVC, LEX algorithm, and the motivation. Based on them, the proposed encryption scheme is presented in Section 3, and the experimental results are provided and discussed in section 4. Finally, some conclusions are drawn in section 5.

II. BACKGROUND AND SCHEME OVERVIEW

A. *Overview of SVC*

A layered stream representation of SVC in terms of temporal, spatial, and SNR resolution is shown in Fig. 1. In SVC encoding scheme, each video stream is encoded in multiple video quality tiers. First tier which provides the basic quality of the video is called "Base Tier" while the other tiers are used to enhance the overall video quality of the tier are called "Enhancement Tiers". It is important to note that the lower tier (Base tier) is more important than the higher tiers (Enhancement tiers) to decode a particular stream. Thus, we leverage the characteristics of SVC, to encrypt different tiers with different domains using different keys.

978-1-4244-4131-0/09/$25.00 ©2009 IEEE

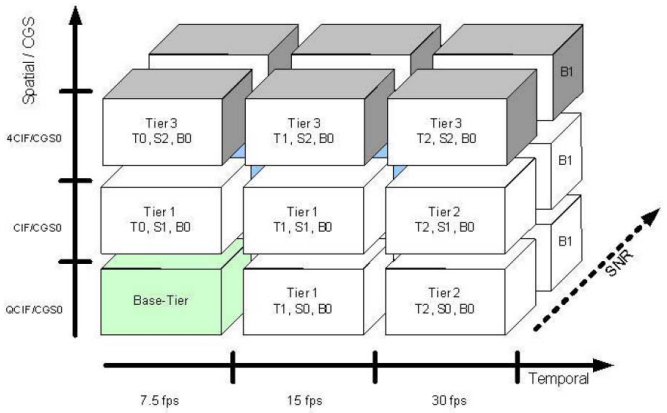

Figure 1. Layered representation of SVC

In base tier, the basic concepts of intra prediction are employed as in H.264/AVC. For enhancement tiers, the redundancy between different tiers is exploited by additional inter-layer prediction concepts that include prediction mechanisms for motion parameters as well as texture data (intra and residual data).

Based on them, our proposed encryption scheme encrypts intra prediction modes, motion vector difference (MVD) value and residual data for both intra-coded and inter-coded data.

In SVC, these parts are encoded using two entropy coding methods: context-adaptive variable-length coding (CAVLC) and context-adaptive binary arithmetic coding (CABAC). In order to keep format compliance, there is a wide agreement that the encryption apply on MVD and residual data's sign.

Additionally, an important feature of SVC is the provision of scalability on the stream level. Bit-steams for reduced resolution/quality can be simply obtained by discarding Network Abstraction Layer (NAL) units from a full SVC bit-stream. In order to make the encryption scheme applicable in bit-stream extraction process, the encryption should be applied in NAL level to ensure that each NAL unit can be decrypted independently. Furthermore, considering security, different NAL units use different keys in our scheme.

B. *Encryption Scheme*

In order to make the encryption in accordance with the scalability of SVC, three requirements need to satisfy. Both the requirements and the corresponding solutions are as follows:

First, all layers including base and enhancement layers in SVC should be encrypted for robustness of video security. Our methods apply to all types of data in the SVC bit-stream.

Second, the encryption scheme should be applicable in bit-stream extraction process. We perform encryption segment by segment with the NAL unit.

Third, the encryption method should be light-weighted in computational complexity. The proposed algorithm encrypts sensitive data such as Intra Prediction Modes (IPM), sign of MVD and residual.

We use Leak Extraction (LEX) as the stream cipher for our scheme. LEX is a stream cipher algorithm based on the round transformation of AES [4]. LEX provides the same key agility and short message block performance as AES while handling longer messages faster than AES. The LEX algorithm is shown in Fig. 2.

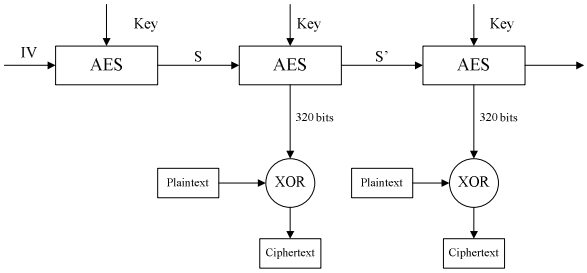

Figure 2. LEX encryption algorithm

III. PROPOSED SCHEME

From the above analysis of the structure of SVC, we know that base tier of SVC bit-stream is more important than the enhancement tiers because the information of base tier is the basis of the enhancement tiers.

According to the encryption requirements, we propose a layered video encryption scheme that encrypts different parts of base tier and enhancement tiers to satisfy the trade-off between the security and computation cost.

A. *Base Tier Encryption*

As described above, the base tier of SVC can be decoded individually and it is the basis for enhancement tiers' decoding. Therefore it should be protected with the highest level of security.

Base tier of SVC is encoded in intra-coded mode. Both the intra-prediction mode and the residual data should be encrypted to gain high security.

1) Intra-Prediction mode encryption: In SVC, the intra-prediction modes are encoded with Exp-Golomb codes [1]. This kind of codeword is composed of R zeros, one '1' and R bits of information (Y). The encryption process is shown in Fig. 3.

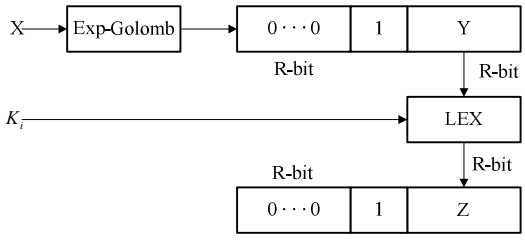

Figure 3. Intra-prediction mode encryption algorithm

That is, X is firstly encoded into a variable-length code with Exp-Golomb coding, and then only the information part Y is

encrypted into Z with a stream cipher. The encryption process is defined as:

$$Z = C_e(Y, K_i) = C_e(Expg_e, K_i) \quad (1)$$

Where K_i is the i-th frame's encryption key, $C_e()$ is the encryption process of a cipher and $Expg_e()$ is the encoding process of Exp-Golomb coding. This encryption algorithm realizes encryption and variable-length coding at the same time, and keeps the codeword's length unchanged.

2) Sign encryption: Fig. 4 shows the residue block encryption algorithm. After transformation and quantization of 4×4 blocks, the produced coefficients are encrypted with sign encryption under the control of key. Then, the encrypted coefficients are encoded with entropy coding (CAVLC/CABAC), which produces the encrypted code stream.

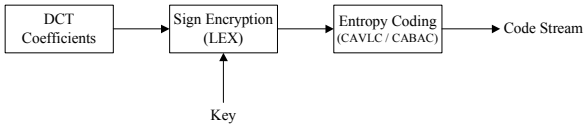

Figure 4. Sign encryption algorithm

B. Enhancement Tier Encryption

Considering the security and computational cost, the proposed scheme applies a light-weighted encryption on enhancement tier, which utilizes the properties of scalability type in each enhancement tier.

S. W. Park et al. [8] presented that the proportion of bits for each domain in the bit-stream comprising one layer is greatly influenced by the scalability type. Experiment results show that the residual data domain occupies most parts in the spatial scalability and the SNR scalability layer, while both motion vector difference (MVD) values and residual data are important components in the temporal scalability layer. Thus, we encrypt the different parts for different scalability layers, respectively.

1) Temporal scalability layer: In the temporal scalability layer, SVC adopts hierarchical B prediction structure, both MVD and residue data should be encrypted to gain high security. The encryption process is described in Fig. 5.

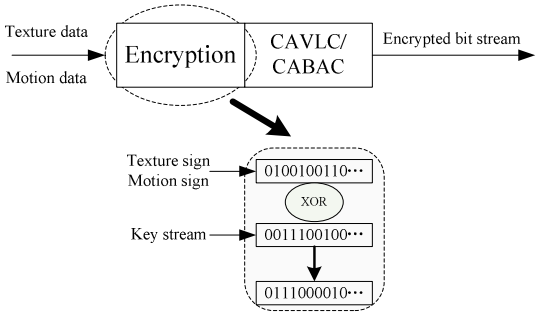

Figure 5. Encryption method for temporal scalability layer

2) Spatial and SNR scalability layer: Spatial scalability is achieved by an oversampled pyramid approach. And SNR scalability can be considered as a special case of spatial scalability for which the picture sizes of base and enhancement layer are identical. There are four steps of this method:

> Step 1: Decide NAL units belonging to a certain spatial/SNR layer,
> Step 2: Filter VCL (Video Coding Layer) NAL units,
> Step 3: Divide residual data into sign and absolute value,
> Step 4: Encrypt the sign data with LEX.

Figure 6. Encryption method for spatial/SNR scalability layer

C. Key Generation and Distribution

In our proposed encryption scheme, the procedure is carried out at the NAL level, both base tier and enhancement tier encryption operations are introduced. In order to maintain security, and scalability function of SVC, the layer-based encryption operations can be controlled by different keys. That is, data streams can be encrypted segment by segment with the NAL unit, which strength the system security, keep the scalability function, and also improve the robustness to transmission errors. The key used by stream cipher is generated by using the NAL_unit_type, dependency_id, temporal_id, and quality_level for each NAL unit.

IV. Experiment Results

We have implemented the proposed method in JSVM 11[2]. SVC test sequence, "Foreman" is used for the experiment. The sequence is encoded by 2 spatial layers (CIF, QCIF), 2 temporal levels (15fps, 30fps) and 2 SNR layers (base SNR layer, enhancement SNR layer). The GOP size is 32, where the first frame of each GOP is intra-coded.

A. Security Analysis

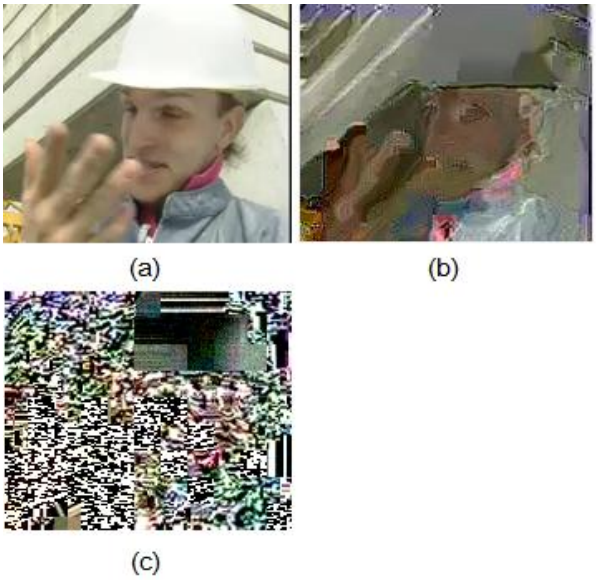

Figure 7. Results of video encryption. (a) is the original video, (b) is the video with only base tier encrypted, and (c) is the video with all tiers encrypted

In our proposed encryption scheme, both the predicted information and the residue data are encrypted, which make the video unintelligible. Also, the adopted stream cipher LEX is based on AES, whose security has been analyzed and confirmed [4]. The encryption results of the sample video are given in Fig. 7. Visual pattern of video encrypted only base tier and all tiers are shown, respectively. The encrypted video is too chaotic to be understood. Thus, this encryption scheme is of high security. Table I shows PSNR of the decoded result in Fig. 7.

TABLE I. PSNR RESULTS OF VIDEO

Encryption data	PSNR Y	PSNR U	PSNR V
Original video	36.74	41.07	42.64
Base tier	29.92	34.16	35.94
Base tier + enhancement tier	12.87	13.68	14.77

B. *Computation Cost*

The data volumes to be encrypted and the cost of the stream cipher are the main computing cost for our proposed scheme. The encrypted data (including intra-prediction mode, signs of MVD and residue data) volumes are fewer compared with the whole volumes. Moreover, the computational cost of LEX is only 40% of AES, the operation time ratio between the stream ciphers and the encoding process is very small. We test the time-efficiency of the encryption/decryption and the results are given in Table II.

TABLE II. TIME-EFFICIENCY OF THE PROPOSED SCHEME

Sequence	Format	Time Ratio	
		Encryption/ Compression	Decryption/ Decompression
City	QCIF 15Hz	0.5%	3.5%
	CIF 30Hz	1.1%	4.9%
Foreman	QCIF 15Hz	0.8%	4.7%
	CIF 30Hz	1.0%	5.6%
Soccer	QCIF 15Hz	0.6%	3.7%
	CIF 30Hz	0.8%	4.4%

C. *Error Robustness*

In this encryption scheme, we encrypt the video stream with the NAL as a segment, and assign each NAL unit with a different key. The proposed scheme can realize synchronization easily when transmission errors happen. That is, if errors happen in a NAL unit, then only the corresponding NAL unit is fault-decrypted, while other ones can still be decrypted correctly.

V. CONCLUSION

In this paper, an efficient layered encryption scheme for Scalable Video Coding (SVC) is proposed. For encryption, we firstly analyzed the characteristics of SVC, described encryption requirements for SVC, and presented the stream cipher algorithm: LEX, then our scheme is described in detail. For base tier, intra-prediction mode and residual sign are encrypted, while enhancement tiers are classified into temporal scalability layer and spatial/SNR scalability layer. Furthermore, we proposed key generation and distribution scheme to guarantee the security and to improve the robustness to transmission errors. Experiment results show that the proposed algorithms satisfy SVC encryption requirements and provide high security, low computation cost as well as robust to transmission errors.

REFERENCES

[1] ISO/IEC JTC 1/SC29/WG11 and ITU-T SG16 Q.6, "Scalable Video Coding – Joint Draft 11, Doc. JVT-X201,"Jul. 2007.

[2] ISO/IEC JTC 1/SC29/WG11 and ITU-T SG16 Q.6, "Joint Scalable Video Model JSVM-11, Doc. JVT-X202,"Jul. 2007.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Trans. On Circuits and Systems for Video Technology, vol.17, no. 9, pp.1130-1120, 2007.

[4] B. Alex, "A New 128-bit Stream Cipher LEX," ECRYPT Stream Cipher Project Report, Available at, http://www.encrypt.eu.org/stream/lex.html , 2005.

[5] Bin B. Zhu, M.D. Swanson, and Shipeng Li, "Encryption and Authentication for Scalable Multimedia: Current State of the Art and Challenges," Proc. SPIE Internet Multimedia Management System V, vol. 5601, pp. 157-170, Oct. 2004.

[6] Y.G. Won, T.M. Bae, and Y.M. Ro, "Scalable Protection and Access Control in Full Scalable Video Coding," LNCS 4283, pp. 407-421, Nov. 2006.

[7] Y.G. Won, S.H. Jin, T.M. Bae and Y.M. Ro, "A Selective Video Encryption for the Region of Interest in Scalable Video Coding," IEEE Region 10 Conference, pp. 1-4, 2007.

[8] S.W. Park and S.U. Shin, "Efficient Selective Encryption Scheme for the H.264/Scalable Video Coding (SVC)," NCM 2008, pp. 371-376, Sept. 2008.