End-to-End Real-Time ROI-based Encryption in HEVC Videos

Mohammed Abu Taha
PPU
Hebron, Palestine
Email: m_abutaha@ppu.edu

N. Sidaty, W. Hamidouche and O. Dforges
IETR/INSA Rennes
Rennes, France
Email: wassim.hamidouche@insa-rennes.fr

J. Vanne and M. Viitanen
TUT
Tampere, Finland
Email: jarno.vanne@tut.fi

***Abstract*—In this paper, we present an end-to-end real-time encryption of Region of Interest (ROI) in HEVC videos. The proposed ROI encryption makes use of the independent tile concept of HEVC that splits the video frame into separable rectangular areas. Tiles are used to extract the ROI from the background and only the tiles forming the ROI are encrypted. The selective encryption is performed for a set of HEVC syntax elements in a format compliant with the HEVC standard. Thus, the bit-stream can be decoded with a standard HEVC decoder where a secret key is only needed for ROI decryption. In Inter coding, tiles independency is guaranteed by restricting the motion vectors to use only unencrypted tiles in the reference frames. The proposed solution is validated by integrating the encryption into the open-source Kvazaar HEVC encoder and the decryption into the open-source openHEVC decoder, respectively. The results show that this solution performs secure encryption of ROI in real time and with diminutive bitrate and complexity overheads.**

Keywords: User identity management, High Efficiency Video Coding (HEVC), tiles, Region of Interest (ROI), selective encryption, quality assessments.

I. Introduction

High Efficiency Video Coding (HEVC) is currently the newest video coding standard issued by the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group [1], [2]. The main objective of the HEVC standardization efforts is to enable 50% bitrate reduction for similar video quality [3] compared to its predecessor H.264/AVC [4]. In the upcoming years, HEVC is expected to replace the previous video coding standards in the emerging applications featuring 4K resolution, high dynamic range, virtual reality, etc. In many of these applications, security and confidentiality of multimedia contents also play an integral part and they have been widely investigated in the last decade [5], [6], [7], [8], [9],[10]. [11], [12], [13], [14], [15]. The most straightforward method to secure video content is to use standard encryption for the whole bitstream or only part of the bitstream related to the *Region of Interest (ROI)*. This approach is called *Naive Encryption Algorithm (NEA)* [6] and it addresses the video bitstream as textual data without paying attention to the structure of the compressed video. However, bit-streams encrypted with NEA are decodable only after a correct decryption event, when only parts of the video are encrypted. This process limits the usage of video content to only users who have the right permission on the encrypted parts. Moreover, NEA solutions are time consuming and not suitable for real-time video applications.

In this paper, we focus on selective encryption that hides or encrypts only the ROI in the video (human faces, personal data) and keeps the rest of the video (background) unencrypted. In our approach, the HEVC video is first split into independent rectangular regions called tiles [13] and then only the tiles belonging to the ROI are encrypted. The selective algorithms encrypt a set of HEVC parameters including *Motion Vector (MV)* differences, MV-signs, *Transform coefficients (TCs)*, TC-signs, as given in [14]. In this work, the encryption is extended to luma and chroma Intra prediction modes (IPMs) in HEVC. The main contribution of this paper is to apply independent tile concept in HEVC for encryption of ROI in a way that prevents the propagation of encryption outside the ROI in Intra and Inter encodings. The encryption is based on the chaotic generator proposed in [16]. Finally, the encryption/decryption processes are implemented in the real-time Kvazaar HEVC encoder [17], [18] and the openHEVC decoder [19], respectively.

The rest of the paper is organized as follows. Section II presents the related work. In Section III, we describe the proposed ROI encryption in HEVC. Experimental results are shown in Section IV. Finally, Section V concludes the paper.

II. Related Work

For the time being a couple of encryption algorithm has been proposed for HEVC video. Shahid et al. [20] proposed a selective encryption scheme by joint encryption and compression system lies on *Context Adaptive Binary Arithmetic Coding (CABAC)* bin string. Hamidouche et.al [6] proposed a fast and secure selective chaos-based crypto-compression system for HEVC and its scalable extension SHVC. Boyadjis et al. [7] proposed an extended Selective Encryption (SE) method for H.264/AVC and HEVC streams. Their approach tackles the main security challenges of selective encryption. The contribution in [7] is the improvement of visual distortion induced by SE. Several works proposed the encryption of ROI in the video. Peng et al. [8] presented an encryption scheme for ROI of H.264 video based on *flexible macroblock ordering (FMO)* and chaos, where the ROI was the human face areas. Dufaux et al. [9] proposed an effective approach to encrypt ROI based on code stream-domain encryption. Work in [11]

ISBN 978-90-827970-1-5 © EURASIP 2018

enables rectangular region privacy by de-identifying faces. This solution guarantees that face recognition software cannot reliably recognize de-identified faces even though part of the facial details are preserved. In [12] the authors investigated the privacy protection in the H.264/SVC (Scalable Video Coding). This solution detects face regions (ROI) first and then encrypts these ROI in the transform domain by scrambling the sign of the non-zero TCs of all SVC layers.

III. ROI Encryption In HEVC

In this section we propose a solution based on the tile concept to protect privacy in the HEVC standard.

A. Tile based Encryption System

The proposed ROI encryption is based on the tile concept introduced in HEVC. The tiles split the video frame into rectangles with integer number of blocks where Intra prediction and the entropy coding dependencies are broken at the tile boundaries. The proposed solution performs a selective encryption of ROI tiles at the CABAC bin string level. The most sensitive HEVC syntax elements are encrypted to deteriorate the visual quality of the ROI. The selective encryption process encrypts only the tiles that contain the ROI whereas the non ROI tiles (background) remain clear (not encrypted). The encryption process encrypts a set of HEVC parameters including MVs, MV signs, (TCs), and TC signs. They are encrypted in a format compliant with HEVC without increasing the bitrate of the encrypted video. In addition to these four parameters, this work investigates the HEVC format compliant encryption of IPMs.

B. Encryption of Intra Prediction Parameters

In HEVC, there are three scanning orders of the quantized TCs and the scanning order is derived for Intra coded blocks from the IPM. The proposed algorithm encrypts the IPM without changing the original scanning order of the IPM (before encryption). This enables the IPM encryption to be format compliant with HEVC. The proposed encryption solution of IPMs is performed as shown by Algorithm 1. First, the IPM elements of HEVC are classified into three sets $Set\_VER \in \{6,7,8,9,10,11,12,13,14\}$, $Set\_HOR \in \{22,23,24,25,26,27,28,29,30\}$, and $Set\_DIA \in \{0,1,2,3,4,5,15,16,17,18,19,20,21,31,32,33,34\}$.
Each set contains the prediction modes that share the same scanning direction *(horizontal, vertical or diagonal)*. The encryption process is carried out using a circular shift operation. Each IPM in a particular set is shifted according to key stream bits. 5-bit stream values required to the encryption process are produced from a chaos based generator. Then, a new IPM position is deduced inside the same set.

The luma and chroma IPMs are encrypted in the same manner. The encryption is fully format compliant, since we keep the scanning directions unchanged. Unlike the encryption of other syntax elements, the encryption of the IPMs is performed before the entropy coding and thus may decrease rate-distortion (RD) performance. The chaotic generator produces the necessary keystreams to obtain ciphered syntax elements. The applied key-stream generator is proposed in our previous work [16]. It is performed by two recursive filters of order three. The first filter contains a discrete Skew tent map and the second one contains a discrete piecewise linear chaotic map. These maps are used as non-linear functions. A new initial vector value is generated in each generator call; this value allows producing different key-stream bits on each generator call. The cryptographic security analysis of the key stream generator is detailed in [16].

Algorithm 1 ROI IPM encryption

Input: *Intra Prediction Mode* **IPM**
Output: *Encrypted Intra Prediction Mode* **E_IPM**
1: $Set\_VER \in \{6,7,8,9,10,11,12,13,14\}$
2: $Set\_HOR \in \{22,23,24,25,26,27,28,29,30\}$
3: $Set\_DIA \in \{0,1,2,3,4,5,15,16,17,18,19,20,21,31,32,33,34\}$
4: Call chaotic generator to produce bit steam **K**
5: **if IPM** > 5 And **IPM** < 15 **then**
6: **E_IPM=***Circular shift*(Set_VER, **IPM, K**)
7: **else if M** > 21 And **IPM** < 31 **then**
8: **E_IPM=***Circular shift*(Set_HOR, **IPM, K**)
9: **else**
10: **E_IPM=***Circular shift*(Set_DIA, **IPM, K**)
11: **end if**

The encryption of syntax elements in ROI bin stream (MV differences, MV-signs, TCs, TC-signs) is given by the following formula:

$$C(i) = P(i) \oplus X(i) \tag{1}$$

where *P(i)* denotes the syntax elements of ROI, *C(i)* the ciphered syntax elements, and *X(i)* the key stream bits. Furthermore, the encryption of luma and chroma IPMs is performed as follows: Let N be the number of elements in the set $V = [1, 2, \cdots, N]$, where $V \in \mathbb{R}^N$, n_b is number of bits produced by chaotic generator. and i is the index of the IPM. The new IPM produced at the ith position of IPM $V[i]$ is given by:

$$V_s[i] = V[(i + n_b) \mod N] \tag{2}$$

The decryption algorithm is performed by inverse operations of (1) and (2).

C. Encryption Propagation in Inter Video Coding

The decoding processes of the background tiles may use samples belonging to the encrypted tiles for Inter prediction. Merge mode in HEVC derives the MVs information from a list of spatial neighbour and temporal candidates. Therefore, these two decoding operations can propagate the encryption from the encrypted tiles to the background tiles when the ROI is not correctly decrypted. In the case of merge mode, we

ISBN 978-90-827970-1-5 © EURASIP 2018

TABLE II: BD-rate and complexity increase of the proposed encryption scheme in intra and Inter coding(4x4 tile configuration).

Resolution	Sequence	Intra coding (4×4 tiles)			Inter coding (4×4 tiles)		
		Bit rate loss (%)	Complexity increase (%)		Bit rate loss (%)	Complexity increase (%)	
		BD-rate	Encoding	Decoding	BD-rate	Encoding	Decoding
2560×1600	PeopleOnStreet	1.09	3.05	1.87	2.74	3.27	2.88
1920×1080	Kimono	2.86	3.16	1.21	10.89	3.87	1.96
	ParkScene	3.13	2.34	1.13	8.39	3.08	1.89
	Cactus	2.77	2.82	2.02	4.73	3.96	2.19
	BQTerrace	2.78	2.19	1.67	14.91	3.54	1.93
	BasketballDrive	2.16	3.16	2.15	12.88	3.78	2.44
1280×720	Vidyo1	2.76	2.13	1.32	11.53	2.60	1.91
	Vidyo3	3.98	2.31	1.41	8.71	2.98	2.07
	Vidyo4	3.48	2.25	1.48	13.51	2.71	1.88
Average		2.77	2.60	1.58	9.81	3.31	2.12

restrict the temporal candidates of the background tiles to be inside the background zone in the reference frame. In order to prevent the propagation of encryption outside the ROI tile, two other non-normative encoding constraints are enforced in the Kvazaar encoder (as shown in Figure 1):

1) The MVs in the reference frame are restricted to point only to the co-located tile of the predicted block.
2) The in-loop filters are disabled across the tile boundaries.

These constraints tend to have a negative impact on the RD-performance, depending on resolution and content of video, but they enable to perform a safe interpolation process at the tile boundaries.

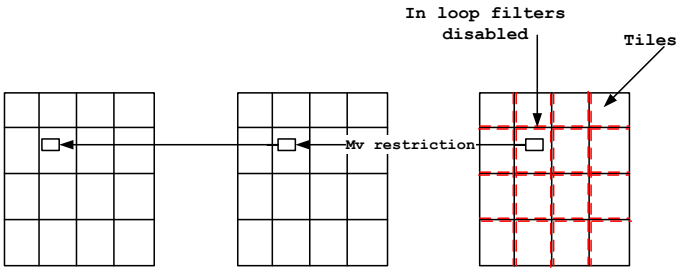

Fig. 1: Mvs and loop filter restrictions.

TABLE I: PSNR and SSIM values between original and encrypted videos (QP = 22).

Sequence	No-Encryption		ROI-Encryption	
	PSNR	SSIM	PSNR	SSIM
PeopleOnStreet	42.8	0.93	11.2	0.23
Kimono	42.2	0.96	9.9	0.22
ParkScene	43.3	0.91	10.7	0.20
Cactus	42.5	0.94	10.4	0.23
BQTerrace	41.8	0.90	10.8	0.24
BasketballDrive	41.5	0.96	10.1	0.23
Vidyo1	45.2	0.92	11.3	0.21
Vidyo3	44.6	0.94	10.9	0.20
Vidyo4	44.7	0.90	11.1	0.22

IV. Experimental Results

The encryption and decryption algorithms are implemented in the real time Kvazaar HEVC encoder and OpenHEVC decoder, respectively. We consider eight video sequences with different resolutions and frame rates from the HEVC common test conditions [21]. These videos are simultaneously encoded and encrypted by Kvazaar at four Quantization Parameter $(QP)values \in \{22, 27, 32, 37\}$ in both Intra and Inter (IPPP) coding configurations. The encrypted videos are encoded with uniform tiling configuration: 4 x 4 (i.e. four horizontal by four vertical repartition). The same encoder configuration but without tiles and encryption is used as an anchor. We consider the Bjontegaard BD-BR metric [22] which refer to the average bitrate difference between two bitrate-PSNR curves.

(a) (b)

Fig. 2: Frame #9 of different HEVC videos, encrypted with the proposed ROI encryption: (a) Correctly decrypted videos. (b) ROI-Encrypted videos.

ISBN 978-90-827970-1-5 © EURASIP 2018

A. Video Quality Measurements

Peak Signal to Noise Ratio (PSNR) and the Structural Similarity (SSIM) are used as objective metrics to evaluate the encrypted videos quality. The PSNR and the SSIM results of original (without encryption) and encrypted schemes is provided in **Table I**. In general, the encrypted sequences quality decrease significantly over those of the original. The average PSNR inside the ROI remains below 11.4 dB for all encrypted sequences. Correspondingly average SSIM values are remain below 0.24. These objective results show that the proposed solutions performs secure and adaptive encryption of ROI in the HEVC video.

In **Figure 2** we can observe that, the proposed encryption method conceals the objective quality of the ROI zone. Videos decoded and decrypted with the correct key are shown on the left side and decoded without decryption are illustrated on the right side.

The encoding process is performed using Inter and Intra coding for the 4 x 4 tile repartition, with MVs limitations and disabling the in-loop filters across the tile edges. The RD losses with Intra and Inter coding configurations is provided in **Table II** for 4 x 4 tile repartition. The bitrate overhead caused by the MVs restriction varies between 1% and 14.91% depending on the coding configuration (Inter and Intra), video content and the number of tiles within the frame (fixed here to 16 tiles by frame).

The BD-Rate loss for 4 x 4 tiles repartition in Inter coding is more than the loss in Intra coding, it reaches 9.81% and 2.77% respectively,. For example, the loss in BD-Rate with *Kimono1 (1920x1080)* video sequence with 4 x 4 tiles using Inter coding configuration is around 10.89%. However in Intra coding it remains low and does not exceed 2.86%.

B. Complexity Evaluation

The encoding and decoding complexities of 4 x 4 tile configuration are reported over the anchor configuration in **Table II**. The processor used in these evaluations has 32-bit multicore Intel Core (TM) i5 processor running at 2.60 GHz with 16GB of main memory. The operating system is Ubuntu 14.04 Trusty Linux distribution.

With considered video sequences, the additional time for encoding and decoding is negligible. For 4 x 4 tile configuration, the encoding time increases by 2.6% in Intra coding and 3.3% in Inter coding. The respective decoding times are 1.6% and 2.1% higher. This low complexity overhead at the encoder/decoder sides is mainly introduced by the encryption/decryption processes as well as the specific processing and bitrate increase related to the tiling repartitions.

These results confirm that the proposed selective encryption of ROI can be performed without noticeable performance compromises. The small computational overhead is crucial especially for portable devices having restricted processing power.

V. Conclusions

This paper proposed a new selective encryption solution for protecting privacy in HEVC video contents. Our solution encrypts only the ROI in the video and keeps the rest of the video unencrypted. The encryption algorithm is based on chaotic generator and the ROI is extracted through independent tile concept of HEVC. The proposed encryption method is performed at the CABAC binstring level so that the encrypted bitstream is decodable with a standard HEVC decoder and a privacy key is only needed in ROI decryption. However, some bit rate overhead is introduced in the HEVC coding process in order to prevent the propagation of the encryption outside the ROI. The proposed end-to-end encryption/decryption is integrated into two open-source software projects: HEVC Kvazaar encoder and OpenHEVC decoder. Experimental results showed that the proposed solution performs a secure protection of privacy in the HEVC video content with a small overhead in bit rate and coding complexity. It also prevents unexpected behavior of the decoder.

VI. Acknowledgement

This work is supported by the European Celtic-Plus project 4KREPROSYS - 4K ultraHD TV wireless REmote PRODuction SYStems and Academy of Finland (decision number 301820).

References

[1] ITU-T and ISO/IEC, "High Efficiency Video Coding, Document, ITU-T Rec. H.265 and ISO/IEC 23008-2 (HEVC)," 2013.

[2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[3] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the Coding efficiency of Video Coding Standards Including High Efficiency Video Coding (HEVC)," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1669–1684, 2012.

[4] I. ITU-T and I. JTC, "1: Advanced Video Coding for Generic Audiovisual Services, ITU-T Rec," *H*, vol. 264, pp. 14 496–10, 2009.

[5] Z. Shahid, M. Chaumont, and W. Puech, "Fast Protection of H. 264/AVC by Selective Encryption of CAVLC and CABAC for I and P Frames," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 5, pp. 565–576, 2011.

[6] W. Hamidouche, M. Farajallah, M. Raulet, O. Deforges, and S. El Assad, "Selective Video Encryption using Chaotic System in the SHVC Extension," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 1762–1766.

[7] B. Boyadjis, C. Bergeron, B. Pesquet-Popescu, and F. Dufaux, "Extended Selective Encryption of H. 264/AVC (CABAC) and HEVC Encoded Video Streams," *IEEE Transactions on circuits and systems for video technology*.

[8] F. Peng, X.-w. Zhu, and M. Long, "An ROI Privacy Protection Scheme for H. 264 Video based on FMO and Chaos," *IEEE transactions on information forensics and security*, vol. 8, no. 10, pp. 1688–1699, 2013.

[9] F. Dufaux and T. Ebrahimi, "Scrambling for Privacy Protection in Video Surveillance Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1168–1174, 2008.

[10] M. Ouamri and K. M. Faraoun, "Robust and fast selective encryption for hevc videos," 2014.

[11] E. M. Newton, L. Sweeney, and B. Malin, "Preserving Privacy by De-identifying Face Images," *IEEE transactions on Knowledge and Data Engineering*, vol. 17, no. 2, pp. 232–243, 2005.

[12] H. Sohn, E. T. AnzaKu, W. De Neve, Y. M. Ro, and K. N. Plataniotis, "Privacy Protection in Video Surveillance Systems using Scalable Video Coding," in *Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*. IEEE, 2009, pp. 424–429.

[13] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth, and M. Zhou, "An Overview of Tiles in HEVC," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969–977, 2013.

[14] M. Farajallah, W. Hamidouche, O. Déforges, and S. El Assad, "ROI Encryption for the HEVC Coded Video Contents," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 3096–3100.

[15] N. Sidaty, W. Hamidouche, and O. Deforges, "A New Perceptual Assessment Methodology for selective HEVC Video Encryption," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017.

[16] M. A. Taha, S. El Assad, O. Jallouli, A. Queudet, and O. Déforges, "Design of a Pseudo-Chaotic Number Generator as a Random Number Generator," in *The 11th International Conference on Communications*, 2016, pp. 401 – 404.

[17] "Kvazaar HEVC Encoder." [Online]. Available: https://github.com/ultravideo/kvazaar

[18] M. Viitanen, A. Koivula, A. Lemmetti, A. Ylä-Outinen, J. Vanne, and T. D. Hämäläinen, "Kvazaar: Open-Source HEVC/H. 265 Encoder," in *Proceedings of the 2016 ACM on Multimedia Conference*. ACM, 2016, pp. 1179–1182.

[19] "HEVC Decoder ." [Online]. Available: https://github.com/OpenHEVC/openHEVC

[20] Z. Shahid and W. Puech, "Visual Protection of HEVC Video by Selective Encryption of CABAC Binstrings," *ieee transactions on multimedia*, vol. 16, no. 1, pp. 24–36, 2014.

[21] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "HEVC Complexity and Implementation Analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1685–1696, 2012.

[22] G. Bjøntegaard, "VCEG-M33: Calculation of Average PSNR Differences Between RD-Curves," Apr. 2001.

ISBN 978-90-827970-1-5 © EURASIP 2018