SELECTIVE VIDEO ENCRYPTION USING CHAOTIC SYSTEM IN THE SHVC EXTENSION

W. Hamidouche, M. Farajallah, M. Raulet, O. Déforges and S. El Assad

IETR Lab CNRS 6164, France

ABSTRACT

In this paper we investigate a selective video encryption in the scalable HEVC extension (SHVC). The SHVC extension encodes the video in several layers corresponding to different spatial and quality representations of the video. We propose a selective encryption solution using a chaotic-based encryption system. The proposed solution encrypts a set of sensitive parameters with a minimum complexity overhead, at constant bitrate and SHVC format compliant.

Experimental results compare the performance of three encryption schemes: encrypt only the lowest layer, all layers, and only the highest layer. The first two schemes achieve a high security level with a drastic degradation in the decoded video, while the last scheme enables a perceptual video encryption by decreasing the quality of the highest layer below the quality of the clear layers.

Index Terms— Selective video encryption, SHVC.

1. INTRODUCTION

Nowadays, the most transferred content over Internet is multimedia data including still image and video. The new video coding standard High Efficiency Video Coding (HEVC) [1] allows up to 50% gain in terms of subjective video quality with respect to the H.264/AVC high profile [2]. Moreover, the scalable extension of the HEVC standard (SHVC) defines tools to enable spatial and quality (SNR) scalability [3]. In the up-coming years, HEVC standard and its scalable extension are expected to be progressively adopted in industry with the perspective to replace the predecessor video compression standard. On the other hand, security and confidentiality of multimedia contents become a challenging research topic, which was widely investigated in the last decade [4, 5]. The most straightforward method to secure video content is to encrypt the whole bitstream using standard encryption algorithms such as Advanced Encryption Standard (AES) . This method called Naive Encryption Algorithm (NEA) treats the video bitstream as text data without considering the structure of the compressed video [5]. However, NEA suffers from several drawbacks. First, the encryption/decryption process becomes time and energy consuming for large scale-data especially video at high resolution (4K and 8K) and high bitrate. Therefore, NEA is not suitable for real time video transmission applications, which have rigid restrictions on delay and energy on mobile devices. Second, NEA prevents untrusted middle-box in the network to perform post-processing operations on the encrypted video bitstream such as transcoding and watermarking. Selective video encryption has emerged as an effective alternative to NEA [6, 7]. Selective video encryption considers the coding structure of the compressed video and encrypts only the most sensitive information in the video bitstream. Authors in [6] studied the impact in terms of both video quality and bitrate of all encryptable parameters in the HEVC bitstream. The encryption of a set of parameters including Transform Coefficients (TCs), TC sign, Motion Vector (MV) difference, MV difference sign and delta Quantization Parameter (dQP), enables a high degradation in the video quality with a minimum increase in bitrate. Shahid et al. [7] proposed a selective encryption solution for the HEVC video with a constant bitrate. The proposed solution encrypts TCs, TC sign, MVs difference and MV sign. The encryption is performed at the level of the Context-Adaptive Binary Arithmetic Coding (CABAC) bin-string (ie. after the binarization process of the CABAC). The binarization of the TCs is performed in the HEVC draft 6 with a combination of Truncated Rice code with an adaptive context p (TRp) and kth-order Exp-Golomb (EGk) code with $k = 0$ (EG0). Authors in [7] proposed an algorithm to encrypt the suffix of only residuals that do not impact the adaptive parameter p after encryption, which fulfills constant bitrate and format compliant encryption requirements. In [7] AES algorithm is used as block cipher to encrypt the HEVC bitstream. Thus, the encryption introduces a delay at both encoder and decoder to prepare the plain-text of 128 bits. This delay is not convenient with the architecture of real time encoding/decoding which have a rigid restrictions on delay and memory [8]. In this paper we investigate a real time selective encryption of the SHVC video. We propose an encryption solution that fulfills the three following requirements: 1) Format compliant encryption: the encrypted bitstream remains compliant with the conforming SHVC syntax. 2) Constant bitrate: the encryption algorithm does not affect the SHVC compression ratio. 3) Secured and fast encryption: achieve a high security level with a minimum additional delay and complexity. The proposed encryption

This work is supported by the European Celtic-Plus project 4KREPROSYS - 4K ultraHD TV wireless REmote PROduction SYStems -

solution uses a chaotic encryption system [9, 10]. Chaotic-based encryption systems are more flexible and modular, and thus more suitable for large scale-data encryption [11]. We also propose a fast algorithm to determine the encryptible bins after the binarization of the TCs in TRp and EG-k ($k = p+1$) codes while satisfying constant bitrate and format compliant requirements. The rest of this paper is organized as follows. Section 2 provides the principles of the SHVC extension with details on the arithmetic encoder and the used chaotic generator. The proposed selective SHVC encryption schemes are described in section 3. The performance of the proposed encryption schemes is assessed and discussed in section 4. Finally, section 5 concludes this paper.

2. SYSTEM DESCRIPTION

2.1. SHVC video extension

The SHVC extension is defined to provide spatial and fidelity scalability with a simple and efficient coding architecture [12]. All technologies defined in the HEVC standard are used in SHVC including accurate Intra/Inter predictions and highly adaptive entropy coding [1]. Moreover, HEVC standard uses the concept of dQP to adapt the QP value at the coding unit level for visual quality optimization and rate control. SHVC extension adopts an inter-layer prediction to take advantage of spatial correlation and improve the rate-distortion performance compared to independently encoding of the layers. The SHVC encoder consists of L HEVC encoders, one encoder to encode each layer with L the number of layers: one Base Layer (BL) and $L-1$ Enhancement Layers (EL). In the case of spatial scalability, the BL HEVC encoder encodes a down-sampled version of the original video and feeds the first EL encoder with the decoded picture and its MVs. The EL encoder encodes a higher resolution video with using the decoded picture from lower layer as an additional reference picture. The inter-layer reference picture is up-sampled and its MVs up-scaled to match with the resolution of the EL layer being decoded. Figure 1 shows an example of the SHVC encoder encoding two layers in spatial scalability configuration.

2.2. CABAC binarization in SHVC

The CABAC engine defined in HEVC remains unchanged at each SHVC layer (see Figure 1). The CABAC engine consists of three main functions: binarization, context modeling and arithmetic coding [13]. First, the binarization step corresponds syntax elements to binary symbols (bin). Second, the context modeling updates the probabilities of bins, and finally the arithmetic coding compresses the bins into bits according to the estimated probabilities. Five binarization methods are used in HEVC namely Unary (U), Truncated Unary (TU), Fixed Length (FL), TRp and EGk codes. The U code represents an unsigned integer Y with a bin string of length $Y+1$

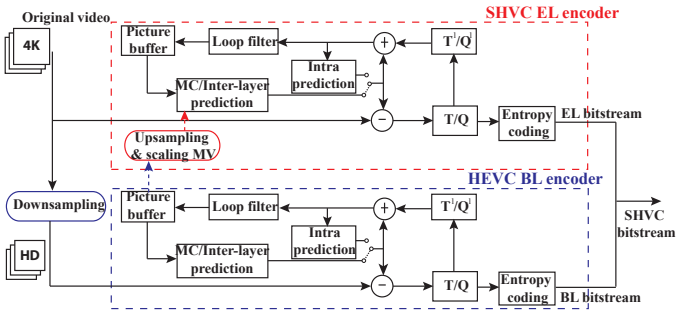


Fig. 1. Block diagram of the SHVC encoder with two layers

composed of Y 1-bins followed by one 0-bin. The TU code is defined with the largest possible value of the syntax element $cMax$ ($0 \leq Y \leq cMax$). When the syntax element value $Y < cMax$, the TU is equivalent to U code, otherwise Y is represented by a bin string of $cMax$ 1-bins. The FL code represents a syntax element Y with its binary representation of length $\lceil \log_2(cMax+1) \rceil$. The TRp code is a concatenation of a quotient $q = \lfloor Y/2^p \rfloor$ and a remainder $r = Y - q2^p$. The quotient q is first represented by TU code as a prefix concatenated with a suffix r represented by the FL code of length p. The EGk code is also a concatenation of prefix and suffix. The prefix part of the EGk code is the U representation of $l(Y) = \lfloor \log_2(\frac{Y}{2k}+1) \rfloor$. The suffix part is the FL code of $Y + 2^k(1-2^{l(Y)})$ with $cMax = k + l(Y)$.
The arithmetic coder can be performed either by an estimated probability of a syntax element (context coded) or by considering equal probabilities of 0.5 (bypass coded).

2.3. Low complexity chaotic generator

In chaotic systems, periodicity of the generated sequence trajectory is a dangerous weakness that must be avoided. We use the chaotic generator of a discrete chaotic sequence implemented based on El Assad et al. patent [10]. It consists in two chaotic maps: Skew Tent Map (STM) expressed by Equation 1 [14], and Discrete Piece-wise Linear Chaotic Map (PWLCM) expressed in Equation 2 [15]. These two maps are connected in parallel and the generated value is the xor or addition between the output of these maps depending on either their values are equal or different, respectively. Each generator is perturbed using a linear feedback shift register. This ensures a very large periodicity for all generated sequences. P_1 and P_2 are the control parameters and are ranging from 1 to 2^N-1 and $2^{N-1}-1$, respectively. $N = 32$ bits is the finite precision. It should be noted that the considered chaotic generator has passed all the tests including those of the national institute of standards and technology [16].

$$X_1[n] = \begin{cases} \left\lfloor \frac{2^N \times X_1[n-1]}{P_1} \right\rfloor & \text{if } 0 < X_1[n-1] < P_1 \\ 2^N - 1 & \text{if } X_1[n-1] = P_1 \\ \left\lfloor \frac{2^N \times (2^N - X_1[n-1])}{2^N - P_1} \right\rfloor & \text{if } P_1 < X_1[n-1] < 2^N \end{cases} \quad (1)$$

$$X_2[n] = \begin{cases} \left\lfloor \frac{2^N \times X_2[n-1]}{P_2} \right\rfloor & \text{if } 0 < X_2[n-1] < P_2 \\ \left\lfloor \frac{2^N \times (X_2[n-1]-P_2)}{2^{N-1}-P_2} \right\rfloor & \text{if } P_2 < X_2[n-1] < 2^{N-1} \\ \left\lfloor \frac{2^N \times (2^N - X_2[n-1]-P_2)}{2^{N-1}-P_2} \right\rfloor & \text{if } 2^{N-1} \leq X_2[n-1] < 2^N - P_2 \\ \left\lfloor \frac{2^N \times (2^N - X_2[n-1])}{P_2} \right\rfloor & \text{if } 2^N - P_2 \leq X_2[n-1] < 2^N - 1 \\ 2^N - 1 & \text{otherwise} \end{cases} \quad (2)$$

3. SHVC SELECTIVE ENCRYPTION SCHEMES

In this section we propose a selective solution to encrypt the SHVC bitstream. This solution encrypts the SHVC bitstream in three different configurations (called here schemes). The first scheme (SE-SHVC-BL) encrypts only the bitstream of the BL. This scheme will also affects the quality of the ELs since the decoded BL picture and its MVs are used as reference for inter-layer prediction at the EL encoders. The second scheme encrypts the bitstream of all SHVC layers (SE-SHVC-All). Therefore, these two schemes will achieve a high secure encryption since in addition to the encryption of the BL, all ELs are implicitly or explicitly encrypted in schemes SE-SHVC-BL and SE-SHVC-All, respectively. The third scheme encrypts only the highest EL (SE-SHVC-EL). Thus, the low quality of the video remains clear and only end-users holding the secret key can visual a high quality of the video. The encryption solution is similar for all SHVC layers and is described in the next sections.

3.1. Encryption space

The proposed encryption solution is SHVC format compliant and does not affect the compression ratio of the SHVC encoder. Therefore, only syntax elements binarized in FL code and then bypass coded can be safely encrypted. The absolute value of MV differences minus 2 is binarization in EG1 code and then bypass coded. Thus, the suffix part of MV differences is encrypted without impacting the compression ratio. The signs of MV difference and TCs are also encrypted since they are binarized in FL code with $cMax = 1$ and bypassed. The absolute value of the dQP is context coded so its encryption will affect its probability and the compression ratio. We propose to encrypt only the dQP sign which is bypassed in the SHVC CABAC. Concerning the TCs, they are bypassed and binarized with a combination of TRp with $p \in \{0, 1, 2, 3, 4\}$ and EGk codes ($k = p + 1$). The suffix of the EGk code can be safely encrypted, while the encryption of the TRp suffix is not format compliant since the encryption can affect the p parameter value and consequently the compression ratio. In this paper we propose an algorithm enabling to accurately determine the bins of the TRp suffix that can be encrypted without changing the update of p value before and after all possible encryption values of the TRp suffix. The absolute value of TC ($b_i \neq 0$) is composed of the base level noted $baseLevel \in \{1, 2, 3\}$ plus the remaining part noted Rem ($b_i = baseLevel + Rem$). The base level value is first signaled in the bitstream with a specific syntax elements and only remaining part Rem different from 0 is binarized in TRp and EGk codes. Algorithm 3.1 provides the positions of the encryptable bins in the TRp suffix.

Algorithm 3.1 Encryptable bins in the TRp suffix of TCs

if ($baseLevel == 1$) **then**
  The whole suffix is encryptable.
else if ($p == 1$) **then**
  if ($baseLevel == 2$ AND ($Rem == 4$ OR $Rem == 5$)) **then**
    No encryption.
  else
    The whole suffix is encryptable.
  end if
else if ($p == 2$) **then**
  if ($Rem \leq 7$ OR $Rem \geq 12$) **then**
    The whole suffix is encryptable.
  else if ($baseLevel == 2$ AND ($Rem == 10$ OR $Rem == 11$)) **then**
    No encryption.
  else
    The first bin of the suffix is encryptable.
  end if
else if ($p == 3$) **then**
  if ($Rem \leq 15$ OR $Rem \geq 24$) **then**
    The whole suffix is encryptable.
  else if ($Rem \leq 19$) **then**
    The first two bins of the suffix are encryptable.
  else if ($baseLevel == 2$ AND ($Rem == 22$ OR $Rem == 23$)) **then**
    No encryption.
  else
    The first bin of the suffix is encryptable.
  end if
else if ($p == 4$) **then**
  if ($Rem \leq 31$ OR $Rem \geq 48$) **then**
    The whole suffix is encryptable.
  else if ($Rem \leq 39$) **then**
    The first three bins of the suffix are encryptable.
  else if ($Rem \leq 43$) **then**
    The first two bins of the suffix are encryptable.
  else if ($baseLevel == 2$ AND ($Rem == 46$ OR $Rem == 47$)) **then**
    No encryption.
  else
    The first bin of the suffix is encryptable.
  end if
end if

3.2. Chaotic encryption process

The encryption process is performed on the fly syntax element by syntax element with using a simple xor and addition operations:

$$c_i = r_i \oplus (k_i + c_{i-1}) \quad (3)$$

where r_i is the encryptable bins of one syntax element (plain), c_{i-1} is the previous encrypted value, k_i is the generated bits from the chaotic generator (dynamic key) and $c_0 = r_0 \oplus (k_0 + IV)$ with IV the Initial Vector. The simple $\oplus$ and $+$ operators are used to produce the confusion and diffusion effects

Class	Scal.	Orig.	SE-SHVC-BL		SE-SHVC-All		SE-SHVC-EL	
		PSNR Y	PSNR Y	ES(%)	PSNR Y	ES(%)	PSNR Y	ES(%)
A	SNR	41.12	8.28	7.27	8.25	15.06	23.69	7.79
	2x	41	9.04	5.62	8.96	16.61	17.72	10.98
	HEVC	41.25	-	-	-	-	8.55	17.83
B	SNR	39.54	9.18	6.23	9.13	15.72	25.77	9.49
	2x	39.6	9.82	4.11	9.66	17	18.65	12.89
	1.5x	39.57	9.11	6.31	9.03	16.41	21.97	10.1
	HEVC	39.6	-	-	-	-	9.29	18.27

Table 1. Video quality and ES analysis at $QP_{EL} = 22$

whereas the generated chaotic pseudo-random dynamic key (k_i) has the same length than the current plain parameter (r_i). At the decoder side, the decryption is performed as follows:

$$r_i = c_i \oplus (k_i + c_{i-1}) \tag{4}$$

where $r_0 = c_0 \oplus (k_0 + IV)$. We consider one chaotic generator by layer to perform an independent encryption of the SHVC layers and enables a correct decryption even when previous EL frames are not decoded. At each layer, the SHVC encoder and decoder must have the same secret key used to initialize the chaotic generator.

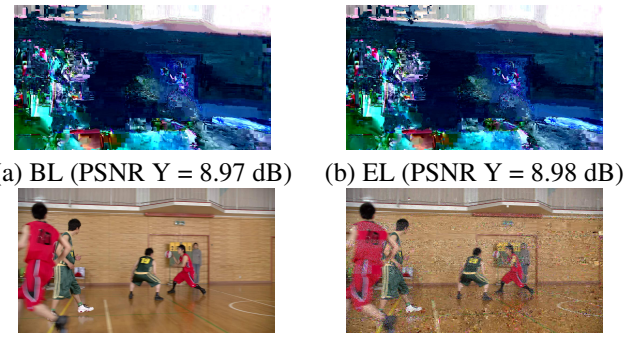
4. RESULTS AND ANALYSIS

4.1. Experimental configuration

The proposed encryption schemes were implemented in the Scalable Reference software Model (SHM) version 4.1 [17]. We consider the common SHVC test conditions including two 2560×1600 video sequences (Class A) and five 1920×1080 video sequences (Class B) [18]. These video sequences are encoded in low delay P configuration (I frame followed by P frames), two layers ($L = 2$) and three scalability configurations: 2x, 1.5x and SNR. We use the Peak Signal to Noise Ration (PSNR) to assess the quality of the decoded video.

4.2. Results

Table 1 gives the average performance in terms of PSNR and Encryption Space (ES: the ration of chipred bits in the SHVC bitstream) of the three proposed encryption schemes for video classes A and B at one particular QP configuration. The PSNR is drastically decreased by schemes SE-SHVC-BL and SE-SHVC-All (PSNR Y lower than 10 dB), encrypting only the BL and all layers, respectively. Moreover, the SE-SHVC-BL scheme encrypts less than 7.5% of the whole SHVC video bitstream in the three scalability configurations. Figure 2 shows the visual quality of the 9^{th} frame of *BasketballDrive* video sequence encrypted by schemes SE-SHVC-BL and SE-SHVC-EL. We can notice that SE-SHVC-BL scheme encrypting only the BL also affects the quality of the EL. The inter-layer prediction using the decoded BL picture and its MVs propagates the errors to the EL. However, the



(a) BL (PSNR Y = 8.97 dB) (b) EL (PSNR Y = 8.98 dB)

(c) BL (PSNR Y = 39.35 dB) (d) EL (PSNR Y= 17.59 dB)

Fig. 2. Visual quality of frame #9 of the $BasketballDrive$ video sequence in SNR scalability configuration ((a) and (b) SE-SHVC-BL, (c) and (d) SE-SHVC-EL)

SE-SHVC-EL scheme by encrypting only the EL, the BL remains clear and the quality of the EL is slightly decreased compared to the quality in the SE-SHVC-BL scheme. This is because most parts of the information is predicted from the BL and only details (encrypted data) are encoded at the EL level. Table 2 gives the Encryption Quality (EQ) and Hamming Distance (HD) of the three proposed schemes in different scalability configurations for *Traffic* video sequence. The EQ calculates the average frequency difference between all possible bytes in the original and the encrypted video frames. To affirm the high security level of the proposed schemes, HD is calculated between the plain and the cipher frames of the *Traffic* video sequence. The HD in bits between the plain and the ciphered frames P and C should be close to 50% (probability of bit changes [19]). Thus, the plain-text sensitivity attack would become an useless attacking method. The obtained results in Table 2 prove the robustness of our proposed selective encryption solution.

5. CONCLUSION

In this paper we have investigated a selective video encryption in the SHVC extension. We proposed a low complexity selective video encryption solution based on a chaotic encryption system. The proposed solution encrypts a set of SHVC parameters at constant bitrate and SHVC format compliant. Experimental results showed that encrypting only the BL achieves a high security level on all SHVC layers with less than 7.5% encrypted data. Moreover, encrypting only the EL enables a perceptual video encryption by decreasing the quality of the highest layer below the quality of other layers.

Schemes	EQ		HD	
	2x	SNR	2x	SNR
SE-SHVC-BL	10942	15499	0.48	0.51
SE-SHVC-All	11056	15439	0.48	0.51
SE-SHVC-EL	2880	1561	0.37	0.32

Table 2. Encryption Quality for Traffic video sequence

6. REFERENCES

[1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, pp. 1648–1667, December 2012.

[2] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparaison 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, pp. 1969–1684, December 2012.

[3] C. C. Chi, M. Alvarez-Mesa, B. Juurlink, G. Clare, F. Henry, S. Pateux, and T. Schier, "Parallel Scalability and Efficiency of HEVC Parallelization Approaches," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, pp. 1827–1838, December 2012.

[4] I. Agi and L. Gong, "An Empirical Study of Secure MPEG Video Transmissions," in *Network and Distributed System Security*, February 1996, pp. 201–210.

[5] L. Qiao and K. Nahrstedt, "Comparaison of MPEG Encryption Algorithms," *Data Security in Image Communication and Networking*, vol. 22, pp. 437–448, December 1998.

[6] Glenn Van Wallendael, Andras Boho, Jan De Cock, Adrian Munteanu, and Rik Van de Walle, "Encryption for High Efficiency Video Coding with Video Adaptation Capabilities," *IEEE Transactions on Consumer Electronics*, vol. 59, pp. 634 – 642, August 2013.

[7] Zafar Shahid and William Puech, "Visual Protection of HEVC Video by Selective Encryption of CABAC Binstrings," *IEEE Transactions on Multimedia*, vol. 16, pp. 24 – 36, January 2014.

[8] Wassim Hamidouche, Mickael Raulet, and Olivier Deforges, "Real time SHVC Decoder: Implementation and Complexity Analysis," in *IEEE International Conference on Image Processing (ICIP)*, October 2014.

[9] Jiri Fridrich, "Symmetric Ciphers based on Two-dimensional Chaotic Maps," *International World Scientific Journal on Bifurcation and Chaos*, vol. 8, no. 06, pp. 1259–1284, 1998.

[10] Safwan El Assad and Hassan Noura, "Generator of Chaotic Sequences and Corresponding Generating System," July 15 2014, US Patent 8,781,116.

[11] Mousa Farajallah, Z Fawaz, Safwan El Assad, and Olivier Dforges, "Efficient Image Encryption and Authentication Scheme based on Chaotic Sequences," in *International Conference on Emerging Security Information, Systems and Technologies*. Barchalona, Spain, August 2013, pp. 150–155.

[12] ISO/IEC-JTC1/SC29/WG11 and ITU-T-SG16, "Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC)," in *SO/IEC JTC1/SC29/WG11 (MPEG) Doc. N12957 or ITU-TSG 16 Doc. VCEG-AS90*, July 2012.

[13] Vivienne Sze and Madhukar Budagavi, "High Throughput CABAC Entropy Coding in HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, pp. 1778–1791, December 2012.

[14] NAKKL Masuda, Goce Jakimoski, Kazuyuki Aihara, and Ljupco Kocarev, "Chaotic Block Ciphers: from Theory to Practical Algorithms," *IEEE Transactions on Circuits and Systems I: Regular Papers,*, vol. 53, no. 6, pp. 1341–1352, 2006.

[15] Shiguo Lian, Jinsheng Sun, Jinwei Wang, and Zhiquan Wang, "A Chaotic Stream Cipher and the usage in Video Protection," *International Elsevier journal on Chaos, Solitons and Fractals*, vol. 34, no. 3, pp. 851–859, 2007.

[16] Rukhin Andrew, Soto Juan, Nechvatal James, Smid Miles, and Barker Elaine, "A statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications," Tech. Rep., DTIC Document, 2001.

[17] "SHVC Reference software model (SHM)," in *https://hevc.hhi.fraunhofer.de/svn/svn_SHVCSoftware/*.

[18] V. Seregin and Y. He, "Common SHM test conditions and software reference configurations," in *document JCTVC-O1009*. Geneva, Switzerland, November 2013.

[19] Xingyuan Wang, Dapeng Luan, and Xuemei Bao, "Cryptanalysis of an image encryption algorithm using Chebyshev generator," *Digital Signal Processing (Elsevier)*, vol. 25, pp. 244–247, 2014.