Sufficient Encryption with Codewords and Bin-strings of H.264/SVC

Mamoona Naveed Asghar
School of Computer Science and Electronic Engineering
University of Essex, Colchester, United Kingdom CO4 3SQ

Mohammad Ghanbari,
Fellow IEEE
School of Computer Science and Electronic Engineering
University of Essex, Colchester, United Kingdom CO4 3SQ

Martin J. Reed
School of Computer Science and Electronic Engineering
University of Essex, Colchester, United Kingdom CO4 3SQ

***Abstract*—Multi-layered scalable bit-stream distribution requires protection for the rendering and transmission of its individual layers. This paper presents a sufficient encryption (SE) scheme for SVC layers which maintains the compression efficiency and the format compliancy of scalable bit-streams, without compromising the security. The purpose of SE is achieved by applying the partial encryption on carefully selected codewords and bin-strings of the CAVLC and CABAC of H.264/SVC respectively. The performance of the scheme is tested on various resolution sequences, which demonstrate the advantages of the scheme when compared to alternative techniques. These advantages include: minimal computational delay by encrypting partial data; no bit-rate escalation by keeping the compression ratio unchanged; and, format compliancy of the bit-stream at the decoder. A comparative security analysis of the scheme confirms that it is suitable for commercial, real-time applications. The minimal increase in processing requirements ensures that the scheme is highly suitable for video distribution to users who have subscribed to differing video qualities on end systems ranging from small handheld devices to those with high computational capability.**

Keywords-AES-CFB, CABAC, CAVLC, H.264/SVC, sufficient encryption

I. INTRODUCTION

Multimedia streaming servers entertain the needs of various heterogeneous devices simultaneously according to device subscribed bandwidth and constraints like picture resolution, video quality and the processing power. The emerging scalable video technology has the capability to fulfill the objective of variable frame rate, resolution and fidelity, thus meeting the needs of varying resolutions of multiple receiving media with different bandwidth characteristics. Distribution of copyrighted digital content over heterogeneous networks is always susceptible to plagiarism attacks by its ease of copying and modification. Therefore, the concerns that exist about their protection are significant. This paper investigates some important issues in the protection of scalable video distribution by presenting a partial encryption scheme for the H.264/Scalable Video Coding (SVC) codec [1], an extension of the state-*of-the*-art H.264/AVC standard [2].

Video content security has been extensively investigated in the past although most of the schemes are based on naïve or partial encryption of the video content. Naïve encryption, also called blind/confidential encryption, requires more computational cost. So it is suitable only for highly confidential defense/military and law enforcement applications. Applications like video storage and real-time video conferencing demand medium level security which is attainable by partial encryption of data. Partial encryption, also referred to as the *sufficient encryption* in digital rights management (DRM) [3], encrypts the selective most influential video data to make the video viewable in a scrambled form rather than be watchable.

To improve the run-time performance and to avoid the computational overhead, partial/selective video content encryption is proposed in this paper. The partial encryption is carried out on the most significant information at different stages of the codec, such as on the original pixels, the transform coefficients, the quantization indexes, the bit-planes, the entropy coder, or the final bit-stream [4]. Encryption alters the video statistics, resulting in the issues of bit-rate overhead and format compliancy at the decoder. To minimize this effect of encryption, it can be applied as part of the entropy coding where all the natural redundancies have already been exploited for maximum compression efficiency. The problem with joint security and compression is to make sure that the security will not affect the compression efficiency [5]. Therefore, entropy coding still needs to be handled with great care, since tampering with the statistical dependency of the symbols, also harms the compression efficiency. H.264/SVC has been developed from H.264/AVC and thus uses the same entropy coding modes: variable length coding (VLC) and binary arithmetic coding (BAC).Both of these modes are used in a context adaptive (CA) manner, leading to the terms CAVLC and CABAC.

We have proposed a sufficient encryption scheme for H.264 scalable layers through the Advanced Encryption Standard in a cipher feedback mode (AES-CFB). The proposed SE is applied separately on the carefully selected variable length codewords of the CAVLC and the bin-strings of the CABAC. The carefully selected codewords

and bin-strings have no affect on the compression ratio and the format compliancy of the scalable bit-streams whilst maintaining medium level security of proposed scheme against attacks. The reason for employing SE is that the scalable content can reach the user in a viewable form, which enhances the user interest in the subscription of upper high quality video layers.

The remainder of this paper is organized as follows. Section II, reviews the prior research in the area of H.264 protection. Section III describes the H.264/SVC codec with CAVLC and CABAC. Section IVdescribesthe proposed SE scheme with the choice of syntax elements for SVC. Section V outlines the validation tests that have been performed, while Section VI makes some concluding remarks.

II. Prior Research on H.264 Standard

Many multimedia selective content encryption methods have been proposed for the security of state-*of-the*-art H.264 standard. Fan et al., [6][7] presented a novel video encryption scheme for H.264/AVC with naïve and partial encryption. Three different block/stream cipher algorithms AES, FLEX (Fast Leak Extraction) and XOR methods are used to encrypt the H.264/AVC stream. The algorithms encrypted the parameters such as motion vector differences (MVD), inter and intra MBs but the proposed security levels increase the bit-rate while the MVD magnitude encryption causes drift and the stream is not decodable at some stages. The proposed study [8] applied the XOR encryption on signs of coefficients, signs of MVD and alteration of DC values, but does not consider using an entropy coder for H.264/SVC. The DC value alterations also change the video statistics and affect the compression ratio.

The research over CAVLC codeword encryption for H.264/AVC is more common in the literature due to its usage in baseline profile for low processing mobile devices. Bergeron et. al., [9] presented a method for the CAVLC codewords of H.264/AVC; the scheme is not compression efficient, as the whole codeword encryption causes bit-rate fluctuation. Wang et. al., [10] proposed a scrambling method for the signs of the trailing ones (T1s) and the MVD for H.264/AVC. The scheme is implemented by a hardware oriented stream cipher Decim v2 [11] with 80-bit secret key. The solution compromises the security by encrypting just a single sign bit. Wang et. al., [12] presented the partial encryption scheme on the codewords of 4×4 and 16×16 intra prediction mode (IPM), Exponential-Golomb for MVD and level_suffix by using RC4 stream cipher. The authors of [13][14] proposed the encryption on the signs of the T1s and the non-zeros (NZ) of the transform coefficients (TC) on a limited number of MBs for each frame in the intra mode of CAVLC. The scheme is implemented by AES-CFB algorithm on selective MBs by using the prediction error. The scheme maintains the compression efficiency for I-frames but does not provide perceptual security. Sohn et. al., [15] implemented the encryption on CAVLC codewords of H.264/SVC. The region of interest (ROI) is selectively encrypted to aid the video surveillance systems for scalable video coding. The method comprises of random sign bit inversion of NZ levels by using XOR method with multiple keys. However the video can be recovered by the brute force attack on sign bits.

CABAC based encryption is not widely recorded in the literature because of its complex structure and the limitation that it cannot be used in the baseline profile for SVC base layer. Lee et. al., [16] proposed a scrambling method for CABAC used in H.264/AVC. The scheme combines the manipulations of CABAC initialization table and integer DCT coefficients of 4×4 and 8×8 sized sub-macroblocks, resulting in bit-rate overhead because of the change in the context model. The study in [17] investigates scalable layer protection with individual layer keys and applies encryption over signs of texture, MVD and FGS. The same selective parameters for encryption are extended for the protection of ROI using CABAC [18]. The technique applied in [19] used the selective encryption on both the codewords and the bin-strings of H.264/AVC. The scheme only encrypts the residual data, which makes the security of video content questionable. Li, et al. [20][21] devised selective encryption technique for H.264/SVC on both entropy coders using a network abstraction layer (NAL). The scheme encrypts the instantaneous decoding refresh (IDR) pictures, sequence parameter set (SPS), picture parameter set (PPS) on individual NAL units [20] and IPM with signs of textures for base layer [21] by using the stream cipher Leak Extraction (LEX) algorithm. Park et al. [22] proposed the scheme of partial encryption of H.264/SVC where the IPM, MVD and texture sign bits are encrypted. However, the IPM encryption in studies [12][20][21][22] affect the compression efficiency by negatively changing the video statistics. A low-quality free preview application was developed in [23][24] by performing transparent encryption/conditional access on the H.264/SVC layers. The scalable enhancement layers are encrypted by using AES, while leaving the base layer un-encrypted. However, the authors point out that the enhancement layers are non-format compliant.

III. Overview of H.264/SVC with Cipher Algorithm

A. H.264 Scalable Video Coding

Scalable Video Coding technology permits devices to send and receive multi-layered bit-streams; it allows the transmission and decoding of partial bit-streams to provide video services with different frame rates, spatial resolutions (picture size) and quality (SNR). An SVC bit-stream has a single base layer and number of enhancement layers comprising of various scalable features. The main difference between CAVLC and CABAC is that the MB type, IPMs, MVs, reference picture indexes, residual TC and several other syntax elements are coded adaptively as CABAC bin-strings, whereas with CAVLC only the residual transform coefficients are adaptively coded [25]. Consequently, CAVLC requires less complexity than CABAC during

encoding of SVC layers while CABAC provides an average of 10 to 14% more compression at the cost of greater complexity.

To make entropy coding computationally efficient, both CAVLC and CABAC use the single infinite extent codeword, called Exp-Golomb code [26] to generate the required code for most of the syntax elements [27]. The encoding parameters (except transform coefficients) for both CAVLC and CABAC are first mapped to symbols and the symbols are then Exp-Golomb coded. The Exp-Golomb codes are constructed by concatenation of a prefix and a suffix parts and are working in zero order (*EG0*) and k^{th} order (*EGk*). The CAVLC entropy coder only supports *EG0* codewords, while CABAC supports both orders.

B. Context Adaptive Variable Length Coding

CAVLC [28] is commonly used in the baseline profile of H.264 for small, handheld, devices. In addition to Exp-Golomb codes, it uses variable length codes (VLC) for the encoding of syntax elements. Depending on the syntax of the parameter, Exp-Golomb mapping is done by one of the three methods: (a) unsigned direct mapping *ue(p),* (b) signed direct mapping *se(p),* and (c) mapped symbols *me(p*). The residual data is encoded by the Huffman-derived VLC codewords by using seven fixed variable length look-up tables (LUT). Selection of one of these tables is based on the magnitude (value) of some of the previously coded coefficients. The run-length coding is performed first for encoding the levels (NZ-TC) and runs (zero) separately. For the first NZ-TC, the VLC0/VLC1 table is used according to the conditions specified in the standard [2]. If the magnitude of NZ lies within the range of the table (short codes), it is encoded by regular mode otherwise the long codes are encoded by the escape mode shown in Fig.1.

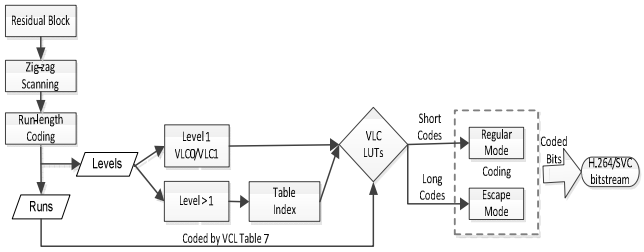
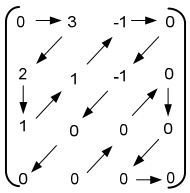


Figure 1. Block Diagram of VLC coding in CAVLC

The CAVLC processes the residual block encoding in five sequential steps: (i) the number of NZ coefficients and T1s *(coeff_token),* (ii) the signs of each trailing ones *(signs of T1s),* (iii) the levels of the remaining non-zero coefficients *(levels),* (iv) the total number of zeros before the last coefficients *(total_zeros),* and (iv) each run of zero *(run_before)*. An example of coefficient encoding is shown in Fig. 2.

C. Context Adaptive Binary Arithmetic Coding

CABAC [29] is one of the entropy coding modes used



Zig-zig Scanning order: 0,3,2,1,1,-1,0,-1,0,0,...
Total NZs (coeff_token) = 6 (3,2,1,1,-1,-1)
levels = 3 (3,2,1)
T1s = 3 (1,-1,-1)
(T1s are 4 but maximum can be 3 only)
total_zeros = 2

Figure 2. Example of 4×4 submacroblock coefficients coded by VLC

by H.264/SVC to achieve high compression and can be easily computed on devices with medium to high computational resources. The H.264 Main profile and the various High profiles, which deal with 4CIF resolution pictures and above, support CABAC. Thus, it appears that the multi-scale video distribution of the future will support CABAC. CABAC encoding (Fig. 3) is based on three steps: i) binarization, ii) context modeling (CM) and iii) binary arithmetic coding (BAC). In binarization the input, non-binary syntax elements, are converted into unique binary codewords, known as bin-strings, for a given syntax element. The bit position in each bin-string, known as a bin, is then passed to one of the two coding mode decisions: regular coding mode and by-pass coding mode. The bins in regular coding mode are passed to the next step, context modeling/probability distribution and then encoded by the regular BAC engine. The bins from the bypass coding mode skip the context modeling step and directly enter in the bypass BAC engine for the encoding process. These bins are related to the sign of MVD information and the signs of transform coefficients levels or for lower significant bins which are assumed to be uniformly distributed.

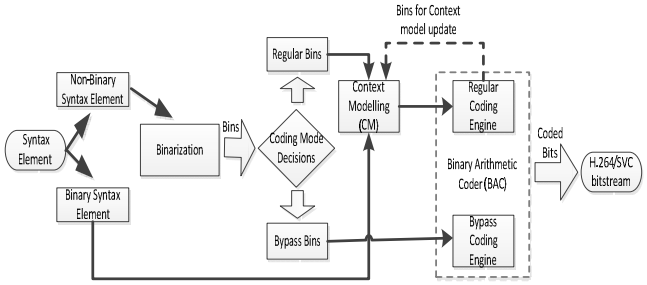


Figure 3. CABAC top view

CABAC uses five binarization schemes according to syntax elements, similar to Huffman trees for binary sequences, which are: (i) unary code, (ii) truncated unary *(TU)* code, (iii) kth order Exp-Golomb *(EGk)* code (iv) fixed-length *(FL)* code and (v) concatenation of the first and third scheme *(UEGk).* There are three situations where concatenations of the four basic types are used: (a) *UFL,* (b) *UEG3, and* (c) *UEG0*.

D. Advanced Encryption Standard

The Advanced Encryption Standard (AES) [30] is based on a modified substitution-permutation network. AES can use the keys of lengths, 128 bits, 192 bits and 256 bits. For both encryption and decryption, the AES algorithm uses a round function that comprises four different byte-oriented

transformations. AES is basically a symmetric key block cipher using a 128-bit block size, but it can be used as stream cipher in Cipher Feedback (CFB), Output Feedback (OFB), and Counter (CTR) modes of operations. The partial encryption requires a smaller number of bits for encryption, so implementing the AES is recommended as a stream cipher [31]. The choice of mode is not critical to the proposed scheme, but to maintain the self-synchronization mechanism for real-time devices, CFB mode is a valid choice. CFB mode uses the same encryption function and encryption key for the decryption process with different sequence of using plain-text and cipher-text block data. In CFB algorithm, the plain-text and cipher-text block data is denoted by $_n$ and $_n$ respectively; where n represents the number of blocks, the encryption and decryption processes can be denoted by the equations below:

$$_n = \ _n XOR \{Encrypt(C_n\text{-}1)\} \quad (1)$$

$$_n = \{Encrypt(C_n\text{-}1)\} XOR \ _n \quad (2)$$

IV. Proposed Sufficient Encryption Scheme

To maintain the characteristics of the SVC bit-streams we have devised an efficient protection mechanism jointly with compression on sensibly selected codewords and bin-strings.

A. Codewords and bin selection for SE

CAVLC and CABAC use multiple parameters for encoding e.g. header information, MB type, inter and intra MB prediction modes, CBP, delta quantization parameters (dQP), MVD and the transform coefficient. The following two constraints are considered in the parameter selection to ensure the effectiveness of proposed scheme.

(1) *Compression efficiency* requires that the SE must not disturb the compression ratio of the encoder by increasing the bit-rate of video file. It can be controlled by keeping the size of the encrypted codewords/bin-string same as the size of input codewords/bin-strings. For CABAC, the condition of keeping the context model unchanged for the given syntax element must also be considered to keep the compression ratio the same.

(2) *Format compliance* means the SE must not change the overall video statistics by changing the absolute values of syntax elements, which would create an error when the SVC encodes the selectively encrypted bit-stream.

To fulfill the above two constraints we can make some recommendations for SE. Some of these recommendations are made on the basis of experimental results (to be described in Section V) while others pertain to the nature of syntax elements. The SE should not be applied on the following codewords and/or bin-strings of CAVLC and CABAC.

- MB type: because it determines intra or inter coding modes of MB, and its encryption results in a non-format compliant bit-stream.
- the intracoded syntax elements (IPM) that are related to neighboring MB syntax elements such as intra DC and AC. This is because it would increase the bit-rate and drift in the values of syntax elements and also the bit-stream will not be decodable at some stage.
- the intercoded syntax elements like MVDs: because this prediction residual is used to predict the future MBs, and alters the video statistics by changing magnitudes. Its use would also increase the bit-rate although the bit-stream can still be decodable.
- delta QP syntax element: because it causes bit-rate fluctuation, either increasing or decreasing the overall bit-rate according to new encrypted dQP values.
- MB header information: because it is used for the prediction of future MBs, this is related to format compliance.
- Coded-block-flag (CBF): because it makes the bit-stream non-format compliant. Every 4×4 block within a MB is encoded if CBP and MBmode are set. The encoded 4×4 block has a CBF syntax element showing the NZs in the current block.

For CAVLC:

- The NZs magnitude codewords: because it disturbs the compression efficiency, resulting in bit-rate overhead.
- The codewords of Total_zeros and run_before: because it causes the bit-rate escalation, and may cause the format compliancy issues on some stage.

For CABAC:

- Unary and truncated unary *(TU)* bin-strings: because they have different codeword lengths and cause the change of bit-rate.
- FL bins: because they have the mandatory header information.

The above discussed syntax elements do not meet the either or both of the compression efficiency or format compliance constraints; therefore, we have eliminated them from our proposed SE scheme. The SE should be applied on the codewords/bin-strings which are equally distributed, since this does not affect the compression efficiency of the codec and in the case of CABAC, they must be encoded by the bypass BAC engine with uniform probability [31]. We found three types of codewords and bin-strings to fulfill our purpose of SE for each method. These are:

Selected codewords for CAVLC

i) Signs of MVD in *se(p)* mapping,
ii) Signs of T1s, and
iii) *EG0* suffix of NZ levels including signs of the NZ levels.

The *se(p)* mapping codeword has separate sign and magnitude bits for MVD. The LSB of the *se(p)* codeword is the sign bit. The signs of the T1s are encoded separately and the *EG0* suffix part (suffixlength) of NZ levels are encoded according to the rules of NZs prefix value [2], while the LSB of suffix length is the sign bit for NZs. The *EG0* suffixes are encrypted both in regular and escape mode of CAVLC.

Selected bin-strings for CABAC

i) *UEG3* suffix,
ii) *UEG0* suffix, and
iii) Signs of the NZ-TC levels.

The *UEG3* suffix consists of the MVD sign bits if two conditions hold i.e. $|MVD| \geq 9$ and $0 < |MVD| < 9$, the sign bits of TC levels and the suffix of *UEG0* can be encrypted only when abs_level > 14. The results (Section V) show that the selected codewords and bins are fully compression friendly, format compliant and do not compromise the security of scalable layers.

B. Codewords and bins selection for SVC Layers

An SVC coded video has a base layer and a number of enhancement layers, depending on how the three scalabilities of temporal, spatial and quality at various resolutions are used. The selected codewords/bins must be compliant with all three scalabilities in SVC coded video, so the bins selection must take into account the specific scalability type. In SVC, every temporal layer requires the change of MVD, dQP and coefficients. Spatial layers require the change of coefficients only and SNR layers require the change of dQP and coefficients. Such SVC layer behavior shows that the coefficient encryption is mandatory for all layers because they change with every scalability option. Hence the codewords of signs of T1s, EG0 suffix and the bin-strings of UEG0 suffix with signs of TC levels are the most suitable parameters for the SVC layer encryption. To keep the sufficient security level of scalable layers, only encrypting the residual data is not enough; the other parameters must be considered to enhance the perceptual security. Therefore the signs of *se(p)* mapping and UEG3 suffix parameters are chosen. Their encryption is also appropriate for temporal scalability which is usually combined with SNR or spatial scalability; hence the MVD sign encryption is meaningful for all three scalabilities. The following equations summarize the relation of syntax elements with scalable characteristics.

C. Codewords and bins encryption by AES-CFB

The SE is implemented on individual SVC coded layers by using AES-CFB, which is a stream cipher; hence it does not alter the number of output bits and also maintains the self-synchronization mechanism between sender and receiver. The CFB mode uses an initialization vector (IV) (fixed in our implementation) and an encryption key (variable) for each data block. The single encryption key will encrypt all three chosen bin-strings on individual layers. So the client will receive only one encryption key to decrypt and watch the subscribed data. Let the three chosen bin-strings be ${}_1B_1$, ${}_2B_2$, ${}_3B_3$ and their encryption represented as ${}_1B_{1,}$ ${}_2B_{2,}$ ${}_3B_3$ then the general cipher process can be presented as:

$$\sum (1B1, \ {}_2B_2, \ {}_3B_3) = \{\sum ({}_1B_1, \ {}_2B_2, \ {}_3B_{3)}\} \ XOR \ \{Encrypt ({}_1B_1\text{-}1, \ {}_2B_2\text{-}1, \ {}_1B_1\text{-}1)\} \quad (3)$$

The encryption process (Fig. 4 and 5) is performed in a unique way which makes the bit-rate consistent. The encryption of the sign bits is straightforward because of the constant codeword/bin sizes. However, the handling of *EG0* suffix parts for both CAVLC and CABAC entropy coder is requires more sophistication as the suffixes have variable length codewords. If the sizes of suffix codewords are changed after encryption, the bit-rate will escalate which

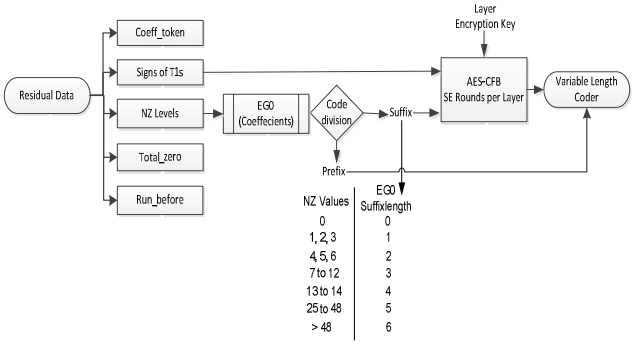


Figure 4. Block Diagram of SE over codewords

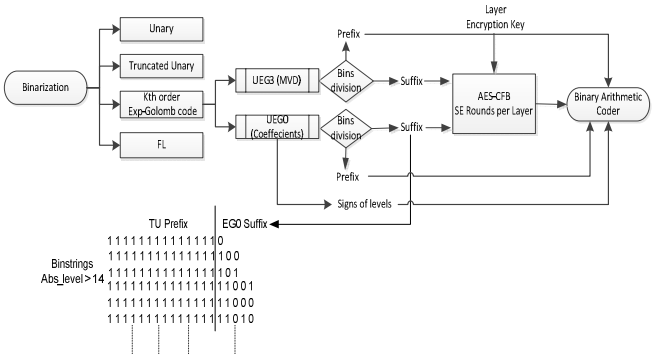


Figure 5. Block Diagram of SE over bin-strings

would not meet the compression efficiency constraint. To make the codewords compression friendly, we first count the suffix bins present in each UEG0 suffix length and then perform the encryption only on the number of existing suffix length fields rather than on the whole suffix. This scheme ensures the encrypted codewords are of the same size as the original codewords.

The decryption process is reversed at the CAVLC/CABAC decoding side. The client is supplied with the same encryption key that was used for enciphering. The encrypted values of the codewords/bin-strings are converted into the original codes/bin values and then passed to the inverse quantization and DCT processes to get the finally de-ciphered and fully decoded bit-stream. The general de-cipher process can be represented as:

$$\sum ({}_1B_1, \ {}_2B_2, \ {}_3B_{3)} = \{Encrypt ({}_1B_1\text{-}1, \ {}_2B_2\text{-}1, \ {}_1B_1\text{-}1)\} \ XOR \ \{\sum ({}_1B_1, \ {}_2B_2, \ {}_3B_3)\} \quad (4)$$

V. EVALUATION

The performance of the proposed SE scheme has been tested with the SVC reference software (Joint Scalable Video Model) JSVM 9.19.10 version encoder. For the evaluation of

results, several test video sequences were chosen with different combinations of features such as motion, colors, objects and texture etc. The experiments were performed on different video resolutions; CIF (352 × 288 pixels/frame) resolutions were used for the SE on CAVLC while 4CIF (704 × 576 pixels/frame) resolution sequences were used for CABAC. Both *CIF* and *4CIF* resolution test sequences were encoded into four layers (one base with three enhancement layers) representing five temporal, two spatial and two SNR scalable levels for both picture resolutions. To aid SVC layer encoding efficiency, CAVLC is usually used for the base layer and CABAC is used for the enhancement layers. To test all the coding modes, the test sequences are encoded with three schemes i.e. all layers are encoded with CAVLC, with CABAC and the combination of both entropy coders. For CABAC all test videos were encoded in a Main/High profile (for base layer encoding) and Baseline profile for enhancement layers. For CAVLC the Baseline profile is used for all layers. The Intra and Inter frames were selectively encrypted in sequence of their occurrence in layers with same GOP size and Intra period i.e. 16. The SE performance is analyzed by taking different QP values, different encoding frame rates and by calculation of computational overhead.

To demonstrate the efficiency of our proposed scheme, we have encoded 90 frames at 30 fps for both *CIF* and *4CIF* resolutions. Table I compares the average PSNR of 90 frames (I+P+B) with and without SE. It shows the suitability of our SE scheme for both Intra and Inter frames of *CIF* and *4CIF* resolution pictures. The average PSNR value of luma is in the lower range for all tested sequences. The impact of SE over sequences is shown in Fig. 6.

TABLE I

COMPARISON OF AVERAGE PSNR (dB) OF 90 FRAMES (I+P+B) AT QP 24

Sequences	Plain Y	SE Y	Plain U	SE U	Plain V	SE V
CIF sequences Encoded with CAVLC						
CREW	38.2	10.9	43.3	23.4	42.6	20.7
MOBILE	35.3	8.8	39.6	15.8	39.1	12.9
MOTHER	41.0	10.5	46.0	20.7	46.9	26.4
4CIF Sequences encoded with CABAC						
CITY	37.1	10.0	44.9	26.8	46.8	29.3
CREW	37.9	11.9	43.0	24.4	44.1	20.9
HARBOUR	37.2	7.1	44.7	23.1	46.2	32.6
4CIF sequences encoded with CAVLC and CABAC						
CREW	38.0	12.7	43.0	24.3	44.2	19.9
ICE	41.2	11.5	48.7	30.2	49.4	25.9
SOCCER	37.1	9.8	44.6	22.1	46.6	22.6

The test sequences were also encoded in the combination of both entropy coders i.e. CAVLC for base layer and CABAC for the three enhancement layers. We performed the experiments at different QP values of 8, 16, 24, 32, 40 and 48 to show the independence of the proposed SE on the QP values for both entropy coders. The graph in Fig. 7 shows the PSNR variance in YUV values with and without SE on the CREW *(4CIF)* sequence, which is encoded with combined CAVLC and CABAC on different QP values. The graph verifies that our SE scheme is independent of QP for both entropy coders, and the average PSNR is still in the lowest range at all QP values.

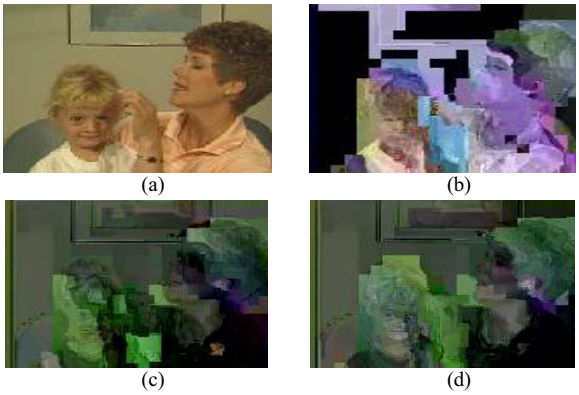


Figure 6. Impact of SE on the PSNR of Mother *(CIF)* sequence encoded with 90 frames (I+P+B) and QP 24. (a) Frame # 70 [Y=41.0, U=46.0, V=46.9] dB. (b) Proposed SE on CAVLC [Y=10.5, U=20.7, V=26.4] dB. (c) Proposed SE on CABAC [Y=13.4, U=25.1, V=24.9] dB. (d) Proposed SE with both CAVLC and CABAC [Y=12.6, U=23.0, V=24.3] dB.

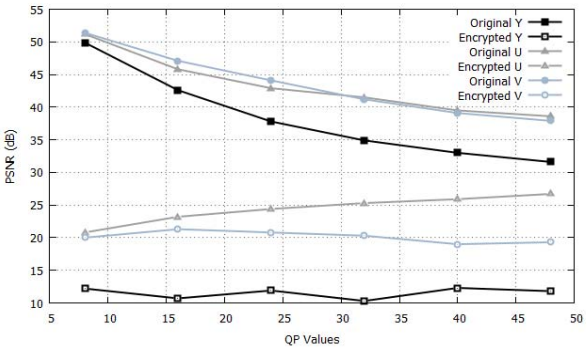
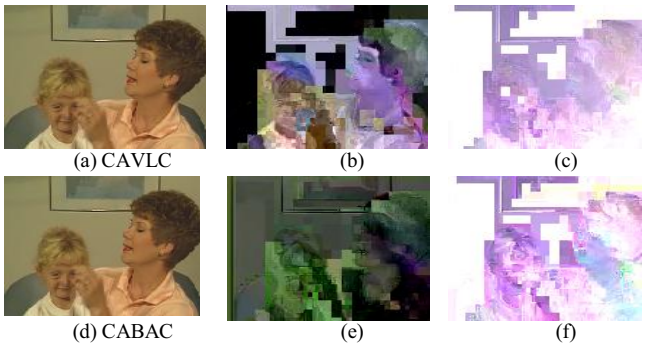


Figure 7. PSNR variance of the *Crew* sequence at different QP values

A. Security Analysis

The robustness of the proposed sufficient encryption scheme against various attacks was also evaluated by the following tests on the bit-streams encoded with the combination of both entropy coders.

i) Known-plaintext attacks: To evaluate the strength of the proposed encryption scheme, we performed experiments on different sequences by guessing the encrypted bits of data on the basis of their known values in plaintext. The sign bits for MVD, T1s, and TC can be 0 or 1 and known to everyone. An attacker can recover them by brute force attack. The NZ-TC suffixes encryption is a significant factor to make the scheme robust to this attack. We substitute the constant bits and determining the PSNR values against such guessing attack. If someone tries to guess and insert the random data, with the intention of improving video quality to make it watchable, they would tend to insert the specific constant bits or random strings of 0 and 1. But the proposed SE scheme is robust against such known-plaintext attacks. As an experiment, in the *Mother (CIF)* video, we replaced the encrypted data with 1's on MVD signs, 0's on T1s, 0's on signs TCs and added a constant integer value of five in the NZ-TC suffixes to get the video. The result is a distorted image for CAVLC and CABAC in Fig. 8(c) and 8(f) respectively.



(a) CAVLC (b) (c)

(d) CABAC (e) (f)

Figure 8. Impact of replacement attack on the *Mother (CIF)* sequence encoded with 90 frames (I+P+B) and QP 24. (a)(d) Frame # 63 [Y=41.0, U=46.0, V=46.9] dB. (b) Proposed SE on CAVLC[Y=10.5, U=20.7, V=26.4] dB. (c) Attack on CAVLC [Y=6.5, U=18.2, V=28.2] dB. (e) Proposed SE on CABAC [Y=13.4, U=25.1, V=24.9] dB. (f) Attack on CABAC [Y=6.5, U=17.0, V=24.3] dB.

ii) Inference attack on encrypted parts: The partial encryption of video, specifically at the entropy coding stage, is a sensitive issue. The attacker may infer the partially encrypted parts from the un-encrypted parts. To analyze this attack, we performed a statistical analysis on partially encrypted video data. An image data distribution can be examined by two statistical measuring parameters, mean (μ) and its standard deviation (σ). The pixels are highly correlated with each other in horizontal, vertical and diagonal directions. So when the encrypted video is decoded, the entropy (data randomness) falls and correlation becomes high because the video frames (texture & edges) are converted into flat regions and produce the artifacts in the image. During SE, pixel values were truncated to a maximum and minimum of 255 and 0 respectively. This causes the spread of dark or very bright colors across the video image. The mean and standard deviation (Table III & IV) was determined for the local neighborhood of each pixel, before averaging across all pixels and all frames of the tested sequence.

TABLE III
μ AND σ OF SE FOR *MOBILE (CIF)* SEQUENCE ENCODED WITH CAVLC WITH 90 FRAMES (I+P+B) ATDIFFERENT QP VALUES

QP Values	μ of Plain file	μ of SE-CAVLC	σ of Plain file	σ of SE-CAVLC
8	125.6199	93.5236	52.0771	66.6339
16	125.6285	100.9666	51.9963	83.6553
24	125.6141	110.6216	51.8272	70.8237
32	125.6508	115.6560	51.4961	62.6046
40	125.6170	110.1820	51.0290	70.9787
48	125.7253	123.7258	48.7148	49.9262

TABLE IV
μ AND σ OF SE FOR *CREW (4CIF)* SEQUENCE ENCODED WITH CABAC WITH 90 FRAMES (I+P+B) AT DIFFERENT QP VALUES

QP Values	μ of Plain file	μ of SE-CABAC	σ of Plain file	σ of SE-CABAC
8	92.8826	37.7784	42.4478	54.5554
16	92.8824	31.4880	42.4378	53.5476
24	92.9298	97.3457	42.3932	45.0803
32	92.9353	25.4347	42.3482	48.7712
40	92.5902	22.4794	42.1049	44.7842
48	92.5778	81.8818	41.8736	42.3048

Note the standard deviation values of the encrypted videos are more than the plain videos for both CAVLC and CABAC encoder, which produces the dark or bright color pictures. The results verify that the luma and chroma values of the whole video are drastically changed by the proposed SE and there is no way to infer the encrypted parts from the un-encrypted parts.

B. *Computaional overhead*

The computational cost is calculated on the basis of the additional processing time required for the encoding and decoding of test sequences with SE on bit-streams encoded with the combination of both entropy coders. The experiments were performed on a machine, Intel Core 2 Duo (3.33GHz) processor with 4GB RAM. Table V shows the encoding and decoding timings of the *ICE (4CIF)* sequence by encoding different numbers of frames to study the suitability of the scheme for real-time transmissions, for example by encoding 10, then 30, up to 90 frames. The additional processing delay (Table V, column no. 4 and 7) is negligible compared to the overall processing delay, verifying the efficiency of the proposed scheme for real-time streaming applications on both the encoder and decoder side.

Table V
THE COMPUTATIONAL OVERHEAD MEASUREMENT (MS) FOR THE *ICE (4CIF)* SEQUENCE WITH CAVLC (BASE LAYER) AND CABAC AT A DIFFERENT NUMBER OF ENCODED FRAMES (I+P+B) AT QP 24

No. of frames	Encoding time without SE	Encoding time with SE	Encoding Delay difference	Decoding time with SE	Decoding time without SE	Decoding Delay difference
10	21975	22042	67	1327.3	1366.6	39.3
30	68472	68597	125	3524.5	3602.8	78.3
50	113936	114185	249	5645.6	5760.9	115.3
70	158965	159356	391	7675.5	7858.1	182.6
90	204994	205531	537	9819.9	10076.3	256.3

C. *Comparison with exisiting wrok*

The brief comparison of our proposed SE schemes with the existing work is performed to verify the efficacy of chosen parameters for partial encryption of H.264 scalable layers data. The existing methods have the disadvantages over compression ratio, format compliancy and/or security of implemented schemes by choosing weak encryption parameters and/or weak cipher algorithms. The only residual data and sign bits encryption in [10][15][17][19] are considered to provide low level perception security, the video can be recovered by brute force attacks. The IPM encryption proposed in [12][21][22] and DC values alteration in [8][16] both change the video statistics in a negative manner. This is because of using absolute values and thus the change in any bit of the absolute value cause the fluctuation in bit-rate.

VI. CONCLUSIONS AND FUTURE WORK

This paper has proposed the sufficient encryption for H.264 scalable layers with sensibly chosen codewords and bin-strings. The security scheme incorporates the partial protection of the scalable layers utilizing AES-CFB. The proposal takes into consideration: the security strength of the scheme, the compression efficiency, the format compliance

and scalability features (Temporal, SNR and Spatial) of the scalable layers. The performance results demonstrate that our scheme is fully implementable on Intra and Inter (I, P & B) frames with all scalable features of SVC. The performance of the proposed scheme is verified by many important factors such as: the quality degradation of individual SVC layer without any bit-rate escalation; a security and statistical analysis on video perception of pre-compression encryption; computational overhead calculation caused by SE; and, comparative analysis with existing work.

The significance of the proposed SE is to resolve the key issues of bit-rate overhead and format compliancy of the bit-stream by carefully handle the syntax elements on the sensitive compression stage of encoder. The proposed SE scheme is suitable for video distribution to users who have subscribed to a differing video quality regarding bandwidth, storage and device rendering capabilities. The proposed scheme can be extended to ROI for bit-rate reduction in video surveillance [15] without any modification. The error resilience issues for the proposed work can be investigated in the transmission scenarios of scalable layers as a future work.

REFERENCES

[1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 17, no. 9, pp. 1103-1120, 2007.

[2] T. Wiegand, G. Sullivan, J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H. 264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576, 2003.

[3] H. D. Engel, R. Kutil, and A. Uhl. "A symbolic transform attack on lightweight encryption based on wavelet filter parameterization", in *Proc. of ACM Multimedia and Security Workshop, MM-SEC '06*, pp. 202-207, Geneva, Switzerland, Sept. 2006.

[4] Y. Mao, and M. Wu, "A joint signal processing and cryptographic approach to multimedia encryption," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 2061–2075, July 2006.

[5] E. Magli, M. Grangetto, and G. Olmo, "Joint source, channel coding, and secrecy," *EURASIP J. on Information Security*, vol. 2007, Article ID 79048, 7 pages, 2007.

[6] Y. Fan, J. Wang, T. Ikenaga, Y. Tsunoo, and S. Goto, "An unequal secure encryption scheme for H.264/AVC video compression standard," *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol. 91, no. 1, pp. 12-21, 2008.

[7] Y. Fan, J. Wang, T. Ikenaga, Y. Tsunoo, and S. Goto, "A new video encryption scheme for H.264/AVC," *Advances in Multimedia Information Processing,* LNCS vol. 4810, pp. 246–255, 2007.

[8] G.B. Algin, and E. T. Tunali. "Scalable video encryption of H. 264 SVC Codec," *J. of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 353-364, May 2011.

[9] C. Bergeron and C. Lamy-Bergot, "Compliant selective encryption for H.264/AVC video streams," in *Proc. of the 7th IEEE Workshop on Multimedia Signal Processing (MMSP '05)*, pp. 1–4, Shanghai, China, October 2005.

[10] J. Wang, Y. Fan, T. Ikenaga and S. Goto, "A partial scramble scheme for H.264 video", in Int. Conf. ASIC'2007, Oct. 2007.

[11] C. Berbain, O. Billet, A. Canteaut, N. Courtois, B. Debraize, H. Gilbert, L. Goubin, A. Gouget, L. Granboulan, C. uradoux, M. Minier, T. Pornin, and H. Sibert, "DECIM v2, eSTREAM report 2006/004 ," 2006. http://www.ecrypt.eu.org/stream/papers.html

[12] D. Wang, Y. Zhou, D. Zhao, and J. Mao, "A Partial Video Encryption Scheme for Mobile Handheld Devices with Low Power Consideration," In *Proc. the 2009 Int. Conf. on Multimedia Information Networking and Security*. Hubei, China, 2009, pp. 99-104.

[13] Z. Shahid, M. Chaumont, and W. Puech, "Fast Protection of H.264/AVC by Selective Encryption," in *Proc. SinFra 2009, Singaporean-French IPAL Symposium, Fusionopolis, Singapore*, 18-20 Feb. 2009.

[14] L. Dubois, W. Puech, J. Blanc-Talon "Smart Selective Encryption of CAVLC for H. 264/AVC Video," *IEEE Int. workshop on Information Forensics and Security (WIF*S), pp. 1-6, Dec. 2011.

[15] 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 *Proc. IEEE Int. Conf. Adv. Video Signal Based Surveillance*, pp. 424–429, Sep. 2009.

[16] H.-J. Lee and J. Nam. Low complexity controllable scrambler/descrambler for H.264/AVC in compressed domain. In K. Nahrstedt, M. Turk, Y. Rui, W. Klas, and K. Mayer-Patel, editors, Proceedings of ACM Multimedia, pp. 93–96, 2006.

[17] Y.G. Won, T.M. Bae, and Y.M. Ro, "Scalable protection and access control in full scalable video coding," in *Proc. 5th Int. Workshop on Digital Watermarking,* LNCS vol. 4283, pp. 407–421, 2006.

[18] Y. Kim, 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.

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

[20] C. Li, X. Zhou, and Y. Zong, "NAL level encryption for scalable video coding," in *proc. PCM*, no. 5353, pp. 496–505, 2008.

[21] C. Li, X. Zhou, and Y. Zong, "Layered Encryption for Scalable Video Coding," IEEE Conf. on Image and Signal Proc., pp. 1–4, Oct. 2009.

[22] S.W. Park, and S.U. Shin. "An efficient encryption and key management scheme for layered access control of H.264/Scalable Video Coding," *IEICE Trans. on Information and Systems*, vol. 92, no. 5, pp. 851-858, 2009.

[23] E. Magli, M. Grangetto, and G. Olmo. "Conditional access techniques for H.264/AVC and H.264/SVC compressed video," *IEEE Trans. Circuits Sys. Video Technol.*, 2008, to appear.

[24] E. Magli, M. Grangetto, and G. Olmo. "Transparent encryption techniques for H.264/AVC and H.264/SVC compressed video," *J. of Signal Proc*, vol. 91, no. 5, May 2011.

[25] G. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding standard: overview and introduction to the fidelity range extensions," *SPIE Conf. on Applications of Digital Image Processing XXVII*, 2004, pp. 454-474.

[26] J. Teuhola, "A compression method for clustered bit-vectors," *Inf. Proc. Lett.,* vol. 7, no. 6, pp. 308-311, 1978.

[27] M. Ghanbari, Standard codecs: Image compression to advanced video coding, 3rd edition, IET Press, London, UK, 2011.

[28] T. C. Chen, Y. W. Huang, C. Y. Tsai, B. Y. Hsieh, and L. G. Chen, "Architecture design of context-based adaptive variable-length coding for H.264/AVC," *IEEE Trans. on Circuits Syst. II, Exp. Briefs*, vol. 53, no. 9, pp. 832–836, Sep. 2006.

[29] D. Marpe, H. Schwarz, and T.Wiegand, "Context-adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 620–636, July 2003.

[30] Federal Information Processing Standards Publication 197, November 26, 2001-ADVANCED ENCRYPTION STANDARD (AES), available from http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

[31] M. Kuchar, "Dispelling the myths of cryptography," *Database and Network Journal*, vol. 30, no. 2, pp. 3-3, 2000.