FAST PROTECTION OF H.264/AVC BY REDUCED SELECTIVE ENCRYPTION OF CAVLC

Loïc Dubois[1,2]*, William Puech*[1] *and Jacques Blanc-Talon*[2]

[1]LIRMM Laboratory, UMR 5506 CNRS, University of Montpellier II
161, rue Ada, 34392 MONTPELLIER CEDEX 05, FRANCE
[2] DGA, Paris, France
loic.dubois@lirmm.fr, william.puech@lirmm.fr, jacques.blanc-talon@dga.defense.gouv.fr

ABSTRACT

In this paper we propose a new approach to protect video sequences while using selective encryption (SE) and reducing the encryption ratio (ER). Several methods of SE have been applied to video codec H.264/AVC in CAVLC mode in order to perform confidentiality, bit-rate, and data-size of protected video-sequences. In our scheme, SE-CAVLC is used but ER is decreased while preserving the confidentiality of the videos. Prediction error of H.264/AVC is used to spread a selective encryption through each frame, which allows the selection of just a part of the macro-blocks to encrypt.

1. INTRODUCTION

Due to the rapid growth in processing speeds and network bandwidths, the digital videos are commonplace and their number multiplies with each passing day. The phenomenal increase in the amount of transmitted and archived data requires it to be protected while ensuring an efficient transparency. Two solutions can answer these problems, namely the data security and the network security. The first solution allows a better control of the processing time and data size. Moreover, multimedia data require to be compressed and encrypted in order to reduce the transmission time. In video processing, full encryption is rarely mandatory because processing time is doubled as against a simple compression. That is why most of the applications use selective encryption (SE) which guarantees data confidentiality and protect high resolution without any duplication of data.

In this paper we present a new approach to SE. It is based on the SE-CAVLC [5, 6] but the SE is applied to the half of the macroblocks to further an experimental analysis. This analysis shows the spread of encryption thanks to the prediction error of the H.264/AVC codec. Two measures of similarity are used in this experimentation: PSNR and SSIM [8]. The final encryption reduces the encryption ratio (ER), which is the ratio between the encrypted part and the whole data. Even if the ER is reduced, the same confidentiality level is conserved.

In section 2, previous work on SE of video codec H.264/AVC is briefly presented. In section 3, after having presented an analysis of the impact of encrypting a single block per frame, we propose a method based on these results which reduces the ER, called Reduced Selective Encryption (RSE). In Section 4 we present experimental results and we show that only half of the encrypted bits are used with this approach while maintaining the same level of confidentiality. In Section 5, concluding remarks and future perspectives about the proposed scheme are given.

2. PREVIOUS WORK

H.264/AVC (also know as MPEG-4 Part 10) is the video coding standard of ITU-T and ISO/IEC. The principal characteristics of H.264/AVC are : the division of each frame in macro-blocks which are encoded separately, a Discrete Cosinus Transform (DCT) followed by a quantization of the macro-block, a prediction between macro-blocks in *intra* and *inter* frames, and an entropy coding using either run length coding (CAVLC) or arithmetic coding (CABAC).

In the literature, several methods have been proposed for the SE of videos. SE (or partial encryption) is a technique aiming at saving computation time or enabling new system functionalities by only encrypting a portion of the compressed bit-stream while still achieving adequate security [3]. SE is applied only to certain parts of the bitstream. Even if SE encrypts only a part of data, videos remains protected and confidential, like with a total encryption system; two of the main objectives of a video encryption. During the decoding step, both encrypted and non-encrypted informations should be appropriately identified and displayed [1].

Different encryption techniques including permutation, the Data Encryption Standard (DES) and the Advanced Encryption Standard (AES) [7] have been used for the SE of video. The candidate domains for SE can be divided into five broad categories. These are named spatial, video codec structure, transform, entropy coding stage and bit-stream. Encryption during the entropy coding module has been investigated by several authors. The use of Huffman entropy coder, as encryption cipher, has been studied in [9], which shows that the bitrate increases in spite of a compliant bitstream. Another method [10] has been introduced for performing encryption by using various Huffman tables, but this technique is vulnerable to plaintext attacks [2]. Entropy coding, based on selective the encryption of MPEG4 video standard, has been studied in [9] wherein DES was used to encrypt fixed length and variable length codes. In this approach, the encrypted bitstream is fully compliant with the MPEG4 bitstream format but it increases the bitrate. Further, AES has also been used in SE-CAVLC [5, 6] while encrypting only a part of the quantized coefficients in various VLC tables. This method keeps a compliant bitstream and secures data. We developed our algorithm from this approach.

3. PROPOSED METHOD

In the first step, the proposed method analyses the encrypted spread of the prediction error while encrypting a single block in the H.264/AVC CAVLC, in the *intra* mode only. SE-CAVLC [5, 6], is used for the encryption in the entropy

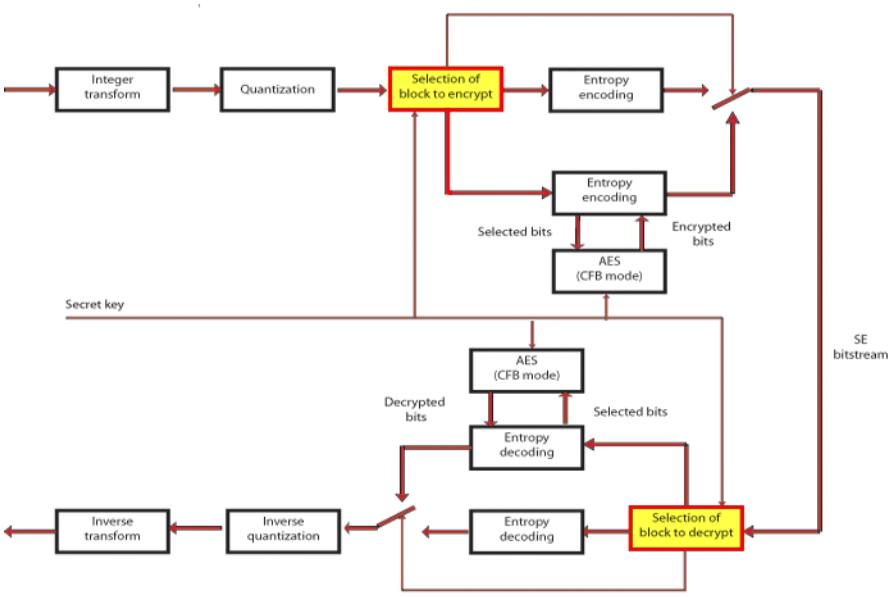


Figure 1: Block diagram of encryption and decryption process in RSE-CAVLC.

coding modules, as shown Fig. 1. The principal change, in comparison with SE-CAVLC [5, 6] is the selection of encrypted block and decrypted block as illustrated in Fig. 1. The encryption level is measured for the ten blocks in the neighborhood of the single encrypted block according to a zigzag scan, as illustrated in Fig. 2. For *chroma*, only three blocks are analyzed due to the subsampling of the *Cr* and *Cb* channels in H.264/AVC. Zigzag-scan is used to measure the first neighboring macro-blocks, because they should be more encrypted than the next macro-blocks and measurements results are more explicit. The H.264/AVC [4] algorithm uses the prediction error between macro-blocks for encoding macro-blocks in order to reduce the bit-stream. A scan, of each previously encoded neighboring macro-block, is done to find the macro-block which gives the smallest prediction error. During the decoding step, a macro-block that has been decoded from an encrypted macro-block, might be heavily distorted, thanks to this prediction error. We use this specificity in order to spread the encryption through the neighboring macro-blocks of a previous encrypted macro-block with SE. The experimental analysis in section 4.1 allows us to know the efficient range of this phenomenon in terms of confidentiality.

Furthermore, SE-CAVLC [6] is performed by using the Advanced Encryption Standard (AES) algorithm in the Cipher Feedback (CFB) mode on a subset of *codewords/bin-strings*. For each macro-block, header information is encoded but not encrypted because it is used for the prediction of the next macro-blocks. After encoding the header, the macro-block data are encoded and selectively encrypted. We are interested in this CAVLC entropy coding method. In CAVLC, SE is performed on equal length codewords from a specific VLC table. In CAVLC, five syntax elements are used to code levels and runs: *coeff_token, signs of trailing ones, remaining non-zero levels, total number of zeros* and *runs of zeros*. To keep the bit-stream compliant, only *signs of trailing ones* and *remaining non-zeros levels* are encrypted. CAVLC uses multiple VLC tables with

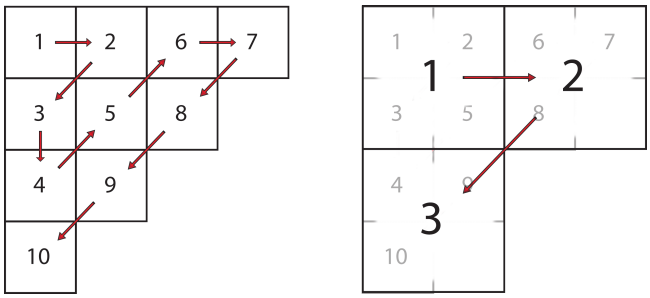


Figure 2: Zigzag scan of the ten closer macro-blocks of the single encrypted block for *luma*, and zigzag scan of the three closer macro-blocks of the single encrypted block for *chroma*

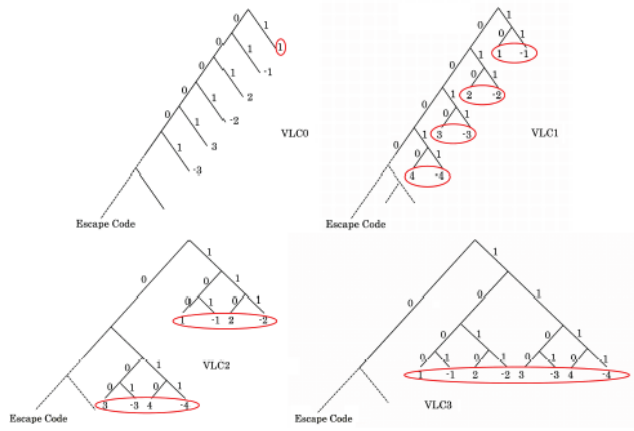
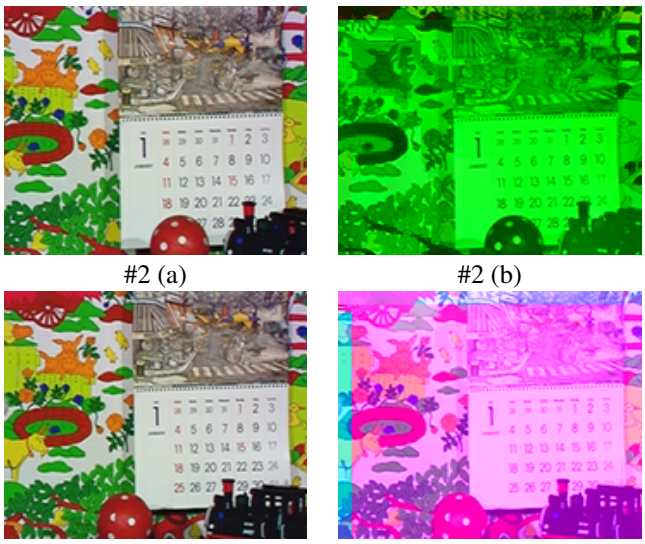


Figure 3: The first four VLC tables used in CAVLC, and the encrypted bits circled in red.

some threshold for incrementing the table. VLC codes, having same code lengths, constitute the encrypted space as illustrated in Fig. 3. For table VLC0, every *non-zero* has different codeword length, consequently we cannot encrypt the *non-zeros* in table VLC0.

In the second step, thanks to the experimental results of the section 4.1, we apply a mapping of encryption in order to reduce the ER. One block, out of two, is encrypted like a chessboard, beginning by the first macro-block because it gives the best spread of encryption, as presented in section 4. The non-encrypted block will appear encrypted with this spread through the prediction error. Global measures of PSNR and SSIM [8] are used to determine if the confidentiality remains sufficient despite of the decrease of the ER.



#2 (a) #2 (b)

#31 (a) #31 (b)

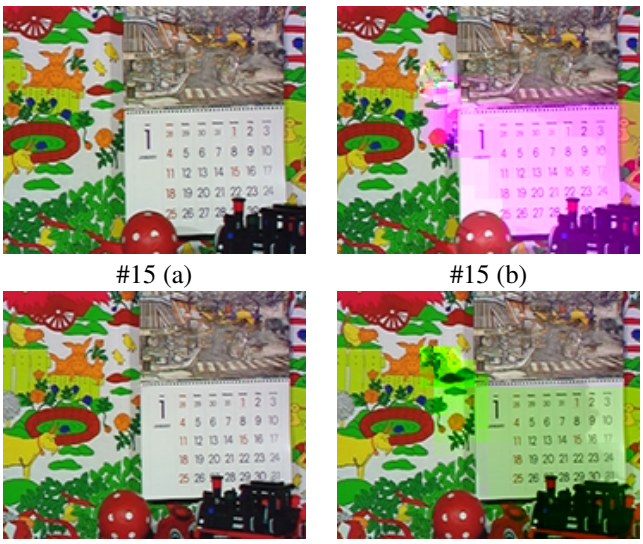
Figure 4: a) Original images of *"Mobile"*, b) Corresponding images where only the first block is encrypted using SE-CAVLC [6].

4. EXPERIMENTAL RESULTS

In this section, we have used four benchmark video sequences in QCIF resolution with a wide range of representative QP between 12 and 42. Each video represents different combinations of motion, color, contrast and objects. We have compressed 100 video frames of each video. We analyzed the spread of encryption for different macro-block positions. Results are presented with the most representative samples. In section 4.1 we present an analysis of the impact of encrypting a single block in CAVLC. In section 4.2, we compare SE-CALVLC and RSE-CAVLC methods on *video sequences*. In terms of encryption, we consider a good confidentiality if global PSNR of *luma* and global SSIM [8] of *luma* are respectively less than 13dB and less than 0.6.

4.1 Analysis of the impact of encrypting a single macro block in CAVLC

This analysis shows a significant impact of encrypting a single block on its neighbors. Fig. 4 and 5 present two frames of



#15 (a) #15 (b)

#45 (a) #45 (b)

Figure 5: a) Original images of *"mobile"*, b) Corresponding images where only the 26th block is encrypted with SE-CAVLC [6].

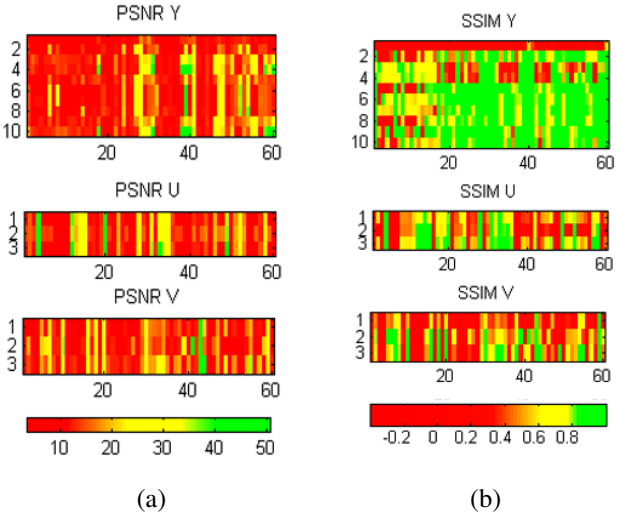


Figure 6: PSNR (a) and SSIM [8] (b) of "Mobile" in QCIF with QP 12 for the 60 first frames, only the first macro-block is encrypted.

the sequence *Mobile* where only a single block is encrypted, a drift can be clearly seen. Fig. 4.a and 5.a present the original frames while Fig. 4.b and 5.b illustrate the impact when only one macro-block is encrypted. Visually, in Fig. 4.b, colors change in the whole image and in Fig. 5.b a color drift appears diagonally behind the encrypted block.

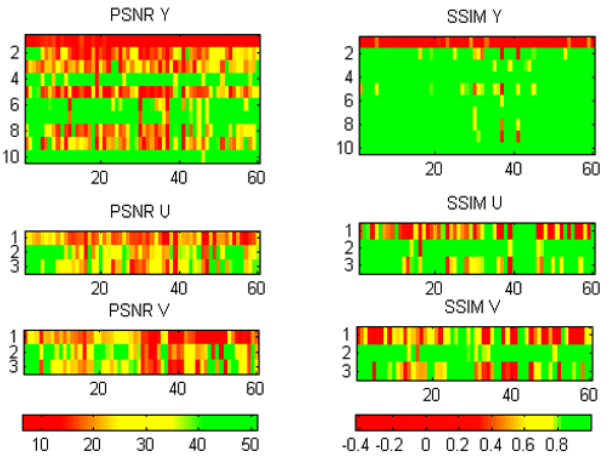


Figure 7: PSNR and SSIM [8] of *"mobile"* in QCIF with QP 12 for the 60 first frames, only the 26th macro-block is encrypted.

This phenomenon needs to be analyzed in order to validate the proposed method presented in section 3. PSNR and SSIM [8] are used in order to measure this drift. Generally, the SE of the first macro-block has the most significant impact on the neighboring blocks, as shown in Fig. 6; the first four macro-blocks are clearly under the encryption threshold. We can explain this by the fact that the first frame macro-block is totally encoded without prediction. In Fig. 7, this phenomenon is softer but still efficient in the first four macro-blocks following the encrypted block.

Also, a single encrypted block has an efficient impact on its neighboring blocks and particularly on the second, the third and the fifth. However, this impact depends on the motion, the camera, the objects and the scene. It may be significant and may spread over a large number of neighboring blocks.

4.2 Reduced Selective Encryption

The analysis achieved in Section 4.1 shows that the prediction error can be used to spread the SE, especially if the first macro-block of each frame is encrypted. In this section, a chessboard encryption is used as proposed in Section 3. RSE-CAVLC yields positive results in terms of confidentially; PSNR of *luma* varies around 10.5 dB while it is around 9 dB with SE-CAVLC [6], as shown in Fig. 9 although ER is reduced by half. Moreover, the relative maximum variation of PSNR is in the same range for both RSE-CAVLC and SE-CAVLC [6]. These results are relatively constant whatever the variation of QP between 12 and 42, and the video scene.

Moreover, this variation of PSNR remains in the same range as for SE-CAVLC [6] in Fig. 8. When RSE is applied on different video sequences, the previous results improve as

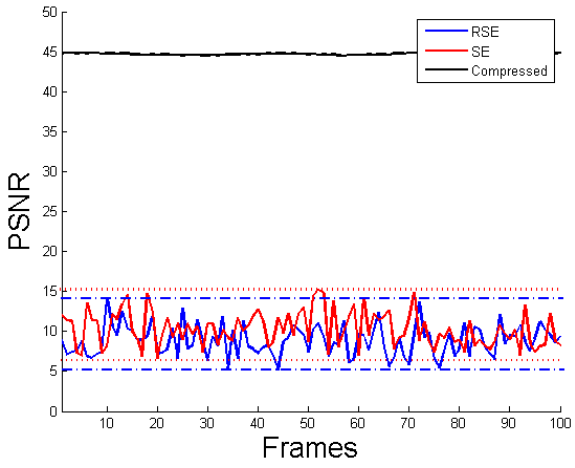


Figure 8: PSNR of *foreman* for 100 frames while using SE [6], RSE and a standard H.264/AVC compression with QP=18.

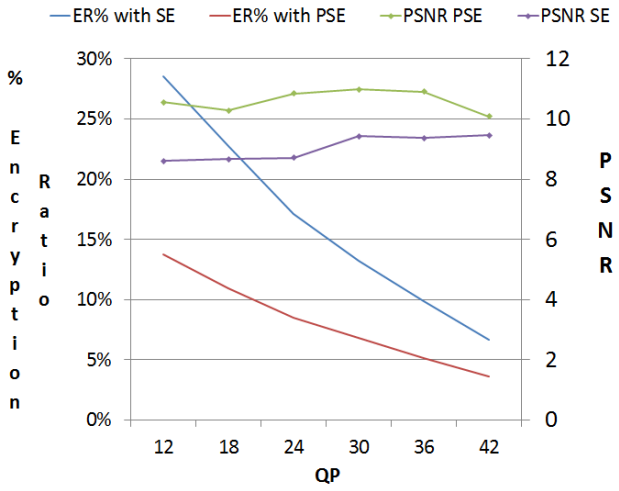
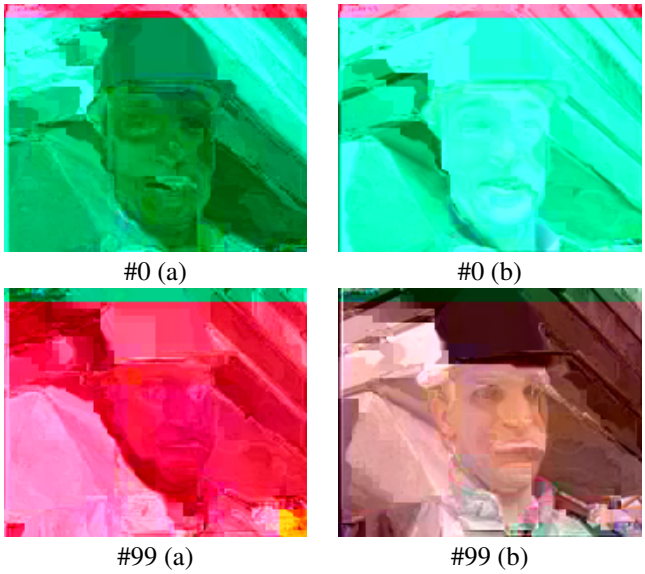


Figure 9: PSNR and ER of *foreman* while applying SE [6] and RSE for a wide range of QP.

Videos in QCIF - QP 18 - 100 frames								
	Foreman		Mobile		City		Football	
	SE [6]	RSE	SE [6]	RSE	SE [6]	RSE	SE [6]	RSE
PSNR Y (dB)	8.67	10.28	8.32	9.77	10.90	12.70	11.48	12.06
PSNR U (dB)	24.14	28.21	10.44	10.89	31.89	33.53	14.85	16.85
PSNR V (dB)	10.16	11.48	9.58	9.84	33.47	35.36	24.28	27.62
SSIM [8] Y (dB)	0.198	0.302	0.04	0.258	0.115	0.198	0.219	0.239
ER	22.76%	10.95%	36.17%	16.68%	26.41%	11.64%	25.33%	11.72%

Table 1: Analysis of ER for SE-CAVLC [6] and RSE-CAVLC for different benchmark video sequences at QP value 18.



#0 (a) #0 (b)

#99 (a) #99 (b)

Figure 10: a) Images of *foreman* in QCIF compressed with QP=12 while using SE-CAVLC [6], b) Corresponding images while applying RSE.

shown in Table 1 with four classic videos in QCIF: *foreman, mobile, city and football*. The PSNR of *luma* and *chroma* remain in the same range as SE-CAVLC [6]. SSIM [8] of *luma* is also lower than 0.6, and confirms the method's effectiveness. Visually, RSE-CAVLC and SE-CAVLC are closer, as presented in Fig. 10, and that confirms the precedent measures of similarity.

5. CONCLUSION

In this paper we showed that RSE can be applied to a video sequence in H.264/AVC CAVLC. The prediction error of H.264/AVC allows to encrypt neighboring macro-blocks of an encrypted macro-block, especially with the first macro-block of each frame. When we combine this phenomenon with a simple chessboard encryption, RSE reduces the encryption ratio (ER) while keeping a good confidentiality. Indeed, compared to the SE-CAVLC[6], the PSNR of the encrypted video sequences increases by 1.5 dB whereas the ER is reduced by one half whatever the video or its QP is.

In future, the proposed approach will be optimized in term of the selection of the encrypted data. The chessboard encryption, used in this article, may be transformed while keeping two mains axes: reduced encrypted data-size and efficient confidentiality. A final improvement may be the implementation of a smart RSE: real-time measurements would dictate the encryption of each macro-block. In the encryption loop, the first macro-block may be encrypted and the similarity of the next macro-blocks may be measured locally in order to decide either to encrypt these blocks, or otherwise. This approach may will reduce the encryption space, and may thus be more efficient and more appropriate whatever the video scene is.

REFERENCES

[1] H. Chen and X. Li. Partial Encryption of Compressed Images and Videos. *IEEE Transactions on Signal Processing*, 48(8):2439–2445, August 2000.

[2] G. Jakimoski and K. Subbalakshmi. Cryptanalysis of Some Multimedia Encryption Schemes. *IEEE Transactions on Multimedia*, 10(3):330–338, April 2008.

[3] T. Lookabaugh and D. Sicker. Selective Encryption for Consumer Applications. *IEEE Communications Magazine*, 42(5):124–129, May 2004.

[4] I. E. G. Richardson. *H-264 and MPEG-4 Video compression*. Wiley, 2003.

[5] Z. Shahid, M. Chaumont, and W. Puech. Fast protection of H.264/AVC by selective encryption of CABAC for I & P frames. *EUSIPCO*, pages 2201–2205, 2009.

[6] 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*, 2011.

[7] A. Ulh and A. Pommer. *Image and Video Encryption - From digital Rights Management to Secured Personal Communication*. Springer, 2005.

[8] Z. Wang, A. C. Bovik, and E. P. Simoncelli. Multiscale Structural Similirarity for Image Quality Assessment. *IEEE Asilomar Conference Signals, Systems and Computers*, pages 1398–1402, 2003.

[9] J. Wen, M. Severa, W. Zeng, M. Luttrell, and W. Jin. A Format-Compliant Configurable Encryption Framework for Acess Control of Video. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(6):545–557, June 2002.

[10] C. Wu and C. Kuo. Design of Integrated Multimedia Compression and Encryption Systems. *IEEE Transactions on Multimedia*, 7:828–839, October 2005.