TRANSPARENT ENCRYPTION FOR HEVC USING BIT-STREAM-BASED SELECTIVE COEFFICIENT SIGN ENCRYPTION

Heinz Hofbauer *Andreas Uhl* *Andreas Unterweger*

University of Salzburg, Jakob Haringer Str. 2,
5020 Salzburg, Austria

ABSTRACT

We propose a selective encryption scheme for HEVC which allows for transparent encryption in a wide range of quantization parameters. Our approach focusses on the AC coefficient signs, since they can be altered directly in the bit stream without entropy reencoding. This allows for fast encryption and decryption while retaining full format-compliance and length-preservation. Furthermore, we show our approach's applicability for a number of use cases by evaluating the quality degradation and robustness against attacks.

Index Terms— HEVC, transparent, encryption, bit stream, coefficient, key space

1. INTRODUCTION

We introduce an encryption scheme based on selective sign encryption of quantized transform parameters, aimed at format-compliant transparent encryption. The principal points of this encryption scheme is format compliance, i.e., the encrypted stream is decodable by a standard-compliant decoder. Digital rights management (DRM), more specifically transparent encryption, is its main field of application.

Perceptual or transparent encryption means that consumers to are able to view a preview version of the video, but in a lower quality, e.g. [1]. While preventing unauthorized consumers from accessing the full version, it is available to authorized consumers. This can be used in a pay-per-view scheme where a lower quality preview version is available from the outset to attract the viewers' interest.

Sufficient encryption aims at preventing a pleasant viewing experience, e.g. [2]. In practice, this means a reduction in quality to a point where the video is heavily distorted, but may still be recognizable. Since the content of a video is still recognizable, sufficient encryption is the middle ground between transparent encryption and *content security*, where no content should be discernible, e.g. [3].

Our proposed encryption scheme is HEVC-specific. Although numerous approaches for DCT-based video coding standards like MPEG-2 Video, MPEG-4 Part 2 and H.264 have been proposed [4, 5, 6, 7, 8, 9, 10], the latter flip all AC coefficient signs to encrypt the content, aiming at full encryption. In contrast, our approach selectively flips AC coefficient signs of the luminance channel, reducing the quality slightly, but noticeably, allowing for transparent encryption.

In addition, our approach is bit-stream based, i.e., it can be applied directly at a bit-stream level without the need to fully decode the video. Van Wallendael et al. [11] as well as Shadid and Puech [12] have investigated HEVC bit stream elements which are suitable for format-compliant bit-stream-based encryption without changes in length, one of which are AC coefficient signs. Our approach selectively encrypts the latter, allowing for transparent encryption, while retaining full format compliance and length preservation.

This paper contributes a new approach for transparent encryption which modifies a fixed percentage of coefficient signs in the bit stream rather than a fixed percentage of the total number of coefficients per block. Quantization parameters (QP) as well as GOP structure heavily affect the number of coefficients in the bit stream. This in turn affects the resulting quality and key space size, thus we will provide a thorough analysis of the visual quality impact and the key space size depending on encoding structure and QP.

This paper is structured as followed: In section 2, we describe our encryption approach. In section 3, we evaluate it with respect to quality and security before concluding the paper in section 4.

2. ENCRYPTION METHOD

Full sign encryption [12] is clearly in the region of sufficient encryption, but even partial sign encryption can introduce strong distortions. Therefore, we encrypt only a part of the coefficients of each block while keeping the parsing overhead minimal. Furthermore, with our approach we only encrypt sign bits in the luminance channel since the distortion introduced by encrypting chroma channels results in chromatic aberration which are more noticeable by the human visual system.

HEVC stores the coefficient signs for each block raw in the bit stream, i.e., without entropy coding. This makes it easy to manipulate them directly without impacting format compliance, while keeping the parsing overhead low. We

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

(a) no encryption (b) 25% encryption

(c) 50% encryption (d) 75% encryption

Fig. 1. Visual example of the sign encryption from the *crew* sequence with *randomaccess* structure and QP 21.

(a) QP 15, no encryption (b) QP 15, full encryption

(c) QP 39, no encryption (d) QP 45, no encryption

Fig. 2. Visual example of original and full sign encryption and different QP parameters.

pseudo-randomly flip a specified percentage p of signs in the bit stream. Since the scan order in HEVC is inverted, i.e., the high-frequency coefficients come first, we only encrypt the first p percent of the sign bits in the bit stream, excluding the DC coefficient sign to avoid extreme drift.

Note that coefficient signs are only stored for non-zero coefficients. Thus, encrypting p percent of the sign bits does not yield identical results to encrypting p percent of all transform coefficients (including zero coefficients). However, our proposed approach is faster as it minimizes parsing overhead and does not require any decoding.

Figure 1 shows examples of partial sign encryption. From the figure it can be seen that for transparent encryption between 25% and 50% of the signs have to be encrypted. Encrypting more than 50% of signs introduces strong distortions and results in sufficient encryption. However, even full sign encryption is not acceptable for content security, as illustrated in fig. 2 (a)–(b).

3. EVALUATION

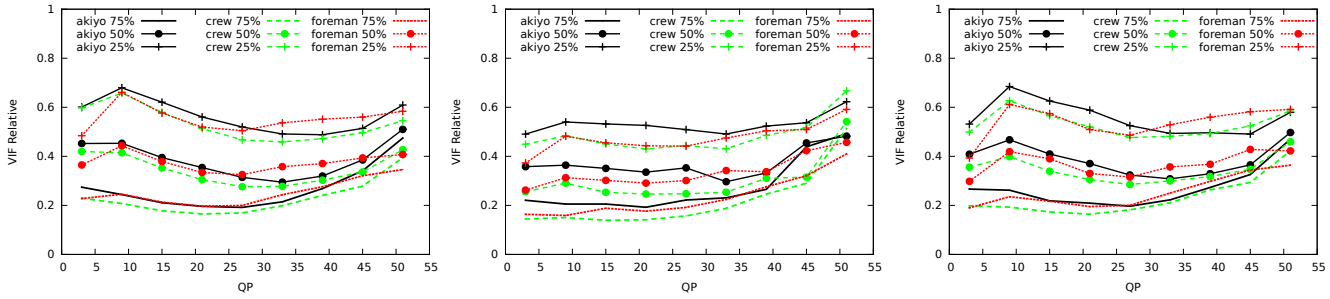
The quality analysis utilizes the visual image fidelity (VIF) image metric by Sheikh and Bovik [13]. The VIF significantly outperforms other image metrics when it comes to block based artifacts, especially in lower quality ranges, as shown by Hofbauer and Uhl [14]. In order to properly evaluate the encryption scheme, different traits of the bitstream need to be taken into account. The prediction structure has a huge influence on the propagation of the error introduced by the encryption. As such, three GOP types are used in this evaluation which reflect a variety of possible application scenarios. The GOP structures chosen are the default reference software configurations: *intra* (I frames only), *lowdelay* (one I frame, followed by groups of four P frames) and *randomaccess* (groups of one I frame followed by 32 B frames).

As a test set we chose the well known high, medium and low motion sequences, *crew*, *foreman* and *akiyo*, respectively. The different motions types were chosen because they influence prediction and the amount of coefficients for encryption.

Furthermore, we have to take into account replacement attacks [15, 16, 17, 18], which can reduce the visual distortion by replacing encrypted bitstream elements by elements which are statistically more likely to not introduce an error. Since the sign distribution is uniform, the easiest attack is to set the coefficients for which the signs are encrypted to zero. Given the differential coding nature, this will introduce less distortion than sign-flipped coefficients. In the subsequent figures, *orig* will refer to the original, i.e., unencrypted, bitstream, while *enc* will refer to the encrypted version.

The replacement attack on average increases the VIF quality by $\bar{Q}_{\text{VIF}} = 0.0306$, with a median of $\tilde{Q}_{\text{VIF}} = 0.0328$ and $\sigma_{Q_{\text{VIF}}} = 0.0199$, which results in an average quality increase by a factor of $\bar{F}_{\text{VIF}} = 1.172$, with a median of $\tilde{F}_{\text{VIF}} = 1.146$ and $\sigma_{F_{\text{VIF}}} = 0.114$. This shows that the quality increase of a replacement attack is not a security risk for this type of encryption.

Figure 3 shows the relative quality reduction of the encrypted sequence compared to the unencrypted sequence over different quantization parameters. Three behaviours, in relation to QP, can be discerned. For lower QP, there is a severe drop in quality with the same encryption type, which is more closely examined in sec. 3.2. For middle-range QP, between



(a) *intra* GOP structure (b) *lowdelay* GOP structure (c) *randomaccess* GOP structure

Fig. 3. Relative quality (VIF) of the attacked and original sequence for the given QP for 75%, 50% and 25% encryption.

15 and 40, there is a relatively stable reduction in quality, within a range of about 0.1. For higher QP, the quality drop becomes less severe.

The drop in quality reduction for high QP is influenced by the lower number of non-zero coefficients and the fact that the quality is already so low that any further impairment is not registered as strongly by the image metric, see the high QP cases without encryption in fig. 2 (c)–(d). Furthermore, given the already low quality of sequences with a QP higher than 40, further sign encryption would lead to a quality so low that it could no longer be used as a preview.

Therefore, we suggest using our encryption approach for bit streams with mid-range QP, between 15 and 40. This is also the range which is considered useful for most applications. Using on our method in this QP range lowers the quality in a way which is suitable for the described low-quality-preview scenario.

Note that the GOP structure has quasi no effect on the results at all. This can be seen from fig. 3 (a)-(c), which show very similar courses in terms of relative quality. Therefore, our approach can be used for all tested GOP structures.

3.1. Key Space Size

Figure 4 shows the average number of encrypted bits per frame. It can be seen that the key space is heavily influenced by the prediction structure of the sequence. Most key bits are compacted into I frames, since the B and P frames have a higher number of zero coefficients which are not used in sign encryption. In fig. 4 the *intra* structure has the highest number of key bits, while *lowdelay* (with only a single I frame) exhibits the lowest number of key bits. These two cases can be used as an upper and lower bound for the actual number of key bits when a different GOP structure is used, as seen in the case of *randomaccess*.

Furthermore, the number of non-zero coefficients decreases with a QP increase, consequently decreasing the number of encrypted signs and therefore the key space. This is only an apparent detrimental phenomenon since the quality in this higher QP range is already so low that any further impairment would results in sufficient rather than transparent encryption.

However, using the lower limit of the *lowdelay* GOP structure for the border case of 25% encrypted signs at QP 40, we only have about 3 bits per frame, which would result in about 720 bits for 10 second sequence at 24fps. This is clearly borderline for a security application. However, a slight increase of I frames as is the case of the *randomaccess* GOP structure, would increase this to about 25 bits per frame and consequently to 6000 bits for the same sequence.

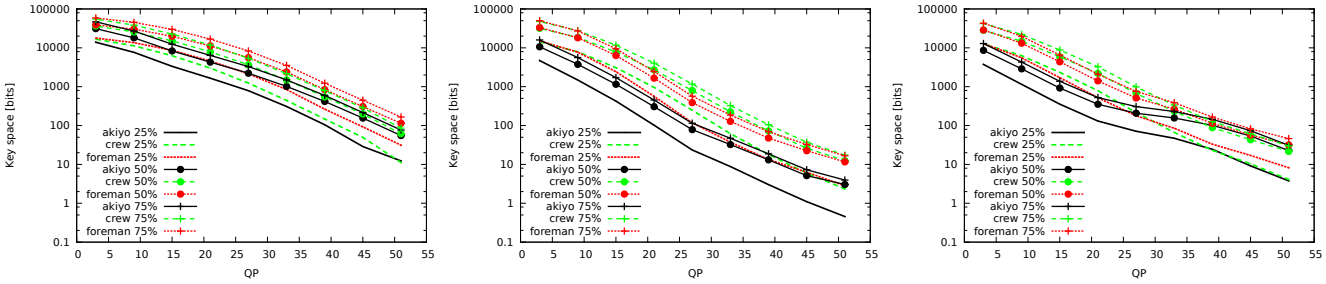
3.2. The Curious Case of High Quality Encryption

As can be seen in fig. 3, the relative quality of the encrypted sequences drops significantly at QP 3, which is against the general trend for lower QP. Hence, we analyzed the low QP (high quality) range in more detail. Figure 6 depicts the relative quality of the *foreman*, *akiyo* and *crew* sequence with the *randomaccess* GOP structure between QP 1 and 15 in steps of 1. We subsequently analyze the cause of the depicted behaviour.

For lower QP, the transform coefficient magnitudes get larger due to the smaller quantization step size. Flipping the signs of these coefficients due to encryption results in very high or very low pixel values in the picture domain, respectively. This may cause clipping to 0 or 255 for some pixels of a block. Since these clipped pixels are used for prediction, further clipping in predicted blocks is more likely to occur.

For high QP, this rarely happens and is thus negligible. The lower the QP gets, the higher the transform coefficient magnitudes get (see above) and the more likely clipping occurs, lowering the overall quality. The most extreme quality drop is at QP 4 which corresponds to a quantization step size of 1 [19]. Figure 5 illustrates this on for the foreman sequence and QP 1 to 8 for 25% encryption.

For lower QP, the encoder is more likely to bypass the transform for some blocks, i.e., it quantizes the residual pixel values directly. Since the residual values are bounded between -255 and 255, as opposed to the transform coefficients (whose magnitude may be larger), the probability of clipping in the image domain significantly decreases when flipping signs, yielding a better visual quality. Since the number of



(a) *intra* GOP structure (b) *lowdelay* GOP structure (c) *randomaccess* GOP structure

Fig. 4. Average per-frame key space of the proposed encryption methods for the given QP.

(a) QP 1 (b) QP 2

(c) QP 3 (d) QP 4

(e) QP 5 (f) QP 6

(g) QP 7 (h) QP 8

Fig. 5. Quality samples for lower QP of frame 25 of the *foreman* sequence with *randomaccess* GOP structure and 25% encryption.

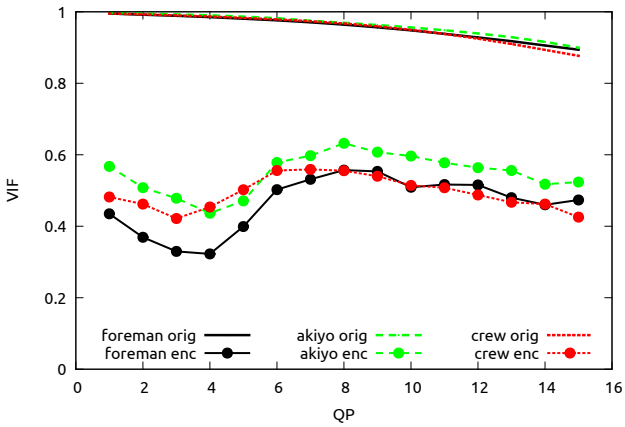


Fig. 6. Details of the low QP for the 25% encryption for the test sequences with *randomaccess* GOP structure.

blocks for which the transform is bypassed increases for lower QP, the overall quality rises again for very low QP below 4.

Due to this effect and the clipping described above, we do not recommend using our approach for very low QP, i.e., very high quality.

4. CONCLUSION

We proposed a bit-stream-based encryption approach for HEVC which pseudo-randomly flips a fixed percentage of sign bits in the bit stream. Due to its design, our approach is easy to implement and suitable for transparent and sufficient encryption, e.g., in a pay-per-view scenario for mid-range QP. We showed that the key space is sufficiently large for this application, allowing for security against attacks.

5. ACKNOWLEDGEMENTS

This work is supported by FFG Bridge project 832082.

6. REFERENCES

[1] Q. Li and I. J. Cox, "Using perceptual models to improve fidelity and provide resistance to valumetric scaling for quantization index modulation watermarking," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 2, pp. 127–139, June 2007.

[2] Thomas Stütz and Andreas Uhl, "Efficient format-compliant encryption of regular languages: Block-based cycle-walking," in *Proceedings of the 11th Joint IFIP TC6 and TC11 Conference on Communications and Multimedia Security, CMS '10*, B. De Decker and I. Schaumüller-Bichl, Eds., Linz, Austria, May 2010, vol. 6109 of *IFIP Advances in Information and Communication Technology*, pp. 81–92, Springer.

[3] Thomas Stütz and Andreas Uhl, "A survey of H.264 AVC/SVC encryption," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 3, pp. 325–339, 2012.

[4] F. Dufaux and T. Ebrahimi, "Scrambling for Anonymous Visual Communications," in *Proceedings of SPIE, Applications of Digital Image Processing XXVIII*. 2005, vol. 5909, SPIE.

[5] F. Dufaux and T. Ebrahimi, "Region-Based Transform-Domain Video Scrambling," in *Proceedings of Visual Communications and Image Processing, VCIP'06*. 2006, SPIE.

[6] F. Dufaux, M. Ouaret, Y. Abdeljaoued, A. Navarro, F. Vergnenegre, and T. Ebrahimi, "Privacy Enabling Technology for Video Surveillance," in *Proceedings of SPIE, Mobile Multimedia/Image Processing for Military and Security Applications*. 2006, vol. 6250, SPIE.

[7] Frederic Dufaux and Touradj Ebrahimi, "H.264/AVC video scrambling for privacy protection," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '08*, San Diego, CA, USA, Oct. 2008, pp. 47–49, IEEE.

[8] Frederic Dufaux and Touradj 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.

[9] Mourad Ouaret, Frederic Dufaux, and Touradj Ebrahimi, "Enabling Privacy For Distributed Video Coding by Transform Domain Scrambling," *SPIE Visual Communications and Image Processing*, vol. 6822, pp. E8222, Jan. 2008.

[10] Lingling Tong, Feng Dai, Yongdong Zhang, and Jintao Li, "Visual security evaluation for video encryption," in *Proceedings of the International Conference on Multimedia*, New York, NY, USA, 2010, MM '10, pp. 835–838, ACM.

[11] G. Van Wallendael, A. Boho, J. De Cock, A. Munteanu, and R. Van de Walle, "Encryption for high efficiency video coding with video adaptation capabilities," in *2013 IEEE International Conference on Consumer Electronics (ICCE)*, 2013, pp. 31–32.

[12] Zafar Shahid and William Puech, "Investigating the structure preserving encryption of high efficiency video coding (HEVC)," in *Real-Time Image and Video Processing 2013*, 2013, number 8656 in Proceedings of SPIE, pp. 86560N–86560N–10.

[13] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, May 2006.

[14] Heinz Hofbauer and Andreas Uhl, "Visual quality indices and low quality images," in *IEEE 2nd European Workshop on Visual Information Processing*, Paris, France, July 2010, pp. 171–176.

[15] M. Podesser, H.-P. Schmidt, and A. Uhl, "Selective bitplane encryption for secure transmission of image data in mobile environments," in *CD-ROM Proceedings of the 5th IEEE Nordic Signal Processing Symposium (NORSIG 2002)*, Tromso-Trondheim, Norway, Oct. 2002, IEEE Norway Section, file cr1037.pdf.

[16] A. Uhl and A. Pommer, *Image and Video Encryption. From Digital Rights Management to Secured Personal Communication*, vol. 15 of *Advances in Information Security*, Springer-Verlag, 2005.

[17] Thomas Stütz and Andreas Uhl, "On JPEG2000 error concealment attacks," in *Advantages in Image and Video Technology: Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology, PSIVT '09*, Tokyo, Japan, Jan. 2009, Lecture Notes in Computer Science, pp. 851–861, Springer.

[18] Heinz Hofbauer and Andreas Uhl, "Selective encryption of the MC-EZBC bitstream and residual information," in *18th European Signal Processing Conference, 2010 (EUSIPCO-2010)*, Aalborg, Denmark, Aug. 2010, pp. 2101–2105.

[19] M. Budagavi, A. Fuldseth, G. Bjontegaard, V. Sze, and M. Sadafale, "Core Transform Design for the High Efficiency Video Coding (HEVC) Standard," *IEEE Journal of Selected Topics in Signal Processing*, 2013, to appear.