A Survey of H.264 AVC/SVC Encryption-Part II

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 3, MARCH 2012

IV. Overview of H.264 [Encryption]

A. Encryption Before Compression

- If encryption is conducted before compression, the most important issue is the influence on H.264 compression performance.
- Thus if the entire visual information should be concealed, these approaches are not the method of choice, but if smaller areas need to be concealed, encryption before compression has been proposed.

- A possible solution is to encrypt the privacy-endangering image areas and then encode the modified image, e.g., a permutation of positions of the pixels in the privacy-endangering areas is proposed in [18].

- [18] P. Carrillo, H. Kalva, and S. Magliveras, “Compression independent object encryption for ensuring privacy in video surveillance,” in *Proc. ICME*, Jun. 2008, pp. 273–276.

- With respect to H.264 the following aspects need to be considered:
- Is decryption after lossy H.264 compression still possible (it is possible with pixel position permutations) and how can the influence on compression performance be minimized, which is more easily achieved when encryption is performed in the compression pipeline (cf. Prof. Min Wu's work).

- *B. Compression [Integrated Encryption]*
- In H.264 a picture is processed in blocks, starting with 16 × 16 macroblocks, which can be further sub-divided in a hierarchical tree fashion down to 4 × 4 blocks. The macroblocks can be grouped in slices, but most commonly a slice consists of the macroblocks of an entire picture (default configuration in most encoders).

- 1) *Intraprediction [Mode Encryption]:* Intraprediction may only take advantage of the previously coded data of the current slice. In intracoding the pixel data of a block can be either predicted on the basis of previous block data (in raster scan order) or transmitted directly (I_PCM mode).

- The prediction mode has to be signalled in the bitstream, modification of the intraprediction modes for encryption has been proposed in [19], [20] –[23].

- [19] S. Lian, Z. Liu, Z. Ren, and H. Wang, “Commutative encryption and watermarking in video compression,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 6, pp. 774–778, Jun. 2007.

- [20] S. Lian, Z. Liu, Z. Ren, and H. Wang, “Secure advanced video coding based on selective encryption algorithms,” *IEEE Trans. Consumer Electron.*, vol. 52, no. 2, pp. 621–629, May 2006.
- [21] S. Lian, J. Sun, G. Liu, and Z. Wang, “Efficient video encryption scheme based on advanced video coding,” *Multimedia Tools Applicat.*, vol. 38, no. 1, pp. 75–89, Mar. 2008.
- [22] Z. Shahid, M. Chaumont, and W. Puech, “Fast protection of H.264/AVC by selective encryption of CABAC,” in *Proc. IEEE ICME*, Jun. 2009, pp. 1038–1041.
- [23] P.-C. Su, C.-W. Hsu, and C.-Y. Wu, “A practical design of content protection for H.264/AVC compressed videos by selective encryption and fingerprinting,” *Multimedia Tools Applicat.*, vol. 52, nos. 2–3, pp. 529–549, Jan. 2010.

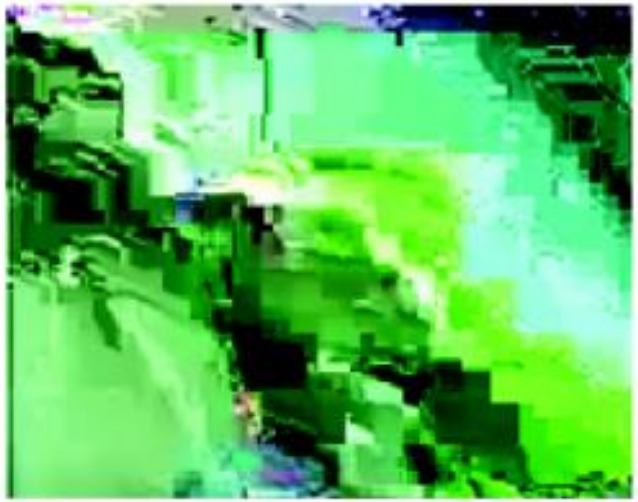
- A visual example of a encryption of the intraprediction modes is shown in Fig. 8(a). The encryption of the intraprediction modes is very similar in all contributions.

- Fig. 8. Visual examples of H.264 encryption. (a) Interprediction mode encryption: IDR picture (figure taken from [1, Fig. 1(a), p. 392]). (b) Secret DCT transforms (figure taken from [78, Fig. 6, p. 896]). (c) Sign encryption and coefficient encryption (CABAC) (figure taken from [60, Fig. 5(c), p. 1040]). (d) MVD, DCT, and Inter-PM encryption (figure taken from [43, Fig. 4, p. 286]). (e) IDR picture encryption (a reconstruction of a P picture is shown). (f) SVC base layer encryption (replacement attack). (g) SVC base layer encryption (replacement attack and level adjustment). (h) Original frame from the *Foreman* sequence.

(a)

(b)

(c)



(d)

(e)

(f)

(g)

(h)

- 2) *Interprediction [Mode Encryption]:*
- In interprediction blocks are predicted on the basis of previously decoded reference pictures. For that end motion estimation and compensation is conducted, a novelty of H.264 is the tree-structured motion estimation and compensation, which employs variable block sizes.

- Some inter frame macroblock subdivisions result in the same number of motion vectors. In [24], it is proposed to perform permutations on the set of inter frame macroblock subdivisions with the same number of motion vectors.

- [24] Y. Li, L. Liang, Z. Su, and J. Jiang, "A new video encryption algorithm for H.264," in *Proc. 5th ICICS*, Dec. 2005, pp. 1121–1124.

- 3) *Motion Vector [Difference Encryption]:*
- Motion estimation and compensation works on a macroblock basis (16×16 blocks), which can be further decomposed to 4×4 blocks. For each of the blocks of a macroblock a motion vector is calculated.

- The motion vector data are subject to further processing before entropy coding, i.e., motion vector prediction, which yields motion vector differences. The modification of motion vectors and motion vector difference data for encryption has been proposed frequently [19], [24]-[29].

- [25] N. Thomas, D. Lefol, D. Bull, and D. Redmil, “A novel secure H.264 transcoder using selective encryption,” in *Proc. IEEE ICIP*, Sep. 2007, pp. IV-85–IV-88.
- [26] S. G. Kwon, W. I. Choi, and B. Jeon, “Digital video scrambling using motion vector and slice relocation,” in *Proc. 2nd ICIAR*, LNCS 3656. Sep. 2005, pp. 207–214.

- [27] E. Magli, M. Grangetto, and G. Olmo, “Conditional access to H.264/AVC video with drift control,” in *Proc. IEEE ICME*, Jul. 2006, pp. 1353–1356.
- [28] Y. G. Won, T. M. Bae, and Y. M. Ro, “Scalable protection and access control in full scalable video coding,” in *Proc. 5th IWDW*, LNCS 4283. Nov. 2006, pp. 407–421.
- [29] M. Grangetto, E. Magli, and G. Olmo, “Conditional access to H.264/AVC video by means of redundant slices,” in *Proc. IEEE ICIP*, vol. 6. Sep. 2007, pp. 485–488.

- The schemes to modify the motion vectors and motion vector differences are diverse, e.g., many propose sign encryption [23], [24], [28], while other encrypt the suffix of the exponential Golomb code [21], [27].

- In order to control the distortion (for perceptual/transparent encryption), it is proposed in [27] and [29] to modify the motion vectors mv in the following way: $mv = mv + \text{round}(\alpha * Z)$, where Z is uniformly distributed on $[-1, 1]$ and α is used to adjust the quality degradation (the larger α, the higher the distortion).

- 4) *[Secret] Transform:*
- The residual data either obtained by intraprediction or interprediction is subject to transformation and quantization. Residual data is subject to 4×4 DCT-based transform. The chroma DC coefficients are further transformed with 2×2 Hadamard transform (in all macroblocks) and the luma DC coefficients are only further transformed with a 4×4 Hadamard transform in case of intraprediction and 16×16 mode.

- The transform coefficients are subject to scalar quantization. It has been proposed to employ different 4×4 transforms with similar properties as the 4×4 DCT-based transform [30]. If the coefficients of these different 4×4 transforms are input to the standard inverse DCT transform, a reduced quality version is decoded as shown in Fig. 8(b).

- In [31], it is proposed to encrypt the quantization parameter.

- [30] S.-K. A. Yeung, S. Zhu, and B. Zeng, “Partial video encryption based on alternating transforms,” *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 893–896, Oct. 2009.
- [31] S. Spinsante, F. Chiaraluce, and E. Gambi, “Masking video information by partial encryption of H.264/AVC coding parameters,” in *Proc. 13th EURASIP*, Sep. 2005.

- 5) *DCT Coefficient [Encryption]:* Many proposals modify the signs of the coefficients. If CAVLC encoding is applied, the context-adaptive coding makes it complex to integrate format compliant and length-preserving encryption of the coefficients other than sign encryption.

- In [32], it is proposed to encrypt the level of intra macroblock DC coefficients as well. The proposal of [33] encrypts non-zero coefficients that have been mapped to exponential Golomb codes prior to CABAC coding.

- [32] S. Lian, Z. Liu, Z. Ren, and H. Wang, “Secure advanced video coding based on selective encryption algorithms,” *IEEE Trans. Consumer Electron.*, vol. 52, no. 2, pp. 621–629, May 2006.

- [33] Z. Shahid, M. Chaumont, and W. Puech, “Fast protection of H.264/AVC by selective encryption of CABAC,” in *Proc. IEEE ICME*, Jun. 2009, pp. 1038–1041.

- The proposals of [34] and [35] encrypt only the least significant bits of a coefficient (for transparent encryption), which significantly reduces the compression efficiency. As a solution the authors propose to code these data independently and encrypt it separately.

- [34] E. Magli, M. Grangetto, and G. Olmo, “Conditional access to H.264/AVC video with drift control,” in *Proc. IEEE ICME*, Jul. 2006, pp. 1353–1356.
- [35] M. Grangetto, E. Magli, and G. Olmo, “Conditional access to H.264/AVC video by means of redundant slices,” in *Proc. IEEE ICIP*, vol. 6. Sep. 2007, pp. 485–488.

- 6) *[Secret] Scan Orders:* Prior to entropy coding in each 4×4 array of coefficients, the non-zero coefficients are mapped to a sequence by the zig-zag scan. A modification of the scan order has been proposed in [36] and [37], i.e., the instead of the zig-zag scan a random permutation is performed.
- In H.264 two entropy coding modes are available, CAVLC and CABAC.

- [36] P.-C. Su, C.-W. Hsu, and C.-Y. Wu, “A practical design of content protection for H.264/AVC compressed videos by selective encryption and fingerprinting,” *Multimedia Tools Applicat.*, vol. 52, nos. 2–3, pp. 529–549, Jan. 2010.
- [37] F. Dufaux and T. Ebrahimi, “H.264/AVC video scrambling for privacy protection,” in *Proc. IEEE ICIP*, Oct. 2008, pp. 1688–1691.

- 7) *[Joint Encryption and] CAVLC:* In CAVLC all data but the residual data is encoded with fixed codewords (not context-adaptive). Only the residual data is encoded context adaptively. Many of the codewords in CAVLC are coded with exponential Golomb codes, which can be encrypted format-compliantly.

- The exponential Golomb code encryption scheme is employed frequently, the intraprediction modes and the motion vector differences are encrypted in this fashion. In [38], the code words of the syntax element run before are permuted in an almost length-preserving fashion.

- [38] C. Mian, J. Jia, and Y. Lei, “An H.264 video encryption algorithm based on entropy coding,” in *Proc. 3rd Int. Conf. IIH-MSP*, 2007, pp. 41–44.

- 8) *[Joint Encryption and] CABAC:* CABAC is employed for coding slice data and macroblock data.
- In order to preserve the compression performance, only bits encoded in bypass mode should be encrypted, these include the suffix of the exponential Golomb coded MVD and coefficient levels.

- Performing format-compliant encryption/bit replacement directly on the compressed bitstream is extremely complex (practically infeasible) as internal states of the coder have to be preserved, otherwise the remaining data is interpreted falsely which may easily lead to format violations.

- Compression-oriented encryption schemes can be combined (see Table II of the cited paper for possible combinations) and Fig. 8(d) for a visual example of the quality reduction for a combination of MVD, DCT coefficient and inter-PM encryption.

- *C. Bitstream [Oriented Encryption]*
- The output of H.264 encoding is a bitstream (coded video sequence) which is given as a sequence of NALUs (network abstraction layer units).

- 1) *NALU [Encryption]:* The fully format-compliant encryption of the NALUs (network abstraction layer units) with simple encryption algorithms, e.g., comparable to packet body encryption in JPEG2000 [39], is not possible due to the violation of syntactical and semantical requirements.

- [39] D. Engel, T. St¨utz, and A. Uhl, “A survey on JPEG2000 encryption,” *Multimedia Syst.*, vol. 15, no. 4, pp. 243–270, 2009.

- In JPEG2000, the encryption of packet body data is possible as only fractional bitplane data is coded, i.e., every decoded bit sequence can be interpreted as fractional bitplane data, while H.264 arithmetically encodes a multitude of syntax elements, not all syntax element values are always possible.

- The decoding of an ill-suited syntax element value (by decoding encrypted data) leads to violations of semantical requirements after arithmetic decoding, e.g., a value out of range.

[40] T. St¨utz and A. Uhl, “Format-compliant encryption of H.264 AVC and SVC,” in *Proc. 8th IEEE ISM*, Dec. 2008.

- However, given the appropriate encryption schemes which avoid the generation of marker sequences, regular NALU processing can be conducted on the compressed and encrypted bitstream. The preservation of the NAL structure and syntax requirements enable the transparent application of packaging methods, [40], which is a prerequisite for efficient transmission.

- The encryption of NALUs and preservation of the NALU header has been proposed in [41]. In [40], it is highlighted that encrypted data can be transparently and efficiently signaled by the application of unspecified NALU types (NUTs) for encrypted NALUs, which then have to be ignored by a H.264 decoder.

- [41] Y. Zou, T. Huang, W. Gao, and L. Huo, “H.264 video encryption scheme adaptive to DRM,” *IEEE Trans. Consumer Electron.*, vol. 5, no. 4, pp. 1289–1297, Nov. 2006.

- 2) *Container Formats:* Several proposals [42], [43] employ *Container* formats and encrypt the H.264 bitstream or fractions of the H.264 bitstream with conventional algorithms, e.g., AES in counter mode. The basic setup is illustrated in Fig. 7.

- [42] R. Iqbal, S. Shirmohammadi, and A. El Saddik, "A framework for MPEG-21 DIA based adaptation and perceptual encryption of H.264 video," *Proc. SPIE Multimedia Comput. Netw.*, vol. 6504, no. 650403, pp. 1–12, 2007.
- [43] R. Iqbal, S. Shirmohammadi, and A. El-Saddik, "Secured MPEG-21 digital item adaptation for H.264 video," in *Proc. ICME*, Jul. 2006, pp. 2181–2184.

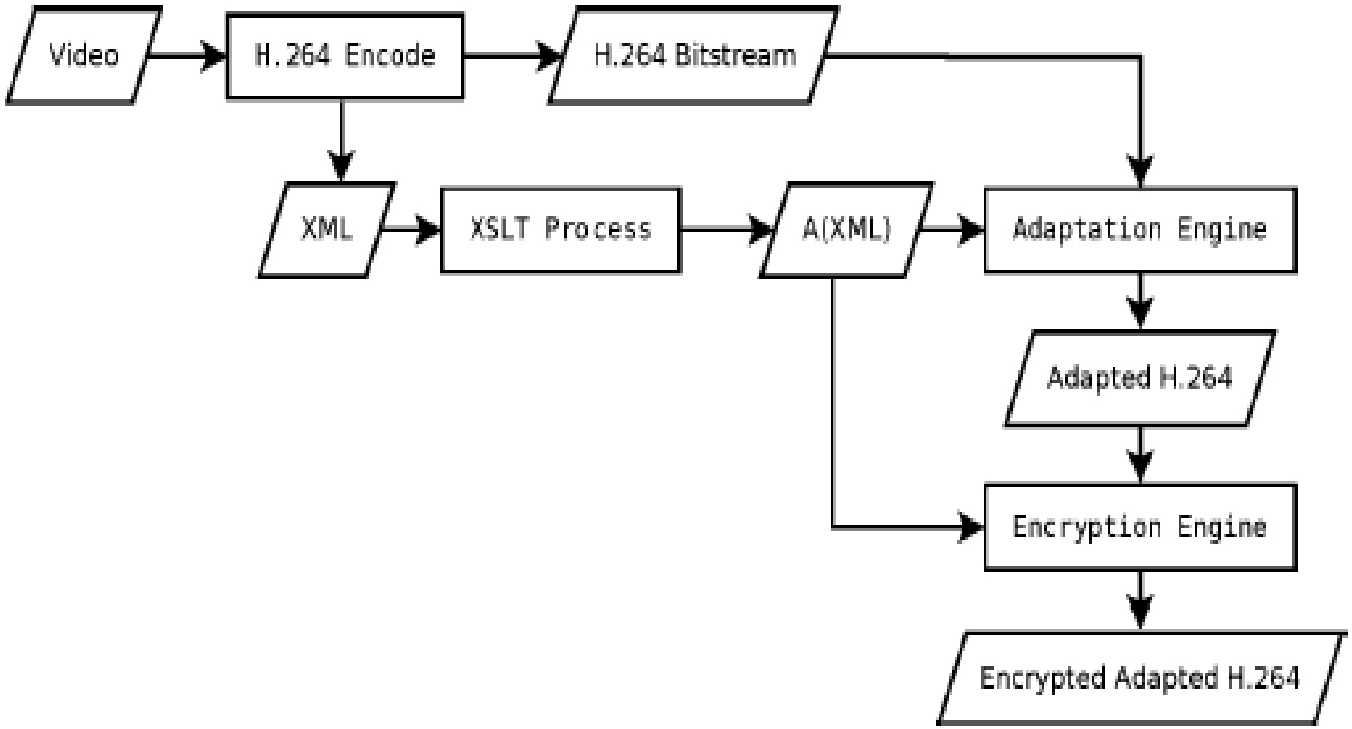


Fig. 7. MPEG-21 encryption and adaptation.

- In the compression process, additionally an XML description of the bitstream is produced. This XML description can be used to adapt the bitstream (see [44] for an evaluation for H.264). The same description can also be used to efficiently identify parts of the bitstream suitable for encryption.

- There is a standard for secure streaming of RTP data, including H.264:AVC/SVC [45]. If *Container* formats are employed functionality can be preserved through explicit signaling in the *Container* format, e.g., all NAL based functionality can be preserved.

- [44] R. Kuschnig, I. Kofler, M. Ransburg, and H. Hellwagner, “Design options and comparison of in-network H.264/SVC adaptation,” *J. Vis. Commun. Image Representation*, vol. 19, no. 8, pp. 529–542, Dec. 2008.
- [45] *ISMA Encryption and Authentication Specification 2.0*, Internet Streaming Media Alliance, San Francisco, CA, Nov. 2007.

- 3) *Partial/Selective Encryption:* Partial encryption of H.264 bitstreams can reduce the amount of data to encrypt and can also lead to performance improvements in a streaming system: results are presented for base layer encryption and IDR picture encryption (of an SVC bitstream), which corresponds to the encryption of a certain subset of the NALUs of the bitstream.

- Fig. 8(e) shows a possible reconstruction if IDR picture encryption is applied, we have replaced the IDR picture by a picture of zero values (a replacement attack) and Fig. 8(f) and (g) shows the result of a replacement attack against base layer encryption (we have replaced the base layer picture by a picture of zero values).

- Another approach is to partially encrypt NALUs, i.e., only parts of a NALU are encrypted. The encryption of the leading fraction of a NALU renders the entire NALU not decodable by standard decoders, which could be employed to effectively implement content confidentiality.

- *D. Overview of SVC [Encryption]*
- 1) *Encryption Before Compression:* There are no dedicated encryption proposals that take SVC-specifics into account.
- 2) *Compression [Integrated Encryption]:* The base layer is encoded similar to AVC, thus all encryption schemes for AVC can be basically employed in the base layer.

- The enhancement layers can employ interlayer prediction, but not necessarily have to, e.g., if interlayer prediction does not result in better compression. The compression integrated encryption approaches for AVC can be applied as well for SVC, e.g., the approaches targeting the coefficient data can also be applied for SVC.

- 3) *Bitstream [Oriented Encryption]:* The approach of [40] takes advantage of SVC to implement transparent encryption after compression. The following approach has been proposed for SVC encryption [46], which all preserve the NALU structure and encrypt almost all of the NALU payload. As the NALU structure is preserved, scalability is preserved in the encrypted domain.

- [46] N. Thomas, D. Bull, and D. Redmill, “A novel H.264 SVC encryption scheme for secure bit-rate transcoding,” in *Proc. PCS*, May 2009, pp. 1–4.