MV Compensation Attack on fast Selective Video Encryptions

Jay M. Joshi[1], Upena D. Dalal[2]
[1] S. V. M. Inst. of Tech., Bharuch-392001, Gujarat, India,
[2] S. V. National Institute of Standards and Technology, Surat, Gujarat, India

ABSTRACT:
Many researchers had published their papers related with privacy and security of multimedia data which combine the property of compression and encryption. In compression many sensitive parameters are available, without these decoding have no meaning. Encrypting only sensitive parameters make encryption faster and reduce complexity overhead of encryption. Such sensitive parameters in video codec are non–zeros in CAVLC and motion vectors, which are used commonly in many papers. Proposed encryption method is shown with the use of standard 128-bit AES algorithm in counter mode and compare with other selective encryption methods. This paper deals with sensitive application which demands strict security levels. This paper shows cryptanalysis of proposed algorithm using MV compensation technique and compare with non-zeros in CAVLC and motion vector encryption.

KEYWORDS: *CAVLC, Motion Vector, Motion Compensation, Selective Encryption, Video Security,*

I. INTRODUCTION

With continuing development of internet and mobile technology, video data is being used more and more widely, in application such as video on demand, video conferencing, broad casting etc. Now video data is closely related to many aspects of daily life, including education, commerce, and politics. In order to maintain privacy, video data needs to be secure before transmission and distribution. Different video applications need different levels of security. In general video applications can be differentiated according to the security level they demand in two categories: entertainment applications like Video on Demand, Video broadcasting, and sensitive video applications, such as telemedicine, military conferences and business meetings. Sensitive video applications usually have strict security requirements equal to those demanded for text encryption. The encryption algorithms have to withstand not only classical cryptanalytic attacks, e.g. ciphertext–only attacks, known–plaintext attacks, or chosen–plaintext attacks, but also the perceptual attacks in order to ensure that no visible information related to the business communication is disclosed.

Video compression removes redundancy, making it hard to guess information about one part of a compressed bitstream from another. Also, compression algorithms often exhibit concentration, a higher dependency of reconstruction quality on some parts of the compression syntax than on others. Combining these properties suggest selective encryption, the idea that a compressed bitstream can be adequately secured from eavesdroppers with only part of the whole bitstream encrypted. Reducing the fraction encrypted can help to reduce the total complexity of implementation which further leads to reduce the encryption and decryption time. Selective encryption has been suggested for a number of specific real time applications. In this paper, slices in a picture are compressed by H.264/AVC using baseline constrained profile with quantization parameter 16 [1]. In this paper, section 2 describes proposed selective encryption. Section 3 represents the result analysis of proposed selective encryption based on quality analysis, speed of encryption and complexity cost and compression efficiency after experimentation on different benchmark videos and compare with other selective parameter encryption. Section 4 checks the robustness of the proposed encryption based on perceptual attack using Motion Vector Compensation technique.

II. PROPOSED ALGORITHM (IDCPMV)

The proposed algorithm described here is based on encryption of lowest nonzero DCT coefficient of CAVLC in I frame with MV in P frame encryption named as IDCPMV. Encryption is done using AES-CTR (Counter mode). The other selective encryption all non-zero values of CAVLC (treated here as 'NZs in CAVLC') described in [2] and [28] and MV encryption described in [2] are compared with the proposed algorithm.

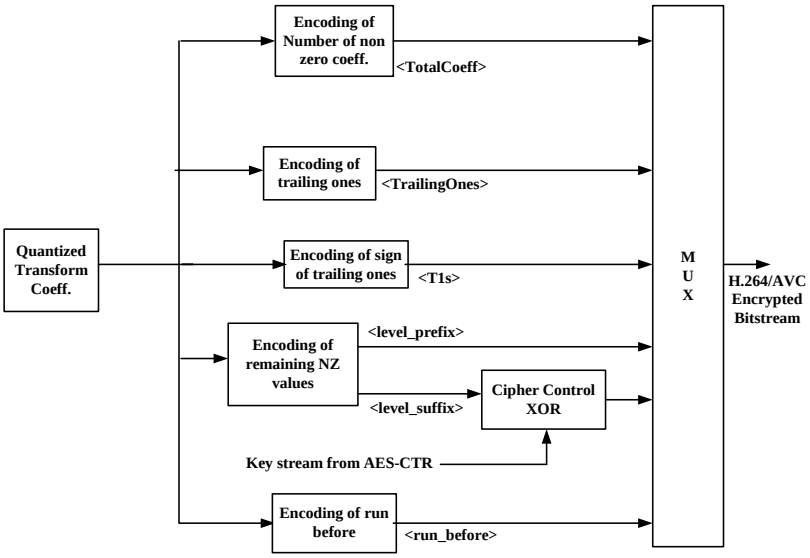


Figure1. Non zero(NZs) coefficients encryption in CAVLC for I frame

The proposed algorithm uses select two different parameters for I and P frames. This encryption uses the characteristic of entropy encoding –CAVLC for I frame. It scans the coefficients in reverse order (from high frequency non zeros to low frequency non zeros). CAVLC is designed to better exploit the characteristics of non-zeros (NZs); it works in several steps [1]. Encoding of total NZs and number of trailing ones (*T1*'s) is done by a single syntax element named *coeff_token*. It is followed by coding of signs of *T1*'s. Remaining NZs are then coded using seven VLC tables. Lastly, total number of zeros and then runs of zeros are coded. To keep the format compliant, which is a required feature for some direct operations; *coeff_token*, total number of zeros and runs of zeros are not encrypted. The suitable syntax element for encryption is the remaining nonzero values (NZs). The nonzero values are encoded by *level_suffix* and *level_prefix* bits. As *level_prefix* bits are predefined format (leading zeros followed by 1); encryption of *level_prefix* is also avoided to preserve format compliance. Thus, *level_suffix* bits are encrypted by 128-bit AES–CTR mode as shown in Fig. 1. The encryption is carried out in I frame in luminance (Y) and chrominance (Cb and Cr) planes. But the encryptions are not carried out in all the level suffix values of macroblock. The Cipher control XOR operation selects only lowest frequency level

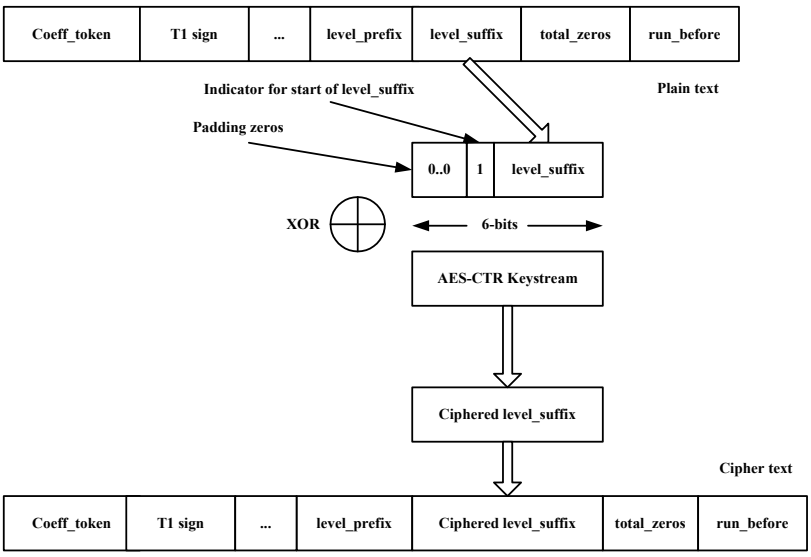


Figure2. Encryption Scheme in NZs in CAVLC for I frame

suffix values. Lowest level suffix value is XORed with 6-bit AES-CTR keystream value in manner shown in Fig. 2. The other level suffix values are medium to high frequencies values, which are not much affected on degradation of image. From the plaintext coded bits *level_suffix* bits are extracted.

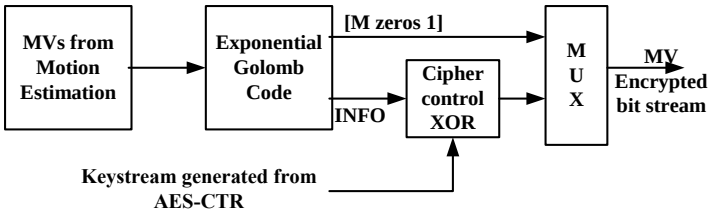
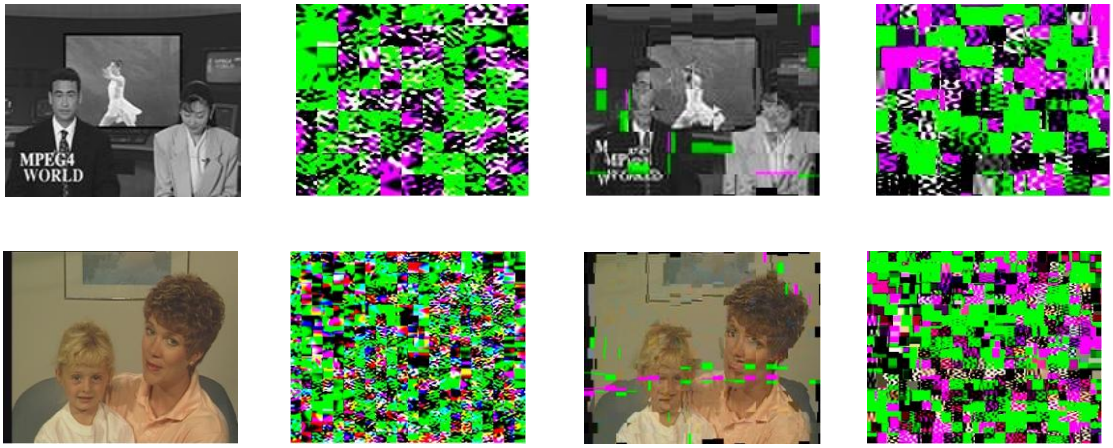


Figure3. Motion Vectors Encryption in P frame

Normally, *level_suffix* values have 0 to 6 bits. Lowest frequency value of *level_suffix* has approximately 4 to 6 bits. If *level_suffix* bits are less than 6 bits, it is preceded by 0...1. The last padding bit 1 indicates the start of *level_suffix* bits. This padded *level_suffix* bits are XORed with 6-bits of AES-CTR keystream and generated cipher *level_suffix*. This ciphered *level_suffix* put in place of plain *level_suffix* which become ciphertext. Reason of this padding and converting into 6-bits put more randomness in pixel intensity



(a) (b) (c) (d)

Figure 4. (a) Plain video frame of News and Mother-daughter sequences; Encrypted frame of (b) News (Y-PSNR=9.425dB, Y-SSIM=0.0552),and Mother-daughter (Y-PSNR=10.448dB, YSSIM=0.061513), using NZs in CAVLC; (c) News (Y-PSNR=16.806dB, Y-SSIM=0.4) Mother-daughter (Y-PSNR=21.60dB, Y-SSIM=0.6409) using MV encryption; (d) News (Y-PSNR=10.5813dB, Y-SSIM=0.077134) Mother-daughter (Y-PSNR=11.3357dB, Y-SSIM=0.1383) using IDCPMV (Proposed) encryption

after decoding without decryption compared to other proposed methods [3-6]. One execution of AES–CTR generates 128-bits keystream, which encrypt twenty one values of level suffix. Therefore, one execution of AES–CTR encrypts at least twenty one mcroblocks (16 X 16 for luminance or 8 X 8 for chrominance) of the frame.In P frame, proposed algorithm encrypts only the Motion vectors (MVs). For each 16 X 16 macroblock contains only two values of MV(x, y) which is encoded using Expo- Golomb code. That contains very few bits in whole NAL unit, which reduces complexity overheads in P frame. Expo- Golomb code produces two part of bitstream [M zeros 1] and *INFO*. The encryption is done only on *INFO* part and other predefined format [M zeros 1] is avoided to preserve format compliance. Motion Estimation engine uses the search area 46 X 46 for luminance. Hence, the range of MVs is -15 to 15. These MVs are applied to Expo- Golomb code after passing

through *se*() map [1] which converts the range of MV in 0 to 30. Expo-Golomb code give M = 0 and *INFO* = NULL (bitstream: < 1 >) for zero MV and M = 4 and *INFO* = 4 (bitstream: < 000011111 >) for thirty MV. In 55% of total MV are zeros which does not carry any bits in *INFO*. *INFO* encryption without any padding bits are proposed by the other authors [4, 7, 8, 9, 10] which does not encrypt 55% of MV bitstream. Proposed method for MV encryption scheme is described in Fig. 3. Encryption engine PME (Partial Motion Encryption) encrypts *INFO* bits of each MV with three bits of AES–CTR keystream. Maximum encrypted *INFO* bits are 3, if the number of bits in *INFO* for one MV value is lower than three bits than it is padded by zeros followed by 1. One execution of AES–CTR can encrypt at least 42 MVs of P frame. The idea behind both methods is to make encryption in every frame in every macroblock and take different parameters for I and P frame to make attacking harder.

Table 1. Average Y-PSNR and Y-SSIM of Encrypted Videos

Video	NZs in CAVLC Encryption		MV Encryption		IDCPMV Encryption	
	Y-PSNR	Y-SSIM	Y-PSNR	Y-SSIM	Y-PSNR	Y-SSIM
News	8.76	0.0454	17.7	0.3967	10.5	0.06434
Bus	9.44	0.0183	15.4	0.23	10.5	0.04711
City	9.98	0.0191	19.7	0.326	11.99	0.05683
Crew	9.89	0.0251	20.2	0.446	11.9	0.07976
Flower	6.88	0.0371	14.9	0.389	7.20	0.07507
Foreman	8.24	0.02914	18.0	0.49	8.96	0.07327
Hall	9.16	0.041	18.3	0.538	10	0.1058
Harbour	8.79	0.0251	16.9	0.282	9.8	0.04652
Mobile	8.6	0.0288	22.0	0.267	9.6	0.04867
Mother-daughter	10.5	0.0951	22.5	0.6	11.6	0.1105
Soccer	9.3	0.02073	16.9	0.3771	10.5	0.06052
Stefan	9.3	0.0316	16.1	0.33	10.3	0.06515
Tempete	9.1	0.0266	17.9	0.3374	10.9	0.05954

Table 2. Encryption Time Ratio (ETR %) of Encrypted videos

Video	NZs in CAVLC	MV	IDCPMV
News	76.7	6.06	9.72
Bus	64.9	7.08	8.76
City	66.4	6.9	8.97
Crew	64.5	6.66	8.54
Flower	72.9	7.51	9.44
Foreman	68.03	7.99	8.79
Hall	75.6	8.36	9.69
Harbour	72.25	7.56	8.9
Mobile	68.9	7.18	8.74
Mother-daughter	75.6	7.72	9.19
Soccer	75.7	7.54	9.01
Stefan	64.6	6.82	8.48
Tempete	74.7	6.85	8.64

III. EXPERIMENTAL RESULT

The standard video sequence from SFU IVB database in CIF format and standard 'News' sequence in QCIF format with SD resolution have been used here. The results in this section are carried out in the way of perception analysis, Compression efficiency and Speed of encryption.

3.1 Perception Analysis

Proposed IDCPMV algorithm provides good level of degradation in video and no perception sequences as shown in Fig. 4 compared with MV encryption. In spatial domain every pixel values are get randomize in all three planes and it is observed abrupt changes at the boundary values of macroblock except MV encryption. In temporal domain, luma and chroma values rise to maximum and get back to minimum values within the one to two frames and this cycle is repeated. Lots of transitions are observed in values of color and brightness in both spatial and temporal domain. Peak Signal to Noise (PSNR) is good metric to measure the perception analysis. Average Y-PSNR (for luminance frame) is shown in Table 1. Proposed algorithm provides slightly higher PSNRs compared to NZs in CAVLC. However, their PSNR values of proposed algorithm is still lower than algorithm described in paper [11-14], which proves more degradation in video and does not give any perceptual information. The PSNR metric suffers from a number of limitations. Hence, perception security is better measured using Structural Similarity Index (SSIM) [27]. Table 4 compares the average SSIM of 150 frames for all standard videos. It is seen from the table that the encrypted videos' SSIMs from the proposed method are lower as compared to [9, 28] as well as MV encryption. Thus, proposed algorithm and NZs in CAVLC encryption achieve a high perception security, which does not give any perception.

Table 3. Change in Compression Ratio(CCR %)

Video	NZs in CAVLC	MV	IDCPMV
News	0.067	0.0023	0.016
Bus	0.103	0.006	0.008
City	0.108	0.009	0.014
Crew	0.113	0.006	0.009
Flower	0.123	0.007	0.008
Foreman	0.115	0.009	0.012
Hall	0.125	0.006	0.008
Harbour	0.107	0.004	0.008
Mobile	0.12	0.005	0.006
Mother-daughter	0.091	0.004	0.013
Soccer	0.099	0.01	0.011
Stefan	0.113	0.005	0.007
Tempete	0.124	0.004	0.007

3.2 Speed of encryption

Encryption algorithm should be efficient so that it does not delay the transmission or access the operations for real time video applications. To measure the encryption speed Encryption Time Ratio (ETR) is the better choice [15]. ETR is the percentage ratio of encryption time with decryption time to encoding time with decoding time. ETR is calculated by considering 250 frames for News sequence in QCIF format and 300 frames for other CIF format videos (See table 2). MV encryption provides lowest ETR. NZs in CAVLC provides very high ETR. Proposed encryption gives slightly higher ETR compared to MV. However, proposed encryption provides lower encryption speed compared to method described in [4, 10, 16, 17, 18, 19, 20] and higher encryption speed compared to the method described in [5, 13, 21, 22, 23] based on encryption time and ETR. If ETR $\leq$ 10%, encryption/decryption operations is very efficient compared with the encoding/decoding and does not affect the real time system [15]. Therefore, MV and IDCPMV encryption algorithms are suitable for real time applications.

3.3 Compression efficiency

Selective encryptions described earlier are applied to video data after compression. Ideally video encryption algorithms should not change the compression ratio or should at least keep the changes in a small range. This is especially important in wireless or mobile applications, in which the channel bandwidth is limited. To evaluate the effects on bandwidth, change in compression ratio (CCR) is used as the ratio between the changed data size and the original data size [15]. IDCPMV provides lower CCR compared to NZs in CAVLC algorithm (See table 3). CCR value is also lower compared to the algorithms presented in [5, 10, 11, 12, 18, 19, 24, 25, 26]. As CCR $\leq$ 10%, proposed algorithm have near constant bandwidth and suitable for wireless and mobile applications [15].

IV. MV COMPENSATION ATTACK

Proposed algorithms described, uses counter mode 128-bit AES algorithm which withstands all the classical attack at bits level. However, at the pixel level it might have to check robustness against various attacks in different environment. For all three techniques, it is assumed that the all encrypted frames are degraded in video quality, and it might be improved by perceptual attack which gives some perceptible information.In these techniques, two properties were taken into account: 1) I frames have less degradation than P frames and 2) the consecutive frames have similar characteristics. MV compensation is done only in encrypted P frame shown in Fig. 4. Attack deals with two different approaches: 1. Ciphered I frame is used as a reference frame for estimating the motion vectors of all the macro blocks and compensation is applied on ciphered P frame. 2. Same work is done using plain I frame as reference frame.

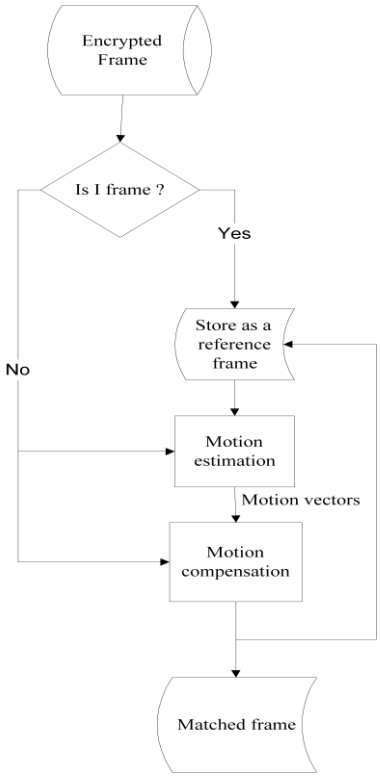


Figure 5. Motion Compensation attack for recovery from encrypted frame

MV compensation attack is shown in Fig. 5. From the previous frame calculate new motion vectors by finding out the position of best matching macroblock (i.e. motion estimation) and change the position of current ciphered microblock of P frame using this new motion vectors (i.e. motion compensation). Selection of reference frame is either cipher I frame (first approach) or plain I frame (second approach). In this attack, there is no recovery in I frame, rather I frame is used as the first reference frame to find out next motion vectors for preceding P frame. The compensated P frame is again taken as a reference frame to find out next motion vectors and so on. Results of attack are shown in Fig. 6. As IDCPMV and NZs in CAVLC encryption reference frame has ciphered I frame for first approach, frames (as shown in Fig. 6(a) and (c)) still contain high perception security after attack. In MV encryption I frame is not encrypted, hence reference frame is plain I frame for first approach also. Slight perception is seen in Fig. 6(b). However, faces of news readers as well as mother and daughter are not recognized. The message 'MPEG4' and 'WORLD' are still garbled. Second approach plain I frame is given as reference frame; hence the results of second approach for IDCPMV and MV are same as MV encryption with first approach (See Fig. 6(b)). Perception of encrypted frame of NZs in CAVLC after attack with second approach is little bit clear. We can recognize the face of male news reader, daughter. Even the word 'MPEG4' is visible in news sequence. Perception security is very low as PSNRs and SSIMs are very high. In all the analysis, IDCPMV achieve good security against MV compensation attack in first and second approach.

V. CONCLUSION

In this paper, all the issues related to speed of encryption, change in compression ratio and security of proposed algorithm are discussed with other selective parameter encryptions using AES-CTR mode of encryption in H.264/AVC bitstream. MV encryption and proposed encryption techniques achieve satisfactory video encryption result with low complexity overhead. Software implementations of these algorithms are fast enough to reach real time applications. From the all analysis in this paper, proposed algorithm IDCPMV provides higher level of perception security compared to other encryption methods. Proposed algorithm has format compliance and does not degrade in video quality after decrypting and decoding. In terms of security, compression efficiency and speed of encryption, proposed encryption technique is suitable for fast encryption for highly confidential video applications in wireless environment. As a future work, it is planned to carry out this work with CABAC entropy coding in Main profile.

REFERENCES

[1]. ITU-T Rec H.264, ITU-T: Advanced video coding for generic audiovisual services, v11, 2009.

[2]. Joshi J, Dalal U., "Enhancing selective ISMACryp video encryption for real time applications in handheld devices", *in Proc TENCON 2011 - 2011 IEEE Region 10 Conference*, 2011; p.p. 274 –278, doi:10.1109/TENCON. 2011.6129107.

[3]. Dubois L, Puech W, Blanc-Talon J. "Smart selective encryption of cavlc for H.264/AVC video", *Information Forensics and Security (WIFS), 2011 IEEE International Workshop on*, 2011; p.p. 1–6, doi:10.1109/WIFS. 2011.6123130.

[4]. Lian S, Sun J, Liu G, Wang Z. "Efficient video encryption scheme based on advanced video coding", *Multimedia Tools and Applications 2008*; 38(1):75–89, 2008; doi:10.1007/s11042-007-0150-7

[5]. Qian Z, Jin-mu W, Hai-xia Z. "Efficiency video encryption scheme based on H.264 coding standard and permutation code algorithm", *in Proc Computer Science and Information Engineering*, 2009 *WRI World Congress on*, vol. 1, 2009; p.p. 674–678, doi:10.1109/CSIE.2009.334.

[6]. Feng Wang L, dong Wang W, MA J, XIAO C, Qiao Wang K., "Perceptual video encryption scheme for mobile application based on H.264", *Journal of China Universities of Posts and Telecommunications 2008*; 15, Supplement(0): p.p. 73–78,2008; doi:http://dx.doi.org/10.1016/S1005-8885(08) 60159-4.

[7]. Bhargava B, Shi C, Wang SY. "MPEG video encryption algorithms", *Multimedia Tools and Applications*, 2004; 24(1), p.p. 57–79, doi: 10.1023/B:MTAP.0000033983.62130.00.

[8]. Lian S, Chen X. "Secure and traceable multimedia distribution for convergent mobile TV services", *Computer Communications* 2010; 33(14), p.p. 1664–1673.

[9]. Boho A, Van Wallendael G, Dooms A, De Cock J, Braeckman G, Schelkens P, Preneel B, Van de Walle R. "End-to-end security for video distribution: The combination of encryption, watermarking, and video adaptation", *Signal Processing Magazine, IEEE* 2013; 30(2), p.p. 97–107, doi:10.1109/MSP.2012.2230220.

[10]. Varlakshmi L, Sudha G, Jaikishan G. "An efficient scalable video encryption scheme for real time applications", *Procedia Engineering*, 2012; 30(0), p.p. 852–860, doi:http:// dx.doi.org/10.1016 /j.proeng.2012.01.937.

[11]. Dufaux F, Ebrahimi T. "Scrambling for privacy protection in video surveillance systems", *Circuits and Systems for Video Technology, IEEE Transactions on*, 2008; 18(8), p.p. 1168–1174,doi:10.1109/TCSVT.2008. 928225.

[12]. Dai F, Tong L, Zhang Y, Li J. "Restricted H.264/AVC video coding for privacy protected video scrambling", *Journal of Visual Communication and Image Representation* 2011; 22(6), p.p. 479 – 490, doi:http://dx.doi.org /10.1016/j.jvcir.2011.05.006.

[13]. Hong GM, Yuan C, Wang Y, Zhong YZ. "A quality controllable encryption for H.264/AVC video coding", *Advances in Multimedia Information Processing - PCM* 2006, vol. 4261, *Springer Berlin Heidelberg*, 2006; p.p. 510–517, doi:10.1007/ 11922162 59.

[14]. Zeng W, Lei S. "Efficient frequency domain selective scrambling of digital video", *Multimedia, IEEE Transactions on* 2003; 5(1), p.p. 118–129, doi:10.1109/ TMM.2003.808817.

[15]. Lian S., *Multimedia Content Encryption: Techniques and Applications*, CRC Press: Taylor & Francis Group, 2009; p.p. 15–16.

[16]. Wang X, Zheng N, Tian L. "Hash key-based video encryption scheme for H.264/AVC", *Signal Processing: Image Communication* 2010; 25(6), p.p. 427–437, doi:http://dx.doi.org/10. 1016/j.image.2010.03.005.

[17]. Raju C, Srinathan K, Jawahar C. "A real-time video encryption exploiting the distribution of the DCT coefficients", *in Proc TENCON 2008 - 2008 IEEE Region 10 Conference*, 2008; p.p. 1–6, doi:10.1109/TENCON.2008. 4766482.

[18]. Kodikara Arachchi H, Perramon X, Dogan S, Kondoz A. "Adaptation aware encryption of scalable h.264/avc video for content security", *Signal Processing: Image Communication* 2009; 24(6), p.p. 468 – 483.

[19]. Qiao L, Nahrstedt K., "A new algorithm for mpeg video encryption", *In Proc. of The First International Conference on Imaging Science, Systems, and Technology (CISST97)*, 1997; p.p. 21–29.

[20]. Tosun A., Feng Wc. "On error preserving encryption algorithms for wireless video transmission", *In Proceedings of the ninth ACM international conference on Multimedia, MULTIMEDIA '01*, ACM:New York, NY, USA, 2001; p.p. 302–308, doi:10.1145/500141. 500187.

[21]. Akkus I, Ozkasap O, Civanlar M. "Secure transmission of video on an end system multicast using public key cryptography", *Multimedia Content Representation, Classification and Security*, vol. 4105, *Springer Berlin Heidelberg*, 2006; p.p. 603–610, doi: 10.1007/ 11848035 80.

[22]. Kumar A, Ghose M. "Extended substitution–diffusion based image cipher using chaotic standard map", *Communications in Nonlinear Science and Numerical Simulation* 2011; 16(1), p.p. 372 – 382, doi:http://dx.doi.org/10. 1016/j.cnsns.2010.04.010.

[23]. Patidar V, Pareek N, Purohit G, Sud K. "Modified substitution diffusion image cipher using chaotic standard and logistic maps", *Communications in Nonlinear Science and Numerical Simulation* 2010; 15(10), p.p. 2755 – 2765, doi:http://dx.doi.org/10. 1016/j.cnsns.2009.11.010.

[24]. Magli E, Grangetto M, Olmo G. "Transparent encryption techniques for h.264/avc and H.264/SVC compressed video", *Signal Processing* 2011; 91(5), p.p. 1103 – 1114, doi:http://dx.doi.org/ 10.1016/j.sigpro.2010.10.012.

[25]. Liu Z, Li X, Dong Z. "A lightweight encryption algorithm for mobile online multimedia devices" *Web Information Systems – WISE* 2004, vol. 3306, *Springer Berlin Heidelberg*, 2004; p.p. 653–658, doi:10.1007/978-3-540-30480-7 67.

[26]. Tang L. "Methods for encrypting and decrypting MPEG video data efficiently", *In Proc. of the fourth ACM international conference on Multimedia, MULTIMEDIA '96*, ACM: New York, NY, USA, 1996; p.p. 219–229, doi:10.1145/244130.244209.

[27]. Wang Z, Bovik A, Sheikh H, Simoncelli E. "Image quality assessment: from error visibility to structural similarity", *Image Processing, IEEE Transactions on* 2004; 13(4):600–612, doi:10.1109/TIP.2003.819861.

[28]. Shahid Z, Chaumont M, Puech W. "Fast protection of h.264/avc by selective encryption of CAVLC and CABAC for I and P frames", *Circuits and Systems for Video Technology, IEEE Transactions on* 2011; 21(5):565–576, doi:10.1109/TCSVT. 2011.2129090.