H.264 based Selective Video Encryption for Mobile Applications

Saranya.P
M.Tech.
Dept. of ECE,
Sri Manakula Vinayagar Engg. College,
Puducherry, India.

Varalakshmi.L.M
Assistant Professor,
Dept. of ECE,
Sri Manakula Vinayagar Engg. College,
Puducherry, India.

ABSTRACT
With high development of computer technology and internet technology, multimedia service has become a new area in internet services today. It has problems like huge amount of data, high speed play, bandwidth limitations etc. Enhancing its security and speed has become an assurance of the video service. Using H.264 to compress and encrypt, videos can solve the speed and security problems in mobile application. Protecting the video information by encrypting selective data is the crucial element. Considering the limited resource and bandwidth of mobile devices, a selective video encryption algorithm, is proposed based on the special features of H.264. In this algorithm, the luminance transform coefficients of residual data are selectively encrypted. Experimental results demonstrate that the proposed algorithm encrypts much less important data and achieves good security and high efficiency.

Keywords
H.264, Video Coding, Partial Encryption, Luminance Transform Coefficients, Stream Cipher**.**

1. INTRODUCTION
Multimedia services are more prominent in our day to day lives. More and more people use mobile handheld devices for transferring sensitive information and also to perform commercial transactions. Multimedia information such as graphics, images, audio and video have been widely used in the portal mobile devices. Security in video conference, video surveillance, pay-TV, etc., becomes a challenging task in video communication especially for wireless mobile device. So an efficient encryption algorithm for multimedia data will become increasingly important. H.264/Advanced Video Coding (AVC) has been widely used because of its high compression rate and network friendliness. The H.264 standard is also known as MPEG-4 Part 10 and is a successor to earlier standards such as MPEG-2 and MPEG-4.

In this paper, aiming at wireless mobile application, a selective encryption of lightweight luminance transform coefficients is proposed. The rest of this paper is organized as follows. Section 2 will introduce the related research on video encryption algorithm. In Section 3, the selective video encryption algorithm proposed is given in detail. In Section 4, experimental results are analyzed. Conclusion is presented in Section 5.

2. RELATED RESEARCH

2.1 Overview of the H.264 Baseline Profile
The H.264 Baseline Profile supports coded sequences containing I- and P-slices. I-slices contain intra-coded macroblocks in which each 16 × 16 or 4 × 4 luma region and each 8 × 8 chroma region is predicted from previously-coded samples in the same slice. P-slices may contain intra-coded, inter-coded or skipped MBs. Inter-coded MBs in a P slice are predicted from a number of previously coded pictures, using motion compensation with quarter-sample (luma) motion vector accuracy. After prediction, the residual data for each MB is transformed using a 4×4 integer transform (based on the DCT) [1] and quantized. Quantized transform coefficients are reordered and the syntax elements are entropy coded. In the Baseline Profile, transform coefficients are entropy coded using a context-adaptive variable length coding scheme (CAVLC) and all other syntax elements are coded using fixed-length or Exponential-Golomb Variable Length Codes. Quantized coefficients are scaled, inverse transformed, reconstructed and filtered with a de-blocking filter before being stored for possible use in reference pictures for further intra- and inter-coded macroblocks.

2.2 Former Encryption Algorithm for H.264
Video encryption algorithms can be classified into three categories [2**]:** naive encryption algorithm, joint compression encryption algorithm and selective encryption algorithms. The former two encryption algorithms are not well suited to encrypt video data in mobile application, because of high computational complexity, more power consumption and extra storage. These shortcomings restrict their applications. On the other hand, selective encryption algorithm has been a good choice for providing good security and at low overhead. Several selective encryption algorithms for H.264 are discussed here.

1. The scheme proposed in [3] encrypts videos by scrambling the intra-prediction mode (IPM) of intra macroblocks. The intra-prediction scramble method [4] is efficient and simple, but random sequence is equal to the IPM in length. So its main security problem results from the length of the pseudo number sequence.

2. Motion Vectors Differences (MVD) containing dynamic information is the important data in H.264 encoding and decoding process. Work in [5] proposed a MVD scrambling scheme, which extracts signs of MVD, XORing them with secret keys to introduce security.

3. Selective encryption of DCT coefficients is proposed in [6]. The scheme is to encrypt DCs and ACs except 0 and 1 by using AES [7] in stream cipher mode with a key. Security of this scheme is high, but selection of data and encrypting operation can reduce its speed. Moreover, it is to use a secret key randomly changing the sign bits of ACs which equals to 1. The signs encryption algorithm extracts ACs which equals to 1 to be binary bitstream and XOR them with secret keys, the encryption space is low, and its visual quality is not high. Encrypting DC and AC coefficients [8] has destroyed data's

statistical characteristic which influences the entropy encoding and so the compression ratio is changed.

4. In entropy coding domain, in [9], a format compliant selective encryption permutated the codeword in MPEG-4. In [10], encryption in combined fixed length code (FLC) and variable length code (VLC) codeword are achieved by using permutation of the codeword and encrypting the index of code table in MPEG-4. In [11], the security issues of the multimedia encryption schemes using multiple Huffman table are analyzed.

5. In bitstream domain [12], studies indicate that encryption in bit stream domain can destroy the structures and syntax readily available, so it is not proper for selective encryption.

3. THE PROPOSED SELECTIVE VIDEO ENCRYPTION ALGORITHM

The major point of selective encryption algorithm is to select proper data to encrypt. Taking the limited power and processor resource of wireless mobile device into account, we select data before compression for encryption which gives higher efficiency and lower cost. In the H.264 standard, a macroblock is predicted either spatially or temporally. Intra prediction is an important technique in image and video compression to exploit correlation within one picture to predict spatially.

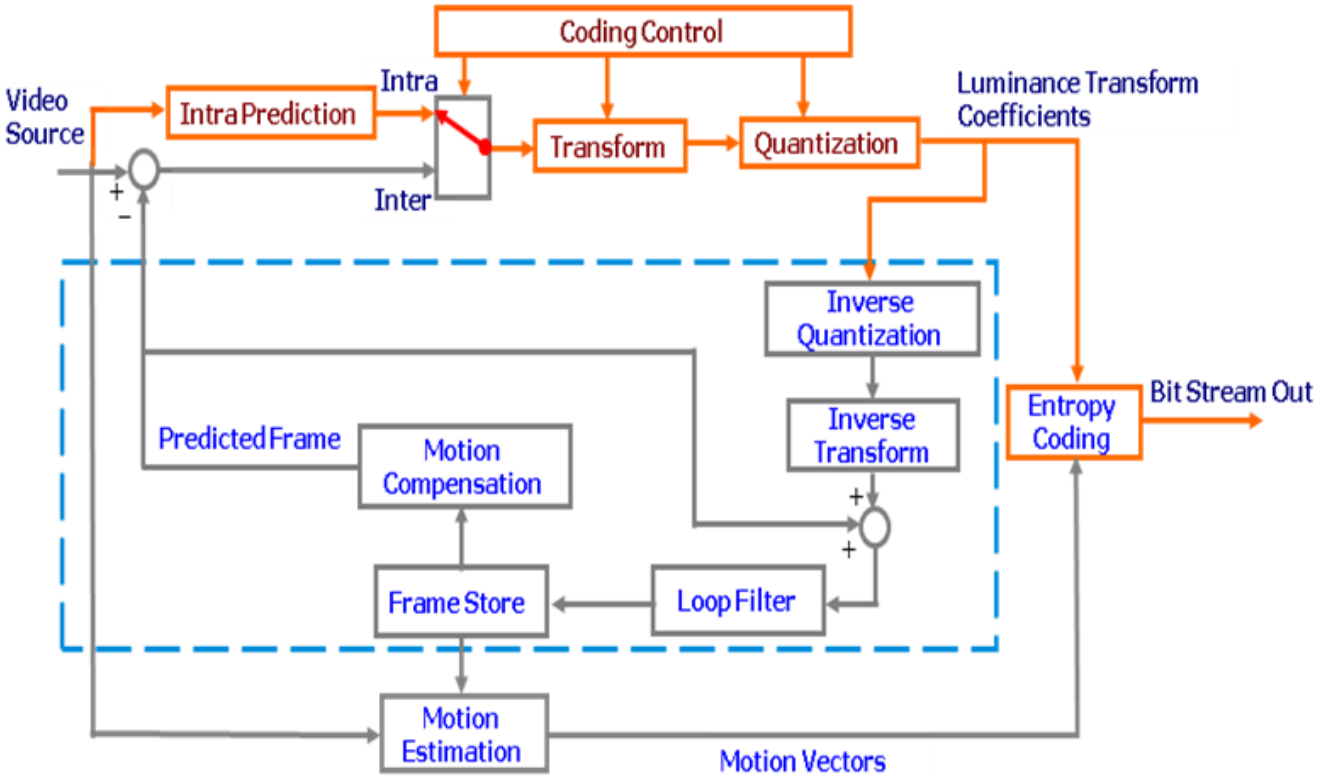


Fig 1: H.264 Encoder Block Diagram

Intra prediction block is formed based on previously encoded and reconstructed blocks and is subtracted from the current block prior to encoding. It predicts the pixels in a MB using the pixels in the available neighboring blocks; each 4×4 luma prediction mode generates 16 predicted pixel values using some or all of the neighboring. There are a total of nine optional prediction modes for each 4 × 4 luma block. The encoder typically selects the prediction mode for each block that minimizes the difference between prediction block and the block to be encoded. After prediction, the luminance residual data for each macroblock is transformed using integer transform and quantized. Then, the quantized luminance transform coefficients are reordered and entropy coded.

The quantized transform residual coefficients before the context-based adaptive variable length coding (CAVLC) are adapted to encrypt. The transform coefficients are the vital information for decoder. If transform coefficients data are encrypted, unauthorized users have difficulty in decoding correctly without the key.

A macroblock is composed of luminance and chrominance samples. To the human eye, it is much more sensitive for difference in amount of light luminance than the actual color of the light chrominance, so only the luminance transform coefficients (LTC) are considered to encrypt in order to minimize the computational burden. There are four kinds of LTC. In Intra 16x16 mode, DC LTC are a 4x4 array (marked by T_1), and AC LTC in each of the 4x4 luminance blocks have 15 vectors (marked by T_2). In all 4x4 blocks (intra or inter), there are 15 AC LTC (marked by T_3) and one DC LTC (marked byT_4). The 4x4 mode is well suited for coding of picture parts with significant details, while the 16x16 mode is more suitable for very smooth areas of the picture. The use frequency of the 4x4 mode is always more than the 16x16 mode in encoder process. In proposed LTCE algorithm, the one DC LTC must be encrypted. Moreover, the DC is independent of the other AC coefficients. The DC encryption and AC encryption are independent of each other, and they can be freely combined.

To ensure the security, traditional encryption ways are used, which includes block ciphers and stream ciphers. Block ciphers tend to simultaneously encrypt groups of characters of a plaintext message using a fixed encryption transformation. Stream ciphers convert plaintext to ciphertext one bit at a time, so they are generally faster than block ciphers at speed, desirable in a multimedia coding and transmission environment, because of limited error propagation.

This paper proposes a single scheme with RC4 stream ciphers. RC4 generates a pseudorandom stream of bits in which the encryption is combined with the plaintext using bit-wise exclusive-or; decryption is performed the same way. To generate the key stream, the cipher makes use of a secret internal state which consists of two parts:

1. A permutation of all 256 possible bytes.
2. Two 8-bit index-pointers (denoted by "i" and "j").

The permutation is initialized with a variable length key, typically between 40 and 256 bits using the key scheduling algorithm (KSA). Once this has been completed, the stream of bits is generated using the pseudo-random generation algorithm (PRGA).

The selective encryption algorithm adopts the RC4 stream ciphers, which are symmetric key ciphers where plaintext bits are combined with the pseudorandom cipher bit streams by the exclusive-or (xor) operation. The plaintext T can produce a ciphertext C used key k and function E by the expression (1).

$$C = T \oplus k \quad (1)$$

In selective encryption algorithm, LTC (T_1, T_2, T_3 andT_4) can be encrypted, which can be freely combined to encrypt. The encryption process can be described as follows

$$C = (T_1 \oplus k_1) + (T_2 \oplus k_2) + (T_3 \oplus k_3) + (T_4 \oplus k_4) \quad (2)$$

Where k_1, k_2 and k_3 are the keys of T_1, T_2 and T_3 respectively. In order to keep the statistical characteristics of entropy coding and compression ratio, the LTC cannot be applied if its value is equal to 0, as given in expression (3). The function F (E) denotes the bit encryption operation done whether or not.

$$F(E) = \begin{cases} \text{do nothing,} & \text{if } T_1 \mid T_2 \mid T_3 \mid T_4 = 0 \\ \text{encryption else} & \end{cases} \quad (3)$$

Encrypting LTC except 0 is suitable for real-time video application, but there may be a fatal mistake. If the ciphertext is just equal to 0, decoder will skip the data and not decrypt it. Decryption stream key will be in confusion, but the slight difference in the keys will cause great difficulties for authorized users to decoder correctly. To solve this problem, a new method is given. The encryption and decryption implementation are described as expression (4), where T_i and C_i denote the ith of plaintext and ciphertext respectively, and k_i is the ith key of the stream cipher.

Encryption:

$$C_i = E_k(T_i)$$
$$T_i = C_i \oplus k_i \quad (4)$$
$$\text{If } C_i = 0$$

Decryption:

$$C_i = C_i \oplus k_i \quad (5)$$
$$T_i = C_i$$
$$\text{If } T_i = 0$$

Once encoder finds that the ciphertext value is equal to 0 after encryption, and then recovers it to the plaintext used by key and original ciphertext value. When decoder decrypts the nonzero number and finds that plaintext is 0, the transmission ciphertext is the correct plaintext.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed algorithm, we use QCIF standard sequence for testing. Considering the low bit rate and low delay constraint under mobile application, the baseline profile of H.264 is used. The format synchronization information is not changed during the encryption process. Thus, any standard H.264 players can decode and play the encrypted video stream.

For the experimental results, three benchmark video sequences have been used for the analysis in QCIF format. Each of them represents different combinations of motion (fast/slow, pan/zoom/rotation), color (bright/dull), contrast (high/low) and objects (vehicle, people, buildings). The video sequences 'bus', contain camera motion while 'foreman' contain high luminance images with smooth motion. 'Mobile' sequence contains a complex still background and foreground motion.

4.1 Security Analysis

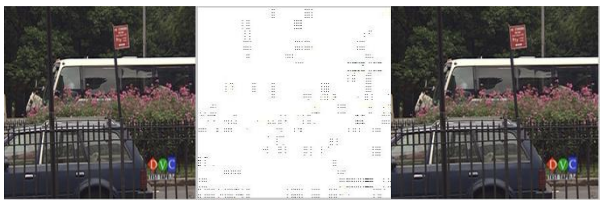
The chief aim of encryption is to make the video unintelligible for unauthorized users. We use two ways to evaluate the effect of encryption, decoder video visual quality and PSNR.

4.1.1 Video Quality Assessment

Taking video sequence of bus for example, Fig. 4.1 illustrates decoded visual quality of the bus sequences in the original frame, corresponding encrypted and decrypted frames.

Original Frame 1 Encrypted Frame 1 Decrypted Frame 1

Fig 2: Sample frame no.1



Original Frame 3 Encrypted Frame 3 Decrypted Frame 3

Fig 3: Sample frame no.3

From the point of visual quality of the encrypted images, these are so indistinct that cannot be identified. RC4 is a very powerful standard cipher and no practical attack has been published on RC4 which also prevents a timing analysis attack when the key is 128bits.

The security of the proposed algorithm is checked against most common ways of attacking the videos like ciphertext-only attack and known-plaintext attack. Ciphertext-only attack is the most difficult attack since the cryptanalyst has access only to the encrypted data. For a video frame of size 176 × 144, the number of luminance macroblocks would be 11×9, the computational complexity of breaking the first step would be $511^{11*9*10}$ where 10 specify the number of XOR operations per block. Hence the overall cost of XOR and permutation makes breaking the video file practically infeasible, making the proposed technique robust to ciphertext-only attack. In the case of known plaintext attack, the unauthorized user has the original video, the corresponding encrypted video and the encryption algorithm. Though PRNG is less secure to known-plaintext attacks, renewing of the key has been done at periodic intervals. Hence it is secure against known-plaintext. Meanwhile, the scheme is secure against exhaustive attack, since it is almost impossible to completely get plaintext which requires about 2^{128} attacks.

4.1.2 The peak signal-to-noise ratio (PSNR)

The peak signal-to-noise ratio (PSNR) is generally used to measure the quality of reconstructed images that have been compressed. Each picture element (pixel) has a color value that can change when an image is compressed and then uncompressed.

To demonstrate the efficiency of our proposed scheme, we have compressed 100 frames as INTRA of each sequence at 30 fps. The encrypted video frames at different quantization parameter (QP) values of bus, foreman and mobile video sequence and their PSNR values are given in Table 1, Table.2 and Table 3 respectively. They are compared with the PSNR obtained for the same video frames without encryption. One can note that with increase in QP, the quality of the encrypted video increases.

Table 1. Comparison of PSNR without encryption (WE) and with selective encryption (SE) for bus sequence at different QP values

QP	PSNR(dB)					
	(Y)		(U)		(V)	
	WE	SE	WE	SE	WE	SE
18	44.28	7.81	45.22	25.92	46.51	25.62
24	38.41	8.56	41.52	27.81	43.24	25.21
30	34.13	9.11	39.17	29.22	40.75	27.09
36	31.58	9.98	36.76	36.63	37.80	31.04

Table 2. Comparison of PSNR without encryption (WE) and with selective encryption (SE) for foreman sequence at different QP values

QP	PSNR(dB)					
	(Y)		(U)		(V)	
	WE	SE	WE	SE	WE	SE
18	44.45	8.60	45.63	25.60	47.42	24.92
24	38.41	8.56	41.52	27.81	43.24	25.21
30	34.93	9.40	39.38	30.01	41.00	26.09
36	30.78	10.31	37.33	37.33	38.10	30.22

Table 3. Comparison of PSNR without encryption (WE) and with selective encryption (SE) for mobile sequence at different QP values

QP	PSNR(dB)					
	(Y)		(U)		(V)	
	WE	SE	WE	SE	WE	SE
18	44.46	8.11	44.14	18.82	44.05	14.17
24	39.58	8.92	41.52	28.64	42.96	28.38
30	35.14	9.14	39.54	31.71	41.24	29.52
36	30.95	10.15	37.08	37.02	38.19	32.22

Comparing the PSNR of all benchmark video sequences at QP value '18' without encryption and with SE, the proposed method is computationally very efficient. Although it depends on the contents of video and the quantization value, but in proportion to the overall computation which a video codec consumes, it is negligible.

4.2 Compression Ratio

Maintaining a good compression ratio is one of vital goals in real-time video system. In this selective video encryption algorithm the number of bit streams obtained after entropy coding are not affected. So there is no change between the compression ratio of encrypted video and unencrypted video. Considering the proposed selective encryption scheme adequately, bit rate is not increased at all.

4.3 Computational complexity

Considering the limited power of mobile device, the algorithm uses stream ciphers which typically execute at a higher speed than block cipher and have lower hardware complexity. Furthermore, the important data in coding process are selected to encrypt, so the computational complexity is rather low to satisfy the mobile application.

5. CONCLUSION

In this paper, in order to provide secured multimedia service for mobile device and wireless application, the selective video encryption algorithm has been proposed based on the special feature of the H.264 video coding standard. The luminance transform coefficients of residual data are selectively encrypted by a stream cipher. Experimental results demonstrate that the proposed algorithm can achieve good security and high efficiency with low complexity. Selective video encryption algorithm is thus suitable for securing multimedia services for wireless application.

6. REFERENCES

[1] C Wang, H-B. Yu, M Zheng. A DCT-based MPEG-2 transparent scrambling algorithm. IEEE Trans. Consumer Electrons. 2003, 49 (4), 1208-1213.

[2] A. Servetti, J. C. De Martin. Perception-based partial encryption of compressed speech. IEEE Trans. Speech Audio Process. 2002, 10 (11), 637-643.

[3] Jinhaeng Ahn, et al. Digital Video Scrambling Method Using Intra Prediction Mode [A], PCM 2004, LNCS vol.3333, 2004:pp. 386-391.

[4] Zheng Liu, Xue Li. motion vector encryption in multimedia streaming. Proceedings of the 10th International Multimedia Modeling Conference. 2004, 1-8.

[5] P. Melih and D. Vadi. A MPEG-2-transparent scrambling technology. IEEE Trans. Consumer Electron. 2002, 48(2), 345-355.

[6] Jidong Wang,et al. A Partial Scramble Scheme for H.264, Video[C] ASIC, 2007. ASICON '07. 7th International Conference on 22-25 Oct. 2007 Page(s):802 –805.

[7] Wang Yajun, Cai Mian, Tang Feng. Design of a New Selective Video Encryption Scheme Based on H.264. Computational Intelligence and Security, 2007 International Conference on15-19 Dec. 2007 Page(s):883 – 882.

[8] C. Narsimha Raju, Kannan Srinathan and C. V. Jawahar, A real-time video encryption exploiting the distribution of the DCT. TENCON 2008. IEEE Region 10 Conference 19- 21 Nov. 2008 Page(s):1 – 6.

[9] Wu C P, Jay Kuo C... Design of integrated multimedia compression and encryption systems. IEEE Trans. On Multimedia. 2005, 7(5), 828-839.

[10] Wen J, Severa M, Zeng W, Luttrell M H, Jin W. A format-compliant configurable encryption framework for access control of video. IEEE Transactions on Circuits & Systems for Video Technology. 2002, 12, (6): 545-557.

[11] Zhou J, Liang Z, Chen Y, Au O C. Security analysis of multimedia encryption schemes based on multiple Huffman table. IEEE Signal Processing Letters, 2007, 14(3): 201-204.

[12] Yinian Mao and Min Wu. A Joint Signal Processing and Cryptographic Approach to Multimedia Encryption. IEEE Trans. on image processing. 2006, 15(7), 2061-2075.