Face Protection by Fast Selective Encryption in a Video

J. M. Rodrigues, W. Puech, P. Meuel, J.C. Bajard and M. Chaumont

LIRMM Laboratory, UMR CNRS 5506, University of Montpellier II
161, rue Ada, 34392 MONTPELLIER CEDEX 05, FRANCE
jose-marconi.rodrigues@lirmm.fr, william.puech@lirmm.fr,
bajard@lirmm.fr, peter.meuel@lirmm.fr, marc.chaumont@lirmm.fr

Keywords: Selective encryption, crypto-compression, advanced encryption standard.

Abstract

In this paper we propose an approach for monitoring human activities in an indoor environment using detected and tracked face and selective encryption. The objective is to encrypt partially the human face in a video sequence. This approach is based on AES stream ciphering using VLC (Variable Length Coding) of the Huffman's vector. The proposed scheme allows scalable decryption of a specific region of image and result in a significant reduction in encrypting and decrypting processing time. It also provides a constant bit rate and keep the JPEG and MPEG bitstream compliance.

1 Introduction

There are two sorts of demand for secure multimedia transmission the full encrypted and the selective encryption. Military and law enforcement for example require full encryption. However, there is a huge spectrum of applications that demands security on a lower level, *i.e.* partial or selective encryption (SE). SE is an approach to reduce the computational requirements in networks with different client device capabilities [4]. One example of SE application in security field is the images acquired by a survey camera. For various reasons, this kind of images must be quickly transmitted and no full encryption is necessary. The security of SE is always lower when compared to full encryption. We have a trade-off between the amount of encrypted data and the necessary available time and memory.

In this paper we propose a new approach for SE of the Huffman coding for video sequence. In our approach we employ the Advanced Encryption Standard (AES) [1] cipher in the Output Feedback Block (OFB) mode, this allows versatility at the decoding stage. In our work the survey camera is fixed, thus it is easier to follow the human motion in the video sequence.

In Section 2, we review the previous research findings. In Section 3, we introduce the proposed method, the selective encryption algorithm as well as the detecting and tracking of human skin process. Finally in Section 4 we show the experimental results when we apply our algorithm in sequence of images and provide the conclusions of this study in Section 5.

2 Previous Works

Several selective encryption methods have been used in compressed images using DCT-based algorithms. Droogenbroeck and Benedett [2] originated encrypt a selected number of AC coefficients. In their method the DC coefficients are not ciphered because they carry important visible information and they are highly predictable. Moreover, in their approach the compression and encryption stages are separated and that requires an additional operating cost. Fisch et al. [3] have proposed a partial image encryption where the data are organized in a scalable bit-stream form. These bit streams are constructed from the DC and some AC coefficients from every image block and then arranged in layers according to their visual importance. Recently A. Said [6] has measured the strength of partial encryption showing attacks that exploit information from non-encrypted bits and availability of side information.

2.1 Selective Encryption of DCT based Images

In the Huffman coding block, the quantized coefficients are coded by the pair {(HEAD), (AMPLITUDE)}. The HEAD contains the controllers provided by the Huffman's tables. The AMPLITUDE is a signed-integer that is the amplitude of the nonzero AC, or in the case of DC is the difference between two neighbor DC coefficients. For the AC coefficients the HEAD is composed by (RUNLENGTH, SIZE), while for the DC coefficients it is made up only by SIZE. Because DCs are highly predictable, they are treated separately, in the Huffman coding. The method depicted in this paper is essentially based on encrypting of some AC coefficients.

For the AC coding, it is made a combining run-length and amplitude information. It aggregates zero coefficients into runs of zeros. RUNLENGTH is a consecutive number of zero-valued AC coefficients which precede nonzero-value in the zigzag sequence. The SIZE is the amount of necessary bits to represent the AMPLITUDE. Two extra codes that correspond to (RUNLENGTH, SIZE) = (0, 0) and (15, 0) are used for symbolize EOB (End-Of-Block) and ZRL (Zero Run Length) respectively. The EOB is transmitted if the last nonzero

coefficient is followed by zeros in the quantized block. The ZRL symbol is transmitted whenever RUNLENGTH is greater than 15 and represents a run of 16 zeros.

2.2 The Advanced Encryption Standard Algorithm

Advanced Encryption Standard (AES) is very powerful standard cipher that operates by performing a set of steps, for a number of iterations called rounds. The enciphering of a plain text X_i in AES is described in Fig. 1. The AES algorithm can support several modes such as Electronic CodeBook (ECB), Cipher Block Chaining (CBC), Output FeedBack (OFB), Cipher FeedBack (CFB) and Counter (CTR). In OFB mode, object of this work, the ciphertext block Y is produced by performing a XOR with Z_i, where $Z_i = E_k(Z_{i-1}), i \geq 1$ and $Y_i = X_i \oplus Z_i$, as illustrated Fig. 2.

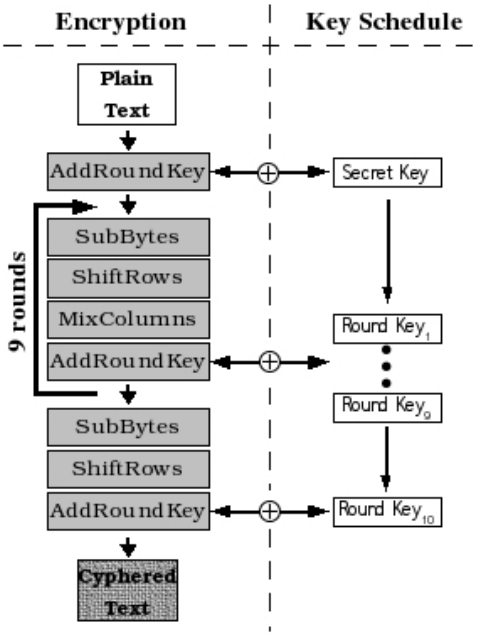


Figure 1: AES general outline.

Although AES is a block cipher, in the OFB, CFB and CTR modes it operates as stream cipher. Each mode has advantages and drawbacks as well. In ECB and OFB modes for example any modification in the plaintext block causes the corresponding ciphered block to be altered, but other ciphered block are not affected. On the other hand, if a plaintext block is changed in CBC and CFB modes, then all subsequent ciphered blocks will be affected. These properties mean that OFB mode treats separately each block. From Fig. 2, we can observe that the encryption function $E_k()$ is used for both encryption and decryption in the OFB mode.

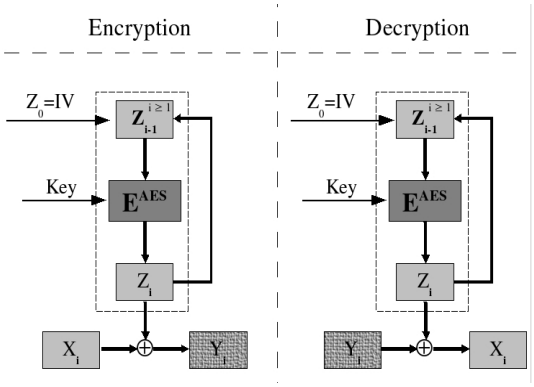


Figure 2: Encryption and decryption in OFB mode.

3 The Proposed Method

Let $E_k(X)$ be the encryption of a n bit data block X using the secret key k with AES cipher in the OFB mode. For practicality, we assume that $n = 128$ and that X is a plaintext. Let $D_k(Y)$ be the decryption of a ciphered text Y using the secret key k.

3.1 The selective encryption of DCT coded images

The proposed method for selective encryption has three stages and is applied in the entropy encoding stage during the creation of the Huffman's vector. The three steps are the construction of the plaintext X_i, the ciphering of X_i to create Y_i and the substitution of the original Huffman's vector with the ciphered information. These operations are performed separately in each quantized DCT block. The following procedures are applied for each block. The homogeneous blocks are not ciphered due to the high quantity of zero AC coefficients.

3.1.1 The construction of plaintext X

For constructing the plaintext X_i, we consider the non-zero AC coefficients of the current block i by accessing the Huffman's vector to create the {HEAD, AMPLITUDE} pairs. From each HEAD number is extracted the length of AMPLITUDE. These values are tested according to the following equation :

$$f(\rho) \;\leq\; L_{X_i} \;\leq\; C, \tag{1}$$

where ρ is the homogeneity of the block, $f(\rho) = 0$ for $\rho \to \infty$ and $C \in \{128,\ 64,\ 32,\ 16,\ 8,\ 4\}$ bits.

As shown in Fig. 3, only the AMPLITUDE's $(A_n, A_{n-1}...A_1)$ are considered to build the vector X_i. The final plaintext length L_{X_i} depends on both the homogeneity of the block ρ and the given constraint C. The constraint C specifies the maximum

number of bits that must be considered in each block. For this work we use $C = 128$. Then, we apply the padding function $p(j) = 0$, where $n \geq j > L_{X_i}$, to fill in the vector X_i with zeros.

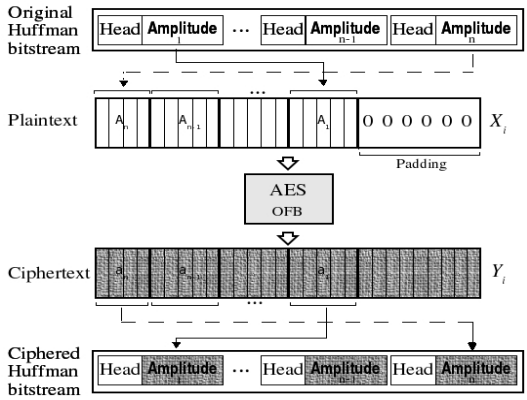


Figure 3: Global overview of the proposed method.

3.1.2 Ciphering of X in the OFB mode of the AES

In the ciphering step, the plaintext X_i is XOR-ed with the output of AES cipher in order to create Y_i. The vector IV ($Initialization\ \ Vector$) is created from the secret key k according to the following algorithm. The secret key k is used as the seed of PRNG (Pseudo-Random Number Generator). The k is divided into 16 portions of 8 bits each. The PRNG produces 16 random numbers that define the order of formation for the IV vector. For example, if the first random number generated is 7, the first byte of the secret key will be copied in the seventh element of the vector IV. After generating the vector $IV = Z_0$, it is ciphered firstly to produce Z_1, then ciphered to produce Z_2, afterward Z_3 and so on, as shown in Fig. 2. Z_i is XOR-ed with the plaintext X_i producing Y_i. The use of OFB mode for ciphering purpose allows for an independent generation of the keystream Z_i.

3.1.3 Substitution of the original Huffman's bitstream

The final step is the substitution of the ciphered information in the Huffman's vector. As in the first step (construction of the plaintext X_i), the Huffman's vector is accessed in sequential order, while the ciphered vector Y_i is accessed in the inversed order. Given the length in bits of each AMPLITUDE $(A_n, A_{n-1}...A_1)$, we start replacing these parts of Y_i with the AMPLITUDE in the Huffman's vector. The total quantity of replaced bits is L_{X_i}.

3.2 The Decryption Procedure

The decryption process proceeds as follows. The secret key is used to build the vector $IV = Z_0$. Z_1 is created by encoding IV, while Z_n is created by encoding Z_{n-1}. Therefore, the same procedure for encryption as described in the previous section is employed at the decryption. The difference consists in the fact of the entry of the ciphering process is the ciphered Huffman's vector. This ciphered vector is also accessed in inversed order of its bits in order to construct the plaintext Y_i. Then, Y_i will be used together with Z_i in AES decoding procedure as shown in Fig. 2. The resulted plaintext vector is split in order to substitute the AMPLITUDE in the ciphered Huffman coding and to generate its corresponding original Huffman vector.

3.3 Detection and tracking of human skin

There is a large spectrum of utilization of SE. The method depicted in Section 3.1 can be applied in a video sequence to selectively encrypt the human face.

The first step of the algorithm is the color transformation from RGB to YCbCr space. Then, we calculate the Discrete Cosine Transform to each component (Y, Cb and Cr) in blocks of 8×8 in order to generate the DC and AC coefficients. We use the DC coefficients of the Cr and Cb components to produce small images that will be used for detecting the human skin according to:

$$\text{if } \sqrt{\left(\frac{DC_{Cr}}{8} - Cr_s\right)^2 + \left(\frac{DC_{Cb}}{8} - Cb_s\right)^2} < S \quad \text{then } \textit{skin detected}, \tag{2}$$

where Cb_s and Cr_s are the reference skin color in $YCbCr$ space and S is the threshold. With a quality factor of 100 %, the values $\frac{DC_{Cr}}{8}$ and $\frac{DC_{Cb}}{8}$ correspond to the mean values of the Cr and Cb component blocks.

The result of the detection step is generally a noisy binary image. In order to remove the noise, we have applied an algorithm of erosion and dilatation [8, 7]. In our approach the survey camera is fixed, thus it makes easy to follow the motion of target in the video sequence.

The denoised binary images address the region that must be encrypted in the original image. Each white pixel in the binary image corresponds to a block of 8×8 pixels in the original image. Finally, the SE method, depicted in Section 3.1, is applied in the Huffman vector of the component Y. It is important to denote that the selective encryption employed in the Huffman vector will not change the final sizes of the original images of the sequence.

4 Experimental Results

For our experiments, we have selected four images $S_4 = (\#083, \#123, \#135 \text{ and } \#147)$ gotten from a sequence of one hundred eighty-six images. Each one of them is in the JPEG format in the baseline sequential mode with a Quality Factor (QF) of 100%. For the encryption we have used the AES cipher in (OFB) stream cipher mode with a key of 128 bits long.

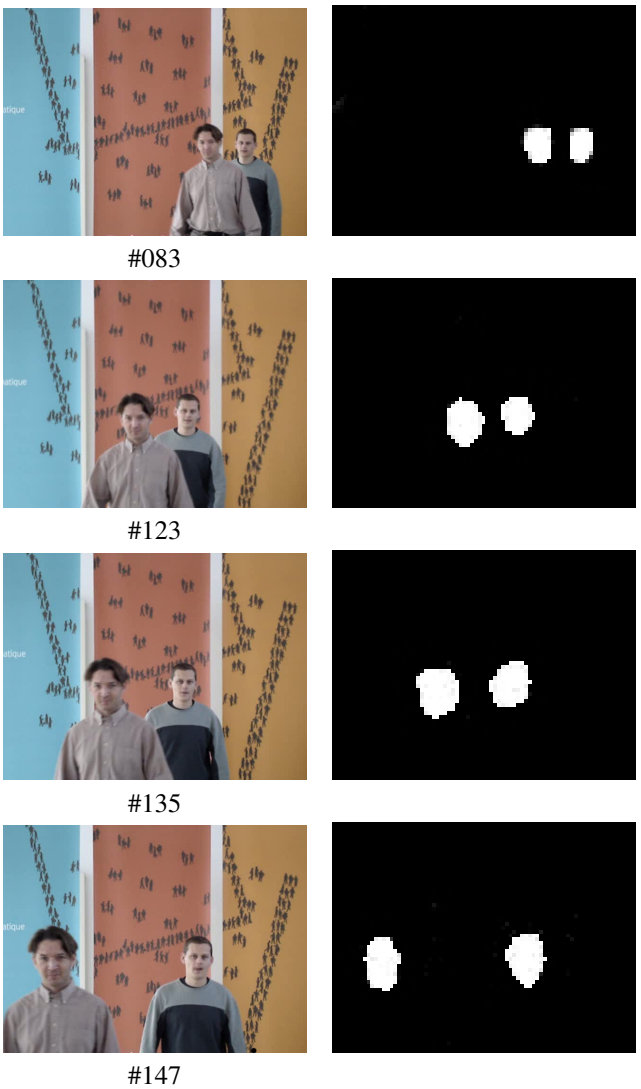


Figure 4: a) Sequence of original images, b) Sequence of binary images.

From each original RGB image, 640×480 pixels, of the sequence S_4, Fig. 4.a, we have generated the $YCbCr$ components as shows as example the Fig. 5. From the DCs of the components Cb and Cr, we have applied the Equation 2 to obtain the binary images. In these binary images we have employed an algorithm of erosion/dilatation to obtain the

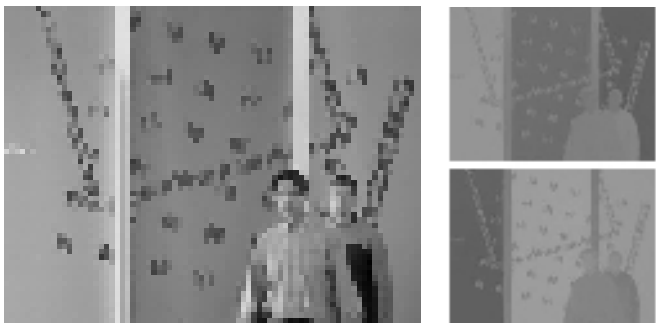


Figure 5: YCbCr of the DC coefficients of image #083.

neat binary image illustrated in Fig. 4.b. The detection of the human skin region, in this case the face is represented by the white pixels. We have mapped one white pixel in then binary image as a block of 8×8 pixels in the original image. Finally, we have applied the SE method depicted previously to generate the selectively encrypted images.

	Total ciphered			
Image	Quant. Blocks	Coeffi cients	Bits	% Block
#083	79	2547	10112	1.65
#123	113	3042	14464	2.35
#135	159	4478	20352	3.31
#147	196	5396	25088	4.08

Table 1: Results of SE employed in sequence of images.

The Table 1 shows the values yielded in each image. For the image #083 we had 79 blocks tracked and detected. That means 2547 AC coefficients encrypted and 10112 bits in the original image. The quantity of block encrypted corresponds to 1.6 % of the total amount of block of the original image. For the image # 123 we have gotten 113 blocks, this was expected because the faces are closer and bigger. We have encrypted 3042 AC coefficients and that represents 14464 bits and 2.35 % of blocks encrypted. The quantity of block for encryption increases because in our example the two people are getting closer to the camera. Analyzing the Table 1, we can conclude that the amount of bits encrypted is very small related with the size of whole image. That makes our method applicable in low power environment like video camera.

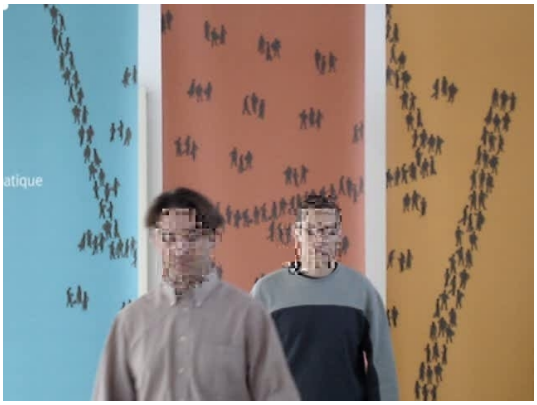
The Fig. 6 shows the final results of tracking, detecting and selectively encrypting the sequence of images.

In order to clearly show our results, we have cropped from image # 123' a sub-image of 216×152 pixels, Fig. 7.

It should be noted that, the security is linked to the ability to guess the values of the encrypted data. For example, from a security point of view, it is preferable to encrypt the bits that look the most random. However in practice, the trade-off is

#083’

#123’



#135’

#147’

Figure 6: Sequence of selectively encrypted images.

(a)

(b)

Figure 7: Region of 216×152 pixels of image #123: a) Original image, b) Protected image.

more difficult to define because the most relevant information, like DC coefficients in a JPEG encoded image, usually are highly predictable [2, 5].

5 Conclusion

In this paper, we have proposed a new scheme for selective encryption of JPEG image sequences by using the AES cipher in OFB mode. Our approach brings several advantages such as portability, constant bit rate, JPEG format compliance, scalable selective encryption and a progressive decryption of the region of interest. By using the proposed algorithm we are able to encrypt an image without affecting at all the JPEG compression rate. In the decoding stage we are able to decrypt selected areas, which makes the proposed method useful for a large range of applications. The experiments have shown that our scheme provides satisfactory image PSNR, sufficient visual security and good information confidentiality results.

References

[1] J. Daemen and V. Rijmen. AES proposal: The rijndael block cipher. Technical report, 2002.

[2] M. Van Droogenbroeck and R. Benedett. Techniques for a selective encryption of uncompressed and compressed images. In *ACIVS'02*, Ghent, Belgium, 2002. Proceedings of Advanced Concepts for Intelligent Vision Systems.

[3] M. M. Fisch, H. Stgner, and A. Uhl. Layered encryption techniques for dct-coded visual data. In *EUSIPCO*, Vienna, Austria, September 2004.

[4] X. Liu and A. Eskicioglu. Selective encryption of multimedia content in distribution networks:challenges and new directions. In *IASTED Communications, Internet & Information Technology (CIIT), USA*, November, 2003.

[5] T. Lookabaugh. Selective encryption,information theory, and compression. In *ASILOMAR*, pages 373–376. Conference on Signals, Systems and Computers, 2004.

[6] A. Said. Measuring the strength of partial encryption scheme. In *ICIP'05*, pages 1126–1129, Genova, Italy, September 2005. The IEEE Inter. Conf. on Image Processing.

[7] J. Serra. *Image Analysis and Mathematical Morphology, vol. 2*, volume 2. London: Academic Press, 1988.

[8] M. Van Droogenbroeck and M. Buckley. Morphological erosions and openings: fast algorithms based on anchors. *Journal of Mathematical Imaging and Vision*, 22(2-3):121–142, May 2005. Special Issue on Mathematical Morphology after 40 Years.