

Secure and low cost selective encryption for JPEG2000

Ayoub Massoudi & Frédéric Lefèbvre
Thomson R&D France, Security Labs
{ayoub.massoudi
frederic.lefebvre}@thomson.net

Christophe De Vleeschouwer & François-Olivier Devaux
UCL, TELE Laboratory, Belgium
{christophe.devleeschouwer
francois.deviaux}@uclouvain.be

Abstract

Selective encryption is a new trend in content protection. It aims at reducing the amount of data to encrypt while achieving a sufficient and inexpensive security. This approach is particularly desirable in constrained communication (real time networking with delay constraints, mobile communication with limited computational power...). In this paper we introduce selective encryption from information theory point of view. We define a set of evaluation criteria for selective encryption algorithms and propose a novel selective encryption algorithm for JPEG2000 compressed images satisfying all these criteria. The main contribution of this proposal consists of reaching the minimum amount of data to encrypt regarding a given level of security and target application requirements. For this purpose, we exploit the R-D optimization performed by JPEG2000 EBCOT algorithm.

1. Introduction

In traditional content access control, Shannon [15] suggested the fully layered scheme. It consists of compressing data (preferably with a perfect compressor that removes all source redundancies) and then encrypting the whole bitstream. In this configuration, all plaintext symbols or bits are assumed to be of equal importance. This is relevant when the transmission of the content is unconstrained. In situations where only few resources are available (high transmission rate, low memory, low power or computation capabilities), this approach seems inadequate as pointed out by Shamir [12]. Recent works explored a new way for content protection, named partial encryption or selective encryption where only parts of the data are encrypted. In [10], T. Lookabaugh pointed out the close link between selective encryption and Shannon's work on communication and security [15]. Shannon highlighted the relationship between source statistics and the ciphertext security [15]. A secure encryption scheme should remove all plaintext redundan-

cies so that no exploitable correlation is observed in the ciphertext. He introduced the unicity distance as the minimum amount of ciphertexts required to permit a computationally unlimited adversary to recover the unique encryption key. It is given by:

$$n_u = \frac{H(k)}{r} \quad (1)$$

Where $H(k)$ is the key entropy and r the plaintext redundancy. Hence, the less redundant compressed data is, the more secure the ciphertext is. In perfect compression configuration, given a plaintext P , let P' be its "perfect" compression. We split P' into two parts P'_1 and P'_2 . Let C_1 and C_2 be the ciphertexts of P'_1 and P'_2 (Figure 1). Perfect compression implies that if we know only P'_1 , then P'_2 is completely unpredictable [10]. Thus, if we consider that only a subset of the compressed data is encrypted (P'_1 or P'_2), the security of the ciphertext is preserved. By encrypting only P'_1 , we get a selective encryption scheme (Figure 2). This result is fundamental in building a cryptographically secure selective encryption algorithm as will be described in this paper.

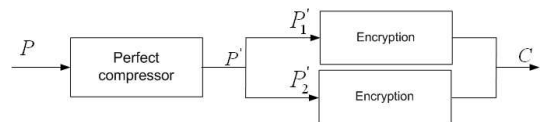


Figure 1. Fully layered system as suggested by Shannon [15]

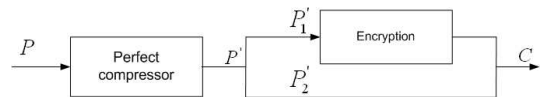


Figure 2. Post compression selective encryption preserves security if a perfect compressor is used

To evaluate and compare selective encryption algo-

gorithms, we propose a set of evaluation criteria.

1.1. Evaluation criteria

The following list contains evaluation criteria often encountered in the literature, as well as new criteria proposed in this work:

Configurability (C) Regarding a target application, configurability allows to dynamically fine tune the encryption algorithms with different encryption parameters.

Visual degradation (VD) Measures the perceptual distortion (preferably configurable) of the cipher image with respect to the plain image. The PSNR is widely used as a metric for this criterion.

Cryptographic security (CS) Many papers on selective encryption evaluate security based only on visual degradation. However, in [17], high visual degradation is achievable while important security leakages were pointed out in [8]. In section 2.1, we define cryptographic security.

Encryption ratio (ER) This criterion measures the ratio between the size of the encrypted part and the whole data size. It is one of the main expected features of selective encryption. It is required to be as small as possible.

Compression friendliness (CF) A selective encryption algorithm is considered compression friendly if it has no or very little impact on compression efficiency.

Format compliance (FC) Any standard decoder should be able to decode the encrypted bitstream without decryption. Format compliance allows preserving some features of the compression algorithm (for example scalability).

Error tolerance (ET) This criterion is not often considered in the literature. The challenge is to design a secure selective encryption algorithm with limited error propagation in case of transmission errors.

1.2. Related work

We can classify selective encryption algorithms in three categories: before compression (Pre), during (In) or after (Post). Figure 3 lists selective encryption algorithms regarding each criterion described above. The challenge is to trade-off all the aforementioned criteria. Most of the pre-compression algorithms are format compliant [20, 17]. [20] Proposes selective bit scrambling and blocks shuffling at lower resolutions of the DWT pyramid. The algorithm does not allow configurability since only full confidentiality with high visual degradation is achievable. In addition, this approach requires high encryption ratio of about 20%. In-compression approaches adversely impact compression

performance [18, 9, 4], this is basically due to modification of the encoder. [4] Proposes a JPEG2000 lightweight encryption based on secret randomized anisotropic wavelet bases. Large key space is achievable yielding high security level. However, the proposed method is only applicable to transparent encryption with low resolution preview. Compression performance is adversely impacted and the full bitstream is not format compliant. Finally, most post-compression schemes are compression friendly [2, 13, 3]. However, none of the proposed methods offer configurable algorithm with tunable visual degradation. [13] selectively encrypts packet data of JPEG2000 codestream. At least the first 20% of the data is encrypted to achieve full confidentiality, this proposal is applicable only to resolution progression and offers no configurability. In [3], the authors suggest to encrypt packet headers in addition to packet data. They argue that packet headers transport discriminant information that can be considered as a fingerprint. This proposal is relevant for full confidentiality where all JPEG2000 packets are encrypted. In a selective encryption scheme, unencrypted part can be considered as a fingerprint. Therefore, encrypting packet headers would increase encryption ratio without significant improvement on security. In order to achieve JPEG2000 format compliance, [16] performs iterative CCP (Codeblock Contribution to Packet) encryption. The number of iterations increases exponentially with CCP length and iterative approach may give a hint to perform side channel attacks.

An important effort has been also made in Part 8 of JPEG2000, namely JPSEC or secure JPEG2000 [7] to provide a standardized framework to implement security tools and services such as selective encryption, authentication, integrity... Our proposal could be implemented in JPSEC framework.

Domain	Ref	C	VD	CS	ER	CF	FC	ET
Pre	[20]	-	H	-	≈20%	≈5%	+	+
	[17]	-	V	-	V	+	+	-
In	[18]	-	H	+	≈15%	-	+	-
	[9]	+	V	-	V	-	+	-
	[6]	-	H	-	?	+	+	-
	[4]	-	L	+	L	-	-	-
Post	[2]	-	H	-	L	+	-	-
	[19]	-	H	+	L	-	-	-
	[13]	-	H	-	20%	+	-	-
	[16]	-	?	-	?	-	+	+
	[3]	-	H	-	?	+	+	-
	Our method	+	V	+	minimized	<5%	+	+

Figure 3. State of the art selective encryption algorithms.

None of the previously proposed schemes tackle all the criteria described above. In section 2 we propose a novel scheme offering a trade-off between these criteria (table 3). In section 3, we present experimental results and comparisons regarding evaluation criteria. We conclude in section 4.

2. Our selective encryption algorithm

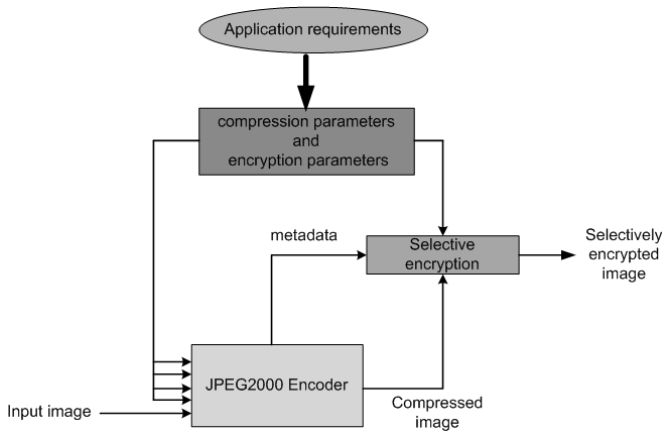


Figure 4. General approach.

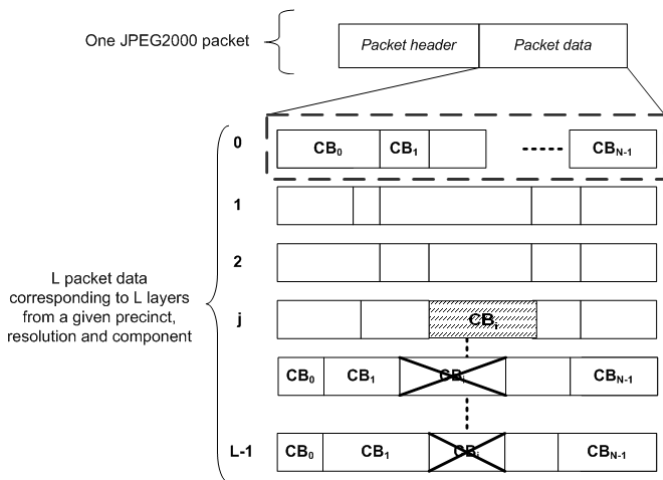


Figure 5. JPEG2000 packet structure. By encrypting contribution of code-block CB_i at a given layer (j), all its contributions to less significant layers become undecodable.

A JPEG2000 codestream is composed of packets; each packet contains the data from a given Resolution (R), Quality Layer (L), Spatial region called precinct (P) and Component (C). Each packet contains a packet header followed by the packet data (figure 5). Packet data contains the the code-

block contributions (CCPs) of each packet (figure 5). Depending on the target application, a subset of these packets is encrypted (figure 4). We define the following encryption parameters:

R_e : List of resolutions to be encrypted. Low resolution subbands concentrate signal energy. High resolution subbands represent details. If high visual degradation is required with no preview, it is recommended to encrypt only low resolution subbands. If low visual degradation is required with a "thumbnail" preview, it is recommended to encrypt only highest resolution subbands.

L_e : List of layers to be encrypted. If a low quality preview is required, we need to encrypt least significant layers. If high visual distortion is required, most significant layers need to be encrypted.

C_e : List of components to be encrypted.

P_e : List of precincts to be encrypted. If a particular region of the image needs to be encrypted, then all packets from precincts that cover the corresponding area should be encrypted.

Each packet in the set $S_e = R_e \times L_e \times C_e \times P_e$ is selectively encrypted. Each packet is composed of independent code-block contributions. Only the minimum amount of bytes in each code-block contribution of each packet from S_e is encrypted. Since CCPs have different lengths, additional data is required. It is included within the bitstream (for example in the COM marker segment of JPEG2000 codestream or the SEC marker segment of a JPSEC main header) to indicate encrypted packets S_e and CCPs lengths.

The next subsections detail how we have computed the minimal number of bits to encrypt in each code-block contribution (section 2.1), how to select code-blocks to encrypt in order to minimize encryption ratio (2.2) and finally the encryption algorithm (2.3).

2.1. Cryptographic security

Very few papers have proposed a serious evaluation of the security of selective encryption algorithms. In most cases, visual distortion (measured using the PSNR) is used as the exclusive criterion for such purpose. However, visual degradation remains a subjective measure. In addition, it has been shown that some selective encryption algorithms that yield important visual distortion may have important security leakages [17, 8]. Cryptographic security should rely on:

- The encryption key (of a well scrutinized encryption algorithm).

- Unpredictability of the encrypted part.

In this work, we use AES-128 as a well-scrutinized encryption algorithm. The only "successful" attacks against AES have been side channel attacks [14]. These attacks should be blamed on the implementation leakages and not to the architecture of AES block cipher. Very few works have been reported on the unpredictability of the encrypted part. In [11], guesswork is used as a measure to evaluate the confidentiality of selectively encrypted messages. It estimates the expected number of guesses an attacker should try before finding the right secret. We investigate the implications of these results on selective encryption of JPEG2000 compressed images. To a certain extent, the JPEG2000 arithmetic coder can be considered as a "perfect compressor". Hence, we can consider that the outputs of this coder are uniformly distributed. Codeblocks are the fundamental coding units in JPEG2000 and since each code-block is encoded independently, we base our approach on code-blocks statistics. We empirically verified that all bytes values in code-block contributions are equally probable ($p_0 = \frac{1}{256} \approx 0.0039$) as illustrated in figure 6.

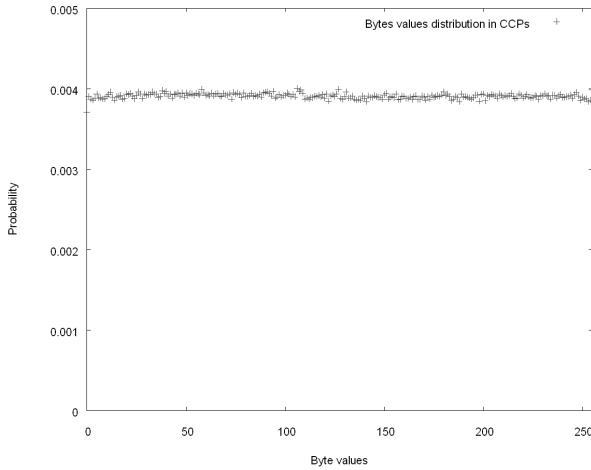


Figure 6. Bytes values of code-blocks contributions are uniformly distributed.

This observation is a necessary condition for uniform distribution assumption. We consider a message M (representing one code-block contribution) composed of n bytes. We arbitrarily choose n_e bytes that will be encrypted ($n_e \leq n$), the encrypted part is represented with message X (figure 7). The remainder of the message is left unencrypted. The encryption ratio is given by:

$$ER = \frac{n_e}{n} \quad (2)$$

We evaluate the difficulty for an attacker to predict (X) in a brute force attack and try to find conditions that make

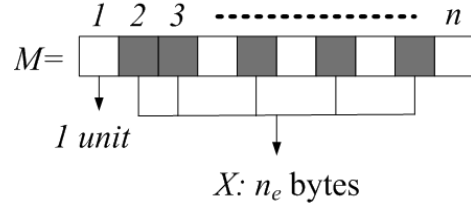


Figure 7. Selective encryption of a message, only grey units are encrypted

brute force attack on the key space easier than optimal brute force attack on the plaintext space. This condition is fundamental. Indeed, if brute force attack on the plaintext is easier, then the cipher and encryption key can be bypassed. The attacker would then prefer to concentrate his effort on guessing the right plaintext. We assume that the attacker knows the length and the location of the encrypted part and is able to recognize when a right guess occurs. For this purpose we use guesswork:

$$W(X) = \sum_{i=1}^{|L|} i \cdot p_i \quad (3)$$

Where $L = \{X_1, X_2, \dots, X_{|L|}\}$ is the language space of X and $p_i = Pr(X = X_i)$. For uniformly distributed symbols, we obtain identically distributed elements in language space which yields:

$$p_i = \frac{1}{|L|} = \frac{1}{|\Sigma|^{n_e}} \quad (4)$$

Where Σ is the alphabet of language L . We get guesswork:

$$W(X) = \frac{1}{|\Sigma|^{n_e}} \cdot \sum_{i=1}^{|\Sigma|^{n_e}} i = \frac{1 + |\Sigma|^{n_e}}{2} \quad (5)$$

Now, if we consider the guesswork on encryption key (of k bits), we get:

$$W(K) = \sum_{i=1}^{2^k} \frac{i}{2^k} = \frac{1 + 2^k}{2} \quad (6)$$

From equations 5 and 6, we can conclude that brute force attack on the message space is harder than key guessing if $W(X) \geq W(K)$. In other terms:

$$n_e \geq \frac{k}{\log_2(|\Sigma|)} \quad (7)$$

In the case of JPEG2000 selective encryption using the AES-128 algorithm, we have ($k = 128$, $|\Sigma| = 256$). This yields a minimum number of bytes to encrypt per code-block contribution:

$$n_e \geq 16 \quad (8)$$

2.2. Encryption ratio

In the previous section, we have shown that cryptographic security requires encrypting only 16 bytes in each code-block contribution from S_e . In this sub-section, we exploit the R-D optimization performed by the EBCOT algorithm in order to select code-block contributions to encrypt. Let us consider an image compressed with R resolutions (resolution 0 is the smallest (LL subband)), L layers (layer 0 is the Most Significant), P precincts and C components. From a selective encryption standpoint, if the contribution of code-block CB_i at layer j from a given precinct, resolution and component is encrypted, all its contributions to layer j and subsequent layers ($j + 1 \dots L - 1$) are no more correctly decodable (figure 5). This is equivalent to truncating the contribution of CB_i at layer $j - 1$. The distortion resulting from this encryption is given by:

$$d(E_{i,j}) = d_{i,j-1} + d'_{i,j} \geq d_{i,j-1} \quad (9)$$

Where $d_{i,j-1}$ is the distortion resulting from truncating CB_i at layer $j - 1$ and $d'_{i,j}$ is the distortion resulting from incorrect decoding of the contributions of code-block CB_i to layers $j \dots L - 1$. The context based nature of the arithmetic coder used to encode quality layers make code-block contributions causal from most significant to least significant layers. In selective encryption literature [13, 3], the distortion $d(E_{i,j})$ is commonly used to measure the security of the algorithm. However, the noisy component $d'_{i,j}$ can be canceled by enabling error resiliency mechanisms at encoding and decoding. Visual distortion is not relevant for security estimation. The distortion resulting from the decoding public part is:

$$d_e(E_{i,j}) = d_{i,j-1} \quad (10)$$

All users have access to the visual quality given by this public part. Thanks to the R-D optimization performed by the EBCOT algorithm, it is not necessary to encrypt more than the contributions of code-blocks to layer j in order to obtain a distortion greater or equal to that of layer $j - 1$. Thus only the most significant layer among L_e is encrypted. It allows an important encryption ratio reduction.

2.3. Encryption algorithm

In Section 2.1, we have shown that for each selected code-block contribution, we only need to encrypt 16 bytes. Then, in section 2.2, we have determined the set of code-blocks to encrypt. In this last sub-section, we use a pattern-constrained encryption method in order to output a format compliant encrypted bitstream. The only constraint we need to observe is that codewords within the interval $[0XFF90, 0XFFFF]$ are forbidden in packet data. A

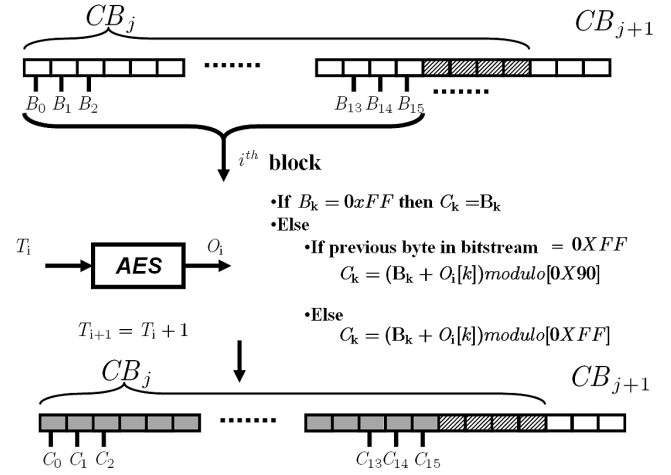


Figure 8. Only 16 bytes are encrypted per code-block contribution.

similar approach was proposed in [5] for full encryption, our approach is adapted for partial to full encryption. To that end, we use a modified AES CTR mode with conditional modular addition as shown in figure 8. For decryption, all additions are replaced by subtraction. The main advantage of this approach is to achieve format compliance without iterating many encryption cycles. This yields important time and memory saving. In addition, iterative approaches may give a hint for attackers to perform side channel attacks such as timing attack.

3. Experimental results and comparisons

For experiments, we select a set of high definition images (1850×2160). We use OpenJPEG library [1] for encoding/decoding and metadata generation. The compression parameters are five layers, five resolutions, three components and six precincts.

3.1. Visual degradation

Our approach allows fine tuning visual degradation with respect to application requirements. For illustration, we consider two tunings to meet two different applications. The first tuning focuses on an application called "thumbnail view only". It means that only a small version (low resolution) of the original content (figure 9-a) can be previewed (without key) by all users while the full bit stream is sent to all users. The full content is only viewable by users who have the corresponding key(s). Only the lowest resolution (resolution 0: LL subband) is public, layer 0 is encrypted at resolutions 1, 2, 3 and 4. The distortion is

11.10dB. Only 0.44% of data are encrypted. The visual degradation is illustrated in figure 9-b. The next application is called "full protection". Any part of the content is only viewable by users who have the corresponding key(s). This application requires hard visual degradation with no scalability. Only the most significant layer (layer 0) is encrypted at all resolutions. The achieved distortion is 8.17dB. Only 0.47% of data are encrypted. The visual degradation is illustrated in figure 9-c. In figure 10, we show how visual distortion presents rapid drop since the first byte encrypted. Indeed, encrypting one single byte per CCP causes important visual distortion. However, encrypting more bytes per CCP does not bring significant change. This is due to the context based nature of the EBCOT combined with the MQ arithmetic coding of the CCPs. The encryption of a single byte compromises the EBCOT decoding process for the entire CCP. Therefore, encrypting 16 bytes per CCP is sufficient to achieve the desired level of visual distortion while guaranteeing cryptographic security.

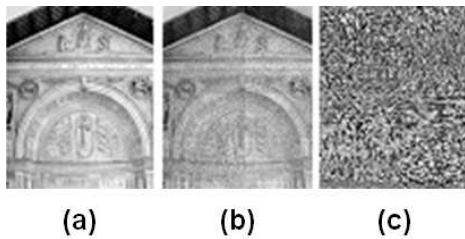


Figure 9. (a): Plain image, (b): Perceptual encryption and (c): Hard encryption.

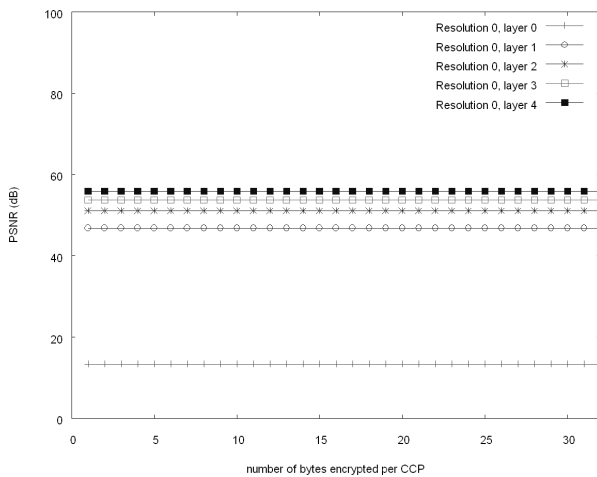


Figure 10. Impact of the number of bytes encrypted per CCP on PSNR.

3.2. Encryption ratio

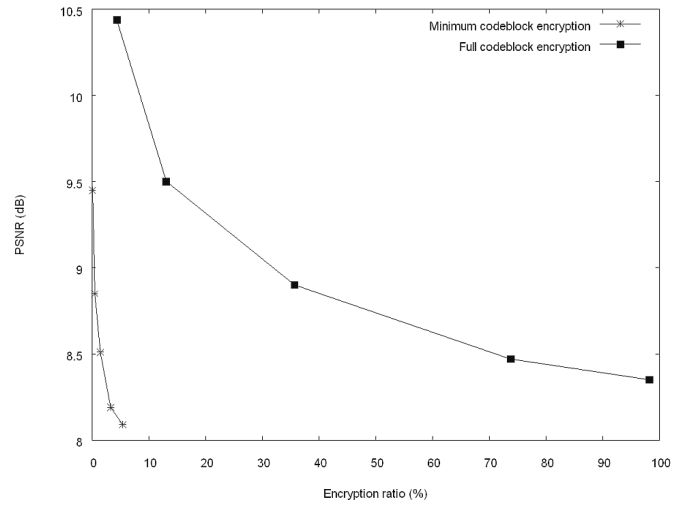


Figure 11. Minimum code-block encryption allows important reduction of encryption ratio.

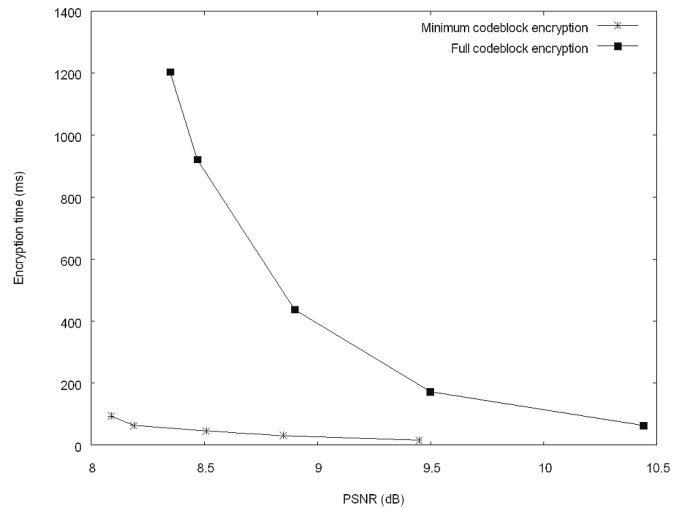


Figure 12. Encryption time.

In [13], a selective encryption algorithm is proposed for resolution progression JPEG2000 compressed images. At least the first 20% of data is encrypted to achieve a distortion of 9dB for full confidentiality. Our proposal is progression independent. In addition, we reach a better level of confidentiality (8.09dB) with only 5.43% of encrypted data and guarantee cryptographic security. In figures 11 and 12, we illustrate respectively visual distortion with respect to encryption ratio and encryption time required to achieve a given visual distortion. All quality layers, precincts and components are encrypted. We compute PSNR, encryp-

tion ratio and encryption time for five points ($R_e = \{0\}$, $R_e = \{0, 1\}$, $R_e = \{0, 1, 2\}$, $R_e = \{0, 1, 2, 3\}$ and $R_e = \{0, 1, 2, 3, 4\}$). We choose these encryption parameters to fit the situation considered in [13]. We observe that with full code-block encryption [13], higher encryption ratio is required to achieve a given distortion compared to minimum code-block encryption (figure 11). For illustration, for [13], we need to encrypt about 30% of data to reach less than $9dB$ visual distortion. However, in our proposal, encrypting only 5.43% yields $8.09dB$ (with $R_e = \{0, 1, 2, 3, 4\}$). For our method, we can see the rapid drop of visual quality with respect to encryption ratio. This allows important time saving as shown in figure 12. To reach a visual distortion of $9dB$ at least, full code-block encryption [13] requires $400ms$ while we achieve only $16ms$ for encryption.

3.3. Compression friendliness

The additional data included in the bitstream introduces a little overhead ranging from 0.1% at least to 5% at most which is acceptable. In addition, encryption impact on compression performance is rarely tackled in the literature which makes comparison difficult.

3.4. Error tolerance

We tested three encryption methods: using block ciphers in chaining mode (AES in CFB mode) as proposed in [13], ECB mode and our method (modified CTR mode). We injected one single bit error in each cipher bitstream. Figure 13 shows error images after decryption. It illustrates error propagation behavior for the three different encryption methods. In CFB mode [13] (figure 13-c), any error that occurs at a given byte in packet data is propagated to all subsequent code-blocks in the precinct, this is due to the use of chaining in CFB mode. In the modified CTR mode (our proposal), any error that occurs at an encrypted byte affects at most that byte and the next one in decryption (if the erroneous byte is $0XFF$). The error remains confined to the code-block where the error occurs. The impact is not perceptible (figure 13-a). AES-128 ECB mode gives intermediate result, at most sixteen adjacent bytes are affected by the error (figure 13-b).

4. Conclusion

Contrarily to state of the art algorithms, the proposed solution achieves configurable, format compliant, compression friendly selective encryption algorithm. It achieves the minimum encryption ratio required to reach a target visual distortion while guaranteeing cryptographically secure selectively encrypted bitstream. This allows achieving impor-

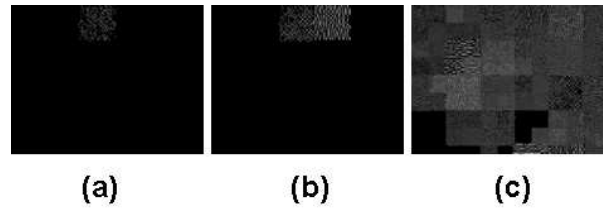


Figure 13. Error images after decoding for one byte error, (a): CTR, (b): ECB and (c): CFB modes.

tant time saving. In future works, we will focus on compressing the metadata needed for decryption.

References

- [1] Openjpeg. open-source jpeg 2000 codec developed by the universite catholique de louvain (ucl), belgium, <http://www.openjpeg.org>.
- [2] H. Cheng and X. Li. Partial encryption of compressed images and video. *IEEE Trans. Signal Processing*, 48(8):2439–2451, 2000.
- [3] D. Engel, T. Stütz, and A. Uhl. Format-compliant jpeg2000 encryption with combined packet header and packet body protection. *MM&Sec*, pages 87–96, 2007.
- [4] D. Engel and A. Uhl. Lightweight jpeg2000 encryption with anisotropic wavelet packets. *ICME*, pages 2177–2180, 2006.
- [5] J. Fang, J. Sunand, and H. Qian. Compliant asymmetric authenticated encryption scheme for jpeg2000 code-streams. *Int. Journal of Computer Science and Network Security*, 6(11):272–276, 2006.
- [6] M. Grangetto, M. Grosso, and E. Magli. Selective encryption of jpeg 2000 images by means of randomized arithmetic coding. *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, pages 347–350, 2004.
- [7] ISO/IEC. Jpsec commission draft 2.0. *ISO/IEC/JTC1/SC29/WG 1, N3397*, 2004.
- [8] S. Li, C. Li, K. T. Lo, and G. Chen. Cryptanalysis of an image scrambling scheme without bandwidth expansion. *Cryptology ePrint Archive*, 2006.
- [9] S. Lian, J. Sun, and Z. Wang. Perceptual cryptography on jpeg2000 compressed images or videos. *Proc. 4th Int. Conference on Computer and Information Technology*, pages 78–83, 2004.
- [10] T. Lookabaugh. Selective encryption, information theory and compression. *Conference Record of the 38th Asilomar Conference on Signals, Sys and Computers*, 1:373–376, 2004.
- [11] R. Lundin, S. Lindskog, A. Brunstorm, and S. Fischer-Hübner. Measuring confidentiality of selectively encrypted messages using guesswork. *In Proc. 3rd Swedish National Computer Networking Workshop*, pages 99–102, 2005.
- [12] Y. Matias and A. Shamir. A video scrambling technique based on space filling curves. *Proc. CRYPTO '87*, 1987.

- [13] R. Norcen and A. Uhl. Selective encryption of the jpeg2000 bitstream. *6th Joint Working Conference on Communications and Multimedia Security, Lecture Notes on Computer Science*, pages 194–204, 2003.
- [14] B. Schneier. Aes timing attack. 2006.
- [15] C. E. Shannon. Communication theory of secrecy systems. *Declassified Report*, 1946.
- [16] T. Stütz and A. Uhl. On format-compliant iterative encryption of jpeg2000. *In Proc. 8th Int. Symp. Multimedia*, pages 985–990, 2006.
- [17] D. V. D. Ville, W. Philips, R. V. De Walle, and I. Lemanhieu. Image scrambling without bandwidth expansion. *IEEE Trans. Circuits Syst. Video Technol*, 14(6):892–897, 2004.
- [18] J. Wen, M. Severa, and W. Zeng. A format-compliant configurable encryption framework for access control of video. *IEEE Trans. Circuits Syst, Video Techno*, 12(6):545–557, 2002.
- [19] C. P. Wu and C. C. J. Kuo. Efficient multimedia encryption via entropy codec design. *Proc. of SPIE Security and Watermarking of Multimedia Content III*, 4314, 2001.
- [20] W. Zeng and S. Lei. Efficient frequency domain selective scrambling of digital video. *IEEE Trans. Multimedia*, 2002.