

## WATERMARKING FOR ENFORCING SECURED MEDICAL IMAGE ACCESS

V. GomathiPriya, R. Sinduja, R. Leelarani  
Sethu Institute of Technology

### ABSTRACT

In this paper, we propose a joint encryption/watermarking system for the purpose of protecting medical images. This system is based on an approach which combines a substitutive watermarking algorithm, the quantization index modulation, with an encryption algorithm: a stream cipher algorithm (e.g., the RC4) or a block cipher algorithm (e.g., the AES in cipher block chaining (CBC) mode of operation). Our objective is to give access to the outcomes of the image integrity and of its origin even though the image is stored encrypted. If watermarking and encryption are conducted jointly at the protection stage, watermark extraction and decryption can be applied independently. The security analysis of our scheme and experimental results achieved on 8-bit depth ultrasound images as well as on 16-bit encoded positron emission tomography images demonstrate the capability of our system to securely make available security attributes in both spatial and encrypted domains while minimizing image distortion. Furthermore, by making use of the AES block cipher in CBC mode, the proposed system is compliant with or transparent to the DICOM standard.

### I. INTRODUCTION

The rapid evolution of multimedia and communication technologies offer new means of sharing and remote access to patient data. In particular, medical imaging is already called to play important roles in applications like telesurgery, telediagnosis, and so on. But at the same time, this ease of transmission and sharing of data increases security issues in terms of [1]

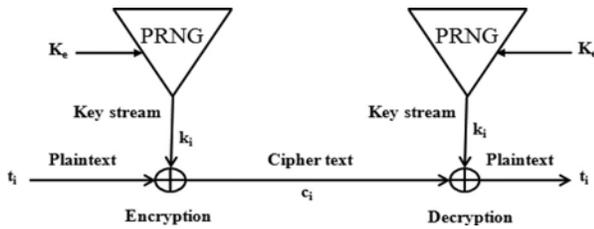
- 1) Confidentiality, which means that only authorized users can access patient data;
- 2) Availability, which guarantees access to medical information in the normal scheduled conditions of access and exercise;
- 3) Reliability, which is based on a) integrity—a proof that the information has not been altered or modified by no authorized persons; and b) authentication—a proof of the information origins and of its attachment to one patient. Reliable pieces of information can be used confidently by the physician.

In any information systems, data confidentiality, integrity, and nonrepudiation services are usually achieved by cryptographic means. DICOM1, the standard of reference for medical images, allows data encryption through the triple DES2, the AES3 . . . , as well as digitally signing a

DICOM object by making use of the DSA4 (see Part 15 of the DICOM standard). However, once decrypted or its digital signature deleted or lost, one piece of information is no longer protected and it becomes hard to verify its integrity and its origin. From this point of view, these cryptographic means, especially encryption, rather appear as an “*a priori*” protection mechanism. Watermarking has been proposed as a complementary mechanism to improve the security of medical images [2]. When it is applied to images, watermarking modifies or modulates the image pixels’ gray-level values in an imperceptible way, in order to encode or insert a message (i.e., the watermark). Thus, it allows us to intimately associate protection data with the information to be protected. watermarking can be used for verifying the reliability of an image by asserting its integrity and its authenticity. For instance, in a transaction, patient name and physician identity can be inserted in the image [3]–[6]. As defined, watermarking is an “*a posteriori*” control mechanism as the image content is still available for interpretation while remaining protected.

Different approaches have been proposed in order to benefit from the complementarity of these two mechanisms in terms of *a priori/a posteriori* protection, essentially in the context of copyright protection. Technically, two categories of methods can be distinguished according to the way watermarking and encryption are merged.

- 1) Joint decryption/watermarking, where watermark embedding is conducted during the decryption process [7]–[10].
- 2) Joint encryption/watermarking (E/W), where watermarking and encryption step processes are merged. In this case, the watermark can be extracted a) in the spatial domain,



**Fig. 1. Encryption/decryption processes of a stream cipher algorithm which**

secret key is  $K_e$ .  $t_i, c_i$ , and  $k_i$  correspond to the plain text bits/bytes, the cipher text bits/bytes, and the secret key stream bits/bytes, respectively.  $k_i$  is issued by a PRNG. i.e., after the decryption process, or b) in the encrypted domain, or c) in both domains [11].

The system we propose in this paper belongs to the second category. It merges a substitutive watermarking algorithm, the quantization index modulation (QIM), and an encryption algorithm which can be a stream cipher algorithm (e.g., RC45) or a block cipher algorithm (e.g., AES). Our objective is to give access to embedded security attributes in the encrypted and spatial domains for the purpose of verifying the reliability of an image. The rest of this paper is organized as follows. In Section II, we independently present the watermarking and the cipher algorithms we used, before introducing their combination in Section III. We then detail our implementation in Section IV. Section V presents some experimental results considering two distinct medical modalities, ultrasound and positron emission tomography (PET), and discusses some constraints of deployment. Before concluding, we analyze the security of the proposed scheme in Section VI.

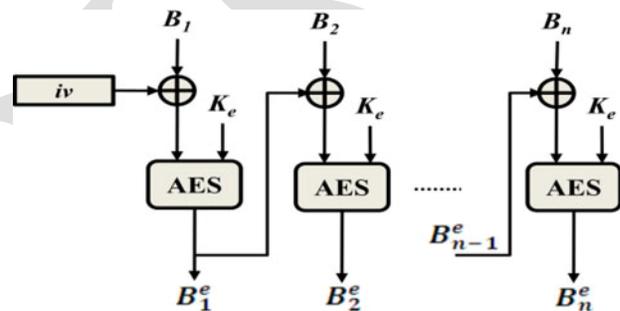
## II. CRYPTOGRAPHIC AND WATERMARKING PRIMITIVES

### A. Cryptographic Primitives

Basically, there exist two types of encryption algorithms: block cipher algorithms and stream cipher algorithms. Block cipher algorithms, like the AES and

the DES, operate on large blocks of plaintext, whereas stream cipher algorithms, like the RC4 or the SEAL6 [12], manipulate stream of bits/bytes of plaintext.

1) *RC4 Stream Cipher Algorithm:* As described in Fig. 1, stream cipher algorithms combine the bits/bytes of plaintext  $T = [t_1, \dots, t_i, \dots, t_n]$  with a secret keystream of bits/bytes  $K = [k_1, \dots, k_i, \dots, k_n]$  issued from a pseudorandom number generator (PRNG), through a XOR operation typically. The keystream generation depends on one secret key  $K_e$ , making stream cipher algorithms as part of symmetric encryption techniques. Thus, bits/bytes of cipher text  $C = [c_1, \dots, c_i, \dots, c_n]$  are usually defined as

$$c_i = t_i \oplus k_i. \quad (1)$$


**Fig. 2. AES Encryption in CBC mode.  $B_i, B_e i$ , and  $K_e$  denote the plaintext block, the encrypted block, and the encryption key, respectively.  $i_v$  is a random initialization vector.**

Some of the main advantages of this type of algorithms are that they are simple and operate at a higher speed than block cipher algorithms [13].

The specificity of such stream cipher algorithm resides in how the bit/byte key stream is generated by the PRNG. The RC4 PRNG is based on two steps.

- 1) “Initialization,” where a table of 256 bytes is filled by repeating the encryption key as often as necessary until to fill this table.
- 2) “Byte key stream generation,” where the elements of the table are combined by applying permutations and additions to generate the key stream.

More details about stream cipher algorithms can be found in [12].

2) *AES in CBC Mode of Operation:* In this paper, we use the block cipher algorithm AES in the cipher block chaining (CBC) mode of operation in order to be compliant with the DICOM standard. The

concept of mode of operation refers to the manner in which blocks of plaintext (sequence of bytes) are treated at the encryption stage (respectively, decryption stage). As depicted in Fig. 2, when the CBC mode is applied, a plaintext block is combined, with the previous cipher text block through a XOR operation before being encrypted with the AES. If we denote  $Be_i$  the encrypted version of a block  $B_i$  and  $Be_{i-1}$  the previous encrypted block,  $Be_i$  is thus given by

$$Be_i = AES(B_i \oplus Be_{i-1}, Ke)$$

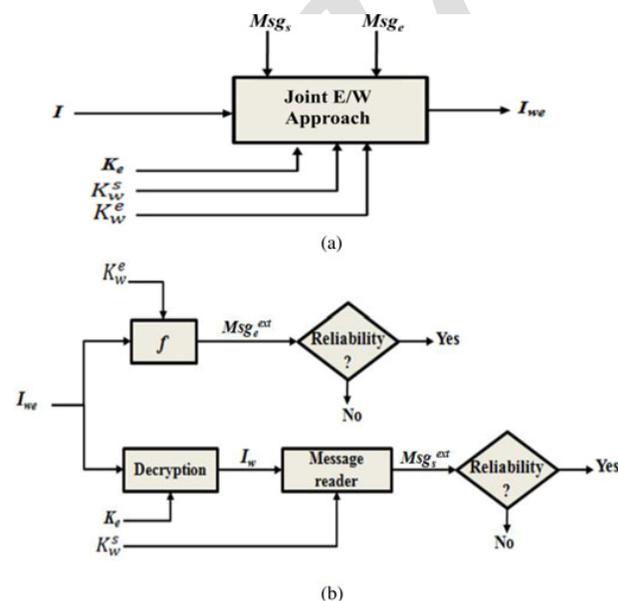
where  $Ke$  is the encryption key. The reader may refer to [14] for a complete description of the AES.

### B. Watermarking Primitive: The QIM Modulation

The QIM, proposed by Chen and Wornell [15], relies on quantifying the components of one image according to a set of quantizers based on codebooks in order to insert a message. More clearly, to each message  $ms_i$  issued from a finite set of possible messages  $Ms = \{ms_i | i=0, \dots, qs\}$ , the QIM associates the elements of a codebook  $Cms_i$  such as

$$Cms_i \cap Cms_j = \emptyset \quad i \neq j. \quad (2)$$

Substituting one component of the image by its nearest element in the codebook  $Cms_i$  thus allows the insertion of  $ms_i$ . Let us consider one image component such as a vector of



**Fig. 3. Architecture of the proposed system.**  $I, I_{we}, I_w, Ke, Ksw,$  and  $Kew$  denote the original image, the watermarked encrypted image, the watermarked

decrypted image, the encryption key, and the watermarking keys for the spatial and encrypted domain, respectively.  $Msg_{eand} Msg_{exte}$  are the embedded and extracted messages in the encrypted domain, respectively.  $Msg_{sand} Msg_{exts}$  denote the embedded and extracted messages in the spatial domain, respectively.  $f$  is the watermarking extraction function in the encrypted domain. (a) Protection. (b) Verification. pixels  $X \in NN$  while dividing the  $NN$ -dimensional space into nonoverlapping cells of equal size. To satisfy (2), each cell is associated with a codebook  $Cms_i, i = 0, \dots, qs$ . As a consequence, one message  $ms_i$  has several representations in  $NN$ . The insertion process is conducted as follows. If  $X$  belongs to the cell which encodes the message to be inserted,  $X_w$  (the watermarked version of  $X$ ) corresponds then to the center of this cell; otherwise,  $X$  is moved to the center of the nearest cell that encodes the desired message. During the extraction step, the knowledge of the cell to which  $X_w$  belongs is enough to identify the embedded message. Notice that such a modulation definitively alters the image. We will come back on this issue in Section V.

## III. PROPOSED JOINT ENCRYPTION AND WATERMARKING SYSTEM

### A. System Architecture and Principles

The purpose of our system is to verify the reliability of an image within the spatial domain as well as the encrypted domain. As illustrated in Fig. 3, it relies on two main procedures: protection and verification. The protection stage [see Fig. 3(a)] jointly conducts the watermarking and encryption of an image  $I$ . It allows us to insert two messages,  $Msg_s$  and  $Msg_e$ , which will be available in the spatial and encrypted domains, respectively. The insertion and the extraction of each message depend on a watermarking key:  $Kew$  for the encrypted domain and  $Ksw$  for the spatial domain. These two messages contain security attributes that will assess the image reliability in each domain. Indeed, each message contains an authenticity code  $AC$ , which identifies the image origin (e.g., about 600 bits by combining the French National Identifier with the DICOM Unique Identifier [16]), and an integrity proof. In the spatial domain, integrity is ensured by making use of a secure hash function (e.g., SHA7) computed on the image bit subset that is not modified by the watermarking process. We call this subset of bits  $nmb$ . So, the message available in the spatial domain,  $Msg_s$ , is defined as follows:

$$Msg_s = \langle AC, SHA(nmb) \rangle \quad (3)$$

In the encrypted domain, integrity is controlled by verifying the presence of a secret pseudorandom sequence of bits generated using a secret watermarking key. As we will see and discuss in Section III-C, the integrity of the watermarked encrypted image is considered as valid if we retrieve these bits at specific locations within the SHA signature of each watermarked-encrypted block bytes. We consider this pseudorandom sequence as a proof of integrity. In consequence, the verification of the image authenticity and integrity in the encrypted domain relies on extracting  $Msge$  given by

$$Msge = \langle AC, PRNG(Kew) \rangle \quad (4)$$

where  $Kew$  represents the watermarking key in the encrypted domain.  $Kew$  initializes the PRNG function.

Anyway, as it can be seen in Fig. 3(b), protection data are made available from the encrypted image or from the decrypted image for a subsequent verification stage. If watermarking and encryption are jointly conducted, watermark extraction and image decryption are two independent processes.

### B. Combination of Encryption and Watermarking

In this section, in order to simplify the presentation of our system, we manipulate 8-bit encoded images.

#### 1) General Principles of Joint E/W Approach:

Let us consider one block of bytes or equivalently a set of contiguous pixels. For this block, our objective is to give access to two messages:  $msi$ , the message available in the spatial domain and  $mej$ , the message available in the encrypted domain. Similar to  $msi$  (see Section II-B),  $mej$  is a message issued from a finite set of possible messages  $Me = \{mej\}_{j=0, \dots, qe}$ .

In order to conduct jointly this double watermarking process and to avoid any interference between them, we propose to adapt the QIM described previously. The basic idea is to decompose each codebook  $Cms_i$  into subcodebooks  $Cmsime_j$  such as

$$Cms_i = \bigcup_{j=0}^{qe} Cmsime_j \quad (5)$$

$$Cmsime_j \cap Cmsime_k = \emptyset, j \neq k. \quad (6)$$

Considering a vector of pixels  $X \in NN$ ,  $msi$  and  $mej$  are then embedded simultaneously by replacing  $X$  with  $Xw$  which corresponds to the nearest element of  $X$  in  $Cmsime_j$ . Using the Euclidian distance,  $Xw$  is given by

$$Xw = \min_k ( \|X - Y_{ik}\| ), Y_{ik} \in Cmsime_j. \quad (7)$$

Making the message  $mej$  available in the encrypted domain depends on the subcodebook construction which is a process intimately linked with the encryption algorithm and also with the watermark extraction algorithm. Considering an encryption algorithm  $E$  and its encryption key  $Ke$ , subcodebooks  $Cmsime_j$  are built so as to verify

$$Cmsime_j = \{ Y \in Cms_i \mid f(Ye, Kew) = mej \}, \\ \text{where } Ye = E(Y, Ke) \quad (8)$$

where  $f$  is the watermark extraction function in the encrypted domain. The choice of the function  $f$  is closely related to the goals to be achieved by our system. We will explain in Section III-C the choice we made. Finally, to sum up this process,  $mej$  is made available in the encrypted domain by modulating pixel values in the spatial domain. This means replacing  $X$  by  $Xw$  its nearest element in the subcodebook  $Cmsime_j$ .

2) Implementation With a Cipher Algorithm (The AES in CBC Mode and the RC4): Depending on the selected cipher algorithm, some other constraints have to be considered when building the subcodebooks  $Cmsime_j$ . In the case,  $E$  corresponds to the AES in CBC mode;  $Ye$  is given by [see (8)]

$$Ye = AES(Y \oplus Xe-1, Ke) \quad (9)$$

where  $Xe-1$  is the previous encrypted block of bytes or set of pixels. So, the construction of  $Cmsime_j$  depends also on the previous encrypted block.

Unlike the AES, the RC4 encrypts each byte separately. When it is used,  $Ye$  is given by [see (8)]

$$Ye = \{ ye_1, \dots, ye_i, \dots, ye_n \}, \text{ with } ye_i = y_i \oplus ki \quad (10)$$

where  $ki$  corresponds to the  $i$ th byte of the keystream  $k$  generated by the RC4 according to the secret key  $Ke$ . Thus, the decomposition of each codebook into subcodebooks depends also on the keystream bytes which are different.

From these constraints, building the subcodebooks before protecting the image is out of interest. In order to reduce computation complexity, it is more realistic to determine subcodebooks at each block to encrypt, it means to build the cells of the subcodebooks  $Cmsime_j$  into the cell of  $Cms_i$  which encodes the desired message in the spatial domain, i.e.,  $msi$ . Moreover, in practice (see Section IV-A), we

sequentially test the elements of  $Cms_i$  until finding the nearest element  $Xwof X$  such as

$$f(Xwe, Kew) = mej, \text{ where } Xwe = E(Xw, Ke). \quad (11)$$

In this study, for sake of simplicity, we work with single-bit messages, i.e.,  $msi = \{0,1\}$  and  $mej = \{0,1\}$ , and consequently with two codebooks  $C0$  ( $ms0 = 0$ ) and  $C1$  ( $ms1 = 1$ ), and four sub codebooks  $C00$ ,  $C01$ ,  $C10$  and  $C11$  derived from  $C0$  and  $C1$ , respectively. More precisely, in one block of byte or in 1-pixel subset, one bit will be embedded in the spatial domain as well as in the encrypted domain.

### C. How Ensuring the Reliability in the Encrypted Domain?

To control the image integrity, one common solution is to compute its digital signature and to embed it. Obviously, if the embedding is not lossless or reversible [17], this signature is computed on the image parts that are left intact by the watermarking process. Because the message available in the encrypted domain results from distortion imposed in the spatial domain, and because the impact of these distortions in the encrypted domain is not predictable, it is not possible to compute the digital signature of the encrypted block and to embed it within itself. To overcome this issue, one alternative consists in verifying the presence of a pseudorandom sequence embedded at the protection process. For instance, such a sequence can be carried by the least significant bit (LSB) of some secretly selected bytes of each watermarked-encrypted block  $Xwe$ . In that case, the watermark extraction function is such as  $f(Xwe, Kew) = LSB(Xwe)$ . In fact, we force these LSBs to be equal to the bits of the pseudorandom sequence by modulating the pixels values in the spatial domain. Unfortunately, with this strategy, the embedded signature is independent of the content, and the verifier has no means to check the link between the pseudorandom it extracts and the rest of the encrypted content. At the same time, the detection rate is rather small. Indeed, we can only detect modification of the pseudorandom sequence. If for example, only 1 bit of this sequence is embedded per watermarked-encrypted block of 8 bytes  $Xwe$ , we have at least 1/128 chance to detect  $Xwe$  has been modified.

To solve this problem of content independence and achieve better detection performance, we propose to verify the presence of the pseudorandom sequence and access to the image authenticity code, within the SHA signatures of the

watermarked encrypted blocks. By doing so, the watermark function  $f$  used to extract the message  $mej$  from  $Xeis$  defined as

$$f(Xwe, Kew) = hk \quad (12)$$

where  $hk$  corresponds to the  $k$ th bit of  $H$ , the SHA-1 signature of  $Xwe$  (i.e.,  $H = \text{SHA}(Xwe)$ ). The choice of the rank  $k$  depends on the secret watermarking key  $Kew$ . Because the "strength" of the SHA-1 is of 80 bits, if 1 bit of  $Xwe$  changes, then there is one-in-two chance that  $hk$  commutes. In that way, the recipient can verify the integrity as well as the authenticity of the image in its encrypted form. It just has to extract  $Msg$  from the SHA signatures of each watermarked-encrypted blocks.

## IV. IMPLEMENTATION OF THE PROPOSED JOINT E/W SYSTEM

As stated earlier, our implementation works with the RC4 or with the AES in CBC mode. In the following, we first describe how codebooks and sub codebooks are built and then detail the different steps of our joint E/W algorithm.

### A. Codebook Construction

The first step consists in constructing the set of  $Cms_i$  codebooks. Let us consider block of  $N$  pixels (or bytes) and, as we stated earlier, the insertion of one bit in both the spatial and encrypted domain (i.e.,  $msi = \{0,1\}$  and  $mej = \{0,1\}$ ). The value of  $N$  depends on the image bit depth and of the adopted cipher algorithm. Indeed, in the case of an image encoded on 16 bits, because the AES works with blocks of 16, 24, or 32 bytes,  $N$  will be equal to 8, 12, and 16 pixels, respectively. In the sequel, for sake of simplicity, we consider 8-bit depth images, i.e.,  $N$  equals the number of bytes in an encrypted block. In our implementation,  $Cms_i$  is built as follows:

$$C_{m_{si}} = \left\{ Y \in \mathbb{N}^N / \left\lfloor \frac{Y_k}{\Delta} \right\rfloor \bmod 2 = m_{si} \right\} \quad (13)$$

where  $\Delta$  represents the quantization step and  $Y_k$  is the  $k$ th byte or equivalently the  $k$ th pixel of the block to encrypt. The choice of  $k$  depends on the secret watermarking key  $Ksw$  in the spatial domain and is different for each pixel block. As designed, only 1 pixel in a pixel block  $X$  is quantized in order to encode 1 bit of the message in the spatial domain (i.e.,  $Msgs$

In order to embed the message  $me_j$  along with  $ms_i$  into one pixel block, we propose to modulate  $l$  LSBs from  $p$  secretly selected pixels other than the pixel at the location  $k$ . Again this process is based on  $K_{sw}$ . By doing so,  $C_{msime} j$  regroups a subset of  $C_{msime} j = 2lp$  elements of  $C_{ms} i$ .

As exposed in Section III-B1 and in (12),  $X$  will be replaced by  $X_w$ , i.e., by its nearest element in  $C_{msime} j$ . In order to reduce the complexity, instead of calculating the whole set of elements of  $C_{msime} j$ , it is preferable to test these different elements depending on their Euclidian distance with  $X$ , starting by its nearest element, until the value of  $X_w$  that satisfies (11) is found.

Based on this strategy, we can determine the probability for not being able to embed  $me_j$  into a block  $X$ , i.e.,  $f(X_w, K_{ew}) = me_j$  after having tested all  $2lp$  elements of  $C_{msime} j$ . Indeed, based on the properties of cryptographic hash functions (see Section III-C), there is one-in-two chance that the change of 1 bit of  $X$  leads to the correct value of  $me_j$ . As a consequence, the probability the embedding of  $me_j$  fails is given by  $PEF = 2^{-2lp}$ . This probability is very small. For instance, in the case  $(l, p) = (2, 2)$ , i.e., we modulate the two LSBs of two pixels; this probability is already about  $PEF \sim 10^{-5}$ .

Similarly, we can also calculate the probability for being able to embed  $me_j$  within  $u$  tests. This probability is given by  $PES(u) = 1 - (0.5)^u$ . As can be seen in Fig. 4,  $PES$  converges rapidly to 1 with the increase of  $u$ . Considering again  $(l, p) = (2, 2)$ , the probability to insert  $me_j$  within two tests equals 0.75. On the average, 1 bit of  $Msg_e$  will be embedded into a pixel block within two tests. As a consequence, the duration of our process is at least two times longer than simply encrypting the image (i.e., without the SHA). We will come back on this issue in Section V-C.

#### B. Algorithm

In the encrypted domain, bits of  $Msg_e$  will be extracted from the SHA-1 signature of these blocks. With the RC4 algorithm, it is possible to work with smaller pixel block dimension due to the fact that it works on stream of bytes (see Section II-A1).

For an image  $I$ , whatever the block dimension, our joint watermarking/encryption approach acts in two steps

1)  $I$  is splitted into no overlapping blocks,  $\{X_i\}_{i=1..U}$ , of  $N$  pixels. In order to form  $Msg_s$  (see Section III-A), we concatenate the image authenticity code  $AC$  with the SHA signature of  $nmb$ , which contains the bits of all the pixels that will not be modified by the insertion, i.e., the nonelected pixels, as well as of the most

significant bits of the selected pixels that will be modulated (see Section IV-A). The message available in the encrypted domain  $Msg_e$  is also built according to (4) using the secret watermarking key  $K_{ew}$ .

2) Messages embedding and encryption are then conducted

jointly, for each block  $X_i$ :

- a) using the subcodebooks  $C_{msime} i$ , one bit  $ms_i$  of  $Msg_s$ , and one bit  $me_i$  of  $Msg_e$ , are jointly inserted  $X_i$  is replaced by  $X_{wi}$ , which belongs to one cell of  $C_{ms} i$  and which verifies:

$$f(X_{wei}, K_{ew}) = me_i \quad (14)$$

where  $X_{we} i$  represents the encrypted watermarked version of  $X_w$

- b) once  $X_{wi}$  computed, it is encrypted through the adopted encryption algorithm (i.e., the stream cipher RC4 or the AES in CBC mode).

As stated before, at the verification stage, extraction can be conducted independently in both the encrypted and spatial domains using the corresponding secret watermarking key  $K_{ew}$  or  $K_{sw}$ . In the encrypted domain, the encrypted image  $I_{we}$  is decomposed in blocks of  $N$  bytes.

Then, the function  $f$  is applied to each block to extract one bit of  $Msg_e$ . In the spatial domain, the message  $Msg_s$  is extracted based on principles of the QIM. Each message is used by next to verify the image reliability in one domain, it means verifying the authenticity code of the image and its integrity by comparing, in the spatial domain, the extracted SHA signature with the recomputed one and, in the encrypted domain, by checking the equality between the extracted and regenerated random sequences. Notice that for 12-bit depth or 16-bit depth images,

The principle of our algorithm remains the same. Differences stand in the codebooks construction and the pixel block dimensions. As an example, for 16-bit encoded image, we work with 8 pixel blocks instead of 16 pixel blocks; the number of bytes remains the same.

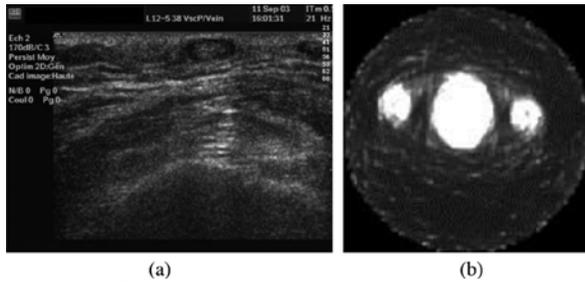


Fig. 5. Samples of our image test sets. (a) Ultrasound image. (b) PET image.

the encrypted domain, by checking the equality between the extracted and regenerated random sequences. Notice that for 12-bit depth or 16-bit depth images, the principle of our algorithm remains the same. Differences stand in the codebooks construction and the pixel block dimensions. As an example, for 16-bit encoded image, we work with 8 pixel blocks instead of 16 pixel blocks; the number of bytes remains the same.

## V. PERFORMANCE EVALUATION AND DISCUSSION

Experiments were conducted on two sets of medical images: 100 ultrasound images of  $576 \times 690$  pixels of 8-bit depth, and 200 PET images of  $144 \times 144$  pixels of 16-bit depth. Some samples of our dataset are given in Fig. 5. Let us recall that for images encoded on 8 or 16 bits, our joint E/W system manipulates blocks of 16 or 8 pixels, respectively (i.e.,  $N = 16$  or  $N = 8$ ).

### A. Image Distortion

We decided to use the peak-signal-to-noise ratio (PSNR) in order to measure the distortion between an image  $I$  and its watermarked and deciphered version  $I_{wd}$

$$\text{PSNR}(I, I_{wd}) = 10 \log_{10} \left( \frac{(2^d - 1)^2}{\text{MSE}} \right) \quad (15)$$

$$\text{MSE}(I, I_{wd}) = \frac{1}{L} \sum_{k=1}^L (I(k) - I_{wd}(k))^2$$

where  $L$  corresponds to the number of pixels of the image  $I$ , and  $d$  corresponds to its depth. Our choice relies on the fact that the algorithm we proposed in Section IV introduces on average the same image distortion in each block, thus spreading it over the whole image. Furthermore, it does not take advantage of a psychovisual model which is helpful to adapt the watermark amplitude locally into the image, making at

the same time the PSNR not appropriate. Even though there exist some models for natural images, none of them have been proved adapted for medical imaging yet. If we still consider our implementation, we can determine the lower bound of PSNR depending on the image depth  $d$ , the number of modulated pixels  $p$ , the number of LSB modulated per pixel  $l$ , and the quantization step  $\Delta$ . Indeed, the maximum distortion one may introduce by modulating  $l$  LSBs of one pixel is  $\delta = 2^l - 1$ . Similarly, the maximum distortion induced by the quantization of 1 pixel is  $\Delta$ . As a consequence, considering a block  $B$  of  $N$  pixels and its decrypted-watermarked version  $B_{wd}$ , the PSNR lower bound is given by

$$\text{PSNR}(B, B_{wd}) \geq 10 \log_{10} \left( \frac{(2^d - 1)^2}{(p\delta^2 + \Delta^2)/N} \right). \quad (16)$$

We give in Fig. 6 the variation of this limit for different values of  $p$  and  $l$  considering  $d = 8/N = 16$  or  $d = 16/N = 8$ , and the smallest possible value of  $\Delta$ , i.e.,  $\Delta = 1$ . In these examples, it can be seen that the PSNR limit is quite high for both 8- and 16-bit depth images.

In practice, with the same parameterization and working with the AES in CBC mode or with the RC4, achieved PSNR values are much greater (about 60 and 105.26 dB for our ultrasound and PET image test sets, respectively), as indicated in Table I. This can be explained by the fact that we do not have to modify all  $p$  pixels in order to make  $m$  available in the encrypted domain.

### B. Capacity

Capacity rates depend on the block size  $N$ . When the AES is used with our implementation, rates achieved in each domain are both of  $1/N$  bits/pixel. As a consequence, capacities are about of 24 000 and 2592 bits for ultrasound and PET images, respectively. While using the AES limits the block size to some specific values, by working with the RC4, it is possible to consider smaller block size. For instance, if  $N = 4$ , the capacity rate becomes of  $1/4$  bits/pixel in each domain. The total amount of bits one can embed is then of 193.5 and 10.125 kb for ultrasound and PET images, respectively. But, this increase of capacity is accompanied with a diminution of the PSNR as shown in Table II with the parameterization ( $l = 2, p = 2, \Delta = 1$ ).

## VI. SECURITY ANALYSIS

The security of our joint E/W system partly relies on the watermarking-encryption relationships we introduced and on the application framework. Let us recall that  $M_{sge}$  and  $M_{gss}$  serve the same purpose which is the protection of the reliability of an image. Thus,

they contain some common pieces of information. In this section, we first start by looking at cryptographic attacks, which aim is to break confidentiality, before focusing on watermarking attacks.

#### A. Cryptographic Attacks

In our joint E/W system, we work with popular encryption algorithms (the AES and the RC4) which security performance is well known. Due to the fact that we do not intrinsically modify them, without the knowledge about the watermarking keys, their performance are preserved against common cryptographic attacks like the ones based on cipher text-only, known plaintext, chosen plaintext, or/and on chosen cipher text attacks. If  $Kew$  or  $Msege$  are known from the attacker, he has no other additional means than a regular cryptographic attack to get  $Ke$  or to have an idea about the clear watermarked image (i.e.,  $Iw$ ). This is due to the fact  $Msege$  is embedded within the SHA signatures of watermarked-encrypted blocks and not directly into the encrypted bit stream. Furthermore,  $Msege$  appears “encrypted” in  $Iwe$  and its presence does not reduce the entropy of the watermarked-encrypted image as compared with the simple encryption of the image. If  $Ksw$  and  $\Delta$  are known, codebooks  $Cms_i$  can be computed but the subcodebooks  $Cms_{ime_j}$  cannot be derived even if  $Iwis$  is known. The attacker has no clues about  $Ke$ . If now, he knows also  $Kew$  or/and  $Msege$ , we retrieve the cryptographic attack based on known plaintext and known ciphertext. Nevertheless, if the attacker complete this set of data with the original image  $I$ , he can get an idea about the subcodebooks and consequently find the encryption key  $Ke$ .

#### B. Watermark Attacks

In [20], Zhou *et al.* have defined three types of watermark attacks we analyze in this section: unauthorized message embedding, unauthorized message detection/extraction, and unauthorized watermark removal.

## VII. CONCLUSION

In this paper, we have proposed a new joint watermarking/ encryption system, which guarantees *a priori* and *a posteriori* protection of medical images. It merges the QIM and a cipher algorithm or a block cipher algorithm. Our system gives access to two distinct messages in the spatial domain and in the encrypted domain, respectively. These two messages are used for verifying the image reliability even though it is encrypted. The AES in CBC mode makes our system compliant with the DICOM standard. Experimental results show that the image distortion is

very low and that the achieved capacity is enough to embed a reliability proof as well as some other data. Obviously, our joint watermarking/encryption system is slower than simply encrypting the image but it provides reliability control functionalities. On the other hand, the execution time for image decryption is not impacted. We have also shown that the way we combine encryption and watermarking does not interfere with the security of the encryption algorithm and that the security of our system depends on the knowledge of the encryption and watermarking keys. Future works will focus on making our scheme more robust to attacks like lossy image compression (e.g., JPEG) and reducing the complexity of our algorithm.

## REFERENCES

- [1] G. Coatrieux, H. Maître, B. Sankur, Y. Rolland, and R. Collorec, “Relevance of watermarking in medical imaging,” in *Proc. IEEE EMBS Int. Conf. Inf. Technol. Appl. Biomed.*, 2000, pp. 250–255.
- [2] G. Coatrieux, C. Le Guillou, J.-M. Cauvin, and C. Roux, “Reversible watermarking for knowledge digest embedding and reliability control in medical images,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 2, pp. 158–165, Mar. 2009.
- [3] U. Rajendra Acharya, D. Acharya, P. Subbanna Bhat, and U. C. Niranjan, “Compact storage of medical images with patient information,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 5, no. 4, pp. 320–323, Dec. 2001.
- [4] W. Pan, G. Coatrieux, N. Cuppens-Bouahia, F. Cuppens, and C. Roux, “Medical image integrity control combining digital signature and lossless watermarking,” *Data Privacy Manage. Autonom. Spontaneous Sec. (LNCS)*, vol. 5939/2010, pp. 153–162, 2010.
- [5] U. Rajendra Acharya, U. C. Niranjan, S. S. Iyengar, N. Kannathal, and L. C. Min, “Simultaneous storage of patient information with medical images in the frequency domain,” *Comput. Methods Programs Biomed.*, vol. 76, pp. 13–19, 2004.
- [6] J. M. Rodrigues, W. Puech, and C. Fiorio, “Lossless crypto-data hiding in medical images without increasing the original image size,” in *Proc. 2<sup>nd</sup> Int. Conf. Adv. Med. Signal Inf. Process.*, Sep. 2004, pp. 358–365.
- [7] R. Anderson and C. Manifavas, “Chameleon: A new kind of stream cipher,” in *Proc. 4<sup>th</sup> Int. Workshop Fast Software Encryption*, Haifa, Israel, Jan. 1997, vol. 1267, pp. 107–113.