A Watermarking Based Medical Image Integrity Control System and an Image Moment Signature for Tampering Characterization

G. Coatrieux, *Senior Member, IEEE*, H. Huang, H.Z. Shu, *Senior Member, IEEE*, L.M. Luo, *Senior Member, IEEE*, Ch. Roux, *Fellow, IEEE*

***Abstract*—In this paper, we present a medical image integrity verification system to detect and approximate local malevolent image alterations (e.g. removal or addition of lesions) as well as identifying the nature of a global processing an image may have undergone (e.g. lossy compression, filtering …). The proposed integrity analysis process is based on non significant region watermarking with signatures extracted from different pixel blocks of interest and which are compared with the recomputed ones at the verification stage. A set of three signatures is proposed. The two firsts devoted to detection and modification location are cryptographic hashes and checksums, while the last one is issued from the image moment theory. In this paper, we first show how geometric moments can be used to approximate any local modification by its nearest generalized 2D Gaussian. We then demonstrate how ratios between original and recomputed geometric moments can be used as image features in a classifier based strategy in order to determine the nature of a global image processing. Experimental results considering both local and global modifications in MRI and retina images illustrate the overall performances of our approach. With a pixel block signature of about 200 bit long, it is possible to detect, to roughly localize and to get an idea about the image tamper.**

***Index Terms* — medical imaging, integrity control, watermarking, image moment, cryptographic hash.**

I. Introduction

ADVANCES in information and communication technologies provide new means to access, share, replicate and manipulate medical images. But if the daily medical practice takes advantage of such an evolution, this facility to handle images also compromises their security. As any pieces of medical information seen and analyzed, medical images have to be reliable or trustworthy, a statement which relies on the outcomes of the image [1][2]: "integrity" that ensures it has not been modified by a non-authorized person; and "authenticity" which asserts its origin and its attachment to one patient. In this work, we focus on verifying the integrity of an image.

G. Coatrieux and Ch. Roux are with the Institut Telecom; Telecom Bretagne; Unite INSERM 1101 Latim, Technopole Brest-Iroise, CS 83818, 29238 Brest Cedex 3 France (e-mail: {gouenou.coatrieux, christian.roux}@telecom-bretagne.eu).

H. Huang was with the Institut Mines-Telecom, Telecom Bretagne, 29238 Brest Cedex 3, France. She is now with Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201203, China (e-mail: hh.sophia@gmail.com).

H.Z. Shu and L.M. Luo are with the LIST Southeast University, China (e-mail: {shu.list, luo.list}@seu.edu.cn).

While it is vital to keep images safe from any damage, it is as much important being able to detect when an image has been modified and in which manner. Indeed, medical images can be modified accidentally, for example during their transmission, or deliberately. In this latter case, images can be tampered with the introduction or the removal of lesions [3]. Also, it must be known that some image processing may lead to similar situations. In telemedicine applications, for instance, lossy image compression is tolerated so as to reduce the amount of information to be transmitted. However, depending on its extent, this process may induce inacceptable information loss and results in a misdiagnosis [4][5], involving at the same time liabilities of physicians.

In a previous work, Coatrieux *et al.*[6] stated that verifying the integrity of a medical image is an analysis process that needs to answer three main questions: is the image identical to its original version? If not, which parts of the image can still be used trustfully for diagnosis? And, finally, from a legal point of view, what is the purpose or the objective, if any, of the image tamper? In order to answer these questions, one can associate three distinct levels of integrity:

-- Level 1 (*L1*): Modification Detection -- an alarm should be given under any kind of image modification;

-- Level 2 (*L2*): Modification Location -- untrustworthy parts of the image have to be indicated; either in a rough way, so as to designate areas still interpretable by the physician;

-- Level 3 (*L3*): Forensics analysis -- the nature of the modification over the whole image or within untrustworthy regions has to be identified (accidental, authorized/non authorized).

In the literature, different strategies have been proposed for verifying image integrity. They include the use of image digests/signatures/hashes [7][8], watermarking [9] and blind forensics methods [10]. The first kind of methods verifies image integrity based on the comparison of hashes computed over the image under investigation or some parts of it with the hashes shared with the image. Such a hash can be computed in different ways. Cryptographic hash functions [7][8] allow verifying the exact identity of the image under investigation with the original image, and can be used to achieve *L1*. They provide the best performance in terms of detection and are extremely difficult to counterfeit. To localize alterations (*L2*), one can compute hashes on independent image areas [11][12]. In order to reduce the image signature length, checksums, which are based on error detection codes, can be used instead of

cryptographic hashes [13]. However, checksums are less efficient in terms of detection. Perceptual hashes are another kind of image digests [7]. They aim at detecting malevolent image content changes but are robust to global image processing such as JPEG and filtering processes. These approaches can be used to partly achieve *L1* and *L3*. Indeed, they only state if yes or not the image is still valuable while considering some specific modifications.

Watermarking is an effective tool for verifying image integrity and authenticity. One common approach consists to insert a specific watermark [14–16]. The non-detection of this latter informs about image integrity loss. In some cases watermarking is combined with image signatures [17][18]. For example, in [13][19–20] a set of signatures is computed from one Region Of Interest (ROI) and then watermarked within Regions Of Non-Interest (RONI). Other watermarking schemes, referred as self-recovery image watermarking, propose to partly restore tampered parts of the image [21-25]. However, these ones only roughly recover the image, i.e. not the image details. Nevertheless, by subtracting the restored image to the observed one, it becomes possible to achieve *L3*. Even though such a solution has some interests for general public multimedia applications, in healthcare it is crucial to retrieve exactly the original signal, otherwise, there still exists a risk that some important information is missing. Furthermore, these methods require high embedding capacity (capacity - amount of information a method can embed) and, sometimes, may not be able to protect images of small dimensions (e.g. the method reported in [25] fails to achieve its objectives for images smaller than 512×512). To gain in performance they sometimes spread the watermark over the whole image with the risk to biased the image interpretation. Recently, Tagliasacchi *et al.* [26] proposed an image hashing algorithm based on the principles of compressive sensing while making the assumption that tampering is sparse in some orthonormal bases. To compute their hash, the image is first partitioned into n blocks of size $N \times N$, from which the average gray value of each block is computed. Then, they quantize a limited number of random projections of this averaged image before Wyner-Ziv coding it to reduce the bit-length. As with the previous approaches, it is possible to get an idea of the modification, but it requires computing iteratively the sparse estimate of the tampering. Even though this hash is rather small (about 100 bytes for a 512×512 image), it is robust to moderate content-preserving transformations (e.g. scaling, rotation ...). More clearly, their hash will remain the same after these transformations. It will also fail to detect non-sparse modification like JPEG compression. As consequence, it only partly fulfills *L3*.

The third strategy refers to blind forensic techniques which do not require any image prior [10]. These techniques aim at identifying the footprint or evidences left by most image modifications. A large group of "blind forensic" methods rely on classifier based mechanisms which use as input some image features [10][27-30] that reveal the statistical nature of image modifications. Most of these solutions allow discriminating non-modified images from images that have been processed (filtered, rotated, compressed and so on [27-29]).

The system we propose is based on the watermarking of a ROI signature (computed on independent pixel blocks) into RONI of the image. As it will be shown, the embedded signature allows us achieving *L1* to *L3* under the constraint of low watermarking capacity performance. Beyond our system objectives, the main contribution of this work stands in a *L3*-signature derived from the geometric moments of pixel blocks used in order: i) to have an idea about the shape of one local modification, by approximating it with its nearest 2D Gaussian function; and, ii) to serve as image feature within a classifier based mechanism which purpose is to discriminate or identify the nature of the global process an image may have undergone. Our solution differs from the one of Tagliasacchi *et al.* and from other self-recovery approaches, because it focuses on approximating the modification rather than reconstructing it, with the objective to discriminate any image modification under the heavy constraint of a small signature length.

In the next Section, we describe the architecture of our system as well as our set of signatures. Two of them, devoted to *L1* and *L2*, are also presented. Section III exposes our *L3*-signature, the way it is computed and used. Then, we report and discuss in Section IV experimental results for both local and global image modifications, before concluding in Section V.

II. Proposed Integrity Control System and *L1*/*L2*-Signatures

A. System Architecture

The basic structure of our system is depicted in Fig. 1. At the protection stage, a signature is extracted from the ROIs of the image or from pixel blocks, and is embedded into the rest of the image (i.e. RONIs). At the verification stage, differences between recomputed and extracted signatures are analyzed in order to achieve the three integrity verification levels: *L1*, *L2* and *L3*. In the following, we propose one specific and independent signature for each integrity level (see Fig. 1). More clearly, *Li* is based on one signature H_i, meaning that the ROI signature H will result in the concatenation of the signatures H_i, i=1..3: $H=[H_1,H_2,H_3]$. H_i is obtained from the ROI by means of an extraction function. We expose in the next sections the extraction functions we retained and developed and also how we conduct *L3* based on *L2*.

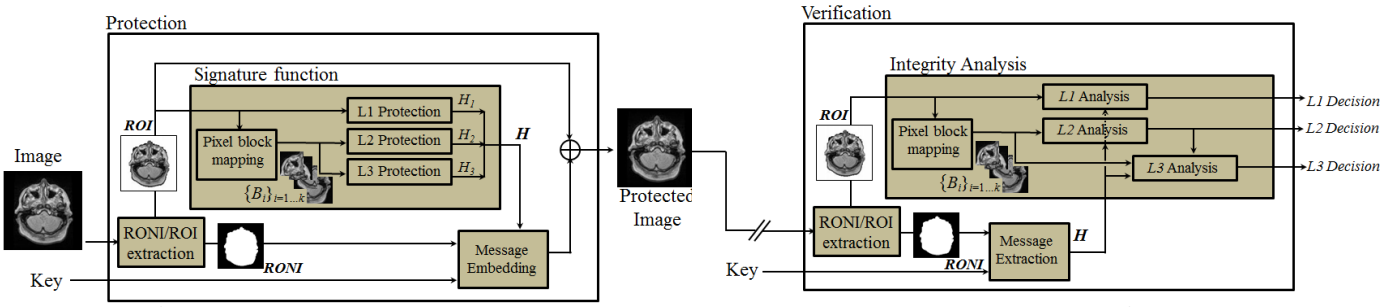


Fig. 1. Principle of our integrity control system. A secret watermarking key is used to guarantee that only the entitled users can verify the image integrity.

It is important to notice that the performance of our system partly depends on the one of the RONI watermarking algorithm used. These methods, as any watermarking schemes, establish a different compromise between three main properties: imperceptibility, capacity and robustness (i.e. ability of the watermark to survive image alterations). The more robust to image distortion the algorithm, the more the image is distorted and the capacity is reduced (i.e. the size of H in our case). RONI watermarking performance relies also on the RONI dimensions. Obviously, it increases along with the dimensions of the RONI and of the image, and may become critical for images of small dimensions. This is the reason why we focus on providing a signature H of dimension $|H|$, expressed in bits, as small as possible. At the same time, the needs of robustness of our signature H_i, i=1..3, are different. Indeed, based on the fact the purpose of H_1 is just to inform us the image has been modified; it can be embedded in a fragile way. Any modification of the image will result in a *L1*-signature different from the original one (i.e. H_1). In case only RONIs are modified, we will know it through H_2 that will indicate all ROIs are authentic. Such an analysis is however possible if H_2 is extracted error free, meaning robustly embedded. That is also the case for H_3. If H_2 and H_3 are extracted with errors the integrity analysis we propose will be biased.

Several RONI watermarking methods from the literature [6][17][25][31-33] can be used in our system. Among them some have been designed for multiple message embedding [31][25]. In [31], Giakoumaki *et al.* conducted message embedding by means of quantization index modulation applied to RONI wavelet coefficients and make use of BCH error correction codes so as to gain in robustness. In case one message does not need to be robust (e.g. H_1), it is embedded within the first wavelet decomposition levels. On the contrary, when high robustness is required, wavelet subbands of low and mid-frequencies are exploited. In their experiments conducted on 8 bit depth encoded echographic images of 256×320 pixels, they embed three messages: one signature (320 bits), one index (364 bits) and some caption text (1456 bits) that survive a lossy JPEG compression of quality factor 85, 90 and 95, respectively. By adapting their scheme to our *Li*-signatures, we assume it is possible to achieve a capacity of 700-800 bits with robustness to a JPEG compression of quality factor 85. OM. Al-Qershi *et al.* [25] provide a similar approach using RS code instead of BCH codes. They achieve high capacity and robustness to salt and pepper noise addition. Notice that for both methods, the RONI comprises some part of the anatomical object observed into the image. This limits the watermark robustness, as it is mandatory not introducing signal anomalies that could err practitioners. This invisibility requirement becomes less strict if the image black background outside the anatomical object is considered (see examples in Fig.4(b) and Fig.5). As example, by simply modulating the relationship between the value of a pixel and the mean of the pixel block it belongs to, [33] we were able to embed 600 bits within the black background of 256x256 12 bit encoded MRIs of the head while being robust to JPEG compression of quality not smaller than 70. On the same kind of images, in [6], by modulating Discrete Cosine Transform Coefficients of pixel blocks we achieved a black background watermark robust at least to a JPEG compression of quality factor 50 or 85, with a capacity of 900 to 1500 bits, respectively. In fact, based on a short experiment about the watermark invisibility we conducted with four radiologists of the CHRU of Rennes [34], we increase the watermark energy while maintaining it smaller than a "hindering-threshold"; threshold under which the watermark is visible for the specialist but does not bother her or him during her or his interpretation of the image. Notice that watermark imperceptibility still needs to be further explored in medical imaging, even for RONI watermarking.

As described, RONI watermarking performance varies from one scheme to another and increases with the image dimensions. Thus, while considering the ROI constituted of k blocks of N×N pixels (e.g. N=64, N=128), we propose in the sequel to ensure *L2* and *L3* with a signature of about 200k bit long for a 12 bit depth encoded image. This will allow us protecting one ROI representing 50% of a 256×256 MRI with less than 1000 bits using 64x64 pixel blocks (see Section IV.C).

B. Detecting (L1) and Localizing (L2) Image Modifications

As stated above, to achieve L1, any modification of the image should cause an alarm. For that purpose, we propose cryptographic hash functions as extraction function. Beyond the fact they provide the best detection performance, they have specific properties like the dispersion property which ensures that two slightly different ROIs will have very different signatures [7]. Thus considering one ROI as a binary message M, the cryptographic hash function (f_1) will provide a binary string of n=$|H_1|$ bit long signature H_1 : $H_1 = f_1(M)$. We decided to use the SHA-1(Secure Hash Algorithm) that yields to a 160 bit signature. Its probability of collision, that is the probability that two ROIs have the same signature, is upper

bounded by $1/2^{160}$ [7].

In order to identify which parts of the image can still be interpreted without risks of misdiagnosis (*L2*), one solution consists to compute a set of signatures from different parts of the ROI. It is then possible to define a map of *k* pixel blocks B_i, i=1…k - see the block mapping function in Fig. 1 - where each block will be independently protected by one signature (i.e. h_2^i,i= 1..k). Integrity is consequently controlled at the block level. At the verification stage, one block B_i is said to be tampered if its recomputed signature $h_2^{i\prime}$ differs from h_2^i. By concatenating this set of signatures, we obtain the signature H_2: $H_2 = f_2([B_1, \ldots, B_k]) = [h_2^1, \ldots, h_2^k]$. In our system, the redundancy provided by error detection codes has been considered for the computation of signatures h_2^i,i=1..k [13]. Considering Hamming error detection codes, here used, a message (equivalently a pixel block) of 2^r-r-1 bit long will be assigned a redundancy or checksum of r bits with an error rate upper bounded by $1/2^r$ [35][13].

Thus, based on H_1 and H_2, it is possible to know if the image has been altered (*L1*) and which pixel blocks B_i,i=1..k, cannot be used trustingly (*L2*). As we will see in the next section, our *L3*-analysis process depends in parts on the output of *L1* and *L2*.

III. *L3*-Signature: Design and Use

As stated previously, *L3* corresponds to a digital forensics analysis which purpose is to identify the nature of the modification over the whole image or within untrustworthy pixel blocks. Our *L3*-signature (H_3) has been designed in order to treat these two situations as two distinct cases. More clearly, differences between H_3 and its recomputed version are analyzed or interpreted differently depending if the modification is identified as global or local. The choice between both cases is made upon the output of *L2* (see Fig. 1). If this latter shows that only some pixels blocks are non-authentic, we assume the modification is local, otherwise we consider the image undergone a global image processing. In the former case, we aim at getting some insights regarding the modification while in the latter we want to identify the nature of the modification. In the following we first introduce our *L3* signature in the case of local modification, and then present how it can be used in a classification based strategy for identifying the nature of global modifications. Notice that as for *L2*, we consider the ROI constituted of *k* blocks B_i, i=1…k, from each is computed one *L3*-signature h_3^i (see Fig. 1) As a consequence, H_3 results from the concatenation of these signatures ($H_3 = [h_3^1, \ldots, h_3^k]$) and the *L3* analysis process is conducted independently on each block.

A. Approximation of Local Modification

Herein, our objective is to know more about one local modification within one pixel block B_i claimed as non-authentic at the output of *L2*. The modification Δ we refer to corresponds to the signal of difference between the original block B_i and its modified version B'_i: $\Delta = B'_i - B_i$. We aim at refining more precisely its position within B_i (see section II-A)

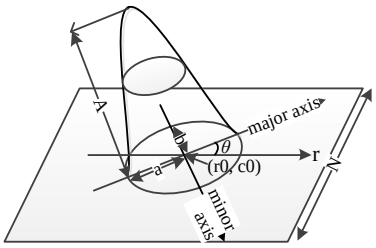


Fig. 2. Gaussian function model in an N×N pixel block

and its dimensions, i.e. its amplitude and size. For that purpose, we suggest to approximate it by means of a modification model: its nearest generalized 2D Gaussian (see Fig. 2) defined as:

$$G_{r_0,c_0}^{A,\sigma_1,\sigma_2}(c',r') = Ae^{-\left(\frac{(c'-c'_0)^2}{2{\sigma_1}^2}+\frac{(r'-r'_0)^2}{2{\sigma_2}^2}\right)} \quad (1)$$

where $c' = c \cdot cos\,\theta + r \cdot sin\,\theta$ and $r' = c \cdot cos\,\theta - r \cdot sin\,\theta$, in the reference coordinates built on the two axis of the Gaussian. Notice that in order to simplify a little this modification model, we make the following hypothesis introduced by [37]: most energy and amplitude of the 2D Gaussian function is concentrated in the range of $\pm 2\sigma$. As a consequence, we consider that the basis ellipse of the Gaussian function in Fig. 2 is constrained with the relation: $a = 2\sigma_1$ and $b = 2\sigma_2$.

The purpose of the *L3* analysis process based on the difference between h_3^i and $h_3^{i\prime}$, computed from B_i and B'_i respectively, is thus to help us to determine the parameters of this modification model so as to make it the nearest as possible of the real local tamper Δ. The parameters we look at are (see Fig. 2): its center of mass, identified by its row and column positions (r_0, c_0); the direction of the major axis (θ); the deviations (σ_1 and σ_2) along the major axis and minor axis and the amplitude (A).

To conduct this task, we decided to estimate each parameter independently. One consequence of such a choice is that our *L3*-signature h_3^i results from the concatenation of six digests d_u^i, u=*1..6*, each derived from B_i by means of a digest function g_u built considering the following properties:

- C_1: g_u should be easy to compute;
- C_2: g_u should be a linear function, so the digest associated to the modification Δ can be achieved easily: $g_u(\Delta) = g_u(B'_i) - g_u(B_i)$.
- C_3: g_u should be proportional to the parameters to be estimated. This will allow us assuming that if $g_u(\Delta) \propto g_u\left(G_{r_0,c_0}^{A,\sigma_1,\sigma_2}\right) \forall u$ then $\Delta \propto G_{r_0,c_0}^{A,\sigma_1,\sigma_2}$.

The digest functions we propose are based on image geometric moments [32]. Beyond the fact that they are easy to compute (C_1) and that they are linear (C_2) – providing easy access to the moments of the modification Δ – they also give us some geometric information about the image shape, thus about the shape of Δ.

The geometric moments of an image intensity function $f(r, c)$ are defined by [37] :

$$M_{nm} = \sum_r^N \sum_c^N r^n c^m f(r,c)\,,\ n,m = 0,1,2,\ldots \quad (2)$$

Among existing moments, geometric moments are the simplest ones, and lower order moments can be taken as global shape

descriptors. In particular, it is possible to estimate the parameters of a generalized 2D Gaussian function $G_{r_0,c_0}^{A,\sigma_1,\sigma_2}$ defined on a pixel block from its geometric moments:

- The Gaussian center of mass (r_0,c_0) can be derived from the two first order moments(M_{01}, M_{10}):

$$r_0 = M_{10}/M_{00}, \quad (3)$$
$$c_0 = M_{01}/M_{00};$$

where M_{00}, the zero order moment, represents the sum of gray value of the image.

- The parameters of the ellipse at the basis of the generalized 2D Gaussian function can be estimated by referring to the image ellipse concept defined by Prokop *et al.*[38] . They suggested that any image object can be described with an ellipse of same first two order moments value and axes directions with the image object. This image ellipse can be determined using the second order central geometric moments (μ_{02}, μ_{11}, μ_{20}), defined as:

$$\mu_{nm}(\Delta) = \sum_x^N \sum_y^N (r - r_0)^n (c - c_0)^m \Delta(r, c) \quad (4)$$

Notice that μ_{nm} can be derived from the corresponding geometric moment M_{nm}[30]:

$$\begin{aligned} \mu_{00} &= M_{00}; \\ \mu_{10} &= \mu_{01} = 0\ ; \\ \mu_{20} &= M_{20} - r_0 M_{10}; \\ \mu_{02} &= M_{02} - c_0 M_{01}.; \\ \mu_{11} &= M_{11} - c_0 M_{10}.; \end{aligned} \quad (5)$$

The direction of the major axis (θ), as well as the axis lengths of the Gaussian basis (a) and (b) can be estimated by [39]:

$$a = \left\{ \frac{2\left[\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}{}^2}\ \right]}{\mu_{00}} \right\}^{1/2} \quad (6)$$

$$b = \left\{ \frac{2\left[\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}{}^2}\ \right]}{\mu_{00}} \right\}^{1/2} \quad (7)$$

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \quad (8)$$

- Once the position and the basis ellipse of the Gaussian are estimated, we still need to estimate its amplitude (A). Considering that the gray volume of the Gaussian function (M_{00}) should be equal to the one of the real modification and the relations between a, b and σ_1, σ_2, the amplitude (A) of the Gaussian function can be derived from the following relation:

$$M_{00}(G) = A \sum_{r'=1}^N \sum_{c'=1}^N e^{-\left(\frac{(r'-r'_0)^2}{2\sigma_1{}^2} + \frac{(c'-c'_0)^2}{2\sigma_2{}^2}\right)} = \frac{1}{2} A \pi a b \quad (9)$$

or equivalently:

$$A = \frac{2M_{00}(G)}{\pi a b} \quad (10)$$

Based on the above relations and on the fact geometric moments are linear (see eq.(2)), we can estimate the parameters of the nearest generalized 2D Gaussian function of a local modification in one pixel block B_i. We simply assume the modification model is of same geometric moments that the real modification Δ. As a consequence our digest functions g_u, u=1..6, correspond to the first geometric moments of order up to two of B_i.

To sum up, for one pixel block B_i, *L3* protection and verification processes are achieved in the following way:

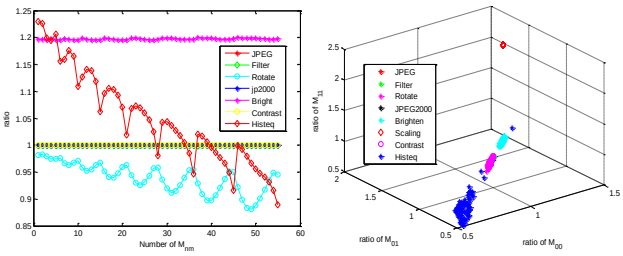
1) *Integrity protection stage*
   - *a-* Compute h_3^i =[$g_1(B_i)$,$g_2(B_i)$,$g_3(B_i)$,$g_4(B_i)$, $g_5(B_i)$,$g_6(B_i)$]=[$M_{00}^{B_i}, M_{10}^{B_i}, M_{01}^{B_i}, M_{11}^{B_i}, M_{02}^{B_i}, M_{20}^{B_i}$].
   - *b-* Embed h_3^i along with other signatures in the RONI.
2) *Integrity verification stage* (B_i is detected modified based on h_2^i)
   a. Compute the *L3*-signature of the observed block B'_i: $h_3^i{}'$=[$M_{00}^{B'_i}, M_{10}^{B'_i}, M_{01}^{B'_i}, M_{11}^{B'_i}, M_{02}^{B'_i}, M_{20}^{B'_i}$].
   b. Compute the geometric moments of the modification Δ ($\Delta = B'_i - B_i$): $h_3^{\Delta} = h_3^i{}' - h_3^i$ =[$M_{00}^{\Delta}, M_{10}^{\Delta}, M_{01}^{\Delta}, M_{11}^{\Delta}, M_{02}^{\Delta}, M_{20}^{\Delta}$] and compute the associated central moments $\mu_{11}^{\Delta}, \mu_{02}^{\Delta}, \mu_{20}^{\Delta}$ (see Eq.(5))
   c. Compute the parameters of the nearest generalized 2D Gaussian function based on Eqs.(3)(6)(7)(8) and (10).

From this procedure, six digests ($M_{00}^{B_i}, M_{10}^{B_i}, M_{01}^{B_i}, M_{11}^{B_i}, M_{02}^{B_i}, M_{20}^{B_i}$) are required per image block As a result, the length of the *L3*-signature, H_3 depends on the number of protected blocks and the image depth. Indeed, for a ROI divided in blocks of 64×64 pixels, for each block, we will need to store 164 bits or 188 bits for an 8 bit encoded image or 12 bit encoded image respectively. If the available embedding capacity allows it, we can embed digests of blocks of smaller dimensions or overlapped. In both cases accuracy of local modification approximation will be better. It must be noticed that the geometric moments of one block can be expressed as a linear combination of those of its sub-blocks [39]. Furthermore, like Tagliasacchi *et al.* [26], it is possible to compress H_3 losslessly and consequently reduce the capacity requirements while improving *L3* performance.

As defined and as we will see in Section IV, in the case the modification is constituted of several pieces into a block, our solution gives an average approximation. Furthermore, this approach remains limited to local modifications. In fact, it suffers from the uni-polarity and linearity assumptions we made about our modification model (see constraint C_2). In case the modification Δ is constituted of different pieces of amplitude positive and negative at the same time, like when the image undergone some filtering or lossy compression, our system will fail to give a good response. Indeed, based on the fact such a process does not impact much $M_{00}^{B_i}$, we will badly estimate the modification position (see Eq.(3)), amplitude (Eq.(10)) and extension (see Eqs.(6)(7)(8)).

B. Global Image Modification Identification

One image is considered as globally modified in our system, if at the output of *L2* all independently protected blocks B_i, *i=1..k*, are declared unauthentic. Our objective is thus to identify the nature of the modification (e.g. filtering, lossy compression, contrast adjustment and so on). As mentioned in introduction, this task can be conducted blindly by means of classifier based mechanisms using as inputs some image features that reveal the statistical nature of image modifications [10] [29]. The reader may refer to [29] for a wide range of image feature sets. The efficiency of such a solution depends



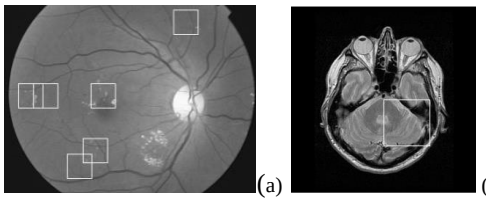
(a) (b)

Fig.3. Variations of ratios between geometric moments of the original image and its modified version in the case of retina images considering different modifications: (a) Ratios of moments of order up to 10th (55 values); (b) Scatter diagram of moments ratios (with the first 2nd order moments).

on: (i) the image feature sensitivity to image modifications and also (ii) their independency to the image content (most image processing impact the image and are partly independent of the image content [10]).

In our system, we follow a similar strategy with the next advantages: 1) based on *L2* and *L1* we know a pixel block is modified while in blind forensics it is first necessary to discriminate original images from modified ones; 2) we can have access to original image feature values by means of watermarking allowing us removing the effect of the image content. Instead of including new image features into H_3 (or equivalently h_3^i), we suggest to use pixel block geometric moments (i.e. h_3^i) as input of a classifier based strategy. As it can be seen in Fig. 3(a), in the case of retina images (see Fig. 4(a)), it appears that the variations of the ratios ∇M_{nm} between the geometric moments of one original pixel block and those of its modified version (i.e. $M_{nm}^{B_i}$ and $M_{nm}^{B'_i}$ respectively - $\nabla M_{nm} = M_{nm}^{B'_i}/M_{nm}^{B_i}$) discriminate fairly well different image processing. This can be better viewed in Fig.3(b) which shows the scatter diagram of three moment ratios --$\nabla M_{00}, \nabla M_{01}, \nabla M_{11}$-- for different modifications. Hence, we propose to train a classifier with the moment ratios as image features so as to discriminate global image modifications of the image. In our system, this classifier is the core of our *L3* analysis process, for global image modification identification.

The determination of the nature of the modification an image may have undergone is a multi-class decision problem (where one class corresponds to one specific type of modification). For that purpose, we built a multi-class classifier based on One-Versus-One (OVO) binary classifiers, each of which discriminates images modified accordingly to two kinds of possible modifications (e.g. JPEG *vs.* Filtering, JPEG *vs.* Contrast adjustment,and so on). Thus, for a given multi-classification problem of M classes, $M(M-1)/2$ One Versus One (OVO) binary classifiers have to be trained, each for every distinct pair of classes. Then, by analyzing the responses of these classifiers, a multi-class conclusion is drawn. Among the different strategies for combining decisions of binary classifiers, the Max-Wins Voting (MWV) is one of the most commonly used approaches [40]. MWV assigns an instance to a class which has the largest votes from all binary classifiers.



(a) (b)

Fig. 4 Image test samples a) Retina image, squared regions are blocks of 64×64 pixels and indicate positions of local modifications (b) MRI of the head JPEG compressed with Q=75, squared regions are blocks of 64×64 pixels and indicate positions of local modifications.

In this work, we use Support Vector Machines (SVM) [41]. This choice stands on the fact SVM provides superior classification performance in many applications [42–44]. One SVM has thus been trained to differentiate pairs of modifications (e.g. JPEG *vs.* Filtering, Rotation *vs.* Scaling …).

It is important to notice that the performance of such an approach depends on the number of geometric moment ratios used per block as well as on the number of protected blocks. Indeed, one signature h_3^i gives us access to the geometric moments of order up to two or equivalently to six moment ratios per $N \times N$ pixel block, all of which can be used to train a classifier. Performance depends also on the nature of the modification and of the sensitivity of geometric moment ratios to these modifications. We come back on these aspects in the next section.

IV. Experimental Results

In these experiments, we assume that *L1* and *L2* procedures have been already performed. We only focus on *L3*. We recall that our system decides if the modification is global or local based on the output of *L2* (see Fig. 1). If all blocks are declared unauthentic, then it tries to identify the nature of the global tampering; otherwise, it looks for approximating local modifications in blocks indicated as tampered.

Different image modalities have been considered (see samples in Fig. 5):

- Magnetic resonance (MRI) of the head: 120 images of 256 × 256 pixels and 12 bits depth from 3 patients, 40 images per patient;
- X-Ray Imaging: 162 mammograms of 4740 × 3540 pixels coded on 12 bits from multiple patients, 1 to 2 images per patient, and 200 abdomen CT images of 512 × 512 pixels coded on 12 bits from 4 patients, 50 images per patient.
- Ultrasound imaging (Echo) of vein: 52 images of 576 × 690 pixels and 8-bit depth from 3 patients, possessing 14, 9 and 29 images respectively.

These image data sets have been used in order to train and test our SVM classifiers for the identification of image processing (see section III.B). Nevertheless, for local modification, experiments have only been conducted on retina images of 8-bit depth (Fig.4(a)) and on 12-bit depth MRIs of the head of 256×256 pixels (Fig.4(b)). Notice also that whatever the imaging modality, the elementary block size is of 64×64 or 128×128 pixels from which 6 geometrics moments

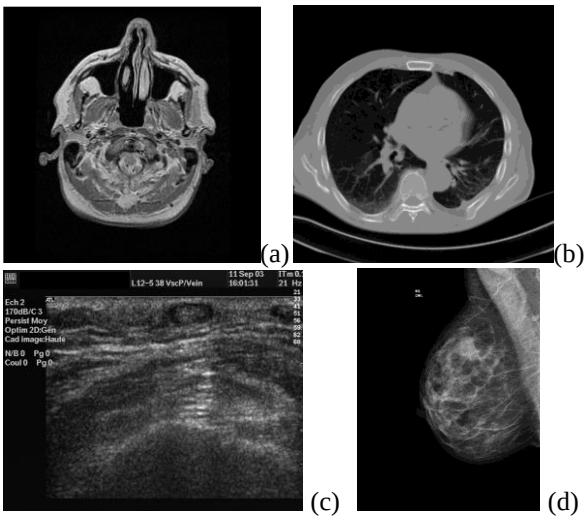


Fig. 5 Samples of our experimental image data sets: (a) MRI (b) CT image; (c) X-Ray image; (d) Echography.

are extracted and watermarked according to the above procedure. The global signature length is discussed in section IV.C.

A. Experiments for Local Modification

Herein, local modification approximation is conducted independently within 64×64 pixel blocks claimed as unauthentic (see Section III.A). Real and synthetic modifications were considered. They correspond to lesion removal or addition. Real modifications were achieved on retina images (see samples in Fig.4(a)). To conduct this task, we build up a system which combines an algorithm for the automatic detection of diabetes signs [45] with an in painting algorithm which automatically fills-in lesions to be masked [46]. Based on the automatic lesion detection we have also built a dictionary so as to insert lesions in image originally wholesome. Modified images and their "realistic" character were evaluated by a physician of our research team. Synthetic modifications allow us introducing some abnormal textures within MRIs. It is based on the generalized 2D Gaussian which amplitude is modulated by a noise uniformly distributed (see example in Fig.7). In MRIs of the head, depending on its location, such a distortion is interpreted as a hypersignal by a radiologist [6].

In order to evaluate the approximation performance of our scheme, we use the Mean Square Error (MSE) so as to measure the distance between the real modification ($\Delta = B'_i - B_i$) and its nearest 2D Generalized Gaussian $G^{A,\sigma_1,\sigma_2}_{r_0,c_0}$:

$$\mathrm{MSE}(\Delta, G^{A,\sigma_1,\sigma_2}_{r_0,c_0}) = \frac{1}{NN}\sum_{r,c=0,0}^{N-1,N-1}\left(\Delta(r,c) - G^{A,\sigma_1,\sigma_2}_{r_0,c_0}(r,c)\right)^2 \quad (11)$$

It gives an idea in average about the distance between the real modification with its approximation looking at them as intensity functions (i.e. $\Delta(r,c)$ and $G^{A,\sigma_1,\sigma_2}_{r_0,c_0}(r,c)$). Our choice stands on the fact our modification model is naive compared to lesions' complexity. In fact, we roughly approximate the modification. In case of a better modification model, other similarity measures like the one of Hausdorff [47] would be more appropriate.

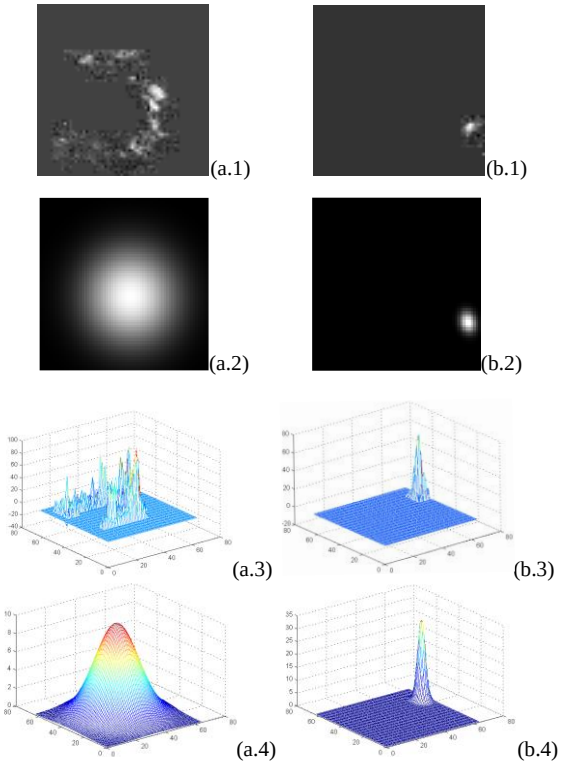


Fig. 6 Real image modifications of the retina image given in Fig. 4(a) and their respective approximations: (a.1)(b.1) (a.2)(b.2) 2D views of two distinct modifications and their approximations respectively; (a.3)(a.4)(b.3)(b.4) 3D views of same real modifications and their approximations respectively.

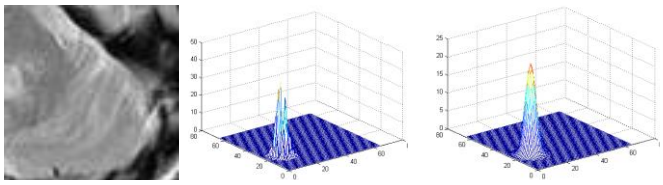


Fig.7 Approximation of a noise modulated Gaussian function added to the MRI image in Fig.4(b). a) 2D view of the zoomed 64×64 pixel block; b) and c) are 3D views of the real modification and of its approximation.

TABLE I IMAGE MANIPULATION AND THEIR PARAMETERS

Modification	Values of parameters				
Scaling up(γ_s %)	1	5	10	25	50
Rotation angle(θ)	1	5	15	30	45
Deviation of Gaussian filter (σ)	0.3	0.5	1.0	2.0	3.0
Contrast enhancement rate (γ_c %)	1	5	8	10	
Brighten rate (γ_b %)	2	5	8	10	
Quality factor(Q)	95	85	80	75	60
Compression rate JP2K (γ_j)	2:1	5:1	10:1	20:1	50:1
Histogram equalization					

For the two modifications shown in Fig.6(a) and Fig.6(b), we got MSE values equal to 101.15 and 4.01 respectively. For the modification in Fig.7, the approximation of the noise modulated Gaussian function has a MSE of 1.46. As it can be seen, it is obvious that the more the real modification is similar to a generalized 2D Gaussian, the more the approximation is correct.

B. Experiments for Global Modification

The next experimental results only concern the four following medical image modalities: MRI, CT-images,

TABLE II DETECTION (P_d), FALSE POSITIVE (P_{fp}), FALSE NEGATIVE (P_{fn}) RATES OF OUR BINARY SVM CLASSIFIERS WITH INPUT SIX GEOMETRIC MOMENT RATIOS ($\nabla M_{00}, \nabla M_{01}, \nabla M_{10}, \nabla M_{02}, \nabla M_{11}, \nabla M_{20}$) ISSUED FROM A BLOCK OF 128×128 OR 64×64 PIXELS. INDICATED FALSE POSITIVE AND NEGATIVE RATES ARE GIVEN FOR 128×128 PIXEL BLOCK ONLY.

Binary SVM Classifier (pair of modifications)	MRI				Mammography				CT				Echography			
	64×64	128×128			64×64	128×128			64×64	128×128			64×64	128×128		
	P_d	P_d	P_{fp}	P_{fn}	P_d	P_d	P_{fp}	P_{fn}	P_d	P_d	P_{fp}	P_{fn}	P_d	P_d	P_{fp}	P_{fn}
jpeg2K *vs.* jpeg	90.50	91.32	16.25	1.11	85.31	86.79	23.95	2.47	84.95	88.36	23.28	0	83.46	84.23	26.15	5.38
jpeg2K *vs.* filter.	95.58	90.00	7.92	12.08	90.99	93.87	2.96	9.24	88.50	91.34	8.96	8.36	87.31	87.69	6.46	16.15
jpeg2K *vs.* rot.	95.83	94.66	10.69	0	95.20	96.34	1.96	5.36	99.50	99.94	0.12	0	92.58	94.35	2.26	9.03
jpeg2K *vs.* scal.	94.75	99.10	0.14	0.14	99.38	99.63	0.62	0.12	99.95	99.85	0.30	0	95.76	95.77	6.92	1.15
jpeg2K *vs.* bright.	98.33	99.38	0	1.81	94.32	94.32	0.62	0	98.95	100	0.15	0	100	100	0	0
jpeg2K *vs.* contr.	97.08	97.08	5.00	0.83	89.69	90.56	6.42	12.47	94.20	97.76	1.94	2.54	79.42	92.88	0.77	13.46
jpeg2K vs. hist.eq.	100	100	0	0	99.38	99.51	0.62	0.37	100	100	0	0	100	100	0	0
jpeg *vs.* filter.	90.33	89.51	12.64	8.33	98.76	98.89	1.48	0.74	99.20	98.13	0.19	2.54	90.38	90.96	0.38	16.69
jpeg *vs.* rot.	98.26	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
jpeg *vs.* scal.	97.33	100	0	0	100	100	0	0	100	100	0	0	98.07	99.23	0	1.54
jpeg *vs.* bright.	100	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
jpeg *vs.* contr.	100	100	0	0	98.52	98.83	0	2.35	97.95	98.96	0	2.09	91.34	90.00	0	10.73
jpeg *vs.* hist.eq.	100	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
filter. *vs.* rot.	89.65	98.51	0	2.99	99.48	99.79	0.10	0.31	99.54	99.88	0	0.25	98.39	99.19	0	1.61
filter. *vs.* scal.	98.42	100	0	0	100	100	0	0	100	100	0	0	96.15	98.65	0	2.69
filter. *vs.* bright.	100	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
filter. *vs.* contr.	100	100	0	0	98.76	98.95	0.74	1.36	98.05	98.28	0	3.43	88.46	92.69	2.69	11.92
filter. *vs.* hist.eq.	100	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
rot. *vs.* scal.	100	98.82	1.67	0.69	97.96	99.26	1.11	0.37	97.55	97.84	3.43	0.90	96.34	97.50	1.15	3.85
rot. *vs.* bright.	100	99.65	0	0.64	99.88	99.94	0	0	100	100	0	0	100	100	0	0
rot. *vs.* contr.	99.67	99.24	0	1.53	97.71	98.58	0.12	2.72	98.95	99.10	0	2.79	99.61	95.96	0.38	0.38
rot. *vs.* hist.eq.	100	99.65	0.69	0	96.23	96.79	3.95	2.47	94.90	98.13	3.73	0	100	98.85	2.31	0
scal. *vs.* bright.	90.00	99.93	0.14	0	99.81	99.88	0	0.25	100	100	0	0	100	100	0	0
scal. *vs.* contr.	89.58	99.93	0	0.14	95.99	98.27	0.49	2.96	98.20	98.43	0	3.23	92.31	95.00	2.31	7.69
scal. *vs.* hist.eq.	100	99.93	0.14	0	95.55	95.74	3.58	4.94	97.40	99.40	1.19	0	100	99.04	1.92	0
bright. *vs.* contr.	92.33	92.85	1.94	12.36	99.38	99.69	0.49	0.12	95.95	96.12	0	7.76	100	100	0	0
bright. *vs.* hist.eq.	100	100	0	0	100	100	0	0	100	100	0	0	100	100	0	0
contr. *vs.* hist.eq.	100	100	0	0	98.09	99.69	0	0.62	100	100	0	0	100	100	0	0

mammograms and ultrasound images. This is due to the fact our retina database test set is too small to train our classifiers. Our classifier (see Section III.B) was built independently on each modality with the aim to detect common medical image processing: contrast and brightness adjustment, Gaussian filtering, scaling, rotation, lossy JPEG and JPEG2000 compression and histogram equalization. The parameter values we consider for each of these image manipulations are shown in Table I. We recall that this classifier is based on binary SVM classifiers that have been built so as to discriminate images modified by means of two distinct global image processes (see Section III-B).

Binary SVM classifiers were trained using only the six geometric moment ratios of one single block of 64×64 or 128×128 pixels centered in the image whatever the image modality (i.e. one signature h_3^i - see section III.A). Two size of block have been considered so as to show the sensitivity of geometric moment ratios to global image processing. The following results are given in average and achieved with classifiers that have been trained several times (at least 10 times) with different fold-cross validation (i.e. training and testing sets are randomly selected at each trial). Notice also that the parameters of the SVM were selected from their Receiver Operating Characteristics (ROC) curves.

The performance indicators we use for our binary classifiers are: the *detection rate* (P_d), that is the ratio between the number of modified images correctly classified and the number of tested images; the false negative (P_{fn}) and false positive (P_{fp}) rates which measure the number of images claimed as tampered with modification of type 1 (e.g. JPEG) while being in reality modified with modification of type 2 (e.g. filtering) and vice versa.

Table II gives the different rates (P_d, P_{fn}, P_{fp}) our binary SVM classifiers achieved with respect to different image modalities. As it can be seen, based on the six first order geometric moment ratios, they can distinguish different modification modes with high accuracy whatever the block size, size which however does not influence detection rates so much. The performance of the classifiers for JPEG2000 *vs.* JPEG and JPEG2000 *vs.* filtering are somehow lower than the others, especially for small modification extent (i.e. for small compression rates). The reason might lie in the fact that these modifications have a similar impact on image moments, thus confusing our classifiers. It is also possible to see that the false positive rate and false negative rate are similar and of small values. This indicates that the classifiers have a good tradeoff in identifying modified images (positive cases) and original images (negative cases). Notice also that working with images from different patients, demonstrate the independence of our feature to the image content. We also verified this by training classifiers with images from one patient; detection rates are very near to those of Table II.

If, now, we focus on identifying the nature of the modification, analyzing the output of our binary SVM classifiers (see section III-B), Table III indicates that the type of the modification can be estimated with a detection rates greater than 90% in general. In the case of echographic images our detector confuses JPEG2000 with JPEG. Indeed, if JPEG is

TABLE III DETECTION RATES OF OUR MULTI-CLASS CLASSIFIER CONSIDERING SIX GEOMETRIC MOMENT RATIOS (∇M_{00}, ∇M_{01}, ∇M_{10}, ∇M_{02}, ∇M_{11}, ∇M_{20}) ISSUED FROM A BLOCK OF 128×128 PIXELS.

Detection rate (%)	MRI	Mammography	CT	Echography
JPEG2000	90.67	70.92	72.10	43.77
JPEG	96.83	99.26	98.90	93.96
Filtering	96.33	99.51	97.60	80.00
Rotation	95.00	97.91	99.20	98.87
Scaling	99.67	98.28	99.60	99.62
Brighten	75.00	99.54	100	100
Contrast	76.04	89.57	83.25	83.49
hist. equ	100	94.48	100	100

omitted in the experiment (i.e. not considered as an attack) then the detection rate goes up to 95%.

C. Discussion

The performance of our complete scheme depends on the available watermarking capacity offered by the RONI watermarking scheme under constraints of robustness and imperceptibility (see Section II.A). This capacity has to be allocated to *L1*, *L2* and *L3*. Let us consider for example one 12 bit encoded MRI of 256x256 pixels with a ROI representing 50% of the image as well as an achievable robust capacity of about 1000 bits (See Section II.A). In daily medical practice, priority should be given to *L1* and *L2*. Considering a cryptographic hash and hamming codes (see section II-B), a few hundred of bits are required so as to encode *L1* and *L2*-signatures, i.e. H_1 and H_2. Our example MRI can be protected by allocating 280 bits to *L1* and *L2*. 160 bits are used to encode H_1 and 100 bits to encode H_2 (using hamming codes) allowing us to protect 5 64×64 blocks with some overlap (e.g. 4 of them cover 50% of the image while the last one is centered on the image), giving access to a good modification location [13]. Considering H_1 is embedded in a fragile way (see Section II.A), the rest of the available robust capacity is assigned to achieve *L3* permitting us to protect 4 of the previous 5 blocks ($|h_3^i|$=188 bits – see Section III.A). To gain in terms of capacity, it is also possible to work with bigger blocks. As illustrated by this small example, different strategies can be drawn so as to allocate the RONI watermarking capacity to H_1, H_2 and H_3, and achieved a protection even for small images. Notice that, as reported in [31], a typical size for many medical image modalities is 512x512 pixels and their size is continuously increasing with technology advances.

More generally, if the capacity allows it, one can work with blocks of smaller size or with overlapped blocks. Both cases will lead to a better local modification approximation as well as to a better detection and identification of global image processing. A hierarchical view of the modification can also be considered based on the fact geometric moments of one block can be expressed as a linear combination of its sub-blocks ones [40].

One can also adjust the size of the *L3* -signature per block, by reducing or increasing the number of geometric moments to embed. For instance, if more order moments are used, we can assume that image processing detection will be improved. The complexity of the local modification model can also vary. For example, to estimate the parameters of a "symmetric" 2D Gaussian (i.e. $\sigma_1=\sigma_2=\sigma$ in Eq.(1)) we just have to embed the 4 first geometric moments instead of 6 (see above).

For global modification, our approach is limited to a priori identified image modification; modifications we know an image may undergone. Nevertheless, one just has to update the *L3*-analysis process of our system with a new classifier if non anticipated modifications have to be considered. Results given above (see section IV.B) consider only 6 geometric moments from one pixel block. They are already high and not so sensitive to the block size (see Table II). They would be better if more blocks were considered.

It is important to notice that our *L3*-signature is not secure. If a pirate knows the image block partition, he or she can take advantage of the linearity and unipolarity properties of our local modification model. Indeed she or he may counteract or reduce the efficiency of our protection by conducting a complex tamper in a block which ensures invariance of the geometric moments (i.e. $h_3^i = h_3^{i\prime}$). Already, by adding a modification which amplitude is at the same time positive and negative, our system will fail to give the right estimation (see section III.A). However, such a lesion removal may not be so obvious to do, especially if the block partition is made secret with block overlapp.

Actually our system badly handles combination of local and global modifications. Let us consider a block tamper constituted of a low pass filter with the addition of a generalized 2D Gaussian as local modification. Because the filtering operation does not impact the zero order geometric moment but the others, we will not correctly estimate the parameter of the generalized 2D Gaussian (see Eqs. (6)(7)(8)(10)). In fact, this error keeps limited if the local modification is of greater energy than the distortion induced by the low pass filter. On the other hand, the local tamper influences all geometric moments and may render impossible the identification of low pass filter operation. Nevertheless, by working on the whole ROI, considering also several blocks, our system can be improved by means of a better decision process. This latter is very simple by now.

V. CONCLUSION

We have proposed a system for verifying the integrity of medical images. This system distinguishes three levels of integrity decision: detection, localization and approximation of the image alteration. For the latter level, we suggested approximating any malevolent local modifications by its nearest 2D generalized Gaussian function whose parameters are derived from the image geometric moments. In case the image is globally processed, these image moments can be also used to identify the type of the modification. Our system can help to find out the motivation of the tampering, but it remains limited to the detection of predefined kinds of image modification or tampering. Future work will focus on identifying a wider range of modifications, single or combined, and especially, their strength when they are global over the image.

REFERENCES

[1] G. Coatrieux, H. Maitre, B. Sankur, Y. Rolland, R. Collorec, "Relevance of watermarking in medical imaging," *in Proc. IEEE EMBS Int. Conf. Inf. Technol. Appl. Biomed.*, 2000, pp. 250–255.

[2] L. Kobayashi and S. Furuie, "Proposal for DICOM Multiframe Medical Image Integrity and Authenticity," *Journal of Digital Imaging*, vol. 22, no. 1, pp. 71-83, 2009.

[3] M. T. Madsen, K. S. Berbaum, A. N. Ellingson, B. H. Thompson, B. F. Mullan, R. T. Caldwell, " A New Software Tool for Removing, Storing, and Adding Abnormalities to Medical Images for Perception Research Studies," *Academic Radiology*, vol. 13, no. 3, pp. 305-312, 2006..

[4] K. Chen, T.V. Ramabadran, "Near-lossless compression of medical images through entropy coded DPCM," *IEEE Trans. Med. Imag.*, 1994, vol. 13, no. 3, pp. 538-548.

[5] E. R. Ritenour and A. D. A. Maidment, "Lossy compression should not be used in certain imaging applications such as chest radiography," *Medical physics*, vol. 26, pp. 1773-1775, 1999.

[6] G. Coatrieux, "Contribution à la sécurité d'images médicales par tatouage," *Ph.D. Thesis, Université de Rennes 1*, Rennes, 2002.

[7] B. Schneier, *Applied cryptography: protocols, algorithms, and source code in C*, Wiley, 1996.

[8] B. Preneel, "MACs and hash functions: State of the art," *Information Security Technical Report*, vol. 2, no. 2, pp. 33-43, 1997.

[9] X. Guo and T. Zhuang, "Lossless Watermarking for Verifying the Integrity of Medical Images with Tamper Localization," *Journal of Digital Imaging*, vol. 22, no. 6, pp. 620-628, 2009.

[10] H. Farid, "Image forgery detection," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16-25, 2009.

[11] C. Li and L. Liu, "An Image Authentication Scheme with Localization and Recovery," *in Proc. Int. Conf. Image and Signal Proc.*, 2008, vol. 5, pp. 669-673.

[12] H. Yang and A. C. Kot, "Binary Image Authentication With Tampering Localization by Embedding Cryptographic Signature and Block Identifier," *IEEE Signal Processing Letters*, vol. 13, no. 12, pp. 741-744, 2006.

[13] G. Coatrieux, H. Maitre, and B. Sankur, "Strict integrity control of biomedical images," in *Proc. of SPIE Security and Watermarking of Multimedia Contents III*, 2001, vol. 4314, pp. 229-240.

[14] M. P. Queluz, "Authentication of digital images and video: Generic models and a new contribution," *Signal Processing: Image Communication*, vol. 16, no. 5, pp. 461-475, Jan. 2001.

[15] X. Qi and J. Qi, "A robust content-based digital image watermarking scheme," *Signal Processing*, vol. 87, no. 6, pp. 1264-1280, Juin 2007.

[16] E. Mwangi, "A Geometric Attack Resistant Image Watermarking Scheme Based on Invariant Centroids," *in Proc. IEEE Int. Symp. Signal Proc. and Inf. Tech.*, 2007, pp. 190-193.

[17] A. Giakoumaki, S. Pavlopoulos, and D. Koutsouris, "A medical image watermarking scheme based on wavelet transform," in Proc. IEEE EMBS Int. Conf. 2003, vol. 1, pp. 856-859 .

[18] 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, 2009.

[19] A. Wakatani, "Digital watermarking for ROI medical images by using compressed signature image," *in Proc. Annu. IEEE Hawaii Int. Conf. System Sciences*, 2002, pp. 2043-2048.

[20] Hyung-Kyo Lee, Hee-Jung Kim, Ki-Ryong Kwon, and Jong-Keuk Lee, "Digital watermarking of medical image using ROI information," *in Proc. IEEE Int. Workshop Healthcom*, 2005, pp. 404-407.

[21] K.-H. Chiang, K.-C. Chang-Chien, R.-F. Chang, and H.-Y. Yen, "Tamper Detection and Restoring System for Medical Images Using Wavelet-based Reversible Data Embedding," *Journal of Digital Imaging*, vol. 21, no. 1, pp. 77-90, Mar. 2008.

[22] X. Zhao, A. T. S. Ho, H. Treharne, V. Pankajakshan, C. Culnane, and W. Jiang, "A Novel Semi-Fragile Image Watermarking, Authentication and Self-Restoration Technique Using the Slant Transform," *in Proc. Int. Conf. Intel. Inf. Hiding Mult. Signal Proc.*, vol. 1, 2007, pp. 283-286.

[23] R. Chamlawi, I. Usman, and A. Khan, "Dual watermarking method for secure image authentication and recovery," *in Proc. IEEE Int. Multitopic Conf.*, 2009, pp. 1-4.

[24] H.-J. He, J.-S. Zhang, and H.-M. Tai, "Self-recovery Fragile Watermarking Using Block-Neighborhood Tampering Characterization," *in Proc. Information Hiding*, 2009, pp. 132-145.

[25] OM Al-Qershi, BE Khoo, "Authentication and data hiding using a hybrid ROI-based watermarking scheme for DICOM images," J Digit Imaging., no. 24, vol1, pp. 114-25, 2011.

[26] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Hash-based identification of sparse image tampering," *IEEE Trans. Imag. Proc.*, vol. 18, no. 11, pp. 2491–2504, 2009.

[27] I. Avcibas, S. Bayram, N. Memon, M. Ramkumar, and B. Sankur, "A classifier design for detecting image manipulations," *in Proc. IEEE Int. Conf. Imag. Proc.*, vol. 4, pp. 2645-2648.

[28] H. Farid and S. Lyu, "Higher-order Wavelet Statistics and their Application to Digital Forensics," in Proc. Int. Conf. Computer Vision and Pattern Recognition, 2003, vol. 8, p. 94.

[29] S. Bayram, I. Avcibas, B. Sankur, and N. Memon, "Image manipulation detection," *Journal Electronic Imaging*, vol. 15, no. 4, pp. 041102-17, Oct. 2006.

[30] H. Huang, G. Coatrieux, H. Shu, L. Luo, Ch. Roux, "Blind Integrity Verification of Medical Images," IEEE Trans. on Inf. Tech. in Biom., Vol. 16, no. 6, pp. 1122-1126, 2012.

[31] A. Giakoumaki, S. Pavlopoulos, D. Koutsouris, "Multiple Image Watermarking Applied to Health Information Management," IEEE Trans. on Inf. Tech. in Biom. vol. 10, no. 4, pp.722-732, 2006.

[32] H. Huang, G. Coatrieux, H. Z. Shu, L. M. Luo, and C. Roux, "Medical image tamper approximation based on an image moment signature," *in Proc. Int. Conf. Healthcom*, 2010, pp. 254-259.

[33] G. Coatrieux, J. Montagner, H. Huang, Ch. Roux, "Mixed Reversible and RONI Watermarking for Medical Image Reliability Protection," *in Proc. IEEE EMBS Int. Conf.*, 2007, pp. 5653 - 5656.

[34] G. Coatrieux H. Maître, Y. Rolland, "Tatouage d'images médicales : perception d'une marque," in Proc. Int. Conf. Sc. Elec. Tech. Inf. Tel., 2003.

[35] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Pearson Education India.

[36] A. Leon-Garcia, *Probability and random processes for electrical engineering*. Addison-Wesley, 1994.

[37] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Th.*, vol. 8, no. 2, pp. 179-187, 1962.

[38] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *Journal CVGIP: Graphical Models and Imag. Proc.*, vol. 54, no. 5, pp. 438-460, 1992.

[39] D. R. Mukundan, K. R. Rao, and K. R. Ramakrishnan, *Moment Functions in Image Analysis*. World Scientific Publishing Co Pte Ltd, Singapore, 1998.

[40] U. H. . Kressel, "Pairwise classification and support vector machines," in *Advances in kernel methods*, 1999, pp. 255–268.

[41] L. Wang, *Support vector machines: theory and applications*. Springer, 2005.

[42] B. Xu, J. Wang, X. Liu, and Z. Zhang, "Passive Steganalysis Using Image Quality Metrics and Multi-class Support Vector Machine," in Proc. *Int. Conf. Natural Computation*, Los Alamitos, CA, USA, 2007, vol. 3, pp. 215-220.

[43] D. Fu, Y. Q. Shi, and W. Su, "Detection of Image Splicing Based on Hilbert-Huang Transform and Moments of Characteristic Functions with Wavelet Decomposition," *Digital Watermarking*, vol. 4283, pp. 177-187, 2006.

[44] D. J. Sebald and J. A. Bucklew, "Support vector machine techniques for nonlinear equalization," *IEEE Trans. Sig. Proc.*, vol. 48, no. 11, pp. 3217-3226, 2000.

[45] G. Quellec, S. R. Russell, and M. D. Abramoff, "Optimal filter framework for automated, instantaneous detection of lesions in retinal images," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, p. 523, 2011.

[46] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Imag. Proc.*, , vol. 13, no. 9, pp. 1200-1212, 2004.

[47] G. Beer, "Topological completeness of function spaces arising in the Hausdorff approximation of functions," Canad. Math. Bull., vol. 35, pp.439-448, 1992.

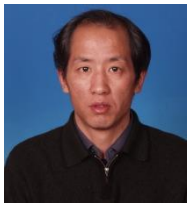
Gouenou Coatrieux (M'06-SM'13) received the Ph.D. degree in signal processing and telecommunication from the University of Rennes I, Rennes, France, in collaboration with the Institute Mines-Telecom, Telecom Paris-Tech, Paris, France, in 2002, in the field of watermarking in medical imaging.

He is currently an Associate Professor in the Department of Information and Image Processing, Institute Mines-Telecom, Telecom Bretagne, Brest, France. He conducts his research in the Laboratory of Medical Information Processing, Institut National de la Santé et de la Recherche Médicale, Brest. His primary research interests concern medical information system security, watermarking, electronic patient records, and healthcare knowledge management.

Dr. Coatrieux is actually an Associate Editor of IEEE TRANSACTIONS IN INFORMATION TECHNOLOGY and *Digital Signal Processing*, and the *Biomedical Engineering and Research*. He is also member of the International Federation for Medical and Biological Engineering "Global Citizen Safety and Security Working Group" and has contributed to the Technical Committee of "Information Technology for Health" of the IEEE Engineering in Medicine and Biology Society.

Hui Huang received the B.S. and M.S. degrees in biomedical engineering from Southeast University, Nanjing, China, in 2004 and 2007, respectively, and the Ph.D. degree in information and image processing from the University of Rennes I, Rennes, France, in collaboration with the Institut Mines-Telecom, Telecom Bretagne, Brest, France, in 2011.

She is currently an Assistant Researcher at Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China. Her research interests include medical information system security and medical image processing.



Huazhong Shu (M'00–SM'06) received the B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1987, and the Ph.D. degree in numerical analysis from the University of Rennes I, Rennes, France, in 1992.

He is currently a Professor in the Department of Computer Science and Engineering of Southeast University, Nanjing, China. His recent work concentrates on the image analysis, pattern recognition, and fast algorithms of digital signal processing.

Limin Luo (M'90–SM'97) received the Ph.D. degree from the University of Rennes I, Rennes, France, in 1986.

He is currently a Professor in the Department of Computer Science and Engineering, Southeast University, Nanjing, China. His current research interests include medical imaging, image analysis, computer- assisted systems for diagnosis and therapy in medicine, and computer vision.

Christian Roux (F'05) received the Aggregation degree in physics from the Ecole Normale Supérieure, Cachan, France, in 1978, and the Ph.D. degree from the Institut National Polytechnique, Grenoble, France, in 1980. In 1982, he joined the Institut Mines-Telecom, Telecom Bretagne, Brest, France, where he became an Associate Professor in 1987 and has been a Professor since 1987. He was a Visiting Professor with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania during 1992–1993, and a Distinguished International Research Fellow with the Department of Electrical Engineering, University of Calgary, Canada, in 1996 and 2003. He is the Founding Director of the Laboratoire de Traitement de l'Information Médicale, Institut National de la Santé et de la Recherche Médicale (INSERM), Telecom Bretagne, UBO, Brest, France. He has published around 160 papers and holds nine patents. His research interests focus on advanced medical information processing, and spatial, temporal, and functional information modeling and analysis in medical images, with applications in various medical domains including orthopedics, gastroenterology, ophthalmology, cardiology, and nuclear medicine.

Dr. Roux was an Associate Editor of the IEEE TRANSACTIONS ON MEDICAL IMAGING during 1993–2000. He is an Editorial Board Member of the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY and the PROCEEDINGS OF THE IEEE and is the Chairman of the GRETSI Program Committee. He served as the President of the IEEE Engineering inMedicine and Biology Society (EMBS) in 2001. He is the Founding Co-Chair of the IEEE EMBS International Summer School, Berder Island, France, and received the IEEE EMBS Award in 2003 and the INSERM Award for basic research in 2006.