Journal of Digital Imaging

A Region-Based Lossless Watermarking Scheme for Enhancing Security of Medical Data

Xiaotao Guo and Tian-ge Zhuang

This paper presents a lossless watermarking scheme in the sense that the original image can be exactly recovered from the watermarked one, with the purpose of verifying the integrity and authenticity of medical images. In addition, the scheme has the capability of not introducing any embedding-induced distortion in the region of interest (ROI) of a medical image. Difference expansion of adjacent pixel values is employed to embed several bits. A region of embedding, which is represented by a polygon, is chosen intentionally to prevent introducing embedding distortion in the ROI. Only the vertex information of a polygon is transmitted to the decoder for reconstructing the embedding region, which improves the embedding capacity considerably. The digital signature of the whole image is embedded for verifying the integrity of the image. An identifier presented in electronic patient record (EPR) is embedded for verifying the authenticity by simultaneously processing the watermarked image and the EPR. Combining with fingerprint system, patient's fingerprint information is embedded into several image slices and then extracted for verifying the authenticity.

KEY WORDS: Watermarking, telemedicine, security, integrity, confidentiality, image authentication, PACS, ROI

INTRODUCTION

In a modern integrated health care environment, digital information systems such as a hospital information system (HIS), picture archive and communication systems (PACS), and EPR system play an even more important role than ever. Compared with its analogy counterpart, the digital representation of medical data has a lot of advantages, such as easy compression and transmission, or image enhancing. On the other hand, with current techniques, it is fairly easy for malicious adversary to intercept or tamper sensitive medical data when the public network (e.g., the Internet) is being used for telemedicine. It is common view that there is an urgent need of security measures in medical information system.[1]

Digital watermarking, which imperceptibly embeds information within a host signal (such as image, audio, or video), is an emerging technique for protecting multimedia data.[2] When applied to medical environments, the watermarked image can still conform to Digital Imaging and Communications in Medicine (DICOM) format. The security information can adhere to the image even if the image format is changed. Furthermore, the property of imperceptibility makes an unauthorized person more difficult to intercept or attack the watermark information hiding in the image.

Motivated by the advantages mentioned above, some researchers already applied watermarking technique to medical data. Zhou et al. presented a watermarking method for verifying the authenticity and integrity of digital mammography images.[3] To embed the digital envelope into the image, the least significant bit (LSB) of one

From the Department of Biomedical Engineering, Shanghai Jiaotong University, Shanghai, 200240, China.

Correspondence to: Xiaotao Guo, Department of Biomedical Engineering, Shanghai Jiaotong University, Shanghai, 200240, China; tel: +86-21-52540226; fax: +86-21-52540226; e-mail: gxt@sjtu.org

Online publication 10 July 2007

doi: 10.1007/s10278-007-9043-6

random pixel of the mammogram is replaced by 1 bit of the digital envelope bit stream. Instead of the whole image data, only partial image data (not including LSB plane) is used for verifying integrity. Other researchers adapted digital watermarking for interleaving patient information with medical images to reduce storage and transmission overheads.[4] Again, the LSBs of image pixels are replaced for embedding. Chao et al. proposed a discrete cosine transform (DCT)-based data-hiding technique,[5] which is capable of hiding those EPR-related data into a marked image. The information was embedded in the quantized DCT coefficients. The drawback of the above watermarking approaches is that the original medical image is distorted in a non-invertible manner, such as bit replacement, truncation, or quantization. It is impossible for the watermark decoder to recover the original image.

As in lossy compression of radiologic images,[6] presently, there exists no legal standards for regulating how much distortion induced by watermarking system can be accepted. To be acceptable, a watermarking system requires thorough clinical validation tests. Such tests must be carried out on a large number of images and should involve a large number of clinicians to assure that the diagnostic accuracy is not jeopardized by such distortion. Coatrieux et al. proposed an alternate approach by separating an image into a protection zone and an insertion zone to avoid compromising any diagnostic capability.[7] Lossless watermarking, which can recover the original image exactly, has drawn lots of interest recently.[8–14] Goljan et al. introduced distortion-free data embedding for images.[8] De Vleeschouwer et al. used a circular interpretation of bijective transformations of the histogram to embed data.[9] Tian proposed a reversible data-embedding method using the difference expansion of pairs of pixel values. [10] To ensure invertibility, a bi-level location map is compressed using JBIG2 method and transmitted as a part of the payload. Alattar generalized Tian's method to arbitrary vectors instead of pairs.[11] Celic et al. presented a reversible data-embedding algorithm by compressing quantization residues.[12] Zhou et al. presented two lossless data-embedding methods to embed digital signature to medical images.[13] The first method was based on compressing the LSBs of randomly selected image pixels; the other method was based on a regular/singular (RS) approach that was introduced in the work of Goljan et al.[8]

However, current reversible watermarking methods did not have region-selecting capability.[8–13] The embedding-induced distortion was distributed in the whole image, instead of some region. If they were applied in medical environments, the watermark extraction and the original image restoration must be performed to ensure that the diagnostic accuracy in the region of interest (ROI) was not compromised by the embedding-induced distortion. This might bring great inconvenience in practical medical applications.

In this paper, to follow up work,[14] we present a novel region-based lossless watermarking scheme, being capable of verifying authenticity and integrity of medical images. Furthermore, the watermark encoder can choose embedding regions at will without introducing any distortion in the ROI. The watermarked image may be used for diagnostic purposes as well as other medical applications, provided that the embedding region does not intersect with the ROI. Our experimental results demonstrate that such scheme can hide large amount of data while keeping distortion level low enough. Any modification in the watermarked image can be detected. Authentication can be enhanced by combining with a fingerprint system.

MATERIALS AND METHODS

The proposed region-based lossless watermarking scheme is based on the difference expansion and region selection.[10,11] Most of the medical images exhibit a high spatial correlation among the values of neighboring pixels. In the smooth area of the image, the difference between the values of two adjacent pixels is rather small. One-bit information b can be embedded into the binary representation of the difference value h by appending b after its LSB. This operation is called difference expansion. A difference value h is said to be expandable if the difference expansion operation does not cause the resultant pixel values underflow or overflow.

After the difference expansion, the expanded difference value might not be expandable. On the decoder side, to check whether the difference value is expandable does not tell whether the

original difference value has been selected for difference expansion. To address this problem, in the work of Tian and Alattar,[10,11] changeable difference values (whose LSB can be modified without causing the resultant pixel values underflow or overflow) are used for the decoder to collect and decode location map. By definition, an expandable difference value is also changeable. A location map that contains the location information of all selected expandable difference values for embedding is compressed using a lossless compression algorithm, such as Joint Bi-level Image Experts Group (JBIG) or an arithmetic compression algorithm. During data embedding, all changeable difference values are modified by either adding a new LSB or modifying its LSB.

One of the drawbacks of this approach is that the overhead of keeping a location map is expensive for low capacity. It is also hard to estimate the embedding capacity because the size of the payload that can be embedded depends on how well their location map can be compressed. Another undesired feature of this approach is that the embedding-induced distortion is distributed in the whole image. During data embedding, all changeable difference values (except those selected for expansion) have to have their LSBs modified, although they do not contribute to capacity size. It is rather difficult to restrict the embedding distortion inside a given region, which is desirable for medical application. For more details, please refer to Tian and Alattar[10,11].

To overcome the drawbacks mentioned above, a new reversible watermarking method based on difference expansion is presented in this paper with medical data as a main concern. It has advantages of being capable of restricting the embedding-induced distortion inside a given region and being able to control embedding capacity easily. For a given watermark payload, a region of embedding (ROE) may be chosen intentionally to prevent introducing any distortion inside the ROI. It is represented by a polygon, whose vertex information is transmitted to the decoder for reconstructing the embedding region. If the total net embedding capacity of all the selected ROEs is smaller than the required payload, another new ROE can be easily added.

Difference Expansion Transform

Let us assume that the original image is a grayscale image I, whose pixel values are from the range $[0, L]$. L depends on the maximum bit depth for representing a pixel. For example, L is equal to 255 for 8-bit ultrasound (US) images. L is equal to 4,095 for 12-bit computed tomography (CT) or magnetic resonance imaging (MRI) images.

Quads

A quad is a vector $u=(u_0, u_1, u_2, u_3)$ formed from 2×2 adjacent pixel values according to a predetermined order. This order may serve as a security key. All quads do not overlap each other, i.e., each pixel exists in only one quad.

Difference expansion transform

The forward difference expansion transform υ=f(u) for the vector $u=(u_0,u_1,u_2,u_3)$ is defined as:

$$\begin{aligned} \upsilon_0 &= \left\lfloor \frac{u_0+u_1+u_2+u_3}{4} \right\rfloor \\ \upsilon_1 &= u_1 - u_0 \\ \upsilon_2 &= u_2 - u_0 \\ \upsilon_3 &= u_3 - u_0 \end{aligned} \tag{1}$$

where $\lfloor \cdot \rfloor$ is the least nearest integer, and $\upsilon=(\upsilon_0, \upsilon_1, \upsilon_2, \upsilon_3)$.[10,11] Note that υ_0 is the average of u, whereas υ_1,υ_2, and υ_3 are the differences between u_1, u_2, u_3, and u_0, respectively.

The inverse difference expansion transform $u=f^{-1}(\upsilon)$ for the transformed quad $\upsilon=(\upsilon_0, \upsilon_1, \upsilon_2, \upsilon_3)$ is defined as

$$\begin{aligned} u_0 &= \upsilon_0 - \left\lfloor \frac{\upsilon_1+\upsilon_2+\upsilon_3}{4} \right\rfloor \\ u_1 &= \upsilon_1 + u_0 \\ u_2 &= \upsilon_2 + u_0 \\ u_3 &= \upsilon_3 + u_0 \end{aligned} \tag{2}$$

Definition

The quad $u=(u_0, u_1, u_2, u_3)$ is said to be expandable if, for all values of b_1, b_2, and $b_3 \in \{0,1\}$, $\upsilon=f(u)$ can be modified to produce $\widetilde{\upsilon} = (\upsilon_0, \widetilde{\upsilon}_1, \widetilde{\upsilon}_2, \widetilde{\upsilon}_3)$ according to Eq. 3 below

without causing overflow or underflow in $\widetilde{u} = f^{-1}(\widetilde{v})$.

$$\begin{aligned} v_0 &= \left\lfloor \tfrac{u_0+u_1+u_2+u_3}{4} \right\rfloor \\ \widetilde{v}_1 &= 2 \times v_1 + b_1 \\ \widetilde{v}_2 &= 2 \times v_2 + b_2 \\ \widetilde{v}_3 &= 2 \times v_3 + b_3 \end{aligned} \tag{3}$$

Note that each of $\widetilde{v}_1, \widetilde{v}_2$ and $\widetilde{v}_3$ is a 1-bit left-shifted version of the original value v_1,v_2, and v_3, respectively. For a quad, 3 bits of information can be reversibly embedded.

The pixel value in the image I are arranged into the set of quads $U=\{u_i, i=1...n\}$. The quads in U can be classified into two groups according to the definition above. The first group S_1 contains all expandable quads whose $v_1 \le T_1$, $v_2 \le T_2$, and $v_3 \le T_3$ (T_1, T_2, and T_3 are predefined thresholds). The second group S_2 contains the rest of the quads. Let us now identify the quads of U using a binary location map M whose entries are 1s and 0s, where symbol 1 indicates S_1 and symbol 0 indicates S_2.

Region of Embedding

A region of embedding is a region where all the quads inside belong to group S_1. Thus, in the location map M, all entries in a ROE are 1s. A ROE can be defined automatically or delineated semi-automatically by the radiologist with minimal interaction and in a standard way. If the total number of expandable quads in a ROE is N_E, $3N_E$ bits can be reversibly embedded in all these quads. However, to let the decoder reconstruct the ROE for extracting information, the description information must be transmitted to the decoder. Let B_c denote the amount of information for describing a ROE. To efficiently reduce B_c, a polygon is used in this study to represent a ROE. A polygon is completely characterized by the number of vertex n_v and the vertex coordinates $v(x, y)$. The net embedding capacity I_E for a ROE is obtained by

$$I_E = 3N_E - \|n_v\| - n_v \times \|v(x,y)\| \tag{4}$$

where $\|x\|$ denotes the bit length for describing x.

Data-Embedding Procedure

1. Compute the hash value H of the original image I. It is computationally difficult to find two images that have the same hash value. Instead of using partial image data, e.g., the bit planes of MSB, to compute the hash value in some watermarking schemes,[3] the whole image data are computed in this study to offer stricter integrity verification.

$$H = H_{MD}(I) \tag{5}$$

where $H_{MD}(\bullet)$ denotes the MD5 hashing function.[15]

2. Produce the digital signature DS based on the above hash value H.

$$DS = RSA_E\left(K_{priv}, H\right) \tag{6}$$

where RSA_E denotes the RAS public-key encryption algorithm,[16] and K_{priv} is the private key at the transmitter site.

3. Suppose the additional data we want to embed is data D (e.g., confidential patient information, or annotation information), the payload P is the summation of D and DS with total length L_{EM},

$$P = D \oplus DS \tag{7}$$

where $\oplus$ denotes a concatenating operator.

4. Scan the image in predefined order (e.g., raster order). Form the set of quads $U=\{u_i, i=1...n\}$ from the image I such that each u_i is formed from non-overlap 2×2 adjacent pixel values. Calculate $v=(v_0, v_1, v_2, v_3)$ using Eq. 1. Use condition 3 and the predefined threshold (T_1, T_2, and T_3) to divide U into two sets $S_1$and S_2. Form the location map, M.
5. Choose a ROE $R^{(i)}$ (initially set at $i=1$) automatically or delineated semi-automatically by the radiologist such that all the quads inside belong to set S_1. The radiologist expert can mark some vertex points for such a ROE. Note in this study that any ROE should avoid intersecting with the ROI.
6. Calculate the net embedding capacity for all $R^{(i)}$ according to Eq. 4. If it is smaller than the payload size L_{EM}, then go back to the last step and add a new ROE until the total net embedding capacity of all the ROEs is greater or equal to L_{EM}.

7. Let n_p denote the number of the selected ROEs. For describing all $R^{(i)}$, $i \in [1, ..., n_p]$, the bitstream B_{ROE} is formed by

$$B_{ROE} = B_0 \oplus B_c^{(1)} \oplus B_c^{(2)} \cdots \oplus B_c^{(n_p)} \quad (8)$$

where B_0 denotes the bitstream for representing n_p. Let L_c denote the total length of B_{ROE}.

8. Encrypt bitstream B_{ROE} into bitstream $\widehat{B}_{ROE}$; the conventional private or public key infrastructure system can be used to ensure the data security.

$$\widehat{B}_{ROE} = RC_E(K_{wm}, B_{ROE}) \quad (9)$$

where RC_E denotes the RC4 stream encryption algorithm, and K_{wm} is the session key used for transmission.

9. From predefined location $L(x, y)$(e.g., the bottom line of I), extract the LSB of the pixel value sequentially with the total length L_c. Concatenate these bits to form bitstream B_{LSB}. From the same location $L(x, y)$, replace the LSB of the pixel value by 1 bit of $\widehat{B}_{ROE}$ sequentially and repeat until all L_c bits are replaced.
10. Combine the watermark payload P and B_{LSB} to form bitstream B.

$$B = B_{LSB} \oplus P \quad (10)$$

11. Encrypt bitstream B into bitstream B_E to ensure the data security.

$$B_E = RC_E(K_{wm}, B) \quad (11)$$

12. Scan the quads inside all ROE in predefined order (e.g., from up to down, from left to right). For each scanned quad u, compute its forward transform, $v = f(u)$, embed 3 bit of information b_1,b_2, and b_3 (which is extracted from B_E sequentially) into v using difference expansion as described in Eq. 3. Compute the inverse transform of the resulting vector $\widetilde{v}$ to produce the watermarked $\widetilde{u}$. Replace the pixel values in u with the corresponding values from $\widetilde{u}$.
13. Repeat the last step until all bits in B_E are embedded; the resulting image is the watermarked image.

Data Extracting and Verifying Procedure

Suppose that the watermarked image is I_w.

1. Extract LSB of the pixel value from predefined location $L(x, y)$ sequentially. Concatenate these bits to form bitstream $\widehat{B}_{ROE}$.
2. Decrypt the bitstream using the secret key to obtain decrypted bitstream B'_{ROE},

$$B'_{ROE} = RC_D\left(K_{wm}, \widehat{B}_{ROE}\right) \quad (12)$$

where RC_D denotes the RC4 stream decryption algorithm, K_{wm} is the session key.

3. From B'_{ROE}, extract the ROE information (including number of the ROEs, number of the vertexes and the corresponding vertex coordinates), then reconstruct all ROE in the image I_w.
4. Scan the quads inside all ROE in predefined order (same as in the embedding procedure). For each scanned quad $\widetilde{u} = (u_0, u_1, u_2, u_3)$, compute its forward transform, $\widetilde{v} = (\widetilde{v}_0, \widetilde{v}_1, \widetilde{v}_2, \widetilde{v}_3)$, extract the LSBs of $\widetilde{v}_1$, $\widetilde{v}_3$, and $\widetilde{v}_2$. Concatenate these bits to form bitstream B'_E.
5. Decrypt B'_E using the secret key to obtain decrypted bitstream B',

$$B' = RC_D\left(K_{wm}, B'_E\right) \quad (13)$$

6. From B', then extract data D', digital signature DS', and embedded LSB bitstream B'_{LSB}, respectively.
7. Decrypt $DS'$to obtain the hash value H'

$$H' = RSA_D\left(K_{pub}, DS'\right) \quad (14)$$

where RSA_D denotes the RSA public-key decryption algorithm, K_{pub} is the public key of the transmitter site.

8. Restore all the modified information to produce the recovered image I'; this involves two main steps

   (a) Restoring all the quads in all the ROEs. Restore the original values of v_1,v_2, and v_3 for v as follows:

$$v_1 = \left\lfloor \frac{\widetilde{v}_1}{2} \right\rfloor, v_2 = \left\lfloor \frac{\widetilde{v}_2}{2} \right\rfloor, v_3 = \left\lfloor \frac{\widetilde{v}_3}{2} \right\rfloor \quad (15)$$

Calculate its inverse transform using Eq. 2 to obtain $\hat{u}$; replace the pixel values in $\tilde{u}$ with the corresponding values from $\hat{u}$.

(b) Restoring the modified LSBs.
From the location $L(x, y)$, replace the LSB of the pixel value in I_w by 1 bit of B'_{LSB} sequentially and repeat until all the bits are replaced.

9. Compute the hash value $\hat{H}$ for the restored image I',

$$\hat{H} = H_{MD}(I') \tag{16}$$

Comparing it with the hash value H' extracted in step 7, if they do not match each other, the image must have been modified with high probability.

10. If some authentication code is embedded in the payload data D', such data can be used for verifying the authenticity of the data. For example, the authentication code in DICOM header or EPR can be simultaneously embedded into the image itself to ensure authenticity. In the following, we also combine the fingerprint technique to enhance the authenticity of image data.

RESULTS

To evaluate the performance of the proposed data-embedding method, firstly, we collected medical images from three different modalities, i.e., CT, MRI, and US. The size of CT, MRI, and US images are 512×512 (12 bits), 256×256 (12 bits), and 640×480 (8 bits), respectively. These images include head CT and MRI images, L-Spine CT images, kidney US images, etc. In the experiments, threshold T_1, T_2, and T_3 are all set equal to T. In Eq. 4, 8 bits are used to represent the number of vertex; 20 bits are used to represent each vertex coordinates.

We examined the embedding capacity outside the ROI under different distortion level. For 8-bpp (bit per pixel) US images, it is easy to reversibly embed more than 1,000 bits of information (outside the ROI) when the threshold T is below 4. For 12-bpp MRI and CT images, it is also easy to embed the same amount of information (outside the ROI) when the threshold T is set properly. To embed more information, one can add more ROEs. If the total embedding capacity for all the ROEs under the threshold T is less than the given payload, one can increase T to create more expandable quads, thus, create more selectable ROEs. When the threshold is set to be a very small value, the perceptual difference between the watermarked image and the original image are hard to observe. When the threshold is increased, the embedding capacity is also increased. When a large amount of information is embedded in the image, some perceptual difference may be observed. There is a tradeoff between capacity and distortion. When the threshold is set the same for the same size image, the embedding capacity depends highly on the nature of the image itself. In general, an image with larger smooth area (outside the ROI) has higher capacity. Fortunately, a large number of medical images have some smooth regions outside the ROI. The actual embedding capacity depends on the total net embedding capacity for all the selected ROEs (refer to Eq. 4). In general, a larger image that has a larger or smoother area (outside the ROI) can bear more information than the same size or smaller image that has a smaller smooth area. We then tested some computed radiography (CR) images, including hand CR and chest CR. These images have larger size, but are much noisy than CT and MR images. Some large ROEs can be found when the threshold is set properly.

For illustrative purpose, we choose a typical 640×480, 8-bpp ultrasound image to demonstrate the data embedding, extracting, and verifying procedure. Figure 1a shows the original ultrasound image. The expandable quads are marked with gray color when the threshold $T = 4$ is shown in Figure 1b. Other color may be chosen as well, as long as it can be distinguished from the background image. Depending on the watermark payload, the threshold T can be chosen properly to minimize the embedding distortion. Three polygons (black color) are chosen as the ROEs. Figure 1c shows the watermarked image. It has been embedded with 6,428 bits of information, consisting of the original image's hash value, "6cac2ace305cea4937a4922647137091," the patient information, etc. No perceptual difference can be observed between the watermarked image and the original one. Figure 1d shows the difference between the watermarked image and

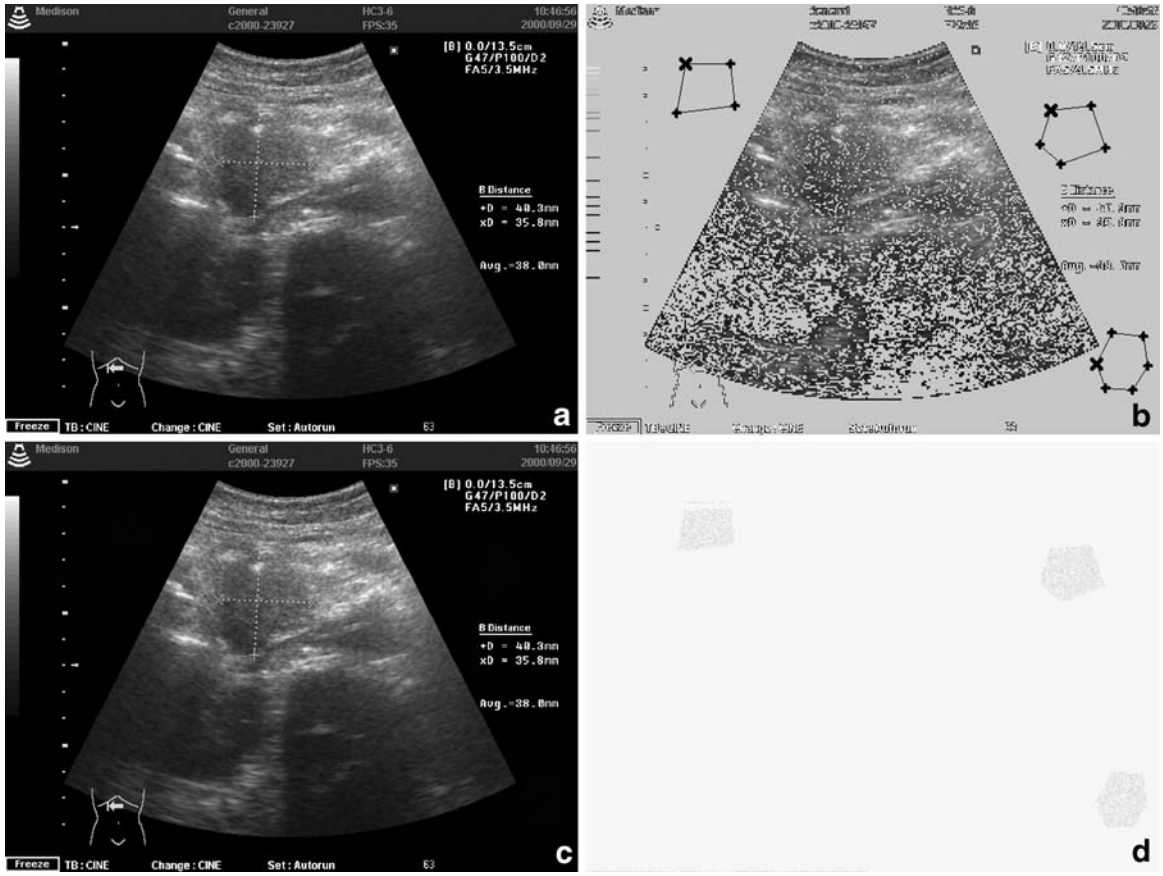

Fig. 1. Watermark embedding. a Original 640 × 480, 8-bpp ultrasound image. b Expandable quads (*gray color*) and three selected polygon ROEs. c Watermarked image. d Difference between the watermarked and the original image (magnified by a factor of 32).

the original one (magnified by a factor of 32). Note that most of the embedding distortion is restricted inside the three ROEs. Some distortion can be observed at the bottom line also; this is due to some LSBs of the pixel values that had been modified in the embedding procedure (step 9). When peak signal to noise ratio (PSNR) is used as a distortion measure, our method can achieve 67.72 dB in this case.

Figure 2 illustrates the watermark extraction and restoration procedure. Figure 2a shows the reconstructed polygon ROEs from the watermarked image. The watermark decoder then forms the quads in all these ROEs to extract all the information. The embedded hash value "6cac2ace305cea4937a4922647137091" and the other information can all be successfully extracted. Figure 2b shows the restored image. The new hash value is computed from this restored image, which is identical to that of the extracted one.

The reversibility of our watermarking scheme can be verified by comparing the recovered image with the original image pixel by pixel, bit by bit, when the original image is available. It can also be verified by comparing the embedded hash value with the new hash value computed from the recovered image. If the watermarked image is not modified, both values must be identical. The recomputed hash value for the restored image (Fig. 2b) is identical to that of the extracted one, demonstrating that our watermarking scheme can recover the original image exactly.

To demonstrate that any modification in the watermarked image can be detected, two scenarios are tested, i.e., tampering in the ROI and non-ROI. In Figure 3a, we replaced some pixel values at the center of the ROI with a black spot

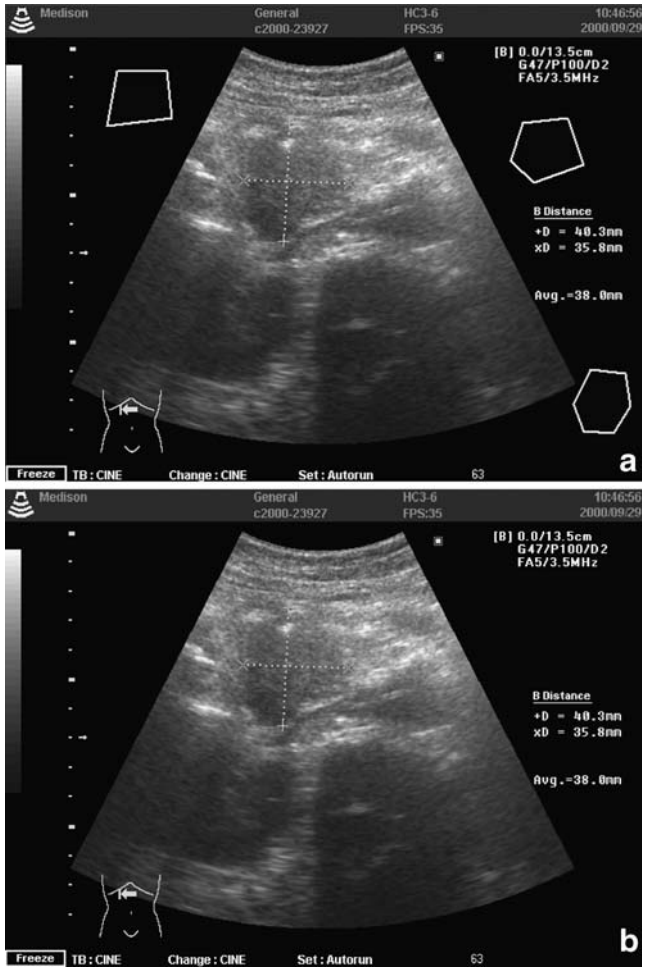

Fig. 2. Watermark extraction and restoration. a The reconstructed polygon ROEs (*white color*) from the watermarked image. b The recovered image: The hash value computed for this image is "6cac2ace305cea4937a4922647137091."

(for illustrative purpose) in the watermarked image. In Figure 3b, we modified the pixel values at the upper right corner such that the original date "2000/09/29" becomes "2000/09/20." In both cases, the watermark decoder can successfully extract the embedded information, including the hash value for the original host image, "6cac2ace305cea4937a4922647137091," and the patient information. After restoration, the recomputed hash value for both recovered image becomes "7597df00317a35def739c940bbc6984b" and "1dbc3dadc400587fac11978433aa47bd," respectively. These hash values are significantly different from the embedded original one. Because MD5 hashing algorithm is sensitive to any modification in input data, our watermarking scheme can detect even a single bit of alteration in the image data.

To guarantee that the image is authentic in PACS, firstly, we extract identification information about the image, such as patient name or hospital name that can also be retrieved in DICOM header or EPR database. We then embed such information into the image. If the image is not tampered, the identification information can be successfully extracted and compared with the corresponding information from the EPR. If they are not matching, some error or falsification can be easily observed. This feature is very helpful for checking errors in the manual entry of patient data.

Biometric information, such as fingerprint, is difficult to forge, providing concrete way for

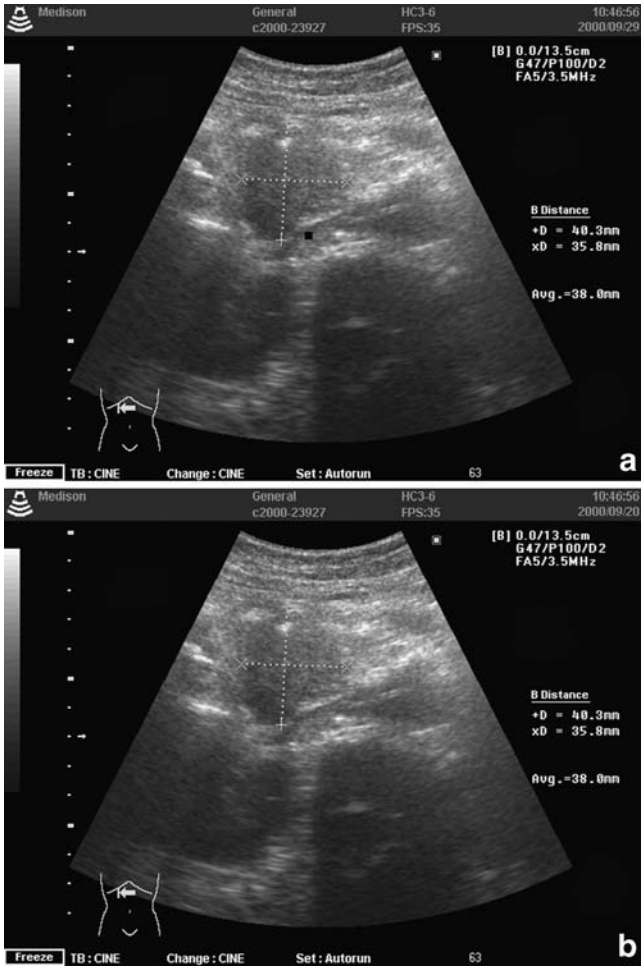

Fig. 3. Integrity verification. a The tampered image in the ROI: Some pixels are modified by a *black spot at the center*. The hash value computed for the recovered image is "7597df00317a35def739c940bbc6984b". b The tampered image in the non-ROI: The date "2000/09/29" at the right–upper corner is modified to "2000/09/20". The hash value computed for the recovered image is "1dbc3dadc400587fac11978433aa47bd".

authentication. In Figure 4, we combine fingerprint authentication system with our watermarking scheme to further enhance the authenticity of medical images. A patient's medical image data often contains several slices of images for CT or MRI modality. Firstly, the patient's fingerprint image is acquired via a standard fingerprint device. The device's output is a binary sequence (about 300 bytes) representing the fingerprint image's eigenvector. We take a dynamic bit allocation strategy to embed such large information while keeping the total distortion level still very low. The embedding capacity of each slice is calculated before embedding. The fingerprint's eigenvector information is distributed to each slice, proportional to its embedding capacity; in general, given a distortion level, a slice with larger capacity will be embedded with more bits. In Figure 5, the watermark is firstly extracted from each slice. The fingerprint image's eigenvector is then reconstructed by combining all the pieces in inverse order and input to the fingerprint authentication system as a reference. During the verifying process, the fingerprint authentication system calculates the eigenvector of the input fingerprint's image and compares it with the reference to decide if the two are matching according to some suitable criteria. Experimental results show that such system can unambiguously identify the right person who is associated with the medical images.

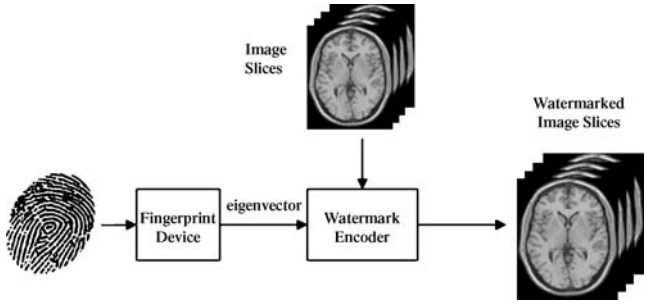

Fig. 4. Fingerprint information embedding process.

DISCUSSION

Comparing our work with other algorithms in the literature, some of the data-embedding methods are lossy,[3–5] i.e., the medical data was distorted in non-invertible manner. However, medical tradition is very strict with the quality of biomedical images. Alteration of the bit-field representing the image is, generally, not allowed. In the current lossless watermarking approaches,[8–14] the embedding distortion is distributed in the whole range of the image, instead of in some region. The watermark decoder has to restore the original image all the time to ensure diagnostic accuracy inside the ROI. This greatly restricts their wide usage in medical application. On the contrary, our method can be used for conveying additional information without restoring the original image every time, thanks to the features that it does not introduce any distortion inside the ROI. The exact recovery of the original image is required only for verifying the integrity of the image, which may be used only when important security issues occur. In addition, the embedding distortion is restricted only in a small area. It can be easily controlled by threshold *T*. Also, thanks to the region-based embedding strategy, the watermark information can always be successfully extracted as long as the modification of the watermarked image is outside the embedding region.

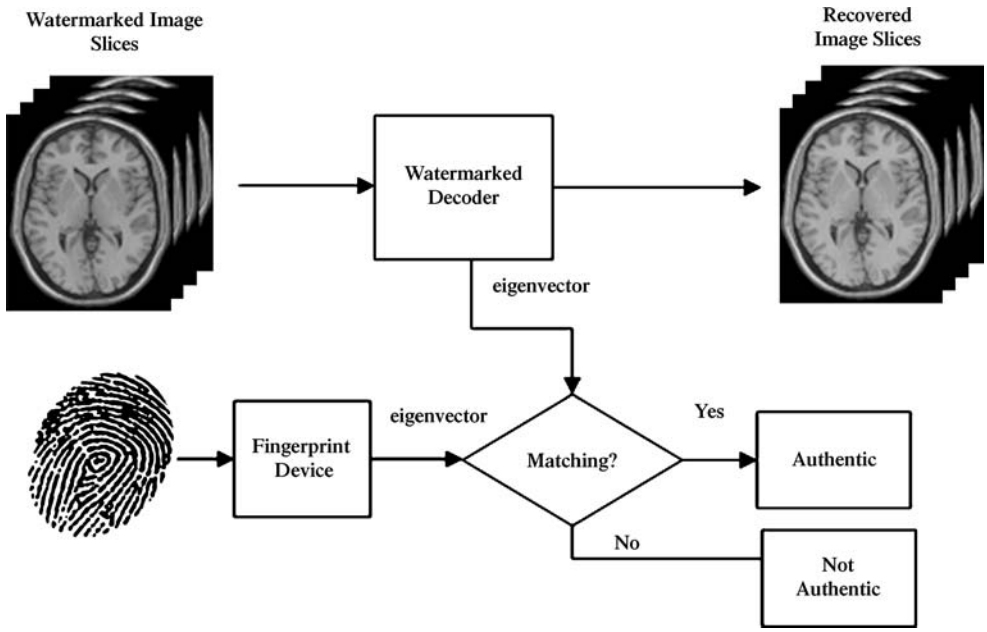

Fig. 5. Fingerprint information extracting and authenticity verifying process.

Our method can also be applied for some categories of medical images (such as captured micrographs or dermatology photography), where the entire field of view is of interest. In this case, some smooth region in the whole image can be selected as the ROEs. Our method may be just applied as the other reversible watermarking schemes.[8–14] For example, authentication code or patient's confidential information may be hidden in the images using our method to reduce storage and transmission overheads. However, caution must be taken before diagnosis. To ensure diagnostic accuracy, the original image restoration must be performed ahead of diagnosis. This is not a serious problem because the computational cost of the additional restoration process is very low in our method.

Although reversible watermark allows exact recovery of the original host image, it is still desirable to keep the embedding distortion to a minimum. One reason is that the ROI may be different for different applications. Ideally, we hope the difference between the watermarked image and the original image is imperceptible. This transparency property allows the watermarked image to be accessed by more than one group of users with more than one application. For example, it allows causal users (e.g., medical school students), who could not access to the watermark extraction and restoration, to incur a less penalty in image quality.

Multiple-layer embedding can be used to increase embedding capacity. For an already-embedded image, the data-embedding procedure can be performed more than once. This recursive embedding is possible because the embedding process is lossless, which means the input image can be recovered exactly after embedding. The secret key may be set different at a different layer, which allows access control at different level. Note that the ROEs at a different layer are allowed to overlap because the data embedding is independent for each layer.

There still exists a limit for the amount of information that can be embedded (outside the ROI) in an image. If the required payload is too large, some information has to be embedded inside the ROI. In this case, some mechanism (e.g., a removable visible watermark) may be used to notify the user that the quality of the image data in the ROI had been degraded by the data-embedding process. The exact recovery of the original image data in the ROI must be performed at this time.

One of the limitations of the proposed method is that the ROI cannot be identified automatically. Currently, it requires manual intervention. The reason is that the ROI might vary significantly for different organ or different modalities. However, if a prior knowledge is known for a particular family of images, this information may be used to reduce manual intervention. For example, if we know that the host image is a CT head image, it is reasonable to deduce that the region outside of the skull might be non-ROI. We may obtain the contour of the skull with the help of segmentation techniques and generate the ROEs automatically. It is interesting to note that the computational complexity of the watermark encoder and decoder in our method is asymmetric. The effort for the watermark embedding process is relatively high. However, once the watermark is embedded, the watermark extraction and restoration process is very simple.

CONCLUSION

In this paper, we addressed the secure issue for medical data. We pointed out some drawbacks of several current watermarking schemes used in medical environment, with emphasis on image quality for diagnosis accuracy. We then proposed a secure lossless watermarking scheme for medical images to ensure that the original image can be exactly recovered. In addition, the scheme has the capability of not introducing any embedding distortion in the region of interest of a medical image. The embedding capacity was evaluated under different distortion level for a large number of medical images. Experimental results indicate that such scheme achieves high embedding capacity with low level of embedding-induced distortion. A medical image's integrity is strictly verified using the digital signature embedded in the image. A medical image's authenticity is verified by simultaneously processing the watermarked image and the EPR. Patient's fingerprint information is further embedded into several image slices for enhancing authenticity. Our scheme can be extended to embed other biometric information into the image slices.

REFERENCES

1. Coatrieux G, Maitre H, et al: Relevance of watermarking in medical imaging. Proc IEEE EMBS Information Technology Applications in Biomedicine, 2000, pp 250–255

2. Petitcolas FAP, Anderson RJ, Kuhn MG: Information hiding—a survey. Proc IEEE 87:1062–1078, 1999

3. Zhou XQ, Huang HK, Lou SL: Authenticity and integrity of digital mammography images. IEEE Trans Med Imag 20:784–791, 2001

4. Rajendra Acharya U, Acharya D, et al: Compact storage of medical images with patient information. IEEE Trans Inf Technol Biomed 5:320–323, 2001

5. Chao HM, Hsu CM, Miaou SG: A data-hiding technique with authentication, integration, and confidentiality for electronic patient records. IEEE Trans Inf Technol Biomed 6:46–53, 2002

6. Wong S, Zaremba L, Gooden D, Huang HK: Radiologic image compression—a review. Proc IEEE 83:194–219, 1995

7. Coatrieux G, Maitre H, Sankur B: Strict integrity control of biomedical images. Proc SPIE Sec Watermarking Multimed Contents III 4314:229–240, 2001

8. Goljan M, Fridrich J, Du R: Distortion-free data embedding for images. The 4th Information Hiding Workshop 2137:27–41, 2001

9. De Vleeschouwer C, Delaigle JF, Macq B: Circular interpretation of bijective transformations in lossless watermarking for media asset management. IEEE Trans Multimed 5:97–105, 2003

10. Tian J: Reversible data embedding using a difference expansion. IEEE Trans CSVT 13:890–896, 2003

11. Alattar AM: Reversible watermark using the difference expansion of a generalized integer transform. IEEE Trans Image Process 13:1147–1156, 2004

12. Celik MU, Sharma G, Tekalp AM, Saber E: Lossless generalized-LSB data embedding. IEEE Trans Image Process 14:253–266, 2005

13. Zhou Z, Huang HK, Liu BJ: Digital signature embedding (DSE) for medical image integrity in a data grid off-site backup archive. Proc SPIE 5748:306–317, 2005

14. Guo X, Zhuang T: A lossless watermarking scheme for enhancing security of medical data in PACS. Proc SPIE 5033:350–359, 2003

15. Rivest RL: The MD5 message-digest algorithm. RFC 1321, 1992

16. Rivest RL, Shamir A, Adleman L: A method for obtaining digital signatures and public-key cryptosystems. Commun ACM 21:120–126, 1978