Authenticating Medical Images through Repetitive Index Modulation Based Watermarking

Chang-Tsun Li and Yue Li

Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
{ctli,yxl}@dcs.warwick.ac.uk

Abstract. In this work we propose a Repetitive Index Modulation (RIM) based digital watermarking scheme for authentication and integrity verification of medical images. Exploiting the fact that many types of medical images have significant background areas and medically meaningful Regions Of Interest (ROI), which represent the actual contents of the images, the scheme uses the contents of the ROI to create a content-dependent watermark and embeds the watermark in the background areas. Therefore when any pixel of the ROI is attacked, the watermark embedded in the background areas will be different from the watermark calculated according to the attacked contents, making the authentication unsuccessful. Because the creation of the watermark is content-dependent and the watermark is only embedded in the background areas, the proposed scheme can actually protect the content without distorting it.

Keywords: Medical image authentication, digital watermarking, data hiding, digital forensics, integrity verification.

1 Introduction

Due to the privacy concerns and authentication needs, many digital watermarking schemes [1, 2, 3, 6, 12] have been proposed to embed authentication data into the contents of medical images. Most methods [1, 2, 6, 11] require the least significant bits (LSBs) of the image pixels to be replaced with the authentication codes or watermarks. Although the distortion due to this kind of "lossy" watermark embedding is usually visually insignificant, medical images with watermarks embedded with this type of irreversible watermarking schemes may not be accepted as feasible evidence in the court of law, should medical disputes occur. Many reversible data hiding schemes [7, 10], although not specifically proposed for the purpose of medical image authentication, have been developed to facilitate reversible data hiding, in which the original images can be recovered after the hidden data is extracted from the watermarked images. A reversible watermarking scheme specifically developed for authenticating medical data has been proposed in [3]. The common problem with these reversible data hiding schemes is that, apart from the actual payload (i.e., the watermark, secret data, authentication codes, etc), side information for reconstructing the exact original image has to be embedded as well. The side information wastes limited embedding capacity and is usually the compressed form of the location map of

M. Sorell (Ed.): e-Forensics 2009, LNICST 8, pp. 153 – 159, 2009.

the original data that is expected to be affected by the embedding process. The waste of embedding capacity reduces the authentication power of the scheme and the resolution of tamper localization, as explained in [8]. Moreover, authentication schemes are also expected to be resistant against attacks, such as the Holliman-Memon counterfeiting attack, the birthday attack and the transplantation attack, by involving the contents in the watermarking process in a non-deterministic manner [5, 8]. Therefore schemes with high payload, high resolution of tamper localization, high security and zero distortion to the ROI are desirable.

2 Proposed Method

It is observed that, apart from the ROI, which represents the actual contents of images, many types of medical images have significant background areas. Exploiting this characteristic, we propose a new scheme, which uses the contents of the ROI to create a content-dependent watermark and embeds the watermark in the background areas without adding any embedding distortion to the ROI. Without loss of generality, we will use mammograms with gray level range [0, 255] in the presentation of this work. Because the background areas contain no information of interest and the gray levels of their pixels fall in the low end of the intensity range, wherein human eyes are not sensitive to variation, a greater degree of embedding can be carry out to strengthen security and/or increase resolution of tamper localization [8].

2.1 Segmentation

The mission of the image segmentation operation is that when given either the original image, I_o, during the *watermarking* process or the watermarked image, I_w, during the *authentication* process as input, the segmentation function should partition the input image into the same bi-level output image, with one level corresponding to the background areas and the other to the ROIs. Figure 1(a) shows a typical mammogram with intensity represented with 8 bits. We can see that it has a dark background with intensity below 30 and a significantly brighter area of a breast (ROI) with the intensity of most pixels above 100. Since we will embed the watermark in the background area and the distortion due to embedding will not raise the intensity significantly, so a threshold between 50 and 100 for partitioning the images is a reasonable value. However, due to the fact that mammograms may be taken at different times with different equipments under various imaging conditions, using a heuristic constant threshold to segment mammograms is not feasible. So we propose to use *moment-preserving thresholding* [11], which is content-dependent, to perform the segmentation task.

Given a gray-scale image, I, with $X \times Y$ pixels, we define the intensity / gray scale at pixel (x, y) as $I(x, y)$. The ith *moment* of an image is defined [11] as

$$m_i = \left(\frac{1}{X \times Y}\right) \sum_{x=1}^{X} \sum_{y=1}^{Y} I^i(x, y) \tag{1}$$

A transform is called *moment-preserving* if the transformed image, I', still has the same moments as I. In the context of binary segmentation, to divided I into two

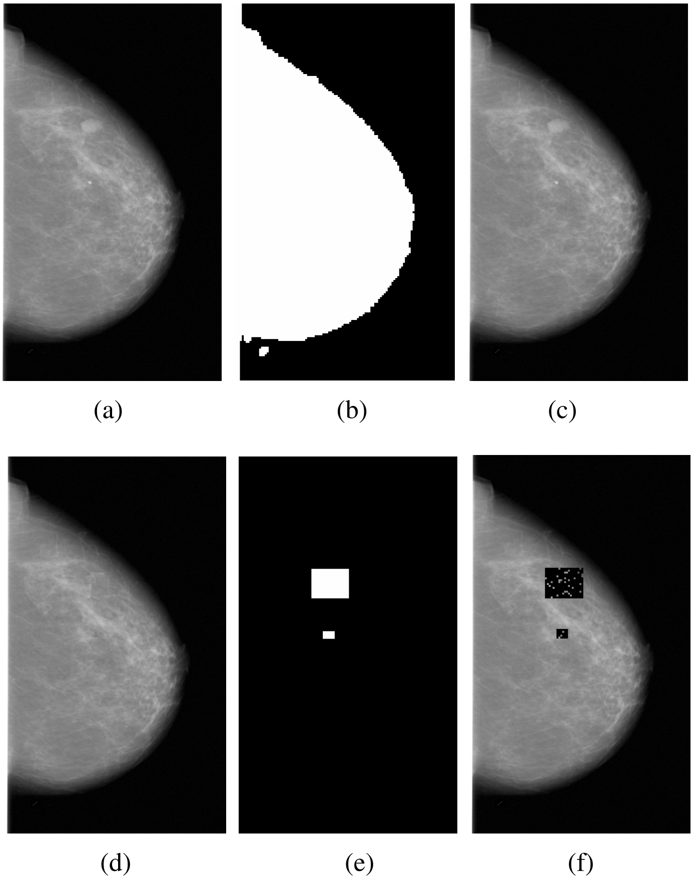

Fig. 1. Mammogram authentication. a) The original mammogram, b) Segmented background and ROIs, c) the watermarked mammogram, d) tampered mammogram, e) the black spots indicate the tampered areas, f) authentication result with the noisy areas indicating tampering.

classes of p_0 and p_1 pixels with gray scale z_0 and z_1, respectively, we can find a threshold t by first solving Eq. (2), as formulated below

$$\begin{cases} p_0 z_0^0 + p_1 z_1^0 = m_0 \\ p_0 z_0^1 + p_1 z_1^1 = m_1 \\ p_0 z_0^2 + p_1 z_1^2 = m_2 \\ p_0 z_0^3 + p_1 z_1^3 = m_3 \end{cases} . \tag{2}$$

Once z_0, z_1, p_0 and $p_1$are obtained, setting the threshold t to a value between the gray scales of p_0 th and (p_0 +1)th pixels will yield segmentation result I' that preserves the first four moments (i.e., m_0 to m_3) of I [11]. From the above description, we know that to make sure the algorithm uses the same segmentation result in both watermarking and authentication processes, when given the original image I_o and watermarked image I_w, respectively, as the input image I, the algorithm should yield the same values for p_0 and p_1. Because the significant gap between the background and ROI in both I_o and I_w, our experiments have proved the feasibility of the use of the moment-preserving thresholding method. The reader is referred to [11] for more details about moment-preserving thresholding.

After moment-preserving thresholding, some pixels with low intensity in the ROI may be classified as background pixels. Moreover, the smoother intensity transition across the boundary separating the background and the ROI may also cause misclassification. To compensate for these two types of misclassifications, a morphological operation of 'dilation' with a disk of radius equal to 5 pixels is applied to the segmented ROI so as to allow the ROI to grow and the background area to shrink.

2.2 Watermarking

The steps as described in the next three subsections have to be taken in order to watermark mammograms.

2.2.1 Establishing Non-deterministic Dependence

In order to authenticate the ROI pixels without distorting them, we involve the gray levels of the ROI pixels in the creation of the watermark (authentication code). As discussed in [5,8], non-deterministic block-wise dependence is crucial in thwarting various types of attacks, such as the cut-and-paste attack, the Holliman-Memon counterfeiting attack, the birthday attack and the transplantation attack. Once the segmentation is completed, the ROI pixels are group into overlapping blocks of multiple pixels. By allowing overlapping, each pixel of the ROI appears in multiple ROI blocks, which in turn gets involved in the watermarking process of multiple background pixels, and as a result, get authenticated by multiple background pixels. If the ROI area is greater than the background, we make the number of ROI blocks the same as the number of background pixels and uniquely associate each background pixel with a ROI block in a key-controlled manner. If the ROI area is smaller than the background, only the same number of background pixels as the ROI blocks are selected as *embeddable*, each uniquely associated with a ROI block. The selection of embeddable background pixels and the association between those background pixels and ROI blocks are also key-controlled. The size of the ROI blocks, as discussed in [8] partially determines the resolution of tamper localization and security, and is to be determined by the user according to the application needs.

2.2.2 Watermark Creation

Suppose we want to embed b bits of watermark (authentication code) in each background pixel. To create the b-bit content-dependent watermark, we first convert the gray levels of the pixels of the corresponding ROI block into binary form, rearrange the bits of each binary gray level under the control of a secret key and perform a bit-wise Exclusive OR operation on those rearranged binary values. Finally, b bits from the result of the Exclusive OR operation are chosen, under the control of the secret key, as the watermark.

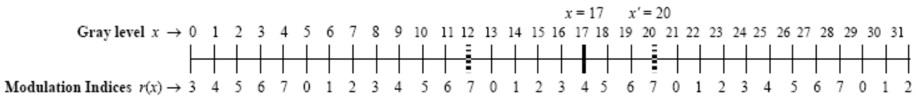
x = 17
x' = 20
Gray level x → 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Modulation Indices r(x) → 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2

Fig. 2. RIM-based watermarking with b equal to 3

2.2.3 Watermark Embedding

Distortion resulted from data hiding is always a major concern [4]. We employ the RIM embedding method reported in [10] to embed the *b*-bit watermark into a background pixel. The idea of embedding the *b*-bit watermark *w* in a pixel, with its gray level equal to *x*, is to assign a modulation index $r(x)$, $r(x) \in \{0,1, 2, ... , 2^b\text{-}1\})$, to each gray level *x* of the image, with gray level 0 corresponding to a key-generated random modulation index $r(0)$, as shown in Figure 2 with $r(0) = 3$. Since the range of $r(x)$ is smaller than the range of the gray levels, the modulation indexes can be *repeatedly* used to index the gray levels. To embed the *b*-bit watermark, the gray level, *x*, of the background pixel is modulated so that the new value *x'* lands on an index $r(x')$ equal to the value of the *b*-bit watermark *w*. Because the indices repeat, watermarking can be achieved by modulating the pixel in question upward or downward, depending on which way results in lower distortion (see Figure 2). The relationship between the gray level *x* of each pixel and its corresponding random modulation index $r(x)$ can be formulated as

$$r(x) = [r(0) + x] \bmod 2^b \tag{3}$$

where “mod” is the modulo operation. For example, suppose $x = 17$ and the 3-bit watermark $w = 7$. We could hide *w* in *x* by changing *x* to either 12 or 20 because $r(12) = r(20) = w = 7$. However, we can see that changing *x* to 20 incurs less distortion, therefore, the algorithm will choose to change *x* to $x' = 20$.

Note that, in order not to provide security gaps to attackers, the inherent characteristics of medical images of various modalities have to be taken into account when watermarking. We observed that, on average, 80% of the pixels in the background areas of mammograms have a gray level of 0. So it is quite easy for an attacker to guess the embedded watermark without knowing the secret key. For example, if the gray level of a background pixel of a *watermarked* mammogram equals 4, the probability that the embedded information equals 4 is 0.8. This is an apparent security gap to be closed. To circumvent this problem, for each zero-valued pixel, we modify its gray level by assigning it a random number in the range of $[0, 2^{b\text{-}1}\text{-}1]$ generated with the secret key. This pre-processing, if necessary, should be carried out before the watermarking process.

2.3 Authentication

To authenticate a watermarked image, the scheme performs the same operations as described in Section 2.1, 2.2.1 and 2.2.2 to calculate the “*original*” *b*-bit watermark *w* for each background pixel. To extract the *embedded* watermark *w'* from each watermarked background pixel x', the scheme simply establishes the correspondence between the elements of the gray level range, with gray level 0 corresponding to a key-generated random modulation index $r(0)$, as described in Section 2.2.3 and shown in Figure 2, and takes the index $r(x')$ corresponding to the gay level of the watermarked pixel x' as the *embedded* watermark w' (i.e., $w' = r(x')$). If $w = w'$, background pixel and the corresponding ROI block are regarded as authentic, Otherwise, they are regarded as manipulated. A bi-level authentication map, with value 255 (0) indicating the authenticity (inauthenticity) of the ROI block could be produced to show the authentication result.

3 Experiments

We have applied the proposed algorithm to various mammograms. Figure 1 demonstrates the process. The size of the images is 460 × 792 pixels. Figure 1(a) to (c) are the original image, segmented ROIs and background, and the watermarked image (with b =3 and the size of the ROI blocks equal to 4 × 4 pixels), respectively. The distortion to the background region is so insignificant that we cannot see it (PSNR = 39.65dB). Figure 1(d) shows the attacked watermarked image, with a small bright spot and a larger mass removed. The two white rectangles in Figure 1(e) show the actually locations where the tampering has taken place. By superposing the authentication map on the attacked image, we can locate the manipulations, as shown in Figure 1(f). For comparison purpose, we also use LSB embedding, in which the b-bit watermark is embedded by directly replacing the b least significant bits of the background pixel with the watermark, in place of the RIM embedding method described in Section 2.2.3. The embedding distortion added to the background area in terms of PSNR is 36.19dB, which is worse then RIM embedding. As discussed in [9], apart from greater distortion, another drawback of LSB embedding is that the very location where the watermark is hidden is known.

4 Conclusions

In this work, we have proposed a novel scheme for authenticating medical images using RIM watermarking technique, which is capable of protecting the region of interest (ROI) without distorting it. The main features of the scheme are:

- By involving the ROI in the creation of a content-dependent watermark and carrying out the embedding in the background area only, the scheme can not only embed higher payload to strengthen security and to increase resolution of tamper localization, but also prevent adding any distortion to the ROI.
- It can resist existing attacks such as the Holliman-Memon counterfeiting attack, the birthday attack and the transplantation attack due to the merit of non-deterministic dependence.
- The balance between security, tamper localization, and embedding distortion can be adjusted by varying the size of the ROI block and the number of watermarkable bits according to the needs of the applications.

References

[1] Bao, F., Deng, R.H., Ooi, B.C., Yang, Y.: Tailored Reversible Watermarking Schemes for Authentication of Electronic Clinical Atlas. IEEE Transactions on Information Technology in Biomedicine 9(4), 554–563 (2005)

[2] Coatrieux, G., Maitre, H., Sankur, B.: Strict Integrity Control of Biomedical Images. In: Proc. Security and Watermarking of Multimedia Contents III, SPIE, vol. 4314, pp. 229–240 (2001)

[3] Guo, X., Zhuang, T.G.: A Region-Based Lossless Watermarking Scheme for Enhancing Security of Medical Data. Journal of Digital Imaging (July 10, 2007), doi:10.1007/s10278-007-9043-6

[4] Ker, A.D.: Locally Square Distortion and Batch Steganographic Capacity. International Journal of Digital Crime and Forensics 1(1), 29–44 (2009)

[5] Kim, H.Y., Pamboukian, S.V.G., Barreto, S.S.L.M.: Authentication Watermarking for Binary Images. In: Li, C.-T. (ed.) Multimedia Forensics and Security, IGI Global (2008)

[6] Kong, X., Feng, R.: Watermarking Medical Signals for Telemedicine. IEEE Transactions on Information Technology in Biomedicine 5(3), 195–201 (2001)

[7] Li, C.-T.: Reversible Watermarking Scheme with Image-independent Embedding Capacity. IEE Proceedings - Vision, Image, and Signal Processing 152(6), 779–786 (2005)

[8] Li, C.-T., Yuan, Y.: Digital Watermarking Scheme Exploiting Non-deterministic Dependence for Image Authentication. Optical Engineering 45(12), 127001-1–127001-6 (2006)

[9] Li, C.-T., Li, Y.: Protection of Digital Mammograms on PACSs Using Data Hiding Techniques. International Journal of Digital Crime and Forensics 1(1), 60–75 (2009)

[10] Thodi, D.M., Rodriguez, J.J.: Expansion Embedding Techniques for Reversible Watermarking. IEEE Transactions on Image Processing 16(3), 721–730 (2007)

[11] Tsai, W.H.: Moment-Preserving Thresholding: a New Approach. Computer Vision, Graphics, and Image Processing 29(3), 377–393 (1985)

[12] Zhou, X.Q., Huang, H.K., Lou, S.L.: Authenticity and integrity of digital mammography images. IEEE Transactions on Medical Imaging 20, 784–791 (2001)