ROI-Based Watermarking Technique for Medical Images via the Use of Alpha Channel

Kishore Sebastian¹, Jenny Maria Johny²

¹Assistant Professor, ²PG Scholar, Department of Computer Science and Engineering, St. Josephs College of Engineering and Technology Pala, Kottayam

Abstract— In this paper, we present a medical image integrity verification system to detect and identify the local image alterations and the nature of a global processing an image may have undergone. The proposed integrity analysis process, is based on the region of interest watermarking with signatures and the use of alpha channel. A new watermarking method via the use of alpha channel is proposed. Also integrity is analyzed at three distinct levels, which are used for the detection of modification, modified location and forensic analysis respectively. Here we first show how the region of interest is automatically extracted using the Region Growing Algorithm. By integrity analysis we can identify whether modifications have occurred and finally recover the image to attain the original image by using the secret sharing scheme. During the process of image recovery, the proposed method can remove the visible watermark from the watermarked image using the data embedded in the alpha channel plane.

Index Terms — Alpha channel, Integrity control, Medical imaging, Region Growing Algorithm, Secret sharing.

I. INTRODUCTION

Advances in information and communication technologies provide new means to access, share, and manipulate the medical images. But if the daily medical practice takes advantage of such an evolution, this facility to handle images also compromises their security. Hence, medical images have to be reliable or trustworthy at all times [1]. Image watermarking has become an important research area in data security, confidentiality and image integrity. The technique of visible watermarking embeds certain visible information (example: company logo) into a host media to declare the copyright or the ownership[2]. Many watermarking techniques were proposed for medical images. Medical image watermarking has various applications in the fields of healthcare, teleradiology, telemedicine, forensic, etc. The medical images, unlike other images, require extreme care when embedding additional data within them such that the additional information must not affect the image quality and readability.

Hence medical images are usually divided into the Region of interest (ROI) and the Region of Non-Interest (RONI). Also the medical records, electronic or not, linked to the medical secrecy and hence, the records must be confidential [3].

In this work, we focus on verifying the damage, it is as much important to be 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. The images can get tampered with the introduction or the removal of lesions [4]. Also, it must be known that some image processing may lead to similar situations. In telemedicine applications, 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 unacceptable information loss and results in a misdiagnosis [5], involving at the same time liabilities of physicians.

II. RELATED WORK

The previous works stated that verifying the integrity of a medical image is an analysis process which involves three distinct levels of integrity [6]:

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 interpret-able by the physician.

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

Different strategies have been proposed for verifying image integrity. They include the use of image digests/signatures/hashes [7] [8], watermarking and blind forensics methods [9]. The image integrity can be verified 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 allow verifying the exact identity of the image under investigation with the original image, and can be used to achieve L1. To localize alterations (L2), one can compute hashes on independent image areas [10] [11]. These approaches can be used to partly achieve L1 and L3. But this can be used only to detect the alterations. Extraction of the ROI can be done both manually as well as automatically. Previously the integrity verification using the integrity control system involved the manual extraction of the ROI. There are several common approaches to the automated extraction of the ROI. Edge-based method uses an edge filter applied to the image classified as edge or non-edge depending on the filter output. The filter output is thresholded at a value of 5: all pixels exceeding 5 are labelled as edge pixels and displayed as black. Connected chains of edge pixels divide the image into regions. Though the extraction seemed easier it wasn’t fully advantageous. The noise distortions could easily create disruptions at the edge hence creating abnormal boundaries thus providing inaccurate edge information [12]. Watermarking is an effective tool for verifying image integrity and authenticity. One common approach consists to insert a specific watermark [13]. The non-detection of this latter informs about image integrity loss. In some cases watermarking is combined with image signatures [14]. A set of signatures is computed from one Region of Interest (ROI) and then watermarked within Regions of Non-Interest (RONI). There are also several other watermarking techniques like self-recovery image watermarking which can be used to partly restore tampered parts of the image [15][16]. But it can only roughly recover the image, i.e. not the image details. For L3 analysis the difference between the restored image and the original image can be computed. But 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. To gain in performance they sometimes spread the watermark over the whole image with the risk to bias the image interpretation reducing the bit-length.

The system we propose is based on the watermark-ing of a ROI signature (computed on independent pixel blocks) into RONI of the image. The major focus is on the recovery of the original image in case the image obtained at the receiver side is found to be altered. The ROI extraction is proposed to be done automatically using the Region Growing Algorithm. Once extracted, a set of three signatures are evaluated and embedded to verify the integrity at three distinct levels, L1, L2, L3.

L3-signature derived from the L2 of pixel blocks is used in order: i) to have an idea about the shape of the local modification, and, ii) to serve as image feature within a classifier based mechanism which purpose is to identify the nature of the global process an image may have undergone.

III. PROPOSED SYSTEM

At the protection stage (fig:1), a signature is extracted from the ROIs of the image, and is embedded into the rest of the image (i.e. RONIs). This watermarked image then goes alpha channel embedding to obtain the alpha channel embedded image. This image is the final protected image which is transferred to the receiver. At the verification stage, the protected image undergoes de-embedding and reconstruction process. The differences between the recomputed and extracted signatures are analyzed in order to achieve the three integrity verification levels: L1, L2 and L3. Each of the Li is verified based on the signature computed for the region of interest [6].
Based on the integrity values obtained and the comparison made, one can detect and identify the alterations that have occurred on the image. Once identified, original image can be recovered by the removal of the alpha channel.

A. Region Growing Algorithm

Automated ROI extraction is done using the RegionGrowingAlgorithm (RGA). The ROI extraction is based on two categories: those pixels of the image have similar greyscale values enter into one category; whereas those with dissimilar values enter into the other category. Region extraction is one of the most critical part in image analysis. In the case of the medical images, the ROI is the region that is prone to diagnosis and the RONI is the insertion zone, onto which the additional data about the ROI is embedded [6] [17].

Algorithm 1:

Input: initial threshold \( t_0 \).

i. Select an initial \( t_0 \): usually mean of the intensity values of the pixels is taken.

ii. Divide the image into 2 regions \( G1 \) and \( G2 \) based on the \( t_0 \) : \( G1 \) consisting of all pixels with gray level values \( >T \), \( G2 \) consisting of pixels with values \( \leq T \).

iii. Compute the average gray level values \( \text{mean1} \) and \( \text{mean2} \) for the pixels in regions \( G1 \) and \( G2 \).

iv. Compute a new threshold value

\[
t = (1/2)(\text{mean1} + \text{mean2})
\]

B. Integrity Analysis

Integrity analysis is done at three levels:

Level 1 (\( L1 \)): Modification Detection:

In order to achieve \( L1 \), any modification of the image if detected will cause an alarm. Cryptographic hash functions is used as the extraction function here. The ROI is considered as a binary message \( M \), onto which the cryptographic hash function \( (f) \) is applied. SHA-1(Secure Hash Algorithm) that yields a 160 bit signature is used.

Level 2 (\( L2 \)): Modification Location:

In order to identify the parts of the image having modified \( (L2) \), one solution consists to compute a set of signatures from different parts of the ROI. A block mapping function is used, i.e. \( k \) pixel blocks \( Bi, i=1...k \), where the signature is computed for each of the blocks. Integrity is consequently controlled at the block level. At the verification stage, one block \( Bi \) is said to be tampered if its recomputed signature differs from its original signature. By concatenating this set of signatures, we obtain the signature \( H2 \). Here, the redundancy provided by the error detection codes has been considered for the computation of signatures, \( i =1..k \). Here the Hamming error detection codes, is used, a message (equivalently a pixel block) of \( 2r-1 \) bit long will be assigned a redundancy or checksum of \( r \) bits with an error rate up-per bounded by \( 1/2r \).

Level 3 (\( L3 \)): Forensics analysis:

\( L3 \)-analysis process depends on the output of \( L1 \) and \( L2 \). From each of the ROI, is computed one \( L3 \)-signature.

i. Based on \( H1 \) and \( H2 \), it is possible to know if the image has been altered \( (L1) \) and which pixel blocks, \( i=1..k \), have been tampered \( (L2) \). Thus \( H3 \) results from the concatenation of each of these signatures. The \( L3 \) analysis helps in identifying the nature of modification that has occurred on the image. [5].

C. Alpha Channel Embedding

The image is appended with an alpha channel plane to form a png image. The appended alpha channel helps in creating a transparency region onto which the recovery data to be used at the recovery stage is embedded into. At the recovery stage, the appended alpha channel is removed once the watermark has been extracted. The removal of the alpha channel brings back the original image without any loss.

The proposed method is based on a secret sharing scheme[20] in which a secret message is transformed into \( n \) shares for keeping by \( n \) participants; and when \( k \) of the \( n \) shares, not necessarily all of them, are collected, the secret message can be recovered lossless. Such a secret sharing scheme is useful for reducing the risk of incidental partial data loss. Hence this scheme is a \((k, n)\)-threshold based secret sharing scheme. Here the image is divided into 2X3 blocks and correspondingly the alpha channel is also segmented into a number of 2X3 blocks resembling the host image. For each of the 2X3 block, a total of 6 shares will be created of which, 2 of the created shares are placed in the same position as that of the image and the other 4 shares are randomized and placed in different positions of the alpha plane.
Before embedding the shares into the alpha plane a value of 238 is added to each of the 6 shares resulting in the new values which fall in the nearly total transparency range of 238 through 254 in the alpha channel plane.

Algorithm 3: (k, n)-threshold secret sharing.

**Input:** a secret s(integer), number of participants n, and a threshold k (k<n). Here k=2

1. Choose randomly a prime number p which is larger than s.
2. Select k - 1 integer values m1,m2, ..., mk-1 within the range of 0 through p - 1.
3. Select n distinct real values x1, x2, ...,xn that represent the shares.
4. Perform the following computation F(xi), called partial shares for i = 1, 2, ..., n
   \[ F(xi) = (s + cx^i_1 + cx^i_2 + ... + cx^{i-1}_k)\mod p \]

Here m1,m2…mk-1 are taken as constant values equal to c. It is necessary to collect at least k shares from the n participants to form k equations to solve these k coefficients in order to recover the secret.

Algorithm 4: Reconstruction.

**Input:** k shares , n participants, prime number p.

1. Solve the k equations in algorithm 3 by Lagrange’s interpolation to obtain d as follows
   \[ s = \frac{1}{k-1} \left[ \left( f(x1)[x2 x3 ... xk-1] + f(x2)[x1 x3 ... xk-1] + ... \right) \mod p \right] \]

   Where,
   \[ T \] is the threshold. If pixel \((x, y)\) satisfies this criterion, then that pixel is added to the region. Otherwise it is neglected.

The self-repairing of the original image is done at the recovery phase. The randomized four shares of each block must be recovered. A value of 238 must be subtracted from the recovered shares to obtain the exact values [20].

IV. EXPERIMENT RESULTS

The system detects if the image has been modified and decides if the modification is global or local based on the output of L2 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 images have been considered (Figure 2):
- Magnetic resonance (MRI) of the head
- X-Ray images
- Ultrasound images

**A. Region Growing**

Region growing is used to automatically select the Regions of interest and the Regions of Non-Interest (Figure 3). It is one of the most efficient algorithms for segmentation. It involves the selection of a seed pixel and the addition of the neighboring pixels into the region based on similarity criterion. Here intensity is used to check similarity. The growing rule of the algorithm is

\[ |I(x, y) - \text{mean}(x, y)| < T \]

Where, \( T \) is the threshold. If pixel \((x, y)\) satisfies this criterion, then that pixel is added to the region. Otherwise it is neglected.

**B. Integrity Analysis**

The analysis involves verifying the integrity of the input image at three distinct levels namely, L1, L2, L3.
Using these integrity values we can find out whether any modification has occurred to the image, the position where the modification has occurred as well as the nature of modification that has occurred on the image (Figure 4).

A. Alpha Channel Embedding

The image is appended with an alpha channel plane to form a png image. The appended alpha channel helps in creating a transparency region onto which the recovery data to be used at the recovery stage is embedded into. At the recovery stage, the appended alpha channel is removed once the watermark has been extracted. The removal of the alpha channel brings back the original image without any loss (Figure 5).

V. Future Work

Though the automated extraction of the regions of interest improves the accuracy of the output image, there are certain constraints with the dimensions of the input image when compared with the manual extraction of the regions. Future work focuses on the automatic extraction of the regions of interest with better accuracy irrespective of the dimensions of the input image.

REFERENCES


