Medical Image Authentication and Self-Correction through an Adaptive Reversible Watermarking Technique

V. Fotopoulos, *Member, IEEE,* M. L. Stavrinou, and A. N. Skodras, *Senior Member, IEEE*

Abstract—**With the advent of Information Technology in the medical world, various radiological modalities produce a variety of digital medical files most often datasets and images. These files as any digital asset should be protected from unwanted modification of their contents, especially as they contain vital medical information. Thus their protection and authentication seems of great importance and this need will rise along with the future standardization of exchange of data between hospitals or between patients and doctors. Watermarking, a technique first introduced for multimedia files, provides a method for authentication and protection and has been recently applied to medical images. In this paper, we propose a novel watermarking technique where the region of non-interest (RONI) of medical Magnetic Resonance Imaging (MRI) images, is used to embed the region of interest (ROI). In this way, any tampering attempt, not only will be detected, but also the image could be self-restored, back to its previous, "original" form by extracting the ROI from the RONI.**

I. Introduction

THE surge of digital radiological modalities in modern hospitals and research institutes around the world, has led to the creation of a vast amount of medical digital assets, like signals and images. Therefore, the need for *authenticity, integrity check*, and *safe transfer* of this type of data will rise. Moreover with the necessity to exchange these medical images between hospitals, the issue of a unified network protocol arises and also issues for their security settings and transfer. Additionally, when a digital medical image is opened for diagnosis, it is important that an automated framework exists to verify the authenticity and integrity of the image itself.

As with the case of multimedia rights protection, a way should be established to verify the integrity, safe transfer and handling of this medical information [1], [2]. Up to now, getting inspiration from the multimedia protection scheme, researchers have implemented various techniques for this purpose in the medical field among which is watermarking [3], [4].

The basic principle of watermarking methods is to add copyright information into the original data, by embedding it into the original image. Then if the image is modified in any sense, it can be detected with the use of the watermark [3]. Initially, the watermark could be simply a unique number, such as the patient's insurance code but as research moves into new paths, a new role has been given to the watermark: to include (apart from hospital digital signatures or copyright information), the electronic patient record, digital documents with diagnosis, blood test profiles or an electrocardiogram signal [5], [6]. By embedding these files into the original image we increase authenticity, confidentiality of patient data and of the accompanying medical documents, availability, and reduce the overall file size of the patients' records [3]. These methods belong to the *data hiding* category. The capacity of the carrier is a very important issue in this field.

Lately, reversible watermarking (RW) schemes have been introduced. According to these techniques the watermark can be fully removed thus leaving the original image intact, ready for diagnosis [7]. A medical image in case of clinical outcome can be divided in two parts, the region of interest (ROI) where the diagnosis focuses and the region of non-interest (RONI), which is the remaining area. In modern RW schemes, the RONI is used as the area where the ROI is inserted [8]. The definition of the ROI space depends on the existence of a clinical finding and its features. Some authors use as RONI the region of background (black area inside an X-ray or MRI image) while others define RONI as any other area of the image that doesn't contain a clinical finding.

In this paper, a reversible RONI watermarking technique has been implemented, for brain MRI scans. In the current work, ROI is defined as a rectangle that contains the whole head shape. In addition to other authors [9], [10] we embedded in the RONI the whole region of interest, in an adaptive reversible way. JPEG2000 compression is used for the ROI. If the RONI gives enough space, the ROI can be compressed in a lossless way, thus making the method fully reversible. If on the other hand, the space is not enough, then the ROI is compressed in the best possible way (in terms of quality) thus making the scheme nearly-reversible (the differences are very small and insignificant). In both cases, if a malicious attack or a simple distortion has happened, the original image can be revealed with good image resolution and the tackling is undoubtedly identified. An area detection algorithm helps measuring the number of RONI pixels where later the information of the ROI is embedded.

The rest of this paper is organized as follows. In section II a brief review of the literature is given while in section III the methodology of embedding, extracting and integrity checking is described. Experimental results are presented in section IV, and conclusions are drawn in section V.

Manuscript received July 4, 2008.

All authors are with the Hellenic Open University, School of Science and Technology, Department of Computer Science, Tsamadou and Aghiou Andreou Str, GR-26222, Patras, Greece (phone: +30 2610 367529; fax: +30 2610367520; e-mails: {vfotop1, m.stavrinou, skodras}@eap.gr).

II. Review of the Literature

Reversible watermarking or data hiding in medical images is gaining great interest lately. Several reversible schemes have been proposed up to now. They can be divided into three major categories based on Feng et al. [11]:

- schemes that apply data compression [12]
- schemes that use difference expansion [13]
- schemes that use histogram bin exchanging [14].

In the first case, in order to recover the original image, the whole image, or part of it, is embedded in the original. Side information is also needed for recovery and has to be embedded as well in the host. In order to increase the embedding capacity, compression is introduced [12], [15], [16], [17]. However this type of reversible watermarking lacks robustness, as any loss of the compressed data may destroy the embedded data [11].

In the second category, Difference Expansion (DE) is applied to embed information. Small values are generated in order to represent the features of the image, taken from an integer transformation, which can be an integer wavelet transformation or another similar function [13]. These schemes are also fragile under attack. Even though, they are pixel-wise, loss of one pixel will not destroy the next pixel and thus the image, however it will destroy the completeness of the location map causing mismatch to all the later pixels.

In the third kind of schemes, histogram bin shifting has been proposed in order to tackle the robustness issue. In that scheme the embedded target is replaced by the histogram of the block [14], [18]. In Vleeschoover et al.'s scheme the original image is segmented into several blocks of pixels and the embedding process follows [18].

In the medical field, the reversible watermarking or data hiding technique is gaining great interest due to the strict importance of medical image security and protection. Some newly developed approaches exist for the Magnetic Resonance Imaging (MRI) images. The MRI modality offers images with extreme clarity of representation of the patients' internal organs and soft tissues and its significance is high in the medical world. Usually with the separation of medical images into regions of interest and regions of non-interest, the usual trend is that the authentication payload is inserted into the RONIs [8], [19]. In this latter work, a mixed reversible scheme was proposed for head MRI images. Two levels of protection schemes were introduced. In the first level, a robust RONI watermarking scheme was applied. Once the ROI is located, a unique identification number (C1) and a digital signature (S1) derived from it are generated and inserted inside the RONI in a robust manner. In the second level, a protection of the image generated in the first level is introduced. A digital signature of the entire image (S2) is computed and inserted with a unique identification number (C2) according to a reversible scheme. This level necessitates the removal of the reversible watermark before the integrity verification. The lossy fragile watermarking is performed with the Least Significant Bit (LSB) scheme. The lossy robust watermarking inserts one bit of the message by modifying in one block B, the relationship between the value of one selected pixel and the mean value of B.

In another image tampered proofing approach, for brain MRI images, belonging to the schemes that apply compression, the host image was divided into blocks of equal size. Then, the recovery and the verification data were created from each block using vector quantization. These latter data were then embedded in the two least significant bits of every block [20].

Moreover, as and in multimedia watermarking the methods can be also divided in those concerning analysis in the frequency or spatial domain. In the first category, the image is transformed into the frequency domain and then some frequency components are being replaced by the watermark. For example Shih and Wu describe a method for robust MRI watermarking that use the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT). In this work, the ROI, which is a rectangle inside the original image with diagnostic importance, is compressed by lossless compression while the rest of the image by lossy compression. Information watermark like a signature image or textual data is embedded inside a rectangle in the RONI that surrounds the rectangular ROI area. The watermarking is performed in the frequency domain [21].

In our approach, we define as ROI the whole head shape image. We consider that, it might be of importance to preserve the whole brain image in as good quality as possible for future diagnostic purposes. And that because, a possible undetected indication that could not be characterized as malicious at the first evaluation could be proven to be otherwise in a future one.

An area detection algorithm automatically detects the edges of the smallest rectangle containing the whole head image and counts the area of ROI and of non-ROI available. From the size of the latter, the kind of compression (lossy or lossless) is decided.

The advantage and novelty of our method is that it is a data hiding technique that can be used to automatically hide the original MRI image in the non ROI area. It can be used to automatically watermark single MRI images or even a sequence of MRI images, stored and used for 3-D reconstruction.

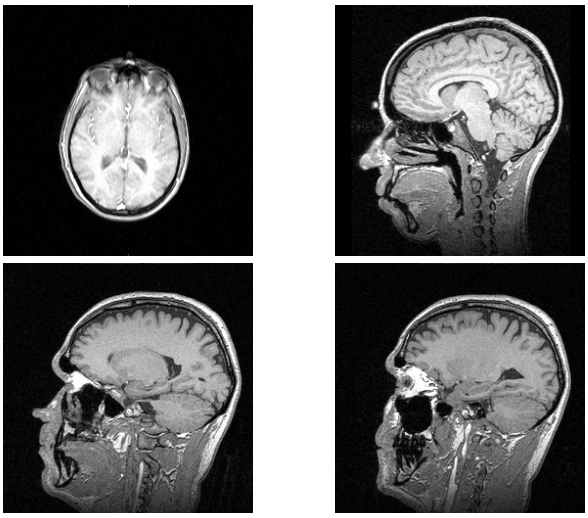
III. Proposed Methodology

A. Embedding

In the proposed method, the MRI is divided in two different regions as stated earlier, the RONI and the ROI. Pixels that belong to the former are suitable carriers for a compressed version of the latter. For this type of segmentation, a simple algorithm scans the image from both sides (left to right and right to left) until it reaches a large intensity value. Large, is defined by a threshold value. For 8-

bit images this threshold should have a value above 15 (because during the hiding phase, the last 4 bits of each pixel that belongs to the RONI will be substituted) but not too large because edges that define the shape of the ROI are smooth. Thus a good value selection would be between 20 and 40.

For each row of the image, the left and the right edges of the ROI (columns) are recorded. For an image of dimensions *MxN*, this gives us two vectors $\boldsymbol{L}$ and $\boldsymbol{R}$ of size M. Similarly, two other vectors $\boldsymbol{T}$ and $\boldsymbol{B}$ of size N are formed, for which the upper and lower edge position is recorded for each column. If we select $l=\min(\boldsymbol{L})$, $r=\max(\boldsymbol{R})$, $t=\min(\boldsymbol{T})$ and $b=\max(\boldsymbol{B})$, then we define a rectangle of which the left upper corner has coordinates (t,l) and the bottom right one is (b,r). This is the rectangle that contains the whole shape. Speaking of ROI, in some of the literature methods, the ROI is selected by experts and may be a small part of the shape (e.g. a shadow inside the brain) or manually by non experts. In the proposed method, a rectangle that contains the whole head shape, is automatically selected for full recovery. Fig. 1 shows some images from the test set used in this work while in Fig. 2, a sample segmentation is shown. It is observed from the segmented image, that there are two thin lines, at the right lower side of the head. Such lines are due to noise phenomena and can be a real problem if they are closer to the image edge, comparing to the real distance of the head from the edge. In that case the ROI is malformed, with a rectangle larger than the one that is really needed. This reduces the slice's hiding capacity by actually reducing the region of non-interest. Such problems may be overcome by a careful selection of the segmentation's threshold or by some kind of morphological processing. In this work, the first way is used.



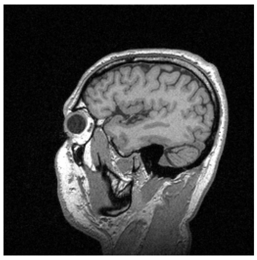


Fig. 1. Typical MRI slice images used for analysis.

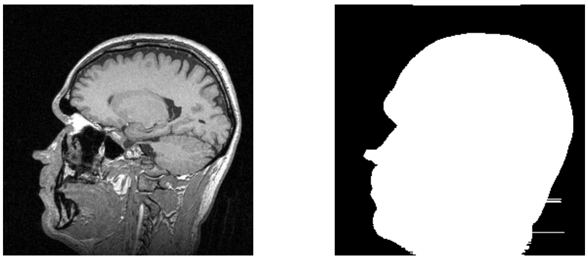


Fig. 2. MRI sagittal slice and its segmented version.

A binary location map of arbitrary shape is formed from the earlier processing. The two regions shown form the RONI and the ROI. Then for each pixel that belongs to the RONI, the intensity is set to zero (the area is cleared). In the next step, the rectangle that contains the patient's head is compressed as a separate image by JPEG2000 tool Jasper [22]. There are two options for this stage; either the capacity of the RONI is such that lossless compression of the rectangle is possible (thus the scheme is fully reversible), or the capacity is not enough. In that case, a desired bit rate is calculated and provided as input to the compression tool, in order to achieve the best possible quality. This bitrate is calculated by (1).

$$\text{Desired Bitrate} = \frac{\text{number of pixels} \in RONI}{2(t-b)(l-r)} \qquad (1)$$

B. Extraction and Integrity Check

During the extraction phase, the same segmentation algorithm is applied on the carrier image. It is certain that the regions produced will be identical to that of the original image, because embedding 4 bits in the cleared RONI area, will produce a maximum intensity of 15, thus a threshold value of 16 will identify all pixels that contain part of the compressed bitstream. Then, the compressed file will be retrieved and decompressed with Jasper. Finally, the saved rectangle area can be compared with the area in the same position of the MRI under investigation in order to perform integrity checking. This comparison could be a single subtraction. If the original ROI was losselessly compressed, then even the slightest change will be revealed. If the compression was lossy, then thresholding with a small threshold value is adequate. Fig. 3 exhibits such a case, where a gray dot is added into the right hemisphere (middle image) and the alteration is fully identified (right image).

TABLE I
EXPERIMENTAL RESULTS FOR THE IMAGES OF FIG.1

Image ID	Number of RONI pixels	Lossless compression filesize	ROI dimensions	Desired bitrate	Lossy compression Filesize	PSNR (dB) between original ROI and lossy compressed version
1	47717	13830	134 x 173 (23182 pixels)	-	-	-
2	27411	32578	215 x 227 (48805 pixels)	0.280	13513	36.099
3	29106	31005	204 x 224 (45696 pixels)	0.318	13806	37.293
4	35371	25827	178 x 212 (37736 pixels)	-	-	-
5	27568	33279	224 x 245 (54880 pixels)	0.251	13694	36.488

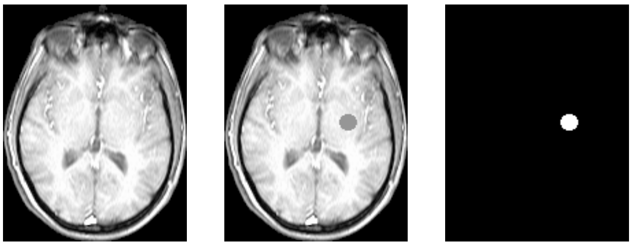


Fig. 3. From left, to right: original ROI, altered version and alteration map.

IV. EXPERIMENTAL RESULTS

Experiments were conducted for a number of MRI slice images. MRI scan was acquired using a Siemens Magnetom Vision 1.5-T scanner (Siemens, Erlangen, Germany) according to a magnetization-prepared rapid gradient-echo sequence (256x256, field of view = 256, time repetition = 9.7 ms, echo time = 4 ms, flip angle = 12^o, thickness = 1 mm). The results for the 5 slices depicted in Fig.1, are given in Table I. The second column, gives the number of pixels available for data hiding while the third column provides the file size of the corresponding lossless bitstream. It is clearly shown that for the fist and fourth image, the capacity of the RONI, is enough in order to losslessly compress and hide the ROI. For the rest of the images, where the area of the RONI does not provide adequate capacity, the desired bitrate is calculated (using the size of ROI that is given in column four) and falls in the range 0.25-0.31 bpp. Column six contains the final filesize produced by Jasper for the lossy compressed versions. These bitrates are good enough, in order to provide an excellent quality compressed ROI, by means of JPEG2000. To justify this claim, the PSNR column shows that the comparison between the original ROI and its J2K compressed version, yields PSNR values in the range of 36-37 dB.

V. CONCLUSIONS

MRI images are used increasingly for pre-operative assessment and planning especially for brain neurosurgical as well as for long term follow–up evaluations. Thus preservation of their integrity and authenticity is of paramount importance for the medical community.

Several reversible watermarking and data hiding schemes have been proposed for that purpose. In our work we present an adaptive reversible watermarking technique on which the embedding capacity and the compression that is followed depend on the number of available pixels of the RONI. In the three usual representations of brain MRI slices, that is sagittal, horizontal and coronal, the number of available RONI pixels increases with the same order. In this work, sagittal and horizontal slices were analyzed, as in those the available RONI pixels are less and depend on the sequential MRI slice number. The method is successfully used for integrity check by using a simple segmentation algorithm, combined with JPEG2000 compression and bit substitution. It can be further expanded in order to watermark reversibly (or nearly reversibly) the whole MRI sequence slices. Furthermore, the bitrate for the ROI compression can be adjusted in order provide some space to hide also some other information like the Electronic Patient’s Record (EPR) etc.

ACKNOWLEDGMENTS

This work was funded by the European Union – European Social Fund (75%), the Greek Government - Ministry of Development - General Secretariat of Research and Technology (25%) and the Private Sector in the frames of the European Competitiveness Programme (Third Community Support Framework - Measure 8.3 programme PENED contract no.03ED832).

Special thanks to the Institute for Advanced Biomedical Technologies (ITAB), Gabriele D’ Annunzio University, Chieti - Pescara, Italy and the Director Prof. Gian-Luca Romani for providing the MRI test set.

REFERENCES

[1] S. Katzenbeisser and F. A. P. Petitcolas, *Information hiding Techniques for Steganography and Digital Watermarking*, Norwood Mass: Artech House, 2000, USA.

[2] V. Fotopoulos and A. Skodras, “Digital image watermarking: An overview”, in *EURASIP Newsletter*, vol. 14, pp. 10-19, Dec 2003.

[3] G. Coatrieux, H. Maitre, B. Sankur, Y. Rolland, and R. Collorec, “Relevance of Watermarking in Medical Imaging”, In Proc. *3rd Conference Information Technology Application in Biomedicine*, Arlington, 2000, pp. 250-255.

[4] G. Coatrieux, L. Lecornu, B. Sankur, and C. Roux, “A review of image watermarking applications in healthcare”, In Proc. IEEE Eng Med Biol Soc., New York, 2006, vol. 1, pp. 4691-4694.

[5] K. A. Navas, S. A. Thampy, and M. Sasikumar, “EPR Hiding in Medical Images for Telemedicine”, in *Proc. of the World Academy of Science, Engineering and Technology*, Rome, 2008, pp. 292-295.

[6] H Münch, U. Engelmann, A. Schröter, and H. Meinzer, “The integration of medical images with the electronic patient record and their web-based distribution”, *Acad. Radiol.*, vol. 11, pp. 661 – 668, June 2004.

[7] J. M. Zain, L. P. Baldwin, and M. Clarke, “Reversible watermarking for authentication of DICOM images”, In Proc. *26th Annual International Conference of the IEEE EMBS*, San Francisco, 2004, pp. 3237-3240.

[8] G. Coatrieux, J. Montagner, H. Huang, and C. Roux, “Mixed Reversible and RONI Watermarking for Medical Image Reliability Protection”, in *Proc. 29th International Conference of the IEEE EMBS*, Lyon, 2007, pp 5653-5656.

[9] J. M. Zain and M. Clarke, “Reversible Region of Non-Interest (RONI) Watermarking for Authentication of DICOM Images, *International Journal of Computer Science and Network Security*, vol. 7, pp.19-28, Sept. 2007.

[10] A. Giakoumaki, S. Pavlopoulos, and D. Koutsouris, “Secure and efficient health data management through multiple watermarking on medical images”*, Med. Bio. Eng. Comput.*, vol. 44, pp. 619-631, Oct. 2006.

[11] J. B. Feng, I. - C. Lin, C. S. Tsai, and Y. P. Chu, “Reversible watermarking: Current Status and Key Issues”, *International Journal of Network Security*, vol. 2, pp. 161-171, May 2006.

[12] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, “Reversible data hiding”, in Proc. International Conference on Image Processing, New York, 2002, pp. 157-160.

[13] J. Tian, “Reversible data embedding using a difference expansion”, *IEEE Transactions on Circuit Systems and Video Technology*, vol. 13, pp. 890-896, Aug. 2003.

[14] C. D. Vleeschouver, J. E. Delaigle, and B. Macq, “Circular interpretation of histogram for reversible watermarking”, in Proc. *IEEE 4th Workshop on Multimedia Signal Processing*, Cannes, 2001 pp. 345-350.

[15] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, “Lossless generalized-LSB data embedding”, IEEE Transactions on Image Processing, vol. 14, pp. 253-266, Feb. 2005.

[16] G. J. Fridrich, M. Goljan, and R. Du, “Invertible authentication”, in *Proc Security and Watermarking of Multimedia content*, San Jose, 2002, pp. 197-208.

[17] G. J. Fridrich, M. Goljan, and R. Du, “Lossless data embedding,-new paradigm in digital watermarking”, *EURASIP Journal on Applied Signal Processing*, vol 2, 185–196, 2002.

[18] C. D. Vleeschoover, J. F. Delaigle, and B. Macq, “Circular interpretation of bijective transformations in lossless watermarking for media asset management”, *IEEE Transactions on Multimedia*, vol. 5, pp. 97-105, Mar. 2003.

[19] G. Coatrieux, H. Maitre and B. Sankur, “Strict Integrity Control of Biomedical Images”, in *Proc. Security and Watermarking of Multimedia Contents III*, San Jose, 2003, pp.689-698.

[20] J. C. Chuang and C. C. Chang, “Detection and Restoration of a Tampered Medical Image”, in. Medical Imaging and Augmented Reality, Berlin Heidenberg: Springer, 2004, pp. 78-85.

[21] F. Y. Shih and Y.-T. Wu, “Robust watermarking and compression for medical images based on genetic algorithms”, *Information Sciences* 175, pp. 200–216, Feb. 2005.

[22] The JasPer Project Home Page, Michael Adams, retrieved 3/7/2008, available: http://www.ece.uvic.ca/~mdadams/jasper