Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090355

Information Hiding for Medical Privacy Protection based on GHM and Canny Edge Detection

Ren Shuai[1,*], *Zhao Xiangmo*[1] *and Zhang Tao*[2]

[1] School of Information Engineering, Chang'an University, Xi'an 710064, China
[2] School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China

Received: 25 Aug. 2014, Revised: 25 Nov. 2014, Accepted: 26 Nov. 2014
Published online: 1 May 2015

Abstract: In order to protect the personal and diagnosis information of patient, GHM and Canny edge detection are employed in this algorithm to process the original CT digital image. The original image will be firstly processed GHM for a global analysis. According to different importance degrees, CT images can be distributed into LL_2, LH_2, HL_2 and HH_2 by GHM multi-wavelet. After that, the distributed sub-images will be processed by Canny edge detection. For example, the edges of LL_2 (with the highest important degree) and HH_2 (with the lowest important degree) are detected by strong Canny method. Then many thin edges can be obtained, which are useless for medical diagnosis, privacy and diagnosis information can be hid by deletion and comparison processing of invalid edges. The experimental results indicate that the robustness and capacity of watermarked CT images can be improved after this algorithm.

Keywords: Diagnosis Gray or Binary Image, Privacy Protection, GHM Multi-Wavelet, Canny Edge Detection.

1 Introduction

Nowadays, personal privacy security has become one of the issues concerned most in digital era. Long after sectors like banking and finance, recently more and more digital auxiliary diagnostic equipment are brought into daily practical operation in medicine field. Individual medical records, carrying ever more sensitive personal information, are already being gathered and stored by the tens of thousands in databanks maintained by hospital networks, health maintenance organizations, and drug companies [1]. As a result, information storage methods of the patients have changed from papery-based into electronic-based.

There are several common image formats in medical field, such as CT, MRI, X-ray, and so on. Some of those electronic-based images are in analog formats and others are in digital formats. Because of the convenience of the digital images, the old analog images are converted into digital ones. One of the problems brought by the popular of digital medical images is the privacy information security, which will make the relationship between both sides of the health care worse. And reasonable method can solve the problem. It is well known that digital products can easily be illegally edited, modified, copied and disseminated. These properties bring great convenience, but at the same time it will inevitably produce the issue of privacy disclosure of patients information. Considering about the medical characteristics, some scholars use the professional mathematical model to analyze the natural system images of human body [2]. This model can simplify the representation of some special medical images, like MRI brain image, so it can find some redundancy to be the storage space. Refer to it, getting the redundancy of medical images will be the first concern of our algorithm, but the algorithm proposed in our research can be used to handle many kinds of medical images, rather than limited to MRI images. Many other scholars achieve the goals of medical privacy protection and authentication by using watermarking [3]. Most of these schemes embed robust watermark into the digital medical images based on frequency-domain methods, such as DCT, DWT, DFT, and so on. And almost all of them select the appropriate coefficients for watermark embedding [4].

Most of the digital medical images are gray scale images, and some of them are the binary images. Based on our previous studies, we know very well about the properties of gray scale and binary images [5]. From [6] we find that GHM multi-wavelet transform has much

* Corresponding author e-mail: maxwellren@qq.com

more advantages than other existing frequency-domain methods in carrier pre-processing. Digital images can be decomposed into sub-images with different frequency characteristics by GHM. And in this paper, Canny edge detection will be applied as an auxiliary tool for GHM to process the medical images, which can find good position to embed secret data [7]. Our algorithm is applied on CT images. So in this paper, CT images will decomposed according to the degree of importance within itself by GHM multi-wavelet transformation. After that, Canny edge detection can be used to extract information of decomposed images.

This paper is organized as follows. Section 2 formally discusses different components of the proposed algorithm—GHM and Canny edge detection methods. Section 3 presents the embedding and the extracting parts of the proposed reversible multiple information hiding algorithm. And section 4 analyzes the security of this algorithm. Experimental results and conclusion are respectively presented in section 5 and section 6.

2 Methodologies

We take the inherent characteristics of medical images into consideration, and employ GHM multi-wavelet and Canny edge detection for data hiding. The main design idea is shown below: CT images are decomposed into sub-images with different energy importance by GHM multi-wavelet transform. Sub-image with lowest energy will be detected by strong Canny edge detection method with small threshold. Then we can get many invalid edges which are useless to medical diagnosis. At last, data hiding can be achieved by invalid edges deletion or xor operation of lowest energy components.

2.1 GHM multi-wavelet transform

GHM multi-wavelet transform is the earliest, most widely used multi-wavelet. It is of some distinctive characteristics simultaneity, such as compactly supported (have better energy compaction), second order approximation ability, integral translation of scaling function orthogonal to each other, high-order vanishing moment (can detect sudden change point more accurately) and symmetry. The energy distribution of GHM transform [8] provides a flexible information hiding strategy. Figure 1 shows GHM transform to Lena image. Energy ratio of four first-order GHM multi-wavelet Transform sub-images is approximately as Table 1 [9].

The energy distribution of 4 components (LL_2, LH_2, HL_2 and HH_2) of lowest resolution sub-images of GHM multi-wavelet transform approximately equals to 4.5:2.2:2.2:1.1. According to the energy distribution characteristics, the lowest energy component HH_2 of the second highest energy area will be employed to be the embedding region. At the same time, the highest energy component LL_2 of the second highest energy area will be the storage unit for some overhead information, which is used to compare the original and modified data. Based on this designing scheme, the capacity of the processed images will be enlarged and with a little difference from the original image structure. The personal and diagnosis information of patient can be hidden in the embedding space and the modified diagnosis map will not show the real final diagnosis. Thus the medical information of patients can be protected.

(a) Original image-Lena (b) Transformed Lena (c) Distribution map

Fig. 1: First order GHM multi-wavelet transform.

Table 1: Energy distribution of first order GHM multi-wavelet transform.

Percentage of LL_1 in the total energy	Energy percentage of secondary sub-images in energy of LL_1 (%)			
	LL_2	LH_2	HL_2	HH_2
97.31	44.76	21.80	22.24	11.20

2.2 Canny edge detection

Being proposed by John Canny in 1986 [10], Canny operator adopts Gaussian filtering to smooth the images, and computes the gradient magnitude by limited difference value in neighborhood, hence the location to edges of Canny is accurate[11]. The parameters of upper and lower threshold of Gaussian filtering are not determined by the features of image edge, but should be given by technician. Those parameters are generally represented by SIGMA, which is also the standard deviation of Gaussian filtering. In the simulation experiment environment (Matlab7.0.0.19920) for this paper, the default value for tools of Canny edge detection is one[12].

3 Proposed hiding and extracting algorithms

3.1 Information hiding and elimination

Information hiding and elimination process based on GHM multi-wavelet and Canny edge detection can be

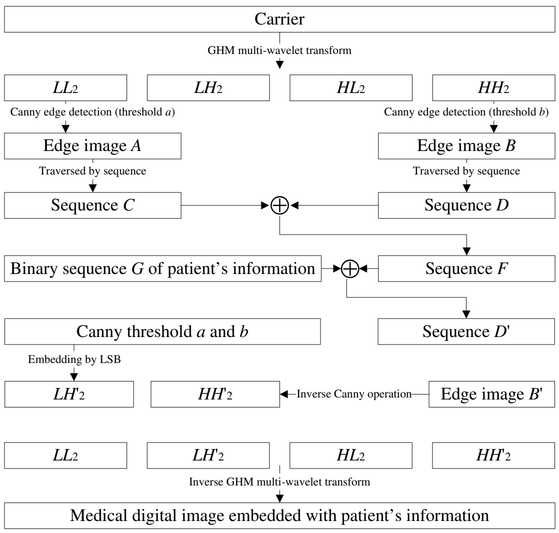


Fig. 2: Flowchart based on GHM and Canny edge detection.

summarized in the following steps, and flowchart is shown in Figure 2:

1) Patient medical images are decomposed into sub-images (LL_1, LH_1, HL_1 and HH_1) by GHM multi-wavelet transform, and four components (LL_2, LH_2, HL_2 and HH_2) of LL_1 are obtained.

2) LL_2 and HH_2 are detected by Canny edge detection method, and a and b are supposed to be the threshold values. The obtained edge images (binary format) are A and B.

3) Image A and B will be traversed respectively in accordance with the sequence from left to right and from top to bottom. The traverse results are two binary sequences, which are represented as C and D.

4) C and D are manipulated by XOR operation, as shown in Formula 1. The result binary sequence is expressed as F

$$F = C \oplus D \tag{1}$$

5) Patient's information is conversed into binary sequence, denoted as G.

6) G is assigned to F according to Formula 2 and a new sequence denoted as D' will be obtained. According to D', new edge image B' can be get by modifying to edge image B of HH_2.

$$D' = C \oplus F \tag{2}$$

7) B' will be put into inverse Canny operation to get a modified sub-image, denoted as HH_2'.

8) Edge threshold values a and b are conversed into binary data, and then embedded into LH_2 by LSB. The changed LH_2 is represented as LH_2'.

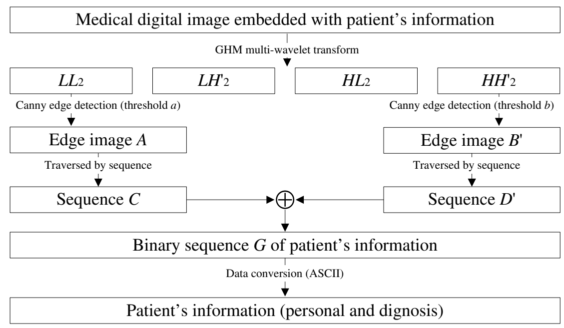


Fig. 3: The extraction of patient's information.

9) Changed medical images with patient's personal and diagnosis information are obtained by GHM inverse transform to LL_2, LH_2', HL_2 and HH_2'.

3.2 Information extracting

Patient's name (username) and threshold a and b (password) of Canny are the key for diagnosis information extracting. The extraction process can be divided into 5 steps, and flowchart is shown in Figure 3.

1) Four components (LL_2, LH_2, HL_2 and HH_2) of LL_1 are extracted from modified patients images by GHM.

2) LL_2 and HH_2 are detected by Canny edge detection method with threshold values a and b. The edge images (binary format) A and B' are obtained.

3) Image A and B' will be traversed respectively in accordance with the sequence from left to right and from top to bottom. The traverse results are two binary sequences, which are represented as C and D'.

4) C and D' are manipulated by XOR operation, as shown in Formula 3. The result binary sequence is expressed as G

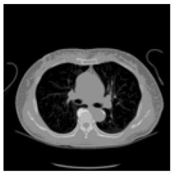
$$G = C \oplus D' \tag{3}$$

5) As a binary sequence, G is conversed into Patient's information (According to ASCII).

4 Security analyses

Sub-images with considerable different amount of energy can be decomposed after GHM. Based on the corresponding relationship between energy and visibility, edges of the associated areas (areas with highest and lowest energy) are modified after Canny edge detection. And the modification is just the information embedding.

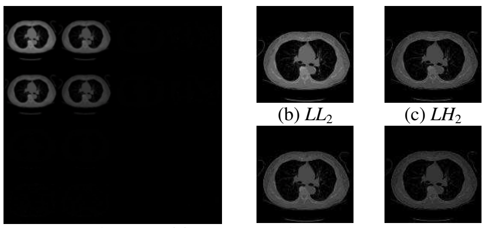
There are some changes between the edge of the original and modified image, thus not only the image shown in computer but also the CT film before the light are all unreal diagnosis image. Only by extracting



Name: Tom
Gender: Male
Date of Birth: yyyy.mm.dd
Diagnosis result: cerebral concussion
Visiting date: yyyy.mm.dd

(a) Brain (original) (b) Patient information

Fig. 4: Original image and patient information.



(a) GHM decomposition (b) LL_2 (c) LH_2 (d) HL_2 (e) HH_2

Fig. 5: Decompostion of patient diagnosis image by GHM.

(a) Edge detection image of LL_2 (b) Edge detection image of HH_2

Fig. 6: Canny edge detection.

according to information hiding inverse algorithm, the original image can be restored and diagnosis information can be obtained, hence the privacy can be preserved.

5 Simulation experiment

In Matlab7.0.0.19920, a 512×512 CT image is considered as the carrier, and the format has been converted from DCM into JPG (Figure 4a). The hidden information is some literal data (Figure 4b).

According to Section 3, medical image is decomposed into sub-images by GHM, shown in Figure 5(a). Than four components (LL_2, LH_2, HL_2 and HH_2) of LL_1 will be mainly studied, which are respectively shown in Figure 5(b), 5(c), 5(d) and 5(e).

LL_2 and HH_2 are respectively processed by Canny edge detection (Matlab7.0.0.19920). The default value of threshold (a and b) is 1, shown in Figure 6.

Images in 6(a) and 6(b) will be traversed respectively in accordance with the sequence from left to right and from top to bottom. The traverse results are two binary

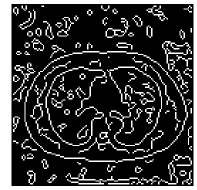


Fig. 7: Modified edge image.

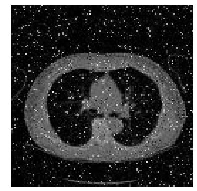


Fig. 8: New HH_2.

Fig. 9: New LH_2 with edge detection threshold values (a, b).

sequences, which are represented as C and D. Than C and D are manipulated by XOR operation, as shown in Formula 1. The result binary sequence is expressed as F. Patient's information is conversed into binary sequence (according to ASCII), which is denoted as G. G is assigned to F according to Formula 2, the obtained sequence is D' (the data is too large to list here).

The edge image B of HH_2 is modified according to D', and the resulted image is B', which is shown in Figure 7. Referring to B', the new HH_2 can be acquired, which is denoted as HH_2' and shown in Figure 8.

Threshold a and b of Canny edge detection are conversed into binary format. Based on LSB, binary code of a and b can be embedded into LH_2, and the covered LH_2 is denoted as LH_2', shown in Figure 9.

LL_2, LH_2', HL_2 and HH_2' are operated by GHM inverse transform, and the medical image covered with privacy and diagnosis information of patient can be obtained, shown in Figure 10.

6 Conclusions

In this paper, a privacy protection algorithm based on GHM and Canny edge detection is proposed. GHM multi-wavelet is applied to analyze the overall energy distribution of medical image and obtain sub-images (to

Fig. 10: Medical image with privacy and diagnosis information of patient.

be used as the embedding region) with energy ratio 4.5:2.2:2.2:1.1. Canny edge detection is employed to get the specific embedding positions (modified data). Actually, the component with lowest energy is modified by XOR operation with highest component. And then the aim of medical image security and patient privacy protection is achieved. In the future, we will find other methods similar to Canny edge detection which can cooperate with GHM to realize the same goal.

Acknowledgments

Our research was funded by two projects, and the names and numbers of these Projects are as follows: 1. Basic Research Project of Shaanxi Province Natural Science Foundation (Grant No. 2013JM8018). 2. The National Natural Science Foundation of China (Grant No. 61402052).

References

[1] R. Cushman, IEEE Technol. Soc. Mag. **15**, 32 (1996).
[2] N. Mahdi Majidi, J. Reza, G. Amirhossein, The 8th International Conference for Internet Technology and Secured Transactions. 582 (2013).
[3] H. Hsiang-Cheh, F. Wai-Chi, C. Shin-Chang, IEEE/NIH Life Science Systems and Applications Workshop. 190 (2009).
[4] R. Asmaa H., H. Abd-Fatth, The 7th International Conference on Informatics and Systems. 28 (2010).
[5] IC Chang, YC Hu, WL Chen, CC Lo, High capacity reversible data hiding scheme based on residual histogram shifting for block truncation coding, Signal Processing, Elsevier **108**, 376-388 (2015).
[6] WH Hsu, Transitioning to a communication-oriented pedagogy: Taiwanese university freshmen's views on class participation, **49**, 61-72 (2015)
[7] X. Qian, S. Varadarajan, C. Chakrabarti, J. Karam, IEEE Trans. Image Process. **23**, 2944 (2014).
[8] Z. Tao, M. De-jun, R. Shuai, Int. J. Digit. Content Technol. Appl. **5**, 210 (2011).
[9] H. Zhuo-jun, M. Zheng-ming, J. Image Graphics. **12**, 1198 (2001).
[10] J. Canny, IEEE Trans Pattern Anal. Machine Intell. **8**, 679 (1986).
[11] C. Jie, W. Zhen-hua, D. Li-hua, Opto-Electronic Eng. **35**, 79 (2008).
[12] O. Laligant, F. Truchetet, IEEE Sig. Process. Lett. **14**, 185 (2007).

Ren Shuai obtained his PhD from Northwestern Polytechnical University of China in 2009. He is an associate professor in School of Information Engineering in Chang'an University. He has been engaged in Information hiding and Network security for 8 years. He published 23 scientific research articles in international publications and 3 are cited by SCI, 8 are cited by EI. He has carried out 5 tasks to study a plan in all, won patent 2. During the last yearhe has written or co-edited for 5 textbooks.

Zhao Xiangmo was born on August, 1966 in Chongqing, China. He is professor and Doctoral tutor in Changan University. His research is focus on intelligent detection and fusion technology of traffic information, theory and application of intelligent transportation system. He has published 97 papers, including 66 SCI, EI and ISTP articles.

Zhang Tao obtained her PhD from Northwestern Polytechnical University of China in 2012. She is a lecture in School of Electronic and Control Engineering in Changan University. She has been engaged in Information hiding and Network security for 8 years. She published 20 scientific research articles in international publications and 1 are cited by SCI, 6 are cited by EI. She has carried out 4 tasks to study a plan in all, won patent 1.