

Mammography Compression in Narrow Band Transmission Environment

Zhiqiong Wang¹, Chen Zhang², Tianshu Wang², Jinhui Zhang¹, Yang Cui¹, He Ma¹ and Yan Kang^{1,*}

¹Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, China

²Medical IT Division, Neusoft Corporation, Shenyang, China

*Corresponding author

Abstract—Mammograms are frequent detection means in breast disease. But the limited bandwidth impedes the communication transmission of mammograms in PACS. In this paper, we investigated on the compression method of four mammograms in DICOM format in narrow band transmission environment. Firstly, according to the similarity of mammograms at the same view, the difference of two images can be obtained by making subtraction between the two images after reversal and reference images. Then, an improved Sobel method was proposed to achieve the edge detection of mammograms and also to extract the region of interest (ROI). Finally, using the run-length coding which is one of the entropy coding to process the lossless compression coding for the ROI of the difference images. In this paper, the compression method proposed is proved on the actual mammograms, and the better compression ratio was obtained at mammograms of CC views and MLO views.

Keywords—sobel method; mammograms; run-length coding; image compression

I. INTRODUCTION

With the continuous improvement of medical imaging technology, the resolution of images is improved as well as the image data are also increasing. Therefore, the problems of management, storage and transmission brought by a large of image data have gradually become a heated issue. Picture Archiving and Communication System (PACS) is one of important components in digital hospital. The birth of DICOM standard makes the PACS system can be realized and developed. But the communication transmission bandwidth is limited. However, the storage space and transmission bandwidth of the system are limited, and the huge medical image data bring difficult and challenge to the development of PACS. Because of the particularity of medical images, the doctors have high request to the image fidelity. Therefore, it is necessary and important to study effective image compression method.

Because of the high incidence and recurrence rate, breast cancer is the most common type of malignant tumor among the female [2]. According to statistics, fifty million people died from it every year. Mammograms are frequent detection ways in breast disease. We often obtain images of different views from different shooting angle and posture. In clinic diagnosis, the total of four images will be shot, two of CC mammography in left and right sides and two of MLO mammography in left and right side. Figure 1 shows the shooting process of CC views and MLO views. Consequently, this paper will research on a compression method of four pieces of mammograms in the

DICOM file format under the narrow band transmission environment.

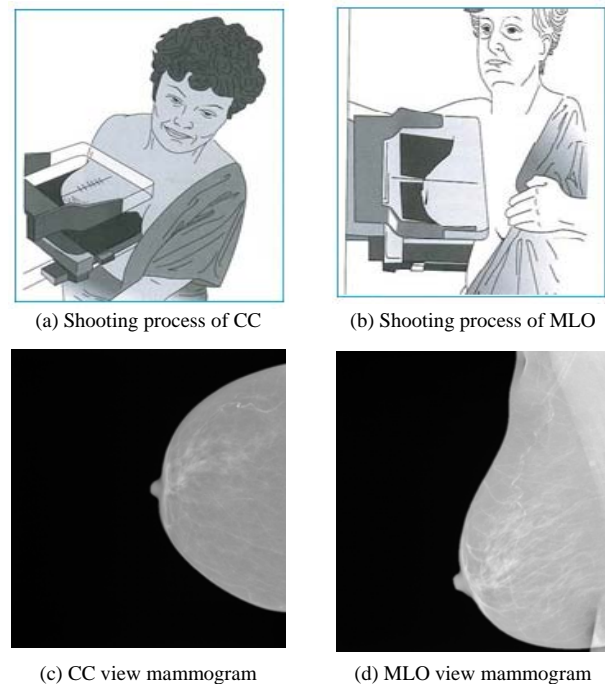


FIGURE 1. SHOOTING PROCESS AND MAMMOGRAMS OF CC AND MLO

II. COMPRESSION METHOD OF MAMMOGRAMS

There is a certain correlation between the four mammograms, so the method proposed in this paper will compress images after processing according to the similarity, so that the overall compression rate can be improved. Before the compression, two reference images are first determined, then the other two are non-reference images. The specific compression process is firstly to obtain the correlation difference of images, that is, to reverse the non-reference images, then making subtraction between the reference images and the reserved non-reference images. Then, extract the region of interest (ROI) in the difference image. Finally, process the lossless compression coding using the run-length coding for the ROI. Breast image compression processing flow chart is shown in Figure 2.

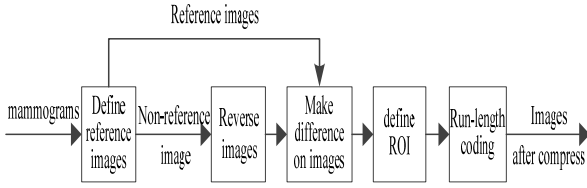


FIGURE II. PROCESSING FLOW CHART OF MAMMOGRAMS COMPRESSION

A. Acquisition of Related Image Difference

Because of the symmetry of the human body tissue, the same view on each side of mammograms is similar, just in different directions. Thus, in this section, four images are divided into two groups, that is, two images of CC mammography for each of the sides and two images of MLO mammography for each of the sides. The process is as follow:

Step1: One side of CC mammography and MLO mammography are used as the transmission of reference image [3];

Step 2: The other side of mammograms, that is, the non-reference images are reversed;

Step 3: To make the subtraction between two images and obtain the difference mammograms.

Figure.3 shows the result of the process at second step. At the third step of the process, we can obtain difference mammography (Figure 4).

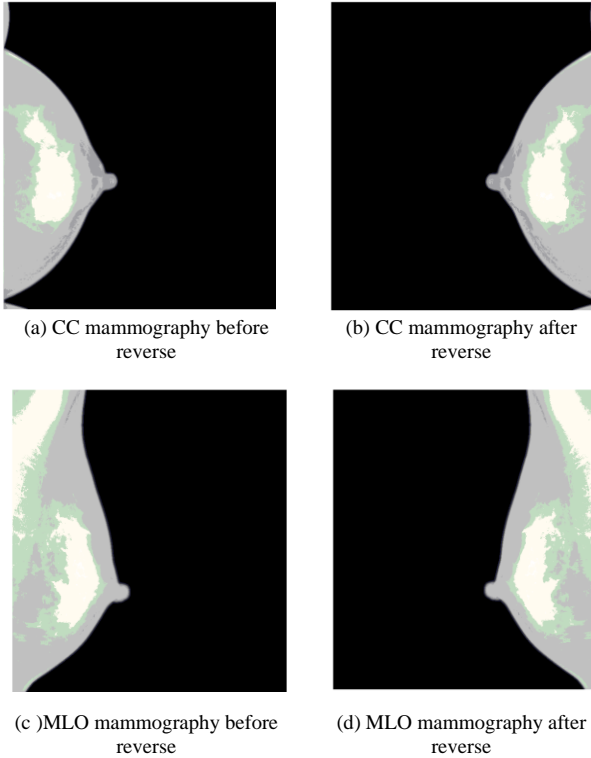
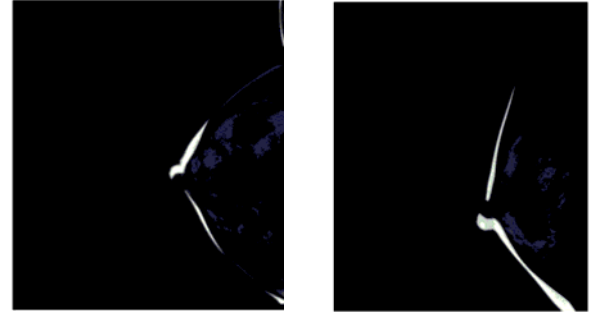


FIGURE III. TWO SIDE OF MAMMOGRAPHY REVERSAL



(a) CC mammography after subtract (b) MLO mammography after subtract

FIGURE IV. TWO SIDE OF MAMMOGRAPHY AFTER SUBTRACT

B. ROI Extraction

In the diagnosis of mammograms, radiologists are mainly interested in the part of breast glandular tissue and not care about the black background region in mammograms [4]. Therefore, we will detect the edge of the difference mammography at first; Then, ROI can be intercepted by the method of interactive edge detection. Finally, the original images are separated into two parts, ROI and the background region. Generally, the purpose of edge detection can be realized by conventional Sobel [5] edge detection. The steps are as follows: Using two masks respectively to filter the images. Then, calculate the square of pixel value after filtering at twice times and adding the results of two images. Finally, the square root of the sum result is calculated [6]. At this point, the gradient of the pixel points is obtained. To compare the threshold, we can acquire edge detecting effects. However, owing to the presence of a background region in black which has lots of noises, the effect of edge detection is not well.

That paper proposed an improved Sobel method to achieve edge detection of mammograms. The improved Sobel method turns the two marks into eight, which means calculating the gradient of the point in all directions and determines the most accurate value ultimately. In this way, we can get relative ideal edge detection result [7]. Figure.5 shows the image neighborhood and mark under the improved Sobel algorithm. Figure 6 shows the result of using improved Sobel method in in the CC or MLO difference mammography. According to the edge detection results, the region of interest is shown in Figure 7.

C. Lossless Compression Based on Run-length Coding

After the original images are divided into ROI and background region, we can lossless compress ROI. Entropy coding [8] is one of the most common methods for lossless compression in image coding. It is a coding method which is based on statistical characteristics for signals. Therefore, the probability of image data and the relevance between data make much difference to choose an appropriate coding method. Analyzing on region of interest data in mammograms, we can find a relatively large range of image data which range from 0 to 2500. In the image, But the difference between each pixel and the pixel value around is not much, and there are more

continuous values. Thus, the method of run-length coding [9] is selected to compress the region of interest. What's more, to further improve the result of encoding compression, reducing the data range and increasing the times of continuous data appeared, we make subtraction between the part of data lines. As far as the background region, which pixel values are almost zero, we remain it without any further processing.

Before four mammograms are calculated through analysis processing, two reference images have been selected while the region of the data in black background is fully preserved. Besides, the data in the background region are almost the same when a set of images are in the same view. So, we only need to transmit the reference images and the compressed data for difference images of ROI, and then using zero-padding method and superposition to process the reference images and difference images of ROI [10]. This transmission process not only can remain the effective image information adequately, but also greatly improved compression ratio of four images. Algorithm 1 shows the compression coding processing. The inputs are four mammograms, which are image A (CC views of the left side), image B (MLO views of the left side), image C (CC views of the right side) and image D (MLO views of the right side). Outputs are the compression results of images C and D. First, we select reference images. After reversing non-reference images, we make a difference between reference images and non-reference images (line 3-4). Then, detect the breast edge using improved Sobel method. Regions of interest (ROI) can be obtained (line 5-6). Finally, lossless compressing the ROI of difference images using run-length coding. We can obtain the compression results (images C and D) and return (line 7-8).

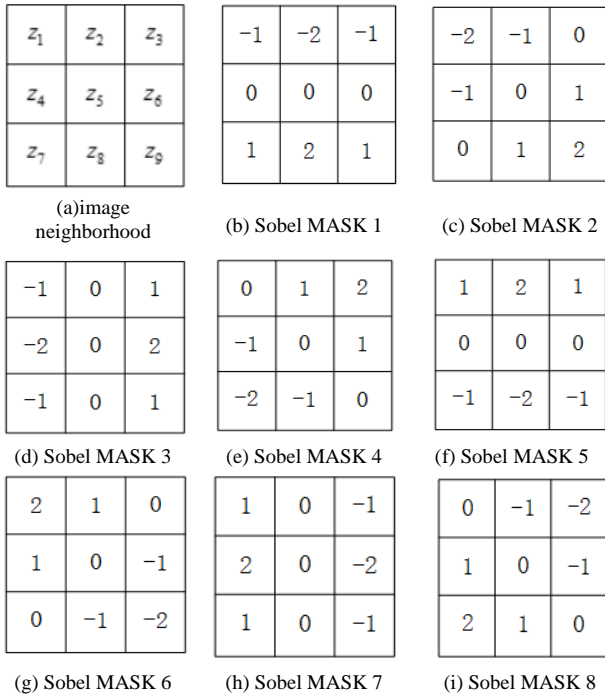
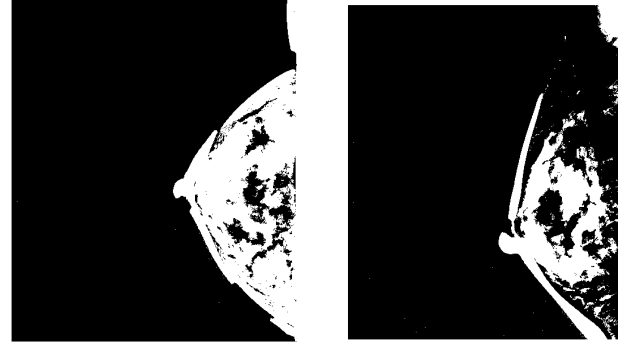
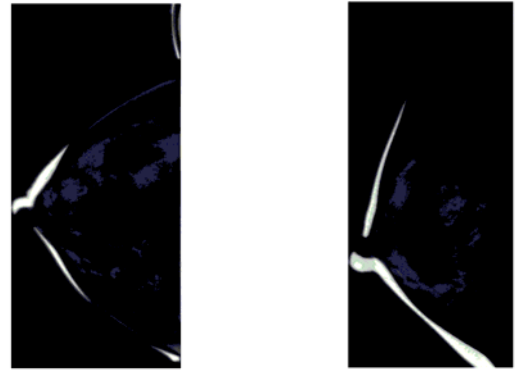


FIGURE V. THE NEIGHBORHOOD AND MASK OF IMPROVED SOBEL METHOD



(a) CC mammography (b) MLO mammography

FIGURE VI. THE EDGE DETECTION RESULT OF A SET DIFFERENCE MAMMOGRAPHY



(c) CC mammography (d) MLO mammography

FIGURE VII. THE EXTRACT RESULT OF THE ROI OF MAMMOGRAPHY

III. EXPERIMENT AND DISCUSSION

In this paper, we use four mammograms in DICOM format as experimental data. All images are come from imaging department of tumor hospital in Liaoning Province. These mammograms are shot by Senographe 2000D full digital mammography camera. Each of the size of the images is 2294*1914 and each pixel values are 16bit. Running environments of the program are Matlab2013 and Microsoft Visual Studio 2010 while the computer configuration is Intel(R) Core(TM) 2 processor, 2.00G main memory and operating system in 32bit.

Experiments on the compression method proposed show that after making subtraction between the reversal non-reference images and reference images, the sizes of difference images are 8781bit and 432bit respectively. The sizes of difference images of CC views and MLO views are 2294*816 and 2294*829 respectively. Results of compression are shown in table 1. Compression ratio of CC views and MLO views can reach on 9.30 and 13.63, which are great compression results.

Algorithm 1: compression coding algorithm

```

1  //input : there are four mammograms that are image
   A(CC views of the left side) ,image B(MLO views of the left
   side), image C(CC views of the right side) and image
   D(MLO views of the right side)

2  //output : compression results of image C and D

3  images A and B are selected as reference images. After
   specular reversing images C and D, images C' and D' are
   obtained;

4  by using two pairs of reverse and reference images which
   are images A, C' and images B, D', we can obtain difference
   images C'' and D'';

5  for images C'' and D'', using improved Sobel method to
   detect breast edge and obtain rectangular regions E and F of
   breast tissue;

6  to subtract rectangular data of regions E and F line by
   line respectively. That is to say, data in first line are retained
   and the second lines of data are replaced by the difference
   between the first line and the second line from the original
   rectangular. By that analogy, we can obtain a new
   rectangular and use it to displace regions E and F;

7  after using the method of run-length coding to lossless
   compress transformed data matrix, results of compression of
   image C and D are obtained.

8  return;

```

TABLE I .THE COMPRESSION RESULT OF MAMMOGRAPHY

Visual type	The size of difference images (bit)	The size of Images after compress ROI (bit)	compression ratio C_r
CC	8781432	944616	9.2963
MLO	8781432	644144	13.6327

IV. CONCLUSION

This paper propose an improved Sobel method to achieve the breast edge detection of the difference images in same views and obtain ROI, which based on the characteristics for four mammograms that great image resolution, the relevance among images, concentrated breast tissue, more regions of background, etc. Then compress the data of ROI by using run-length coding method. Therefore, the compression of mammograms in Narrow band Transmission Environment can be achieved. Experiments under real data indicate that using the method can obtain great compression results.

ACKNOWLEDGMENT

This research was supported by the National High Technology Research and Development Plan (863 Plan) under Grant No. 2012AA02A606.

REFERENCES

- [1] K Thangavelu and T Krishnan, "Lossless color medical image compression using adaptive block-based encoding for human computed tomographic images," *International Journal of Imaging Systems and Technology*, vol.23, no.3, pp. 227-234, Sep, 2013.
- [2] D. S. Wener Borges, D. P. Anselmo Cardoso, Aristofanes Correa "Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees," *Expert Systems With Applications*, vol. 42, pp. 8911-8928, June. 2015.
- [3] T Yoshida, "Background differencing technique for image segmentation based on the status of reference pixels," *International Conference on Image Processing (ICIP 2004)*, Singapore, 2004, pp.3487-3490
- [4] Hong Byung-Woo and Sohn Bong-Soo, "Segmentation of Regions of Interest in Mammograms in a Topographic Approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no.2, pp.129-139, Jan. 2010.
- [5] DH Shen, XE Xu, LC Zhang, "Application of Improved Sobel Algorithm in Medical Image Edge Detection," *3rd International Conference on Mechatronics and Control Engineering (ICMCE)*, Zhuhai China, 2014, pp. 151-154.
- [6] G Chaple and RD Daruwala, "Design of Sobel Operator based Image Edge Detection Algorithm on FPGA," *3rd International Conference on Communications and Signal Processing (ICCSP)*, Melmaruvathur India, 2014, pp. 165-169.
- [7] Gao Wenshuo, Yang Lei, Zhang Xiaoguang, Liu Huizhong, "An Improved Sobel Edge Detection," *3rd IEEE International Conference on Computer Science and Information Technology*, Chengdu China, 2010, pp. 67-71.
- [8] TH Chan and A Grant, "Network Coding Capacity Regions via Entropy Functions," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5347-5374, Sep. 2014.
- [9] Zhan Wenfa, Shi Bing, Zha Huaizhi, "A Scheme of Test Data Compression and Power Reduction Based on Common-Run-Length Coding (CRLC)," *International Conference on Affective Computing and Intelligent Interaction*, Taiwan, 2012, pp.215-223.
- [10] R Soundararajan and AC Bovi, "RRED Indices: Reduced Reference Entropic Differencing for Image Quality Assessment," *IEEE Transactions on Image Processing*, vol.21, no.2, pp.517-526, Feb. 2012.