COMPUTATIONAL IMAGING AND VISION

Digital Mammography

Nijmegen, 1998

Nico Karssemeijer, Martin Thijssen,
Jan Hendriks and Leon van Erning (Eds.)

Kluwer Academic Publishers

DEVELOPMENT OF A NEW ALGORITHM FOR DETECTION OF MAMMOGRAPHIC MASSES

TOMOKO MATSUBARA, HIROSHI FUJITA [1], TAKESHI HARA [1], SATOSHI KASAI [1], OSAMU OTSUKA [1], YUJI HATANAKA [1] AND TOKIKO ENDO [2]

Department of Information Processing, Nagoya Bunri College, Japan
e-mail : tomoko@nagoya-bunri.ac.jp
[1] *Department of Information Science, Faculty of Engineering, Gifu University, Japan*
[2] *Department of Radiology, Nagoya National Hospital, Japan*

1. Introduction

The rate of breast cancer occurrence is increasing and it is estimated that breast cancer will be the top cause of Japanese women's cancer mortality quite soon. The examination by mammography is now becoming a major diagnostic tool for finding the cancers at an early stage. However, the number of the expert doctors is not enough in the visual interpretation of mammography and a computer-aided diagnosis (CAD) system to aid physicians is deeply required. Therefore, many research groups have developed automated schemes for detecting masses [1-6] and clustered microcalcifications on mammograms.

We have also developed automated-detection schemes for the masses on digital mammograms [7-10]. The purpose of this study is to improve our methods by adding new processes for detecting the masses existing around thick mammary gland and reducing the false-positive candidates.

2. Methods

The automated scheme for the detection of breast masses is shown in *Fig.1*. Each step in this scheme is shortly explained below.

2.1. IMAGE DIGITIZATION AND EXTRACTION OF BREAST REGION

The digital images used as the object of analysis in this study are mammograms of 25.4 cm$\times$20.3cm in size which are digitized with a sampling interval of 0.1mm and a density resolution of 12 bits. The border of the breast (skin line) on the digital mammogram is automatically extracted and then smallest rectangle region containing the breast area is cut off [11]. The image matrix is reduced to produce the image with sampling interval of 0.4mm for the detection of masses.

2.2. SEGMENTATION AND CLASSIFICATION

Pectoralis muscle region included in the mammograms obtained from MLO view is firstly extracted. All digital mammograms are classified by histogram-analysis technique into four categories, i.e., (i) fatty, (ii) glandular and fatty, (iii) dense, and (iv) high dense that does not fit to diagnose. For the case of (ii), the breast region is further segmented into glandular and fatty parts by binalization.

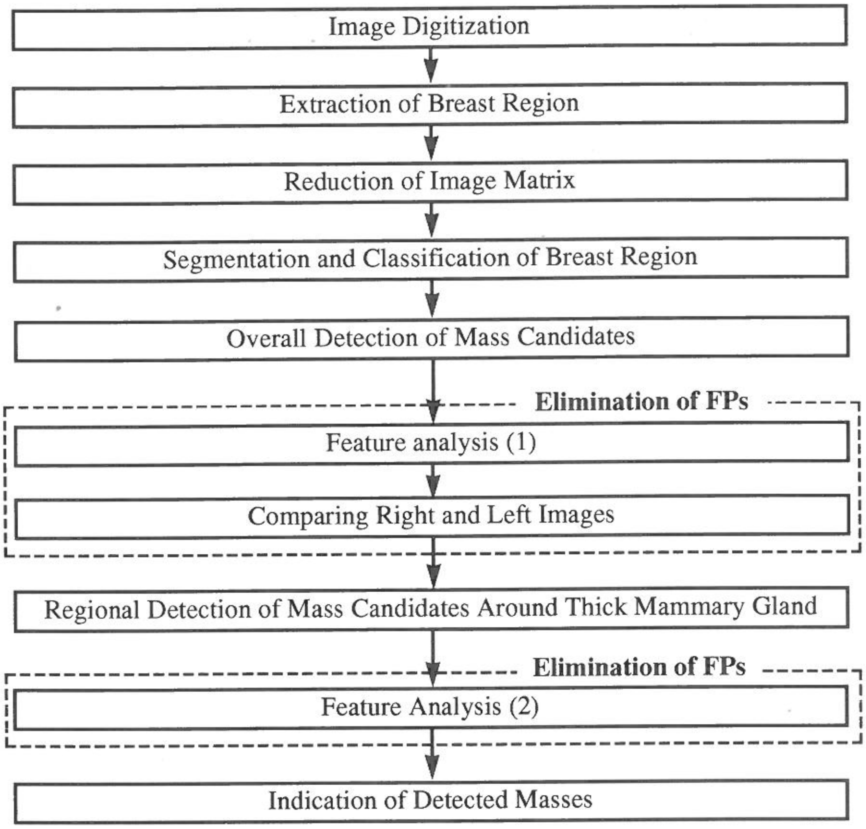

Figure 1. Detection process for breast masses

2.3. OVERALL DETECTION OF MASS CANDIDATES

Because the masses are similar to circular patterns and appeared as white regions on mammograms, the low photographic density areas, as first candidates of masses, are detected by using thresholding technique for each of the three categorized images and segmented parts with separate thresholding conditions.

The regions whose sizes are smaller than a defined one (5.6mm) are eliminated. The much lower density regions are extracted from the candidates selected by the analyses with their circularity, standard deviation of pixel values and size. By this process, it is possible to extract the "true" mass area precisely from the region with normal tissue [12].

2.4. FEATURE ANALYSIS (1)

2.4.1. *Gray Level Difference and Co-occurrence Matrix*

The feature analyses by contrast in gray level difference and by entropy, inverse difference moment and angular second moment in co-occurrence matrix are preformed to eliminate the FP candidates [13].

2.4.2. *Discriminant Analysis*

The main aim in this process is to eliminate linear-shape FP candidates, such as blood vessels and parts of mammary gland region.

The features used are as follows: (1) length-to-width ratio, (2) minimum width, (3) circularity, (4) average contrast in the candidate, (5) average contrast in the central part of the candidate, (6) average of the standard deviations of pixel-value distributions in the equal distance from the gravity, (7) roughness(+) in the central part of the candidate, (8) roughness(-) in the central part of the candidate, (9) percentage of the gradient-component ratio in constant directions determined by nipple position, (10) standard deviation for the central part of the candidate in unsharp-mask processed image and (11) gradient ratio for each direction obtained using the gravity.

The candidates are classified by discriminant analysis with Mahalanobis generalized distance [14].

2.5. COMPARING RIGHT AND LEFT IMAGES

It is possible to eliminate FPs by comparing right and left images because the architecture between normal right and left mammograms is generally symmetry.

The candidates are classified by correlation coefficient and mean of absolute difference of pixel values, and by mean absolute difference and correlation coefficient of density gradient with the corresponding regions in the fellow of the paired images which are aligned by the matching method [15].

2.6. REGIONAL DETECTION OF MASSES AROUND THICK MAMMARY GLAND

The main aim in this process is to detect the specific masses existing around thick mammary gland regions because it is difficult to detect these masses by simple extraction of low density regions with thresholding technique described in section 2.3.

The candidates are detected by the following four features: (1) similarity with the four reference patterns shown in *Fig. 2*, (2) four concentrating features by using density gradient within the areas shown in *Fig. 3*, (3) average and (4) standard deviation of pixel values within ROI [16].

2.7. FEATURE ANALYSIS (2)

The same processes in sections 2.4 and 2.5, and also the other feature analyses of edge's shapes of binary images within mass candidate region by different three thresholds and standard deviation are applied to the candidates detected in the previous section [16].

The residual candidates after this feature analysis for eliminating FPs are indicated as "true" masses.

3. Results

The database for testing this new scheme consists of 231 digitized mammograms (99 abnormal and 132 normal), which include 39 true masses existing around thick mammary gland regions. All of these masses were not able to be detected in our previous method. When the scheme was tested with this database, it achieved a true-

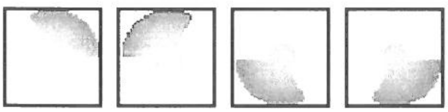

Figure 2. Shapes used by template matching technique

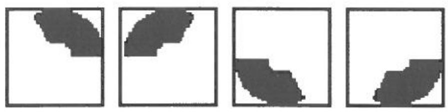

Figure 3. Black areas are used to calculate concentrating features by using density gradient

positive fraction of 96% (63%) with an average of 2.6 (1.1) FP detections per image. It was possible, at 96% true-positive fraction, to detect 35 masses in 39 masses, all of which were not able to be detected in our old method. In addition, the number of FPs was about 4.0 per image to achieve the sensitivity of more than 90% in our previous method.

We are also working on classifying the masses into benign and malignant [17-19].

4. Conclusion

Our improved scheme for the mass detection, in which the processes for detecting mass candidates existing around thick mammary gland and for eliminating FPs were newly added, identified 96% of true masses at an average of 2.6 FP detections per image. The performance of our scheme was improved and it was confirmed that the effectiveness of our newly added processes. For the future work, it is necesary to add the new algorithms for decreasing FPs.

This research was supported in part by grants from the Grant-In-Aid for Cancer Research from the Ministry of Health and Welfare, the Hori Information Science Promotion Foundation, and TEPCO Research Foundation.

References

[1] Lai SM, Li X, Bischof WF (1989) On techniques for detecting circumscribed masses in mammograms.IEEE Trans. Med. Imaging 8, pp 377-386.
[2] Yin FF, Giger ML, Doi K, et al.(1991) Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. Med. Phys. 18, pp 955-963.
[3] Yin FF, Giger ML, Vyborny CJ, et al.(1993) Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. Invest. Radiol. 28, pp 473-481.
[4] Yin FF, Giger ML, Doi K, et al.(1994) Computerized detection of masses in digital mammograms: Investigation of feature-analysis techniques. J. Digital Imaging 7, pp 18-26.
[5] Yin FF, Giger ML, Doi K, et al.(1994) Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. Med. Phys. 21, pp 445-452.
[6] Zheng B, Chang YH, Gur D (1995) Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. Acad. Radiol. 2, pp 959-966.
[7] Hara T, Fujita H (1995) Template matching of gray-scale images using a genetic algorithm. The Trans. of IEICE D-II J78-D-II, pp 385-388.
[8] Fujita H, Endo T, Matsubara T, et al. (1995) Automated detection of masses and clustered microcalcifications on mammograms. Proc. SPIE 2434, pp 265-269.
[9] Matsubara T, Fujita H, Endo T, et al. (1996) Development of mass detection algorithm based on adaptive thresholding technique in digital mammograms. In: K Doi et al (eds.), Digital Mammography, Elsevier, Amsterdam, pp 391-396.
[10] Matsubara T, Fujita H, Endo T, et al. (1997) Development of a high-speed processing algorithm for mass detection based on thresholding technique in mammograms. Med. Imaging Tech. 15, pp 1-13.
[11] Kato M, Fujita H, Hara T, et al. (1997) Improvement of automated breast-region-extraction algorithm in a mammogram CAD system. Med. Imaging and Inform. Sci. 14, pp 104-113.
[12] Matsubara T, Kasai S, Seki K, et al. (1998) Development of a computer-aided diagnostic system for mammogram: Improvement of the method of extracting low-density regions during automated mass detection. J. Jpn. Assc. Breast Can. Scr. 7, pp 87-101.
[13] Ostuka O, Kasai S, Hara T, et al. Elimination of false-positive mass candidates in a mammogram CAD system. submitted.
[14] Kasai S, Fujita H, Hara T, et al. Elimination of linear-shape false-positive candidates in an automated detection algorithm for mammographic masses. submitted.
[15] Kasai S, Fujita H, Hara T, et al. Elimination of false-positive candidates by comparing right and left mammograms in an automated mass detection algorithm. submitted.
[16] Kasai S, Fujita H, Hara T, et al. Detection algorithm for masses around thick-mammary-gland regions on mammograms. submitted.
[17] Ueda H, Fujita H, Endo T, et al. (1995)Automated detection of spicules in mass lesions on mammograms. Med. Imaging and Inform. Sci. 12, pp 68-73.
[18] Goto M, Endo T, H Fujita (1996) Studies on diagnostic logic for classifying benign and malignant masses on mammograms. Jpn. J. of Med. Elect. and Biol. Eng. 34, pp 62-67.
[19] Tani Y, Hara T, Fujita H, et al. Development of mass classification system on mammograms. submitted.