

A Contour-based Mass Segmentation in Mammograms

Yu Zhang, Noriko Tomuro, Jacob Furst, Daniela Stan Raicu
DePaul University
College of Computing and Digital Media, School of Computing
Chicago, IL
{jzhang2, tomuro, jfurst, draicu}@cs.depaul.edu

Abstract

This paper presents a novel segmentation method for identifying mass regions in mammograms. This work is a part of an on-going project whose aim is to build a Computer-Aided Diagnosis (CADx) system that classifies suspicious cancer masses in mammograms as benign or malignant. Segmentation of suspicious mass regions is an important pre-processing step to achieve high accuracy results, because the regions of interest (ROI) marked by radiologists are oftentimes imprecise. For each ROI, we first apply an enhancement function to increase the grey-level contrast of the image, then a filter function to reduce noise. Next, for each pixel in the ROI, we compute the energy feature based on the co-occurrence matrix of the pixel. Finally we extract a contour of the mass from the energy feature image using an edge-based segmentation technique. We present some examples of our segmentation method and discuss future work.

1 Introduction

In 2008, it was estimated that 182,460 new cases of breast cancer would be diagnosed among women in the United States. Breast cancer is the second leading cause of cancer related deaths for women in the U.S. after lung cancer [1]. Breast cancer is treatable when discovered early, so early detection is the best protection. At present, the most effective method for breast cancer early detection is mammography screening [1]. A mammogram is an X-ray picture of the breast that can reveal abnormalities at earlier stage. Microcalcifications and masses are two most common types of suspected abnormalities in mammogram images [2]. Radiologists analyze the mammogram images and detect any abnormal growths. Based on the level of suspicion of the abnormality, radiologists usually recommend a routine follow up, or a biopsy to confirm the result [2].

The problem with mammography screening is that the error rate is high. At present, the sensitivity of most mammography reports is in the 68–92% range [3]. Many Computer-Aided Diagnosis (CADx) systems have been developed as a second opinion to assist radiologists [4]. The research presented in this paper is part of an ongoing project to develop a reliable CADx system to classify malignant or benign of suspicious masses in mammograms.

Our CADx system classifies a mammogram image annotated with a Region of Interest (ROI) marked by a radiologist in five stages: ROI extraction, mass segmentation, feature extraction, feature selection and classification. At the classification stage, we plan to apply an optimal feature set to multiple classifiers, and construct an ensemble of classifiers. Figure 1 below depicts the schematic framework of our CADx system.

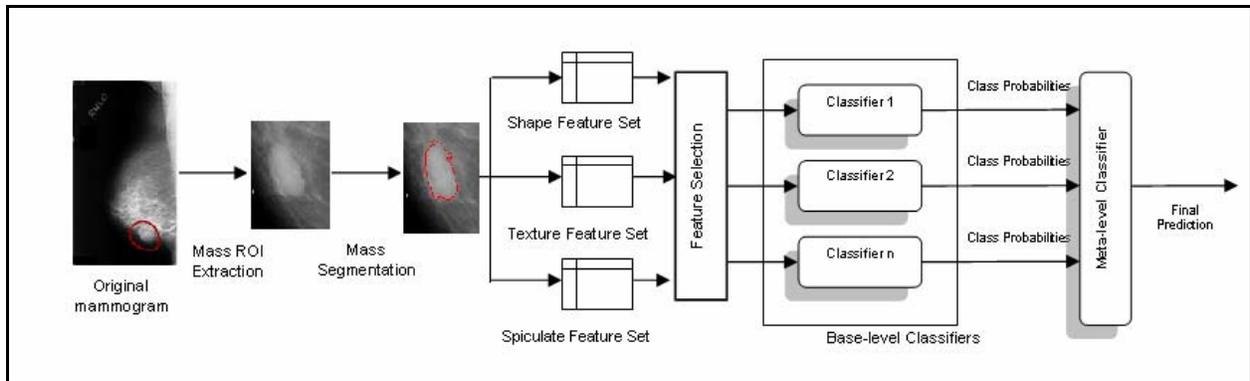


Figure 1. Framework of computer-aided diagnosis (CADx) system

In this paper, we present a novel contour-based mass segmentation method to improve classification. In particular, we compute *energy* (one of Haralick [5] descriptors) at pixel-level from a mass ROI. Then we extract connected edges of the energy values to detect a possible mass region. The outer boundary of detected region is used as the final segmented outline of the mass. This paper is organized as follows. In Section 2, we review related works of mass segmentation. In Section 3, we present the methodology. In Section 4, the preliminary results are presented. In Section 5, the conclusion and future work are discussed.

2 Mass Segmentation

Mass is one of two common types of abnormalities in mammogram images. Masses are thickenings of breast tissue which appear as lesions, with the size ranging from 3mm to 30mm. The shape and margin of masses are two important criteria to distinguish malignant from benign masses [2]. Usually a poor defined shape is more likely to be malignant than a well-circumscribed mass. Margin is the border of a mass. Ill-defined margins or spiculated lesions are much more likely to be malignant [2, 7].

Mass segmentation separates a mass from its background and captures the contour of the mass. Only after finishing the mass segmentation, shape features and spiculation levels of a mass can be computed for classifying benign and malignant masses. Previous studies have shown that improving the accuracy of mass region segmentation can significantly improve the performance of classifications [8].

Digital Database for Screening Mammography (DDSM) [9] from University of South Florida is the largest publicly available resource for the mammogram analysis research community. In DDSM images, suspicious regions (ROIs; including masses and microcalcifications) are marked by experienced radiologists, and BI-RADS [6] information is also available for each abnormal region. However, ROIs are usually marked by a circle for the mass location, and do not trace the accurate outline of the mass. Therefore, more precise mass segmentation is necessary and essential prior to classification.

In Medical Image Processing, many mass segmentation methods for mammogram images have been developed [8, 10, 11, 12, 13]. There are two common mass segmentation approaches: 1) region-based methods and 2) contour-based methods. In region-based methods, mass regions are iteratively grown by comparing all neighboring pixels and including the pixels with similarity to the respective regions. The similarity is measured with different types of properties,

such as pixel intensity or computed texture features. In contour-based methods, segmentation is commonly done by techniques based on edge detection.

3 Dataset and Methodology

3.1 Data Description

In this work, all mass ROI images were extracted from the aforementioned Digital Database for Screening Mammography (DDMS) [9] from University of South Florida. In our experiment, 649 mass instances were selected from the images digitized by LUMYSIS, which had the highest resolution comparing to the other two types of digitizers. Among the 649 instances, 303 instances were benign masses; 346 instances were cancer.

3.2 Mass Segmentation Methods

Boundaries of suspicious regions marked by radiologists were represented as chain code in the original mammograms in DDMS. For each mass suspicious region, using MatLab we extracted a rectangle region which including the mass boundary and its surroundings as mass ROI. After the extraction, three steps were performed to segment the mass ROIs and detect the contour of the masses.

Step1: Image Enhancement

Since most mammogram images are in low grey contrast, linear or nonlinear contrast stretch functions are usually applied to mammograms. Also, filtering functions are often applied to mammogram image to remove noise. In our experiment, a linear function was used to increase the image contrast first, then an average filtering was applied to reduce noise. We also tested different filtering parameters, and an optimal parameter was applied. Figure 2(a-c) shows an example original image (a), contrast adjusted image (b) and filtered image (c).

Step 2: Compute pixel-based texture feature – energy image

Haralick descriptors [5] are commonly used texture features, which are by often used in image classification and segmentation. They are computed from gray-level co-occurrence matrices [14]. These texture features can be extracted from an image at three levels: global level (the entire organ), local level (a region within the organ), and the pixel level. Energy is one of Haralick descriptors, which measures the heterogeneity of an image. In our experiment, the energy texture feature was extracted at pixel level. We first obtained a co-occurrence matrix with angle=0, distance=1 for each pixel [14] using a neighborhood of size 7x7. Then from the co-occurrence matrix, we computed the energy value as the counts of occurrences of repeated pairs within a co-occurrence matrix [5]:

$$Energy = \sum_i^M \sum_j^N P_{ij}^2$$

where M and N represent the co-occurrence row and column sizes, and P_{ij} represents the probability of pixel pair (i, j) occurrences. Figure 2(d-f) shows examples of pixel-level energy

feature images obtained from the original mass ROI (d), from the contrast adjusted image (e) and from the filtered image (f).

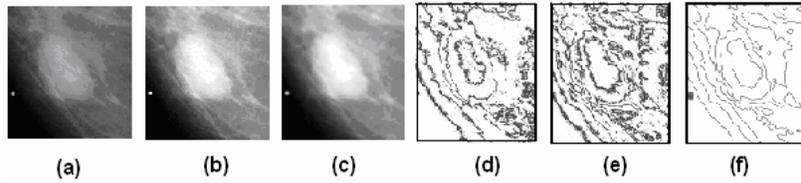


Figure 2. Original mass ROI, contrast adjusted image and filtered image, related pixel-level energy feature images and extracted contours

Step 3: Find the contour of the mass region

From those energy feature images, a binary image was extracted by using a threshold. Then using morphological operation, connected edges were used as boundaries to detect possible mass regions. Next, noise was removed and the mass was identified. Finally the outer boundary of the detected region was used as the final segmented contour of the mass. Figure 3 shows each step of mass segmentation: (a) the contrast adjusted and filtered mass ROI; (b) the energy feature image; (c) the binary image; (d) the possible mass region; (e) the detected mass; (f) the contour of segmented mass.

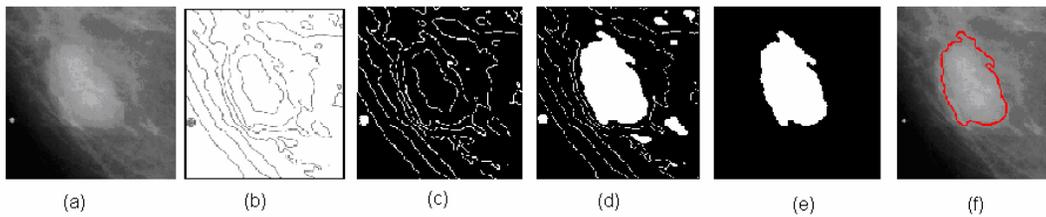


Figure 3. Mass segmentation

4 Preliminary Results

In our experiment, we applied the algorithm on the dataset from DDSM as described earlier. The preliminary investigation indicated that by visual inspection the contours of the automatically segmented masses were in general more accurate than the contours of radiologists marked regions. Figure 4 shows four successfully segmented examples. Each example is shown with a set of three images in a row: (a) the original image, (b) the radiologists marked mass region, (c) the contour obtained by our mass segmentation method.

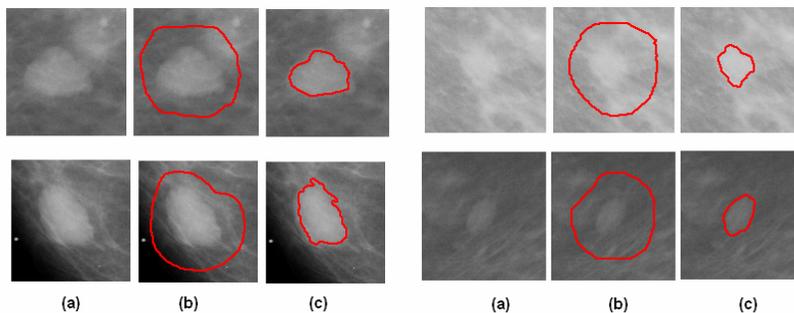


Figure 4. Four examples of contour extracted on mass ROIs

5 Conclusions and Future Work

To improve the classification accuracy of benign and cancer masses in mammograms, we developed a novel method for segmenting the mass ROIs. Although our work is still at an investigation stage at present, the preliminary result seems promising. For the next task, we will experiment other image enhancement and filtering methods as well as parameter values for mass segmentation, and combine our segmentation method with other segmentation algorithms. After segmentation, we will compute texture features, shape features and spiculation feature of the segmented masses, and feed those features to a classification system.

References

- [1] National Cancer Institute, "American Cancer Society Cancer Facts & Figures 2008" (<http://www.cancer.org>).
- [2] D. J. Winchester, D. P. Winchester, C. A. Hudis and L. Norton , *Breast Cancer(Second Edition)* , 2007, Springer.
- [3] B. C. Yankaskas, M. J. Schell, R. E. Bird, D. A. Desrochers. "Reassessment of Breast Cancers Missed During Routine Screening Mammography: A Community-Based", *American Roentgen Ray Society*, vol. 177, 2001, pp. 535-541.
- [4] T. Mu; A. Nandi; R. Rangayyan, "Classification of Breast Masses Using Selected Shape, Edge-sharpness, and Texture Features with Linear and Kernel-based Classifiers ", *Journal of Digital Imaging*, Volume 21, Number 2, June 2008 , pp. 153-169(17)
- [5] R. M. Haralick, "Statistical and structural approaches to texture", *Proceedings of the IEEE*, vol. 67, pp. 786-804, 1979.
- [6] C.J. D'Orsi, L.W.Bassett and W.A. Berg et al, *Breast Imaging Reporting and Data System: ACR BI-RADS-Mammography (ed 4)*, 2003, American College of Radiology, Reston, VA.
- [7] H.D. Cheng, X. J. Shi, R.Min, L.M. Hu, X.P. Cai and H.N.Du, "Approaches for automated detection and classification of masses in mammograms", *Pattern Recognition*, 2006.
- [8] L. Jiang, E. Song, X. Xu, G. Ma and B. Zhang, "Automated Detection of Breast Mass Spiculation Levels and Evaluation of Scheme Performance", *Acad Radiol*, 2008.
- [9] M. Heath, K. Bowyer, D. Kopans, R. Moore and W. P. Kegelmeyer, "The Digital Database for Screening Mammography", *Proceedings of the Fifth International Workshop on Digital Mammography*, 2001.
- [10] O. Marlagelada, *Automatic Mass segmentation in mammographic images (Doctor Dissertation)*, 2007, Universitat de Girona.
- [11] M.A. Kupinski and M.L. Giger, "Automated Seeded Lesion Segmentation on Digital Mammograms", *IEEE Transaction on Medical Imaging*, vol.12, 1998.
- [12] W. Xu, S.Xia, M. Xiao and H. Duan, "A Model-based Algorithm for Mass Segmentation in Mammograms", *Engineering in Medicine and Biology 27th Annual Conference*, 2005.
- [13] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono and E. Colini, "Breast Mass Segmentation in Mammographic Image by an Effective Region Growing Algorithm", *Advanced Concepts for Intelligent Vision Systems Conference*, 2008.
- [14] R. Susomboon, D.S. Raicu, and J.D. Furst, "Pixel-Based Texture Classification of Tissues in Computed Tomography", *DePaul CTI Research Symposium*, 2006.