Segmentation of the Breast Region with Pectoral Muscle Removal in Mammograms

Zhili Chen
zzc09@aber.ac.uk
Reyer Zwiggelaar
rrz@aber.ac.uk

Department of Computer Science
Aberystwyth University
SY23 3DB, UK

Abstract

Breast region segmentation is an essential prerequisite in computerised analysis of mammograms. It aims at separating the breast tissue from the background of the mammogram and it includes two independent segmentations. The first segments the background region which usually contains annotations, labels and frames from the whole breast region, while the second removes the pectoral muscle portion (present in Medio-Lateral Oblique (MLO) views) from the rest of the breast tissue. In this paper we propose a fully automated breast region segmentation method based on histogram thresholding, edge detection in scale space, contour growing and polynomial fitting. Subsequently, pectoral muscle removal using region growing is presented. To demonstrate the validity of our segmentation algorithm, it is extensively tested using over 240 mammographic images from the EPIC database. The qualitative evaluation of experimental results indicates that the method can accurately segment the breast region in a large range of digitised mammograms, covering all density classes.

1 Introduction

Breast region segmentation is an important prerequisite in computerised analysis of mammograms. It aims at excluding the background from further processing. The precise segmentation of the breast region with a minimum loss of breast tissue facilitates the search for abnormalities, the modelling of parenchymal tissue and accurate registration. There have been various approaches to segmentation of the breast region in mammograms [1-6]. The methodologies described in these approaches are summarised in [7], which provides a breakdown into histogram, gradient, polynomial modelling, active contours, and classification based methods. The developed methodology, which is presented in this paper, takes a number of these approaches and combines them into a robust methodology.

2 Breast Region Segmentation

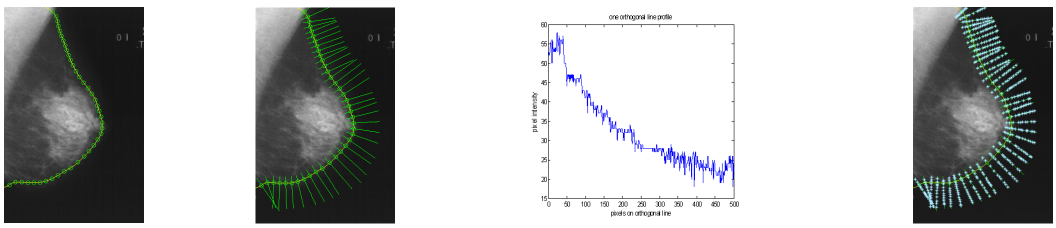
The proposed segmentation incorporates histogram thresholding, edge detection, active contour and polynomial fitting. The original images (Figure 1(a)) to be segmented contain left and right MLO mammograms and need to be split into individual mammograms.

A global threshold, after Gaussian smoothing the histogram, is determined using the minimum between the peaks of the background and the breast tissue (Figure 1(b)). The resulting binary image contains a number of objects. We use a Connected Component Labelling (8-connected) algorithm [7] to remove the labels in the background region and the annotations in the frame from the whole image. Subsequently, we isolate the frame (near the edges of the image) and smooth the remaining region applying a Gaussian low-pass filter. We then split the union region into two separate breast regions to form two binary masks (Figure 1(c)).

(a) original mammographic image (b) binary mammographic image (c) binary masks

Figure 1: Global thresholding. Labels and annotations removal.

The approximate segmentation (Figure 2(a)) is refined using scale-space based edge detection. Firstly, we evenly place 40 points on the mask boundary (Figure 2(a)). For each point a corresponding orthogonal line is obtained (Figure 2(b)). The length of one orthogonal line is 500 pixels (100 pixels inside the mask, 400 pixels outside the mask). One orthogonal line profile (Figure 2(c)) is illustrated to show the lack of a distinct edge. We then perform edge detection to search probable breast boundary points by convolving the pixels on orthogonal lines with a derivative of Gaussian kernel at multiple scales [8]. We use a range of small scales to increase sensitivity to the low contrast breast boundary. Edge detection starts at a relatively coarse scale within the range to suppress noise, and ends at a fine scale to improve accuracy. Probable breast boundary points are achieved by detecting minima (Figure 2(d)).

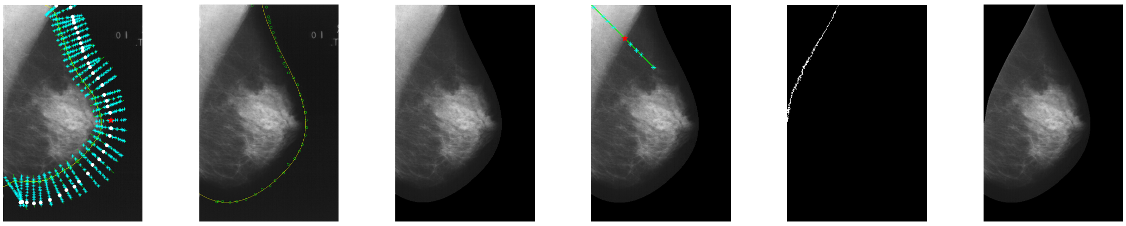


(a) mask boundary (b) orthogonal lines (c) one profile (d) probable points

Figure 2: Overview of edge detection in scale space.

The first step of contour growing is finding the starting orthogonal line and selecting the seed point from all the probable breast boundary points on this line. The contour will be grown in either direction from the seed point. We give priority to choosing the orthogonal line close to the x axis direction as the starting line. Subsequently, we use an edge strength measure to search the seed point along the orthogonal line in the direction from outside to inside the breast. Ideally, the seed point could be found at the boundary point whose edge strength is the first local maximum. If no such a seed exists on this starting line, other alternatives close to this line will be used to search the seed point dynamically. After the seed point is obtained a contour growing process starts based on a contour growing measure combining different criteria. For a seed point, probable breast boundary points obtained using edge detection in scale space on the neighbour orthogonal line along the contour growing direction are regarded as a set of candidate growing points for searching a new seed point. The contour growing measure is calculated for all candidate points to decide the new seed point with the minimum measure value.

The contour growing measure is defined by a weighted function, following the typical snake additive model formulation [9]. This measure includes intensity, edge strength and angle information. Once 40 seed points have been obtained from all the probable breast boundary points on 40 orthogonal lines, contour growing is finished, and these 40 seed points comprise an initial breast boundary (Figure 3(a)). After we obtain an initial breast boundary comprising 40 points, we first order them to solve the misordering due to the intersection of orthogonal lines, and we combine close points into one point. Subsequently, a cubic polynomial fitting is used to yield a smooth and continuous contour as the final breast boundary (Figure 3(b),(c)).



(a) (b) (c) (d) (e) (f)

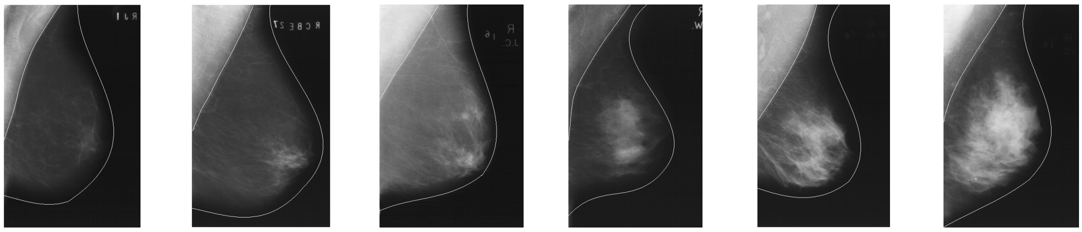
Figure 3: (a) The first seed point (red star) and the initial breast boundary (white circles). (b) Cubic polynomial fitting. (c) Breast region. (d) Seed point (red star). (e) Grown region. (f) Pectoral muscle removal.

The breast region obtained above is the union region of the breast and the pectoral muscle. We use a region growing method to remove the pectoral muscle. Firstly, we place a seed point close to the border between the pectoral muscle and the breast instead of placing a seed point inside the pectoral muscle region [7]. Specifically, we draw a line (slope equal to 1) from the first pixel of the non-curved side into the breast, and then we detect edges on this line in scale space using the method mentioned earlier. The seed point is then chosen from these detected edge points using a measure incorporating aspects of edge strength and edge position (Figure 3(d)). After that, a region is grown from the seed point based on similarity with the region's mean intensity. In traditional region growing, the region is iteratively grown until the intensity difference between the region's mean and new neighbouring pixel is larger than a specific threshold. In this paper, we use a new termination criterion to efficiently avoid undersegmention of inhomogeneous regions. Region growing starts with a critical initial threshold of intensity difference, the threshold increases in the growing process. This process stops when the region is very close to the edges of the image (Figure 3(e)). We use linear smoothing to refine the pectoral muscle boundary, which accurately preserves the boundary feature (Figure 3(f)).

3 Experimental Results

Our method has been tested on over 240 mammograms from the EPIC (European Prospective Investigation on Cancer) mammogram database instead of the commonly used MIAS database, because it contains a large collection of sequential mammographic images. All mammograms were digitised at 8-bit resolution and the size is equal to 5671×3788 pixels.

To demonstrate the validity of our algorithm it has been tested on mammograms with different breast tissue densities: SCC (Six Class Categories) 1 to 6, a quantitative classification of mammographic densities introduced by Boyd *et al.* [10]. For evaluation the segmentation results were visually rated as four categories: accurate, nearly accurate, acceptable and unacceptable for application in CAD (Computer-Aided Diagnosis) systems. Accurate or nearly accurate was rated according to whether the segmentation result was matched to the real border exactly or nearly exactly. Otherwise, the result was rated as acceptable if minor pixels near the breast border were mis-segmented, because those pixels are not relevant and do not provide significant information for CAD purposes. Larger deviations were rated as unacceptable. For the breast background segmentation 66.5% are accurate, 25% are nearly accurate, 6.9% are acceptable, and 1.6% are unacceptable. For the pectoral muscle removal, we have obtained 62.5% accurate, 25.4% nearly accurate, 5.6% acceptable, and 6.5% unacceptable segmentations. Figure 4 shows representative segmentation results of mammograms with 6 different densities ranging from SCC 1 to 6, indicating that the methodology performs robustly with respect to density types.



(a) SCC1 (b) SCC2 (c) SCC3 (d) SCC4 (e) SCC5 (f) SCC6

Figure 4: Mammogram segmentation results covering SCC1 to SCC6.

In some cases (1.6% of the breast background segmentations and 6.5% of the pectoral muscle segmentations are classified as unacceptable) the method does not obtain what could be considered an acceptable segmentation. For the breast region segmentation, those are mainly related to the extremely low contrast between the breast tissue near the boundary and the background region resulting in an inaccurate binary mask. Furthermore, a significant amount of noise in the image leading to a poor placement of the initial seed point of contour growing and the non-uniform breast intensity distribution yield under-segmented results. For the removal of pectoral muscle, a layered pectoral muscle formed in the mammogram acquisition process or an underexposed area inside the pectoral muscle could produce strong edges which penalise the accurate selection of seed point of region growing and cause under-segmented results. Moreover fuzzy contrast between the muscle and the breast tissue leads to over-segmented results.

We compared the results with previous studies where similar visual evaluation criteria were used. Bick *et al.* [1] tested their algorithm on 740 digitised mammograms, and 97% of the segmentation results were visually rated as acceptable. Méndez *et al.* [2] tested their algorithm on 156 digitised mammograms, and segmentation results were deemed to be accurate or nearly accurate in 89% of the mammograms. The method presented by Chandrasekhar and Attikiouzel [3] was tested on all the images from the MIAS database, and it provided about 94% acceptable segmentation results. In the work presented by Ojala *et al.* [4] the percentages of acceptable and accurate or nearly accurate cases for the 20 test images were 90% and 55% respectively. Raba *et al.* [7] tested more than 320 images and obtained 98% nearly accurate segmentation results, and the muscle subtraction results were nearly accurate in 86% of all the extractions. The experimental results obtained by our method are 98.4% acceptable results and 91.5% nearly accurate results which include accurate results for the breast background segmentation. For the pectoral muscle segmentation, we obtain 93.5% acceptable results and 87.9% nearly accurate results.

Future work will focus on further evaluating our method using a larger number of mammograms from the EPIC database and additional databases, including full field digital mammograms and improving our method to resolve unacceptable segmentation cases. The existing problems we have considered are as follows: the binary mask plays an important role in later segmentation steps. However, it is obtained as an approximate segmentation step with reliance upon a simple global histogram thresholding. The weighting factors of contour growing measure are established empirically, further experiments should be carried out in order to estimate the influence of each factor. Some constraints such as size, direction and shape should be involved in region growing measure to regularise resulting regions.

4 Conclusions

An approach to segmentation of the breast region with pectoral muscle removal in mammograms has been proposed based on histogram thresholding, edge detection in scale space, contour growing, polynomial fitting and region growing. Initial segmentation results on more than 240 mammograms have been qualitatively evaluated and have shown that our method can robustly obtain an acceptable segmentation in 98.4% and 93.5% for breast-boundary and pectoral muscle separation in mammograms with different density types and preserve the tissue close to the breast skin line effectively.

References

[1] U. Bick and M. L. Giger. Automated segmentation of digitized mammograms. *Academic Radiology*, 2(1):1-9, 1995.

[2] A. J. Méndez *et al*. Automatic detection of breast border and nipple in digital mammograms. *Computer Methods and Programs in Biomedicine*, 49(3):253-262, 1996.

[3] R. Chandrasekhar and Y. Attikiouzel. Automatic Breast Border Segmentation by Background Modeling and Subtraction. In *IWDM*, pages 560-565, 2000.

[4] T. Ojala *et al*. Accurate segmentation of the breast region from digitized mammograms. *Computerized Medical Imaging and Graphics*, 25(1):47-59, 2001.

[5] M. A. Wirth and A. Stapinski. Segmentation of the breast region in mammograms using active contours. In *VCIP*, pages 1995-2006, 2003.

[6] H. E. Rickard *et al*. Self-organizing maps for masking mammography images. In *IEEE EMBS*, pages 302-305, 2003.

[7] D. Raba *et al*. Breast Segmentation with Pectoral Muscle Suppression on Digital Mammograms. In *PRIA*, pages 471-478, 2005.

[8] J. Canny. A computational approach to edge detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8(6):679-698, 1986.

[9] R. Martí *et al*. Breast Skin-line Segmentation Using Contour Growing. In *PRIA*, pages 564-571, 2007.

[10] N. Boyd *et al*. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian national breast screening study. *Journal of the National Cancer Institute*, 87(9):670-675, 1995.