Breast Skin-Line Segmentation Using Contour Growing

Robert Martí, Arnau Oliver, David Raba, and Jordi Freixenet

Computer Vision and Robotics Group, University of Girona
Av. Lluís Santaló 17071, Spain
{marly,aoliver,draba,jordif}@eia.udg.es
http://vicorob.udg.es

Abstract. This paper presents a novel methodology to obtain the breast skin line in mammographic images. The breast edge provides important information of the breast shape and deformation which is posteriorly used by other processing techniques, typically mammographic image registration and abnormality detection. The proposed methodology is based on applying edge detection algorithms and scale space concepts. The proposed method is a particular implementation (application focused) of a growing active contour with common considerations. Quantitative and qualitative evaluation is provided to show the validity of the approach.

1 Introduction

Breast cancer is one of the most devastating and deadly diseases in women [1]. X-ray mammography remains currently the most effective method for early signs of breast cancer. Although the estimation of the breast skin-line (the boundary between breast tissue in the mammogram and the background) has not received much attention in the field of mammographic image analysis, it should be regarded as an important initial step for achieving an specific task. This includes the delimitation of the region of interest for the detection of abnormalities (microcalcifications and/or masses) in Computer Aided Detection systems or the estimation of breast deformation for image registration. In addition, the removal of non interesting regions in images would also reduce image storage and transmission sizes.

Most of the methods found in the literature are based on combining histogram thresholding techniques (which provides a fair initial estimate of the breast area) with other more elaborated approaches. In contrast, the aim of this paper is to investigate the feasibility of applying an edge detection approach for extracting the breast skin-line. This is based on the combination of edge detection using scale-space representation and active contours concepts.

The paper is structured as follows. The following section explains in more detail the difficulty of extracting the breast skin-line, as well as describes typical approaches for such task, underlining the main key works. Subsequently, the proposed method is described in Sect. 3 and results are shown in Sect. 4. The paper ends with the conclusions and future work.

J. Martí et al. (Eds.): IbPRIA 2007, Part II, LNCS 4478, pp. 564–571, 2007.

2 Skin-Line Segmentation

Although skin-line segmentation might be naively regarded as a simple task, obtaining an accurate segmentation is not often straightforward. This is mainly due to the projective nature of the acquisition process. Even though compression is applied to the breast, the thickness of the breast is not constant and decreases along the skin boundary toward the nipple, this fact decreases film contrast in this area. Besides, an added difficulty is the amount of noise found in the background. This might be due to different reasons involving the acquisition (i.e. scattering) and to the digitalisation process. This latter factor is more pronounced in mammographic films which have been digitised using a film scanner compared to full field digital acquisition systems, in which the heterogenous nature of the background is less prononounced. However, the number of film mammograms being used for diagnosis is still important, as well as the contrast problem still persists in the digital world.

A common approach to breast skin-line segmentation is thresholding algorithms. Their main drawback is that they do not account for the non-homogeneity of the mammographic background and usually low contrast parts of the breast are being considered background. This is partially solved by using post-processing strategies. Thus, in the works of Sallam et al. [2] and Yin et al. [3], the post-processing operations include morphology and line smoothing. In contrast, in the work of Méndez et al. [4] a tracking of the boundary is done using gradient information. In addition, Wirth [5] used active contours and fuzzy classification. A kindly different approach was adopted by Chandrasekhar [6], who modelled the breast background using a polynomial form and, subsequently, subtracted this from the original image obtaining the breast profile. A deeper review of these techniques can be found in [7]. In 2006, Pan et al. [8] presented a novel approach based on incorporating phase, amplitude and orientation from multiscale analysis obtaining very successful results.

3 Method

The idea behind the proposed method is based on finding the skin-line by using a contour growing technique. The growing process is stated following similar concepts of attraction and regularisation found in active contours. The method starts by computing an scale space representation of the image in order to perform edge detection using different scales. Subsequently, an initial seed point lying in the skin-line contour is located based on a robustly estimation process. Using this seed point, a contour growing process starts based on enlarging and adapting a contour using different criteria. Basically, and following the simile of active contours or snakes we adapt the concept of attraction forces (which make the contour enter a region) and what we refer to regularisation forces which penalises rapid curvature and position changes. An overview of the method is illustrated in the Fig. 1.

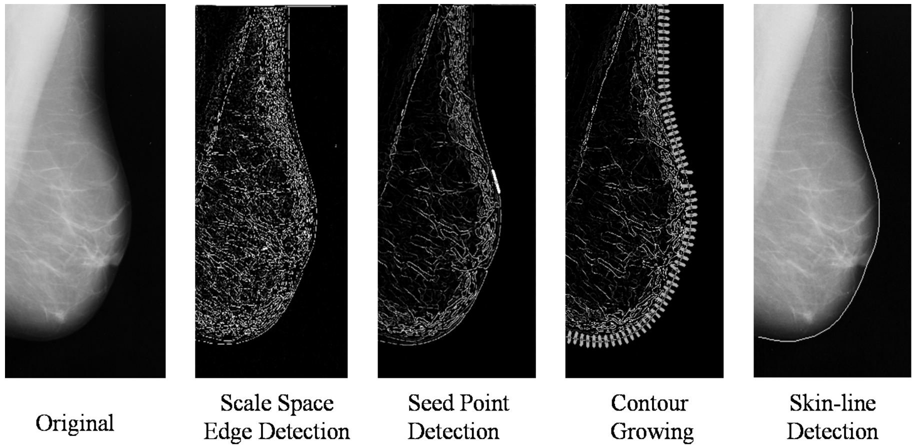


Fig. 1. Overview of the proposed skin-line segmentation method

3.1 Skin-Line Detection in Scale Space

The scale-space representation [9] describes an image as its decomposition at different scales. This is achieved by the convolution of the image with a Gaussian smoothing function at various scales (given by the σ value of the Gaussian function). This representation has been used in conjunction with edge detection in order to automatically extract edges at their optimum scale. If a small scale is used, the edge localisation is accurate but results are sensitive to noise. On the other hand, edges at larger scales have a better tolerance to noise but poor edge localisation. The motivation of using scale space edge detection is given by the nature of the breast skin-line: a low contrast edge often affected by noise. Various approaches to automatic scale selection have been proposed [9]. A simple and common approach is to select as the optimum scale the one which obtains a maximum response from scale invariant descriptors. This is in general given by normalised derivatives, for instance Lindeberg [9] defines

$$L_{norm} = \sigma^{\alpha/2}(L_x^2 + L_y^2) \tag{1}$$

as an edge strength measure for scale σ. L_x is the convolution of the image function with a first derivative Gaussian function. Here α is a parameter used as an additional degree of freedom for edge and ridge detection. A typical value of 1 is generally used in the definition of normalised derivatives for edge detection. Edge points are obtained detecting zero-crossing points of the second derivative in the scale-space representation. The final edge strength of a zero crossing will be given by the maximum normalised strength measure along the different scales. This maximum scale is regarded as the edge scale at that particular point.

3.2 Seed Point

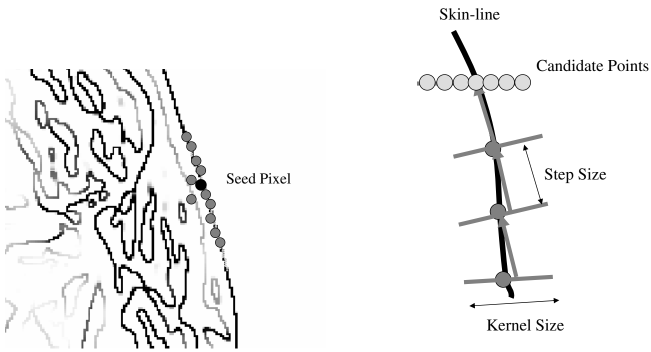
The first step of the method focuses on finding the starting point (or seed point) from which the contour will start growing. Special care has to be taken on estimating this point which directly affects the accuracy of the segmentation. As stated before, mammographic image segmentation presents difficulties mainly due to the low contrast in the skin-line and to the non-homogeneous background. From our experience this lower contrast is less severe for points close the nipple. Therefore a seed point can be easily detected in points at this area. An initial guess of a seed point is obtained as the first local maxima of the gradient in the scale space representation along the x axis at half the height of the image. Obviously, this first estimation lacks of robustness if this first local maxima does not correspond to the skin-line. That could be the case if the point lies inside the breast area (due to a low contrast of the skin-line) or in the background (due to noise, label and other image artifacts). A more robust approach is adopted based on analysing the position of various seed points at close the same position (at a small range in the y coordinate). The final seed point is obtained using a least median error estimation. Edge direction will also provide an important information in the contour growing process. Therefore the estimation of the initial angle it is also important. In this case a similar least median error estimation is adopted for the angle measure. Figure 2a shows an example of seed detection.

3.3 Contour Growing

Once the seed point has been obtained a contour growing process starts based on the combination of different criteria. For each point, a set of candidate growing points are obtained situated in a normal line along the gradient direction. A measure of affinity or cost C_i is computed for all points and the value with the minimum cost is taken as the next growing point. This iterative process is illustrated in the Fig. 2b.

As one may note from the figure, the growing scheme incorporates several parameters which need to be defined. These include a kernel size K, normal to the previous point, and a growing step S. Those values have been empirically determined (i.e. typical values are $K = 51$ and $S = 20$) and kept constant trough all the experiments.

Also from the experiments we noted that a more robust approach should be used for the process of selecting the next candidate point as it was often affected by noise and outliers. Instead of evaluating only a set of normal points at a given distance and obtain the candidate with a minimum cost over C_i, several sets of points on the normal are evaluated at different positions close to the desired position of the candidate point. A set of cost functions C_i^k is then obtained for each set of normal points. Using this approach the candidate point will be the one with the minimum cost over all the different sets C^k. One should note that in the different cost functions, the same (or nearly the same) point can lie in a shifted position. In order to make those cost functions comparable the cost functions are iteratively right and left shifted. The global minimum cost for each



(a) (b)

Fig. 2. Contour growing scheme: (a) Initial seed point and (b) contour growing process

point is obtained as the minimum using those shifted functions and the original cost. Figure 3 shows the minimum cost function of candidate points with and without cost shifting. Note that transportation effects have been minimised when costs are shifted allowing a better estimation of the minimum cost.

Candidate points are obtained from the zero crossing points along the normalised gradient using the scale space representation described earlier. The cost of choosing a candidate point i is given by a weighted function C_i, following the typical snake additive model formulation. This includes gradient, intensity, contour curvature and position information. Hence, the contour tends to grow finding areas of increasing intensity keeping minimal position and direction changes.

$$C_i = \alpha G_i + \beta D_i + (1 - \alpha - \beta) A_i \tag{2}$$

where G_i refers to an attraction factor (i.e. intensity or gradient), while the other two respond to regularisation terms penalising position differences (D_i) and direction changes (A_i). The factors α and β are scalar constants which will weight the importance of each term. As in many other approaches using weighted cost functions, it is important to obtain a good estimation of those factors in order to achieve a satisfactory segmentation. The selection of those factors will be later discussed in the paper (see the results section). Different attraction factors can be stated based on the represented information and how it is computed. Here two commonly used attraction factors are evaluated based on gradient and intensity information.

$$G_i = 1 - \exp(-1/f) \tag{3}$$

where f is the gradient or intensity image function, depending on the factor used. Gradient is obtained from the gradient of the zero crossing pixels while intensity information is given by the median intensity value in a local small window (i.e. $5x5$ pixels).

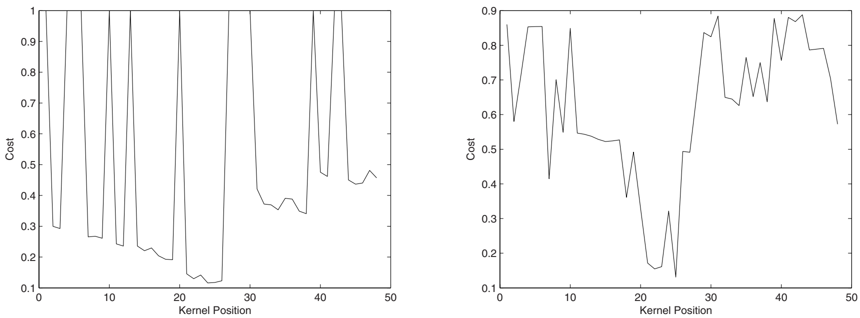


Fig. 3. Cost functions for robust candidate selection: (a) cost without shifting and (b) with cost shifting

The segmented breast skin-line should be continuous without having abrupt changes. This obviously corresponds to the continuous nature of the breast. A way to ensure this continuity is to impose some regularisation conditions to the contour growing process. This continuity assumption might not hold in all cases (i.e. when the nipple appears in the skin-line) but in this case the attraction factors described earlier will be able to adapt the contour to those changes. The first regularisation factor D_i biases the cost to points closer to the centre of the kernel of size K. This means that between two similar points the factor will select as a better point the one with a closer distance to the kernel centre. This factor is independent of the image contents and is given by,

$$D_i = \exp(-1/abs((i-1)-(K-1)/2)/((K-1)/2)) \tag{4}$$

The last regularisation term is defined computing the curvature change in a local neigbourhood. Curvature values at each pixel are obtained with a similar approach as used in [10]. Curvature (or directional change) between two pixels i and j is defined by the scalar product of their normal vectors. Hence, the curvature measure of a given pixel i is obtained by computing the scalar product between i and its neighbouring pixels,

$$A_i = \frac{1}{N}\sum_{j=1}^{N} \exp(-d_{ij}^2)(1-\cos(\phi_i - \phi_j)) \tag{5}$$

where ϕ_i is the angle of the normal at a pixel i. N is the number of points in a local neighbourhood and d_{ij} is the Euclidean distance between points i and j. The distance factor is used here to weight the curvature of each point j, in order to incorporate a bias to points closer to i.

4 Results and Discussion

In this section we show initial results obtained using the proposed skin-line segmentation algorithm. Evaluation has been carried out in several experiments using different mammographic databases: the MIAS [11] and the DDSM database [12]. A total of 65 images were segmented from the MIAS database and compared to manually segmented images, regarded here as ground truth. Similarly, 24 images were evaluated from the DDSM database. All images were randomly selected. Evaluation results have been computed using *completeness* and *correctness* measures for both databases. Those measures are related to the True Positives (TP), True Negatives (TN), False Negatives (FN) and False Positives (FP) values: $Completeness = TP/(TP + FN)$ and $Correctness = TP/(TP + FP)$. For the MIAS database we have obtained mean correctness and completeness values of 0.9697 (std: 0.0507) and 0.9547 (std: 0.0618), respectively. For the DDSM case the mean correctness and completeness values were 0.9524 (std: 0.0557) and 0.9744 (std: 0.0103), respectively. Special care has to be taken when looking at those values as they tend to be too optimistic. For instance, it is accepted that values over 0.95 can be considered as good segmentation but also results below 0.90 were often regarded in our experiments as unacceptable. Those results are slightly lower compared to other approaches [5] for the case of the MIAS database but also interestingly better for the DDSM, which has been generally perceived as more difficult to segment due to the larger amounts of noise. Moreover, one has to keep in mind that those are initial results and are likely to be improved in the future. In addition to the DDSM and MIAS database and although not quantitatively evaluated, we have tested the algorithm using full field digital mammograms from our local database. As expected, and due to the less noise found in the background, the segmentation results were all considered satisfactory.

In some cases the algorithm does not obtain what could be considered an acceptable segmentation. Those are mainly related to a large amount of noise in the image which lead to a poor estimate of the initial seed point and to non-uniform breast intensity distribution which yields undersegmented images. On the other hand, it is also important to notice that the performance of the algorithm does not substantially depend on the database used which usually has been reported with other approaches [5]. The weighting factors of the growing criteria described in the methodology section were established empirically, experiencing that extreme values of any of the factors did not obtain satisfactory results and that the attraction factor (intensity and gradient information) were the most important in order to reach more accurate segmentation. However, additional experiments will be carried out in order to asses the information apported by each factor.

5 Conclusions

A novel approach to the segmentation of the skin-line in digital mammograms has been presented based on a novel contour growing technique using scale-space edge detection and attraction and regularisation terms. Although we have presented initial evaluation results these have shown that our method can robustly obtain an accurate segmentation in most of cases using different databases. Future work will focus in further evaluating our method using a larger number of cases and additional databases, including full field digital mammograms. In addition, this evaluation will be compared to other recent approaches [8] for which we have been unable to include in this work.

Acknowledgments. Work partially supported by MEC grant nbs. TIN2005-08792-C03-01 and TIN2006-08035.

References

1. American Cancer Society: Breast cancer: facts and figures. 2003-04. ACS (2003)
2. Sallam, M., Bowyer, K.: Registration and difference analysis of corresponding mammogram images. Medical Image Analysis 3(2), 103–118 (1999)
3. Yin, F., Giger, M., Doi, K., Vyborny, C., Schmidt, R.: Computerized detection of masses in digital mammograms: automated alignment of breast images and its effect on bilateral-subtraction technique. Med. Phys. 21(3), 445–452 (1994)
4. Mendez, A., Tahoces, P., Lado, M., Souto, M., Correa, J., Vidal, J.: Automatic detection of breast border and nipple in digital mammograms. Comput. Methods Programs Biomed. 49(3), 253–262 (1998)
5. Wirth, M., Nikitenko, D., Lyon, J.: Segmentation of the Breast Region in Mammograms using a Rule-Based Fuzzy Reasoning Algorithm. ICGST International Journal on Graphics, Vision and Image Processing,vol. 05(2) (2005)
6. Chandrasekhar, R., Attikiouzel, Y.: Gross segmentation of mammograms using a polynomial model. In: Proc. Eng. Med. and Biol. Soc. 3, 1056–1058 (1996)
7. Raba, D., Oliver, A.: Breast segmentation with pectoral muscle suppression on digital mammograms. In: Iberian Conference on Pattern Recognition and Image Analysis, Estoril, Portugal, pp. 471–478 (2005)
8. Pan, X., Brady, M., Highnam, R., Declerck, J.: The use of multi-scale monogenic signal on structure orientation identification and segmentation. In: Astley, S.M., Brady, M., Rose, C., Zwiggelaar, R. (eds.) IWDM 2006. LNCS, vol. 4046, pp. 601–608. Springer, Heidelberg (2006)
9. Lindeberg, T.: Edge detection and ridge detection with automatic scale selection. International Journal of Computer Vision 30(2), 117–154 (1998)
10. Deschênes, J., Ziou, D.: Detection of line junctions and line terminations using curvilinear features. Pattern Recognition Letters 21, 637–649 (2000)
11. Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., Taylor, P., Betal, D., Savage, J.: The Mammographic Image Analysis Society digital mammogram database. In: International Workshop on Digital Mammography, pp. 211–221 (1994)
12. Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, P.: The digital database for screening mammography. In: International Workshop on Digital Mammography (2000)