Automatic Nipple Localisation using Local Curvature Modelling

Raquel Gil
rgilrodr7@gmail.com
Reyer Zwiggelaar
rrz@aber.ac.uk
Harry Strange
hgs08@aber.ac.uk

Dept. of Computer Science,
Aberystwyth University,
Aberystwyth, Wales, UK,
SY23 3DB

Abstract

The estimation of the location of the nipple in mammographic images forms an important step as a pre-processing stage, which is used as a landmark for registration. In addition, the location of the nipple could be used to divide mammographic images into regions which can be used by CAD systems and is linked to the visual quadrant assessment. We describe a novel approach to the detection the nipple in mammographic images. The developed approach incorporates the identification of the fatty region at the breast-background boundary and local curvature modelling of the detected region. The evaluation of the developed approach, based on 294 mammograms from the MIAS database, indicated that for about half the cases the nipple was detected within 5mm of the ground truth, while around 85% was within 15mm. These results are comparable with state of the art methods.

1 Introduction

Mammographic image analysis plays an important role in the early detection of breast cancer which is one of the leading causes of cancer, with studies indicating that 1.38 million women were diagnosed with the disease in 2008, accounting for nearly a quarter of all cancer cases worldwide [5]. Although incidence statistics remain high, mortality rates for breast cancer are dropping, thanks in part to improved early detection.

Numerous methods exist to help to automatically analyse mammographic images and so aid in the improvement of early detection rates. Often analysis is performed on a large volume of images taken under different physical imaging conditions and along with differences in breast anatomy between cases can lead to large variation in appearance. Registration is often used to minimise this variation and the location of the nipple can be used as a landmark for this registration process [4]. Further to this, the nipple position along with the pectoral muscle location can be used to split the mammogram into quadrants. This quadrant based analysis can be used by CAD systems and is linked to visual quadrant assessment.

There are various approaches in the literature to automatic nipple detection in mammograms. The earliest method estimated the nipple position by combining information on the maximum height of the breast border, maximum gradient, and maximum second derivative of gray-levels [9]. Other methods include those that seek to make use of the Radon transform [1, 8], genetic algorithms [7], or multi-step texture and rule based analysis [11] to identify the location of the nipple. The most recent state of the art method uses a simple heuristic to estimate the location of the nipple [6].

The proposed method locates the nipple by analysing the local curvature of the breast-background interface as found by using a fuzzy *c*-means clustering approach. This novel method is able to identify nipple location with an average distance from ground truth being estimated as 8.9mm which is comparable with the current state of the art.

The remainder of the paper is structured as follows. In Section 2, the proposed methodology for nipple location estimation based on local curvature analysis is presented. Section 3 presents the results of the proposed methodology using the MIAS database. Finally, conclusions are drawn in Section 4.

2 Methodology

The input images to the methodology described below are mammographic images with the breast boundary and pectoral line removed using the algorithm described in [3]. The pectoral line is estimated from the boundary segmentation. The proposed methodology for nipple detection uses five main steps to extract the estimated nipple position. The method is based on the intuition that the contours of the boundary of the breast within a restricted region of interest will reveal the position of the nipple. By modelling the local curvature of the breast boundary within a fixed window, the nipple location can be estimated.

2.1 Graylevel Segmentation

The first step is to cluster the input image using the fuzzy c-means algorithm [2] from which the boundary clusters can be determined. The fuzzy c-means algorithm is a fuzzification of the classic *k*-means algorithm that seeks to find a partitioning of the dataset by iteratively partitioning the dataset based on each data points distance to a set of *k*-centroids. Specifically, *k*-means seeks to minimise the squared error function

$$\Psi(\mathbf{Y}) = \sum_{j=1}^{k} \sum_{i=1}^{t} \parallel \mathbf{y}_i - \mathbf{c}_j \parallel_2^2 \tag{1}$$

where $\mathbf{Y}$ is the set of all data points, $\mathbf{y}_i$ is the i-th data point of $\mathbf{Y}$, and $\mathbf{c}_j$ is the j-th cluster centre. Fuzzy c-means adapts this clustering by allowing points to have a degree of membership to each cluster centre such that points on the edge of a cluster will belong to that cluster to a lesser degree than those nearer the centre of the cluster. Specifically, fuzzy c-means changes the computation of the centroid of each cluster to allow for degree of membership. In this experiment, $k = 9$ clusters are used based on trial experiments and previously published results [2].

2.2 Breast-Boundary Region Segmentation

The clusters corresponding to the boundary of the breast will contain the information required to localise the nipple position. As such, the second step is concerned with identifying those outermost clusters. Since the pectoral line is given, or in the case of craniocaudal mammograms the vertical line corresponding to the edge of the image by the thoracic wall, the clusters are sorted by maximum distance to the pectoral or thoracic line. The three clusters furthest away from the thorax are taken to be the boundary clusters.

2.3 Region of Interest Identification

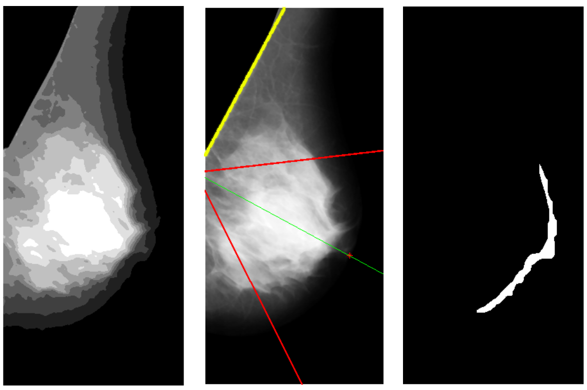
Once the breast boundary has been identified, a region of interest is defined within which the nipple is localised. This region of interest is found by calculating the distance between all points of the furthest cluster (as found in Step 2) and the pectoral line. Connecting the furtherst point to the pectoral line gives rise to a line perpendicular to the pectoral. Now, from this perpendicular line a subset of the final three boundary clusters is taken within the range $[-\alpha, \alpha]$, with $\alpha = 30°$ as shown in Figure 1 (b). All remaining operations will be performed on the cluster regions that fall within this region of interest.

2.4 Morphological Filtering

The three cluster contours found using Steps 2 and 3 are combined into a single contour within which the nipple can be localised by using morphological processing. The first step is to use morphological dilation with a disk structured element of size 21×21 to ensure that the three contours are combined into one (this can be checked by ensuring that the image has only a single connected component). This dilated region is then thinned to a line using morphological thinning. This step can be conceptually thought of as taking the average of the three cluster contours.

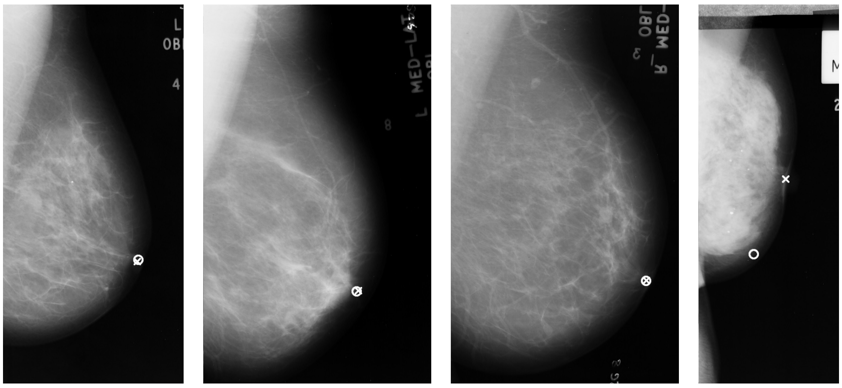
2.5 Nipple Localisation using Curvature Analysis

The final step is to search through the region identified in Step 4 to locate the nipple. This is done by initially fitting a circle through the curve created by the contour pixels. These



(a) (b) (c)

Figure 1: Figures showing the various steps of the algorithm. (a) Shows the results of performing fuzzy c-means clustering on an input mammogram. The clusters along the breast-background interface are used to identify the nipple location. (b) The nipple is searched for within a local region of interest. (c) The result of performing morphological filtering on a local region of the cluster contours on the breast-background interface.



(a) (b) (c) (d)

Figure 2: Results of performing the proposed nipple localisation method on four cases from the MIAS database. A circle indicates the estimated position of the nipple and a cross is the ground truth position. (a-c) show how the proposed method can perform well at identifying the nipple position. (d) shows a failure case where the nipple has not been correctly located.

pixels can be thought to correspond to a arc of a large circle, so a circle is fitted through these points. The point on the breast region contour (found using the previous step) that has the largest distance from the centre point of the circle is then taken to be the nipple location.

3 Results

A subset of 294 images from the MIAS database [10] are used in this work for experimentation. Images where the nipple fall outside of the image region (as is the case for some mammograms of large breasts) are excluded from this study. The proposed methodology was programmed in the MATLAB 2011 programming environment running on a Dell Optiplex 755 with 4GB RAM.

3.1 Qualitative Results

Figure 2 shows results of the proposed method on four cases, three of which show the method performing well (a-c) and one showing a case where the nipple is not identified correctly (d). In the successful cases the estimated nipple location (shown with a circle) closely matches that of the ground truth (shown with a cross). The failure case, shown in Figure 2 (d), occurs because the pectoral region was not correctly identified. As such, the nipple falls outside of the region of interest and so a phantom nipple location is found. This identifies one of the drawbacks of the proposed technique, if the pectoral region is not correctly identified then the nipple may not be correctly located.

3.2 Quantitative Results

To assess the quantitative performance of the proposed approach to nipple localisation, the estimated nipple positions are compared against ground truth positions. The Euclidean distance between the estimated and ground truth positions then gives rise to a measure of accu-

Distance	% Accuracy	Num. Images
0 to 5 mm	49.7%	146
5 to 10 mm	25.1%	74
10 to 15 mm	9.9%	29
15 to 20 mm	3.4%	10
20 to 25 mm	4.4%	13
more than 25 mm	7.5%	22

Table 1: Table of results of the distributions of images falling within specific distance thresholds. In the majority of cases the proposed method is able to detect the nipple within 10mm of the ground truth.

racy. To assess the accuracy of the proposed approach, the distances were sorted into 5mm bins such that each bin represents the percentage of images within the database where the accuracy of the proposed nipple location method falls within the given range.

The results are summarised in Table 1. The majority of nipples can be located within 10mm of the ground truth with accuracy improving such that around 85% can be estimated within the 15mm range and around 92% can be estimated within the 25mm range. Compared with the most recent state of the art method [6] the proposed approach is able to achieve a higher accuracy both overall and within the smaller ranges. The method in [6] achieves an accuracy of 80% within the 15mm range and only 30.4% are estimated correctly within the 5mm range.

The proposed approach is, on average, able to detect the nipple position within 8.9mm $\pm$11mm of the ground truth. This is an improvement over the method described in [6] where on average the nipple is estimated within 11.03mm $\pm$12.8mm.

Figure 3 shows a 2-dimensional histogram of estimated nipple positions with relation to the ground truth position. The distribution of estimated nipple positions is generally Gaussian with a tendency to under estimate the position along the horizontal axis. That is, the proposed approach will estimate the nipple as being closer to the chest wall than it actually is. The outliers in Figure 3 are generally due to the pectoral line being incorrectly estimated. If the pectoral line is incorrectly estimated then the local window within which the boundary curvature is estimated may not contain the nipple. As such, the nipple will be incorrectly located.

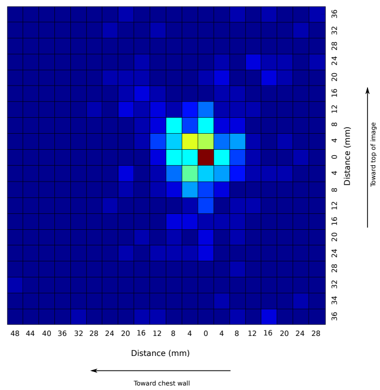


Figure 3: Distribution of the estimated nipple position relative to the ground truth.

4 Conclusions

This paper has presented a novel approach to nipple localisation on digital mammograms using local curvature analysis. The method utilises the contour of the clustered breast-background interface to identify the location of the nipple. Experiments on the MIAS database show that the proposed method is able to estimate the nipple within 8.9mm of the ground truth and around 85% of the cases can be estimated within 15mm of the ground truth.

One drawback of the proposed approach is its dependency on the accurate segmentation and identification of the pectoral line. If the pectoral line is incorrectly estimated then the local region within which the nipple is searched for may not contain the nipple. However, in the cases where the pectoral region is correctly identified, the proposed method is able to accurately identify the nipple location. Further experimentation on different databases could further reveal the effectiveness of the proposed approach. As well as this, using the estimated nipple positions as landmarks for registration of mammograms will help to assess how well the proposed method works as a pre-processing step to registration.

References

[1] J Chakraborty, S Mukhopadhyay, R M Rangayyan, A Sadhu, and P M Azevedo-Marques. Automatic localization of the nipple in mammograms using gabor filters and the radon transform. In *Proceedings of SPIE 8670 Medical Imaging*, 2013.

[2] Z Chen, E R Denton, and R Zwiggelaar. Local feature based mammographic tissue pattern modelling and breast density classification. In *BMEI*, 2011.

[3] Zhili Chen and Reyer Zwiggelaar. A combined method for automatic identification of the breast boundary in mammograms. In *Proceedings of the 5th International Conference on BioMedical Engineering and Informatics*, pages 123–127, 2012.

[4] S Engeland, P Snoeren, J Henriks, and N Karssemeijer. A comparison of methods for mammographic registration. *IEEE Transactions on Medical Imaging*, 22(11):1436–1444, 2003.

[5] J Ferlay, H R Shin, F Bray, D Forman, C Mathers, and D M Parkin. GLOBOCAN 2008 v1.2, cancer incidence and mortality worldwide: IARC CancerBase No. 10. (http://globocan.iarc.fr accessed November 2012).

[6] M Jas, S Mukhopadhyay, J Chakraborty, A Sadhu, and N Khandelwal. A heuristic approach to automated nipple detection in digital mammograms. *Journal of Digital Imaging*, 2013.

[7] M Karnan and K Thangavel. Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications. *Computer Methods and Programs in Biomedicine*, 87:12–20, 2007.

[8] S K Kinoshita, P M Azevedo-Marques, R R Pereira, J A H Rodrigues, and R M Rangayyan. Radon-domain detection of the nipple and the pectoral muscle in mammograms. *Journal of Digital Imaging*, 21(1):37–49, 2008.

[9] A J Méndez, P G Tahocesb, M J Ladoa, M Soutoa, J L Correab, and J J Vidala. Automatic detection of breast border and nipple in digital mammograms. *Computer Methods and Programs in Biomedicine*, pages 253–262, 1996.

[10] J. Suckling, Parker J, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage. The mammographic images analysis society digital mammogram database. In *Digital Mammography*, pages 375–378, 1994.

[11] C Zhou, H P Chan, C Paramgul, M A Roubidoux, B Sahiner, L M Hadjiiski, and N Petrick. Computerized nipple identification for multiple image analysis in computer-aided diagnosis. *Medical Physics*, 31(10):2871–2882, 2004.

An Efficient Gland Detection Method Based on Texture and Morphological Transformation

Jie Shu[1]
psxjs3@nottingham.ac.uk
Hao Fu[1]
Hao.Fu@nottingham.ac.uk
Guoping Qiu[1]
Guoping.Qiu@nottingham.ac.uk
Mohammad Ilyas[2]
Mohammad.Ilyas@nottingham.ac.uk

[1] Computer Science
University of Nottingham
Nottingham, UK

[2] Queens Medical Center NHS Trust
University of Nottingham
Nottingham, UK

Abstract

The diagnosis of digitized tissue specimen can be improved by the computerized image analysis method. In this paper, we present a method of automatically detecting the gland structures in H-DAB stained images of colon cancer. This is a robust method which detect the glands by utilizing the texture information and morphological characteristics. We also introduce another visual feature based regressor to verify if the predicted gland is true or not. Experiments on a publicly available dataset show that our approach outperforms the state-of-the-art.

1 Introduction

Histopathology is a study of the expression of disease through the microscopic examination on a stained tissue specimen or biopsy by pathologist. Gland is an important structure in the specimen which contains essential information in the disease detection especially in the cancer detection. The diagnosis of the specimen is mainly to explore the changes of gland architecture and the distribution of cancerous nuclei in gland.

Most of the previous papers segment gland in Hematoxylin and Eosin (H&E) stained image. In H&E stained image, the components of gland are colored with distinguishable colors, and the stroma which is surrounding the gland are colored differently from the others. Several papers [4, 5] consider the color to be the distinctive cue for segmenting glands. Unfortunately, this is not suitable for the Hematoxylin-Diaminobenzidine H-DAB stained tissue image, as it only colors the nuclei into two classes, the cancerous nuclei and the normal nuclei. Therefore, the previous color dependant gland detection methods might not work on this kind of stained images.

Accordingly, texture or feature based methods are more flexible to be applied on different kinds of color stained tissue images. These methods [1, 2, 7] directly performed on the