

Automatic diagnosis of masses by using level set segmentation and shape description

Arnau Oliver, Albert Torrent, Xavier Lladó & Joan Martí
 Dept. of Computer Architecture and Technology
 University of Girona
 Girona, Catalonia (Spain)
 email: {aoliver,atorrent,llado,joanm}@eia.udg.edu

Abstract—We present here an approach for automatic mass diagnosis in mammographic images. Our strategy contains three main steps. Firstly, region of interests containing mass and background are segmented using a level set algorithm based on region information. Secondly, the characterisation of each segmented mass is obtained using the Zernike moments for modeling its shape. The final step is the diagnosis of masses as benign or malignant lesions, which is done using the Gentleboost algorithm that also assigns a likelihood value to the final result. The experimental evaluation, performed using two different digitised databases and Receiver Operating Characteristics (ROC) analysis, proves the feasibility of our proposal, showing the benefits of a correct shape description for improving automatic mass diagnosis.

Keywords—Mass Diagnosis; Mammography; Shape; Zernike Moments;

I. INTRODUCTION

Breast cancer is still a significant health problem in the world. It is estimated that between one in eight and one in twelve women will develop breast cancer during their lifetime [1]. Up to date, there is no guaranteed way to prevent breast cancer, and efforts are focused on the detection of abnormalities at an early stage that is fundamental for improving survival rates. Mammography is still the most effective and reliable method for an early breast cancer detection, since it allows the detection of subtle scale signs such as masses.

Radiologists tend to describe a mass according to its shape, margin characteristics, optical density, and other associated findings. The mass shape is one of the most important features to diagnose a mass as being benign or malignant. Normally, a benign process is associated with the presence of circular or oval shapes, while in contrast spiculated masses are more probable to be the sign of a malignant process. Figure 1 shows three benign and three malignant masses (the two first columns show typical cases of benign and malignant processes while the last one shows atypical cases).

The diagnosis of masses is a challenging task, even for an expert radiologist, since the mass morphological aspects can be very subtle and difficult to diagnose. In fact, it is reported that less than 30% of all breast biopsies show a malignant

pathology [2], [3]. Hence, computer-aided approaches are being proposed to support radiologists in this discrimination of benign and malignant lesions [4]. These approaches are usually structured in three steps: segmentation of the region of interest (ROI – a crop of the mammogram containing the lesion), characterisation of the segmented masses, and a posterior classification.

According to these steps, our approach for automatic mass diagnosis can be described as follows: the segmentation step is based on a region based level set segmentation, recently used in mass segmentation by Yuan et al. [5]. The characterisation step presents the main novelty of our paper since, to our knowledge, we present the first attempt to use the Zernike moments [6] for describing the mass shape. The final classification step is performed in terms of the Gentleboost algorithm [7].

The rest of this paper is structured as follows. The following three sections introduce the level set segmentation, the Zernike moments, and the Gentleboost algorithm, respectively. Experimental results performed in two different databases are presented in Section V. The paper finishes with the conclusions and further work.

II. ROI SEGMENTATION USING A LEVEL SETS APPROACH

The first step of our approach for mass diagnosis is the segmentation of a mass. As already mentioned, we use a geometric active contour based approach for this task. However, note in Figure 1 that the segmentation of the mass contour is not a trivial task, even for expert radiologists. Notice also that to perform the diagnosis we do not use all the image but ROIs.

In this paper we use the active contour approach of Chan and Vese [8] which relies on intrinsic properties of the region to be segmented, instead of using gradient information. Hence, the energy function modelling the evolving contour C is mathematically defined as

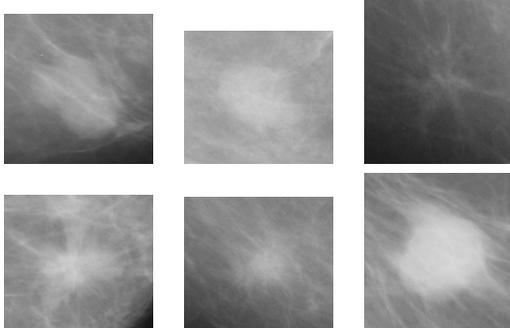


Figure 1. The upper row shows three mass examples being benign masses, while the bottom row shows three mass examples being malignant masses. Note that although round masses without spicules tend to be benign process, they can also be malignant process (last figure of bottom row). On the other hand, the presence of spicules does not necessarily be associated with malignant process (last figure of upper row).

$$\begin{aligned}
E(c_1, c_2, C) &= \mu \text{length}(C) + \\
&+ \lambda_1 \int_{in(C)} |f(x, y) - c_1|^2 dx dy + \\
&+ \lambda_2 \int_{out(C)} |f(x, y) - c_2|^2 dx dy \quad (1)
\end{aligned}$$

where μ , λ_1 , and λ_2 are fixed weight parameters. The first summand prevents the contour from converging to a small area, while the second and the third terms measure the homogeneity of the inner (with mean c_1) and outer (with mean c_2) regions of the image $f(x, y)$, respectively. This equation is solved by level set theory, where the 2D contour is represented as the zero level set of the 3D function ϕ .

In order to avoid the intrinsic re-initialising term of level sets, Li et al [9] proposed to add the following term in the model:

$$\nu \frac{1}{2} \int_{\Omega} (1 - \|\nabla \phi_t\|)^2 dx dy \quad (2)$$

where ν is another weighting parameter and the integration is over the whole image space (Ω). Therefore, as shown in the work of Yuan et al. [5], the differential equation that defines the model is:

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= \delta_{\epsilon} [\mu \kappa - \lambda_1 (f(x, y) - c_1)^2 + \lambda_2 (f(x, y) - c_2)^2] \\
&+ \nu \text{div} \left[\left(1 - \frac{1}{\|\nabla \phi\|} \right) \nabla \phi \right] \quad (3)
\end{aligned}$$

where δ_{ϵ} is the Dirac measure and κ the curvature of the contour that incorporates also the regularisation term shown in Eq. (2).

After the segmentation, the ROI may contain several disconnected objects. Hence, for obtaining a single object mass we selected the largest one using the connected component labeling (CCL) algorithm. Moreover we apply

a *close* morphologic operation to obtain a smooth contour. Note that we used a small round morphologic structural element (size 3) to avoid removing possible spiculation.

III. ROI DESCRIPTION USING THE ZERNIKE MOMENTS

Once the mass is segmented from the background, we describe its shape using the Zernike moments, which have been previously used as object descriptors in several pattern recognition and image retrieval applications with significant results [6]. The Zernike moments are a set of descriptors obtained using complex kernel functions based on Zernike polynomials, which are defined with an order p and a repetition q (being $|q| \leq p$, $|p - q|$ *even*) as:

$$Z_{pq} = \frac{p+1}{\pi} \int \int_{x^2+y^2 \leq 1} V_{pq}^*(x, y) g(x, y) dx dy \quad (4)$$

where $g(x, y)$ represents the analysing (segmented) mass shape and V^* is the complex conjugate of function V usually defined in polar coordinates as:

$$V_{p,q}(\rho, \theta) = \sum_{\substack{k=|q| \\ |p-k| \text{ even}}}^p \frac{(-1)^{\frac{p-k}{2}} \frac{p+k}{2}!}{\frac{p-k}{2}! \frac{k-q}{2}! \frac{k+q}{2}!} \rho^k e^{iq\theta} \quad (5)$$

The main characteristics of the Zernike moments are their ability to describe a shape with minimum information redundancy (due to their orthogonality property), the robustness in noisy environments, and the fact that they are invariant to an arbitrary rotation and multi-level representation of the describing shape. In this work we extracted the 13 first order moments, which resulted in a set of 49 features per mass.

IV. ROI CLASSIFICATION BY MEANS OF THE GENTLEBOOST ALGORITHM

Once the segmented masses are translated to feature vectors, the learning step is performed using the Gentleboost classifier. Boosting algorithms are based on the simple idea that the sum of weak classifiers can produce a strong classifier. In the Gentleboost algorithm [7] the weak classifiers (h_t) are simple regression stumps with one of the features (x), so at each round t the feature with less error is selected:

$$h_t(x) = a \delta(x_i > th) + b \quad (6)$$

where th is a threshold determining if pattern x belongs to the object class, x_i is the i 'th dimension of x , and a and b are parameters selected to minimise the error of the classifier (a is the regression slope and b the offset):

$$e = \sum (w (y - (a (x_i > th) + b))^2) \quad (7)$$

At each round the training data weights (w) are updated, increasing in the following round the possibility of classifying correctly the previous incorrectly classified points. In the GentleBoost algorithm the data weights are updated using:

$$w_{t+1} = w_t e^{y \cdot h_t(x)} \quad (8)$$

When testing a new data, the final classifier is computed using the weak classifier created at each round of the boosting. Therefore, the testing data is classified according to the sign of the sum of weak classifiers:

$$H(x) = \sum h(x) \quad (9)$$

The absolute value of $H(x)$ shows the confidence of the classified data.

V. RESULTS

The described approach has been evaluated using two different digitised database: the MIAS and the DDSM. We extracted all the diagnosed masses of the MIAS database [10], which constituted a set of 57 ROIs, 37 diagnosed as benign masses, while the 20 remaining were malignant. The spatial resolution of the images was $50\mu\text{m} \times 50\mu\text{m}$ and the optical density was linear in the range $0 - 3.2$ and quantised to 8 bits. On the other hand, we used a large subset of 818 ROIs extracted from the DDSM database [11], 397 diagnosed as benign and 421 as malignant. The mammograms of the DDSM were digitised using different mammograms: a DBA M2100 ImageClear ($42 \times 42 \mu\text{m}$ pixel resolution), a Howtek 960 ($43.5 \times 43.5 \mu\text{m}$ pixel resolution), a Lumisys 200 Laser ($50 \times 50 \mu\text{m}$ pixel resolution), and a Howtek MultiRad850 ($43.5 \times 43.5 \mu\text{m}$ pixel resolution). All the images were 12 bits per pixel. In both databases, the ROIs were manually extracted from the respective annotations, hence the mass was always centred at the ROI (see Figure 1 for examples). Note that in the case the mass is not centred but fully included in the ROI the level set segmentation approach is also able to provide good results.

The evaluation of our experiments was done by using a cross-validation scheme and Receiver Operating Characteristics (ROC) analysis [12]. In a N -folder cross-validation all the images are divided into N different groups, from where $N - 1$ is used to train the classifier, while the remaining group is used for testing. This procedure is repeated N times until all the groups are used for testing. Note that using this methodology each image appears in the test set only once. In detail, we used $N = 10$ for the DDSM database and $N = 57$ for the MIAS one, which is the degenerative case where all the images except the testing one are used in the training (leave-one-out methodology). Note that this is necessary due to the small number of ROIs available.

The evaluation of the results is performed in terms of ROC analysis, widely used in the medical field. Note that the used classifier provides a numerical value related to the

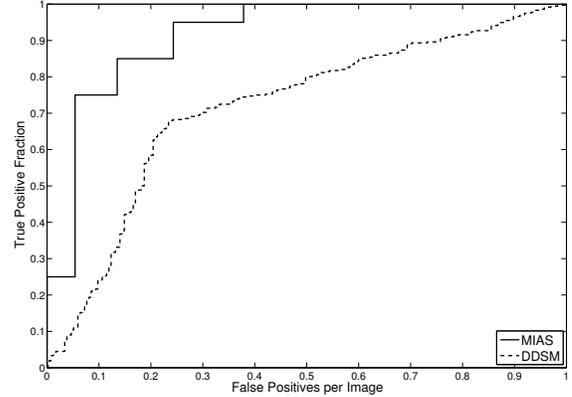


Figure 2. ROC analysis of our approach. We obtained better results when testing the MIAS database (solid line) than testing the DDSM one (dotted line).

membership of each class. Therefore, varying the threshold of this membership it is possible to generate a ROC curve, representing the true positive rate (the number of malignant ROIs detected divided by the total number of malignant ROIs) as a function of the false positive rate (the number of malignant ROIs missed divided by the overall number of benign ROIs). The area under the curve (known as A_z) is an indication for the overall performance of the observer, and is typically used to analyse the performance of the algorithms.

Figure 2 shows the ROC curve of the proposed approach evaluated in both databases. For the MIAS database we obtained $A_z = 0.89$, while for the DDSM we had an $A_z = 0.72$. For instance, at a false positive rate of 0.24, 95% of the malignant cases were correctly classified for the MIAS database. In contrast, at the same false positive rate, the sensitivity for the DDSM database was 69%. As expected, better results were obtained when using the MIAS database, since fewer cases were used in this experiment and also due to the fact that the DDSM contains a larger mass variability. However, we would like to point out that the results were obtained using only shape information alone, and in some cases, this information is not enough to correctly diagnose the case (see last column of Figure 1).

Table I shows a comparison between different approaches for mass diagnosis. However, we want to clarify that not all the methods used the same databases and number of ROIs and therefore our aim is only to provide a general trend of the performance of our approach with respect to other strategies. Note that the results obtained using the MIAS database outperformed most of the approaches, although our results were obtained using less number of cases. In contrast, the results obtained using the DDSM database were lower than those shown in the table. Note, however, that we used a larger database (except for the work of Varela et al. [13]). Moreover, we want to stress that we

Table I
COMPARISON OF MASS DIAGNOSIS APPROACHES, DETAILING THE
NUMBER OF CASES USED AND THE OBTAINED RESULTS.

Authors	Cases	Az
Guliatto et al. (2008) [15]	111	0.94
Delogu et al. (2007) [16]	226	0.78
Varela et al. (2006) [13]	1076	0.81
Lim and Er (2004) [17]	343	0.87
Sahiner et al. (2001) [14]	249	0.87
Mudigonda et al. (2000) [18]	39	0.85
Our approach – MIAS	57	0.89
Our approach – DDSM	818	0.72

only used shape information to characterise the masses, while the other works combine several feature descriptors.

VI. CONCLUSION

A new proposal for mass diagnosis have been presented in this paper. Our approach starts segmenting a ROI by using a level set formulation based on the internal properties of the regions instead of using gradient information. Once the mass is segmented, we used the Zernike moments for characterising the shape, and the Gentleboost algorithm for classifying the mass as benign or malignant. The experimental results which have been obtained using two different databases show the validity of our approach, proving the benefits of obtaining a good shape description.

Further work is directed to the use of other features that radiologists take into account when diagnosing a case, like the mass optical density or features related to the abruptness of the margin [14], [13]. Moreover, we want to test our approach using a digital database, since it is well known that this technology improves the contrast between the different internal structures.

ACKNOWLEDGEMENT

This work was supported in part by Ministerio de Educación y Ciencia of Spain under Grant TIN2007-60553 and by the CIRIT and CUR of DIUIE of Generalitat de Catalunya under Grant 2008SALUT00029.

REFERENCES

- [1] American Cancer Society, "Breast cancer: facts and figures. 2007-08," *ACS*, 2007.
- [2] W. E. Barlow, C. Chi, P. A. Carney, S. H. Taplin, C. D'Orsi, G. Cutter, R. E. Hendrick, and J. G. Elmore, "Accuracy of screening mammography interpretation by characteristics of radiologists," *J. Natl. Cancer Inst.*, vol. 96, no. 24, pp. 1840–1850, 2004.
- [3] K. Ghosh, L. J. Melton, V. J. Suman, C. S. Grant, S. Sterioff, K. R. Brandt, C. Branch, T. A. Sellers, and L. C. Hartmann, "Breast biopsy utilization: A population-based study," *Arch. Intern. Med.*, vol. 165, no. 14, pp. 1593–1598, 2005.
- [4] M. Elter and A. Horsch, "CADx of mammographic masses and clustered microcalcifications: a review," *Med. Phys.*, vol. 36, no. 6, pp. 2052–2068, 2009.
- [5] Y. Yuan, M. L. Giger, H. Li, K. Suzuki, and C. Sennett, "A dual-stage method for lesion segmentation on digital mammograms," *Med. Phys.*, vol. 34, no. 11, pp. 4180–4193, 2007.
- [6] A. Khotanzad and Y. H. Hongs, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 5, pp. 489–497, 1990.
- [7] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *Ann. Stat.*, vol. 38, no. 2, pp. 337–374, 2000.
- [8] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [9] C. Li, C. Xu, C. Gui, and M. D. Fox, "Level set evolution without re-initialization: a new variational formulation," in *IEEE Conf. Comput. Vision Patt. Rec.*, vol. 1, 2005, pp. 430–430.
- [10] J. Suckling, J. Parker, D. R. Dance, S. M. Astley, I. Hutt, C. R. M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. L. Kok, P. Taylor, D. Betal, and J. Savage, "The Mammographic Image Analysis Society digital mammogram database," in *Int. Work. Dig. Mammography*, 1994, pp. 211–221.
- [11] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. J. Kegelmeyer, "The Digital Database for Screening Mammography," in *Int. Work. Dig. Mammography*, 2000, pp. 212–218.
- [12] C. E. Metz, "Evaluation of digital mammography by ROC analysis," in *Int. Work. Dig. Mammography*, 1996, pp. 61–68.
- [13] C. Varela, S. Timp, and N. Karssemeijer, "Use of border information in the classification of mammographic masses," *Phys. Med. Biol.*, vol. 51, no. 2, pp. 425–441, 2006.
- [14] B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and L. M. Hadjiiski, "Improvement of mammographic mass characterization using spiculation measures and morphological features," *Med. Phys.*, vol. 28, no. 7, pp. 1455–1465, 2001.
- [15] D. Guliatto, R. M. Rangayyan, J. D. Carvalho, and S. A. Santiago, "Polygonal modeling of contours of breast tumors with the preservation of spicules," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 14–20, 2008.
- [16] P. Delogu, M. E. Fantacci, P. Kasae, and A. Retico, "Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier," *Comput. Biol. Med.*, vol. 37, no. 10, pp. 1479–1491, 2007.
- [17] W. K. Lim and M. J. Er, "Classification of mammographic masses using generalized dynamic fuzzy neural networks," *Med. Phys.*, vol. 5, no. 31, pp. 1288–1295, 2004.
- [18] N. R. Mudigonda, R. M. Rangayyan, and J. E. Desautels, "Gradient and texture analysis for the classification of mammographic masses," *IEEE Trans. Med. Imag.*, vol. 10, no. 19, pp. 1032–1043, 2000.