Advances in Computational & Experimental Engineering & Science

A Microcalcification Detection System for Digital Mammography using the Contourlet Transform

J. A. Manzano-Lizcano[1], C. Sánchez-Ávila[2], L. Moyano-Pérez[3]

Summary

Mammograms depict most of the significant changes in breast disease. In this paper, a computer aided diagnosis (CAD) system is presented. First, the mammographic image is enhanced using an orientation space analysis based on a contourlet transform. Then a multirresolution analysis based on the dyadic wavelet transform and wavelet transform modulus maxima is computed, and the resulting image is classified by a support vector machine (SVM). Performance results are presented.

Introduction

Breast cancer incidence varies across countries, but in most cases this type of cancer is the second cause of mortality for the female population [1]. The most efficient technology for the early detection of breast cancer is the mammography analysis, which outperforms auto-exploration and manual exploration by the specialist. It is estimated that the actual prevention by screening programs fails to detect around 25% of the cancers that are visible in a retrospective analysis [2]. Screening programs have drawbacks like the high cost, inexpert radiologists and visual fatigue.

There is a big research effort in identifying calcifications and masses in the mammography images, apart of other possible abnormalities, due to the fact that these two types of objects are the best indicators of a possible early stage of breast cancer. Calcifications are small calcium structures showed in the images as more brilliant structures that the surrounding tissue. Their spatial density and structure of calcifications is useful in categorizing them as either punctuate, branching, linear, spherical, fine, coarse, cylindrical, smooth, jagged, and regular in shape and size or heterogeneous [3]. In digital mammography the smallest calcifications are between 0.1 and 0.2 millimetres and we will call them *microcalcifications.* They appear isolated or as part of a group. This cumulus are usually formed by five or more microcalcifications in a volume of 1 cm^3, each of them no more of 0.5 mm and separated from the rest less than 5 mm. Detection is difficult due to the noise level of the images, being our work their identification and classification.

Detection of a breast tumour allow specialist to classify it in benign or malignant, but it is more difficult to distinguish between a benign or a malign microcalcification, so a lot

[1] Telefónica de España, Madrid, Spain. joseavelino.manzanolizcano@telefonica.es
[2] ETSI de Telecomunicación. Universidad Politécnica de Madrid. Spain. csa@mat.upm.es
[3] Guidant España, lmoyano@eresmas.com.

Proceedings of the 2004 International Conference on Computational & Experimental Engineering & Science 26-29 July, 2004, Madeira, Portugal

Advances in Computational & Experimental Engineering & Science

of unnecessary biopsies are made in screening programs due to their high false positive rate. In our opinion the critics [4], [5] are due to this fact, pointing out also the risks of inadequate treatments in case of false positives including the amount of psychological stress to the patient. A system that can help to reduce false positives and simultaneously giving a high rate of malignant microcalcifications detection will be very useful to improve actual screening programs at a reasonable cost. In fact, the specialist faces difficulties to make a diagnostic from the mammography and 10% to 30% of lesions are not detected in the first review [6]. Computer assisted diagnosis (CAD) systems can be very useful to the doctors in detecting subtle microcalcifications inside mammographic images. Some of these systems have the objective of improving their visibility to help the radiologist, and others make an automatic diagnostic to be used as a “second opinion”. This later is the case of our proposed system.

General description of the proposed system

The developed CAD system for the automated detection and classification of microcalcifications in digitalized mammographic images is based on multiscale differential operators. Is a second opinion aid system for diagnostic, with a first image preprocessing stage, followed by a microcalcification detection stage and a third classification-diagnosis stage of the mammographic image in two categories, benign or malignant. The pre-processing stage uses contourlet transform analysis. Detection stage works with dyadic wavelet transform and uses modulus maxima for contour detection. The wavelet function used is B-spline. Finally, classification assigns a diagnosis to the mammographic image, based in a set of morphological characteristics of the binary regions formerly detected that are fed to a SVM (support vector machine) classificator.

Preprocessing stage

This stage is divided in two blocks. In the first pre-processing block a set of operations is applied to the image in order to prepare it for the detection stage. These operations are performed to enhance the borders of the microcalcifications and reduce the high frequency noise. The first one is histogram equalization, to improve the contrast of the image, especially in high levels of grey. Then, a texture pre-classification in several categories allow us to use different noise reduction thresholds in a contourlet transform denoising stage. Contourlet transform was developed [7] as a sparse efficient decomposition for two-dimensional signals that are piecewise smooth away from sooth contours. In our problem we consider microcalcifications as objects defined by discontinuities or edges along one-dimensional smooth contour. Two-dimensional wavelets are good at catching point discontinuities or zero-dimensional discontinuities, but contourlets involve basis functions that are oriented at any power of two’s number of directions with flexible aspect ratios [8]. As we can see in the next figure, contourlets can represent a contour with much fewer coefficients than wavelets. Similar to wavelet-based models, contourlet-based models take into account scale and space, but as a true two-

Advances in Computational & Experimental Engineering & Science
Copyright 2004 Tech Science Press

dimensional representation contourlets allow us to also model the coefficient's dependencies across directions or orientations.

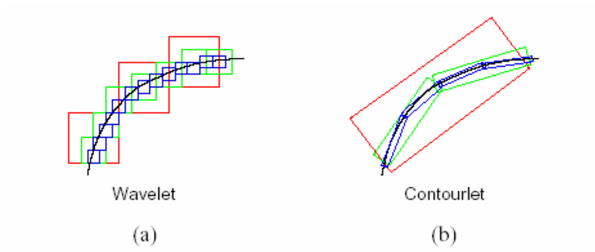

Figure 1. Wavelet and contourlet representation for images.

Contourlets are implemented by the pyramidal directional filter bank (PDFB) which decomposes images into directional subbands at multiple scales. In terms of structure the contourlet transform is a cascade of a Laplacian Pyramid [9] and a directional filter bank.

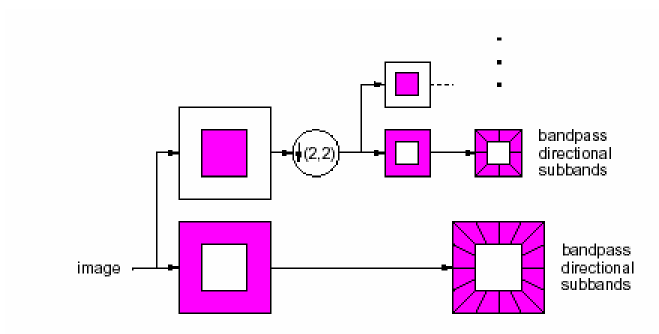

Figure 2. Pyramidal directional filter bank.

In our implementation, the second pre-processing block uses contourlet transform to reduce image noise and enhance the oriented features that determine the structure of the objects in the image. After carrying the direct contourlet transform, the coefficients in the transformed domain above a threshold are retained. The noise threshold depends of the variance of the input image. Then, an inverse contourlet transform reconstruct the approximation image. Due to non-oriented characteristic of the noise, the reconstructed image shows better signal to noise ratio.

Detection stage

Differential operators have been widely used for the geometric description of the images [10]. After the pre-processing stage we can use this representation for the border extraction of the microcalcifications and creation of a binary image that can be classified. In this way, we consider the B-spline wavelet of order n

Advances in Computational & Experimental Engineering & Science

$$\psi^n(x) = 2\beta^n(2x+1) - 2\beta^n(2x) + \beta^n(2x-1) \tag{1}$$

and using the relations

$$\frac{1}{2}\beta^n\left(\frac{x}{2}\right) = \sum_{k=-\infty}^{\infty} h_k \beta^n(x-k) \quad \text{and} \quad \frac{1}{2}\psi^n\left(\frac{x}{2}\right) = \sum_{k=-\infty}^{\infty} g_k \beta^n\left(x+\frac{1}{2}-k\right) \tag{2}$$

being

$$g_0 = 1,\ g_1 = -1,\ g_k = 0, \qquad k \neq 0, 1 \tag{3}$$

$$g_{-1} = 1,\ g_0 = -2,\ g_1 = 0\ g_k = 0, \qquad k \neq -1, 0, 1 \tag{4}$$

$$h_k = \begin{cases} \dfrac{1}{2^{n+1}}\begin{pmatrix} n+1 \\ \dfrac{n+1}{2}+k \end{pmatrix}, |k| \leq \dfrac{n+1}{2} \\ 0, \text{in other case} \end{cases} \tag{5}$$

we can write, from the pre-processed image $I_p\ (x, y)$ the approximation an detail images of I_p for $j = 1, 2, \ldots, J$ using the recursive algorithm

$$\begin{aligned} S_{2^j} I_p &= S_{2^{j-1}} I_p * (h, h)_{\uparrow 2^{j-1}} \\ W^1_{2^j} I_p &= S_{2^{j-1}} I_p * \left(g^{(2)}, d\right)_{\uparrow 2^{j-1}} \\ W^2_{2^j} I_p &= S_{2^{j-1}} I_p * \left(d, g^{(2)}\right)_{\uparrow 2^{j-1}} \\ W^3_{2^j} I_p &= S_{2^{j-1}} I_p * \left(g^{(1)}, g^{(2)}\right)_{\uparrow 2^{j-1}} \end{aligned} \tag{7}$$

where $S_{2^{j-1}} I_p = S_{2^{j-1}} I_p * (h, g)_{\uparrow 2^{j-1}}$ denotes separable convolution of rows and columns of the image I_p with the filters $h_{\uparrow 2^{j-1}}$ and $g_{\uparrow 2^{j-1}}$ respectively; $g^{(1)}$ and $g^{(2)}$ are the filters defined in (3) and (4) and d represents the Dirac filter.

Once obtained the horizontal, vertical and diagonal detail images, and the approximation image in successive scales, the wavelet modulus is calculated from the three detail images as follows

$$I_M(x, y) = \sqrt{\left|W^1_{2^j} I_p(x, y)\right|^2 + \left|W^2_{2^j} I_p(x, y)\right|^2 + \left|W^3_{2^j} I_p(x, y)\right|^2} \tag{8}$$

and will be used for border localization. Implementation of the wavelet transform has been done using wavelets originated from m (1 or 2) order derivatives of B-spline functions of order n. In each scale information from the three detail images and

Advances in Computational & Experimental Engineering & Science

corresponding approximation image, and also modulus image I_M is used to detect the borders we are looking for.

Classification stage

Support Vector Machines (SVM) are learning systems that use as hypothesis space linear functions in a space of characteristics of high dimension, trained with a classical algorithm from the optimization theory, adding a control of the learning process derived from the statistical learning theory of Vapnik-Chervonenkis [11], for solving problems of classification, regression and estimation. In this approach, the four problems of efficiency of learning, efficiency of evaluation, overfitting and parameter adjustment of the learning algorithm are solved in an elegant manner in SVM approach [12].

In our implementation, we have chosen as kernel the polynomial of degree 3, the input to the SVM classifier is preceded by a binary tree classifier, and the vector is a group of the following characteristics:

- Number o microcalcifications
- Area of the biggest microcalcification, A_{max}.
- Maximal relative linearity: D_{i1}/D_{i2}, being D_{i1} and D_{i2} the distances average and maximal, respectively, from the border to the centre of the *i*-th microcalcification.
- Average circularity, where we can express circularity as P_i^2 / A_i being P_i the perimeter of *i*-th microcalcification and A_i its area.
- Average rectangularity, where rectangularity can be expressed in the way A_i / A_{ri} being A_{ri} the area of the minimal rectangle that includes the *i*-th microcalcification.
- Average number of microcalcifications per cumuli.

Results

The database used for this work comes from the *Puerta de Hierro* Madrid hospital and is a set of 43 digitalized mammographic images, 25 of them corresponding to benign pathologies and 18 to breast cancer. Due to the good performance of the detection stage, only few microcalcifications are not detected and the classifications results are 95.35% of true positives, a 2,33% of false positives and a 2,33% of false negatives. Detection of the maximum possible number of microcalcificactions is very important for the success of the system, being very critical the correct adjustment of noise thresholds in the contourlet pre-processing stage.

Conclusions and further developments

The proposed system combines several state of the art image processing techniques, namely contourlet transforms for the noise removal of the mammographic images and border detection with the wavelet transform modulus maxima lines. Also a SVM design

Advances in Computational & Experimental Engineering & Science

of the classificator is a proven more robust approach to the classification stage. Further improvements under research are the study of the geometric characteristics of the transformed coefficients in the contourlet transform, to determine another border detection strategy, and also the possibility to define alternative geometric characterization of the microcalcifications different from the border concept itself, in order to develop different vectors to enter in the classification SVM stage.

References

1 N. Karssemeijer, "Computer-Assisted Reading Mammogramms", *European Radiol*., vol. 7, pp. 743-748, 1997.

2 C. Chang, A. Laine. "Coherence of Multiscale Features for Enhancement of Digital Mammograms". *IEEE. Trans. on IT in Biomedicine*. Vol. 3 Nº1. March 1999.

3 J. S. Suri, S. K. Setarehdan, S. Singh (Eds.) *Advanced algorithmic approaches to medical image segmentation.* Cap 9. p. 441-540. Springer 2002.

4 Gøtzshe and Olsen. "Confusion over mammography screening intensifies". *The Lancet*. (2001; 358: 1340-42) October 2001.

5 L. W. Basset, D. H. Bunnell, R. Jahanshahi, R.H. Gold, R. D. Arndt, J. Linsman, "Breast cancer detection: one versus two views", *Radiology*, 165, 95-97, 1987.

6 L. Li, W. Qian, L. P. Clark. "X-Ray medical image processing using directional wavelet transform". *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*. vol. 4, p. 2251-2254. Atlanta. 1996.

7 M. N. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multirresolution Image Representation. *IEEE Transactions on Image Processing.* To appear. Draft october 2003.

8 D. D. Y. Po and M. N. Do. Directional Multiscale Modeling of Images using the Contourlet Transform. *IEEE Transactions on Image Processing.* To appear. Draft, july 2003.

9 P. J. Burt and E. H. Adelson. "The Laplacian Pyramid as a compact image code". *IEEE Transaction on Communications.* Vol. 31, Nº 4, pp 532-540, april 1983.

10 S. Mallat. *A wavelet tour of signal processing*. Academic Press. 1998.

11 V. Vapnik, A. Chervonenkis. "On the uniform convergence of relative frequencies of events to their probabilities". *Theory of Probability and its Applications*, 16(2):264-280, 1971.

12 N. Cristianini, J. Shawe-Taylor. *Support Vector Machines and other kernel-based learning methods.* Cambridge University Press. 2000.