Available online at www.sciencedirect.com
ScienceDirect
Computerized Medical Imaging and Graphics 32 (2008) 304–315
Computerized Medical Imaging and Graphics
www.elsevier.com/locate/compmedimag

Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection

Alfonso Rojas Domínguez [1], Asoke K. Nandi *

Department of Electrical Engineering & Electronics, The University of Liverpool, Brownlow Hill, Liverpool L69 3GJ, United Kingdom

Received 11 April 2007; received in revised form 23 January 2008; accepted 28 January 2008

Abstract

A method for automatic detection of mammographic masses is presented. As part of this method, an enhancement algorithm that improves image contrast based on local statistical measures of the mammograms is proposed. After enhancement, regions are segmented via thresholding at multiple levels, and a set of features is computed from each of the segmented regions. A region-ranking system is also presented that identifies the regions most likely to represent abnormalities based on the features computed. The method was tested on 57 mammographic images of masses from the Mini-MIAS database, and achieved a sensitivity of 80% at 2.3 false-positives per image (average of 0.32 false-positives per image).

Keywords: Breast cancer; Breast masses; Mammography; Tumor detection; Image analysis

1. Introduction

Breast cancer is the most common form of cancer in the female population, affecting one of approximately 11 women at some stage of their life in the Western world [4,19]. As with any form of cancer, early detection of breast cancer is one of the most important factors affecting the possibility of recovery from the disease. Early detection of breast cancer can be achieved through mammography screening programs assisted by computers [7]. Over the past one and a half decades, several researchers have studied and proposed methods for computer-aided detection and classification of abnormalities related to breast cancer in mammograms.

Several methods have been proposed for the detection of tumors in mammograms. In general, these methods are a combination of image processing techniques and pattern recognition algorithms. Some researchers have approached the problem of tumor detection by devising methods to detect all types of masses (circumscribed, spiculated, and ill-defined), while others have focused their efforts on the detection of masses of a particular type or with a diagnosis of malignancy. A brief review of the methods that are more relevant to this study is provided below.

Kegelmeyer et al. [12] investigated the use of a computer vision method as a second reader for the detection of spiculated lesions on screening mammograms. Their method first assigns to each pixel in an image a probability of being suspicious based on five features: analysis of local oriented edges and four Laws' texture energy measures [14]. The probability of each pixel is computed via a binary decision tree. Then, the final selection of suspicious regions is performed through spatial filtering and edge detection. Their method achieved 97% per-image sensitivity and 82% per-case specificity (from a range of zero to five false positives (FPs) per image). Some disadvantages of the method presented by Kegelmeyer et al. are that the algorithm is computationally expensive and was designed to detect only spiculated lesions; and that their study included a relatively small dataset with only 36 positive cases and 49 negative cases.

Polakowski et al. [24] presented a model-based vision algorithm to detect and classify masses in digitized mammograms. Their algorithm consists of five modules, two of which perform the detection of masses and the reduction of FPs prior to classification. Their first module, the focus of attention module, uses a difference of Gaussians (DoG) filter followed by thresholding to select regions of interest (ROIs). Their second module, the index module, produces an approximate mask of the mass in each ROI and reduces the number of false ROIs based on the area, contrast and circularity of the masks. Polakowski et al. tested

* Corresponding author. Tel.: +44 1517944525; fax: +44 1517944540.
E-mail addresses: Alfonso.Rojas@gmail.com (A. Rojas Domínguez), A.Nandi@liv.ac.uk (A.K. Nandi).
[1] Supported by the National Council of Science and Technology (CONACYT) of Mexico.

0895-6111/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compmedimag.2008.01.006

their algorithm on a database containing 272 images (including 36 malignant and 53 benign mass images), and reported a sensitivity of 92% in locating malignant ROIs (33 out of 36 malignant masses correctly located), with an average of 2.36 FPs per image after the first two modules. Including the rest of their modules (in charge of feature selection, feature extraction, and classification), the number of FPs was reduced to 1.8 per image. Disadvantages of this method are that the parameters of the algorithm are dependent on characteristics of the database used; the study was focused on finding only medium-sized masses; and due to the use of all the data for training and testing, the results reported may be optimistically biased.

Christoyianni et al. [5] presented a method for fast detection of circumscribed masses in mammograms. Their method performs classification of mammographic regions into tumorous and healthy tissue via a radial basis function neural network, and then provides the location of the masses by employing sub-image windowing analysis. For image preprocessing, Christoyianni et al. applied a sharpening filter to maximize the contrast between masses and local background. They reported results of recognition of abnormal tissue of 90.9%, 62.5%, and 33.3% in fatty, glandular, and dense tissue, respectively. The obvious disadvantage of this method is that it is highly sensitive to the background tissue. A second disadvantage is that it was designed to detect only circumscribed masses.

Mudigonda et al. [18] presented a mass detection method that performs segmentation of objects based on isointensity contours and texture flow-field analysis. Their study included 43 masses (30 benign and 13 malignant) and 13 normal cases from the Mini-MIAS database [28]. Circumscribed and spiculated masses were present in both the benign and malignant classes, but no ill-defined masses were included. Their method consists of two stages. The first stage is in charge of the detection of masses, and the second stage performs rejection of FPs via pattern classification. The reported detection accuracy after the first stage is 74% (11 missed cases). The second stage achieved 81% sensitivity at 2.2 FPs per image. However, considering that 11 masses were missed in the detection stage, the overall detection sensitivity of their method reduces to 60% (26 masses detected out of 43 in the test set) at 2.2 FPs per image.

Zheng and Chan [29] proposed a segmentation method using the discrete wavelet transform and a multiresolution Markov random field as part of their detection algorithm, and used a binary decision algorithm to select suspicious areas based on features generated from the segmented regions. Their study included 322 mammograms from the MIAS database with a total of 37 masses. Well-defined circumscribed masses and ill-defined masses were included, but not spiculated masses. The sensitivity of the algorithm was reported as 97.3% but no operating point was defined. Instead, only the average number of FPs was reported as 3.92 per image.

A density-weighted adaptive contrast enhancement (DWCE) filter was presented by Petrick et al. [21] as part of a mass-detection algorithm including Laplacian-of-Gaussian edge detection and morphological feature classification. The DWCE filter was later used with improved results in other mass detection algorithms [22,23]. The revised procedure [23] included a local border refinement procedure and characterization of structures based on morphological and textural features. Petrick et al. tested this algorithm on a dataset including 175 malignant masses and 149 benign masses. Their reported detection rate was 84% of individual malignant masses and 70% of benign masses (77.47% of masses detected overall), with an average of 1.5 marks per mammogram.

In this paper, a method for the detection of masses in mammograms is proposed. This method incorporates some of the advantages of the algorithms described in the literature (such as a modular design, employment of features for reduction of false positives, etc.) while avoiding some of the identified drawbacks (such as dependence of the algorithm parameters on the database, a high computational complexity, etc.). Our interest is focused on the detection of masses, either benign or malignant, and whether well-defined circumscribed, spiculated, or ill-defined masses.

The proposed method is divided into three main stages. The first stage is a *mammogram enhancement* procedure that has the objective of improving the segmentation of the distinct structures in the mammogram when performed via simple conversion to binary images at multiple threshold levels (Section 3.1). Our enhancement algorithm is different from others in the literature in that it computes the parameter of the enhancing function in an adaptive manner, based on local statistics of the pixels in the mammographic image.

The second stage consists of *segmentation and feature extraction* steps (Section 3.2). In this stage, regions are segmented and several shape and gray-level characteristics of the regions are computed.

Finally, in the third stage, a ranking system is employed to select suspicious regions (Section 3.3, *region selection*). The ranking system is a novel approach to the problem of region selection (i.e., the elimination of FPs) that does not require training, and implements a type of on-the-fly feature selection.

The rest of this paper is organized as follows: in Section 2, we describe the materials employed in this study (database and software). In Section 3, we present in detail our methodology for mass detection. Experimental results and discussions are presented in Section 4. Finally, conclusions and directions for future work are given in Section 5.

2. Materials

The database of mammograms used in this study is known as Mammographic Image Analysis Society (MIAS) Mini Mammographic Database [28]. In the Mini-MIAS database, the MIAS Database (an earlier version digitized at 50 μm pixel size) has been downsampled to 200 μm pixel size and clipped/padded so that every image is of size 1024 × 1024 pixels. The MIAS database has been used in several research studies previously [5,6,15,18,21,27,29,30] and could be considered as a benchmark database. The database is available from: http://peipa.essex.ac.uk/info/mias.html.

The database is arranged in pairs of films, where each pair represents the left (even filename numbers) and right (odd filename numbers) mammograms of a single patient, in mediolateral

Table 1
Characteristics of the MIAS database of mammograms

(1) MIAS database reference number.
(2) Nature of background tissue:
Fatty
Fatty-glandular
Dense-glandular
(3) Class of abnormality present:
Calcification
Well-defined/circumscribed masses 25 cases, 21 benign
Spiculated masses 19 cases, 11 benign
Other, ill-defined masses 15 cases, 7 benign
Architectural distortion 19 cases, 9 benign
Asymmetry, 19 cases, 6 benign
Normal, 207 cases
(4) Severity of abnormality:
Benign 54/115 cases
Malignant 61/115 cases
(5) Abnormality location: x,y image coordinates of center of abnormality.
(6) Approximate radius (in pixels) of a circle enclosing the abnormality.

oblique (MLO) views. There are a number of variables that characterize each mammogram, as shown in Table 1.

There is a total of 322 cases; of these, 207 are normal, while the other 115 possess one or more abnormalities with diagnosis according to the statistics in Table 1. One of the cases with a circumscribed mass has missing information on the location of the mass (MIAS reference number mdb059). This version of the database is not suitable for experiments on detection of microcalcifications because of the resolution to which it has been digitized. Ideally, the resolution for experiments on microcalcifications should be 50 μm per pixel. However, the database is useful for experiments on detection of the other types of abnormalities listed above.

The mammograms were digitized from film, and because of this, artifacts are present in some of the images that can affect the processing, but not severely. Besides digitization artifacts, the presence of labels on the film, as well as perforations and tape on some of the films and the incorrect placement of the film on the scanner when it was digitized can also affect the processing of the images. It is assumed that mammograms obtained in more recent screening programs, through the use of conventions for X-ray imaging of the breast region and modern equipment, these problems will disappear. Thus, we do not discuss methods for the elimination of the artifacts or other unexpected elements.

The complete method presented in this paper was implemented in MATLAB [16] version 7, and makes extensive use of the Image Processing Toolbox (version 5.2). Some of the functions for computation of shape properties were taken from Chapter 7 of Nixon and Aguado [20].

3. Methodology for mass detection

The methodology used consists of three main stages. In the first stage the images are enhanced to make all structures equally detectable (or *approximately* equally detectable). In the second step, the enhanced images are segmented into distinct regions through thresholding at multiple levels, and a number of features are computed from each one of the regions present at each segmentation level. The third stage is the selection of suspicious regions based on their features and employing a ranking system. Each of these stages is described in detail below.

The pectoral muscle in most MLO views of mammograms appears as the predominant density region, and can affect negatively the results of image processing methods (see, for example, reference [18]). For this reason, the region representing the pectoral muscle was manually removed from all the images prior to any further processing. The identification and removal of the pectoral muscle, as well as the identification of the breast boundary, are problems on their own that have received the attention of researchers in the past but are disregarded in this study. The reader interested in these problems is referred to references [1,8,9,11].

3.1. Mammogram enhancement

One of the difficulties that has to be overcome for a successful detection of masses in mammograms is caused by the difference in brightness of the objects in the mammograms. This difference is due to the difference in density of the tissue that absorbs the X-ray radiation. This effect is the useful property of the mammographic technique, since it allows the separation of objects with different densities; but it also affects negatively regions with lower average density. Density is not completely uniform across the whole breast, and the average image intensity is lower at the boundaries of the breast [2]. Depending on the detection algorithm, this effect may increase the difficulty of detecting masses that are located near the breast boundary. Furthermore, when the density of the parenchyma is high, masses with a lower density appear with low negative contrast in the mammograms. Often such masses are missed by algorithms that use the brightness level of the structures as their main feature for detection. Other factors that increase the difficulty of obtaining a successful detection are a low signal-to-noise ratio at the edges of masses (particularly of malignant masses) and complex structures in the background of the mammogram.

In an attempt to alleviate the situations described above, an image enhancement procedure is proposed. The objective of this procedure is to increase the contrast between mammogram structures and their background, while providing a relatively uniform intensity to all of the structures. Previous studies [17,21,23,26] have demonstrated the benefits of applying a contrast enhancement processing driven by the local characteristics of the mammograms. In our contrast enhancement routine, statistical measures of the pixel intensities in local neighborhoods are employed to automatically set the parameters of the transform applied to each pixel. In order to obtain effective enhancement over all mammographic structures of different sizes, the enhancement routine is wrapped into a multiscale processing framework. Application of the enhancement procedure over one image is illustrated in Fig. 1. The procedure consists of four main steps:

(1) Eliminate background and scale down the image, producing a multi-scale representation (three scales were used, each

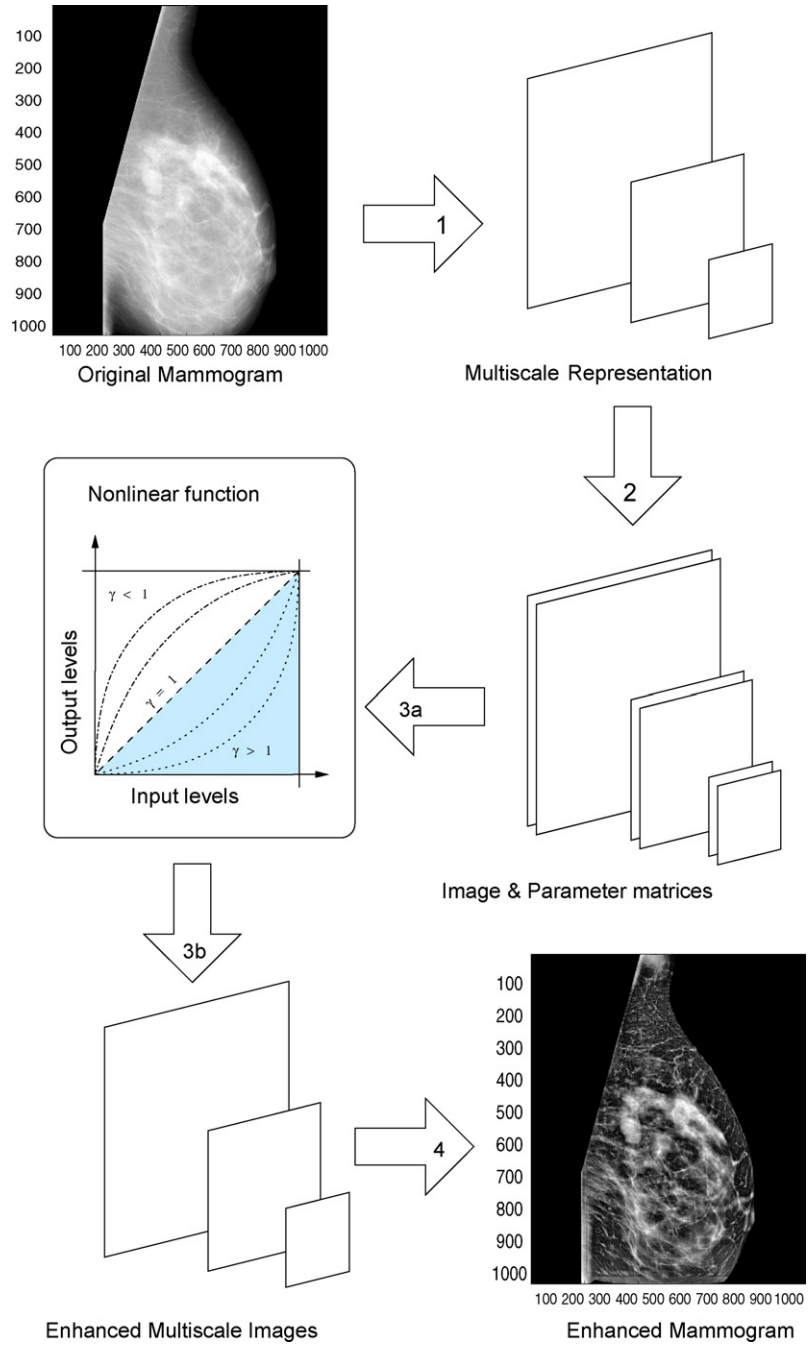

Fig. 1. Enhancement procedure. Refer to the text for an explanation of each of the steps illustrated.

one producing an image half the size of the image in the previous scale).

(2) Produce a parameter matrix for each scale.
(3) Each scale is processed in blocks, which are sent to the nonlinear function together with the corresponding parameter value. Application of the nonlinear function produces one enhanced image per scale.
(4) Scale up the images back to the original scale, and combine them (linear combination) to produce the final result.

The background of the mammograms is eliminated by substraction of the opened image from the original image (a morphological operation known as the top-hat operation). However, prior to the substraction, a Gaussian filter is applied to the opened image to soften any edges and better preserve small structures. The structuring element used in the top-hat operation is a disk of radius $r_{top} = 80$ pixels. This structuring element was chosen after experimentation and examination of the distribution of the masses size (Fig. 2). To clarify, notice that most of the masses (about 88%) have a size below the size of the chosen structuring element (radius of 80 pixels). These masses will not be substracted by the top-hat operation, and will be preserved for later detection. The masses that are bigger than the size of the structuring element will be (partially) substracted. The size

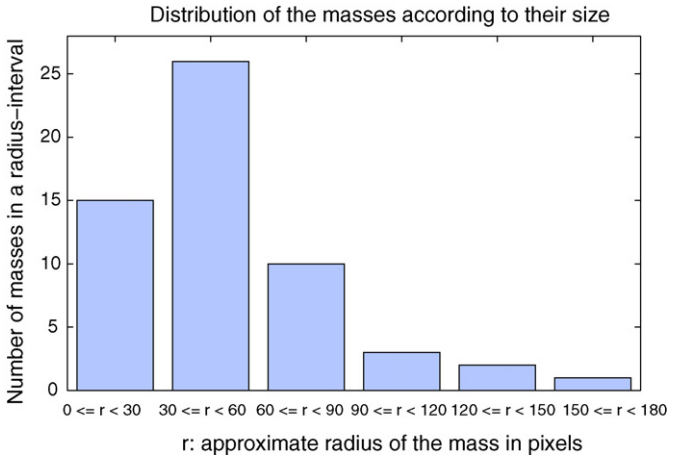

Fig. 2. Distribution of masses in database according to size.

of a radius of 80 pixels corresponds approximately to the inflexion point in the histogram of mass-sizes (only a few masses are bigger than the structuring element selected, and most of these are in fact much bigger than the structuring element). Because of this, and due to the intensity profile of large masses, the central part of these will survive the operation, thus enabling their detection by the algorithm.

A nonlinear function is applied on each block of pixels. The nonlinear function used is a mapping given by: $y = [(x - l_x)/(h_x - l_x)]^{\gamma}(h_y - l_y) + l_y$, where l_x and h_x are the lowest and highest intensity values of the input matrix x, respectively, and, similarly, l_y and h_y are the lowest and highest intensity values of the output matrix y. Notice that, if the intensity limits of the input and output matrices are set to the range [0, 1], the nonlinear function is reduced to: $y = x^{\gamma}$. The parameter γ is given by the value of the element at the centre of the block in the corresponding parameter matrix. The parameter matrix is produced based on on local statistical measures in the mammograms. This is further explained below.

The moving-window application of the nonlinear function is illustrated in Fig. 3. The image matrix is processed one block of $[N \times N]$ pixels at a time; each block is centered on the pixel at location (i, j), where $i = \{1, 2, \ldots, \text{Rows}\}$, $j = \{1, 2, \ldots, \text{Cols}\}$, and [Rows × Cols] is the size of the image being processed. The size of the processing window, $N = 31$, was chosen after examination of the structures present in the mammograms, their size (see Fig. 2), and how they were modified by the scaling function. The size of the processing window is not critical, but some considerations must be observed. One condition on the size of the processing window is that it should be an odd number. This is computationally attractive because windows of such sizes are centered on individual pixels, thus simplifying the application of the function. Notice as well that the processing window has a relative size, depending on the scale of the image that is being processed. In our multi-scale representation, the scale is increased by a factor of 2 at each iteration (the image becomes half the original size), and thus the size of the processing window is *relatively* increased by the same factor (although the actual size of the window remains constant). It was observed that it is not required that structures (such as masses)

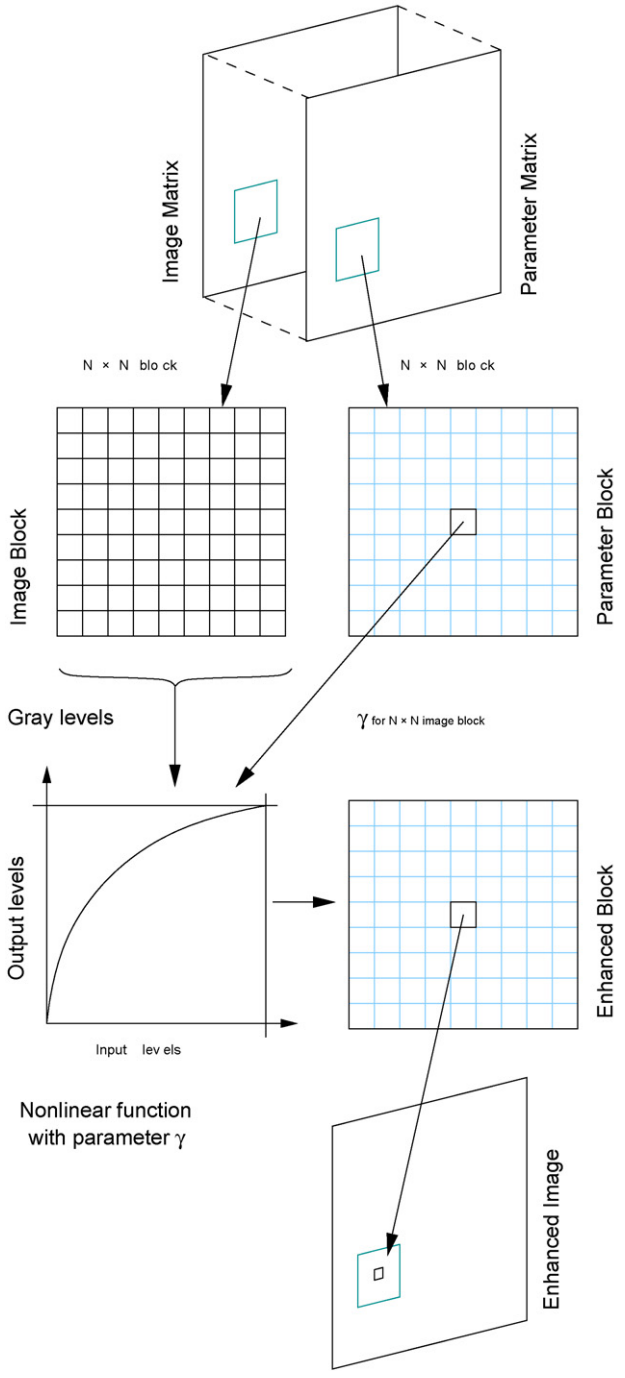

Fig. 3. Block processing with nonlinear function. Each $[N \times N]$ block is centered on the pixel at location (i, j). The value of the parameter γ for each block is stored in a separate matrix. The center pixel of the output of the nonlinear function becomes pixel (i, j) of the enhanced image.

be completely contained by the window for the enhancing operation to be efficient. This is because the form of the function is dynamically adjusted based on on the statistical measures of the image, which are computed in smaller neighbourhoods than the size of the processing window (see below).

The values of the parameter γ used by the function on each block is stored in a second matrix, the parameter matrix. The values in the parameter matrix are obtained by mapping the

Table 2
Mapping from the statistical measure space (μ_B, σ_B) into the parameter space (γ)

Condition	Value range	Numerical value
$\mu_B < \bar{\mu}$ and $\sigma_B < \bar{\sigma}$	$\gamma > 1$	1.2
$\mu_B < \bar{\mu}$ and $\sigma_B \geq \bar{\sigma}$	$\gamma \approx 1$	0.8
$\mu_B \geq \bar{\mu}$ and $\sigma_B < \bar{\sigma}$	$\gamma < 1/2$	0.4
$\mu_B \geq \bar{\mu}$ and $\sigma_B \geq \bar{\sigma}$	$\gamma < 1$	0.6

μ_B and σ_B represent the mean and standard deviation of the pixels in a block, respectively. $\bar{\mu}$ and $\bar{\sigma}$ represent the average mean and standard deviation of the pixels in the whole breast region.

local statistical measures (mean and standard deviation) of the image matrix. First the statistical measures are computed in small neighborhoods ([5 × 5] pixels), and then the rules given in Table 2 are applied. To complete the computation of the parameter matrix, a Gaussian filter is applied to smooth the boundaries between the adjacent regions with different values. The conditions in Table 2 correspond to different mammographic tissues. Smooth bright regions, like the tissue found in the interior of a mass, show a large mean value and a small standard deviation; heterogeneous regions, like the glandular tissue in the mammograms, possess a relatively large standard deviation and mean values, etc. As illustrated in Fig. 1, the shape of the curve mapping the intensity levels in the nonlinear function is specified by the parameter γ. If $\gamma < 1$ (strictly, $0 \leq \gamma < 1$), the mapping is weighted toward higher output values. If $\gamma > 1$, the mapping is weighted toward lower output values. When $\gamma = 1$ the mapping is linear. Thus, the range of values assigned to γ in Table 2 were selected in such a way that the mapping improves the contrast of the different tissues. For example, in the case of smooth bright

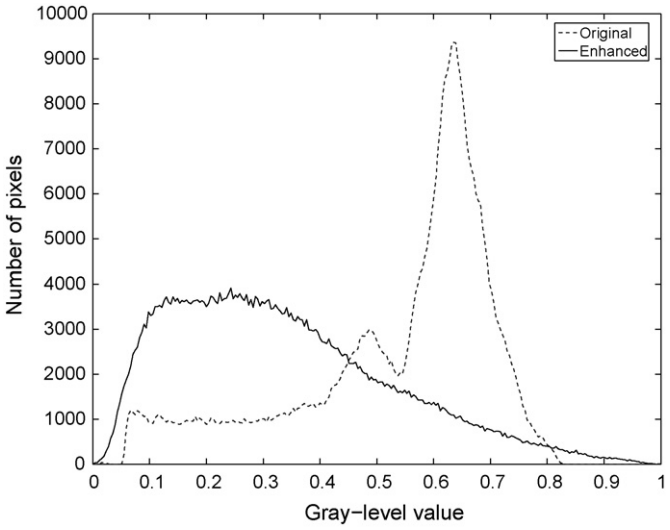

Fig. 5. The histogram of the original image (dashed) and of the enhanced image (solid). The value for zero gray-level has been removed for improved visualization of the details of interest.

regions the mapping should be weighted toward higher output values so that the intensity of pixels with relatively low intensity is increased; this improves the smoothness of the region by making the intensity values more uniform. The specific numerical values were selected experimentally. The range of the input levels is the range of values in the block of pixels being processed; the range of the output levels was set to [0,1].

Fig. 4 illustrates the effect of the enhancement routine. It can be observed that all structures at different scales are easily distinguishable. The histograms of the original and enhanced images

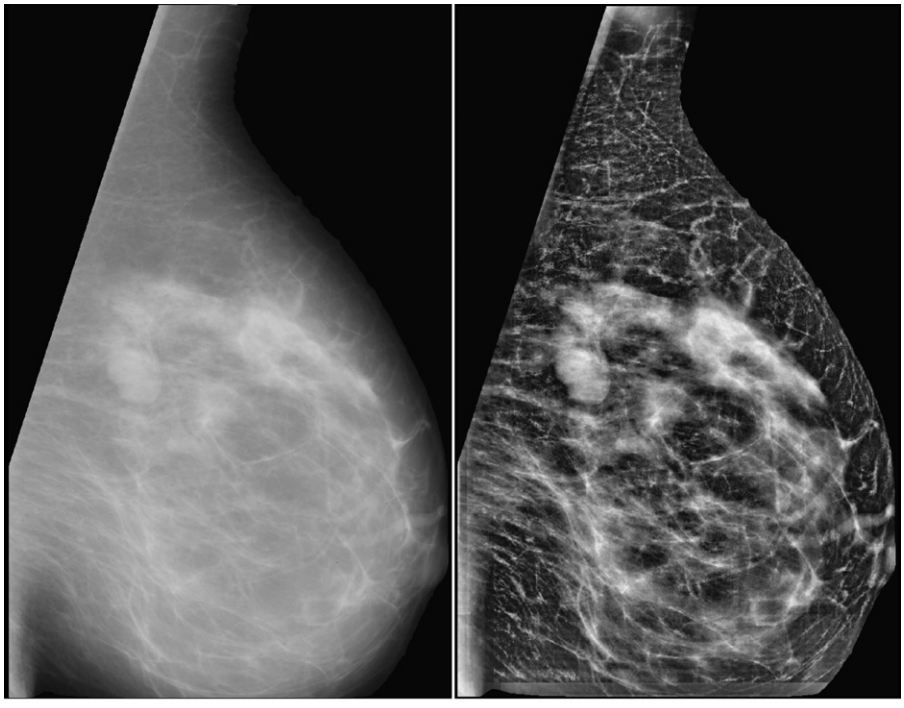

Fig. 4. Example of the effect of the proposed enhancement routine. A: Original mammogram. B: Enhanced version.

are compared in Fig. 5, where the pixel-count of level zero has been eliminated to enable the appreciation of the other levels. From the histogram of the enhanced image, it is easy to understand why this image is more suited for a multilevel thresholding segmentation process than the original image: in the histogram of the enhanced version, as one moves from the top of the gray-level scale towards the lower values, there is a gradual increase in the number of pixels at a given gray-level. This supports the idea of employing a decreasing threshold value to separate portions of the image in the form of regions that will increase in size as the threshold value is decreased. In comparison, if the histogram of the original version of the image is examined, one can observe that the number of pixels reaches a maximum much earlier than in the enhanced version; the effect is that many of the regions being segmented merge at an earlier stage than desired during the multilevel thresholding process.

The enhancement routine proposed is suitable and very effective on most mammograms. The exceptions are mammograms that present homogeneous regions with very high density. In such cases, the abnormalities can be obscured by the parenchyma and show very similar intensity levels. The enhancement algorithm provides limited (though still appreciable) improvement in such cases.

3.2. Segmentation and feature extraction

The enhanced images are converted to binary images through thresholding at different values starting from the top level. At each iteration, a binary image is produced by setting the value of pixels that are above the current threshold in the mammogram to one, while all other pixels are set to zero. The process continues until the whole breast region is segmented or until the threshold value reaches a chosen minimum. This segmentation technique has been used previously by Mudigonda et al. [18] where it was called density slicing. It was found that for the enhanced images in this study, with gray values in the range [0, 1], 30 levels with a step size of 0.025 were adequate to segment all the mammograms. That is, for each mammogram, the algorithm produced 30 binary images with the segmented regions at the corresponding level of segmentation: $1 - 0.025l$, where $l = 1, 2, \ldots, 30$.

Once the segmentation procedure described above is completed, the binary images are filtered with a Gaussian smoothing filter (parameters $\mu = 9$ pixels and $\sigma = 5$ pixels) to eliminate noise (any isolated pixels) and split regions that are joined by single pixels or by a small group of pixels. Next, all the holes in the existing regions (if there are any), are filled automatically. Finally, all regions with a pixel count of less than 150 pixels are eliminated because they are well below the size of the smallest mass in the Mini-MIAS database and it is unnecessary to process them. The elimination of regions with pixel counts of less than 150 pixels was performed to save computing time. The specific value of 150 pixels was chosen after considering that the smallest mass in our database has an approximate radius of 17 pixels, which, if the mass was a perfect circle, would roughly correspond to a pixel count of 900 pixels. Fig. 6 illustrates the set of regions extracted from the mammogram in Fig. 4 B. For illustration purposes, only seven levels (but including the whole range of segmentation values) are presented, each in a different colour, and each region is indicated by its contour.

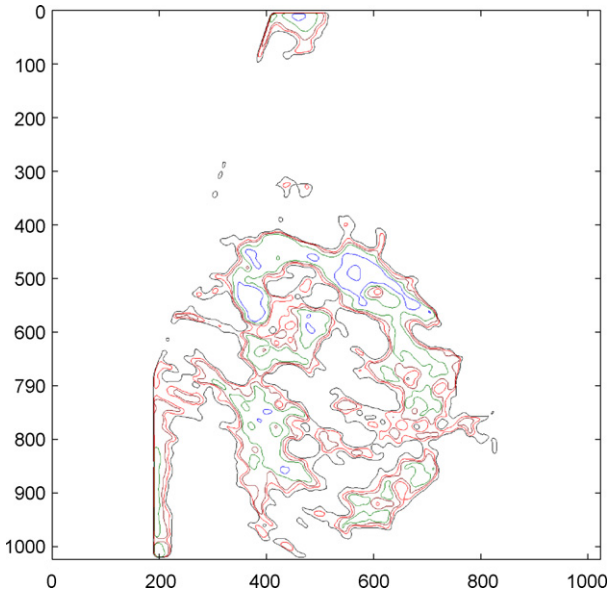

Fig. 6. Contours of regions segmented at multiple threshold levels.

To complete this stage of the detection method, a set of properties of the remaining regions are computed and stored together with the binary image containing all regions at the corresponding segmentation level. To define which pixels belong to the same object and which are part of two adjacent objects, four-connectivity was employed. The properties obtained from each region are listed in Table 3.

Most of the properties are shape descriptors; the basic shape properties are *area* and *perimeter*. The *minor axis length* and *major axis length*, as well as the *orientation* and *eccentricity* are computed from an ellipse with the same second moments as the region. The *equivalent diameter* is of the circle with the same

Table 3
Properties extracted from the regions segmented

1. Area
2. Perimeter
3. Major axis length
4. Minor axis length
5. Eccentricity
6. Orientation
7. Equivalent diameter
8. Solidity
9. Extent
10. Compactness
11. Dispersion-I
12. Dispersion-II
13. Mean gradient within region
14. Mean gradient of boundary
15. Gray value variance
16. Edge distance variance
17. Mean intensity difference
18. Fractal dimension

area as the region. *Solidity* is the proportion of the pixels in the convex hull that are also in the region; another similar measure is *extent*, which is the proportion of pixels in the bounding rectangular box of a region that are also in the region. Details about these and other shape properties can be found in Chapter 11 of Gonzales et al. [10].

Compactness of a shape is the ratio of its squared perimeter to its area. *Dispersion*, also known as irregularity, measures the density of the region. *Dispersion-I* is defined as the ratio of the major chord length to area, while *Dispersion-II* is defined as the ratio between the radius of the minimum circle enclosing the region and of the maximum circle that can be contained in the region. The implementation of these properties is described in Chapter 7 of Nixon and Aguado [20].

Properties 13–17 were used by Zheng and Chan [29], and before them by Qian et al. [25] in a study of feature extraction for mass detection, and are formally defined by them.

The *mean gradient within region* is the mean value of the gradient of the pixels inside the region; the *mean gradient of boundary* is the mean value of the gradient of the pixels on the boundary of the region. *Gray value variance* is a measure of the smoothness of the region. *Edge distance variance* is a measure of rotational symmetry and shape. *Mean intensity difference* is a measure of the difference in gray values outside and inside the region. All of these measures were computed using the gradient and intensity values of the enhanced mammogram.

Finally, the *fractal dimension* is a measure of the extent of self-similarity of an object when examined at different scales. This measure was computed using an adaptation of the method of Caldwell et al. [3], from the original mammogram. The fractal dimension of a 3-D surface can be obtained using

$$A(\epsilon) = \lambda \epsilon^{2\text{-D}} \tag{1}$$

where $A(\epsilon)$ represents the surface area measured with a square of side ϵ, λ a scaling constant, and D is the fractal dimension, which is related to the slope of the plot of $\log(A(\epsilon))$ versus $\log(\epsilon)$. In particular, Caldwell et al. considered the 3-D surface of the image (i.e., interpreting the gray value as height) to be a set of rectangular columns, where the top of each column is a square with side of length ϵ. They computed the area of the surface as

$$A(\epsilon) = \sum_{x,y} \epsilon^2 + \sum_{x,y} \epsilon \left[|I(x, y) - I(x+1, y)| + |I(x, y) - I(x, y+1)| \right] \tag{2}$$

where I represents the intensity of a pixel. In the present study, the fractal dimension was computed locally using a square block-processing window with a side of length equal to 40 pixels. The area of this window was calculated repeatedly with square neighborhoods of side $\epsilon = 3, 5, 7$ and 9 pixels. The fractal dimension for all pixels in each processing window was set to the slope of the corresponding plot of $\log(A(\epsilon))$ versus $\log(\epsilon)$. Finally, the fractal dimension of any region overlapping two or more of the processing windows was computed by using the average of the fractal dimension in the overlapping region of the windows.

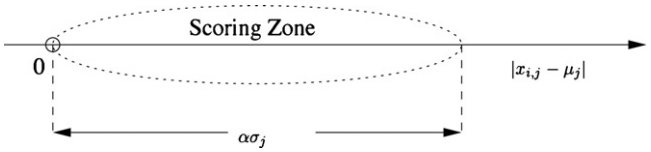

Fig. 7. The concept of a scoring area: if the absolute value of the difference between property j of a region and the mean of property j over all masses is inside the scoring zone, the region receives a score of 1.

3.3. Selection of suspicious regions

The selection of suspicious regions is performed by means of a ranking system. Ideally, all the properties of the regions representing abnormalities would be located within a relatively compact range of values, whereas the properties of regions corresponding to normal tissue would be outside this range. Thus, the detection of masses would be achieved via a selection of the appropriate threshold values for each of the features. In practice, there will always be outliers from both of the classes (i.e., normal and suspicious regions), and the range of values describing suspicious and unsuspicious regions will overlap, often significantly, making the selection of the optimal threshold values a difficult task or the use of simple linear discrimination rules altogether unsuitable. The ranking system used in this study is based on the assumption that even when not all the properties of a suspicious region will be concentrated around a certain value and within a fixed range, most of them will be so concentrated. By considering how many of the properties of each given region are concentrated around a reference value and within a fixed range (let this be called the *scoring zone*), a rank can be assigned to each region. The reference value for each property is the mean value of that property computed over the set of masses, and is located at the center of the scoring zone. The range defining the extent of the scoring zone is the standard deviation of the property (again, computed over the set of masses) times a regularization factor α; this is illustrated in Fig. 7. The reference means and standard deviations were obtained using the ground truth of the masses in the Mini-MIAS database as a guide: a contour of each mass was manually drawn and the properties of the mass regions were computed to obtain the desired statistical measures. Note that, before the computation of the rank or any statistical measure, the values of all the properties were scaled to be within the range [0, 1].

To compute the rank of the i-th region, the absolute value of the difference between the set of properties $\bar{x}_i$ and the set of means $\bar{\mu}$ is compared against the limit of the scoring zone, $\alpha\bar{\sigma}$, where $\bar{\sigma}$ is the set of standard deviations. A score of 1 is given to the region for each difference that is lower or equal to the corresponding limit (i.e., whose value is inside the scoring zone), and zero for each difference that is larger than the limit (outside the scoring zone). The scores are stored in a vector of 1s and 0s, $\bar{Z}_i$, that works like a scoring card: the vector shows how many and which properties of each region are inside the scoring zone. The length of the vector is equal to the number of properties being used. To compute the final rank of each region, the scoring system has to be compensated for differences in the dispersion of the statistical distribution of the properties. That

is, we would like to reflect the fact that some of the properties have more discriminating power than others. This is achieved by dividing each element in the vector $\bar{Z}_i$ by the corresponding value of $\bar{\sigma}$, and then computing the Euclidean norm of the result; the rank of the i-th region is the norm, mathematically expressed as

$$\begin{aligned} \bar{Z}_i &= [|\bar{x}_i - \bar{\mu}| \leq \alpha\bar{\sigma}], \\ \text{rank}_i &= \|\bar{Z}_i/\bar{\sigma}\|, \end{aligned} \tag{3}$$

where we use $[\cdot]$ to clarify that $\bar{Z}_i$ receives the outcome of the test condition $|\bar{x}_i - \bar{\mu}| \leq \alpha\bar{\sigma}$, which is 1 if the condition is true and zero otherwise. The similarity between this ranking scheme and the definition of the normalized Euclidean distance between two multidimensional random vectors given by

$$D_i = \sqrt{\sum_j \frac{(x_{i,j} - \mu_j)^2}{\sigma_j^2}} \tag{4}$$

should be evident. The difference is that, in the ranking rule, the numerator of the elements in the sum of (4) are first converted to one of two possible values: the value of the numerator is set to one if the corresponding difference is less than or equal to a threshold ($\alpha\sigma_j$ for the j-th property) or to zero if the difference is larger than the threshold.

Once the ranks of all regions are computed, the algorithm selects the ones with high ranks up to a desired number of regions. Note that the choice of the number of regions that are to be returned by the algorithm does not affect the processing time, because the ranks of all the regions must be processed before any number of them is selected.

4. Results and discussion

4.1. Results

The algorithm for mass detection was tested on a selection of mammograms from the Mini-MIAS database. The number of masses in the Mini-MIAS database is 59, which includes 25 circumscribed masses, 19 spiculated masses, and 15 ill-defined masses, as shown in Table 1. However, one of the cases is missing ground truth information and was not considered in the test (case mdb059). Another mammogram (case mdb005) shows two abnormalities which appear as a single object in the MLO view; thus, it was considered as a single mass. The detection method was tested, therefore, on 57 masses including circumscribed, spiculated and ill-defined masses.

In order to quantify the effect of the proposed enhancement procedure a contrast measure was computed over the original and the enhanced mammograms considering only the portions of the image containing the mass in each case. The contrast measure is [17]:

$$C = \frac{f - b}{f + b} \tag{5}$$

where f is the mean gray-level of the foreground region (mass) and b is the mean gray-level of the background region (tissue surrounding the mass). Foreground and background were defined using the manually segmented regions of the masses. The ground truth of the mass is the foreground region; the background was obtained by subtracting the ground truth from a dilated version of itself, producing a band of pixels approximately 20 pixels wide around the ground truth region. Using the contrast of the two sets of mammograms (original and enhanced), a ratio of the enhanced contrast to the original contrast was obtained. The summary statistics of this comparison are given in Table 4.

Table 4
Percentiles of the ratio of contrast measure of enhanced mammograms to contrast of original mammograms

Min.	First qt.	Median	Mean	Third qt.	Max.
1.526	3.134	3.901	4.652	5.874	14.472

The average of the ratio between the contrast measure of the enhanced mammograms to the contrast of the original mammograms is 4.65, with a standard deviation of 2.41. The minimum increase in contrast obtained is 1.52 times the original contrast. These results show that, according to the measure defined above, the proposed enhancement routine does in fact improve the contrast of the regions of interest.

For the detection step, a true positive (TP) was recorded for a segmented region when the region overlapped the centroid of a mass, represented by a circular area with a radius of five pixels (the center of each mass is provided in the Mini-MIAS database). Otherwise, the region was considered as a false positive (FP). A similar definition of a TP was used by Mudigonda et al. [18] and part of the method of Petrick et al. [22]. Since mass-detection algorithms work in the absence of a ground-truth, an automatic mechanism for detecting oversegmented regions

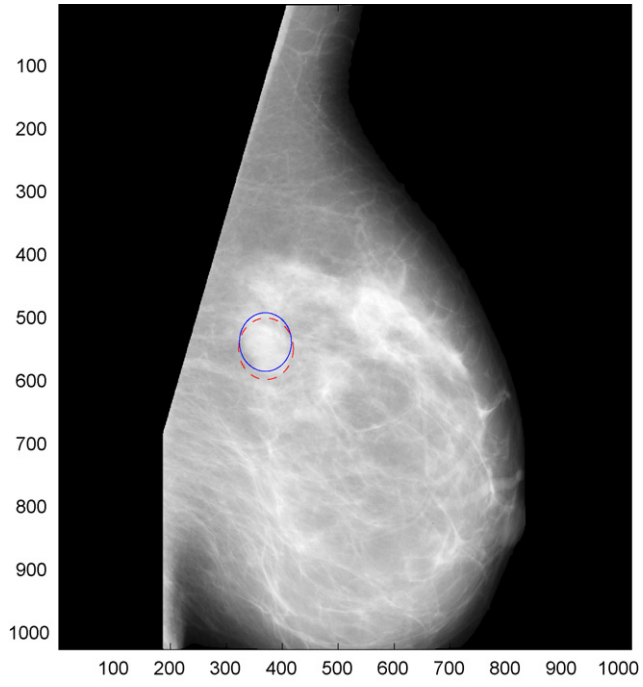

Fig. 8. Final result of the detection method for case mdb019. The solid line indicates the region as determined by the algorithm; the dashed line is the ground-truth from the Mini-MIAS database.

Table 5
Subsets of properties

Subset	Properties (see Table 3)
A	7–18
B	13–18
C	7–12

cannot be included. However, special care was taken to ensure that (excessive) oversegmentation did not occur in the regions labeled as TPs, by visual inspection of each TP recorded.

The output of the algorithm is the location of the region(s) selected, indicated by a circle of radius equal to the major axis length of the selected region(s). Fig. 8 illustrates this for the mammogram of case number mdb019 with one region selected, which corresponds to the mass. There are no FPs in this example. The output of the algorithm is shown with a solid line; the dashed line is the ground-truth as given in the Mini-MIAS database, and is shown for comparison.

The algorithm was tested with four sets of properties from Table 3 to test their discrimination power. One set included all the properties, whereas the other three included a subset of these, as shown in Table 5. Subset *A* included all properties except the very basic shape descriptors. Subset *B* included only the measures corresponding to gray-level characteristics. Subset *C* included only the more advanced shape descriptors.

The value of the parameter α for the ranking system described above was chosen after experimentation over a range of values. The procedure to select the value of α consisted of running the region-selection step of the mass-detection algorithm with different values between 0 and 3 in steps of 0.1. The value that produced the best performance overall was selected. In this way, the value of α was fixed to 1.9.

Fig. 9 presents a plot of the true positive (TP) fraction achieved using each of the four sets of properties versus the number of FPs per image, in the range 1–21. Only one mass was not detected within this range of FPs using subset *C*. The plots obtained using all the properties and subset *A* are almost identical and show the highest performance. The plots of subsets *B* and *C* show poorer performance, that of subset *C* being the worst. For a TP fraction of 0.8, the number of FPs per mammogram is 2.3 when using the set of all properties as well as when using subset *A*. Note that the number of FPs reported is the maximum per mammogram, not the average over all the mammograms (the average is 0.32 FPs per image).

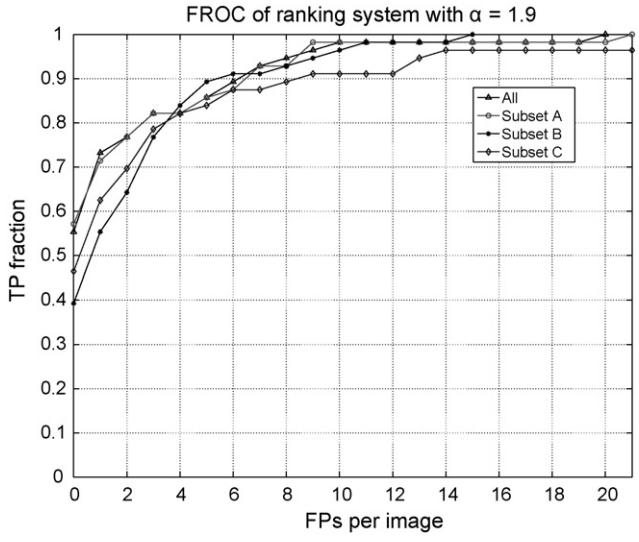

Fig. 9. True positive (TP) fractions versus false positives (FPs) per image.

4.2. Discussion

The mammogram enhancement procedure proposed in this paper works under two assumptions. The first assumption is that an object in a given image exists only if the contrast (that is the difference in the gray-level) between the object and its background is large enough; or there is an edge delimiting the boundary of an object whose brightness is the same as its background. The second assumption is that masses in mammograms are objects that share some characteristics. The first assumption implies that contrast and edge information should be enhanced and used in combination to improve mass detection. The second assumption indicates that not every structure must be equally enhanced, or the segmentation of masses will not improve. Having these two assumptions as guides, the actual implementation of the enhancement algorithm would differ greatly between researchers depending on their expertise, their programming skills, preferred processing techniques, etc. Here, one implementation has been presented, which nevertheless contains aspects that can be improved. For example, at this stage our implementation has not been optimized for speed. This optimization is necessary if the algorithm is to be used in a screening program. Furthermore, the inclusion of adaptive methods for the automatic selection of the critical parameters of the algorithm, such as the regularization factor α, could make the algorithm robust.

The detection of masses used in this study follows the general scheme of first finding all possible distinguishable regions, and then sorting out which of them actually represent masses in the mammograms. This scheme has the disadvantage that a very large number of regions must be processed, which is costly in computing time and resources. The clear advantage is that the initial sensitivity is high; other advantages are that the design of the algorithm is simple and the implementation does not require complex computations. The most time-consuming operation is the computation of the properties of all the regions segmented. The time used in this operation can be reduced by either reducing the initial number of regions or the number of properties used. Other schemes with different advantages should be investigated in combination with the enhancement procedure proposed here. For example, Zheng and Chan [29] make use of a fractal dimension measure to eliminate areas of mammograms with no masses, thus significantly reducing the number of regions to be processed.

The ranking system used in this study works at two levels simultaneously: the scoring system itself works at a general level; it is a rough or fuzzy way to determine the best candidates (that is, the most suspicious regions), like a sieve with large holes. The scaling of the scoring vector by the vector of standard deviations is a compensation step that works at a second and

finer level, like a second sieve with smaller holes. The ranking system, as a whole, then performs a type of on-the-fly feature selection for each one of the images processed. Some of the properties computed from the regions are redundant (for example, the measure of compactness includes perimeter and area), but this amount of redundancy was allowed given the design of the selection method. Kobatake et al. [13] used a method similar to our ranking system to select malignant masses. Their method computes the Mahalanobis distance measure between an input vector $\bar{x}$ and the mean vector of two categories (1: malignant mass and 2: others), D_1 and D_2. The detection rate of their system is controlled by comparing the ratio D_1/D_2 with a threshold value. Our method is different, in that it uses data from only one category instead of two; we assume a diagonal covariance matrix, and include an intermediate step that converts the distances to binary values. In other words, the method for region selection that has been presented is not a (hard) classifier.

5. Conclusion and future work

A computer-aided method for the detection of masses in mammograms has been presented. With a performance of 80% of all types of masses in the test database being successfully detected at 2.3 FPs per image (average 0.32 FPs per image), this algorithm compares well with other methods in the literature. Combining this algorithm with other detection methods, refining the system for FP reduction, and including a feature selection step are being considered for future work.

References

[1] Bick U, Giger ML, Li L, Qian W, Clarke LP. Digital mammography: computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms. Acad Radiol 1997;4(11):724–31.
[2] Bick U, Giger ML, Schmidt RA, Nishikawa RM, Doi K. Density correction of peripheral breast tissue on digital mammograms. Radiographics 1996;16(6):1403–11.
[3] Caldwell CB, Stapleton SJ, Holdsworth DW, Jong RA, Weiser WJ, Cooke G. Model-based detection of spiculated lesions in mammograms. Phys Med Biol 1990;35(2):235–47.
[4] Cancer Research UK. Breast cancer factsheet; February 2004.
[5] Christoyianni I, Dermatas E, Kokkinakis G. Fast detection of masses in computer-aided mammography. IEEE Signal Process Mag 2000;17(1):54–64.
[6] Christoyianni I, Koutras A, Dermatas E, Kokkinakis G. Computer aided diagnosis of breast cancer in digitized mammograms. Comput Med Imaging Graph 2002;26(5):309–19.
[7] Di Maggio C. State of the art of current modalities for the diagnosis of breast lesions. Eur J Nucl Med Mol Imaging 2004;31(Suppl. 1):S56–69.
[8] Ferrari RJ, Rangayyan RM, Desautels JEL, Borges RA, Frère AF. Automatic identification of the pectoral muscle in mammograms. IEEE Trans Med Imaging 2004;23(2):232–45.
[9] Ferrari RJ, Rangayyan RM, Desautels JEL, Borges RA, Frère AF. Identification of the breast boundary in mammograms using active contour models. Med Biol Eng Comput 2004;42(2):201–8.
[10] Gonzalez RC, Woods RE, Eddins SL. Digital image processing using MATLAB. New Jersey, USA: Prentice Hall; 2004.
[11] Karssemeijer N. Automated classification of parenchymal patterns in mammograms. Phys Med Biol 1998;43(2):365–78.
[12] Kegelmeyer WP, Pruneda JM, Bourland PD, Hillis A, Riggs MW, Nipper Jr ML. Computer-aided mammographic screening for spiculated lesions. Radiology 1994;191(2):331–7.
[13] Kobatake H, Murakami M, Takeo H, Nawano S. Computerized detection of malignant tumors on digital mammograms. IEEE Trans Med Imaging 1999;18(5):369–78.
[14] Laws KI. Textured image segmentation. Ph.D. thesis. Los Angeles, California, USA: University of Southern California; 1980.
[15] Liu S, Babbs CF, Delp EJ. Multiresolution detection of spiculated lesions in digital mammograms. IEEE Trans Med Imaging 2001;10(6):874–84.
[16] MATLAB. Using MATLAB. Natick, MA, USA: The MathWorks Inc.; 2006.
[17] Morrow WM, Paranjape RB, Rangayyan RM, Desautels JEL. Region-based contrast enhancement of mammograms. IEEE Trans Med Imaging 1992;11(3):392–406.
[18] Mudigonda NR, Rangayyan RM, Desautels JEL. Detection of breast masses in mammograms by density slicing and texture flow-field analysis. IEEE Trans Med Imaging 2001;20(12):1215–27.
[19] National Cancer Institute of Canada. Canadian cancer statistics 2006; 2006.
[20] Nixon M, Aguado A. Feature extraction & image processing. Oxford, UK: Newnes; 2001.
[21] Petrick N, Chan H-P, Sahiner B, Wei D. An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. IEEE Trans Med Imaging 1996;15(1):59–67.
[22] Petrick N, Chan H-P, Wei D, Sahiner B, Helvie MA, Adler DD. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996;23(10):1685–95.
[23] Petrick N, Sahiner B, Chan H-P, Helvie MA, Paquerault S, Hadjiiski LM. Breast cancer detection: evaluation of a mass-detection algorithm for computer-aided diagnosis—experience in 263 patients. Radiology 2002;224(1):217–24.
[24] Polakowski WE, Cournoyer DA, Rogers SK, De Simio MP, Ruck DW, Hoffmeister JW. Computer-aided breast cancer detection and diagnosis of masses using difference of gaussians and derivative-based feature saliency. IEEE Trans Med Imaging 1997;16(6):811–9.
[25] Qian W, Li L, Clarke LP. Image feature extraction for mass detection in digital mammography: influence of wavelet analysis. Med Phys 1999;26(3):402–8.
[26] Rangayyan RM, Shen L, Shen Y, Desautels JEL, Bryant H, Terry TJ. Improvement of sensitivity of breast cancer diagnosis with adaptive neighborhood contrast enhancement of mammograms. IEEE Trans Inf Technol Biomed 1997;1(3):161–70.
[27] Sheshadri HS, Kandaswamy A. Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms. Comput Med Imaging Graph 2007;31(1):46–8.
[28] Suckling J, Parker J, Dance D, Astley S, Hutt I, Boggis C. The mammographic image analysis society digital mammogram database. In: Exerpta Medica, International Congress Series, vol. 1069. 1994. p. 375–8.
[29] Zheng L, Chan AK. An artificial intelligent algorithm for tumor detection in screening mammogram. IEEE Trans Med Imaging 2001;20(7):559–67.
[30] Zwiggelaar R, Parr TC, Schumm JE, Hutt IW, Taylor CJ, Astley SM. Model-based detection of spiculated lesions in mammograms. Med Image Anal 1999;3(1):39–62.

Alfonso Rojas Domínguez was born in Mexico City, Mexico, in 1978. He received the degree of BSc in telecommunications engineering from the National Autonomous University of Mexico (UNAM) in 2002, the degree of MSc with distinction in intelligence engineering from the University of Liverpool, UK, in 2003, and the degree of PhD in electrical engineering an electronics, from the same university, in 2007. His research interests include automated image processing and analysis, machine learning, pattern recognition and genetic programming.

Prof. Asoke K Nandi received the degree of PhD from the University of Cambridge (Trinity College), UK, in 1979. He joined the Imperial College, UK, as a lecturer in 1987. In 1991, he joined the Electronic and Electrical Eng. Dept. in the University of Strathclyde, UK, as a senior lecturer; he was appointed a reader in 1995 and a professor in 1998. In March 1999, he moved to the University of Liverpool, UK. Currently, he is the Head of the Signal Processing

and Communications Research Group with interests in the areas of nonlinear and non-Gaussian signal processing, communications and machine learning research. With his group he has been carrying out research in blind source separation, development and applications of machine learning, machine condition monitoring, signal modelling, communication signal processing, and biomedical signal processing. Prof. Nandi is the author of two books—Automatic Modulation Recognition of Communications Signals (Boston, MA: Kluwer Academic, 1996) and Blind Estimation Using Higher-Order Statistics (Boston, MA: Kluwer Academic, 1999).