Boundary modelling and shape analysis methods for classification of mammographic masses

R. M. Rangayyan[1,2] N. R. Mudigonda[1] J. E. L. Desautels[1,3]

[1]Department of Electrical and Computer Engineering , University of Calgary, Calgary, Canada
[2]Department of Radiology, University of Calgary, Calgary, Canada
[3]Alberta Cancer Board, Calgary, Canada

Abstract—*The problem of computer-aided classification of benign and malignant breast masses using shape features is addressed. The aim of the study is to look at the exceptions in shapes of masses such as circumscribed malignant tumours and spiculated benign masses which are difficult to classify correctly using common shape analysis methods. The proposed methods of shape analysis treat the object's boundary in terms of local details. The boundaries of masses analysed using the proposed methods were manually drawn on mammographic images by an expert radiologist (JELD). A boundary segmentation method is used to separate major portions of the boundary and to label them as concave or convex segments. To analyse the shape information localised in each segment, features are computed through an iterative procedure for polygonal modelling of the mass boundaries. Features are based on the concavity fraction of a mass boundary and the degree of narrowness of spicules as characterised by a spiculation index. Two features comprising spiculation index (SI) and fractional concavity (f_{cc}) developed in the present study when used in combination with the global shape feature of compactness resulted in a benign/malignant classification accuracy of 82%, with an area (A_z) of 0.79 under the receiver operating characteristics (ROC) curve with a database of the boundaries of 28 benign masses and 26 malignant tumours. SI alone resulted in a classification accuracy of 80% with A_z of 0.82. The combination of all the three features achieved 91% accuracy of circumscribed versus spiculated classification of masses based on shape.*

Keywords—*Mammography, Breast cancer, Breast masses, Shape analysis, Concavity, Convexity, Spiculation index, Tumour classification*

Med. Biol. Eng. Comput., 2000, **38**, 487–496

1 Introduction

MANY STUDIES have shown that early detection through periodic mammographic screening of asymptomatic women can reduce breast cancer mortality. In suspicious cases, the radiologist may recommend a biopsy. Recent studies have shown that mammography is sensitive in cancer screening and diagnosis, but with a high false-positive rate (STERNS, 1996). The true-positive and false-positive rates of mammography vary in different age groups; the sensitivity of mammography is higher in women over the age of 50 years (MUSHLIN *et al.*, 1998). Considering the traumatic nature and cost of biopsy, it is desirable to develop computer-based methods to distinguish accurately between benign masses and malignant tumours. Such methods may help in performing initial screening or second reading of mammograms, and may lend objective tools to help radiologists in analysing difficult cases and in deciding on biopsy recommendations.

Computer detection of breast tumours is difficult due to the nature of mammographic images and features; in the present study we focus on the equally important problem of their classification as benign or malignant based on their morphology after boundaries have been delineated manually by an expert radiologist (JELD) specialised in mammography. The objective is to develop shape-based features that can closely track shape differences between benign masses and malignant tumours as perceived by the radiologist.

Breast abnormalities present varying diagnostic information on mammograms. The diagnostic features vary in terms of shape, density, and textural content. Most benign masses are homogeneous and possess well-defined edges; malignant tumours typically have fuzzy or ill-defined boundaries. Benign masses possess smooth, round, or oval shapes with possible macrolobulations, as opposed to malignant tumours which typically exhibit rough contours with microlobulations, spiculations, and concavities. Many studies have therefore focused on analysing the shapes of mammographic masses.

Correspondence should be addressed to Prof. R. M. Rangayyan; email:ranga@enel.ucalgary.ca

First received 21 December 1999 and in final form 18 May 2000
MBEC online number: 20003501

ACKERMAN and GOSE (1972) analysed breast lesions on xeroradiograms and investigated the use of four measures of malignancy: calcification, spiculation, roughness, and area-to-perimeter ratio. Their spiculation and roughness measures required the location of the centre of the lesion as a reference point for computing radial projections. The centre of the lesion was simply defined as the average position of the left-to-right and top-to-bottom borders of the rectangle bounding the lesion. Given a suspicious area on a xeroradiogram, their computer-aided classification methods obtained similar operational characteristic curves to those of the radiologist. In another study on xeroradiograms, ACKERMAN *et al.* (1973) used 36 radiographic properties of lesions to estimate the probability of malignancy. Using the properties in an automated clustering scheme, they achieved a false-negative (FN) rate of zero at a false-positive (FP) rate of 45%.

POHLMAN *et al.* (1995) used measures of tumour circularity and surface roughness to classify breast tumours. By using a logistic regression model, they reported an area (A_z) of 0.9 under the receiver operating characteristics (ROC) curve. The shape features were based on the radial distances of a mass boundary from its centroid. The surface roughness was calculated as the percentage of angles with multiple boundary points. KILDAY *et al.* (1993) developed a set of seven shape features based on tumour circularity and radial distance measures (RDM) from the centroid to the points on the boundary. The features included compactness, mean of RDM, standard deviation of RDM, entropy of the RDM histogram, area ratio, zero crossings, and boundary roughness. A three-group classification of breast tumours as fibroadenomas, cysts, and cancers was performed by using the features in a linear discriminant function. They reported a classification accuracy of 51% using the leave-one-out method.

The measures to characterise surface roughness based on the centroid of a mass (POHLMAN *et al.*, 1995; KILDAY *et al.*, 1993) work well with a generally round boundary. However, in the case of complex shapes, the centroid may lie outside the tumour region and may not be a valid point to measure the distances to the boundary.

BRUCE and KALLERGI (1999) studied the effect of the resolution of the images in detection and classification of mammographic mass shapes as round, lobular, or irregular, using the same shape features as proposed by KILDAY *et al.* (1993) along with wavelet-based scalar energy features. Markov random field methods were employed to extract mass regions. Features computed from the regions extracted in images of two different resolutions (220 μm and 180 μm) resulted in similar classification trends. The best overall classification rate of 75.9% was reported by using wavelet-based features computed from manually segmented mass regions.

RANGAYYAN *et al.* (1997) used moments of distances of contour points from the centroid, compactness of the boundary, Fourier descriptors, and chord-length statistics to characterise the roughness of tumour boundaries. While circumscribed versus spiculated classification of masses was achieved at accuracies of up to 94.4%, benign versus malignant classification accuracy obtained by using only shape factors based upon contours was limited to about 76%, with a database of 54 masses.

HUO *et al.* (1995) extracted mass regions using region growing methods and proposed two spiculation measures obtained from an analysis of radial edge gradient information surrounding the periphery of the extracted regions. Classification studies performed using the features yielded an area of 0.85 under the ROC curve. POLAKOWSKI *et al.* (1997) developed a model-based vision algorithm using difference-of-Gaussian filters to detect masses and computed nine features based on size, circularity, contrast, and Laws texture features. A multilayer perceptron neural network was used for the classification of breast masses as benign or malignant. They reported a detection sensitivity of 92% for identifying malignant masses with 1.8 false positives per image. BRZAKOVIC *et al.* (1990) proposed an automated detection and Bayesian classification scheme using tumour size, shape, and intensity changes in extracted regions. Other methods developed to detect distortions in mammographic images as a result of the presence of masses have also included steps to follow the morphological signs or orientations of masses during various stages of detection (KARSSEMEIJER, 1995; KEGELMEYER, 1993; TARASSENKO *et al.*, 1995).

KOBATAKE and YOSHINAGA (1996) developed skeleton analysis methods using the iris filter to detect spicules of lesions. Further, they used a modified Hough transform to extract radiating lines from the centre of the mass region to discriminate between star-shaped malignant tumours and non-malignant masses. Masses with less than three spicules were classified to be non-star-shaped. They reported a classification accuracy of 74% in detecting malignant tumours. It should be noted that the classification accuracies reported in many of the studies reviewed above were obtained using limited databases. Other studies on detection and analysis of breast tumours have quantified features based on tumour edge sharpness (RICHTER and CLARIDGE, 1991; HUO *et al.*, 1995; RANGAYYAN *et al.*, 1997; SAHINER *et al.*, 1998) and textural information (KOK *et al.*, 1994; PETROSIAN *et al.*, 1994; SAHINER *et al.*, 1996; WEI *et al.*, 1997); these methods will not be reviewed here as the scope of the present work is limited to analysing the shape of the mass contour or boundary.

Most of the shape analysis methods that have been applied to breast mass discrimination have focused on computing global measures characterising the boundary's shape. Such methods are relatively insensitive to important local changes manifested by spicules and microlobulations. While the majority of benign masses on mammograms are well circumscribed, some do possess stellate or spiculated distortions. On the other hand, while most malignant tumours are spiculated, some circumscribed malignant tumours are also encountered. The present study concentrates on the development of methods that can characterise the above-mentioned exceptions in shapes of masses and classify them appropriately. Discrimination between the microlobulations in malignant tumours and the macrolobulations of benign masses requires detailed analysis of local characteristics of mass boundaries. Recently, MENUT *et al.* (1998) performed parabolic modelling of tumour boundaries and used the narrowness and width of individual parabolic segments for classification. A benign/malignant classification accuracy of 76% was achieved with a database of 54 cases (28 benign and 26 malignant) using four features computed based on the mean and variance values of the narrowness and width of the parabolic segments.

We use the same data set in the present study as that used by MENUT *et al.* (1998) and RANGAYYAN *et al.* (1997), and propose a method of shape analysis that treats a mass boundary as a union of piece-wise continuous and locally-salient concave and convex parts (MUDIGONDA *et al.*, 1999). A convex part is defined as a segment of the boundary that encloses a portion of the mass, while a concave part is one formed by the presence of a background region within the segment. In order to identify and analyse the concavities and convexities possessed by a mass boundary in a localised manner, we propose methods based on contour segmentation and polygonal modelling. The segments are used to compute the angles subtended by the individual concave and convex parts at their vertices. This information is used in characterising the narrowness and depth of spiculations or indentations.

2 Boundary analysis and modelling

2.1 *Segmentation of concave and convex parts*

The first step in the proposed method is to identify and label the prominent concave and convex parts of a mass boundary. For this purpose, the mass boundary is segmented into a set of piecewise continuous curves by locating points of inflection on the boundary.

Let $\vec{U}_\circ = (x(n), y(n))$, $n = 0, 1, 2, \ldots, N-1$, be the vector of the (x, y) coordinates of N points on the boundary of a mass. The points of inflection on the boundary are defined by

$$\vec{U}_\circ{}' \times \vec{U}_\circ{}'' = 0$$
$$\vec{U}_\circ{}' \times \vec{U}_\circ{}''' \neq 0 \tag{1}$$

where $\vec{U}_\circ{}'$, $\vec{U}_\circ{}''$, and $\vec{U}_\circ{}'''$ are the first, second, and third derivatives of $\vec{U}_\circ$. Solving (1) is equivalent to solving the system of equations given by

$$x''(n)y'(n) - y''(n)x'(n) = 0$$
$$x'(n)y'''(n) - x'''(n)y'(n) \neq 0 \tag{2}$$

where $x'(n)$, $y'(n)$, $x''(n)$, $y''(n)$, $x'''(n)$, and $y'''(n)$ are the first, second, and third derivatives of $x(n)$ and $y(n)$, respectively.

The portions of a mass boundary between successive inflection points were modelled as parabolas in the work of MENUT *et al.* (1998), who described and solved the following problem. Difficulty lies in boundary segmentation because the contours of masses are, in general, not smooth. Many false or irrelevant inflection points appear on relatively straight parts of a tumour boundary when $x''(n)$ and $y''(n)$ are not far from zero. To solve this problem, derivatives at each contour point were computed by considering weighted and averaged differences of a certain number of pairs of points on either side of the point under consideration. Variable numbers of pairs of points were used to compute derivatives that resulted in varying numbers of inflection points for a given contour. A hysterisis procedure was then applied to the resulting data, plotted as the number of inflection points detected against the number of pairs of differences used, to determine the optimal number of pairs of differences that provides the most appropriate inflection points. In practice, this is equivalent to locating the first straight segment on the curve that indicates a fixed number of inflection points over a range of number of pairs of differences used; details of this procedure may be found in MENUT *et al.* (1998).

After performing segmentation of the boundary as explained above, the individual segments between successive inflection points are labelled as concave or convex parts. A convex part is defined as a segment of the boundary that encloses a portion of the mass, while a concave part is one formed by the presence of a background region within the segment. Fig. 1 shows a 630×730-pixel section of a mammogram (pixel size = 50 μm) with a circumscribed benign mass overlaid with the contour drawn by a radiologist specialised in mammography (JELD); the black and white portions represent the concave and convex parts, respectively. Fig. 2 shows the result of concavity/convexity labelling of the boundary of a spiculated malignant tumour (a 600×770-pixel section of a mammogram with pixel size = 62 μm).

Once the boundary of a mass is identified in terms of its concave and convex parts, the fractional concavity of the mass boundary is computed for use as a feature in pattern classification. The boundaries used in the present work were manually drawn, and include artefactual, minor modulations that could lead to inefficient representation for pattern classification. A polygonal modelling procedure is proposed in the following subsection to extract features that can effectively characterise the shape complexities represented by the individual concave and convex parts identified as above.

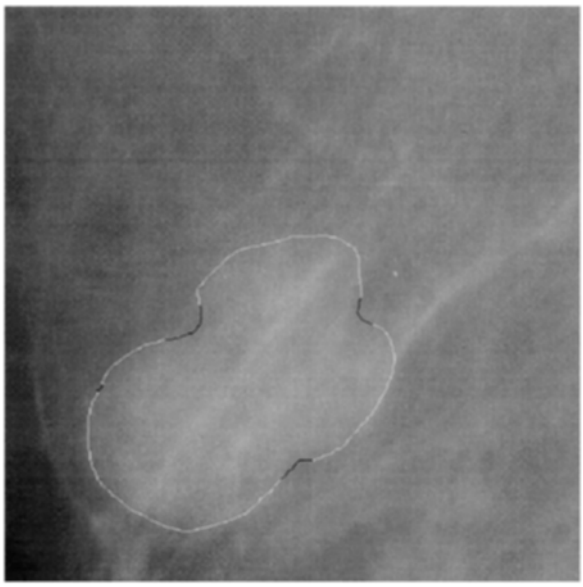

Fig. 1 *Concave and convex parts of the boundary of a benign mass. The concave parts are shown in black and the convex parts in white. The image size is 630 × 730 pixels or 31.5 × 36.5 mm with a pixel size of 50 μm*

2.2 *Polygonal modelling of mass boundaries*

VENTURA and CHEN (1992) presented an algorithm for segmenting two-dimensional curves in which the number of segments is prespecified to initiate the process, in relation to the complexity of the shape. This is not a desirable step when dealing with the complex or spiculated shapes of breast tumours (RANGAYYAN *et al.*, 1997). In the present work, the polygon formed by the points of inflection is used as the initial input to the polygonal modelling procedure. This step helps in

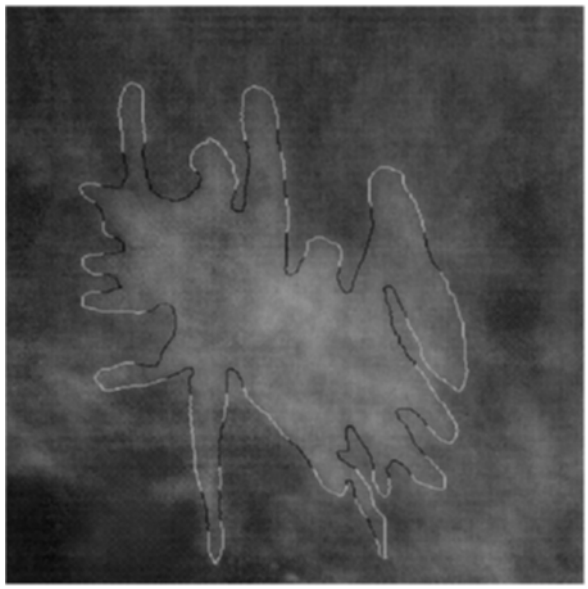

Fig. 2 *Concave and convex parts of the boundary of a spiculated malignant tumour. The concave parts are shown in black and the convex parts in white. The image size is 600 × 770 pixels or 37.2 × 47.7 mm with a pixel size of 62 μm*

automating the polygonalisation algorithm: the proposed method does not require any interaction from the user in terms of the starting number of segments.

Given an irregular boundary **C** as specified by the set of its x and y co-ordinates, the polygonal modelling algorithm starts by dividing the boundary into a set of piecewise continuous curved parts by locating the points of inflection on the boundary as explained in the preceding subsection. Each segmented curved part is represented by a pair of linear segments based on its arc-to-chord deviation. The procedure is iterated subject to predefined boundary conditions so as to minimise the error between the length of the initial boundary and the cumulative length computed from the polygonal segments.

The details of the modelling algorithm developed for application to mammographic mass boundary analysis are as follows. Let $\mathbf{C} = (x(n), y(n))$, $n = 0, 1, 2, \ldots, N-1$, be the boundary characterised by the (x, y) co-ordinates of N points. Let SC_{ij} ($SC_{ij} \in \mathbf{C}$), $i = 1, 2, \ldots, M$, be M curved parts, each of them containing a set of boundary points, at the start of the jth iteration, such that $SC_{1j} \bigcup SC_{2j} \bigcup \ldots \bigcup SC_{Mj} \equiv \mathbf{C}$.

1. In each curved part represented by SC_{ij}, the arc-to-chord distance is computed for all the points, and the point on the curve with the maximum arc-to-chord deviation (d_{max}) is located.
2. If $d_{max} \geqslant 0.25$ mm (5 pixels in our images with a pixel size of 50 μm), the curved part is segmented at the point of maximum deviation to approximate the same with a pair of linear segments, irrespective of the length of the resulting linear segments.
   If $0.1 \text{ mm} \leqslant d_{max} < 0.25$ mm, the curved part is segmented at the point of maximum deviation subject to the condition that the resulting linear segments satisfy a minimum-length criterion, which is specified as 1 mm.
   If $d_{max} < 0.1$ mm, the curved part (SC_{ij}) is considered to be almost linear and is not segmented any further.
3. After performing steps 1 and 2 on all the curved parts of the boundary available in the current jth iteration, the resulting vector of the polygon's vertices is updated, after sorting and rearranging in clockwise order.
4. If the number of polygonal segments following the jth iteration equals that of the previous $(j-1)$th iteration, the algorithm is considered to have converged and the polygonalisation process is terminated. Otherwise, the procedure (steps 1 to 3) is repeated until the algorithm converges.

The criterion for choosing the threshold for arc-to-chord deviation is based on the assumption that any segment possessing less deviation is insignificant in terms of its degree of spiculation for further analysis (as indicated by the radiologist involved in this work). Fig. 3 shows the points of inflection (denoted by *), and the initial stage of polygonal modeling (thick straightline segments) of the boundary of the spiculated malignant tumour in Fig. 2. Fig. 4 shows the final result of polygonal modelling of the same boundary. The algorithm converged after four iterations, as shown in the convergence plot of Fig. 5. The final result of application of the polygonal modelling algorithm to the circumscribed benign mass boundary in Fig. 1 is shown in Fig. 6.

The number of linear segments required for the approximation of a boundary increased with its shape complexity, in the range 20–400 for the database used. However, the number of iterations required for the convergence of the algorithm did not vary much for different mass boundary shapes, remaining within the range 3–5. This is due to the fact that the relative complexity of the boundary to be segmented is taken into consideration during the initial pre-processing step of locating the points of inflection; hence the subsequent polygonalisation process is robust and computationally efficient. The algorithm performed well and

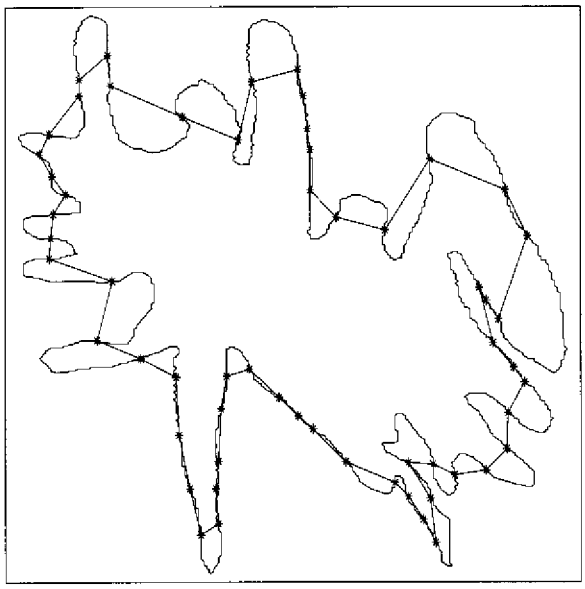

Fig. 3 *Points of inflection (as indicated by *) on the boundary of the spiculated malignant tumour shown in Fig. 2. Thick lines represent the initial polygonal approximation. Number of sides = 58*

delivered satisfactory results on various irregular shapes of spiculated cases of benign and malignant masses in our database. The techniques proposed can successfully divide a tumour's boundary into concave and convex parts and further approximate each part by a set of linear segments leading to a polygonal model of the whole boundary.

RANGAYYAN *et al.* (1997) developed a region-based edge profile acutance measure for classification of breast masses in the same database. To compute acutance, a version of the polygonalisation algorithm proposed by VENTURA and CHEN (1992) was implemented so as to require user input for the number of segments to use in the model. The algorithm proposed in the present paper is completely automated. The result is used for the extraction of shape-based features, as described in the following section.

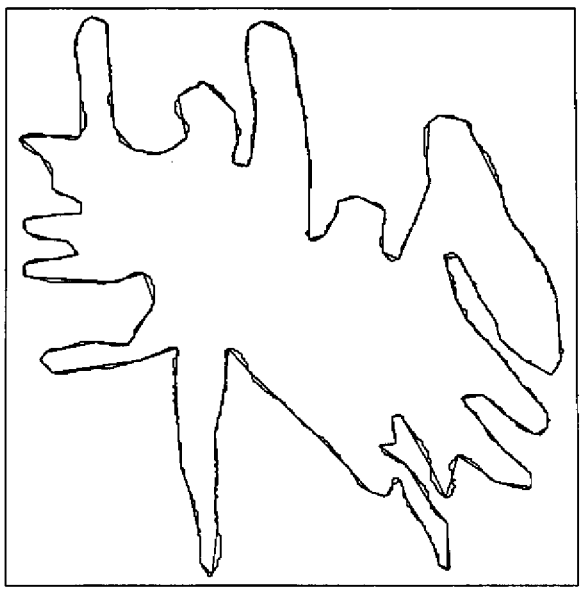

Fig. 4 *Final result of polygonal modelling of the boundary of the spiculated malignant tumour shown in Fig. 2. Number of sides = 146*

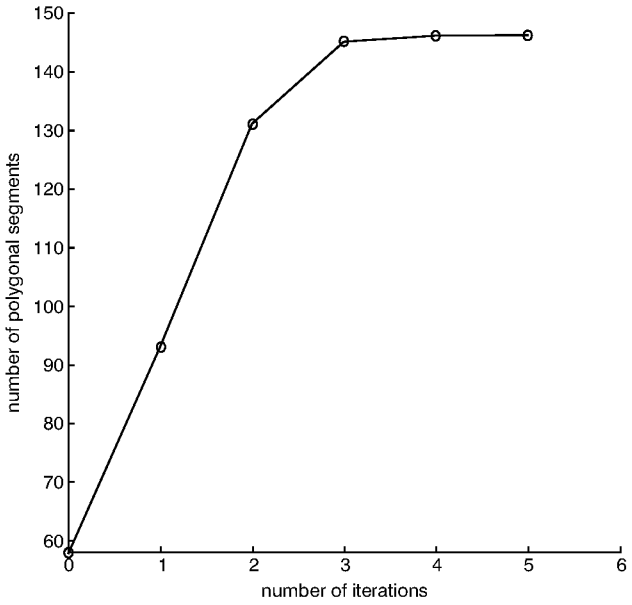

Fig. 5 *Convergence plot of the iterative polygonal modelling procedure for the spiculated malignant tumour in Fig. 2*

3 Feature extraction

The features computed in the present work from the individual segments of a mass boundary consist of measures representing the concave or convex fraction of the total boundary length and an index of spicularity.

3.1 *Concavity and convexity fractions*

Most benign mass boundaries have major portions of convex macrolobulations. Some benign masses may have minor concavities and spicules. On the other hand, malignant tumours typically possess both concave and convex segments as well as microlobulations and prominent spicules. Fractional concavity and fractional convexity values of each mass boundary were computed to characterise and quantify these properties.

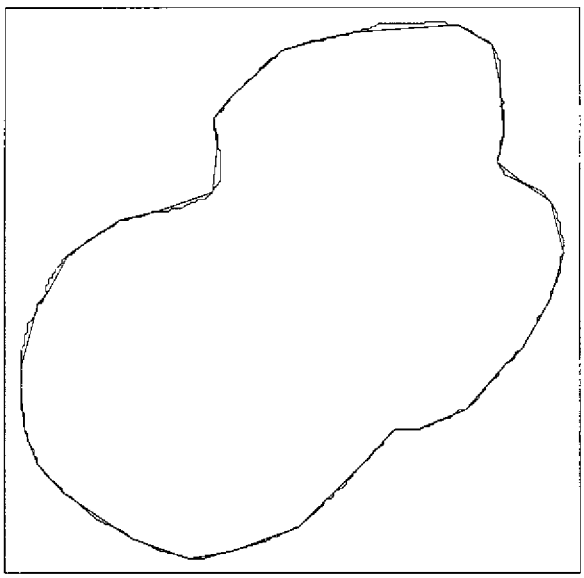

Fig. 6 *Final result of polygonal modelling of the boundary of the benign mass shown in Fig. 1. Initial number of sides = 14. Final number of sides = 36. Number of iterations = 4*

Let S_i, $i = 1, 2, 3, \ldots, M$, be the lengths of M segments obtained from the polygonal model of a mass boundary. The total length of the boundary T_l is computed as the cumulative length of the M segments: $T_l = \sum_{i=1}^{M} S_i$. Let CC_i, $i = 1, 2, 3, \ldots, P$, be the lengths of P concave segments, and CV_i, $i = 1, 2, 3, \ldots, Q$, be the lengths of Q convex segments. The cumulative concave length (CC_l) and convex length (CV_l) are computed as $CC_l = \sum_{i=1}^{P} CC_i$, and $CV_l = \sum_{i=1}^{Q} CV_i$.

The fractional concavity (f_{cc}) and convexity (f_{cv}) values are computed by normalising the respective cumulative lengths by the total boundary length so that the sum of the concavity and convexity fractions for a given boundary equals unity: $f_{cc} = CC_l/T_l$, and $f_{cv} = CV_l/T_l$. The fractional parameters are normalised with respect to the total length of the boundary, thus making them independent of the size of the boundary. The f_{cc} and f_{cv} parameters are complementary to each other; hence, only f_{cc} is used as a feature in pattern classification experiments.

3.2 *Spiculation index*

It is known that invasive carcinomas, due to their nature of infiltration into surrounding tissues, form narrow, stellate distortions at their boundaries. Based on this feature, we propose a spiculation index (SI) to represent the degree of spiculation of a mass boundary. In order to emphasise narrow spicules and microlobulations, we apply a function that enhances the contribution of narrow spicules in the computation of SI. For each curved part of a mass boundary or the corresponding polygonal model segment, the ratio of its length to the base width can represent its degree of narrowness or spiculation. A nonlinear weighting function is proposed based upon the segment's length S and angle of spiculation θ that delivers progressively increasing weighting with increase in the narrowness of spiculation of the segment. Initially the spicule candidates are identified as portions of the boundary delimited by pairs of successive points of inflection. The proposed polygonal modelling method is used to compute the parameters S and θ for each spicule as described by the following procedure.

If a spicule is modelled with M polygonal segments, then there exist $M - 1$ angles at the points of intersection of the successive polygonal segments. Let SL_i, $i = 1, 2, \ldots, M$, be the polygonal segments, and Θ_j, $j = 1, 2, \ldots, M - 1$, be the angles subtended. Then, the segment length (S) and the angle of narrowness (θ) of the spicule under consideration are computed as follows:

1. If $M = 1$, the portion of the boundary that has been delimited by successive points of inflection is relatively straight as can be seen in Fig. 3. Such parts are merged into the spicules that include them, thus enhancing the lengths of the corresponding spicules without affecting their angles of spiculation. The merging process discards the redundant points of inflection lying on relatively straight parts of the boundary. The above point can be verified by comparing the initial points of inflection present on the boundary in Fig. 3 with the points of inflection that are retained to compute SI in the corresponding boundary shown in Fig. 7, specifically in the spicule with the angle of spiculation labelled as 116°.
2. If $M = 2$, then the length of spicule is $S = SL_1 + SL_2$, and the angle subtended by the linear segments at the point of intersection represents the angle of narrowness (θ) of the spicule.
3. If $M > 2$, then the length of the spicule is $S = \sum_{i=1}^{M} SL_i$. Further, for estimating the angle of narrowness, an adaptive threshold is applied by using the mean of the set of angles Θ_j, $j = 1, 2, \ldots, M - 1$, as the threshold (Θ_{th}) for

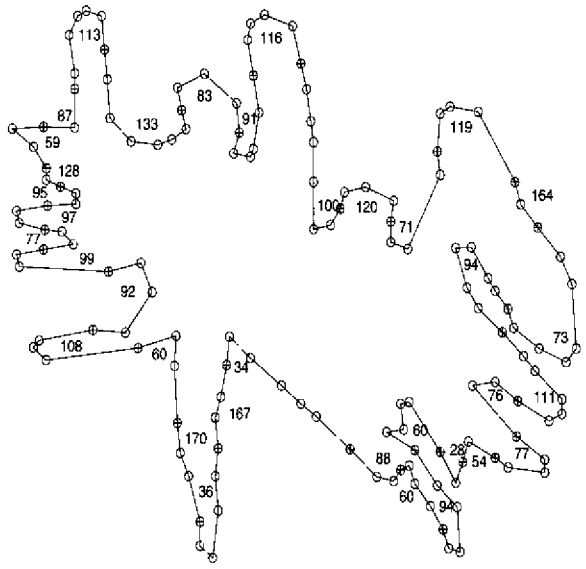

Fig. 7 *The polygonal model used in the procedure to compute SI for the spiculated malignant tumour shown in Fig. 2 (with the corresponding complete polygonal model in Fig. 4). The ⊕s correspond to the points of inflection retained to represent the starting and the ending points of spicule candidates, and ○s indicate points of intersection of linear segments within the spicules in the corresponding complete polygonal model. The numbers inside each spicule are angles in degrees as computed for the derivation of SI*

rejecting the insignificant angles. The mean of the angles that are less than or equal to Θ_{th} is considered as an estimate of the angle of narrowness of the spicule.

Fig. 8 illustrates the computation of S and θ using the above procedure for two different examples of spicules with $M = 2$ and $M = 5$, respectively.

Fig. 7 shows the spicule candidates used in the computation of SI for the boundary of the spiculated malignant tumour in Fig. 2 and the corresponding polygonal model shown in Fig. 4. The angles of spiculation computed are indicated in Fig. 7 for all the spicules. It is to be noted that the figure is not in the true scale and hence visual assessment of angles of spicules does not correspond well with the computed and labelled values. However, it can be observed that most of the angles computed for spicules depicting microlobulations are acute; on the other hand, the angles computed are obtuse for macrolobulations. The procedure described above adapts to the complexity of each spicule and delivers reliable estimates of the lengths and angles of

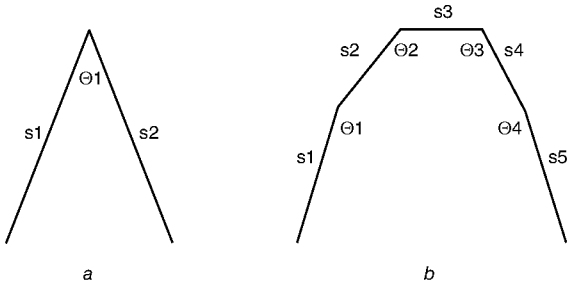

Fig. 8 *Computation of segment length S and angle of spiculation θ for two examples of spicules with number of segments M = 2 and M = 5 respectively. Θ_{th} is the threshold computed to reject insignificant angles present in the spicule with M = 5. (a) M = 2, θ = Θ1, S = s1 + s2. (b) M = 5, $\Theta_{th} = (\Theta 1 + \Theta 2 + \Theta 3 + \Theta 4)/4$, Θ2 < Θ_{th}, Θ3 < Θ_{th}, θ = (Θ2 + Θ3)/2, S = s1 + s2 + s3 + s4 + s5*

narrowness of spicules required for computing SI, following polygonal modelling. The computation of SI for a given mass boundary is described as follows.

Let S_i and θ_i, $i = 1, 2, 3, \ldots, N$, be the length and angle of N sets of polygonal model segments corresponding to the N spicules of a mass boundary. Then, SI is computed as

$$SI = \frac{\sum_{i=1}^{N}(1 + \cos\theta_i)S_i}{\sum_{i=1}^{N} S_i} \tag{3}$$

The factor $(1 + \cos\theta_i)$ in eqn 3 modulates the length of each segment (possible spicule) according to its narrowness. Spicules with narrow angles between 0° and 30° get higher weighting, as compared to macrolobulations which usually form obtuse angles and hence get lesser weighting.

The majority of the angles of spicules of the masses in our database, computed by using the procedure described above, were found to be in the range of 30° to 150°. The function $(1 + \cos\theta)$ in eqn 3 is progressively decreasing within this range, giving reduced weighting to segments with larger angles of spiculation. Relatively flat segments having angles ranging between 150° and 180° receive the least weighting, and hence are treated as the least significant segments.

The denominator in eqn 3 serves as a normalisation factor to take into account the effect of the size of the contour; it ensures that SI represents only the severity of the spiculated nature of the mass boundary, which in turn may be linked to the invasive properties of the mass or tumour under consideration. Thus circumscribed mass shapes should have lower SI values, and sharp, stellate shapes with acute spicules should have higher SI values.

3.3 *Compactness*

To compare the performance of the proposed features with a global shape measure, the compactness of each tumour boundary was computed. Compactness is a simple measure of the efficiency of a contour to contain a given area, and is commonly defined as P^2/A, where P and A are the contour perimeter and area, respectively. A malignant tumour with a number of concavities or spicules could be expected to possess a higher complexity value than a smooth and round benign mass. To restrict the range of the parameter to (0,1) and to obtain increasing values with increase in complexity of the shape, a modified measure of complexity was used, given by (SHEN *et al.*, 1993; RANGAYYAN *et al.*, 1997) $C = 1 - 4\pi A/P^2$. With this expression, $C = 0$ for a circle, and increases with the complexity of the contour to a maximum value of 1. A previous study of our group (RANGAYYAN *et al.*, 1997) has indicated that the parameter can provide benign versus malignant classification accuracy comparable to those given by other shape factors based on Fourier descriptors, moments, etc. (with the database used in the present study). Hence C was chosen for comparative analysis in the present study.

4 Image database

Thirty-nine mammographic images including 16 circumscribed benign, four circumscribed malignant, 12 spiculated benign, and seven spiculated malignant masses were selected from the Mammographic Image Analysis Society (MIAS, UK) database (SUCKLING *et al.*, 1994). The spatial resolution of the images is 50 μm × 50 μm. Although most of the malignant tumours encountered in mammography are spiculated and a majority of the benign masses are well circumscribed, the MIAS

database has a relatively large number of spiculated benign cases. To augment the numbers of the two types of malignant tumours, 15 images (containing three circumscribed and 12 spiculated tumours) from Screen Test: Alberta Program for the Early Detection of Breast Cancer, digitised at a spatial resolution of 62 μm × 62 μm, were added to the study. This increases the number of cases examined to 54, with 28 benign and 26 malignant cases, belonging to the circumscribed and spiculated categories. All diagnoses were proven by pathologic examination of resected tissue. Sections of interest of the mammographic images were displayed on a Sun SPARC Station 2 and the boundary of each mass was traced and input to the computer by an expert radiologist specialised in mammography (JELD) using *XPAINT* for *X – Windows* (RANGAYYAN *et al.*, 1997).

5 Results of pattern classification

For each mass or tumour in the database we computed the C, f_{cc}, and SI parameters. Fig. 9 shows the scatter plot of the three features. All the three features have the distinction of reflecting lesser values for circumscribed benign cases and higher values for spiculated malignant cases. The computed features in the BMDP 7M step-wise discriminant analysis program (BROWN and ENGELMAN, 1988) were used to perform pattern classification. The program realises a jack-knife validation procedure using the leave-one-out algorithm. The pattern classification task was carried out in two parts using the proposed features. In the first part, benign versus malignant classification of masses (that is perceived to be significant by radiologists) was performed. In the second part, circumscribed versus spiculated classification of masses was performed. For the first part, the performance of the features was validated using ROC methodology. ROC plots were obtained using the BMDP package by varying the cut points for benign and malignant prior probabilities between 0 and 1 in steps of 0.1. The procedure does not affect the discriminant ratings of the variables and influences only the computation of the constant term in the discriminant function, thus resulting in varying classification accuracies. The area (A_z) under the ROC curve was computed using the trapezoidal rule.

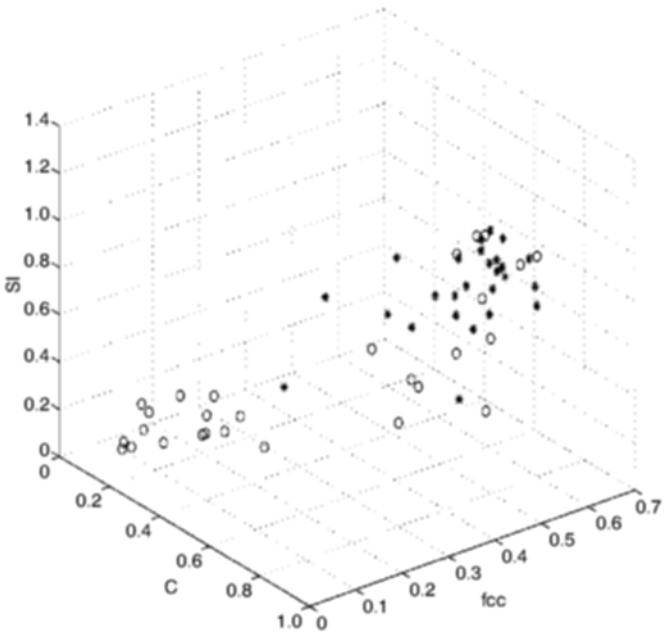

Fig. 9 *Three-dimensional feature-space plot: ○ for benign and * for malignant, SI–spiculation index, f_{cc}–fractional concavity, and C–modified compactness*

5.1 *Benign versus malignant classification*

Table 1 provides details on the benign versus malignant classification performance of various combinations of the proposed features. All of the features, individually and in different combinations, could effectively discriminate circumscribed benign masses from spiculated malignant tumours. The parameter C, being a global measure of shape complexity, failed to classify almost all spiculated benign masses, and classified four out of seven circumscribed malignant tumours correctly.

Concavity analysis is sensitive to the presence of spicules in a mass boundary. Fractional concavity (f_{cc}) increases with the number of spicules and their length; however, it does not take into account the degree of spicularity of the individual spicules present in the boundary. Circumscribed malignant tumours inherently contain a large number of microlobulations in their boundaries that could look like alternating concave and convex segments; hence f_{cc} could distinguish six out of the seven circumscribed malignant cases correctly, and resulted in a high sensitivity of 88.5%. However, f_{cc} does not look into the characteristics of the spicules in terms of their depth and narrowness; hence it failed to distinguish spiculated benign masses from spiculated malignant tumours and resulted in a poor specificity of 60.7% with $A_z = 0.75$.

Since the degree of spiculation of boundary segments is characterised by the proposed spiculation index, a significant portion of the spiculated benign cases (5 out of 12) with large convexities and a few narrow spicules were correctly classified as benign by SI. SI provided an improved specificity of 75% in classifying the benign masses in the database used. This is an encouraging result with the present methods as none of the other shape parameters developed in the past could be effective in separating the spiculated benign cases in the MIAS database. SI, however, failed in four of the seven circumscribed malignant cases (although with narrow classification margins), and resulted in the best classification accuracy of all combinations of features with $A_z = 0.82$. The misclassified cases, although malignant, do not have prominent spicules in their boundaries. It can be observed from Table 1 that combining SI with the other features generally yielded improved classification results with high values of areas under the ROC curves, some of which are shown in Fig. 10. Since the spicules are emphasised based on their angles of narrowness, SI is particularly sensitive to the stellate or star-like distortions of malignant tumours; the detection of such distortions has been the focus of many automated detection studies (KEGELMEYER, 1993; KARSSEMEIJER, 1995; KOBATAKE and YOSHINAGA, 1996).

We obtained a benign versus malignant classification accuracy of 82%, with $A_z = 0.79$, by combining f_{cc} and SI, which are sensitive to local variations in a boundary, with the global shape feature C. A total of eight FPs out of 28 benign cases and two FNs out of 26 malignant cases were observed with the

Table 1 Number of cases correctly classified as benign or malignant by the various features studied: Circ.–circumscribed, Spic.–spiculated; C–modified compactness, f_{cc}–fractional concavity, SI–spiculation index, and A_z–area under the ROC curve

Features	Benign		Malignant		Percentage accuracy			
	Circ.	Spic.	Circ.	Spic.	Benign	Malignant	Total	A_z
SI	16/16	5/12	3/7	19/19	75.0	84.6	79.6	0.82
f_{CC}	15/16	2/12	6/7	17/19	60.7	88.5	74.1	0.75
C	15/16	1/12	4/7	19/19	57.1	88.5	72.2	0.76
C & SI	16/16	5/12	3/7	19/19	75.0	84.6	79.6	0.80
f_{cc} & SI	16/16	4/12	3/7	19/19	71.4	84.6	77.8	0.80
C & f_{cc}	15/16	1/12	5/7	19/19	57.1	92.3	74.1	0.72
C, f_{cc}, & SI	16/16	4/12	5/7	19/19	71.4	92.3	81.5	0.79

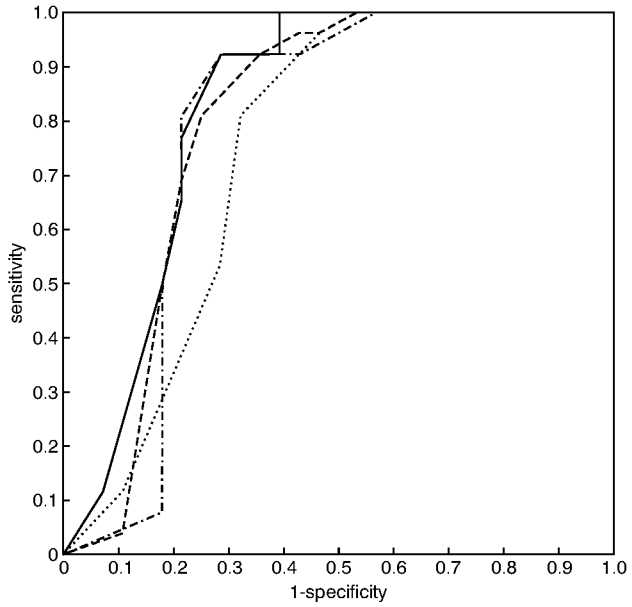

Fig. 10 *ROC plots for SI (——), f_{cc} (. . . .), combination of SI and f_{cc} (- - - -), and combination of all three features (- · - · -): SI–spiculation index, f_{cc}–fractional concavity, and C–modified compactness*

combination of all the three features. The feature *SI* of the present study has clearly outperformed all the other shape-based features developed by our group so far. In the past, using the same dataset of 54 cases, RANGAYYAN *et al.* (1997) and MENUT *et al.* (1998) reported benign/malignant classification accuracies of no more than 76% using various combinations of shape-based features (ROC studies were not performed). *SI*, f_{cc}, and *C* individually resulted in benign/malignant classification accuracies of 80%, 74%, and 72%, respectively.

In the linear discriminant analysis model, the criterion to optimise the feature weights is based on maximising the variance between the classes while minimising the variance within each class. Also, the size of the class influences the computation of the feature weights. Although our database is evenly divided between benign masses and malignant tumours (28 benign and 26 malignant), it includes an unusual proportion of spiculated benign cases from the MIAS database (12 out of 28 benign cases in a total of 54 cases). Considering the above, the performance of the proposed features could be regarded to be very good. Although the improvement of 6% achieved in the present study relates to a gain of only three cases as compared to our previous studies in a small and atypical database of 54 cases tested, the trends with high values of specificity observed in the present study could lead to significant gains on a larger database. Addition of other features based upon density variations (RANGAYYAN *et al.*, 1997; MUDIGONDA *et al.*, 2000) and textural information may result in improved benign versus malignant discrimination. Due to the exclusive nature of the present work, we are unable to make a direct comparison of the results obtained in the present study with the results of other researchers that are briefly discussed in the introduction section as part of literature review.

The focus of the present study is on benign/malignant classification of masses using the proposed shape-based features once the mass regions have been delineated either manually or by automated methods. Hence, no mammographically apparent examples of other dense regions such as fibro-glandular tissues were considered in the classification task. The proposed features have focused on characterising the shape complexity that is locally present in spicules of masses; therefore, the success of the features with the boundaries of masses that are detected by automated mass detection methods depends on the accuracy of the methods in the detection of spicules. Many automated detection algorithms (BRZAKOVIC *et al.*, 1990; HUO *et al.*, 1995; KOBATAKE and YOSHINAGA, 1996; POLAKOWSKI *et al.*, 1997) presented in the literature have focused on the detection of radiating spicules of malignant tumours and have used shape-based features to identify tumour regions. It is in this context we feel optimistic that the proposed features, when used in combination with density-based and texture-based features, could help to attain improved benign versus malignant classification accuracies.

5.2 *Circumscribed versus spiculated classification*

We verified the effectiveness of the proposed features to classify masses as circumscribed or spiculated, although such a classification is not important to radiologists; the results are presented in Table 2. The best classification accuracy of 90.7% was achieved using the combination of features *C* and f_{cc}. The combination of all the three features also obtained similar results. The global shape-based feature *C* was observed to be more effective in circumscribed versus spiculated classification of masses compared to *SI* and f_{cc} which characterise local shape complexities. RANGAYYAN *et al.* (1997) reported the best circumscribed versus spiculated classification accuracy of 94.4% using a total of nine shape-based features including *C* with the same database of 54 cases. The parameter *C* alone was reported to have achieved an accuracy of 88.9%, which was also observed in the present study.

The parameter *SI*, which was found most effective in benign/malignant classification of masses, performed poorly in circumscribed versus spiculated classification, compared to the other features. *SI* misclassified five circumscribed (all malignant) and five spiculated (all benign) cases, and achieved an accuracy of 81.5%. Visual inspection of the shapes of the misclassified masses reveals the exceptions or deviations present in the shapes of some of these masses that contradict their categorisation into the respective groups based on shape: the boundaries of some circumscribed masses contain a few spicules of moderate-to-severe nature as identified by the radiologist, and the spiculated shapes have a few major macrolobulations along with prominent spicules. *SI*, being sensitive to local variations in shape, failed to identify these shapes appropriately as circumscribed or spiculated; however, the parameter could classify some of these exceptional masses into the appropriate benign or malignant groups and thus achieved a high benign versus malignant classification accuracy of masses.

Table 2 Number of cases correctly classified as circumscribed or spiculated by the various features studied: Circ.–circumscribed, Spic.–spiculated; C–modified compactness, f_{cc}–fractional concavity, and SI–spiculation index

Features	Circ.	Spic.	Percentage accuracy		
			Circ.	Spic.	Total
C	17/23	31/31	73.9	100.0	88.9
f_{cc}	16/23	30/31	69.6	96.8	85.2
SI	18/23	26/31	78.3	83.9	81.5
C& f_{cc}	18/23	31/31	78.3	100.0	90.7
C & *SI*	17/23	31/31	73.9	100.0	88.9
f_{cc} & *SI*	17/23	27/31	73.9	87.1	81.5
C, f_{cc}, *SI*	18/23	31/31	78.3	100.0	90.7

5.3 *Sensitivity of the shape features to the presence of noise*

To study the effects of noise present in the boundary in the form of microscopic variations on the computation of the proposed shape factors (f_{cc} and SI), an experiment was conducted that artificially introduced noise in the boundaries. The boundary of a mass or tumour was initially disturbed by randomly perturbing the x and y co-ordinates of the individual points on the boundary in the range ± 4 pixels. The disturbance thus caused resulted in undesirable discontinuities in the boundary, which were eliminated by low-pass filtering the Fourier descriptors (GONZALEZ and WOODS, 1992) of the boundary data.

A closed contour can be treated as a periodic function with a period equal to the length of the boundary. Each contour with noise was represented as a set of complex numbers with the real and imaginary parts being formed by the x and y co-ordinates, respectively, of the individual boundary points. Fourier descriptors were then computed and the boundary was reconstructed after retaining harmonics only up to 30% of the maximum harmonic component present in the Fourier representation. The amplitude level at the selected cut-off frequencies was found to be about -60 dB, as observed from the corresponding magnitude spectra. The number of harmonics retained ensured that the boundary contained the intended microscopic variations while removing the discontinuities that arise following the process of causing perturbations to the boundary.

The boundaries reconstructed as above contained noise in the form of high-frequency oscillations. Regardless, the inflection points detected from the noisy boundaries by the proposed method closely matched the corresponding points detected from the corresponding original boundaries, in terms of their number and their relative locations. The result was observed to be good even with some of the highly complex boundaries of masses in our database, and enabled us to come to the conclusion that the boundary segmentation algorithm, as described in Section 2, is not sensitive to the presence of artifactual modulations in the boundary.

The experiment described above was performed on a total of 10 boundaries including five benign and five malignant cases. The boundaries were selected to represent both circumscribed and spiculated shapes in the benign and malignant categories with varying degrees of shape complexity. The proposed shape factors were computed for all the ten boundaries after introducing noise. The differences between the values computed using the boundaries with noise and the values computed using the corresponding original boundaries were analysed. The means of the differences were determined to be 8% and 13% of the corresponding means for the original 10 contours for f_{cc} and SI, respectively. Based on an analysis of the full database of 54 cases including 28 benign and 26 malignant cases, the separation between the benign and the malignant group means was observed to be 45% and 69% of the corresponding means for the 54 cases for f_{cc} and SI, respectively. Therefore, we note that variations in shape factors due to the presence of noise in the form of high-frequency oscillations may not drastically affect benign/malignant classification.

6 Conclusions

We have proposed new methods of shape analysis to analyse and classify breast tumours. The boundaries of masses used in the present study were manually drawn on digitised mammograms by an expert radiologist. The methods include boundary segmentation algorithms to identify major concave and convex portions of the boundary, and features that are sensitive to local variations in the boundary computed through a polygonal modelling algorithm. The proposed shape measures quantify the extent of the spiculated nature of the boundary and the degree of narrowness of the spicules present in the boundary. The fractional concavity (f_{cc}) measure achieved 88% sensitivity in identifying malignant tumours but resulted in a poor specificity of 61%. The spiculation index (SI) achieved the best specificity of 75% in classifying the benign masses and also yielded a good sensitivity of 85%. SI resulted in a high benign versus malignant classification accuracy with an area (A_z) of 0.82 under the receiver operating characteristics (ROC) curve. The two features in combination with a modified measure of compactness as a global measure of boundary complexity (C) resulted in the best sensitivity of 92% with a specificity of 71%. The three features together achieved an overall benign versus malignant classification accuracy of 82% with $A_z = 0.79$, an improvement of 6% over other shape features applied to the same database of 54 cases (RANGAYYAN *et al.*, 1997; MENUT *et al.*, 1998). The combination of all three features achieved 91% accuracy in circumscribed versus spiculated classification of masses based on shape. The database used in the present study includes an unusual proportion of spiculated benign cases: the proposed measures should provide better classification results with a larger database including a more typical mixture of different types of masses and tumours.

The focus of the present study has been to develop shape-based features that specifically characterise well-defined variations in shapes that are localised within the spicules present in the boundaries of benign and malignant masses of the circumscribed and spiculated categories. Inter-observer variability in drawing mass contours or computer detection of mass contours in the presence of noise in the images could lead to variations in the computed shape features. It is desirable to analyse mass boundaries drawn by several radiologists as well as boundaries detected by automatic methods in the presence of varying levels of noise; however, our team does not have the resources to conduct such a study.

Recently we have developed a method for the detection of mass boundaries (MUDIGONDA *et al.*, 1999a; 2000) using multi-resolution and density-variation-based principles. We are currently improving our mass detection algorithm to be able to detect major portions of spicules present in the boundaries of masses and intend to apply the proposed measures to classify masses based on the boundaries detected by our method. The methods will be tested with a database of digitised mammograms being acquired from Screen Test: Alberta Program for the Early Detection of Breast Cancer.

Acknowledgment—This project is supported by grants from the Alberta Heritage Foundation for Medical Research (AHFMR), the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Biomedical Engineering Graduate Programme of the University of Calgary.

References

ACKERMAN, L. V., and GOSE, E. (1972): 'Breast lesion classification by computer and xeroradiograph', *Cancer*, **30**, pp. 1025–1035

ACKERMAN, L. V., MUCCIARDI, A. N., GOSE, E. E., and ALCORN, F. S. (1973): 'Classification of benign and malignant breast tumours on the basis of 36 radiographic properties', *Cancer*, **31**, pp. 342–352

BROWN, M. B., and ENGELMAN, L. (1988): 'BMDP statistical software manual' (University of California, Berkeley, CA)

BRUCE, L. M., and KALLERGI, M. (1999): 'Effects of image resolution and segmentation method on automated mammographic mass shape classification', *Proc. SPIE*, **3661**, pp. 940–947

BRZAKOVIC, D., LUO, X. M., and BRZAKOVIC, P. (1990): 'An approach to automated detection of tumours in mammograms', *IEEE Trans. Med. Imaging*, **9**, pp. 233–241

GONZALEZ, R. C., and WOODS, R. E. (1992): 'Digital image processing' (Addison-Wesley, Reading, MA)
HUO, Z., GIGER, M. L., VYBORNY, C. J., BICK, U., LU, P., WOLVERTON, D. E., and SCHMIDT, R. A. (1995): 'Analysis of spiculation in the computerised classification of mammographic masses', *Med. Phys.*, **22**, pp. 1569–1579
KARSSEMEIJER, N. (1995): 'Detection of stellate distortions in mammograms using scale space operators' in BIZAIS, Y., BARILLOT, C., and PAOLA, P. D. (Eds): 'Information processing in medical imaging' (Kluwer Academic Publishers, Netherlands), pp. 335–346
KEGELMEYER, Jr. W. P. (1993): 'Evaluation of stellate lesion detection in a standard mammogram data set', *Int. J. Pattern Recognit. Artif. Intell.*, **7**, pp. 1477–1493
KILDAY, J., PALMIERI, F., and FOX, M. D. (1993): 'Classifying mammographic lesions using computerized image analysis', *IEEE Trans. Med. Imaging*, **12**, pp. 664–669
KOBATAKE, H., and YOSHINAGA, Y. (1996): 'Detection of spicules on mammogram based on skeleton analysis', *IEEE Trans. Med. Imaging*, **15**, pp. 235–245
KOK, S. L., BRADY, J. M., and TARASSENKO, L. (1994): 'The detection of abnormalities in mammograms' in GALE, A. G., ASTLEY, S. M., DANCE, D. R., and CAIRNS, A. Y. (Eds): 'Proc. 2nd Int. Workshop on Digital Mammography', York, 10–12 July, pp. 261–270
MENUT, O., RANGAYYAN, R. M., and DESAUTELS, J. E. L. (1998): 'Parabolic modeling and classification of breast tumours', *Int. J. Shape Modeling*, **3**, pp. 155–166
MUDIGONDA, N. R., RANGAYYAN, R. M., DESAUTELS, J. E. L., and MENUT, O. (1999a): 'Segmentation of breast masses in mammograms: a multi-resolution and hierarchical density propagation approach' in LEMKE, H. U., VANNIER, M. W., INAMURA, K., and FARMAN, A. G. (Eds): 'Proc. Computer Assisted Radiology and Surgery', Paris, June, p. 1014
MUDIGONDA, N. R., RANGAYYAN, R. M., and DESAUTELS, J. E. L. (1999): 'Concavity and convexity analysis of mammographic masses via an iterative segmentation algorithm' in MENG, M. (Ed.): 'Proc. Canadian Conf. Electrical and Computer Engineering', Edmonton, May, pp. 1489–1494
MUDIGONDA, N. R., RANGAYYAN, R. M., and DESAUTELS, J. E. L. (2000): 'Segmentation and classification of mammographic masses', *Proc. SPIE*, in Medical Imaging 2000: Image processing K. M. Hanson (Ed.) Proc. SPIE vol. 3979, pp. 55–67
MUSHLIN, A. I., KOUIDES, R. W., and SHAPIRO, D. E. (1998): 'Estimating the accuracy of screening mammography: a meta-analysis', *Am. J. Preventive Med.*, **14**, pp. 143–153
PETROSIAN, A., CHAN, H. P., HELVIE, M. A., GOODSITT, M. M., and ADLER, D. D. (1994): 'Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis', *Phys. Med. Biol.*, **39**, pp. 2273–2288
POHLMAN, S. K., POWELL, K. A., OBUCHOWSKI, N., CHILCOTE, W., and GRUNDFEST, B. S. (1995): 'Classification of breast lesions based on quantitative measures of tumour morphology'. IEEE Eng. in Med. and Biol. Soc. 17th Ann. Int. Conf., Montreal, Canada, p. 2.4.2.3
POLAKOWSKI, W. E., COURNOYER, D. A., ROGERS, S. K. DESIMIO, M. P., RUCK, D. W., HOFFMEISTER, J. W., and RAINES, R. A. (1997): 'Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency', *IEEE Trans. Med. Imaging*, **16**, pp. 811–819
RANGAYYAN, R. M., EL-FARAMAWY, N. M., DESAUTELS, J. E. L., and ALIM, O. A. (1997): 'Measures of acutance and shape for classification of breast tumours', *IEEE Trans. Med. Imaging*, **16**, pp. 799–810
RICHTER, J. H., and CLARIDGE, E. (1991): 'Extraction of quantitative blur measures for circumscribed lesions in mammograms', *Med. Inform.*, **16**, pp. 229–240
SAHINER, B., CHAN, H. P., PETRICK, N., WEI, D., HELVIE, M. A., ADLER, D. D., and GOODSITT, M. M. (1996): 'Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images', *IEEE Trans. Med. Imaging*, **15**, pp. 598–611
SAHINER, B., CHAN, H. P., PETRICK, N., HELVIE, M. A., and GOODSITT, M. M. (1998): 'Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis', *Med. Phys.*, **25**, pp. 516–526
SHEN, L., RANGAYYAN, R. M., and DESAUTELS, J. E. L. (1993): 'Detection and classification of mammographic calcifications', *Int. J. Pattern Recognit. Artif. Intell.*, **7**, pp. 1403–1416
STERNS, E. E. (1996): 'Relation between clinical and mammographic diagnosis of breast problems and the cancer/biopsy rate', *Canadian J. Surg.*, **39**, pp. 128–132
SUCKLING, J., PARKER, J., DANCE, D. R., ASTLEY, S., HUTT, J., DOGGIS, C. R. M., RICKETTS, I., STAMATAKIS, E., CERNEAZ, N., KOK, S. L., TAYLOR, P., BETAL, D., and SAVAGE, J. (1994): 'The Mammographic Image Analysis Society digital mammogram database' in GALE, A. G., ASTLEY, S. M., DANCE, D. R., and CAIRNS, A. Y. (Eds): 'Proc. 2nd Int. Workshop Digital Mammography', York, 10–12 July, pp. 375–378
TARASSENKO, L., HAYTON, P., CERNEAZ, N. J., and BRADY, M. (1995): 'Novelty detection for the identification of masses in mammograms'. Proc. 4th Int. Conf. Artificial Neural Networks, Cambridge, 26–28 June, pp. 442–447
VENTURA, J. A., and CHEN, J. M. (1992): 'Segmentation of two-dimensional curve contours', *Pattern Recognit.*, **25**, pp. 1129–1140
WEI, D., CHAN, H. P., PETRICK, N., SAHINER, B., HELVIE, M. A., ADLER, D. D., and GOODSITT, M. M. (1997): 'False-positive reduction technique for detection of masses on digital mammograms: global and local multiresolution texture analysis', *Med. Phys.*, **24**, pp. 903–914

Authors' biographies

RANGARAJ (Raj) MANDAYAM RANGAYYAN received his PhD in electrical engineering from the Indian Institute of Science, Bangalore, India, in 1980. He is a Professor with the Department of Electrical and Computer Engineering (and an Adjunct Professor of Surgery and Radiology) at the University of Calgary, Calgary, Alberta, Canada, His current research projects are on mammographic image enhancement and analysis for computer-aided diagnosis of breast cancer; region-based adaptive image processing; and knee-joint sound signal analysis for noninvasive diagnosis of cartilage pathology. He was awarded the Research Excellence Award of the Department of Electrical and Computer Engineering as well as that of the Faculty of Engineering of the University of Calgary for 1997; the Killam Resident Fellowship of the University of Calgary for 1999; and the IEEE Third Millennium Medal in 2000.

NAGA RAVINDRA MUDIGONDA received his BE (1987) from Osmania University, India, and MS (1989) from the Indian Institute of Technology, Chennai. He is currently a PhD candidate in the Department of Electrical and Computer Engineering, University of Calgary. He was associated with Philips Medical Systems India Limited from 1989 to 1997. His research interests focus on medical image processing.

JOSEPH EDWARD LEO DESAUTELS obtained his MD from the University of Ottawa, Ontario, Canada, in 1955, and completed post-graduate training in radiology at the Henry Ford Hospital, Detroit, MI. He was a Staff Radiologist at the Foothills Hospital and a Clinical Professor with the Faculty of Medicine, the University of Calgary, from 1970 to 1994. He is currently a Reference Radiologist to the Alberta Program for the Early Detection of Breast Cancer and an Adjunct Professor of Electrical and Computer Enigneering at the University of Calgary, and is interested in computer applications in mammography.