Feature Extraction from the Turning Angle Function for the Classification of Contours of Breast Tumors

Rangaraj M. Rangayyan, Denise Guliato*, Juliano Daloia de Carvalho, and Sérgio Anchieta Santiago

***Abstract*— Malignant breast tumors and benign masses appear in mammograms with different shape characteristics: the former usually have rough, spiculated, or microlobulated contours, whereas the latter commonly have smooth, round, oval, or macrolobulated contours. Features that characterize shape roughness and complexity can assist in distinguishing between malignant tumors and benign masses. Signatures of contours may be used to analyze their shapes. We propose to use the turning angle function of contours of breast masses to derive features that capture the characteristics of spicules and shape roughness as described above. We propose methods to derive an index of spiculation (SI_{TA}), index of convexity (CI_{TA}) and a measure of fractal dimension (FD_{TA}) from the turning angle function. The methods were tested with a set of 111 contours of 65 benign masses and 46 malignant tumors. Classification accuracies of 0.92, 0.93, and 0.91, in terms of the area under the receiver operating characteristics curve, were obtained with SI_{TA}, CI_{TA}, and FD_{TA}, respectively.**

I. ANALYSIS OF CONTOURS AND SIGNATURES

A. Shape analysis of breast tumors

Breast tumors and masses appear in mammograms with different shape characteristics: malignant tumors usually have rough, spiculated, or microlobulated contours, whereas benign masses commonly have smooth, round, oval, or macrolobulated contours [1], [2]. Measures that can quantitatively represent shape roughness and complexity can assist in the classification of malignant tumors and benign masses [3], [4]. Objective features of shape complexity, such as compactness (C), fractional concavity (F_{cc}), spiculation index (SI), a Fourier-descriptor-based factor (FF), fractal dimension (FD), moments, chord-length statistics, and wavelet transform modulus-maxima have been developed to distinguish benign masses from malignant tumors using pattern recognition methods for computer-aided diagnosis (CAD) of breast cancer [3], [4], [5], [6], [7], [8], [9], [10]. However, atypical cases of macrolobulated or spiculated benign masses, as well as microlobulated or well-circumscribed malignant tumors create difficulties in pattern classification [3], [4]. Regardless, in comparative analyses of several features of shape, edge-sharpness, and texture for the classification of breast masses and tumors, shape factors such as F_{cc}, FF, and SI have been observed to lead to higher classification accuracies than measures related to texture and density variation [7], [5].

Notwithstanding the relative success of measures of shape in the classification of breast tumors and masses, obtaining precise and artifact-free boundaries of masses from mammograms remains to be a difficult problem [7], [11], [12]. Computer-detected contours may be expected to contain inaccuracies and artifacts due to the limitations of the procedures for the detection and segmentation of masses in mammograms. For these reasons, some studies on shape-based discrimination of masses have been based on contours of masses drawn manually on mammograms by radiologists [3], [4], [5]. However, manually drawn contours may contain noise related to hand tremor; they are also affected by intra-observer and inter-observer variations. Sahiner et al. [7] performed a comparative evaluation of the performance of several shape and texture features using computer-detected boundaries of masses: the shape factor FF was found to be the most efficient feature for discriminating between benign masses and malignant tumors.

B. Signatures of contours

Signatures of contours may be used to analyze their shapes. The most commonly used method to transform a two-dimensional (2D) contour into a one-dimensional (1D) signature is in terms of the radial distance from each contour point to the centroid of the contour, expressed as a function of the index of the contour point. Given a contour with N points $\{x(n), y(n)\}$, $n = 1, 2, \ldots, N$, the signature $s_d(n)$ is defined as $s_d(n) = \sqrt{[x(n) - \bar{x}]^2 + [y(n) - \bar{y}]^2}$. Here, $(\bar{x}, \bar{y})$ is the centroid of the contour, with the coordinates given by the averages of the corresponding coordinates of all of the contour points. A benign mass that is round or macrolobulated will have a smooth signature [13]. On the contrary, a malignant tumor that is spiculated or microlobulated will have a rough and jagged signature [13]. The 1D signature of a contour as above may be used to derive the fractal dimension (FD) to represent the complexity of the contour [8].

Another type of signature $s_c(n)$ may be defined as $s_c(n) = x(n) + j\, y(n)$. Fourier descriptors and normalized shape factors to characterize roughness may be derived from $s_c(n)$ [13], [3].

Pohlman et al. [14] defined the signature of a given contour of a breast mass as the radial distance to the contour from its centroid, expressed as a function of the angle of the radial line in the interval $[0°, 360°]$. Such a function could

* To whom all correspondence should be addressed.

R. M. Rangayyan is with the Department of Electrical and Computer Engineering, Schulich School of Engineering, and Department of Radiology, University of Calgary, Calgary, AB, Canada T2N 1N4. ranga@ucalgary.ca

D. Guliato, J. D. de Carvalho, and S. A. Santiago are with the Faculty of Computation, Federal University of Uberlândia, Uberlândia, MG, Brazil. guliato@ufu.br

This work was supported by the Conselho Nacional Desenvolvimento Científico e Tecnológico of Brazil, and a Catalyst grant from Research Services, University of Calgary. We thank Fábio José Ayres, University of Calgary, for assistance with the ROC procedures.

be multivalued for an irregular or spiculated contour. The signature computed in this manner would also be undefined, in certain ranges of the angle, for a contour for which the centroid falls outside the region enclosed by the contour.

A major advantage with the use of 1D signatures is the reduction in dimensionality from the corresponding 2D contours. Signatures may be filtered or processed for the reduction of noise and artifacts in the contour.

C. Fractal analysis

Fractal analysis may be used to study the complexity and roughness of 1D functions, 2D contours, and images [15], [16], [17], [18], [19], [20]. Fractal analysis may be applied to classify breast masses based on the complexity of their contours. Matsubara et al. [21] obtained 100% accuracy in the classification of 13 breast masses using FD. The method required the computation of a series of FD values for several contours of a given mass obtained by thresholding the mass at many levels; the variation in FD was used to categorize a given mass as benign or malignant. Pohlman et al. [14] obtained a classification accuracy of more than 80% with fractal analysis of signatures of contours of masses. Nguyen and Rangayyan [8] estimated the FD of a set of 111 contours of breast masses and tumors using the ruler and the box-counting methods applied to the 2D contours as well as their 1D signatures ($s_d(n)$ as described in Section I-B). The best classification performance with $A_z = 0.89$ was obtained with the ruler method applied to the 1D signatures of the contours.

D. Turning angle function as a signature

The turning angle function $T_s(s(n))$ of a contour s is defined as the angle, measured in the counterclockwise direction, of the tangent to the contour at $s(n)$, with reference to the x axis, expressed as a function of the contour index n. The turning angle function is also known as the tangent function [22], [23]. Fig. 1 and Fig. 2 show the turning angle functions for the contours of a benign mass and a malignant tumor, respectively. For a convex contour, such as the case in Fig. 1, the turning angle function is, in general, a monotonically increasing function (see Fig. 1 (b)). For a contour with concave and convex portions, the turning angle function begins to decrease at the beginning of a concave portion, and keeps on decreasing until the direction of the tangent to the contour changes at the beginning of the next convex portion. In the turning angle function in Fig. 2, a spicule in the contour is related to a portion bounded by a pair of successive significant drops in the angle. The turning angle function of a contour may be used as a signature to represent its shape characteristics [22], [23].

The objective of this paper is to present the application of turning angle functions for the analysis of contours of breast masses. In particular, we present methods to derive an index of spiculation to estimate FD, and to compute an index of convexity from a given turning angle function. To evaluate the performance of the proposed features in terms of the efficiency in the classification of breast masses, we compare the results with those provided by SI using the method of Rangayyan et al. [4] and FD as obtained by Nguyen and Rangayyan [8], in terms of the area A_z under the receiver operating characteristics (ROC) curve.

II. DATA USED: CONTOURS OF BREAST MASSES

Mammograms of 20 cases were obtained from Screen Test: the Alberta Program for the Early Detection of Breast Cancer [5], [24], [25]. The mammograms were digitized using the Lumiscan 85 scanner at a resolution of 50 μm with 12 b/ pixel. Fifty-seven regions of interest (ROIs), of which 37 are related to benign masses and 20 are related to malignant tumors, were obtained [5].

Mammograms containing masses were also obtained from the Mammographic Image Analysis Society (MIAS, UK) database [26], [27] and the teaching library of the Foothills Hospital (Calgary) [4], [3]. The MIAS images were digitized at a resolution of 50 μm. The Foothills Hospital images were digitized at 62 μm per pixel. This set includes 28 benign masses and 26 malignant tumors.

Contours of the masses in the images described above were drawn by an expert radiologist specialized in mammography. The dataset used in the present study includes 111 contours, with typical and atypical shapes of 65 benign masses and 46 malignant tumors. The diagnostic classification was based upon biopsy. (The present work employs the same dataset as that used by Nguyen and Rangayyan [8]).

III. METHODS: THE TURNING ANGLE FUNCTION

The turning angle function of a given contour was obtained as described in Section I-D. Fig. 1(b) and Fig. 2(b) illustrate the turning angle functions of the contours of a benign mass and a malignant tumor, respectively. It is readily seen that while the former is a nearly monotonically increasing function, the latter has many decreasing and increasing segments related to the spicules present in the contour of the tumor. The examples indicate that the turning angle function may be used to represent the complexity as well as the variations present in the shapes of breast masses and tumors.

A. Index of spiculation from the turning angle function

In order to derive an index of spiculation (or lobulation) from a turning angle function (referred to as SI_{TA}), the length of each possible spicule is multiplied by $(1 + \cos\psi)$, where ψ is the angle of the spicule (as obtained from the turning angle function). The weighted lengths of the spicules are summed, and normalized by twice the sum of their unweighed lengths.

B. Fractal dimension of the turning angle function

Nguyen and Rangayyan [8] obtained the best performance in the classification of breast masses and tumors by the application of the ruler method to 1D signatures ($s_d(n)$ as described in Section I-B) of breast masses, as compared to the box-counting method applied to the 2D contours or their 1D signatures, as well as the ruler method applied to the 2D contours. In the present work, the ruler method was applied to the 1D turning angle functions of the contours of the breast masses (referred to as FD_{TA}).

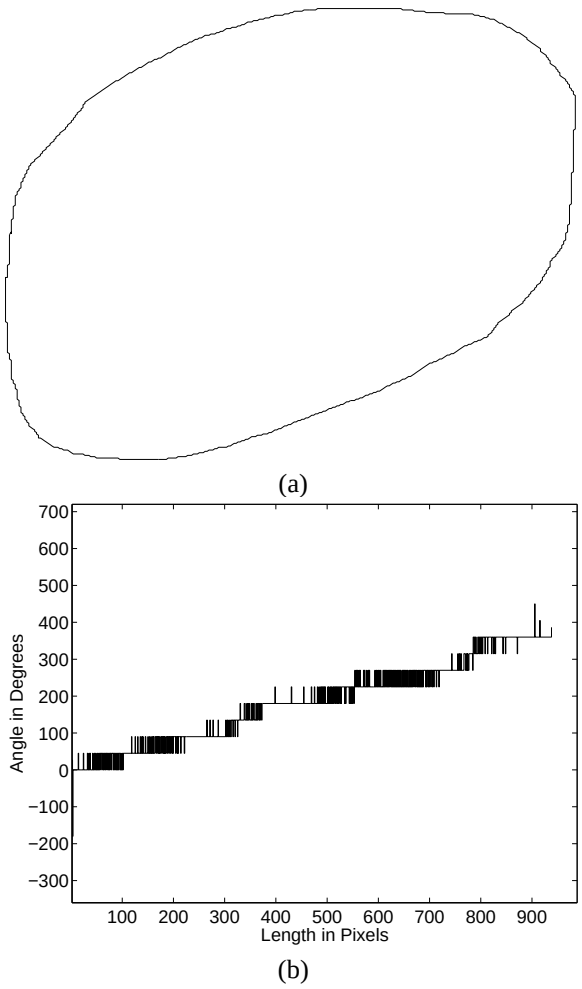

Fig. 1. (a) A benign mass with a relatively smooth and convex contour. (b) Turning angle function of the contour.

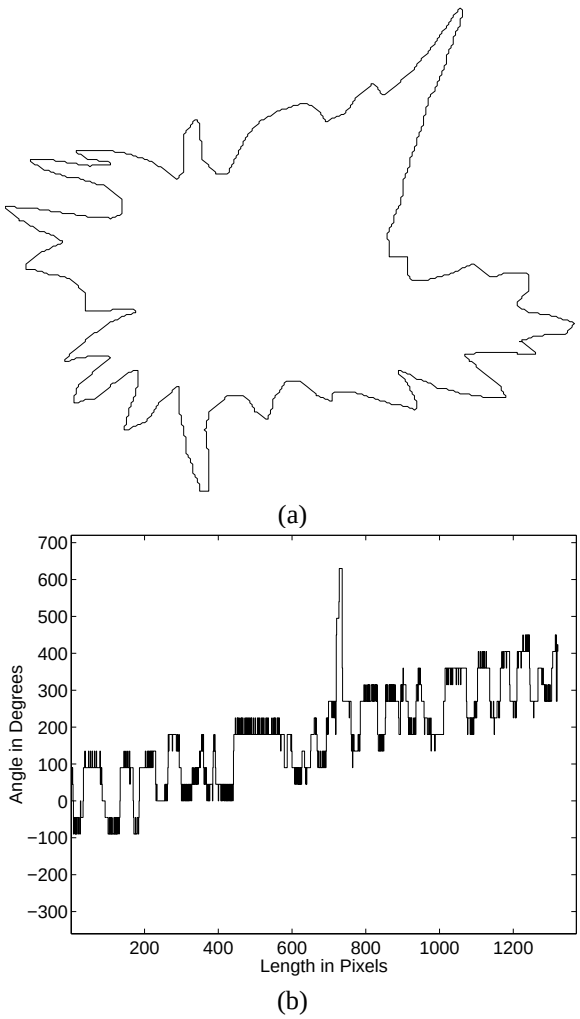

Fig. 2. (a) A malignant tumor with a spiculated contour including concave segments. (b) Turning angle function of the contour.

C. Index of convexity from the turning angle function

Hand-drawn contours, such as those shown in Fig. 1(a) and Fig. 2(a), could contain artifacts and noise related to hand tremor and other limitations. As a consequence, the turning angle function could be expected to contain several small segments that are insignificant in the representation of the contours for further analysis. For this reason, it is necessary to filter the turning angle function in a selective manner, so as to remove the artifacts and noise while preserving the significant details, as shown in Figure 3.

To obtain the index of convexity CI_{TA}, the filtered turning angle function is smoothed to remove irrelevant information. The smoothed function is obtained by replacing each monotonically increasing or decreasing section of the filtered turning function by a representative segment and its corresponding turning angle. The new segment length is obtained by summing all related individual segment lengths in the increasing or decreasing section, and the new turning angle is obtained by computing the average of the relative turning angles of the corresponding segments. The results of the smooth turning angle function are shown in Fig. 4. Note that the smoothed turning angle function of a convex contour is a constant function, as illustrated in Figure 4(a); on the other hand, the filtered and smoothed turning function of a contour with concavities has several variations as shown in Fig. 4(b). The index of concavity IC_{TA} based on the turning angle function is defined as:

$$IC_{TA} = 1 - \left(\frac{\alpha + \beta}{4S}\right) \tag{1}$$

where α represents the presence of convex regions in the contour, β represents the presence of concave regions in the contour, and S is the sum of all segments of the smoothed turning angle function.

The terms α and β are obtained as follows:

$$\alpha = \sum_{i=1}^{N} (\ (1 + cos(\theta(i)))\ S_a(i)) \tag{2}$$

where $S_a(i)$ is the sum of the lengths of two adjacent segments joined by the drop in angle $\theta(i)$ obtained from the smoothed turning angle function (see Figure 5; for the

illustration shown, $S_a(i) = (L1+L2)$), and N is the number of drops in angles.

$$\beta = \sum_{j=1}^{M}(\ (1 + cos(\phi(j)) \ S_b(j)) \tag{3}$$

where $S_b(i)$ is the sum of the lengths of two adjacent segments joined by an increasing angle $\phi(i)$ obtained from the smoothed turning angle function (see Figure 5; for the case illustrated, $S_b(j) = (L2 + L3)$), and M is the number of steps with increasing angles.

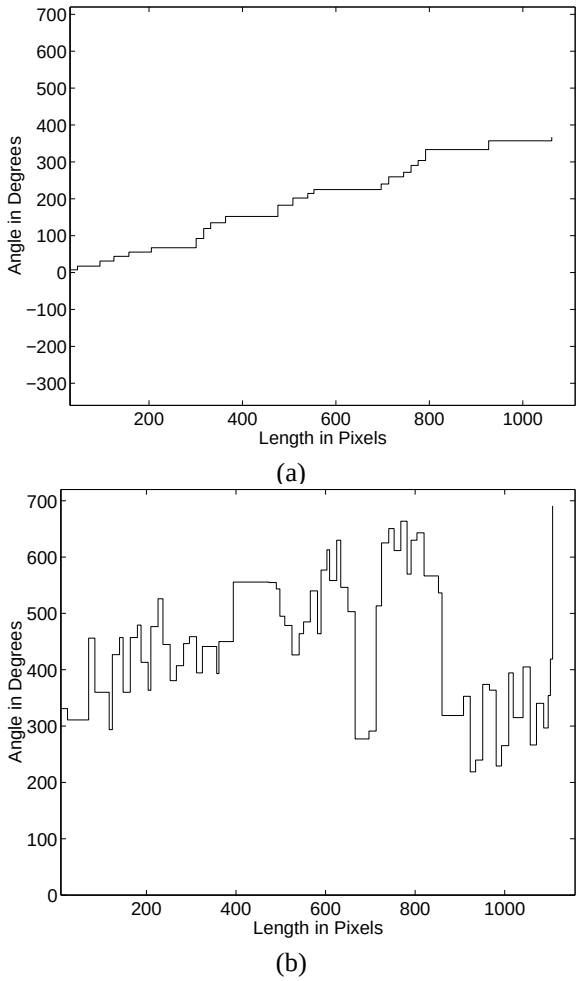

Fig. 3. Filtered turning function with artifacts and noise eliminated, corresponding to: (a) the turning angle function in Figure 1(b); (b) the turning angle function in Figure 2(b).

D. Pattern classification experiments

The conditional probability density functions of feature vectors derived from the proposed methods, assumed to be Gaussian, were estimated for the two classes of benign masses and malignant tumors. Using Bayes formula, a discriminant function was composed, and the leave-one-out method was used in estimating the classification accuracy [28]. A sliding threshold was applied to classify the feature vectors, and receiver operating characteristics (ROC) curves

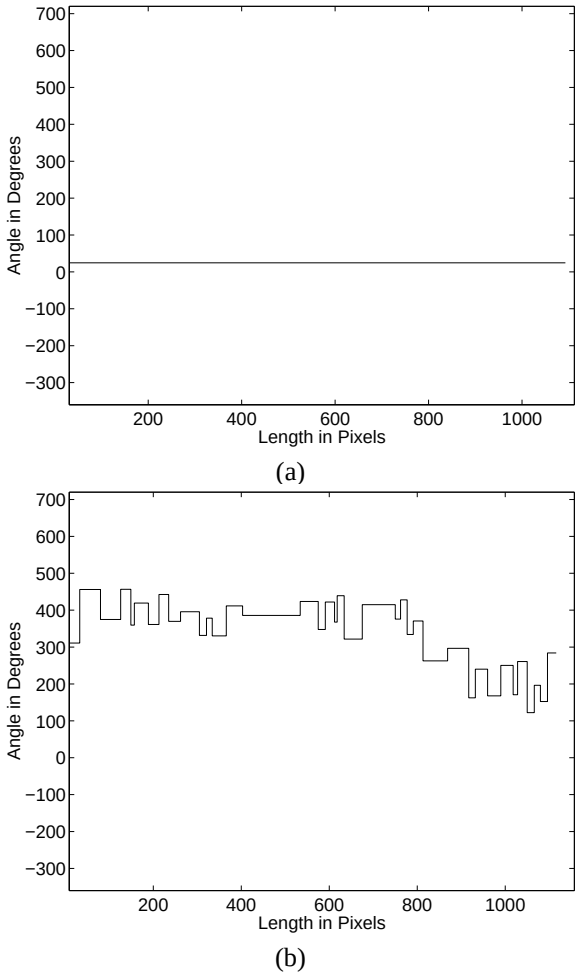

Fig. 4. Smoothed turning angle functions: (a) of the benign mass with convex contour shown in Figure 1; (b) of the malignant tumor with spiculated contour shown in Figure 2.

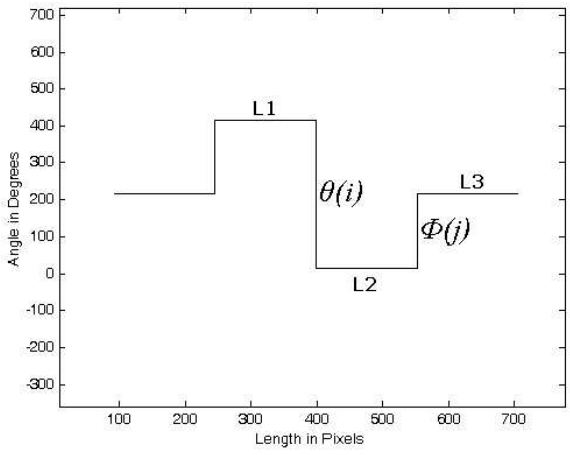

Fig. 5. Relation between a drop and an increase in the turning angle and their adjacent segments.

TABLE I

CLASSIFICATION PERFORMANCE OF VARIOUS SHAPE FEATURES.

Feature	A_z
SI of Rangayyan et al. [4]	0.90
FD of Nguyen and Rangayyan [8]	0.89
SI and FD	0.91
SI_{TA}	0.92
CI_{TA}	0.93
FD_{TA}	0.91
SI_{TA} and FD_{TA}	0.91

[29] were generated. The area A_z under each ROC curve was computed to serve as a measure of the classification performance of the corresponding feature vector.

IV. RESULTS AND DISCUSSION

The methods were tested with a set of 111 contours of breast masses; see Section II for details regarding the data used. The features SI_{TA}, CI_{TA} and FD_{TA} obtained using the methods described in Section III were evaluated in terms of the area A_z under the ROC curve. Table I provides the values of A_z obtained for some of the feature combinations tested. The results provided by the spiculation index SI obtained using the method of Rangayyan et al. [4], as well as by FD using the ruler method applied to 1D signatures of contours by Nguyen and Rangayyan [8] are also shown in the table. It is seen that the shape features proposed in this paper provide, the best results.

V. CONCLUSION

We have proposed methods to obtain shape features from the turning angle functions of contours. The features are useful in the analysis of contours of breast masses and tumors because of their ability to capture diagnostically important details of shape related to spicules and lobulations. The proposed features have provided high classification accuracies in discriminating between benign breast masses and malignant tumors. The methods should be useful in computer-aided diagnosis of breast cancer.

REFERENCES

[1] Homer MJ. *Mammographic Interpretation: A Practical Approach*. McGraw-Hill, Boston, MA, 2nd edition, 1997.

[2] American College of Radiology, Reston, VA. *Illustrated Breast Imaging Reporting and Data System (BI-RADS$^{\mathrm{TM}}$)*, third edition, 1998.

[3] Rangayyan RM, El-Faramawy NM, Desautels JEL, and Alim OA. Measures of acutance and shape for classification of breast tumors. *IEEE Transactions on Medical Imaging*, 16(6):799–810, 1997.

[4] Rangayyan RM, Mudigonda NR, and Desautels JEL. Boundary modelling and shape analysis methods for classification of mammographic masses. *Medical and Biological Engineering and Computing*, 38:487–496, 2000.

[5] Alto H, Rangayyan RM, and Desautels JEL. Content-based retrieval and analysis of mammographic masses. *Journal of Electronic Imaging*, 14(2):023016:1–17, 2005.

[6] Bruce LM and Adhami RR. Classifying mammographic mass shapes using the wavelet transform modulus-maxima method. *IEEE Transactions on Medical Imaging*, 18(12):1170–1177, 1999.

[7] Sahiner BS, Chan HP, Petrick N, Helvie MA, and Hadjiiski LM. Improvement of mammographic mass characterization using spiculation measures and morphological features. *Medical Physics*, 28(7):1455–1465, 2001.

[8] Nguyen TM and Rangayyan RM. Shape analysis of breast masses in mammograms via the fractal dimension. In *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (CD-ROM)*, pages 4, paper number 1852, Shanghai, China, September 2005. IEEE.

[9] Huo Z, Giger ML, Vyborny CJ, Wolverton DE, and Metz CE. Computerized classification of benign and malignant masses on digitized mammograms: A study of robustness. *Academic Radiology*, 7(12):1077–1084, 2000.

[10] Huo Z, Giger ML, and Vyborny CJ. Computerized analysis of multiple-mammographic views: Potential usefulness of special view mammograms in computer-aided diagnosis. *IEEE Transactions on Medical Imaging*, 20(12):1285–1292, 2001.

[11] Mudigonda NR, Rangayyan RM, and Desautels JEL. Detection of breast masses in mammograms by density slicing and texture flow-field analysis. *IEEE Transactions on Medical Imaging*, 20(12):1215–1227, 2001.

[12] Wei D, Chan HP, Helvie MA, Sahiner B, Petrick N, Adler DD, and Goodsitt MM. Classification of mass and normal breast tissue on digital mammograms: multiresolution texture ananlysis. *Medical Physics*, 22(9):1501–1513, 1995.

[13] Rangayyan RM. *Biomedical Image Analysis*. CRC Press, Boca Raton, FL, 2005.

[14] Pohlman S, Powell KA, Obuchowski NA, Chilcote WA, and Grundfest-Broniatowski S. Quantitative classification of breast tumors in digitized mammograms. *Medical Physics*, 23(8):1337–1345, 1996.

[15] Mandelbrot BB. *The Fractal Geometry of Nature*. WH Freeman and Company, San Francisco, CA, 1983.

[16] Peitgen HO, Jürgens H, and Saupe D. *Chaos and Fractals: New Frontiers of Science*. Springer, New York, NY, 2004.

[17] Deering W and West BJ. Fractal physiology. *IEEE Engineering in Medicine and Biology Magazine*, 11(2):40–46, June 1992.

[18] Schepers HE, van Beek JHGM, and Bassingthwaighte JB. Four methods to estimate the fractal dimension from self-affine signals. *IEEE Engineering in Medicine and Biology Magazine*, 11(2):57–64, June 1992.

[19] Fortin C, Kumaresan R, Ohley W, and Hoefer S. Fractal dimension in the analysis of medical images. *IEEE Engineering in Medicine and Biology Magazine*, 11(2):65–71, June 1992.

[20] Goldberger AL, Rigney DR, and West BJ. Chaos and fractals in human physiology. *Scientific American*, 262:42–49, February 1990.

[21] Matsubara T, Fujita H, Kasai S, Goto M, Tani Y, Hara T, and Endo T. Development of new schemes for detection and analysis of mammographic masses. In *Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS'97)*, pages 63–66, Grand Bahama Island, Bahamas, December 1997.

[22] Arkin EM, Chew LP, Huttenlocher DP, Kedem K, and Mitchell JSB. An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:209–216, March 1991.

[23] Latecki LJ and Lakämper R. Application of planar shape comparisons to object retrieval in image databases. *Pattern Recognition*, 35(1):15–29, 2002.

[24] Alberta Cancer Board, Alberta, Canada, www. cancerboard. ab.ca/ screentest. *Screen Test: Alberta Program for the Early Detection of Breast Cancer – 2001/03 Biennial Report*, 2004.

[25] Alto H, Rangayyan RM, Paranjape RB, Desautels JEL, and Bryant H. An indexed atlas of digital mammograms for computer-aided diagnosis

of breast cancer. *Annales des Télécommunications*, 58(5-6):820–835, 2003.

[26] The Mammographic Image Analysis Society digital mammogram database. http://www.wiau.man.ac.uk/ services/MIAS/ MIASweb.html, accessed June, 2004.

[27] Suckling J, Parker J, Dance DR, Astley S, Hutt I, Boggis CRM, Ricketts I, Stamatakis E, Cerneaz N, Kok SL, Taylor P, Betal D, and Savage J. The Mammographic Image Analysis Society digital mammogram database. In Gale AG, Astley SM, Dance DR, and Cairns AY, editors, *Proceedings of the 2nd International Workshop on Digital Mammography*, volume 1069 of *Excerpta Medica International Congress Series*, pages 375–378, York, England, July 1994.

[28] Duda RO, Hart PE, and Stork DG. *Pattern Classification.* Wiley, New York, NY, 2nd edition, 2001.

[29] Metz CE. Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, VIII(4):283–298, 1978.