# Classification of Mammograms Tumors Using Fourier Analysis

**Osama R.Shahin and Gamal Attiya,**

University of Helwan Egypt, University of Menoufia Egypt

**Summary**
Breast cancer is one of the most common cancers among women in the developing countries. It has become a major cause of death. In this work a new algorithm for classifying mammograms by using an evolutionary approach known as signatures- distances from the centroid to all points on the boundary of the region of interest (ROI) as a function of a polar angle θ. The signature of a closed boundary is a periodic function, repeating itself on an angular scale of 2π. Then encode and describe this closed boundary to arbitrary function through 1-D (radial) Fourier expansion coefficients. The method was tested over several images from the image databases taken from Breast Imaging Reporting and Data System BIRADS developed by the American College of Radiology, and from MIAS (Mammogram Image Analysis Society, UK), that provides a standardized classification for mammographic studies. The implementation of the algorithm was carried out using MATLAB codes programming and thus is capable of executing effectively on a simple personal computer with digital mammogram as accumulated data for assessment. In this paper, we describe the formatting guidelines for IJCA Journal Submission.

***Key words:***
*Microcalcifications, Mass lesions, Signature, Fourier expansion coefficients.*

## 1. Introduction

Breast cancer is currently the top cancer in women worldwide, both in the developed and the developing world. The majority of breast cancer deaths occur in low- and middle-income countries, where most of the women are diagnosed in late stages due mainly to lack of awareness and barriers to access to health services. 7.6 million people worldwide died from cancer in 2008, Approximately 70% of cancer deaths occur in low- and middle-income countries [1].Breast cancer has placed itself on top of the list of health problems for women in Egypt, representing 35.1% of all female cancer cases, according to the National Cancer Institute in Cairo [2]. No effective way to prevent the occurrence of breast cancer exists. Therefore, early detection is the first crucial step towards treating breast cancer. It plays a key role in breast cancer diagnosis and treatment. This process requires the selection of the effective features for the process of detection and classification. Over recent years there has been much research into the application of breast cancer detection and classification with numerous different features. Many of these involved geometrical and statistical features with many problems that associated with these analyses [3, 4]. The fundamental problems which faced the last approaches were the dependency of these approaches on the image translation, scaling and rotation. So the needed of a robust shape descriptor become the important factor to solve such problem. The work presented in this paper was concerned specifically with encoded the boundary of the tumor by Fourier expansions which are certainly translation invariant. In our experimental study, we will use the digital mammogram images that were provided from online mammogram database (MIAS database) [5]. Firstly, for each image the location for each suspicious area will be decided and secondly a number of significant features will be computed. After running the algorithm, the results that obtained by the proposed features – coefficients of Fourier expansion-will be compared with the standard results which depend on the geometric and texture features that discussed in [3, 4]. The remainder of the paper is organized as follows: we discuss the main mammographic abnormalities in section II. The methodology was described in section III. Experimental results are presented in section IV. Finally, conclusions are drawn in section V.

## 2. Mammographic Abnormalities

There are about eight typical kinds of abnormalities revealed with a conventional mammogram. An experienced radiologist is highly tuned to the appearance of abnormalities in breast X-rays, and most of the time has a pretty good idea whether a suspicious abnormality is likely to be malignant or not. Typical mammographic findings from breast cancer screening mammograms would include asymmetrical breast tissue, asymmetric density, architectural distortion, mass, microcalcifications, interval changes compared with previous films, adenopathy, and other miscellaneous

Manuscript received February 5, 2014
Manuscript revised February 20, 2014

findings. Usually, a mammographic abnormality is followed by additional imaging studies, such as ultrasound, and if the lesion still appears suspicious it may be sent for biopsy [6].

### 1.1.Characteristics of Mass Lesion

In terms of shape, if it is round, oval, or slightly lobular, the mass is probably benign. If the mass has a multi-lobular contour, or an irregular shape, then it is suggestive of malignancy. 'Margin' refers to the characteristics of the border of the mass image. When the margin is circumscribed and well-defined the mass is probably benign. If the margin is obscured more than 75% by adjacent tissue, it is moderately suspicious of malignancy. Likewise, there is moderate suspicion if the margin having many small lobes. If the margin is indistinct or speculated (consisting of many small 'needle-like' sections) then there is also high suspicion of malignancy. 'Density' is usually classified as either fatty, low, or high. The mass is probably benign for fatty and low densities, moderately suspicious of malignancy for high densities [7].

### 1.2 Microcalcifications

Microcalcifications are one of the main ways breast cancer mammographically detected when it is in the very early stages. Microcalcifications are actually tiny specks of mineral deposits (such as calcium) they can be distributed in various ways. Sometimes microcalcifications are found scattered throughout the breast tissue, and they often occur in clusters. Frequently, microcalcification deposits are due to benign causes. However, certain features and presentations of microcalcifications are more likely to be associated with malignant breast cancer Figure1. Three categories of calcifications have been identified by the “The American College of Radiology (ACR) BIRADS” [8]

(a) Typically benign

(b) Intermediate concern

(c) High probability of malignancy

The summary of BIRADS Classification of Calcifications summarized in Table 1

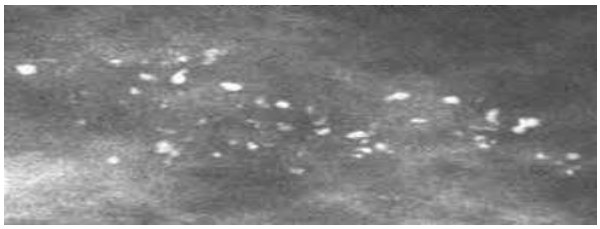


Fig. 1: Microcalcifications

### 1.3Mass Lesions

Breast cancer is characterized with the presence of a mass accompanied or not accompanied by calcifications [9]. There is a possibility of a cyst, which is non-cancerous collection of fluid to resemble a mass in the film. The identical intensities of the masses and the normal tissue and similar morphology of the masses and regular breast textures makes it a tedious task to detect masses in comparison with that of calcifications [10]. The location, size, shape, density, and margins of the mass are highly beneficial for the radiologist to evaluate the probability of cancer. A majority of the benign masses are well circumscribed, compact, and roughly circular or elliptical whereas the malignant lesions are characterized by blurred boundaries, irregular appearances and are occasionally enclosed within a radiating pattern of linear speckles [11]. Nevertheless some benign lesions may also possess speculated appearances or blurred peripheries.

### 1.4 Roundness

The shape of particles is an important property in determining their history and behavior [12]. For this reason it is important to have an objective quantitative measure of particle shape so that the changes in shape as well as the differences between the shapes of different populations can readily be identified. One of the important methods that used in describing the shape is roundness. Roundness is a measure of the extent to which the edges and comers of a particle has been rounded. A number of authors have proposed methods using the Fourier transform to determine roundness [13, 14]. The basic method is to obtain coordinates on the edge of the profile of the fragment being measured also the centre point is determined, and all the edge coordinates converted to polar coordinates using the centre point as origin. The Fourier expansion coefficients of the vector of radii (r (θ)) are then calculated and the roundness determined from these coefficients.

## 3. Methodology

The automatic massive lesions classification algorithm can be summarized in the following points:

a- Detect the region of interests (ROIs): These regions can be easily detected in an image if the area has sufficient contrast from the background. In this phase, the algorithm in [15] was applied. Once the Region of Interest (ROI) is automatically extracted, A Sobel filter was applied on the image to detect the edges of the region of interest (ROI). The ROI is the tumor of the digital mammogram and the goal is to isolate this area from the image. A dilation operation was performed after filtering to connect edges. Dilation was followed by filling the remaining holes of the ROI.

**Table 1. .Summary of BIRADS Classification of Calcifications [8]**

| | Type of calcification | Characteristics |
|---|---|---|
| **Typically benign** | Skin | Typical polygonal shape. |
| | Vascular | Parallel tracks or linear tubular. |
| | Coarse or pop-corn like | Involving fibro adenomas. |
| | Rod-shaped | Large rod usually > 1mm. |
| | Round | Smooth, round clusters. |
| | Punctuate | Round or oval calcifications. |
| | Spherical or lucent centered | Found in debris collected in ducts, or necrosis areas. |
| | Rim or egg-shell | Found in wall of cysts. |
| | Milk or calcium | Calcium precipitates. |
| | Dystrophic | usually large > 0.5mm in size |
| **Intermediate concern** | Indistinct or amorphous | Appear round and hazy uncertain morphology |
| **High risk** | Pleomorphic or heterogenous | irregular in shape, size and < 0.5mm raises suspicion |
| | Fine, linear or branching | Thin, irregular that appear linear from a Distance |

b- Determine the centroid for any (ROI).

c- Boundary Signature: Calculate the distances from the centroid to the all points on the boundary of the region of interest (ROI) as a function of a polar angle θ. A signature is the representation of a 2-D boundary as a 1-D function. The signature of a closed boundary is a periodic function, repeating itself on an angular scale of 2π. Such distance called Radial Distance (RD), see Figure 2.

d- Evaluation of Fourier Expansion: Encode this closed boundary by an arbitrary function through 1-D (radial) Fourier expansion coefficients. One simple and neat way to encode and describe a closed boundary to arbitrary and accuracy is through a 1-D (radial) Fourier expansion.

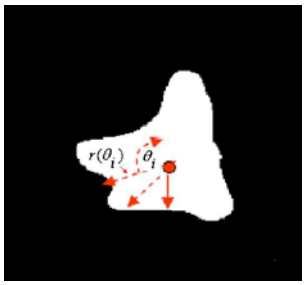



**Fig. 2: Radial Distance Measure (RDM)**

The signature can be expressed in real or complex form as follows:

$$r(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta) \quad (1)$$

$$r(\theta) = \sum_{n=-\infty}^{\infty} C_n e^{(in\theta)} \quad (2)$$

The shape can be parametrically encoded by the real Fourier expansion coefficients $\{a_n, b_n\}$ or by the complex coefficients $\{C_n\}$.These coefficients can easily be calculated through use of the orthogonality relations for Fourier series [14] . The real coefficients for a radial signature are given by:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} r(\theta) \cos(n\theta) d\theta \quad (3)$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} r(\theta)\sin(n\theta)d\theta \quad (4)$$

Whereas the complex coefficient $\{C_n\}$ is given by:

$$C_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} r(\theta)e^{(-in\theta)}d\theta \quad (5)$$

Typically, a good approximation to the shape can be encoded using a relatively small number of parameters, and more terms can be included if higher accuracy is required. The use of radial Fourier expansions can, however, become problematic on complicated boundary shapes, particularly those in which the boundary 'meanders back' on itself .The signature function r(θ) may not be single valued, there being two or more possible radial values for a given value of θ. In such cases, the choice of which value of r(θ) to select is somewhat arbitrary and the importance of these unavoidable ambiguities will depend on the specific application. In general, however, strongly meandering boundaries. The Fourier descriptors calculated according to Equations [3-5] are certainly translation invariant. This follows because the radial distance in the signature is calculated with respect to an origin defined by the centroid coordinates of the boundary. Multiplication of the signature by an arbitrary scale factor is reflected in the same scale factor multiplying each of the individual Fourier coefficients. A form of scale invariance can thus be achieved most simply by dividing the signature by its maximum value (thus fixing its maximum value as one). The extraction features that can be calculated from the last two steps are:

Number of zero crossing: the mean value of the radial distances r(θ) can be taken as a reference axis , so it is easy to find the number of points which the r(θ) passing through this axis.

The Fourier expansion coefficient $\{a_n, b_n\}$ also used as important feature because these values increased with the complexity of the boundary (signature) and conversely.

Find the mean value and the variance for the first 200 Fourier expansions, Where:

$$mean(a_n) = \mu(a_n) = \frac{1}{m}\sum_{n=1}^{m} a_n \quad (6)$$

$$mean(b_n) = \mu(b_n) = \frac{1}{m}\sum_{n=1}^{m} b_n \quad (7)$$

$$variance(a_n) = \frac{1}{m-1}\sum_{n=1}^{m}(a_n - \mu(a_n))^2 \quad (8)$$

$$variance(b_n) = \frac{1}{m-1}\sum_{n=1}^{m}(b_n - \mu(b_n))^2 \quad (9)$$

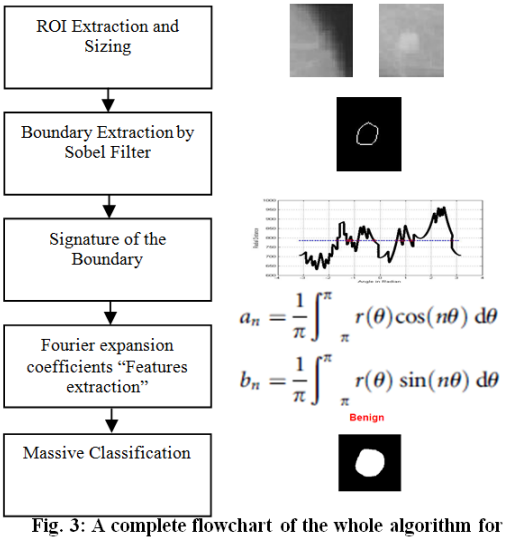



**Fig. 3: A complete flowchart of the whole algorithm for the breast masses classification.**

Classify the massive lesion according to the values of last extraction features that obtained in the previous step. A flowchart of the whole algorithm for the breast masses classification is shown in Figure3

## 4. Results

The method suggested for the detection of micro-calcifications and mass lesions from mammogram images was described above. The mammograms of our study were analyzed to detect possible clusters of micro-calcifications and mass lesions taken from MIAS (Mammogram Image Analysis Society, UK), and BIRADS databases. Figure4 shows an example of an original image containing a mass lesion, and the results of the classification procedure. Figures [5] show these resulted signatures. Table 2 shows the feature values that described above. The value of each feature at its corresponding image is also illustrated. The results show high classification rate. The summarizes of the features range was depicted in Table 3, thus the feature-set formed is well suited for any CAD system. Also these features have an additional advantage of involving less computational complexity, translation invariant. This follows because the radial distance in the signature was

calculated with respect to an origin defined by the centroid coordinates of the boundary. This can prove helpful in increasing efficiency of any CAD system. To check the validity of these features we compare it by the geometrical and statistical features for the same images were calculated in Tables [4, 5] for predefined size. The value of each features in geometric and statistical not enough to discriminate between the different types of masses conversely for the proposed features.

**Table 2: Statistics Values for the First 200 Fourier Expansion Terms** *B assign for Benign, M assign Malignant

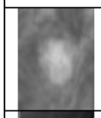
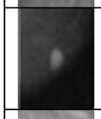
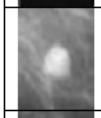
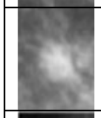
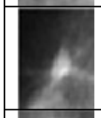
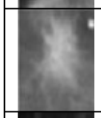
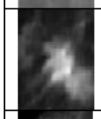
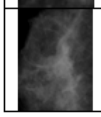
| ROI | $\mu(a_n)$ | $\mu(b_n)$ | $var(a_n)$ | $var(b_n)$ | *Mass Type |
|---|---|---|---|---|---|
|  | 3.3807 | 5.2194 | 41.4072 | 19.1919 | B |
|  | 2.0976 | 1.9374 | 15.5916 | 7.9063 | B |
|  | 3.5214 | 3.4259 | 22.1500 | 14.6525 | B |
|  | 5.2194 | 4.1874 | 52.6881 | 23.4905 | B |
|  | 10.7955 | 10.3994 | 748.6206 | 301.4120 | M |
|  | 10.7955 | 11.3316 | 555.3416 | 145.5688 | M |
|  | 9.7815 | 11.3316 | 347.2834 | 177.2130 | M |
|  | 10.9853 | 11.7909 | 296.2697 | 642.5353 | M |

**Table 3: Summary of Results for the Proposed Features**

| Massive Lesion | Mean value | Variance value | Number of zero crossing |
|---|---|---|---|
| **Benign** | 3.4595 | 24.4383 | 17 |
| **Malignant** | 11.5516 | 373.9002 | 23 |

**Table 4. Summary of Geometric Features for the massive lesion in Database**

| | | | | |
|---|---|---|---|---|
| Benign | Average Area | Average Perimeter | Average Roundness | Average Compactness |
| | 120.1001 | 35.4511 | 1.0491 | 1.0223 |
| Malignant | Average Area | Average Perimeter | Average Roundness | Average Compactness |
| | 591.8001 | 105.1529 | 0.6562 | 0.8062 |

**Table 5. Summary of Statistical Features for the massive lesion in Database**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Benign | μ | σ2 | σ | Smoothness | Skewness | Uniformity | Entropy |
| | 20.7337 | 1067.5767 | 29.3707 | 0.9979 | 7.6738 | 2.2407. | -577.8862 |
| Malignant | μ | σ2 | σ | Smoothness | Skewness | Uniformity | Entropy |
| | 40.7406 | 5295.6981 | 70.7061 | 0.9997 | 2.5889 | 2.1115. | -489.6207 |

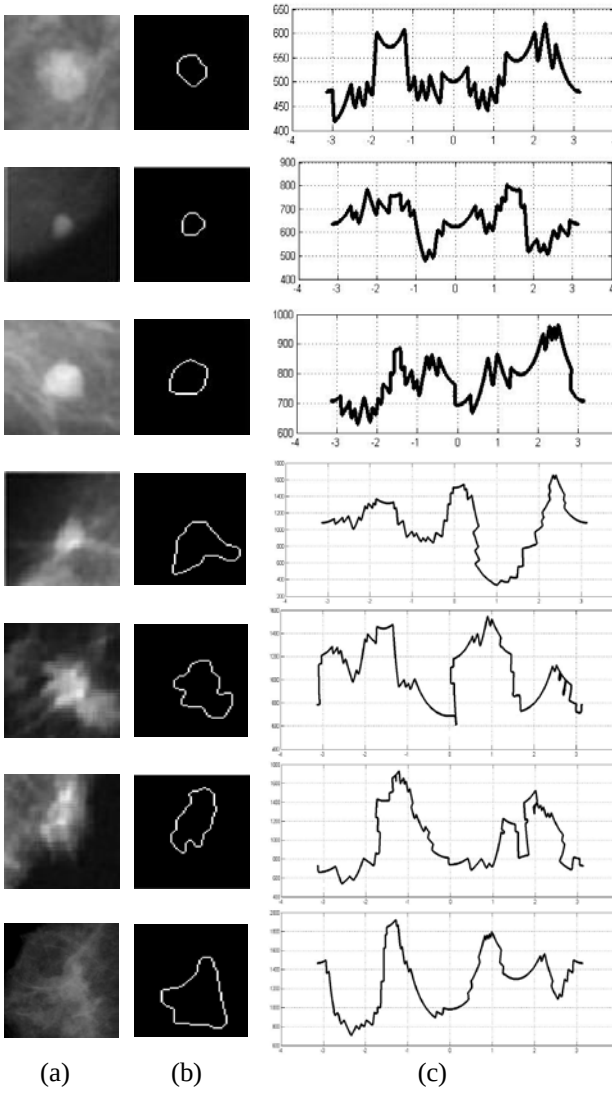



**Fig.4: (a) Mammographic ROI , (b) Boundary of ROI, (c) Signature pattern.**

## 5. Conclusion

As a preliminary investigation toward the implementation of a complete CAD system for the early detection of breast cancer, in this paper we present an algorithm for the classification and features extraction of microcalcifications. The fundamental problems which faced the geometric and statistical features were the dependency on the image translation, scaling and rotation. So the needed of a robust shape descriptor become the important factor to solve such problem, that solved by using Fourier expansion coefficients which are certainly translation invariant.

## 6. References

[1] World health organization, http://www.who.int/cancer/en.
[2] National Cancer Institute in Cairo, http://www.nci.edu.eg/
[3] Hala Al-Shamlan and Ali El-Zaart, “Feature Extraction Values for Breast Cancer Mammography Images”, International Conference on Bioinformatics and Biomedical Technology 2010.
[4] D. C. Hope E. Munday S. L. Smith., “ Evolutionary Algorithms in the Classification of Mammograms”, Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)
[5] MAIS database, http:// peipa.essex.ac.uk/info/mias.html [Accessed: September 3, 2011].
[6] Bijay Ketan Panigrahi Assistant, Manas Ranjan “Feature Extraction for Classification of Micro-calcifications and Mass Lesions in Mammograms”, IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.5, May 2009 255 Manuscript
[7] Sampat MP, Whitman GJ, Stephens TW, Broemeling LD, Heger NA, Bovik AC, Markey MK., “ The reliability of measuring physical characteristics of spiculated masses on mammography”, Br J Radiol. Dec 2006; 79 Spec No 2:S134-40.
[8] American College of Radiology. The ACR breast imaging reporting and data system (BI-RADS) , http://www.acr.org/departments/stand_accred/birads/contents.html. [Accessed: September 20, 2011]
[9] American Cancer Society, “Cancer Facts and Figures 2011”, Atlanta, GA: American Cancer Society, 2011
[10] Feig SA, Yaffe MJ, “Digital Mammography, Computer-Aided Diagnosis and Telemammography”, The Radiologic Clinics of North America, Breast Imaging, , vol 33, 6, 1205- 1230, January 1995
[11] Phil Evans W., “Breast Masses Appropriate Evaluation”, The Radiologic Clinics of North America, Breast Imaging, vol. 33, no. 6, pp. 1085-1108, January 1995.
[12] Gather R. Drevin, “ Using Entropy to Determine the Roundness of Rock Particles” Proceedings of ICSP 2000.
[13] Bala Muralikrishnan,J. Raja “ Computational Surface and Roundness Metrology" Springer, Dec 3, 2008.
[14] Chris Solomon,SchToby Breckon " Fundamentals of Digital Image Processing A Practical Approach with Examples in Matlab” , John Wiley & Sons, Ltd 2011.
[15] Mencattini, A., Salmeri, M., Casti, P., Raguso, G. “ Automatic breast masses boundary extraction in digital

mammography using spatial fuzzy c-means clustering and active contour models", IEEE International Workshop on Medical Measurements and Applications (MeMeA) , Bari, Italy, pp.632-637 , May 30-31, 2011.