International Journal of Electronics Engineering, 1(1), 2009, pp. 13-18

A Novel Image Processing Based Approach for Determining Size of Breast Tumors

Syed Abdaheer. M. & Ekram Khan

Department of Electronics Engineering, A.M.U., Aligarh, U.P., INDIA

Abstract: *In this work, we consider an image of tumor identified breast and we present a novel approach for analysis of the size of breast tumors by using Closed Circle Curvature Fitting method (CCCF). The tumor area of the image is first segmented for extracting tumors with the help of optimal threshold techniques. The segmented images undergo for morphological treatment to remove unwanted high frequency components present in the image. The edges are extracted from filtered images through Laplacian edge detection method. The polar coordinates are derived from the centroid value of edge extracted images. Radius and area of the circle are then calculated. Our next task is to fit this extracted images with circle for getting numerical details about tumors. The accuracy of the proposed method and details about tumor like area and size are obtained with the help of area matching method. By comparing with elliptical curvature fitting method, our proposed method has superiority in terms of zero single precision error, computing cost and consistency.*

Keywords: *Breast Tumors, Closed Circle Curve Fitting (CCCF), Digital mammogram image, optimal threshold, Edge Detection.*

1. Introduction

Breast cancer is the most common malignancy among women, and it is also generally the world wide leading cause of mortality in women each year. The incidence of breast cancer in India is on the rise and is rapidly becoming the number one cancer in females pushing the cervical cancer to the second spot. The seriousness of the situation is apparent after going through recent data from Indian Council of Medical Research (ICMR). The rise is being documented mainly in the metros, but it can be safely said that many cases in rural areas go unnoticed. Due to deadly effect of this disease, the importance of analyzing its characteristics for further diagnostic purpose has increased. It is generally observed that the shape of the breast can be changed significantly with respect to its position and size of the tumors.

Due to terrific nature of the disease, we are motivated to develop a simple algorithm to help the doctors during the diagnostic period. Many studies have applied the concept of fractals to analyze cells, tumors, and other regions of interest (ROI) in biomedical images. N. J. Lee analyzed and compared the size and shape of the tumors with the help of elliptical closed curve fitting methods [1]. Kikuchi *et al*. investigated the change in fractal dimension at different stages of ovarian tumor growth [2]. Miguel Alem-an-Flores. *et al.* also analyzed about shape of breast lesions with the help of digital mammogram images [3], [4].

Generally elliptical system is used to analyze the size of the object, where the major and minor axis are used to measure the radius and area of the object. But this method is costing the computational difficulty and also it needs proper scaling of objects.

The purpose of this work is to improve accuracy of the tumor analysis in simplified way [5]. The approach to tumor shape analysis presented here involves approximating the shapes of tumors by fitting them in circle and then categorizing the tumors based on the dimensions of the circle and the closeness of the approximation. Thus, the shapes of these masses projected onto two-dimensional images would appear round or near circular when segmented apart from the rest of the image. So for that, our proposed algorithm segments the tumor image by separating foreground images from background image. In segmented images, of the tumors appear in a variety of shapes. Referring to the assumption that tumors suggested for biopsy tend to be round to some degree, and have edges that form closed curves, the tumor shape algorithm is intended to be used to define a criterion by which true positives are generated by the automatic tumor identification systems. We refer to the tumor shape-fitting algorithm as the Closed Circle Curve Fitting (CCCF) system. The proposed our novel method is based on circle fitting method [6]. In fact, we adopt area matching technique for finding the degree of closeness between tumors and circle. Our proposed system has been divided into three parts. The first part is to separate the regions containing tumors from the image. Thus identified regions are segmented with the help of optimal threshold techniques [7], [8]. The second part of our proposed system is to extract the edges of the tumors for further analysis with the help of basic edge detection and boundary detection method. Then the third

*Corresponding author: sss.viji @gmail.com, ekhan@lycos.com.

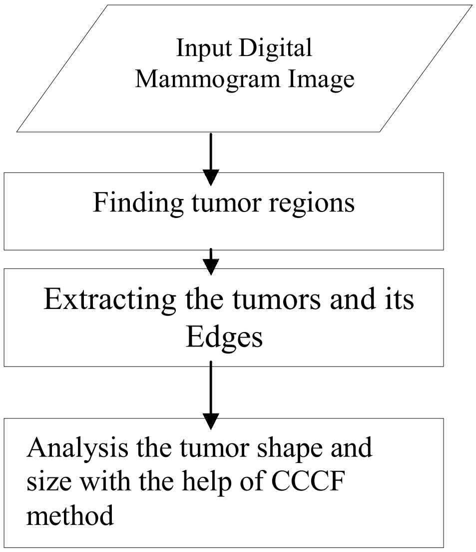


Figure 1: Flow Chart of Proposed System

part will end up with analysis of tumors with closed circle curve fitting algorithms. The equivalent flow diagram is shown below in Fig. 1.

The paper is organized as follows, Section. 2. Provides the brief review about the existing curve fitting algorithms. Section.3 explains about identifying and extracting tumors from digital mammogram images. Section 4 describes about our proposed algorithms. Computer simulations and results are included in section. 5. Conclusion and suggestion for further work has been discussed in Section 6.

2. EXISTING METHODS

Generally ellipse is used for analyzing 2D and 3D points as desired in various size of image analysis applications [9], [10], [11]. The ellipse is one of the most common geometric features for the application of image processing [12]. In the past, fitting problems have usually been solved through the least squares method (LSM) with respect to the effective implementation and acceptable computing costs [11], [13]. The main alternative methods for the detection and analysis of geometric features are Hough transforms and the moment method [9]. But all the above methods have drawback of large computing time and memory requirement with single precession error. Sung Joon Ahn et.al groups defined modified elliptical curve fitting algorithm [14]. Even though it might be suffered by repeated iteration, and unavailability of nearest point on the ellipse and Jacobian matrix. Radim Hali.et.al also defined elliptical curve fitting algorithms [15]. But this method is suffered from single precession error. In order to avoid the above said draw backs and complex calculations; we proposed a closed circular curve fitting algorithms for analyzing size and shape of the tumors.

3. EXTRACTING TUMOR IDENTIFIED REGIONS

In this section we identified and extracting the tumors area through optimal threshold segmentation method [7], [8]. The foreground and background are separated using optimal threshold technique which initially takes four 'almost' arbitrary pixel values as background, whereas the rest of the image is regarded as object (these points are only almost arbitrary, since they should be situated in the vicinity of the respective corners of the image in order to insure stable results) [7], [8],[16]. At the next step, the means for the background as well as the object gray-level values are computed using equation (1) and (2) as follows

$$\mu_B^t = \frac{\sum (i,j) \in back\ ground\ f(i,j)}{\neq back\ ground - pixels} \tag{1}$$

$$\mu_o^t = \frac{\sum (i,j) \in object\ f(i,j)}{\neq object - pixels} \tag{2}$$

from which the threshold is computed as

$$T = \frac{\mu_B^t + \mu_o^t}{2} \tag{3}$$

This procedure is iteratively repeated until value of T at two successive iteration becomes the same i.e. $T^{(t+1)} = T^{(t)}$

Tumors identified areas can easily be extracted by thresholding the image with the above threshold with respect to its background image.

4. PROPOSED METHOD

After extracting the tumors from digital mammogram images the shape of the tumors have been identified with the help of closed curve edges [4]. The shape of tumors is more likely to be elliptical or circle forms. The closed circle curve fitting algorithms survived as a post processing of automatic tumor detection method [9]. The CCCF system comprised various sections like edge detection method, centroid calculation, curve fitting with area matching[6].

4.1. Circle Quantities

Since the structural analysis of the CCCF system is based on circle quantities, we considered two quantities to describe the circle [6]. These are area and radius of the circle. Any 2D circle can be completely described given its area and its radius [9]. The area of the circle is given as pi times the square product of its radius. Where area is

$$Area = \pi * r^2 \tag{4}$$

$$circumference = 2 * \pi * r \tag{5}$$

where the eccentricity of the circle is zero. The radius of the circle has been calculated from its polar co-ordinate with respect to its Cartesian co-ordinates.

4.2. Edge Detection Methods

The pixels defining the shape of the tumor are of interest. The most of the algorithms used to approximate tumor shape operate only for pixels lying on the edge of the tumors. All pixels inside the edge boundary therefore need to be removed. Thus, after segmenting foreground image (tumor) from its background, the next operation performed is the edge detection. The edges in an image can be found by computing the Laplacian on that image. An image can be represented as a two dimensional surface $I(x, y)$, where $I(x, y)$ represents pixel intensity at pixel co-ordinates (x, y). Edges in the image exist in places where the intensity of the pixels is varying. More precisely, edges exist where the partial derivate $\frac{dI(x,y)}{dx}$ or $\frac{dI(x,y)}{dy}$ are not equal to zero. Thus edges detected can be detected by computing the derivative of $I(x, y)$, with respect to x and y or the gradient of $I(x, y)$, $\nabla I(x,\ y) = \frac{dI(x,y)}{dx} + \frac{dI(x,y)}{dy}$. The gradient will have negative or positive values at edges, but the absolute value of the gradient is the only thing of interest. We take the dot product of $\nabla I\,(x, y)$ with itself. Thus the edge-detected images are essentially the Laplacian of the original image $I(x, y)$, i.e. $I(x, y)$, i.e. $\nabla^2 l\ (x, y)$.

The Laplacian is seldom used by itself for edge detection because, as second order derivatives, it is unacceptably sensitive to noise, its magnitude produces double edges and it is unable to detect edge directions.

With respect to above discussion as a background, the basic idea behind edge detection is to find places in an image where the intensity changes rapidly. Fig. 2 shows an example of the four images of breast tumors Fig. 2(a-d) and its corresponding edge detected images Fig. 2(e-h).

4.3: Centroid Calculation

The Cartesian coordinates of a centroid are the mean value of the coordinates of the set of vertices. Fig. 3 contains a simple illustration of the centroid corresponding to a set of three points Ti, where $i = 1, 2, 3$.

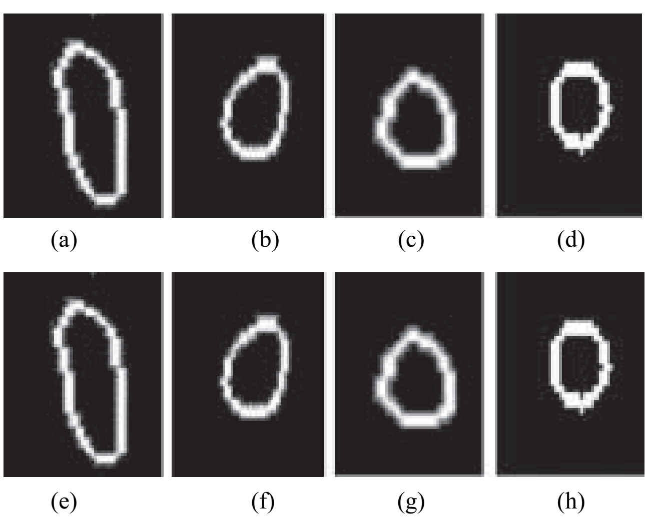


Figure 2: (a), (b), (c) and (d) are Breast tumors; (e), (f), (g) and (h) Corresponding Edge Extracted Images

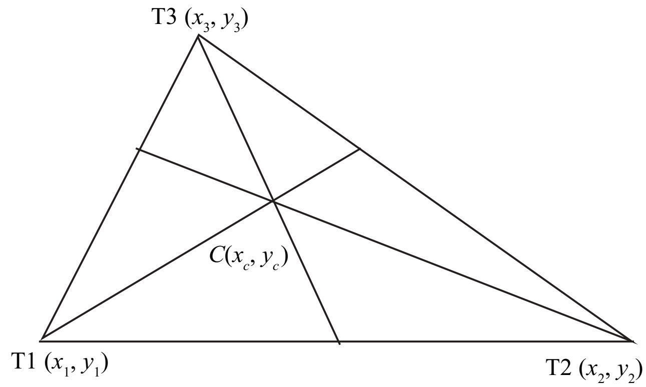


Figure 3: Calculation of Centroid

The centroid of an image is a point at the intersection of the three medians of the image [17]. One of the basic properties of centroid is that it divides the medians in to a 2:1 ratio. The part of the median nearest the vertex is always twice as long as the part nears the midpoint of the side. If the co-ordinates of the image are known, then the centroid of the respective image has been calculated. If we call three vertices of the images as $T_1(x_1, y_1)$, $T_2(x_2, y_2)$, $T_3(x_3, y_3)$, then the co-ordinates of geo-center $C(x_c, y_c)$ would be

$$xc = \frac{x1 + x2 + x3}{3} \tag{6}$$

$$yc = \frac{y1 + y2 + y3}{3} \tag{7}$$

In general, for any given N number of points N the centroid may be calculated with the help of following equations:

$$xc = \frac{\sum_{i=1}^{N} mixi}{\sum_{i=1}^{N} mi} \tag{8}$$

$$yc = \frac{\sum_{i=1}^{N} miyi}{\sum_{i=1}^{N} mi} \tag{9}$$

where m_i's are the weighting factors, which represents the area of the particular image.

The vertices $T(x_i, y_i)$, are the coordinates of the pixels forming the shape of tumor. $C(x_c, y_c)$, represents the location of the tumor and serves as a unique and non arbitrary point of origin for constructing the fitted circle.

4.4. Mapping Cartesian Coordinates to Polar Coordinates

With the centroid serving as a reference position for the analysis of image such as tumor, we now create the matching element for shape and size to the respective tumor with circle as close as possible. For this purpose we convert the above Cartesian co-ordinates in to polar format. The radii of the circles have been drawn using polar coordinates.

In order to implement our proposed idea's, we required to convert all the data points from Cartesian plane to single point in a polar plane.

$$T(x_i, y_i) = r_i \angle \theta_i \tag{10}$$

The 2-D polar coordinate system involves the distance from the origin and an azimuth angle. Fig.4 shows the 2-D polar coordinate system, where r is the distance from the origin (centroid in our case) to point T and θ is the azimuth angle measured from the horizontal (X) axis in the counterclockwise direction. Thus, the position of point T is described as (r, θ) where r & θ are the 2-D polar coordinates.

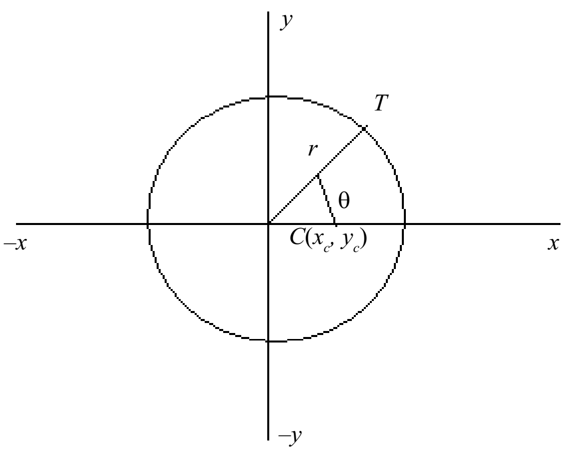


Figure 4: Plot Cartesian Co-ordinates in to Polar Co-ordinates

The respective relationship between Cartesian to polar coordinates is derived from following equations.

$$r = \sqrt{x^2 + y^2} \qquad \theta = \tan^{-1}(y/x) \tag{11}$$

where "r" represents the distance of the point "T" from its center "θ" represents azimuth angle of the distance in anticlockwise directions.

$$\textit{where} -\pi \le \theta < \pi \quad \textit{or} \quad x = r.\cos\theta \quad y = r.\sin\theta \tag{12}$$

4.5. Area Matching

For measuring the amount of fitness between two objects, we have to measure the area matching between these two objects. The area of the tumor is computed by counting the number of pixels of the image containing tumors. It is simply the number of white pixels in the original overlay (block and white) image.

$$A_{tumor} = \sum_{i}^{m,n} I(x, y) \tag{13}$$

where $I(x, y) = 1$ for white pixels, $I(x, y) = 0$ for block pixels. The area of the circle can be computed as:

$$Area = \pi * r^2 \tag{14}$$

The amount of fitness is measured with the help of following formulas:

$$\text{Area matching} = \frac{Area\ of\ circle}{Area\ of\ tumors} = \frac{Area}{A_{tumor}} \tag{15}$$

If the Area matching close to 1, it is an indication that tumors are closely fitting with constructed circles. The area match indicator is a factor, which indicates the closeness of fit.

Fig.5 displays the concept of above said area matching technique and shows how circle is fitted with the analyzing image. The fitted circle is shown in red. The shape of the tumor is represented by the point set *Ti* connected by green lines. The center of the circle and centroid of the tumor are positioned to coincide. The smaller the area matching values of fitted circle with tumors displays closeness of fitness between two objects. A circle area matching value is equal to one means the target has an identical shape as the circle. The higher the shape conformity value, the more dissimilar the target shape is from the fitted circle.

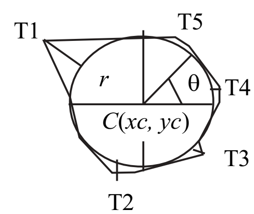


Figure 5: Area Matching between Fitted Circles with Original Tested Tumors

Generally eccentricity is the kind of parameters which can make a difference between circular and non circular shape of objects. Especially the eccentricity values of circular object is zero and thus values find the amount of closeness of fitting image with circle i.e. it define the accuracy of our method.

5. EXPERIMENT AND RESULT

The proposed technique is implemented by using Matlab 7.1 software on Window XP plateform. The necessary data base has been drawn from ICMR (Indian cancer Medical Research Institute) India and university of Florida, USA [18], [19]. We have analyzed the performance of proposed CCCF method in terms of six parameters. These parameters are sufficient to characterize the tumor in terms of its shape and size.

With the help of proposed CCCF method, we find the values of centroid, eccentricity, diameter, and area of each tumor. The eccentricity value gives the idea about shape of the tumors such that whether it is a circular or non circular form like elliptical form. For elliptical shape or non circular form of tumors, the eccentricity values lies between "0.5 to 1". And also we calculate the accuracy in terms of area matching between fitted circles with tested tumors. It is evident from our experimental results that area matching is constantly varied in the range of 90% to 93.16%. Here we compared our proposed CCCF algorithms with ECCF algorithms in terms of accuracy. We exposed accuracy of our proposed method with ECCF method in terms of area matching parameters. Table 2 compares the area matching values of CCCF and ECCF methods. The corresponding graphs are displayed in Fig. 6. Fig. 7 compares the average accuracy of two methods in terms of area matching.

Table 1
Original Tumors and its Measured Parameters

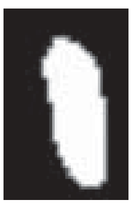
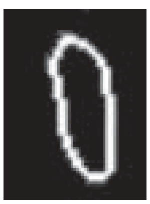
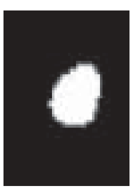
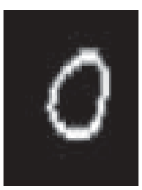
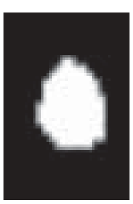
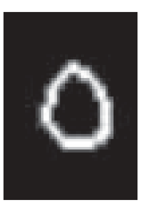
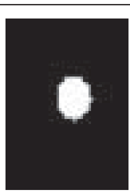
Sl. No	*Original Tumors*	*Corresponding Edge Detected version of Original Tumors*	*Corresponding Experimental Values*
1.			1. Centroid: [25.3637, 50.1242] 2. Radius [pxw]: 29.4082 3. Diameter: 58.8165 4. Eccentricity: 0.9387 5. Area of Circle: 2.7107e + 003 6. Area of Tumors [px]: 2964 7. Area matching: 0.9167
2.			1. Centroid: [30.618, 30.5809] 2. Radius [pxw]: 20.1045 3. Diameter: 43.3051 4. Eccentricity: 0.6153 5. Area of Circle: 1.4729e + 003 6. Area of Tumors[px]: 1581 7. Area matching: 0.9316
3.			1. Centroid: 34.6024, 28.9412 2. Radius [pxw]: 18.3117 3. Diameter: 36.6255 4. Eccentricity: 0.4795 5. Area of Circle: 1.0534e + 003 6. Area of Tumors[px]: 1157 7. Area matching: 0.9075
4.			1. Centroid: 24.5742, 22.2428 2. Radius [pxw]: 15.9395 3. Diameter: 31.9789 4. Eccentricity: 0.3584 5. Area of Circle: 800.6763 6. Area of Tumors[px]: 869 7. Area matching: 0.9214

Table 2
CCCF versus ECCF

Sl. No.	*Tested Tumors*	*Area Matching In CCCF Method*	*Area Matching in ECCF Method [1]*
1	Tumor 1	0.9167	0.7697
2	Tumor 2	0.9105	0.9075
3	Tumor 3	0.9316	0.9052
4	Tumor 4	0.9214	0.9507
	Average	0.92005	0.883275

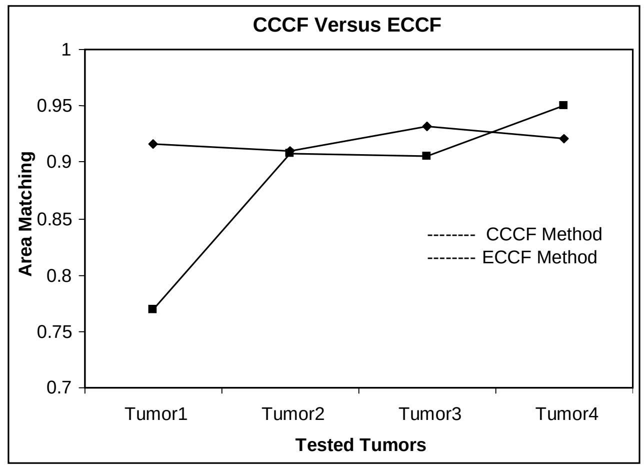


Figure 6: Comparison of CCCF Vs ECCF

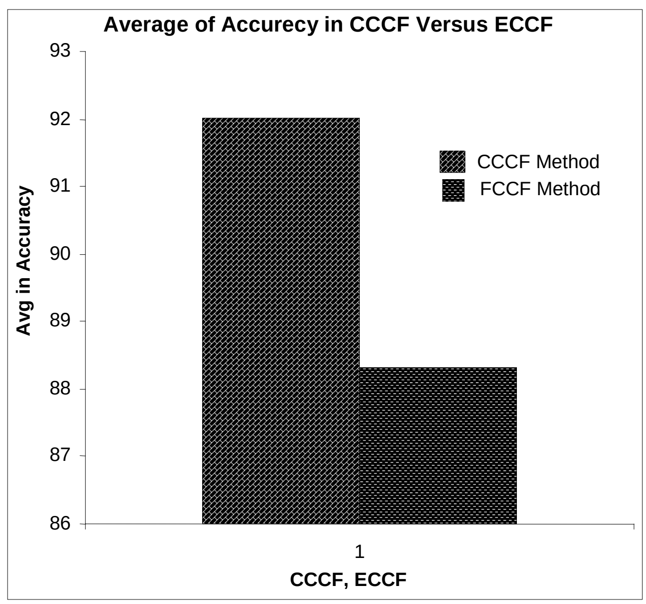


Figure 7: Average Accuracy of CCCF versus ECCF

From fig. 6 and 7, it is evident that our proposed CCCF method has superiority than ECCF method in terms of accuracy i.e. area matching. For some tumor shapes, ECCF system has better performance than the CCCF system, but the proposed CCCF system has consistency in the range of 90% to 93%.Therefore, on an average basis, the proposed

CCCF method has 92% accuracy, whereas ECCF has only 88.3275% accuracy. In addition, our proposed algorithm is simpler to design and implement as compared to ECCF algorithm

The main objective of CCCF system is to extract the information about the tumors for helping the specialist for further diagnosis. Doctors can easily understand about numerical measurement of tumors for data mining. The quantity of tumors area and circumference are always approximate. Eccentricity will confirm the shape about the tumors.

6. CONCLUSION

In this paper, we have proposed a novel method of determining the shape and size of breast tumors. The effectiveness of proposed CCCF system was verified for varies types of breast tumors with various sizes. The simulated results have proved the efficiency of the CCCF system. Comparing with elliptical based analysis our proposed system has superiority in terms of cost of calculations, accuracy and consistency. Especially our CCCF system is not hanging with single precession error, which will appear in ellipse based analysis. In addition to this, our proposed method define an easiest way to find and extract the tumors area in a digital mammogram image with the help of adaptive threshold based segmentation methods. In proposed CCCF method, tumor shapes are also identified by measuring tumors eccentricity values. By utilizing the above said CCCF method, doctors can easily understand the numerical data about various breast tumors and also they can utilize the things for data mining.

REFERENCES

[1] N. J. Lee., "Computer-Aided Diagnostic Systems for Digital Mammograms," *Master of Science Dissertations*, Louisiana State University, Louisiana, (2006).

[2] Kikuchi K., Kozuma S., Sakamaki K, Saito M, Marumo G., Yasugi T., and Taketani Y. Fractal Tumor Growth of Ovarian Cancer: Sonographic Evaluation. *Gynecologic Oncology*, **87**, (2002), 295–302.

[3] Miguel Alem-an-Flores.et.al," Shape Analysis for Ultrasound Breast Lesion Evaluation," **25**, (1999), 22–27.

[4] Harris, J. R., Morrow, M. and Bonadonna, G. "Cancer of the Breast", in *Cancer: Principles and Practice of Oncology*, De Vita, V. T. Jr., Hellman,S. and Rosenberg, S.A. (eds.), J. B. Lippincott Company, Philadelphia, (1993) .

[5] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed., Springer, Berlin, (2003).

[6] N. I. Chernov and G. A. Ososkov, Effective Algorithms for Circle Fitting, *Computer Physics Communications,* **33**, (1984), 329–333.

[7] T. Pun, "A New Method for Gray-level Picture Thresholding using the Entropy of the Histogram," *Signal Processing***,** **2**, (1980), 233–237.

[8] Frank Chung-Hoon Rhee *et al.*, "A Fast Numerical Method for Finding the Optimal Threshold for Image Segmentation," **22**, (1990), 984–989.

[9] Yin, R. K. K., Tam, P. K. S. and Leung, N. K.: Modification of Hough Transform for Circles and Ellipses Detection using 2-D Array. *Pattern Recognition*, **25**, (1992), 1007–1022.

[10] W. Gander, G. H. Golub and R. Strebel, Least-squares Fitting of Circles and Ellipses, *BIT,* **34**, (1994), 558–578.

[11] G. H. Golub and C. Reinsch, Singular Value Decomposition and Least Squares Solutions, *Numerische Mathematik,* **14**, (1970), 403–420.

[12] R. C.Gonzalez, R. C. P. Wintz. Digital Image Processing, Addioson-Wesley Publishing Company. USA, (1987).

[13] Gander, W.: Least Squares with a Quadratic Constraint. *Numerische Mathematik*, **36**, (1981), 291–307.

[14] Sung Joon Ahn *et al*., Ellipse Fitting and Parameter Assessment of Circular Object Targets for Robot Vision, *IEEE International Conference Proceedings*, **1**, (1996), 525–531.

[15] Radm Hali. *et al*., "Numerically Stable Direct Least Squares Fitting of Ellipses," **9**, (1996), 221–229.

[16] N. Otsu. "A threshold Selection Method from Grav-level - Histogram," *IEEE Trans. Sysi. Man Cybern*.**,** **8**, (1979), 62–66.

[17] W. Wernick, "Triangle Constructions with Three Located Points," *Math. Mag.*, **55**, (1982), 227–230.

[18] Cancer Research in ICMR Achievements in Nineties – Survey Reports.

[19] *http://marathon.csee.usf.edu/Mammography/Database.html*