Breast Tumor Shape and Size Analysis for Growth Estimation Using Mammography

Sonali Bhadoria¹, C.G. Dethe², Shrikanta Patra³, Ashutosh Chaubey⁴

¹sonali_nakade@yahoo.com, ²cgdethe@yahoo.com

Abstract: Analysis of tumor size is very important input for the doctors in deciding the stage of the cancer and surgical approach. This affects treatment decisions, and hence plays a significant role. This paper proposes the effective methodology for analysis of shape and size of the tumors present in the breast using mammograms. It is a powerful tool to assist the doctors in the treatment related to breast cancer.

General Terms: Mammogram, Thresholding; Area

1. INTRODUCTION
Breast cancer is the leading cause of cancer deaths among women in many countries. Breast cancer incidence and death rates generally increases with age according to studies. During 2000-2004, 95% of new cases and 97% of breast cancer deaths occurred in women aged 40 and older [1],[2]. Amongst the many imaging techniques available for breast cancer detection mammography is considered as the most effective method. Breast tumors can be very easily extracted from these mammograms and can be further used for analysis. Fig 1 shows an example of a mammogram which is the x-ray image of the breast.

![Example of a mammogram](image)

Most of the research done related to mammograms aims towards the computer aided detection of mammogram tumors which gives the classification of mammograms. In addition of this better knowledge of size and shape of the tumor growth will benefit the evaluation of screening programs [3], as well as the interpretation of clinical trials and epidemiological studies. It has been observed that The mean time a tumor needed to grow from 10 mm to 20 mm in diameter was estimated as 1.7 years, increasing with age.[4]

2. SIGNIFICANCE OF SIZE IN MEDICAL TERMS
To stage cancer, the American Joint Committee on Cancer first places the cancer in a letter category using the tumor, nodes, metastasis (TNM) classification system. Here we concentrate on T that indicates tumor size. The stage of a breast cancer describes its size and the extent to which it has spread. The letter T is followed by a number from 0 to 4, which describes the size of the tumor and whether it has spread to the skin or chest wall under the breast. Higher T numbers indicate a larger tumor and/or more extensive spread to tissues surrounding the breast.

**Stage I**: early stage breast cancer where the tumor is less that 2 cm, and hasn't spread beyond the breast(T1)

**Stage II**: early stage breast cancer in which the tumor is either less than 2 cm across and has spread to the lymph nodes under the arm; or the tumor is between 2 and 5 cm (with or without spread to the lymph nodes under the arm); or the tumor is greater than 5 cm and hasn't spread outside the breast (T2).

**Stage III**: locally advanced breast cancer in which the tumor is greater than 5 cm across and has spread to the lymph nodes under the arm; or the cancer is extensive in the underarm lymph nodes; or the cancer has spread to lymph nodes near the breastbone or to other tissues near the breast(T3)

**Stage IV**: metastatic breast cancer in which the cancer has spread outside the breast to other organs in the body(T4)
3. METHODOLOGY
We have done the experimentation on MAIS data base. Entire work flow is shown in the fig 2.

![Flowchart of the methodology](image)

**Fig 2: Flowchart of the methodology**

i) Initially mammogram image was segmented in two stages: a) Preprocessing which is thresholding the image to remove normal breast tissue b) Postprocessing which involves removing of unwanted part like pectoral muscles and other labeling.

ii) In second stage we find no of tumors present in the mammogram and extracts various shape feature to analyze the properties of the tumors.

A. Image Segmentation

Image segmentation is a very first step in processing the mammogram image for further analysis. [5]. Segmentation of mammograms can be done in two basic steps. In the first step we aim to remove the unwanted breast part and in the second we try to remove the pectorial muscles. This gives us the exact ROI. It is required to segment the mammographic images into various texture regions representing different tissue types. The main objective behind the segmentation of the medical image is to separate the tumor from the background. [6]. The increase in size of medical image database, has led to the use of computers in facilitating their processing and analysis. Estimation of the volume of the whole organ, parts of the organ and/or objects within an organ i.e. tumors is important step in the analysis of medical image [7].

There are various problems which we face during the segmentation as mentioned here. Raw digital mammograms are medical images that are difficult to interpret [8]. Lesser radiation dose of mammographic images degrades the contrast and the overall visibility of the micro calcifications and tumorous mass regions from the surrounding tissue. Due to presence of edges of low signal-to-noise ratio and complicated structured background, detection of a subtle mass on a mammogram is a difficult task [9]. There is a problem in segmentation of mammogram to a simple fatty and non-fatty set of regions due to large differences in parenchymal type appearances and variability of image acquisition parameters [10]. Pectoral muscle detection is a challenging task because it is not very well differentiated from the surrounding breast tissue. There is very small intensity variation of the pectoral muscle and the tumor tissue for each mammogram image. Due to the presence of pectoral muscle detection procedures gets biased, and hence should be removed during mammogram pre-processing [11].

![Original Image](image) ![Thresholded Image](image)

**Fig 3: 9a) Original Image b) Thresholded Image**

a) Preprocessing

The preprocessing phase aims at locating the region containing tumor and eliminating the normal breast tissue from the mammogram.[12] In this way, the size of the image to be segmented is reduced, and the complexity of the Post-segmentation phase is decreased. That intensity ranges of pectoral muscles also
Breast Tumor Shape and Size Analysis for Growth Estimation Using Mammography

falls in the range of tumor so that part cant be removed in this stage. There are various ways to achieve preprocessing

**Thresholding:** There is a significant intensity difference in normal breast tissue (background) and tumor (foreground). Perfect threshold can easily separate foreground from background and can be calculated from Otsu’s method or mean max method. This is one of the very simple way segmentation however it cant give us the accurate results. Fig 3 shows the result of thresholding.

**ii) Contrast Stretching:** Contrast stretching is the simplest method used for enhancement. In this we adjust the histogram to achieve a greater separation between the foreground and background gray-level distributions. Accurate intensity used can give very good results. Here contrast is adjusted as shown in fig 4(b) and then subtracted from the original image as show in in fig 4 (c). Complement is of the result is shown in fig 4(d).

**iii) K mean Clustering:** K-mean clustering is one of the method for post processing which forms a cluster of equal intensity region. Output is shown in figure 5.

*b) Post processing*

This phased aims to extract the exact ROI by removing pectoral muscles and other labeling. Pectoral muscles are the regions in mammograms that contain brightest pixels. These regions must be removed before detecting the tumor cells so that mass detection can be done efficiently. Pectoral muscles lie on the left or right top corner depending on the view of the image. We must detect the position of the pectoral muscles (left top corner or right top corner) before removing it [15]. Extracting the pectoral muscle [16, 17, 18] is particularly important in automated mammogram image assessment. Segmentation of the pectoral muscle is a non-trivial, complex and demanding task. It is also complicated due to a number of factors. Firstly, the muscle edge is not a straight line, but can be convex, concave or a mixture of both. Secondly, muscle edge though may appear to be visually continuous; the edge exhibits variations in texture and sharpness [19].

We propose a method of region labeling to remove the unwanted region. Binary form of output of preprocessing can be given to the input of post processing unit to get the different labeled regions. We can get lot many regions which are not of use. For this we need to use the filters which give the regions of significant area. Fig 6(c) shows the different labeled regions. ROI is tumor so we can select this manually using GUI or we can use some filters. Area based filter may work on some images but when nodule size is very small it will not work. Here we get the exact ROI as shown in 6 (d).
Breast Tumor Shape and Size Analysis for Growth Estimation Using Mammography

![Image](image_url)

Fig 6 mdb002 (a) Original Image (b) Preprocessed Image (c) Region separated Image (d) ROI

### B. Analysis

**Table 1**

<table>
<thead>
<tr>
<th>IMAGE</th>
<th>Area</th>
<th>Orientation</th>
<th>Extent</th>
<th>Centroid1</th>
<th>Centroid2</th>
<th>Solidity</th>
<th>Eccentricity</th>
<th>MajorAxis</th>
<th>MinorAxis</th>
<th>Equiv Dia</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42534</td>
<td>77.23400336</td>
<td>0.443719</td>
<td>461.4246</td>
<td>613.1744</td>
<td>0.665863</td>
<td>0.6505391</td>
<td>299.3395</td>
<td>227.3402</td>
<td>232.7144</td>
</tr>
<tr>
<td>2</td>
<td>5434</td>
<td>19.35193206</td>
<td>0.4576</td>
<td>435.7799</td>
<td>545.8417</td>
<td>0.644144</td>
<td>0.7624502</td>
<td>116.3534</td>
<td>75.31063</td>
<td>83.17923</td>
</tr>
<tr>
<td>3</td>
<td>5349</td>
<td>-71.825432</td>
<td>0.195934</td>
<td>537.2956</td>
<td>712.8949</td>
<td>0.35915</td>
<td>0.9190483</td>
<td>210.7098</td>
<td>83.05013</td>
<td>82.52611</td>
</tr>
<tr>
<td>4</td>
<td>11620</td>
<td>-88.009735</td>
<td>0.960886</td>
<td>295.4138</td>
<td>71.42332</td>
<td>0.992509</td>
<td>0.7940841</td>
<td>159.2949</td>
<td>96.82013</td>
<td>121.6349</td>
</tr>
<tr>
<td>5</td>
<td>22441</td>
<td>66.70600396</td>
<td>0.675242</td>
<td>351.0287</td>
<td>398.1726</td>
<td>0.876191</td>
<td>0.6057013</td>
<td>191.146</td>
<td>152.0934</td>
<td>169.0348</td>
</tr>
<tr>
<td>6</td>
<td>11056</td>
<td>-24.2517645</td>
<td>0.588461</td>
<td>388.1998</td>
<td>706.428</td>
<td>0.823845</td>
<td>0.5408504</td>
<td>133.0395</td>
<td>111.9012</td>
<td>118.6463</td>
</tr>
<tr>
<td>7</td>
<td>32251</td>
<td>-84.2534553</td>
<td>0.425515</td>
<td>555.6014</td>
<td>645.306</td>
<td>0.678126</td>
<td>0.8371597</td>
<td>308.2206</td>
<td>168.5839</td>
<td>202.6407</td>
</tr>
<tr>
<td>8</td>
<td>59762</td>
<td>-55.4876946</td>
<td>0.452648</td>
<td>594.1752</td>
<td>709.3074</td>
<td>0.705757</td>
<td>0.8516729</td>
<td>329.6086</td>
<td>172.7332</td>
<td>225.0334</td>
</tr>
<tr>
<td>9</td>
<td>5349</td>
<td>-71.825432</td>
<td>0.195934</td>
<td>537.2956</td>
<td>712.8949</td>
<td>0.35915</td>
<td>0.9190483</td>
<td>210.7098</td>
<td>83.05013</td>
<td>82.52611</td>
</tr>
<tr>
<td>10</td>
<td>1155</td>
<td>44.58315795</td>
<td>0.288606</td>
<td>574.6926</td>
<td>607.9472</td>
<td>0.517473</td>
<td>0.9110542</td>
<td>75.35104</td>
<td>31.57745</td>
<td>38.34829</td>
</tr>
</tbody>
</table>
Once we get this ROI various shape features can be extracted from the images. We can even find out number of nodules present in the image and their size. For 10 sample images having single tumor following features were extracted.

Output given by the code is shown below for mdb002:

Single Tumor
Analysis of Tumor:
Area = 22441
Orientation = 66.7060
Extent = 0.6752
Centroid1 = 351.0287
Centroid2 = 398.1726
Solidity = 0.8762
Eccentricity = 0.6057
MajorAxisLength = 191.1460
MinorAxisLength = 152.0934
EquivDiameter = 169.0348

2. Growth Estimation:-

As discussed earlier, size, shape of the tumor plays very important role. We can monitor or compare the growth of the tumors after and before the treatment by using these parameters. Let's assume subscript 1-Before treatment and 2-After treatment

Growth = (Area1-Area2)*100/A1......................................(1)

Similarly change in diameter, location, orientation can be found out. We can compare the tumors parameters as shown in fig 7. This comparison is based on the 10 samples used.

REFERENCES

[6]. Sonali Bhadoria, Preeti Agrawwal, CG Dethe, Renu Vig, ’Role of Segmentation in Medical Imaging: A Comparative Study
Breast Tumor Shape and Size Analysis for Growth Estimation Using Mammography

(Proceedings of the International Conference on Recent Trends in Engineering and Technology, 2010)


[11]. Styliani Petroudi and Michael Brady,”Breast Density Segmentation Using Texture”, Wolfson Medical Vision Laboratory, Oxford University, Oxford, OX@ 7DD, United Kingdom.


