CHAPTER 6
EXTRACTION OF CANCER CELL REGIONS AND IMAGE POST PROCESSING

6.1. Post processing methodology

Earlier we discussed about preliminary steps which are needed for any image processing process. Above algorithms will help us to get clear view of image with foreground and background separation.

6.1.1. Image Orientation

Orientation of medical diagnostic mammogram can be identified automatically by

- Obtaining mammogram data representing a multi-dimensional region bounded by a structure of tissue,
- Obtaining locations associated with multi-dimensional region as a function of the mammogram data.
- Obtaining an orientation of the tissue structure as a function of the locations;

Where determining the orientation comprises iteratively performing a morphological operation on the locations, a line being an output of process. The internal line is spaced away from the bounding tissue structure along the entire length by the regions of flow.

The thinned images are analyzed in different orientation levels.

Mammogram image of Breast cancer is always taken different in angle positions of detector module. So post processing of mammogram image analysis need the local orientation of a pixel in a mammogram image. This was computed by using its 16 x 16 neighborhood and this was slightly modified in this research by dividing the image into 16 x 16 blocks and the local orientation for each block centered at pixel $I(i,j)$ (Hong et al 1998).
After the preprocessing steps of mammogram of tumor image, pixel I is defined as a 16*16 matrix, where I (i, j) represents the intensity of the pixel at the i\textsuperscript{th} row and j\textsuperscript{th} column.

Mean and variance of I defined as

\[
M(I) = \frac{1}{16^2} \sum_{i=0}^{16-1} \sum_{j=0}^{16-1} I(i,j)
\]

And

\[
\text{VAR}(I) = \frac{1}{16^2} \sum_{i=0}^{16-1} \sum_{j=0}^{16-1} (I(i,j) - M(I))^2
\]

Respectively

An orientation image O is defined as an 16×16 image, where O (i,j) represents the local ridge orientation at pixel (i,j). Local ridge orientation is usually specified for a block rather than every pixel; an image is divided in to a set of 16×16 non overlapping blocks and a single local ridge orientation is defined for each block. Note that in a mammogram image, after orientation there is no difference between local orientation of images, since ridge oriented by 90° and the ridge oriented by 270° in a local neighborhood cannot be differentiated each other.

6.1.2. Image Triangulation

Breast tumor segmentation and analysis is an important step for doctors in deciding the stage of cancer and to proceed for further treatment. Segmentation of image is a crucial step in image processing which further helps in classification of image based on the features extracted. The segmentation technique in most of the approaches uses similar kind of algorithm for segmentation of region of interest.

Identifying tumor regions with clear edges is a critical step in image post processing step of image processing algorithm. If we use edge detection method for finding clear edges of mammogram image it will not give efficient result in image with too many
edges and non-trivial in producing closed boundary and curve’s. This thesis presents a new approach for preprocessing and segmenting out the infiltration and tumor regions from digital mammograms using two techniques involving iterative and non-iterative algorithms of Delaunay triangulation. The preprocessing involves segmentation and extending image through the intensity of image pixel. The iterative algorithm for segmentation works to get an idea of shape of infiltration/tumor in the breast. The proposed algorithm uses voronoi properties to partition an image into regions of similarity followed by Delaunay triangulation.

![Diagram of Delaunay triangulation](image)

**Figure 22.** (a) Non-Delaunay triangulation (b) Delaunay triangulation.

In the preprocessing steps X ray film Mammogram (Gray scale) is converted in to Digital Mammogram Image. The exact pixel value depends on the range of optical densities that the scanner is capable of finding the density difference of a pixel. Delaunay triangulation is rotation and translation invariant because it consists of the direction and location only relative to its some neighboring pixel. Delaunay triangulation has important characteristics

1. The Delaunay Triangulation of a non-degenerate set of points is unique.
2. A circle through the three points of a Delaunay triangle contains no other points.
3. If the circle is unique it is called circum circle of pixel.
DT net has two key characteristics as follows:

1. For each triangle in DT net, no data point is found inside the circumscribing circle.
2. Max-min angle. The smallest angle in DT net is larger than the smallest angles in other triangulation nets.

![Delaunay triangulation and Voronoi diagram](http://en.wikipedia.org/wiki/Delaunay_triangulation)

**Figure 23. (a) left-Delaunay triangulation (b) right-Voronoi diagram**


Algorithm Steps:

1. Take the input image as thinned Binary image.
2. Given a set S of pixels \(p_1, p_2, \ldots, p_N\). We can compute the Delaunay Triangulation by Voronoi diagram.
3. The voronoi diagram decomposes the image into number or regions around each pixel, such that the entire pixel in the region around \(p_i\) closest to \(p_j\) than they are to any points in S.
4. Let P be a circle free set. Three points p, q and r of P define a Delaunay triangle if there is no further point of P in the interior of the circle which is circumscribed to the triangle p, q, r and its center lies on a voronoi vertex.
5. If any triangle has interior point the triangle is replaced by new triangle.
6. Every time ensure that Triangulation is unique. This algorithm works by growing a current triangulation triangle by triangle.
In each iteration, the algorithm seeks a new triangle which attaches to the boundary of current triangulation.

- Vertices in the Voronoi diagram corresponds to infiltrations or tumors in Delaunay Triangulation, while Voronoi cells correspond with vertices of DT.
- After the Voronoi diagram for a set of points is constructed, Delaunay Triangulation is produced by connecting any two sites, whose Voronoi polygons share an edge.
- More specifically: let P be a circle free set. Three points, p, q and r of P define a Delaunay triangle if there is no further point of P in the interior of the circle which is circumscribed to the triangle p, q, rand its center lies on a Voronoi vertex.

![Figure 24. Delaunay construction (Source from Abbas Cheddad et al., 2008).](image)

P q r defines a Delaunay triangle when the center (S) of the circle circumscribed to p, q, and r is a Voronoi vertex

Convex hull

It is the smallest convex set that contains a set X of points in the Euclidean plane. Here convex set is a Euclidean plane, in which every pair of point is within the object, every point on the straight line segment that joins them is also within the object.

Empty circle property

Two sites p_i and p_j are connected by an edge in the Delaunay triangulation, if and only if there is an empty circle passing through p_i and p_j. (One direction of the proof is trivial from the circum circle property. In general, if there is an empty circum circle passing
through \( p_i \) and \( p_j \), then the center \( c \) of this circle is a point on the edge of the Voronoi diagram between \( p_i \) and \( p_j \), because \( c \) is equidistant from each of these sites and there is no closer site.

The contrast improvement by triangulation based on local information. Weak regions of the image are enhanced more than strong regions. This method is beneficial to experts when manually defining the edges for diagnosing purposes. Next section, we are going to discuss about the Euclidean Distance calculation between foreground and background of an image.

The advantage of this technique is, it works on the histogram of the image instead of the entire image. Hence, it is effective for large sized mammograms reducing the load on algorithm. This is a fully automated and unsupervised process. No parameter is needed for segmentation which is an advantage over other popular segmentation methods like k-means and watershed. The Voronoi and Delaunay segmentation are region growing method which looks for similarity in the images and segregates outs the high intensity region from the entire image. There is no need to set any parameter.

This step in Mammogram image processing algorithm is to get faster output (average execution time = 0.4 sec) as our method recursively draws a line to split the vertices into two sets and for each set, Delaunay triangulation is computed followed by merging two sets along the splitting line. Second, though our algorithm is apt for segmentation in mammograms, it can very well handle segmentation of other types of medical images.

6.1.3. Euclidean Distance Transformation

The distance transform is an operator normally applied only to binary images. The result of the transformation is a gray scale image that looks similar to the input image, except that the intensities of points are changed to show the distance to the closest boundary from each point.

Squared EDT method works well in an image pixel distance finding in terms of intensity but it involves the time consuming calculation such as square, square root and
the minimum over a set of floating point numbers. To overcome these difficulties in finding of distance through Distance transformation between pixels we are going to use two scan recursive algorithms by using 3×3 neighborhoods analysis. This algorithm only requires two image scans that are forward and backward raster scans in a binary image.

![Figure 25. Numerical representation of distance transforms.](image)

**Figure 25. Numerical representation of distance transforms.**

**Characteristic:**

1. The 8 neighbors of pixel p be denoted by q1, q2, .... q8. Thus N1(p) = {q1, q3, q5} and N2(p) = {q5, q7, q8}.

<table>
<thead>
<tr>
<th>q2</th>
<th>q3</th>
<th>q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1</td>
<td>p</td>
<td>q5</td>
</tr>
<tr>
<td>q8</td>
<td>q7</td>
<td>q6</td>
</tr>
</tbody>
</table>

f: A Thinned Binarized Image.

F: The set of Object pixels.

F₀: The set of Background boundary pixels.

o: The set of background boundary pixels

Q: The set of Foreground pixels which already have minimum squared Euclidean Distance through a Triangulation algorithm.
H (p, q): The Difference of the Squared Euclidean Distance of p and q (op^2-oq^2).
q\in N_1 \cup N_2.

G (p, q): The difference of the relative coordinates of p and q.

R (p) the relative coordinates. R_x, R_y of pixel p, which records the horizontal and vertical background pixel distance between p and closest background pixel. It is initialized as all (0, 0).

H (p, q) = \begin{cases} 
2R_x (q) +1 & \text{if } q \in \{q_1, q_5\} \\
2R_y (q) +1 & \text{if } q \in \{q_3, q_7\} \\
2(R_x (q) + R_y (q) +1 & \text{if } q \in \{q_2, q_4, q_6, q_8\} 
\end{cases} \quad (19)

G (p, q) = \begin{cases} 
(1, 0) & \text{if } q \in \{q_1, q_5\} \\
(0, 1) & \text{if } q \in \{q_3, q_7\} \\
(1, 1) & \text{if } q \in \{q_2, q_4, q_6, q_8\} 
\end{cases} \quad (20)

Algorithm:

Forward

1. Pixel belongs to the Binary image p\in F
2. Pixel value of Binary image may be infinite. f(p) = \infty
3. Select 3\times3 neighboring pixel
4. If q\in N_1, q = \{q_1, q_2, q_3, q_4\}
   4.a. f(p) = \min(f(p), f(q) + H(p,q))
   4.b. R (p) = R (q) + G (p, q)
5. If q\in N_2, q = \{q_5, q_6, q_7, q_8\}
   5.a. f(p) = \min(f(p), f(q) + H(p,q))
   5.b. R (p) = R (q) + G (p, q)
6. E (p) = \sqrt{f(p)}
Case 1: The smallest squared Euclidean distance obtained from $q_1$.

When $q_1 \in Q$ and $o \in O$

\[ o \right (0,0) \]

<table>
<thead>
<tr>
<th>$q_1(R_x, R_y)$</th>
<th>$P(R_x, R_y)$</th>
<th>$q_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_8$</td>
<td>$q_7$</td>
<td>$q_6$</td>
</tr>
</tbody>
</table>

Using above conditions $R(q_1) = ((R_x-1), R_y)$. Squared Euclidean Distance at $q_1$ is $oq_1^2 = (R_x-1)^2 + R_y^2$. Squared Euclidean Distance at $p$ will be $op^2 = R_x^2 + R_y^2$. Since $p$ is located on the one pixel away from $q_1$.

The difference of the SED of $p$ and $q_1$ according to equation (1) is

\[ H(p, q_1) = (op^2 - oq_1^2) \]

\[ = 2(R_x-1) + 1 \]

This satisfies the first condition of $H(p, q_1)$.

So $R_x(q_1) = R_x - 1$ and $G(p, q_1) = (1, 0)$

\[ R(p) = R(q_1) + G(p, q_1) \]

\[ R(p) = (R(q_1) + 1, R_y(q_1)) \]

Above producer can be repeated for finding the Smallest Euclidean Distance is obtained from $q_2, q_3, q_4$.

Same way, nearest background pixel located on the right top side also taken care for forward scan and relative coordinate $R(p)$ was calculated.

The Smallest Euclidean Distance of $p$ can also be obtained by another four neighborhoods $q_5, q_6, q_7$ and $q_8$ by reverse scan and $R(p)$ was calculated.
A double scan (first top-down, left-right and second bottom-up, right-left) is required to update them in distance.

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Binary Image

Distance Image

In this step the binary image was again converted in to gray scale image with different intensity level than input gray scale image. It helps to find the area of cancer and spreading out of cancer cell in the mammogram image. Thus two scan algorithm using neighborhood Euclidean distance transformation algorithms use only a small image neighborhood and work with in the image itself and do not need any memory. Each pixel value is changed as a distance value shows the mammogram image with high level features to find the cancer cell area.