Textural Features Based Breast Cancer Detection: A Survey

1 Anupa Maria Sabu, 2 D. Narain Ponraj, 3 Dr. Poongodi
1, 2 Dept. of ECE, Karunya University, Coimbatore, India
3 Dept. of ECE, PPG College of Technology

roopa.maria09@gmail.com, narampons@gmail.com, poongodiravikumar@yahoo.co.in

ABSTRACT

Breast cancer is a major cause of cancer deaths among women. Early detection plays an important role for improving breast cancer prognosis. Mammography is used to demonstrate the presence of breast cancer and to identify the size and location of tumor cells. Texture analysis refers to a procedure or a model that characterizes the spatial variation within the image by extracting information. In this paper we discussed various methods of texture analysis for mass detection and microcalcification in mammography and we also identified the texture features obtained.

Keywords: Breast cancer, Mammography, GLCM, GLRLM, LBP, Gray level difference statistics, Laws’ texture measure, Fractal based texture analysis.

1. INTRODUCTION

Cancer is a type of diseases that causes the cells of the body to change its characteristics and cause abnormal growth of cells. Most types of cancer cells eventually become a mass called tumor. Breast cancer is one of the major causes of death in women when compared to all other cancers [1]. Mammography is one of the imaging modality in early breast cancer detection typically through detection of characteristic masses and micro calcifications. Micro calcification is considered to be an important sign of breast cancer [2]. Visualization and detection of cancer cells in mammography play a crucial role in reducing the rate of mortality from breast cancer. Mammography is a low dose X-Ray procedure that allows the visualization of internal structure of the breast [3]. The Digital Mammography is shown in figure 1.

![Figure 1: Digital Mammography](image)

One method to identify the disease pattern in breast cancer is by means of texture analysis in mammography. Textures are one of the important characteristics for identifying objects and region of interest of various kinds of images [4]. Texture classification has achieved considerable interest for the past few decades and a large number of techniques have been introduced for the texture classification [5]. Texture classification is one of the four problem domains in the field of texture analysis. The other three are texture segmentation, texture synthesis and shape recovery from the texture [6]. An ideal texture descriptor must capture essential perceptual features of texture characteristics and must be invariant to changes in the viewpoint [7].

![Figure 2: Illustration of components of mammogram](image)

Texture analysis is widely used for computer vision and image processing for classification and segmentation of image based on a local spatial variation of intensity or color [8] [9]. Existing texture analysis can be classified into statistical and structural methods. Statistical approach computes different properties and are suitable if texture primitive sizes are comparable with the pixel sizes. In structural texture analysis method, the texture region is defined to have a constant texture if a set of local statistics or other local properties of the image are constant, slowly varying or approximately periodic [10]. Majority of the existing texture analysis assumes that the texture images are acquired from the same viewpoint (same scale and same orientation). This gives a limitation of these methods. In many practical applications, it is very difficult or impossible to ensure that images captured have the same translations, rotations or scaling between each other. Texture analysis should be ideally invariant to viewpoint [11]. But recently model based method is also employed which includes autoregressive model, Gaussian Markov random fields, Gibbs random fields, World medal, wavelet model, multichannel Gabor model and steerable pyramid, etc. These models provide more powerful tools for invariant texture analysis. In statistical methods, texture is described by a collection of statistics of selected features. The statistics are broadly classified into first-order statistics, second-order statistics, and higher-order statistics. In structural methods, texture is viewed as consisting of many textural elements (called Texel) arranged according to some placement rules. In model based a
texture image is modeled as a probability model or as a linear combination of a set of basic functions. The coefficients of these models are used to characterize texture images.

2. BRIEF REVIEW

Texture analysis is important for application of computer image analysis for classification, detection segmentation of an image based on intensity and color. Traditionally texture analysis can be broadly classified into two they are statistical and structural approach. In structural approach the texture can be represented by primitive (micro texture) and spatial arrangement of this primitive (macro texture). In the statistical approach represents the texture indirectly by the non-deterministic properties that govern the distributions and relationships between the gray levels of an image. There have been eight statistical approaches are used for the measurement and characterization of texture analysis [9]. They are Autocorrelation functions, optical transforms, digital transforms, textural edginess, structural elements, spatial gray tone co-occurrence probabilities, gray tone run lengths, and autoregressive models.

3. GREY LEVEL CO-OCCURRENCE MATRIC

GLCM is a statistical texture measure. GLCM collect information about pixel pairs, hence it is of second order statistics [12] [13]. GLCM is a tabulation of frequencies or how the pixel brightness values in an image occur. The matrix is constructed at a distance of d = 1 and for direction of θ given as 0°, 45°, 90° and 135°. A single direction might not give enough and reliable texture information. For this reason, the four directions are used to extract the texture information.

Thirteen textural features were extracted using GLCM energy, entropy, contrast, local homogeneity, correlation, and shade, and provenance, sum of squares, sum average, sum entropy, difference entropy, sum variance and difference variance. [14][15].

Gray level co-occurrence matrices capture properties of a texture but they are not directly useful for further analysis, such as the comparison of two textures. The co-occurrence matrix computation applied to a mammogram is shown in figure 3.

Computer image processing techniques are used to enhance the image which is followed by segmentation of region of interest. Then the texture feature will be extracted from ROI. The texture feature is used to classify the ROI as masses or non-masses. The extraction of texture feature can be done by using grey level co-occurrence matrices. Different steps involved in the texture analysis using GLCM is shown in figure 4.

4. GREY LEVEL RUN LENGTH MATRIC

Gray level run length matrix is a statistical texture descriptor. This method involves the counting the number of pixels that have the same intensity in a particular direction [16]. Run length is the number of adjacent pixels that have
same gray level intensity in a particular direction. The texture feature extracted from run length matrices produces great classification results. The texture features that can be extracted by GLRLM includes Short Run Emphasis (SRE), Long Run Emphasis (LRE), Low gray level Run Emphasis (LGLRE), High Gray Level Run Emphasis (HGLRE), Short Run Low Gray Level Emphasis (SRLGLE), Short Run High Gray level Emphasis (SRHGLE), Long Run Low Gray Level Emphasis (LRLGLE), Long Run High gray Level Emphasis (LRHGLE), Gray Level No uniformity (GLN), Run length no uniformity (RLN), Run Percentage (RPERC) [17][18].

5. LOCAL BINARY PATTERN OPERATOR

LBP operator combines the characteristics of statistical and structural texture analysis. The LBP operator is used to perform gray scale invariant two-dimensional texture analysis. The LBOP operator labels the pixel of an image by thresholding the neighborhood (i.e. 3 × 3) of each pixel with the center value and considering the result of this thresholding as a binary number [19]. When all the pixels have been labeled with the corresponding LBP codes, histogram of the labels are computed and used as a texture descriptor.

Given a pixel in the image LBP code can be computed by comparing it with its neighbors

$$LBP_{p,r} = \sum_{x=0}^{P-1} S(g_x-g_0)2^x, g_x > g_0 \quad 0, g_x < g_0$$ (1)

Where, $g_0$ is the gray level value of the central pixel, $g_x$ is the value of its neighbors. $P$ is the number of involved neighbors. After LBP pattern of each pixel is identified histogram is used to represent the texture image:

$$H(k) = \sum_{x=0}^{P-1} f(LBP_{p,r}(x,y), k, K = [0,K])$$ (2)

$$f(x,y)= \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases}$$

$K$ is the maximal LBP pattern value [20]. The most important properties of LBP features are computational simplicity and tolerance against the monotonic illumination changes [21]. The basic LBP operator cannot be used for the dominant features of large scale structures. So it is extended to facilitate the analysis of textures with multiple scales by combining neighborhoods with different sizes.

6. GREY-LEVEL DIFFERENCE STATISTICS

The Gray level difference method is based on the histogram of absolute difference between pairs of gray level features like edges, spots, line, ripples. Then the variance of measures are computed by first the image is convoluted with a 6. GREY-LEVEL DIFFERENCE STATISTICS

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$K$ is the maximal LBP pattern value [20]. The most important properties of LBP features are computational simplicity and tolerance against the monotonic illumination changes [21]. The basic LBP operator cannot be used for the dominant features of large scale structures. So it is extended to facilitate the analysis of textures with multiple scales by combining neighborhoods with different sizes.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$K$ is the maximal LBP pattern value [20]. The most important properties of LBP features are computational simplicity and tolerance against the monotonic illumination changes [21]. The basic LBP operator cannot be used for the dominant features of large scale structures. So it is extended to facilitate the analysis of textures with multiple scales by combining neighborhoods with different sizes.

7. LAWS’ TEXTURE MEASURES

The texture energy measures developed by K.I Laws have been used for many diverse applications. These measures are computed by first the image is convoluted with a number of small masks. Each of the masks will extract the features like edges, spots, line, ripples. Then the variance of each feature has to be computed over a moving window [23]. This will provide the local texture energy. This is shown in figure8. The texture feature like average gray level, edges, spot, ripple and waves in texture can be obtained [23].

Figure 5: LBP applied to mammogram

Figure 6: (a-c) ROI with masses (d-f) ROI without masses

Figure 7: Texture measure from Gray level difference statistics.

Figure 8: Texture analysis using laws texture energy measure
8. FRACTAL BASED TEXTURE ANALYSIS

Fractal and multi Fractal analysis have also been used for benign/malignant identification and micro calcification detection. In Fractal based texture analysis the fracture will measure the geometric complexity; which will describe the spatial pattern of textures [24]. The Fractal based texture analysis involves the calculation of the intrinsic Fractal dimension of input datasets [25]. The Fractal refers to the complex pattern that recurs at various scales. In Fractal texture analysis the Fractal dimension and Fractal signature provide a good measure of descriptive value of a region. [26]. The Fractal dimension is a statistical quantity which measure roughness of a geometric region [24]. The Fractal dimension provides the degree of linear independence and correlation between the available features. For an image, change in the gray level surface needs to be measured in different scales. The change in a measured area with different scale can be used as Fractal signature. Fractal based texture analysis for cancer and non cancer cases are shown in figure9.

![Fractal texture analysis](image)

(a) (b)

Figure 9: Fractal texture analysis of (a) cancer (b) non cancer on mammogram.

9. RESULT AND DISCUSSION

Table1.Summarizes six texture descriptors and the texture features that can be obtained from the descriptors. GLCM provides thirteen texture features. But the features that obtained from GLCM cannot be used directly in the further analysis. GLRLM characterizes the coarseness of texture in the specified direction. Eleven run length feature can be obtained from GLRLM. LBP provides rotation invariant two-dimensional texture analysis. Gray level Statistics provide the texture measures like mean, standard deviation, entropy, contrast. Laws’ Texture measures the average gray level, edges, spot, ripple and waves in the texture. Fractal Based texture analysis measures the geometric complexity.

<table>
<thead>
<tr>
<th>Textures descriptors</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLCM</td>
<td>contrast, energy, homogeneity ,correlation shade, provenance, sum of squares, sum average, sum entropy, difference entropy, sum variance and difference variance.</td>
</tr>
<tr>
<td>GLRLM</td>
<td>SRE, LRE, LGLRE, HGLRE, SRLGLE,</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Texture measure</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>Spatial structure of local image texture.</td>
</tr>
<tr>
<td>GREY-LEVEL DIFFERENCE STATISTICS</td>
<td>Mean, standard deviation, entropy, contrast.</td>
</tr>
<tr>
<td>LAWS’ TEXTURE MEASURES</td>
<td>Average gray level, edges, spot, ripple and waves in texture</td>
</tr>
<tr>
<td>FRACTAL BASED TEXTURE MEASURE</td>
<td>Fractal Dimension, Fractal signature</td>
</tr>
</tbody>
</table>

10. CONCLUSION

Texture analysis is a method to classify benign and malignant masses and to identify the micro calcification in mammography. This will help to identify the disease pattern of breast cancer in mammography. This paper discussed various texture analysis approaches for the detection of masses and micro calcification in mammography and identifies the texture features obtained from each texture descriptor.

REFERENCES


[20] Zhenhua Guo, Lei Zhang, Member, and David Zhang, ”A Completed Modeling of Local Binary Pattern Operator for Texture Classification” IEEE Transactions on Image Processing,2010


APPENDIX

Let $p(i,j)$ be the normalized co-occurrence matrix $R_g$ be the discrete gray levels of the image $P_{(i,j)}(i,j)$ be the marginal probabilities of rows and columns respectively. Then the formulae used in GLCM are as follows:

1). Contrast: $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{k=-r}^{r} \sum_{l=-f}^{f} p(i,j) (i-r) (j-f)$
2). Correlation: \[
\frac{\sum_{i,j} (f(i,j) - \bar{f})(g(i,j) - \bar{g})}{\sigma_f \sigma_g}
\]

3). Sum of squares: \[
\sum_{i,j} (f(i,j) - \bar{f})^2
\]

4). Sum of average: \[
\sum_{i,j} f(i,j)
\]

5). Sum entropy: \[
\sum_{i,j} R(i,j) \log R(i,j)
\]

6). Difference entropy: \[
\sum_{i,j} (R(i,j) - f(i,j))^2
\]

7). Sum variance: \[
\sum_{i,j} \text{var}(f(i,j))
\]

8). Difference variance: \[
\sum_{i,j} \text{var}(f(i,j) - R(i,j))
\]

Let \( P \) be the number of pixels in an image. \( p(i,j) \) be the \((i,j)\)th element of run length matrix for a specified angle \( \theta \) and a specified direction \( d \). \( N_r \) is the number of different run lengths that occur. Formulae that are used in gray level Run length matrix are as follows:

1). Short Run Emphasis: \[
\frac{\sum_{i,j} f(i,j) P(i,j)}{\sum_{i,j} f(i,j)^2}
\]

2). Long Run Emphasis (LRE): \[
\frac{\sum_{i,j} f(i,j) P(i,j)}{\sum_{i,j} f(i,j)^2}
\]

3). Gray level non uniformity: \[
\frac{\sum_{i,j} f(i,j) P(i,j)}{\sum_{i,j} f(i,j)^2}
\]

4). Run Level Non-Uniformity: \[
\frac{\sum_{i,j} f(i,j) P(i,j)}{\sum_{i,j} f(i,j)^2}
\]