[Note]

Breast Tissue Classification Using Statistical Feature Extraction Of Mammograms

Holalu Seenappa Sheshadri and Arumugam Kandaswamy

Department of Electronics & Communication Engineering
PSG College of Technology, Coimbatore, Tamilnadu 641004, India
(Received April 10, 2006, In final form June 2, 2006)

Abstract : In this paper authors have made an attempt to classify the breast tissue based on the intensity level of histogram of a mmammogram,Statistical features of a mammogram are extrcted using simple image processing techniques.The proposed scheme uses texture models to capture the mammographic appearance within the breast. Parenchymal density patterns are modeled as a statistical distribution of clustered filter responses in a low dimensional space. The statistical features extracted are the mean,standard deviation, smoothness, third moment, uniformity and entropy which signify the important texture features of breast tissue.Based on the values of these features of a digital mammogram, the authors have made an attempt to classify the breast tissue in to four basic categories like fatty, uncompressed fatty, dense and high density. This categorizaton would help a radiologist to detect a normal breast from a cancer affected breast so as to proceed with further investigation.This forms a basic step in the detection of abnormal breast under computer aided detection system. The results obtained out of the proposed technique has been found better compared to the other existing methods. The accuracy of the method has been verified with the ground truth given in the data base(mini-MIAS database) and has obtained accuracy as high as 78% This is a basic step in the development of a CAD for mammo analysis being developed at the department of ECE in support with thePSG Research centre at Coimbatore-India.
Key words : Breast tissue classification, Feature extraction, Moments, Mammograms

1 INTRODUCTION

Texture segmentation has long been an important topic in image processing. Basically, it aims at segmenting a textured image into several regions with the same texture features. An effective and efficient texture segmentation method will be very useful in applications like the analysis of aerial images, biomedical images and seismic images as well as the automation of industrial applications [1]. Like the other segmentation problems, the segmentation of textures requires the identification of proper texture-specific features with good discriminative power. Generally speaking, texture feature extraction methods can be classified into three major categories, namely, statistical, structural and spectral [2]

In statistical approaches, texture statistics such as the moments of the gray-level histogram, or statistics based on gray-level co-occurrence matrix are computed to discriminate different textures [3]. For structural approaches, "texture primitive" (i.e. the basic element of texture), is used to form more complex texture pattern by grammar rules which specify the generation of texture pattern. Finally,in spectral approaches, the textured image is transformed into frequency domain. Later the extraction of texture features can be done by analyzing the power spectrum also.Various texture descriptors have been proposed in the past. In addition to the aforementioned methods, Law's texture energy measures, Markov random field models, texture spectrum etc. are some the other texture descriptors

Even though several methods of tissue classification is available in literature, not a single method is applicable for all the biomedical imaging modality. Hence the authors have made an attempt to classify the breast tissues using statistical moments based on histograms.

For a biomedical image like mammogram, the characteristics of the pixels in the texture pattern are not similar everywhere from a global view point. To circumvent the above-mentioned issue, in this study, based on the fact that each interior point in a texture region must possess similar properties with its neighbors, a new statistical method is proposed [3, 4].

The key idea in the proposed technique is to classify the image into interior pixels and boundary ones, the interior pixels stand for the interior parts of texture regions, then the segmentation can be achieved by applying region growing on the interior pixels. To implement this idea, a novel approach based on the statistical properties of the intensity histogram is employed. By the use of statistical moments based on histogram the architectural distortion of a mammogram can be easily detected.This in turn would help a radiologist to investigate further the accurance of any abnormal tissue or unwanted growth in the breast.This typical classification of breast tissues has also been approved by FDA under the ACR-BIRDS to detect the amount of cancer affected area under regular screening mammograms.

1.1 Statistical Approaches

In general, any image processing and analysis applications would require a particular feature for classification /segmentation. Mainly texture features and statistical features are of more significant in pattern recognition area. A frequently used approach for texture analysis is based on statistical properties of intensity histogram. One such measures is based on statistical moments. The expression for the n^{th} order moments about the mean is given by

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \tag{1}$$

Where z_i is a random variable indicating intensity, $p(z_i)$ is the histogram of the intensity levels in a region, L is the number of possible intensity levels and

$$m = \sum_{i=0}^{L-1} z_i \; p(z_i) \tag{2}$$

is the mean (average) intensity. [4-6.] We have considered similarly other five more descriptors and their details are as shown below in Table 1. [5][7]. These moments have been computed with an matlab function and evaluated for various mammogram images Here the authors have considered the ROI of the images for the experimental analysis.

2 MATERIALS AND METHODS

The mathematical model for the moments to compute the six texture features of a mammogram are as listed in Table 1. Matlab functions to implement the same have been developed and tested over certain mammograms. The programs were tested over ten selected images from mini-MIAS data base. [7]. The results obtained are as tabulated in Table 2. The basic classification based on the values of the texture parameters are as shown in Table3. The original image and their corresponding histograms for basic classifications are also shown in Figures (1-4). From the above results it can be inferred that the statistical features extracted from the mammogram images are useful parameters for tissue classification [8]. This in turn also help in the diagnosis of palpable or micro calcifications or tumors in the breast. The classification of tissues based on the image features may be employed for early detection of breast cancer. Experimental results show a visual clue to any radiologist about the abnormality in the growth of breast.

Table 1 Descriptors of texture based on the intensity histogram of a region.

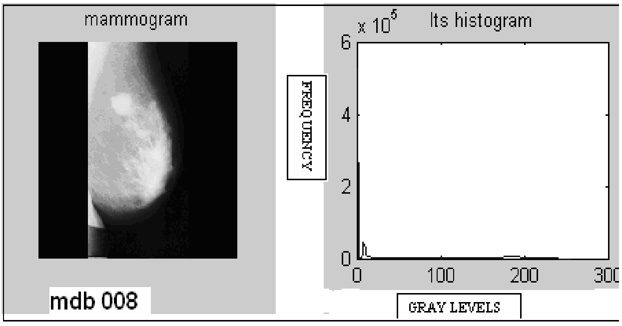
Moment	Expression	Measure of texture
Mean	$m = \sum_{i=0}^{L-1} z_i \; p(z_i)$	A measure of average intensity
Standard deviation	$\sigma = \sqrt{\mu_2 \; (z)} = \sqrt{\sigma^2}$	A measure of average contrast
Smoothness	$R = 1 - 1/(1+\sigma^2)$	Measures the relative smoothness of the intensity in a region.
Third Moment	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 \, p(z_i)$	Measures the skew ness of a histogram
Uniformity	$U = \sum_{i=0}^{L-1} p^2(z_i)$	Measures the uniformity. of intensity in the histogram
Entropy	$e = -\sum_{i=0}^{L-1} p(z_i) \; log_2 \; p(z_i)$	A measure of randomness

Table 2 Texture measures for the mammogram images (Examples).

Image Samples	Average intensity	Average contrast	smoothness	Third moment	Uniformity	Entropy
Mam 1	39.6760	42.8696	0.0275	0.6056	0.1663	4.7401
Mam 2	47.9076	1.9005	0.0736	6.2341	0.1910	4.6683
Mam 3	43.7049	46.3144	0.0319	0.4708	0.2156	4.4888
Mam 4	43.3234	40.3894	0.0245	0.2425	0.1030	5.4656
Mam 5	43.3946	40.4359	0.0245	0.2419	0.1036	5.4638
Mam 6	62.3899	68.4661	0.0672	2.1793	0.2332	3.2310
Mam 7	68.0774	71.3436	0.0726	1.6967	0.2472	3.0586
Mam 8	61.9692	74.2953	0.0782	3.7407	0.2058	4.9878
Mam 9	55.0435	81.8304	0.0934	8.8683	0.2557	4.4263
Mam 10	43.1755	69.3156	0.0688	6.1621	0.3507	3.9049

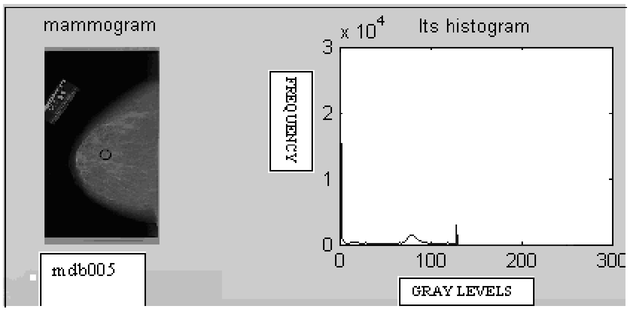
3 SIMULATED RESULTS

As the Mini MIAS database consist of 332 mammograms of different categories, it has been selected for the testing of performance of the proposed algorithms. As per the literature of the above database the mammograms have been grouped under only three categories like, fatty, glandular and dense. According to the recent research results of University of Calgary, (Biomedical Engineering Research Group), these have been further classified based on some statistical features in to the four classes as Uncompressed fatty, fatty, Nonuniform, and high density [9]. This classification would help a radiologist to determine the breast anatomy (fibroglandular tissue) affected due to Estrogen secretion. when a patient is under harmone replacement therapy (HRT) [10]. The HRT results in a number of unpredictable tissue changes which may not be detected by simple visual inspection of mammograms.Hence the proposed



a) Uncompressed fatty (mdb 008) b) Its Histogram

Fig.1 Mammogram (mdb 008) and its histogram [Uncompressed fatty breast]

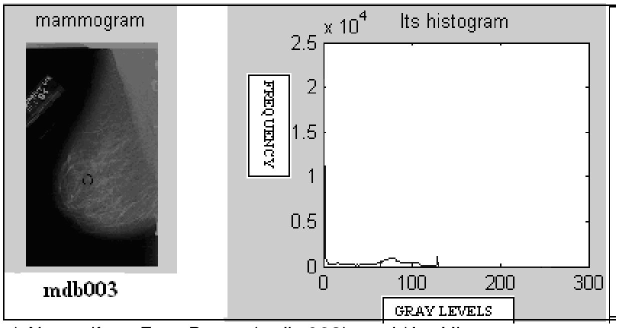


a) Fatty breast (mdb 005) b) Its Histogram

Fig.2 Mammogram (mdb 005) and its histogram [Fatty breast]

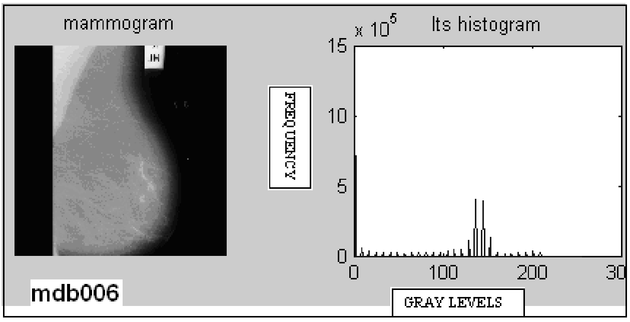
Table 3 Classification of mammogram images based on texture features.

Texture Feature → Tissue Categories ↓	Average intensity	Average contrast	smoothness	Third moment	Uniformity	Entropy
Uncompressed fatty tissue	43.7049	46.3144	0.0319	0.4708	0.2156	4.4888
Fatty tissue	62.3899	68.4661	0.0672	2.1793	0.2332	3.2310
Non uniform tissue	55.0435	81.8304	0.0934	8.8683	0.2557	4.4263
High density tissue	43.1755	69.3156	0.0688	6.1621	0.3507	3.9049



a) Non uniform Fatty Breast (mdb 003) b)Its Histogram

Fig.3 Mammogram (mdb 003) and its histogram [Non uniform fatty breast]



a) High density breast (mdb 006) b) Its Histogram

Fig.4 Mammogram (mdb 006) and its histogram [High dense breast]

method would help a radiologist in such detection of changes in breast tissues.

The following Figures 1-4 gives simulated results of the proposed method.

Note that the Figures (1-4) shown represent the four categories of breast tissues as examples These images are taken from mini-MIAS database. The simulated results agree with the classification as mentioned by the database. The histogram plot represent the frequency of pixel brightness against the number of gray levels as can be seen from the Matlab simulated plots.This forms a tool for segmentation of a digital mammogram.

4 CONCLUSION

The method employed here has given better performance. The results have been validated by visual inspection by an expert radiologist. Certain mammogram images (60) of different abnormalities randomly chosen from the data base have been selected for our experiment and the algorithm applied over it have classified them in accordance to the category as stated above. Further the results of our experiments show very clearly the appearance of any abnormality in the growth of breast tissue. Also this algorithm would simplify the time and computational complexity in analysis of any given mammogram.The proposed method of classification agree with the standard classification as prescribed by ACR-BIRADS (American college Of Radiology-Breast Imaging Reporting And Data systems). The only difficulty is in the selection of ROI for the application of the algorithm.

Our future work include the design and development of an expert system for real time mammogram image analysis. The system so designed would give the radiologist an idea about the exact shape and size of any tumor present in the breast.The other statistical methods of estimation of breast tissue density like the use of Gaussian mixture model are also being considered.

REFERENCES

[1] Hingham R, Brady M, “Mammography Image Analysis”, Kluwer Academic Publishers, 1999.

[2] Daniel B, Md Kopans, “Breast Imaging”, Lippincot Williams, 1997.

[3] Byng J.N, Boyd N.F. Fisheld, et al., The Qualitative analysis of mammographic densities, Physics in Medicine and Biology, Vol. 39, No. 10, pp. 1629-638, 1994.

[4] Karssemeijer N and te Brake G, “Detection of stellate distortions in mammogram”, IEEE Trans. on Medical Imaging, Vol. 15 (, No. 5, pp. 611-619, 1996.

[5] Linguraru M and Brady J, “A non-parametric approach to detecting microcalcifications”, International Workshop on Digital Mammography, Lecture Notes in Computer Science, Springer Verlag Berlin Heidelberg, 2002.

[6] W.K.Pratt. “Digital Image processing-PIKS inside Third Edition JohnWiley, 2002.

[7] R. C. Gonzalez. “Digital Image processing using Matlab” Pearson publication, 2005.

[8] N Karssemeijer’. Automate classification of parenchymal patterns in mammograms. Physics in Medicine and Biology, Vol. 43, No. 2, pp. 365-378, 1998.

[9] Suckling. J et al., ‘The Mammographic Image Analysis Society Digital Mammogram Database’ Exerpta Medica. International Congress Series Vol. 1069, pp. 375-378, 1994.

[10] B. A. Eskin, S. O. Asbell, and Lori Jardines. ‘Breast disease for primary care physicians’ Parthenon publishing 1998.