APPLICATION OF WAVELETS IN DETECTION AND CLASSIFICATION OF MICROCALCIFICATIONS IN DIGITAL MAMMOGRAMS - SOME RECENT ADVANCES

*Soma Biswas** *Kaustav Nandy*†

Abstract **- Microcalcification clusters in mammograms are an important early sign of breast cancer. Varying densities of parenchymal tissue make visual detection of these masses very difficult. Computer-aided diagnosis using wavelet transform is recently being used to improve the diagnostic accuracy and efficiency of screening mammography. Wavelet coeffcients describe the local geometry of an image in terms of scale and orientation apart from being flexible and robust with respect to image resolution and quality. This paper surveys some recent developments in the application of wavelets in enhancement, detection and classification of microcalcifications in mammograms. Possible areas of future research have also been discussed.**

I.Introduction

Today, breast cancer is the most frequent form of cancer in women. Early detection and treatment of breast cancer increases the chances of survival to a considerable extent. Of all the diagnostic methods currently available for detection of breast cancer, mammography is the only reliable and practical method capable of detecting breast cancer in its early stage. Among the early indicators of breast cancers, non palpable masses and microcalcifications are the primary signs. In general, mammographic image analysis can be divided into three steps, *enhancement of mammographic features*, *detection and localization of suspicious areas* and *classification.* In order to improve the diagnostic accuracy and efficiency of screening mammography, computer- aided diagnosis (Fig. 1) has been introduced into the screening process. Various wavelet based methods have been employed for detection and classification of clustered microcalcifications with the unified aim of developing a CAD system[1].

Wavelet analysis is appropriate for detection and classification of microcalcifications as it decomposes the image into well localized, interpretable components that make local features in the image easily accessible. It decomposes an image into coefficients that describe the local geometry of the image in terms of scale and orientation. It has the additional advantage of being flexible with respect to image resolution and robust with respect to varying image quality. The strong mathematical basis and the different properties of the wavelet transform have been utilized for the purpose of enhancement, detection and classification of microcalcifications.

The paper is arranged as follows. In Section II, the different wavelet-based detection and enhancement methods and in Section III, the different wavelet-based microcalcification classification methods are discussed. In Section IV, conclusion is given along with directions for futher research.

II. Detection and Enhancement methods

Microcalcifications appear in digital mammograms as groups of small localized granular bright spots due to their higher X-ray attenuation compared to the normal breast tissue. In some methods, these have been treated as singularities in the digitized image and singularity detection procedures have been used for detecting microcalcifications. Wavelets demonstrate superior performance in the detection of singularities compared

*M.Tech student, Department of Electrical Engineering, Indian Institute of Technology, Kanpur e-mail: soma@iitk.ac.in

†M.Tech student, Department of Electrical Engineering, Indian Institute of Technology, Kanpur e-mail: kaustav@iitk.ac.in

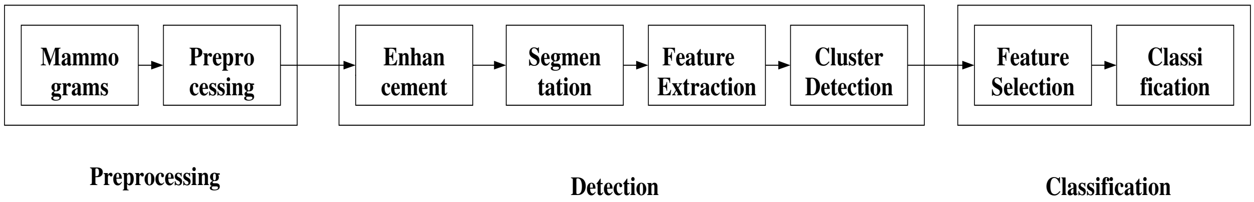

Figure 1: Block Diagram of typical CAD system

to several other mathematical tools. In [2] microcalcifications are detected using new wavelets issued from Matzinger polynomial. These wavelets have a higher Sobolev regularity index compared to the classical wavelets (Daubechies). It is well known that wavelets with high Sobolev regularity index are better able to detect singularities. It has been shown experimentally that these new wavelets outperform the classical wavelets for the detection purpose specially with smaller support width. This multiresolution analysis with the new wavelets provide very small detail coefficient when analyzing normal areas and give detail coefficients with high modulus for the abnormal areas. Use of high Sobolev regularity reduces the number of false alarms that usually cause a lot of wastage of time.

In some methods, wavelets have been combined with different artificial intelligence techniques in order to develop a CAD system. By incorporating the expert knowledge of radiologists, the computer-based systems provide a second opinion in detecting abnormalities and making diagnostic decisions. Detection suspicious areas on the basis of certain features extracted from the images at different resolutions is also frequently used. [3] develops a algorithm in which all the suspicious regions are singled out for further analysis in the preprocessing step . It has been observed from many mammograms that the image roughness of the region containing masses are different from that of normal tissue. So the mammograms are subdivided into several blocks and the fractal dimension (which indicates the roughness of a region) of each block is calculated. The blocks that have a very smooth or very rough surface are discarded. In the second step a DWT based multiresolution Markov Random Field (MMRF) segmentation algorithm is used. By taking only the LL subband images (containing only the low frequency information of the image at that resolution) of DWT for MRF segmentation, the algorithm effectively merges the unnecessary small regions and produce only the sizable regions for subsequent processing. A robust dogs and rabbits clustering is used to initialize the segmentation algorithm in the coarsest level. Segmentation results are propagated according to the self similarity mapping relations between the different levels until the finest level is reached. This is very efficient in removing image noise introduced by veins and fibres to generate clean images for subsequent analysis. Finally features like *area*, *compactness*, *mean gradient within current region*, *mean gradient of region boundary*, *gray level variance*, *edge intensity variance* and *mean intensity difference* are generated from the segmentation step and a binary decision algorithm is used to pin point suspicious areas. Using this method both well defined and ill defined masses are detectable(Fig. 2). The procedure shows a very high detection accuracy with low false alarm rate.

[4] presents a CAD system in which mixed features consisting of wavelet features and grey level statistical features are used for segmenting potential microcalcification pixels in the mammograms. To generate wavelet features each mammogram image is decomposed upto four levels using the separable 2-D wavelet transform (Daubechies orthogonal wavelet). Reconstruction of images is performed in each level by setting the transform values of the other levels to zero of which only those corresponding to levels two and three are retained as they contain the most meaningful information. The two gray level statistical features that have been used are *median contrast* and *normalized gray level value*. These are fed to a multilayer neural network which classifies each pixel in the original mammogram into a microcalcification pixel or a normal pixel. The classified pixels are then grouped into potential microcalcification objects by their spatial connectivity. To eliminate these false detections due to noise a second detection step is used which uses Karssemeijer's criteria. Here these potential individual microcalcification objects are classified as true or false based on a set of thirty one features extracted from these potential individual microcalcification

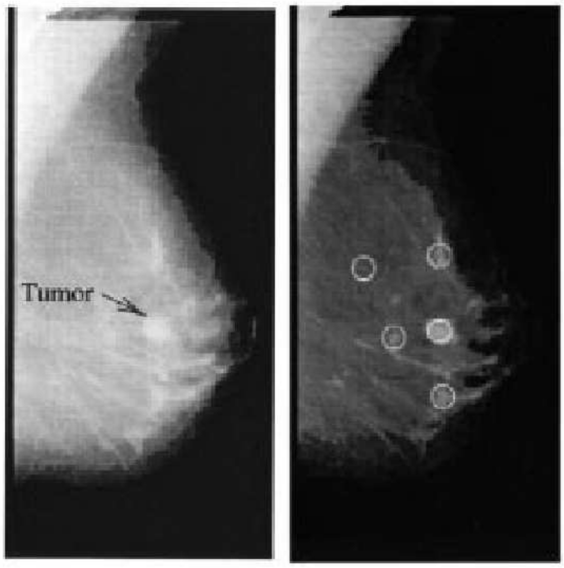

Figure 2: The detected masses in a mammogram

objects.

Detection of spiculated lesions using feature extraction has also been adopted in [5] where a linear phase nonseparable 2-D wavelet transform has been used. At each resolution, for every pixel, four features are extracted from the low-frequency wavelet coefficients, namely *mean pixel brightness*, *standard deviation of pixel brightnesses*, *standard deviation of gradient orientation histogram* and *standard deviation of the folded gradient orientations.* Pixels in normal areas have similar gradient orientations, while pixels near spiculated lesions tend to have gradient orientations in many directions resulting in larger standard deviation of gradient orientations. This also results in a flatter gradient orientation histogram near a lesion pixel compared to that near a normal pixel. Lesions usually have higher density, or appear brighter in mammograms and so these features have also been incorporated. Detection is performed from the coarsest resolution to the finest resolution using a binary tree classifier. The method overcomes the difficulty of choosing a neighborhood size apriori to capture tumors of varying sizes and also has the advantage of reduced computational requirement.

Since radiologists spend a lot of time investigating images lacking any abnormality, much research has also been carried out for detection of clinically normal mammograms. [6] presents one such method for detecting clinically normal tissue in digitized mammograms using a multiresolution statistical model. Here the word "normal" has been used to identify tissue regions that do not contain calcifications, calcified veins, or image aberations. If a small region deviates significantly from the global model, it is marked as potentially suspicious, and, if a region is in agreement, it can be discarded. Here symmlet wavelet with 12 coefficients have been used as it roughly resembles the intensity profile of the calcifications. Two important points that can be noted regarding this method are that the statistical analysis has been applied to independent subspace images and that the focus is on identification of statistical properties of normal tissue at multiple resolution. In essence this technique merges two powerful analysis techniques, namely classical signal detection theory and multiresoluton decomposition.

Visibility of microcalcifications is often degraded by the high frequency texture of the breast tissue. Much research has been carried out for increasing the visibility of the microcalfications which makes the detection easier. In [7] [8], the background mammogram texture have been modelled using a separable and a non separable Markov process and a optimum matched filter is developed for detection of objects in presence of such background noise. Here these matched filters are implemented using biorthogonal wavelet transform with B-spline wavelet bases for the optimum detection of microcalcifications. Here undecimated wavelet transform is used (conventional filter bank implementation without downsampling) as the features are displayed at full size which makes the features easily visible and also because it is easier to combine the detected pixels from each sub-band when they are at the same resolution. Output of each matched filter of a specific frequency band (here HH HL+LH) is first thresholded to produce a 1(detect) or 0(no detect) result. Binary output from all channels are then combined which is then thresholded to give the final detection decision. In the second stage an accurate segmentation of the boundaries have been performed. The detected pixel sites in the HH and LH+HL are dilated then weighed before computing the inverse wavelet transform. By this method the individual microcalcifications are greatly enhanced in the output image to the extent that straightforward thresholding can be applied to segment them.

Since microcalcifications appear in the high frequency components of the wavelet decomposition,

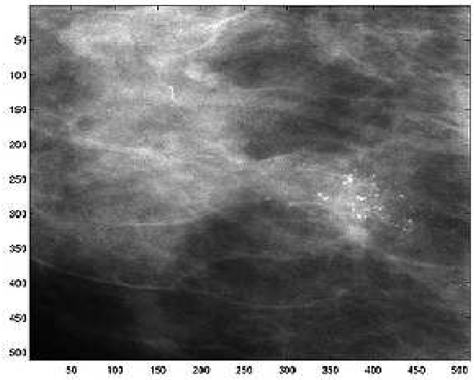

Figure 3: Typical mammogram

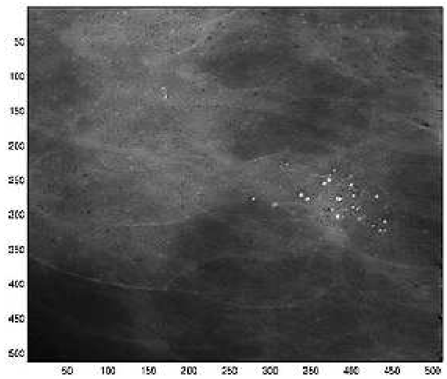

Figure 4: Enhanced mammogram

the background information can be removed by setting the low frequency LL portion of the decomposition to zero. On reconstruction and histogram equalization the visibility of the microcalcifications is greatly increased compared to the background. This approach has been adopted in [9].

[10] presents a undecimated discrete wavelet packet transform based scheme for segmentation and extraction of tumors. This transform helps to generate an image space in which a specific object is easy to be discriminated. After detection of suspicious abnormal areas a multiscale region based segmentation method is used to obtain the exact boundaries of the detected regions.

Due to the fuzzy nature, low contrast and low distinguishability of the microcalcification from their surroundings the detection step can be preceeded by enhancement of the mammogram. So much research has been conducted for enhancement of mammograms. [11] presents a new algorithm for enhancement of microcalcifications in mammograms. A novel method has been developed to discretize the 2-D Continuous Wavelet Transform which allows the wavelet filterbank to adapt optimally to the size and shape of microcalcifications. This discretization permits arbitrary scales and orientation while maintaining the reconstruction property of the continuous wavelet transform (Fig. 3 and 4). [12] accomplishes mammographic feature enhancement by over complete multiresolution representations. The orientation and the frequency selectivity of the wavelet transform have been exploited to make mammographic features more obvious through localized contrast gain. For analysis three multiresolution representations have been used, the Dyadic Wavelet Transform, ϕ - transform (Frazier - Jawerth Transform) and the Hexagonal Wavelet Tranform. The transform coefficients within wavelet frames are modified for contrast enhancement. Both local and global enhancement techniques have been used here for the purpose. Several other wavelet based methods have been used for enhancement of mammograms [13][14].

III. Classification methods

After detection of the microcalcifications the next step is to classify them. The primary features that indicate malignancy of the tumor are related to the tumor's density, size, shape and borders. The classification is usually done based on their shape, size or texture. The malignant mass often infiltrates into the surrounding tissue structure and hence the boundary may comprise of fine linear strands extending irregularly outwards from the central mass. Hence the shape of the tumor can tell us whether the mass is malignant or benign. Till date much research has been focused to develop an automated classification system.

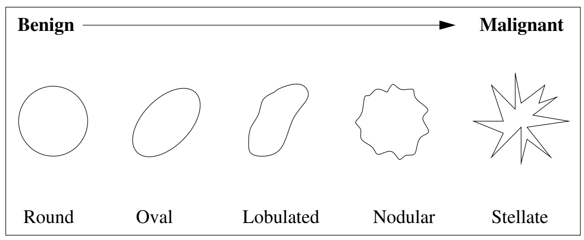

Figure 5: Classification of mammographic masses

Many researchers have applied the method of shape analysis for classification of microcalcifications. For example various shape descriptors have been used for this purpose. In [15] multiresolution techniques (specifically the *Discrete Wavelet Transform* modulus maxima (mod-max) method) have been applied to the problem of shape classification. Here seven uniresolution features like *tumor circularity*, *radial distance mean* and *boundary roughness index* and and three new multiresolution shape features have been used to classify masses as round, nodular or stellate (Fig. 5) because they frequently correspond to the physical conditions of cyst, fibroadenomas and carcinomas. The radial distance measures for rounded masses posses relatively smooth low magnitude variations. These variations progressively increase for the nodular and the stellate masses. The multiresolution features numerically characterize the signal variations present in the radial distance measures. Then the DWT mod-max representation is utilized to detect the sharp signal variations and extract the numerical descriptors. If the local maxima can be tracked from scale 2^J to scale 2^1 and converge to the abscissa x_o, the magnitude(k), variance(σ) and Lipschitz order(α) associated with the sharp variation located at x_o are calculated. The numerical descriptors used are

1. *Variation ratio mean*

$$(k/\sigma)_{avg} = \frac{1}{M}\sum_{i=1}^{M}\frac{k(i)}{\sigma(i)}$$

where M is the number of dominant sharp signal variations.

2. *Variation ratio standard deviation*

$$(k/\sigma)_{std} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{k(i)}{\sigma(i)} - (k/\sigma)_{avg}\right)^2}$$

3. *Lipschitz sum*

$$\alpha_{sum} = \sum_{i=1}^{M}\alpha(i)$$

Feature analysis is carried out via linear discriminant analysis (LDA). The classification algorithm is based on distance (Euclidean) measures. The distance of each feature vector was measured relative to the mean feature vector of each class. The feature vector is classified as a member of the class for which its distance measure is minimized. By using the multiresolution features, the overall classification rates were increased from 72% to 83%.

In cases where mammographic images suffer from poorly defined microcalcification features, classification based on extraction of shape features may not be accurate. So some researchers have tried to investigate other properties like intensity variations and texture information in the area of interest as a means for classification of microcalcifications. [16] uses wavelet decomposition to represent the local texture of the gray level image of the digitized microcalcification area along with second order gray-level histogram statistics to represent global texture and density variations for classification of microcalcifications.

Another approach for microcalcification classification has been adopted in [17]. Here a supervised classifier based on wavelet transform decpmposition has been proposed. The wavelets that have been used are the Haar and Daubechies 4 (Db4). Decomposition of the mammograms are done only upto the first level and among the 4 sets of coefficients only the low frequency ones are retained since they have the capability for sufficient and satisfactory representation of texture. After this only some of the biggest coefficients among these low pass ones are retained and are used as signature vector for the calcification of that class. A prototype for each class is derived in this way. The test mammogram is similarly decomposed and the same number of coefficients are retained. For determination of the classification a nearest neighbour classifier is used based on Euclidean distance. This way the mammogram belongs to the class for which the distance is minimum.

IV. Conclusion and Future Work

Automated breast cancer detection and classification has been studied for a long time and the CAD mammography systems have gone from crude tools in the research laboratory to commercial systems. It is a well approved fact that these CAD systems definitely improve the performance of the radiologists when they are provided the output of the CAD system. This article clearly shows that multiresolution decomposition using wavelet analysis provides an efficient method for mammographic mass detection, enhancement and analysis. Having good localization in both time and frequency, wavelets offer a better means of representing functions with near discontinuities. Moreover wavelets offer a natural way to reduce noise that is scale or space dependent.

Although by now much progress has been achieved, there are still remaining challenges and directions for future research. These include development of better enhancement and segmentation algorithms and designing better feature detection and selection algorithm. One of the most important aspect regarding the wavelet based systems is the proper choice of wavelets depending on the nature of application. This is an area of improvement that requires more attention. Moreover identification of the cofficients that best represent the mammographic masses should be done. This enables the systems to be more efficient and accurate. In order to reduce both the number of false positives and false negetives several different types of features, sometimes, with clinical information should be used. However one drawback of mammograms is that they are 2-dimension representation of 3-dimension data. The spatial distribution of the microcalcification clusters is inevitably distorted in the projected 2-dimension mammograms and will lose the depth and location of the imaged structures. Recently 3-dimension visualization and analysis of mammograms have arroused great interest and to develop new and better image processing algorithms for 3-dimension is still a big challenge.

Reference

[1] W. B. Richardson Jr., Applying Wavelets to Mammograms, *IEEE Engineering in Medicine and Biology*, pp. 551-560, Sep/Oct 1995.

[2] G. Lemaur, K.Drouiche and J.DeConinck, Highly Regular Wavelets for the Detection of Clustered Microcalcifications in Mammograms, *IEEE Trans. Med. Imag.*, vol. 22, pp. 393-401, March 2003.

[3] Lei Zheng and Andrew K. Chan, An Artificial Intelligent Algorithm for Tumor Detection in Screening Mammogram, *IEEE Trans. Med. Imag.*, vol. 20, pp. 559-567, July 2001.

[4] Songyang Yu and Ling Guan, A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films, *IEEE Trans. Med. Imag.*, vol. 19, pp. 115-126, Feb. 2000.

[5] S. Liu, C. F. Babbs and E. J. Delp, Multiresolution Detection of Spiculated Lesions in Digital Mammograms, *IEEE Trans. Image Proc.*, vol. 10, pp. 874-884, June 2001.

[6] John J. Heine, Stanley R. Deans, D. Kent Cullers, Richard Stauduhar and Laurence P. Clarke, Multiresolution Statistical Analysis of High-Resolution Digital Mammograms, *IEEE Trans. Med. Imag.*, vol. 16, pp. 503-515, Oct. 1997.

[7] Robin N. Strickland and Hee Hahn, Wavelet Transforms for Detecting Microcalcifications in Mammograms, *IEEE Trans. Med. Imag.*, vol. 15, pp. 218-229, April 1996.

[8] R. N. Strickland and H. Hahn, Wavelet Transform Methods for Object Detection and Recovery, *IEEE Trans. Image Proc.*, vol. 6, pp. 724-735, May 1997.

[9] T.C. Wang and N.B. Karayiannis, Detection of Microcalcifications in Digital Mammograms Using Wavelets, *IEEE Trans. Med. Imag.*, vol. 17, pp. 498-509, August 1998.

[10] X. Zhang, Multiscale Tumor Detection And Segmentation in Mammograms, *Proceedings IEEE International Symposium on Biomedical Imaging*, 2002, pp. 213-216, July 2002.

[11] P. Heinlein, Johann Drexl and Wilfried Schneider, Integrated Wavelets for Enhancement of Microcalcifications in Digital Mammography, *IEEE Trans. Med. Imag.*, vol. 22, pp. 402-413, March 2003.

[12] Andrew F. Laine, S. Schuler, J. Fan and W. Huda, Mammographic Feature Enhancement by Multiscale Analysis, *IEEE Trans. Med. Imag.*, vol. 13, pp. 725-740, December 1994.

[13] W. Qian, L. P. Clarke, B. Zheng, M. Kallergi and R. Clark, Computer Assisted Diagnosis for Digital Mammography, *IEEE Engineering in Medicine and Biology*, pp. 561-569, Sep/Oct 1995.

[14] A. Laine, J. Fan and W. Yang, Wavelets for Contrast Enhancement of Digital Mammography, *IEEE Engineering in Medicine and Biology*, pp. 536-550, Sep/Oct 1995.

[15] Lori Mann Bruce and Reza R. Adhami, Classifying Mammographic Mass Shapes Using the Wavelet Transform Modulus-Maxima Method, *IEEE Trans. Med. Imag.*, vol. 12, pp. 1170-1177, December 1999.

[16] A. P. Dhawan, Y. Chitre, C. Kaiser-Bonasso and M. Moskowitz, Analysis of Mammographic Microcalcifications Using Gray-Level Image Structure Features, *IEEE Trans. Med. Imag.*, vol. 15, pp. 246-259, June 1996.

[17] C. B. R. Ferreira and D. L. Borges, Analysis of mammogram classification using a wavelet transform decomposition, *Pattern Recognition Letters 24*, pp. 973-982, 2003.