

# Automatic Seed Point Selection Algorithm to Detect Breast Lesions in Ultrasound Image

Telagarapu Prabhakar

Department Electronics & Communication Engineering,  
GMR Institute of Technology, Rajam, Andhra Pradesh, India.

**Abstract:** *Ultrasound* has been one of the most powerful techniques for imaging organs and soft tissue structure in the human body. It has been used for breast cancer detection because of its non-invasive, sensitive to dense breast, low positive rate and cheap cost. But due to the nature of ultrasound image, the image suffers from poor quality caused by speckle noise, low contrast and blurred edges. These make the automatic segmentation and classification of interested lesion difficult. One of the frequently used segmentation method is region growing method. A seed point is the starting point for region growing and its selection is very important for the segmentation and classification result. If this seed point is selected outside the region of interest, final segmentation and classification would be definitely incorrect. So, mostly the seed point is manually selected in order to avoid incorrect segmentation and classification. And also, most of the past works have concentrated only on automatic seed point location and segmentation of single lesion in U.S breast images. Proposing a method which was able to locate the seed point automatically within the multiple lesion regions by using the tan function, textural features of the lesion regions and Support Vector Machine and also able to automatically segment the lesion regions properly. Using this proposed method, able to located seed point within multiple lesion regions automatically. The results show that the proposed method was able to automatically locate the seed points within lesion region automatically for 98.42 % of images in our data set.

**IndexTerms** - Seed point, Breast lesion, Automatic segmentation, Support Vector Machine.

## I. INTRODUCTION

Breast cancer is the most common cancer amongst women. Breast cancer can be diagnosed using invasive and non-invasive method. To avoid invasive method in diagnosis of cancer, radiologists go for mammogram technique along with ultrasound (U.S) imaging method which is considered as a golden standard [1]. When compared to other diagnosing method, U.S method is used mostly because of its non-invasive, non-ionizing radiation, cost effective and painless procedure [2]. U.S images are of poor quality due to inherent speckle noise, shadow effect and low contrast. These make the segmentation of the lesion region very difficult [3, 4]. Breast lesion region in ultrasound images are classified depending upon the contour, shape, size and textural features of the segmented region. Seed point is the initial steps in segmentation of lesion regions in the image. If the seed point is located outside the lesion region, it leads to wrong segmentation which results in misclassification of the lesion regions. To avoid this, most of the time the seed point is located manually [5]. In order to prevent this manual intervention, the seed points must be found automatically, so that the complete process of segmentation will be automatic. In [5] seed point is automatically selected using the textural features and the spatial characteristic of the lesion region. In [6] seed point is automatically located inside the lesion region based on the co-occurrence feature and run length feature. The run length features are calculated around the points selected by the co-occurrence features. Depending upon the run length features of the selected point and its neighborhood points, the seed point is selected. Based on the empirical domain knowledge of the breast structure and clustering method, the lesion region were able to be segmented automatically [7]. In this paper the proposed algorithm able to locate the seed point automatically within the lesion region by using the tan function, textural features and neural network. Using this method, able to locate the seed point within multiple lesion regions.

## II. METHODOLOGY

### a) Preprocessing

U.S images are mostly degraded by the inherently existing speckle noise. Speckle noise is a multiplicative noise, which is directly proportional to the pixel intensity value. The degraded image affected by the speckle noise can be expressed as

$$G(x,y)=I(x,y)*N(x,y) \quad (1)$$

Where  $G(x,y)$  is the degraded image,  $I(x,y)$  is the uncorrupted image and  $N(x,y)$  is the noise multiplied with the image. Here used various filters like Wavelet Thresholding, Wiener, SRAD, Mean and Median filter to reduce the speckle noise and other noises in the image. The peak signal to noise ratio (PSNR) have been found for all these filters and out of all these filters, the best filter is found to be wiener filter. So every input image is pre-processed using wiener filter.

### b) Processing with tan function

The filtered image is processed by a tan function. Tan function is applied to every pixel of the image and a new image is obtained. Since most of the lesion region lies below half the maximum value, so apply the tan function only the lower half of the pixels value is stretched. If the input image is  $I(x, y)$ , then the tan processed image is given by the expression

$$TN(x, y) = \tan(I(x, y)*(pi/2)) \quad (2)$$

**c) Training phase**

1) Feature extraction: The statistical features are obtained from the breast lesion region as well as from the normal region. The statistical features like mean, standard deviation (SD) and entropy is found for both the normal and lesion regions. Since most of the lesion region have low mean when compared to normal region, by using the mean as one of the features. Similarly, most of the breast lesion regions are more homogenous when compared to normal region, since here extracting entropy and SD.

2) SVM training: All the extracted features from the lesion as well as the normal region are used to train the Support Vector Machine (SVM) with the corresponding output. SVM tends to find a hyperplane that can separate the input samples. In the case that the original data cannot be separated by a hyperplane, SVM will transform the original data into a feature space of higher dimension by using the kernel function. The popular kernels are Linear kernel, Polynomial kernel of degree d, Gaussian radial basis function (RBF), Neural nets (sigmoids). In this work linear kernel is used. Training of the network continues until a desired minimum error is obtained between the targeted output and the obtained output.

**d) Testing phase**

In this step, features like mean, entropy and S.D are obtained from every 18\*18 pixels of the testing image. These features are feed as input to the trained network. If the inputs are similar to lesion features, then a corresponding output of 1 is produced. Similarly, if it is similar to normal region, output of 0 is produced.

**e) Finding the Seed point**

Some of the boundary connected normal regions in U.S images have similar features compared to lesion region. These regions must be deleted, so that the seed point can be found effectively within the lesion region. But some time, the lesion regions are also connected to the boundary of the image. When deleting, the lesion region which is connected to the boundary must not be deleted. Careful observation was made on 40 tan processed images. In all these images, the lesion region are surrounded by high pixel intensity when compared to the normal region. Depending upon this observation, the regions are evaluated. Then, the seed point is found within the lesion region.

**III. RESULTS AND DISCUSSION**

The proposed algorithm was simulated on 90 different Breast Ultrasound images obtained from Precision Diagnostic, MEDALL in Chennai. The lesion regions length and width are manually marked by the radiologist. If the seed point is found within the lesion region, it is considered as true positive (TP) else considered as false positive (FP). Qualitative analysis also has been done based on the location of the seed point near to the centre of the lesion region. Different images were used to train the SVM classifier and 90 different images have been tested. The time taken for automatic segmentation is 9 seconds which is found to be time consuming when compared with manual delineation.

**a) Preprocessing**

To evaluate the filters, first a standard image is obtained and speckle noise is added with the image with mean of 0 and with different variance. Speckle noise is a multiplicative noise which is inherently present in the image. After adding the speckle noise, the image is filtered using the various filters as shown in Table 1.

. TABLE 1 COMPARISON OF FILTER TECHNIQUES WITH PSNR VALUES

Filters	Denoised Images PSNR values in dB (Variance=0.02)
Wavelet Thresholding	73.2431
SRAD	70.3285
Wiener	24.4507
Mean	23.6050
Median	22.0156

Finally peak-signal to noise ratio (PSNR) is calculated between the standard image without noise and with various filtered image as shown in the Table 1. Depending upon the PSNR value, the best filter is found to be wiener filter and this filter is used in the program.

**b) Processing with Tan Function**

For effective distinguishing of the lesion region from the normal region, every pixel value of the input image is processed with tan function. The input image and its histogram is shown in the figure 1 and 2, similarly the tan processed image and its histogram is shown in the figure 3 and 4.



Figure 1 Input image

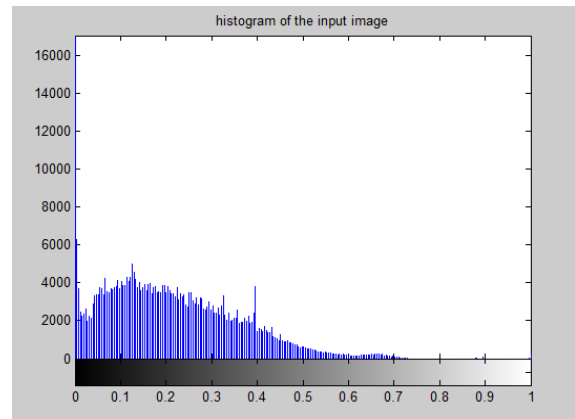


Figure 2 Histogram of the Input image



Figure 3 Tan processed image

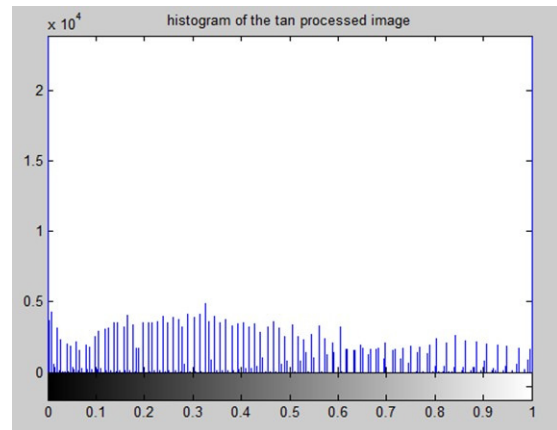


Figure 4 Histogram of the tan processed image

**c) Seed Point Selection**



Figure 5 Input image

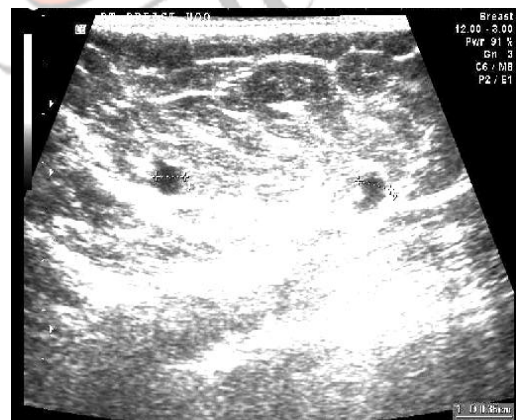


Figure 6 Tan processed image

The features like mean, SD and entropy for lesion region and normal region is extracted for 5 different images. The extracted features are shown for normal region. The extracted features from the training images are feed as input to the SVM classifier. From every 18\*18 pixel of the testing image, the features are found and fed as the input to the newly trained SVM. From the SVM the seed points are found inside the lesion regions. Using the region growing method, the lesion regions are automatically segmented from the seed points. The various stages of our algorithm are shown in the Figure 1 to 6. Quantitative results are shown in the Table 2, where TP represent true positive and FP represent false positive.

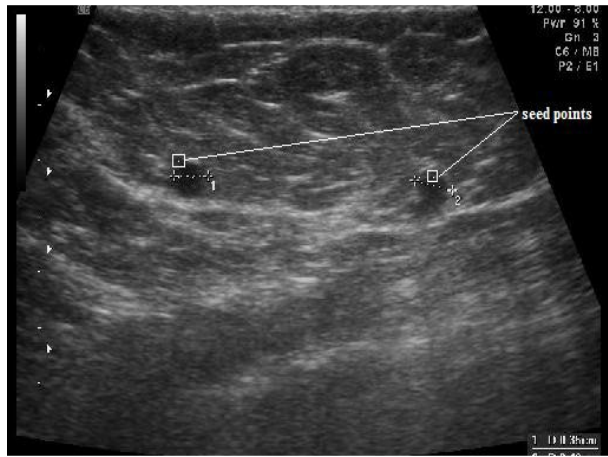


Figure 5 Automatic Seed pointed Image

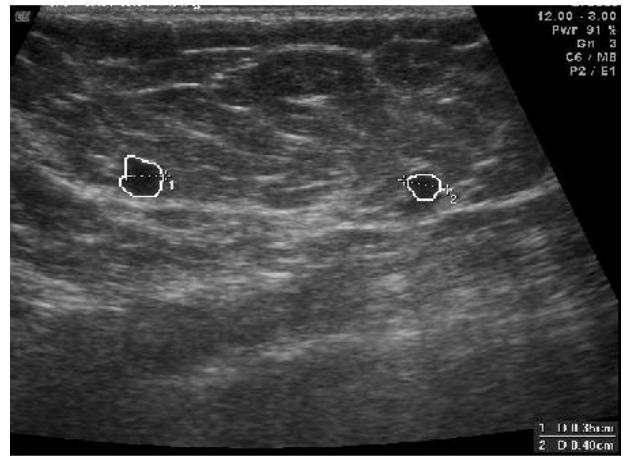


Figure 6 Automatic Segmented Image

Table 2 Performance of SVM Classification

Metric	SVM Classification Accuracy
TRUE POSITIVE (TP)	<b>98.42 %</b>

#### IV. CONCLUSION

The aim of this paper is to find seed points in multiple-lesion regions without the manual intervention which is a difficult task. The proposed method was able to find the seed point automatically within multiple-lesion regions without the intervention of the sonographer with the TP rate of 97.5% for our database. The proposed method is robust and consumes very less time when compared to manual intervention.

#### REFERENCES

- [1] AnantMadabhushi and Dimitris N. Metaxas. Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesion. 2003. IEEE Transactions on Medical Imaging, Vol. 22, No. 2.
- [2] GuitaRahbar, Angela C. Sie, Gail C. Hansen, Jeffrey S. Prince, Michelle L. Melany. Benign versus Malignant Solid Breast Masses: US Differentiation. 1999. Radiology.
- [3].Hong-Ting Chen, Yi-Ping Lien, Po-Ting Liu, Dar-Ren Chen and Yu-Len Huang. Characterization of Benign and Malignant Solid Breast Masses Using Vascular Morphology in 3D Power Doppler Ultrasound Images. 2010. 3rd International Conference on Biomedical Engineering and Informatics.
- [4]Hsieh-Wei Lee, King-Chu Hung, Bin-Da Li, Sheau-Fang Lei, and Hsin-Wen Ting. Realization of high octave decomposition for breast cancer feature extraction on ultrasound images. 2011. IEEE Transactions on circuits and system. Vol 58, NO. 6.
- [5]Juan Shan, H. D., Cheng, Yuxuan Wang. A novel automatic seed point selection algorithm for breast ultrasound images. 2008. IEEE.
- [6]Neb Duric, Cuiping Li, Carri Glide-Hurst, Peter Littrup, Lianjie Huang and Jessica Lupinacc. Breast Imaging with Ultrasound Tomography: Clinical results at the Karmanos Cancer Institute. 2008. International Conference on BioMedical Engineering and Informatics.
- [7]Ruey-Feng Chang, Wen-Jie Wu, Chih-Chi Tseng, Dar-Ren Chen, and Woo Kyung Moon. 3-D snake for US in margin evaluation for malignant breast tumor excision using mammotome. 2003. IEEE transactions on Information Technology in Biomedicine, Vol. 7, No. 3.
- [9]Simona Moldovanu, LuminitaMoraru. Mass Detection and Classification in Breast Ultrasound Image Using K-means Clustering Algorithm. 2010. IEEE. Vol. 25, No. 3.
- [10]Yi Zheng, James F. Greenleaf and John J. Gisvold. Reduction of Breast Biopsies with a modified self-organizing map. 1997. IEEE transactions on neural network. Vol. 8, No. 6.