LIKELIHOOD FUNCTION ANALYSIS FOR SEGMENTATION OF MAMMOGRAPHIC MASSES FOR VARIOUS MARGIN GROUPS

Lisa Kinnard[a,b,c], Shih-Chung B. Lo[a], Erini Makariou[a], Teresa Osicka[a,d], Paul Wang[c], Matthew T. Freedman[a], Mohamed Chouikha[b]

[a]ISIS Center, Dept. of Radiology, Georgetown University Medical Center, Washington, D.C., USA
[b]Department of Electrical and Computer Engineering, Howard University, Washington, D.C., USA
[c]Biomedical NMR Laboratory, Department of Radiology, Howard University, Washington, D.C., USA
[d]Department of Electrical Engineering and Computer Science, The Catholic University of America, Washington DC, USA

ABSTRACT

The purpose of this work was to develop an automatic boundary detection method for mammographic masses and to observe the method's performance on different four of the five margin groups as defined by the ACR, namely, spiculated, ill-defined, circumscribed, and obscured. The segmentation method utilized a maximum likelihood steep change analysis technique that is capable of delineating ill-defined borders of the masses. Previous investigators have shown that the maximum likelihood function can be utilized to determine the border of the mass body. The method was tested on 122 digitized mammograms selected from the University of South Florida's Digital Database for Screening Mammography (DDSM). The segmentation results were validated using overlap and accuracy statistics, where the gold standards were manual traces provided by two expert radiologists. We have concluded that the intensity threshold that produces the best contour corresponds to a particular steep change location within the likelihood function.

1. INTRODUCTION

In a CAD_x system, segmentation is arguably one of the most important aspects – particularly for masses – because strong diagnostic predictors for masses are shape and margin type [2,9]. The margin of a mass is defined as the interface between the mass and surrounding tissue [2]. Furthermore, breast masses can have unclear borders and are sometimes obscured by glandular tissue in mammograms. A spiculated mass consists of a central mass body surrounded by fibrous projections, hence the resulting stellate shape. For the aforementioned reasons, proper segmentation - to include the body and periphery - is extremely important and is essential for the computer to analyze, and in turn, determine the malignancy of the mass in mammographic CAD_x systems.

Over the years researchers have used many methods to segment masses in mammograms. Petrick [7] et al. developed the Density Weighted Contrast Enhancement (DWCE) method, in which series of filters are applied to the image in an attempt to extract masses. Comer et al. [1] segmented digitized mammograms into homogeneous texture regions by assigning each pixel to one of a set of classes such that the number incorrectly classified pixels was minimized via Maximum Likelihood (ML) analysis. Li [5] developed a method that employs k-means classification to classify pixels as belonging to the region of interest (ROI) or background.

Kupinski and Giger developed a method [4], which uses ML analysis to determine final segmentation. In their method, the likelihood function is formed from likelihood values determined by a set of image contours produced by the region growing method. This method is a highly effective one that was also implemented by Te Brake and Karssemeijer in their comparison between the discrete dynamic contour model and the likelihood method [9]. For this reason we chose to investigate its use as a possible starting point from which a second method could be developed. Consequently in our implementation of this work we discovered an important result, i.e., the maximum likelihood steep change. It appears that in many cases this method produces contour choices that encapsulate important borders such as mass spiculations and ill-defined borders.

2. METHODS

2.1 Initial Contours

As an initial segmentation step, we followed the overall region similarity concept to aggregate the area of interest [1, 4]. Used alone, a sequence of contours representing the mass is generated; however, the computer is not able to choose the contour that is most closely correlated with the experts' delineations. Furthermore, we have devised an ML function steep change analysis method that chooses the best contour that delineates the mass body as well as its extended borders, i.e., extensions into spiculations and areas in which the borders are ill-defined or obscured. This method is an extension of the method developed by Kupinski and Giger [4] that uses ML function analysis to select the contour which best represents the mass, as compared to expert radiologist traces. We have determined that this technique can select the contour that accurately represents the mass body contour for a given set of parameters; however, further analysis of the likelihood function revealed that the computer could

choose a set of three segmentation contour choices from the entire set of contour choices, and then make a final decision from these three choices.

The algorithm can be summarized in several steps. Initially, we use an intensity based thresholding scheme to generate a sequence of grown contours (S_i), where gray value is the similarity criterion. The image is also multiplied by a 2D trapezoidal membership function (2D shadow), whose upper base measures 40 pixels and lower base measures 250 pixels (1 pixel = 50 microns). The image to which the shadow has been applied is henceforth referred to as the "fuzzy" image. The original image and its fuzzy version were used to compute the likelihood of the mass's boundaries. The computation method is comprised of two components for a given boundary: (1) formulation of the composite probability and (2) evaluation of likelihood.

In addition, we chose to aggregate contours using the original image. This accounts for the major difference from that implemented by the previous investigators. Since smoother contours were not used, the likelihood function showed greater variations. In many situations, the greatest variations occurred when there was a sudden increase of the likelihood, and this was strongly correlated with the end of the mass border growth. This phenomenon would be suppressed if the fuzzy image was used to generate the contours. The fuzzy image was used mainly to construct the likelihood function.

2.2 Composite Probability Formation

For a contour (S_i), the composite probability (C_i) is calculated:

$$C_i|S_i = p\left(f_i(x,y)\middle|S_i\right)\times p\left(m_i(x,y)\middle|S_i\right) \qquad (1)$$

The quantity $f_i(x,y)$ is the area to which the 2D shadow has been multiplied, $p(f_i(x,y)|S_i)$ is the probability density function of the pixels inside S_i where 'i' is the region growing step associated with a given intensity threshold. The quantity $m_i(x,y)$ is the area outside S_i (non-fuzzy), and $p(m_i(x,y)|S_i)$ is the probability density function of the pixels outside S_i. Next we find the logarithm of the composite probability of the two regions, C_i:

$$Log\left(C_i|S_i\right) = \log\left(p\left(f_i(x,y)\middle|S_i\right)\right) + \log\left(p\left(m_i(x,y)\middle|S_i\right)\right) \qquad (2)$$

2.3 Evaluation of Likelihood Function

The likelihood that the contour represents the fibrous portion of the mass, i.e., mass body is determined by assessing the maximum likelihood function:

$$\arg\max\left(Log\left(C_i|S_i\right)\right), S_i, i = 1,..n \qquad (3)$$

Equation (3) intends to find the maximum value of the aforementioned likelihood values as a function of intensity threshold. It has been assessed (also by other investigators [4]) that the intensity value corresponding to this maximum likelihood value is the optimal intensity needed to delineate the mass body contour. However, in our implementation it was discovered that the intensity threshold corresponding to the maximum likelihood value confines the contour to the mass body. In our study many of these contours did not include the extended borders. We, therefore, hypothesize that the contour represents the mass's extended borders may well be determined by assessing the maximum changes of the likelihood function, i.e., locate the steepest likelihood value changes within the function:

$$\frac{d}{di}\left(Log\left(C_i|S_i\right)\right), S_i, i = 1,\ldots,n \qquad (4)$$

Based on this assumption, we have carefully analyzed the behavior of maximum likelihood function. The analysis reveals that we have successfully discovered that the most accurate mass delineation is usually obtained by using the intensity value corresponding to the first or second steep change locations within the likelihood function immediately following the maximum likelihood value on the likelihood function.

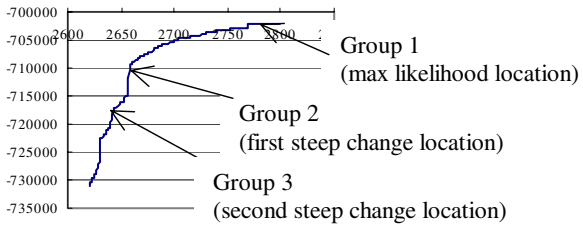

Figure 1: A likelihood function with steep change indicators

2.4 Steep change definition

The term "steep change" is rather subjective and can defined as a location between two or more points in the function where the likelihood values experience a significant change. In some cases the likelihood function increases at a slow rate. The algorithm design accounts for this issue by calculating the difference between likelihood values in steps over several values and comparing the results to two thresholds. The difference equation is given by:

$$h(t) = f(z - wt) - f(z - w(t+1)), \quad t = 0,\ldots,N \qquad (5)$$

where f is the likelihood function, z is the maximum intensity, w is the width of the interval over which the likelihood differences are calculated (e.g. – for w=7 differences are calculated every 7 points), and N is the total number of points in the searchable area divided by w. If the calculation in question yields a value greater than or equal to a given threshold, then the intensity corresponding to this location is considered to be a steep change location. The threshold algorithm occurs as follows:

If $(h(t)_{ML} \geq ML_{T1})$; t=0,…,m
Then choice 1 = intensity where that condition is satisfied
If $(h(t)_{ML} \geq ML_{T2})$; t=m,…,z
Then choice 2 = intensity where that condition is satisfied

where $h(t)_{ML}$ is the steep change value given by equation (5), ML_{T1} and ML_{T2} are pre-defined threshold values, m is the location in the function where the choice 1 condition is satisfied, and z is the location in the function where the choice 2 condition is satisfied. Once the condition is satisfied for the first threshold value (ML_{T1}) then its corresponding intensity value is used to produce the segmentation contour for the first steep change location. Once the condition is satisfied for ML_{T2} then its corresponding intensity value is used to produce the segmentation contour for the second steep change location.

2.5 Validation

The segmentation method was validated on the basis of overlap and accuracy [8,10]:

$$Overlap = \frac{N_{TP}}{N_{FN} + N_{TP} + N_{FP}} \qquad (6)$$

$$Accuracy = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (7)$$

where N_{TP} is the true positive fraction, N_{TN} true negative fraction, N_{FP} is the false positive fraction, and N_{FN} is the false negative fraction. The gold standards used for the validation study were mass contours, which have been traced by expert radiologists.

Our experiments produced contours for the intensity values resulting from three locations within the likelihood functions: (1) The intensity for which a value within the likelihood function is maximum (group 1 contour) (2) The intensity for which the likelihood function experiences its first steep change (group 2 contour) and (3) The intensity for which the likelihood function experiences its second steep change (group 3 contour). We have observed that the intensity for which the likelihood function experiences its first steep change produces the contour trace that is most highly correlated with the gold standard traces, regarding overlap and accuracy.

3. EXPERIMENTS AND RESULTS

Here we describe the database used, describe the experiments, provide visual results obtained by the algorithm, as well as report the results obtained by the ANOVA test.

3.1 Database

For this study, a total of 122 masses were chosen from the University of South Florida's Digital Database for Screening Mammography (DDSM) [3]. The films were digitized at resolutions of 43.5 or 50 μm's using either the Howtek or Lumisys digitizers, respectively. The DDSM cases have been ranked by expert radiologists on a scale from 1 to 5, where 1 represents the most subtle masses and 5 represents the most obvious masses. The images were of varying subtlety ratings. The first set of expert traces was provided by an attending physician of the GUMC, and is hereafter referred to as the Expert A traces. The second set of expert traces was provided by the DDSM, and is hereafter referred to as the Expert B traces.

3.2 Experiments and Results

As mentioned previously, the term "steep change" is very subjective and therefore a set of thresholds needed to be set in an effort to define a particular location within the likelihood function as a "steep change location". For this study the following thresholds were experimentally chosen: ML_{T1}=1800, ML_{T2}=1300, where ML_{T1}= threshold for steep change location 1 for the likelihood function, and ML_{T2} = threshold for steep change location 2 for the likelihood function. We performed a number of experiments in an effort to prove that the intensity for which the likelihood function experiences the first steep change location produces the contour trace, which is most highly correlated with the gold standard traces regarding overlap and accuracy.

First we present segmentation results for two malignant cases followed segmentation results for two benign cases. Each figure contains an original image, traces for Experts A and B, and computer segmentation results for groups 1, 2, and 3. Second, we present data that plots the mean values for various margin groups for both overlap and accuracy measurements. The plots present data for the spiculated and ill-defined groups of malignant masses, and ill-defined and circumscribed groups of benign masses. Data was not presented for the other categories because there was not a sufficient amount of cases.

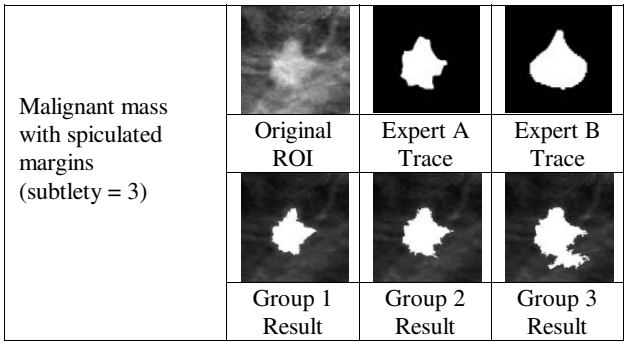

Figure 2: Segmentation Results: Spiculated Malignant Mass

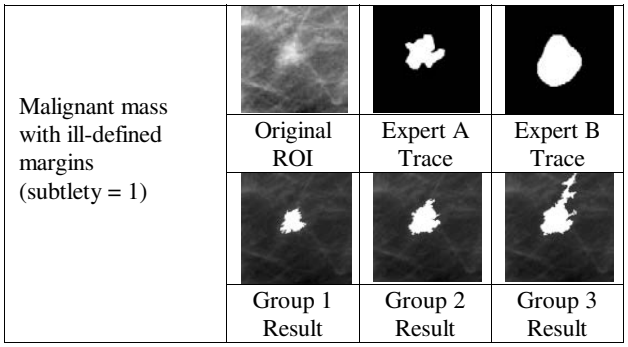

Figure 3: Segmentation Results: Ill-defined Malignant Mass

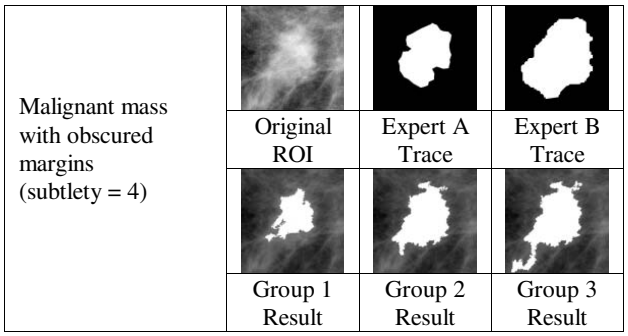

Figure 4: Segmentation Results: Obscured Malignant Mass

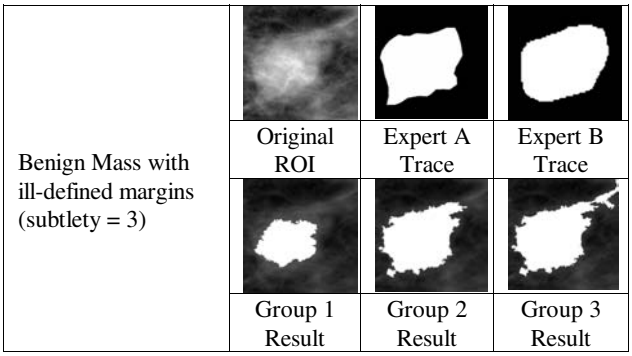

Figure 5: Segmentation Results: Ill-defined Benign Mass

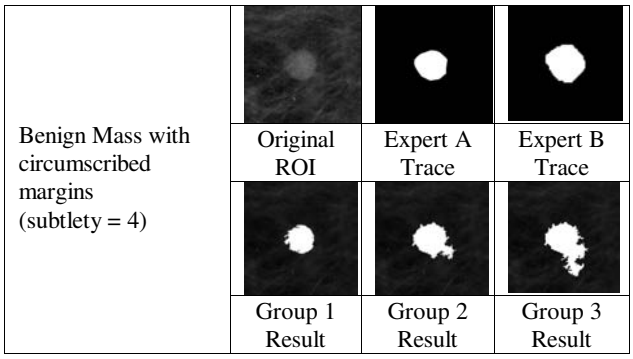

Figure 6: Segmentation Results: Circumscribed Benign Mass

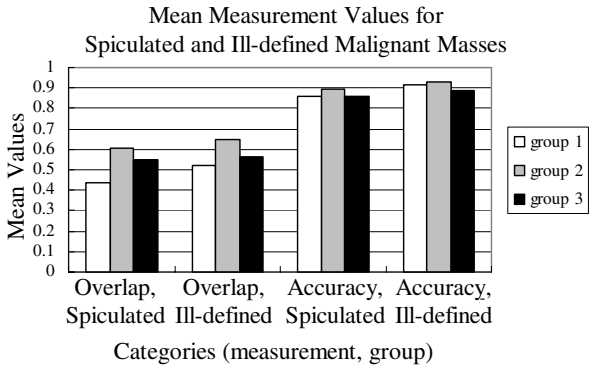

Figure 7: Mean Measurement Values (Malignant Masses)

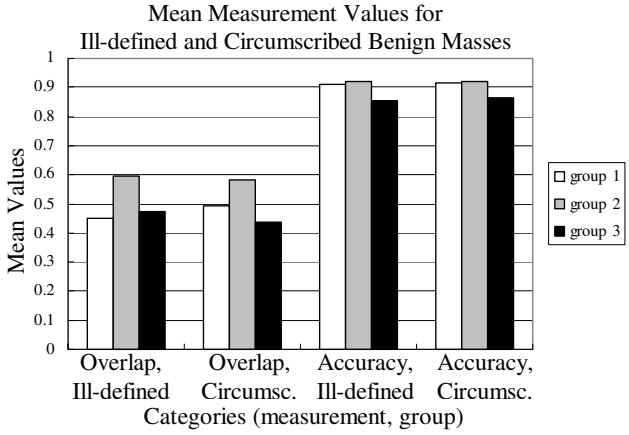

Figure 8: Mean Measurement Values (Benign Masses)

4. DISCUSSION AND CONCLUSION

The visual results (see Figures 2-6) reveal that the group 2 trace appears to delineate the masses better than the group 1 and group 3 contours in most cases. Visually, it appears that the method has performed equally well on all margin groups. This is an encouraging result because some of the more difficult masses to segment are typically those that are spiculated, obscured, and those that have ill-defined borders. The plots shown in Figures 7-8 confirm that the group 2 trace performs better than the other groups on the basis of overlap and accuracy for all margin groups, therefore supporting our visual observations.

In future work, a worthwhile study would be to test gather more data for all margin groups in an effort to see if the various groups require different parameter values to maximize the algorithm's robustness. Our ultimate goal is to optimize its performance for those masses falling in the ill-defined and obscured margin groups because segmentation of masses falling into those categories is exceedingly difficult.

5. REFERENCES

[1] M.L. Comer, E.J. Delp, "The EM/MPM algorithm for segmentation of textured images: Analysis and further experimental results", *Proceedings of the 1995 IEEE ICIP*, Lausanne, Switzerland, September 16-19, 1996.

[2] J.R. Harris, M.E. Lippman, M. Morrow, S. Hellman, "Diseases of the breast", Lippincott-Raven Publishers, Philadelphia, PA, pp. 80-81, 1996.

[3] M. Heath, K.W. Bowyer, D. Kopans et al., "Current status of the digital database for screening mammography", Digital Mammography, Kluwer Academic Publishers, pp. 457-460 , 1998.

[4] M.A. Kupinski, M.L. Giger, "Automated Seeded Lesion Segmentation on Digital Mammograms", *IEEE Trans. on Med. Imag.*, vol. 17, no. 4, pp. 510-517, 1998.

[5] L. Li, Y. Zheng, L. Zhang, R. Clark, "False-positive reduction in CAD mass detection using a competitive classification strategy", *Med. Phys.*, vol. 28, pp. 250-258, 2001.

[6] J.E. Martin, "Atlas of mammography: histologic and mammographic correlations (second edition)", Williams and Wilkins, Baltimore, MD, p. 87, 1988.

[7] N. Petrick, H-P Chan, B. Sahiner, D. Wei, "An Adaptive Density-Weighted Contrast Enhancement Filter for Mammographic Breast Mass Detection", *IEEE Trans. on Med. Imag.*, vol. 15, no. 1, pp. 59-67, 1996.

[8] J. Suckling, D.R. Dance, E. Moskovic, D.J. Lewis, S.G. Blacker, "Segmentation of mammograms using multiple linked self-organizing neural networks", *Med. Phys.*, vol. 22, pp. 145-152, 1995.

[9] G.M. te Brake, N. Karssemeijer, "Segmentation of suspicious densities in digital mammograms", *Med. Phys.*, vol. 28, no. 2, pp. 259-266, 2001.

[10] B. Van Ginneken, "Automatic segmentation of lung fields in chest radiographs", *Med. Phys.*, 27, pp. 2445-2455, 2000.

6. ACKNOWLEDGMENTS

This work was supported by US Army Grant numbers DAMD17-01-1-0267, DAMD 17-00-1-0291, DAAG55-98-1-0187, and DAMD 17-03-1-0314.