

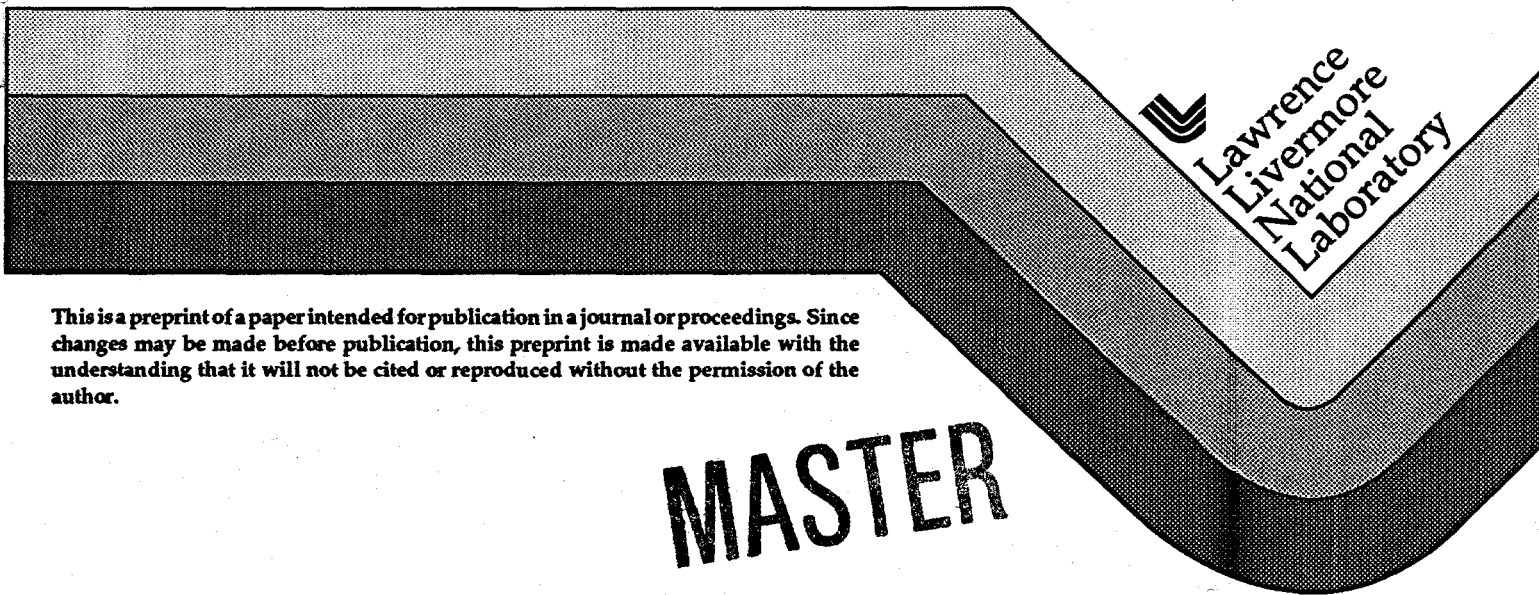
UCRL-JC-113102
PREPRINT

**Automated analysis for microcalcifications
in high resolution digital mammograms**

L.N. Mascio, J.M. Hernandez, and C.M. Logan

**This paper was prepared for submittal to the
SPIE Medical Imaging 1993
February 18, 1993
Newport Beach, CA**

January 1993



This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Automated analysis for microcalcifications in high resolution digital mammograms

Laura N. Mascio, Jose M. Hernandez and Clinton M. Logan

Lawrence Livermore National Laboratory, Biology and Biotechnology Research Program,
L-452, Livermore, CA 94551-9900

ABSTRACT

Digital mammography offers the promise of significant advances in early detection of breast cancer. Our overall goal is to design a digital system which improves upon every aspect of current mammography technology: the x-ray source, detector, visual presentation of the mammogram and computer-aided diagnosis capabilities. This paper will discuss one part of our whole-system approach -- the development of a computer algorithm using gray-scale morphology to automatically analyze and flag microcalcifications in digital mammograms in hopes of reducing the current percentage of false-negative diagnoses, which is estimated at 20% [1]. The mammograms used for developing this "mammographers assistant" are film mammograms which we have digitized at either 70 μ m or 35 μ m per pixel resolution with 4096 (12 bits) of gray level per pixel. For each potential microcalcification detected in these images, we compute a number of features in order to distinguish between the different kinds of objects detected.

1. DIGITIZING MAMMOGRAMS FOR ANALYSIS

The spatial resolution and dynamic range (gray-levels per pixel) required to adequately represent microcalcifications in a digitized film mammogram is an open question. Several studies.[2, 3, 4] suggest that some loss of information occurs at a sampling density of 100 μ m, and it is generally accepted that at least 12 bits of detection resolution are also necessary to prevent loss of information. Yet, we have found no published work on entire mammograms digitized at 100 μ m or less and with 12 bit detection resolution. Therefore, we digitized standard film/screen mammograms using an uncommon digitizer, described in *Equipment* below, which can provide a spatial resolution as low as 35 μ m per pixel and with 4,096 gray levels per pixel over an area the size of a full mammogram.

These sampling specifications increase computational and data management complexity, since a full mammogram can be as large as 50 MB at the highest resolution. Therefore, in order to demonstrate the computer-automated detection capabilities at these sampling specifications, we had to increase computer capacity as described below.

2. EQUIPMENT

Conventional film mammograms were converted to digital format using a digitizer that was designed for precision industrial radiography by DuPont (Wilmington, DE). It can be configured for a sampling density of 35 μ m over a 7 inch by 17 inch area, or 75 μ m over a 14 inch by 17 inch area. The detection resolution is 12 bits per pixel, or 4,096 gray levels. A full mammogram can be digitized and transferred to a workstation in about one minute.

Digitized images were analyzed on a DECstation 5000/200 (Digital Equipment Corporation, Maynard, Massachusetts) running the Ultrix operating system, version 4.2. The system was configured by maximizing the on-board random-access memory (RAM) to

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

128 MB and partitioning the hard drive with 300 MB of contiguous swap space. The software package used for analysis is SCIL-Image (Biological Detection Systems, Pittsburgh, Pennsylvania). This package contains a C-interpreter built onto a comprehensive image processing library and uses the XWindow System for portability and remote analyses.

3. ANALYZING MAMMOGRAMS

We have developed a microcalcification-detection algorithm which operates on digitized mammograms by combining morphological image processing with arithmetic processing. Initial results show that it is a good method for detecting all kinds of microcalcifications and provides a sound platform for discriminating between suspicious and innocuous ones. The algorithm can analyze an entire mammogram in about 10 minutes using the computer hardware and software described above. This includes all computation time; there are no pre-processing or post-processing steps. The entire analysis is automated so there is no human interaction required for processing. The result is a normal mammogram with circled microcalcifications -- interpretation and diagnosis is left for the mammographer.

For a practical demonstration, only a small part of a full digitized mammogram will be used in the figures. Figure 1 shows the small area of one mammogram that will be used in the following discussion.



Figure 1: A portion of a digitized mammogram. According to a radiologist, this mammogram contains many benign microcalcifications and one cluster of three that should be monitored.

To begin, the microcalcification-detection algorithm applies two high frequency analyses to the original digital image and combines the results so that high frequency information common to both analyses are enhanced and that which is common to only one method is de-emphasized. The two analyses are called round high-emphasis and texture gist [5] and can be explained as follows.

1. Round high-emphasis: This technique is a variation of the general method called unsharp masking [6]. The theory for unsharp masking is that a low-pass filtered version of an image is subtracted from the image, leaving high frequency information. In the present case, the general theory is modified by using a round kernel with a diameter of 5 pixels to spatially average the original digitized image. The resulting image is then spatially averaged again with another round kernel of diameter 5. Recursively applying this operation with a small-diameter kernel serves to thoroughly smooth the image while sufficiently maintaining the edges. The result is a low-pass-filtered version of the original image which preserves round edges. This low-frequency result is then subtracted from the original image leaving the high-frequency components of the image. See Figure 2.

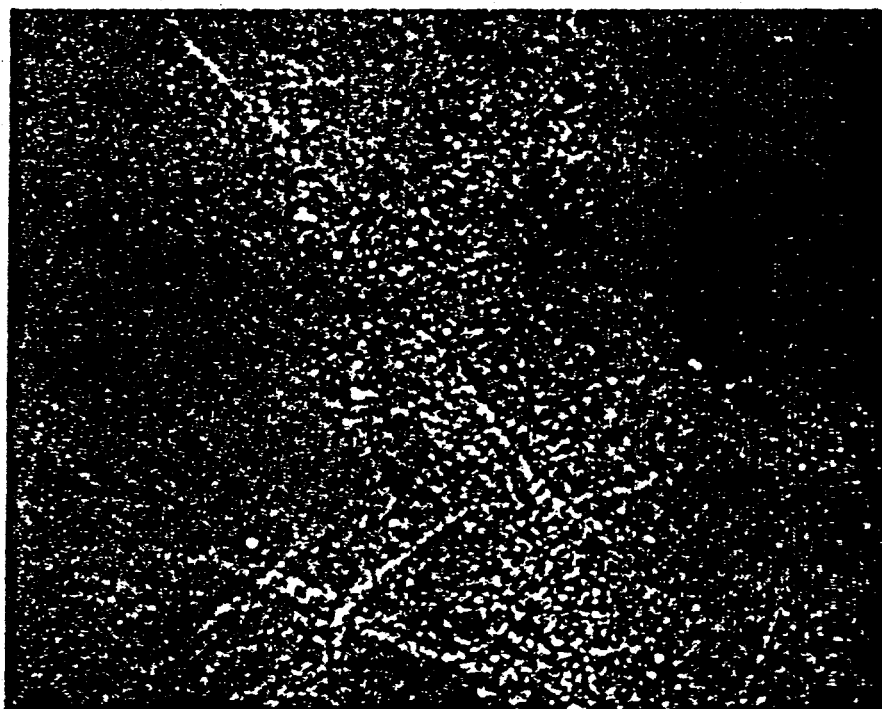


Figure 2: This is the image resulting from applying round-high emphasis to the original portion of the digitized mammogram shown in Figure 1.

2. Texture gist: This method uses gray-level morphological operators [see 7; 8] to yield the upper and lower envelopes of an image. The operators are the minimum

filter (MIN) which is a gray-level erosion, and the maximum filter (MAX) which is a gray-level dilation. The upper envelope is then $\text{MIN}(\text{MAX})$ and the lower envelope is $\text{MAX}(\text{MIN})$. The two envelopes are averaged and the result is subtracted from the original image. See Figure 3.

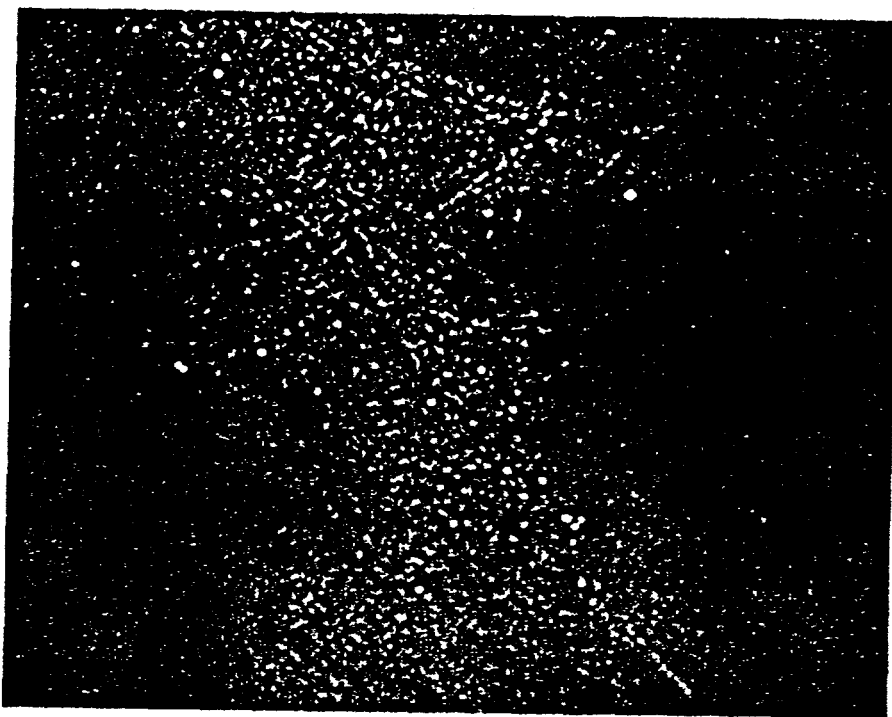


Figure 3: This is the image resulting from applying the texture gist algorithm to the original portion of the digitized mammogram shown in Figure 1.

Both of these methods extract high frequency information from an image, but each emphasizes different aspects of frequency. The first emphasizes any detail in the image which changes sharply in intensity and is larger than several pixels in size. It will highlight streaks or thin lines in the image as well as bright spots. It will not emphasize very small, textured detail in the image which is not highly contrasting with the surroundings. The second method emphasizes detail in the image which is small and textured. Spatial frequency is the governing criterion for this technique. The texture gist will not emphasize lines or streaks, but it will highlight larger, brighter spots in addition to the flecks which make up the texture of tissue or other substances. Adding these images together yields an image which is brightest in locations containing detail common to both. See Figure 4.

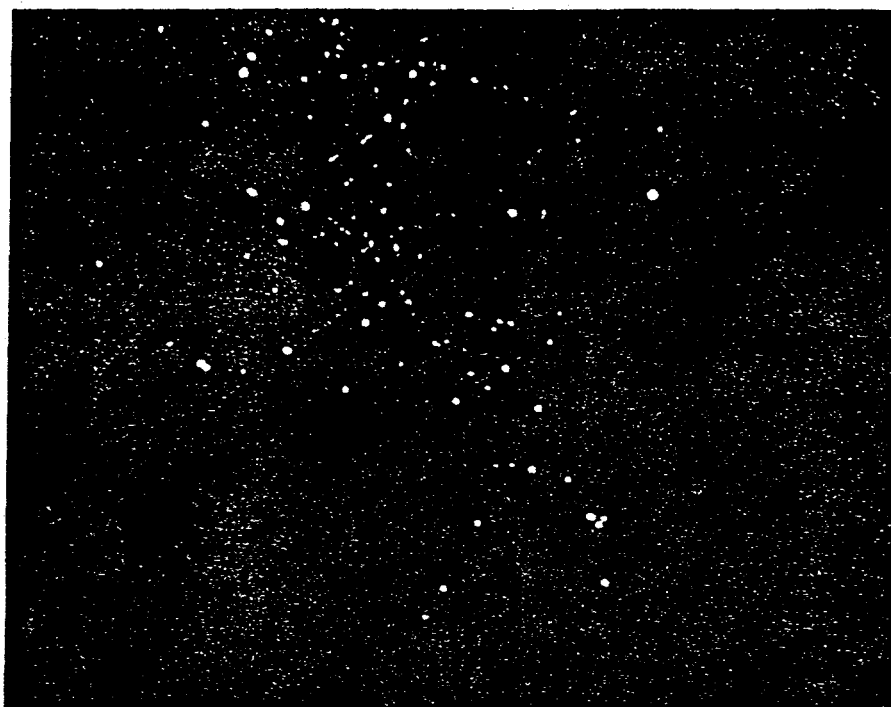


Figure 4: The result of adding Figure 2 and Figure 3 and mapping the result back to 8 bits.

The high frequency result is a gray-scale image, as was the original image, but now, the low-frequency detail is highly de-emphasized in the image. The next step is to segment the image so that all potential objects will be given an equal value and everything else is given a value of zero. This process is called thresholding. In order to automate thresholding, it is necessary to determine some statistics of an image so that a mathematical method can be used to compute a threshold value. We use a technique known as the isodata method [9]. This method uses a histogram of the data (number of pixels vs. intensity) and iteratively selects a value which divides the histogram so its two halves are "balanced" around the chosen value. Applying this value to the image yields a segmentation result in which all microcalcifications are masked with a value of 1 and everything else in the image is black (value of 0). See Figure 5.

At this point, all potential microcalcifications can be circled and presented to a mammographer for diagnosis, as shown in Figure 6. We have done this with local radiologists and have had some good feedback. The algorithm detected several microcalcifications which may have been missed by a radiologist. In addition, however, it found many false-positives in certain cases where tissue texture mimics the frequency information of microcalcifications. This necessitated further discrimination to reduce the number of false positives.

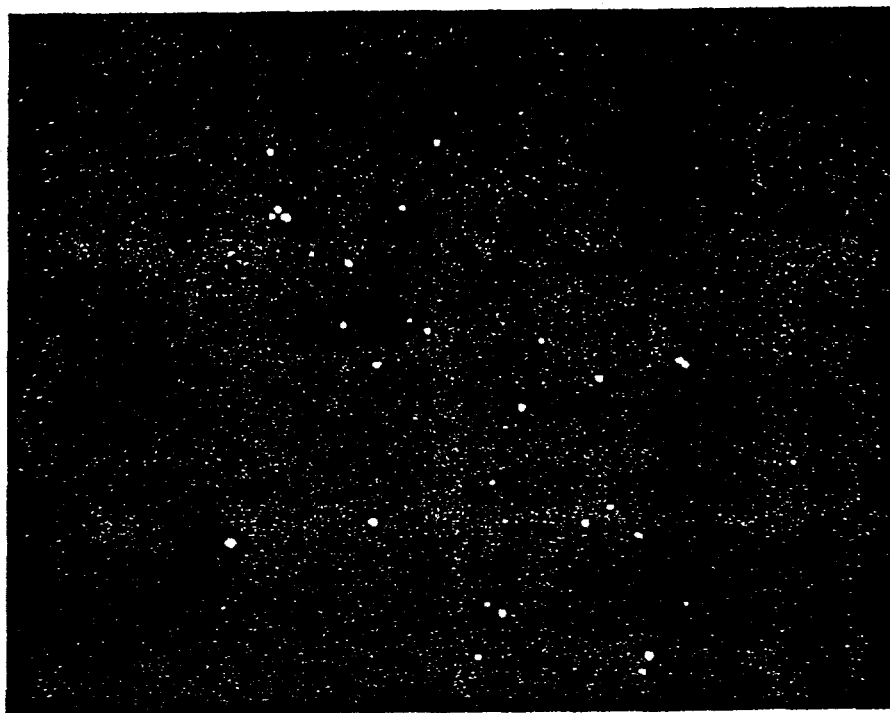


Figure 5: Image resulting from applying a threshold to Figure 4.

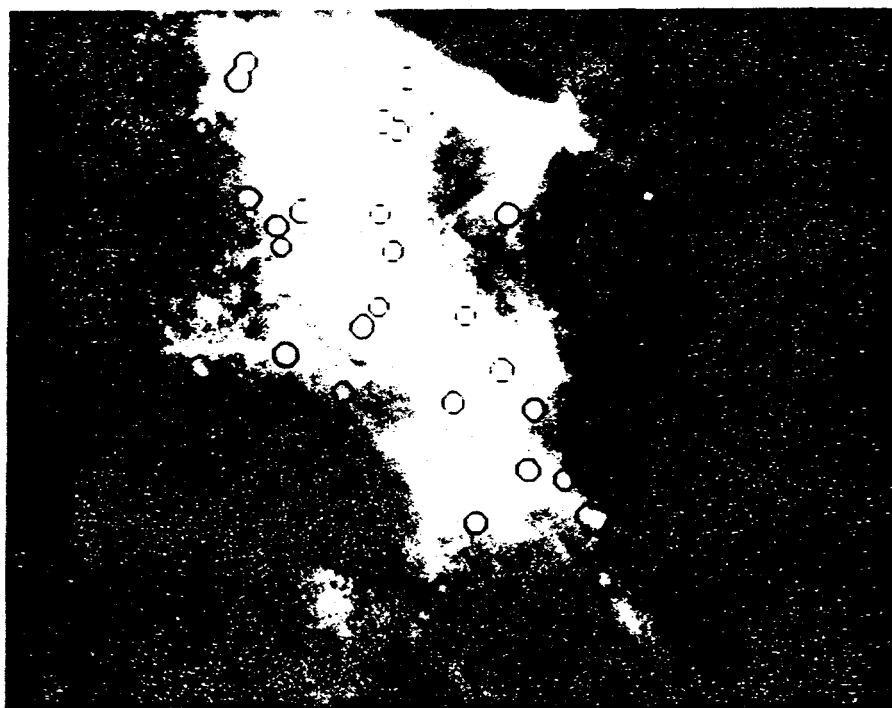


Figure 6: Resulting image with microcalcifications circled. All are almost certainly benign except possibly the cluster of three.

In order to reduce the number of false positives, the segmentation masks (results from Figure 5) are used to extract the potential microcalcifications from the original mammogram for feature extraction. Each object detected must be classified into one of three categories: innocuous microcalcification, suspicious microcalcification or artifact.

Upon conducting a literature search, we found that several groups had independently arrived at different features which they deemed to be important in distinguishing between malignant and benign microcalcifications. Often the importance of the features selected was not substantiated, or substantiating data was not conclusive. In order to determine which combination of features will give the best separation between malignant and benign feature clusters, we have first compiled many of the features used by different groups and have added a few features ourselves. With these measurements, discriminant analysis can be applied to determine which combination of features gives conclusive information as to the diagnosis of a calcification. Therefore, we are measuring all of the following features on calcifications with known diagnosis:

Davies, et al. [10] measured:

- Area, A
- Mean gray level
- Ratio of area to the square of the maximum linear dimension
- Shape parameter, $S = P^2 / 4\pi A$, where P is the perimeter of the object
- Edge Strength = mean value of the Roberts gradient, evaluated for each pixel-edge lying on the object's perimeter.

Chan, et al. [11] used:

- Fourier Power Spectrum
- Contrast (exclude objects with very high contrast)

Additional features we have selected (with input from radiologists):

- Number of microcalcifications in a cluster
- Distance to nearest neighbors (number of calcifications per unit area)
- Aspect Ratio
- Distance to dermal layer

Using discriminant analysis, various features of many known (diagnosed) objects can be analyzed to find the features that give greatest class separation. Then, an object classifier can be built so that only the discriminating features need be computed for each microcalcification in order to classify it as suspicious, innocuous, or as an artifact. After classification, only calcifications classified as having some degree of suspicion are highlighted and presented to a mammographer for subsequent diagnosis. While these results are not available at publication time, we hope to present the results during the conference.

4. CONCLUSION

We are developing a computer algorithm to automatically locate microcalcifications. This computational approach can best be thought of as a "mammographers associate" which will objectively and reproducibly detect microcalcifications and flag them for the mammographer. While this method does not, as yet, attempt to make a diagnosis

concerning the microcalcifications, it will assist mammographers in finding very subtle ones and recognizing the pattern formed by all the microcalcifications. It also will draw attention to those that might be overlooked because a more prominent feature draws attention away from an important object.

5. ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

6. REFERENCES

1. Carolyn Kimme-Smith, "National Cancer Institute Breast Imaging Workshop, Sept 1991," *American Journal of Roentgenology*, Vol 158, p. 268 (1992).
2. H. P. Chan, C. J. Vyborny, H. MacMahon, C. E. Metz, K. Doi, E. A. Sickles, "Digital Mammography: ROC Studies of the Effects of Pixel Size and Unsharp-Mask Filtering on the Detection of Subtle Microcalcifications," *Investigative Radiology*, Vol 22, pp. 581-589, 1987.
3. W-F. Bahaa, S-L Olson, et al., "Algorithm for the Detection of Fine Clustered Calcifications on Film Mammograms," *Radiology*, Vol 169, pp. 333-7, 1988.
4. C. Kimme-Smith, L-W. Bassett, R-H. Gold, L. Gormley, "Digital Mammography: A Comparison of Two Digitization Methods," *Invest. Radiology*, Vol 24, pp. 869-75, 1989.
5. P.W. Verbeek, H.A. Vrooman and L.J. Van Vliet, "Low-Level Image Processing by Max-Min Filters," *Signal Processing*, Vol 15, pp. 249-58, 1988.
6. A.K. Jain, *Fundamentals of Image Processing*, Prentice Hall, Englewood Cliffs, N.J., pp. 249-50, 1989.
7. C.R. Giardina, E.R. Dougherty, *Morphological Methods in Image Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
8. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.
9. T.W. Ridler and S. Calvard, "Picture Thresholding Using an Iterative Selection Method," *IEEE Transactions on Systems, Man and Cybernetics*, Volume SMC-8, No. 8, pp. 630-2, August 1978.
10. D. Davies, D. Dance, C. Jones, "Automatic Detection of Clusters of Calcifications in Digital Mammograms," *SPIE Medical Imaging IV: Image Processing* Vol. 1233, pp. 185-191, 1990.
11. H.P. Chan, K. Doi, C. Vyborny, R. Schmidt, C. Metz., K.L. Lam, T. Ogura, Y. Wu and H. MacMahon, "Improvement in Radiologists' Detection of Clustered Microcalcifications on Mammograms," *Investigative Radiology*, Vol 25 No 10, pp. 1102-10 (1990).