GRANT NO: DAMD17-94-J-4292

TITLE: Improved Mammographic Technique for Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Heang-Ping Chan, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, Michigan 48109-1274

REPORT DATE: 8/9/95

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
The goal of the proposed research is to improve the sensitivity of cancer detection in mixed or dense breasts through optimization of mammographic techniques. We propose to develop a novel exposure equalization system that preferentially reduces the incident x-ray intensity in the peripheral region of the breast, thereby alleviating the problem of limited latitude of x-ray detectors. Optimal imaging technique can then be developed for improving image quality throughout the entire breast.

The studies performed in the first year of the funding period confirmed that designing an x-ray equalization filter device for nearly patient-specific equalization in mammographic imaging is a viable approach. We conducted several studies in order to design the filter device: (a) testing of the breast shape classification method, (b) analysis of the exposure profile of breast images in order to design the exposure equalization filter, (c) development of a fully automated breast image segmentation and edge detection program on a personal computer (PC) to prepare for on-line analysis of patient breast shape. The results of these studies will be used for the implementation of the filter device as planned in the future years.
The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to stay within the lines to meet optical scanning requirements.

**Block 1. Agency Use Only (Leave blank).**

**Block 2. Report Date.** Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

**Block 3. Type of Report and Dates Covered.** State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

**Block 4. Title and Subtitle.** A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

**Block 5. Funding Numbers.** To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

- C - Contract
- G - Grant
- PE - Program
- ELement

**Block 6. Author(s).** Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

**Block 7. Performing Organization Name(s) and Address(es).** Self-explanatory.

**Block 8. Performing Organization Report Number.** Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

**Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es).** Self-explanatory.

**Block 10. Sponsoring/Monitoring Agency Report Number. (If known)**

**Block 11. Supplementary Notes.** Enter information not included elsewhere such as: Prepared in cooperation with..., Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the old report.

**Block 12a. Distribution/Availability Statement.** Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- DOD - See DoD 5230.24, "Distribution Statements on Technical Documents."
- DOE - See authorities.
- NTIS - Leave blank.

**Block 12b. Distribution Code.**

- DOD - Leave blank.
- DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- NASA - Leave blank.
- NTIS - Leave blank.

**Block 13. Abstract.** Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

**Block 14. Subject Terms.** Keywords or phrases identifying major subjects in the report.

**Block 15. Number of Pages.** Enter the total number of pages.

**Block 16. Price Code.** Enter appropriate price code (NTIS only).


**Block 20. Limitation of Abstract.** This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.
FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

X Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

X In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

X In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

X In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC–NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Chan Heang Pong 8/9/95
PI – Signature  Date
(4) Table of Contents

(4) Table of Contents ................................................................. 4
(5) Introduction ........................................................................... 5
(6) Body ....................................................................................... 5
   (a) Testing of breast shape classification algorithm .................. 6
   (b) Determination of filter thicknesses ...................................... 6
   (c) Automated edge detection .................................................. 8
(7) Conclusion ............................................................................... 15
   Tasks to be performed next year ............................................. 15
(8) References ............................................................................... 16
(9) Appendix .................................................................................. 17

<table>
<thead>
<tr>
<th>Accession For</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTIS CRA&amp;I</td>
</tr>
<tr>
<td>DTIC TAB</td>
</tr>
<tr>
<td>Unannounced</td>
</tr>
<tr>
<td>Justification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>By</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Availability Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dist</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>A-1</td>
</tr>
</tbody>
</table>
(5) **Introduction**

Breast cancer is one of the leading causes of death among women. There is considerable evidence that early diagnosis and treatment significantly improve the chance of survival for patients with breast cancer (refs. 1-6). At present, x-ray mammography is the only diagnostic procedure with a proven capability for detecting early-stage, clinically occult breast cancers (refs. 5-8). Although mammography has a high sensitivity for detection of breast cancer when compared to other diagnostic procedures, studies indicate that radiologists identify only 70 to 90% of the lesions present (refs. 4-6, 9-11). The miss rate is particularly high in dense breasts (refs. 12, 13).

One of the difficulties in interpretation of mammograms by radiologists is caused by the limited latitude and contrast sensitivity of mammographic screen/film systems. Mammographic abnormalities related to early breast cancers include clustered microcalcifications, spiculated and irregular masses, areas of parenchymal distortion, and skin thickening (refs. 14,15). These abnormalities are often subtle and low contrast. Therefore, low energy radiation and high contrast screen/film systems are recommended for mammographic imaging in order to increase the contrast between the lesion and the background tissue. Despite the use of vigorous compression during examinations (ref. 16), the low-energy x-ray beam results in a wide dynamic range (the ratio of the maximum to the minimum x-ray exposure at the detector) for the radiation penetrating the breast. This range can be greater than 100 (ref. 17). On the other hand, high-contrast film provides a narrow latitude which is about 10 for a typical mammographic system (refs. 18, 19). As a result, thick and glandular regions of the breast are often imaged at the toe of the sigmoid-shaped sensitometric curve of the screen/film system; whereas thin peripheral regions are imaged at the shoulder. The contrast and signal-to-noise ratio (SNR) of mammographic features are greatly reduced in these regions due to decreased film gradient and increased noise. The contrast sensitivity of the human visual system also drops rapidly as the film density increases (refs. 20-22). Kopans (ref. 12) found that 70% of breast cancers in women with dense breasts are in the periphery of the mammary parenchyma adjacent to the subcutaneous fat or retromammary fat. The poor image quality in the peripheral region thus imposes a serious limitation on the sensitivity of cancer detection in breasts with dense fibroglandular tissue.

We propose a practical and cost-effective exposure equalization method for reducing the dynamic range of the x-ray image. The shapes of compressed breasts of the patient population will be analyzed and classified into a finite number of groups. A shaped filter for attenuating x-rays in the peripheral region of the breasts will be fabricated for each group. For a given patient, the breast shape during compression will be classified into one of these groups and the filter for the selected group will be used for this patient. With this technique, the dynamic range of the x-ray intensities incident on the recording system will be reduced and the entire image can be recorded in the high contrast region of the film. The improved image quality can be achieved without additional radiation dose to the patient. Furthermore, a very high-contrast mammographic technique may be developed in combination with exposure equalization to further improve the signal-to-noise ratio (SNR) of the subtle lesions. We expect that the optimized technique will significantly improve the detectability of cancers in mixed and dense breasts and increase the efficacy of mammography as a screening and diagnostic tool for breast cancers.

(6) **Body**

In the first year (7/11/94-7/10/95) of this grant, we have performed the following studies:
(a) Testing of breast shape classification algorithm

A key element of the external equalization filter method is the hypothesis that compressed breast shapes can be classified into a limited number of groups. We performed a preliminary study to test this hypothesis. Over 70 randomly selected craniocaudal (CC) view and more than 100 lateral (LAT) view mammograms were digitized and analyzed. The anterior borders of the breasts were detected using an edge detection algorithm developed for breast images (ref. 23). The detected border was rotated and translated on the x-y plane to minimize difference in the positions and orientations of the borders among different mammograms. An analytical curve was fitted to the breast border with least squares criterion. Within the different types of curves that we tested, a polynomial of the form,

\[ f(x) = ax^2 + bx^3, \]

(1)

provided the best fit and required only a small number of parameters, i.e., 2, for shape classification. This was true for both the CC and LAT views. We found that the breast borders of 105 LAT mammograms could be classified into 6 groups based on the a and b parameters obtained by curve fitting. Similarly, the breast borders of 75 CC mammograms could be classified into 4 groups based on a and b. An average border shape was determined for each group and the deviation of the individual borders within a group from the average border was calculated. The standard deviation of the groups varied from about 1 mm to 4 mm.

We have performed a study to test the classification of compressed breast shapes based on the trained classifier described above. A test set of 55 CC view and 68 LAT view mammograms were randomly selected from patient files in the Department of Radiology at the University of Michigan. The mammograms were digitized and edge detection was performed. A detected border was translated and rotated on the x-y plane and curve fitting using eqn. (1) was performed to determine the a and b parameters. The breast shape was then classified, based on the values of a and b, into one of the groups determined in the training process. We found that 98% of the 55 CC view mammograms can be classified into one of the 4 groups with a root-mean-square difference of less than 2 mm, and that 97% of the 68 LAT view mammograms can be classified into one of the 6 groups with a root-mean-square difference of less than 2 mm. Fig. 1(a) and 1(b) show the training and test classification of the breast borders in group 1 of the CC view mammograms, while Fig. 2(a) and 2(b) illustrate the corresponding results in group 1 of the LAT view mammograms. The error bars indicate one standard deviation of the distribution of breast border shapes in the group.

The significance of this study lies in the fact that the results support our hypothesis that a small number of pre-fabricated filters will be sufficient to allow selection of a nearly patient-specific filter for each breast being examined. It indicated the feasibility of our approach to exposure equalization.

(b) Determination of filter thicknesses

For each group of mammograms, we estimate the average exposure profile of the breasts along an approximately radial line from the chest wall to the breast border as follows. The digital pixel values are converted to optical densities by use of the calibration curve of the digitizer. The optical density profiles are further converted to a relative-exposure profile by use of typical sensitometric curves of the mammographic films. This exposure profile represents the total, i.e., primary plus scattered, radiation recorded by the film. Because we cannot make measurements of scattered radiation for each patient, a scatter fraction profile measured by phantom studies (ref. 24) is used for an approximate scatter correction. The exposure profile after scatter correction is assumed to be that due to penetration of primary x-rays through the breast. An average exposure profile of the primary beam is then calculated for dense and mixed breasts within the same group. Based on this average exposure profile of the breasts, we can
derive the exposure attenuation profile required of a filter such that the exposure in the peripheral region of a dense breast can be reduced to the average level in the central breast area.

Due to the variations in breast thickness and variations in the gradients of thickness at the compressed breast border, it is possible that the relative exposures in the outer border regions may be quite different for two breasts that have nearly identical outer contours. To determine the magnitude of this problem, we performed a preliminary study in which we compared the relative x-ray exposure at the compressed breast outer border regions for a set of mammograms that were all classified within one group based on the detected compressed breast borders.

A computer program, EXP_NORMAL, was written which converts digitized film pixel values into relative x-ray exposures along normals to the fitted breast borders. Some details concerning this program follow: The program reads in the digitized mammogram data for all images within a given class of breast border shapes. It computes the locations of 17 equally spaced normals along the average fitted border for the class of breast shapes. The normals are each separated by 10 mm along the border, and each normal consists of 21 points with a spacing between points of 1 mm. The general equation of the normals is:

\[
y = y_1 + \frac{-1}{2ax_1 + 3bx_1^2} \cdot (x - x_1)
\]  

(2)

where a and b are the fitted coefficients of the average breast border and \( x_1 \) and \( y_1 \) are points on the border. The pixel values at each point along the normals are converted to optical densities assuming a linear transformation, e.g.

\[
OD(i,j) = \frac{-3.945}{4095} \times \text{image}(i,j) + 3.945,
\]  

(3)

where 3.945 is the maximum optical density, and we use 12-bit (0 to 4095) digitization. These optical densities are then transformed into relative x-ray exposures using an optical density vs. relative x-ray exposure film-characteristic-curve provided to us by William Moore, Ph.D. of Kodak. From the set of images, the program computes the average relative exposures at each point along the normals and the standard deviations of those exposures. The results are output in a format that can be plotted on a Macintosh computer using the commercial program Cricket Graph. Another program, AUTO_3D_NORAL, was also written to output the average values along the normals in a format that could be plotted in 3-D by the commercial program DeltaGraph.

Results

Figure 3 shows the computed average compressed breast border for a given class of compressed breast shapes as well as the 17 normals to the average border all superimposed on a digitized mammogram of a breast that is a constituent of the class. We applied histogram equalization to the digitized mammogram to enhance the gray-scale contrast so the entire breast out to the edges could be visualized.

Plots of the average relative exposure at the film plane as a function of position along several of the normals to the average breast border are shown in Fig. 4. On the abscissa of this plot, normal position 11 corresponds to the breast border, and positions less than 11 are outside the breast border, while those greater than 11 are inside the breast.

A 3-D representation of the mean exposure values of the normals to the average compressed breast border is shown in Fig. 5.

The variability of the exposure values along the normals within the set of images that form the class is demonstrated in Figs. 6A-E, below. In these figures, the mean relative exposure
is plotted (as in Fig. 4) along with the mean ±1 standard deviation. Approximately 68% of the relative exposures along the normals will fall within the mean ±1 standard deviation curves.

Discussion

The equalization filters will be designed to flatten out the exposure profiles shown in Figs. 4 and 5. As shown in these figures, the profiles along the normals for the images within this class are fairly similar in shape. However, there are differences in the slopes of the profiles and the maximum exposure levels. Future studies using a larger number of images and image classes will be performed to determine the extent to which the slope of the equalization filter thickness must be customized as a function of position along the detected boundary of the compressed breast.

The relative exposures that are plotted in the figures were computed from optical densities on mammograms. Both primary and scattered radiation contributed to the film exposures and optical densities. The filter will absorb primary radiation incident upon the breast. Therefore, to design a more effective filter, it is necessary to isolate the primary radiation contribution to the total exposure at the film plane. We have been writing a Monte Carlo routine to accomplish this. This Monte Carlo routine using the MCNP4A code developed at Los Alamos National Laboratory, Los Alamos, New Mexico. In this code, we are simulating all aspects of the imaging chain. We employ measured mammography spectra measured by Robert Jennings, Ph.D. of the BRH. Breasts of various thicknesses, compositions and shapes are modeled. We have developed a code to describe the grid. Scatter to primary ratios at various sites inside and outside the compressed breast will be computed from the relative amounts of energy absorbed at corresponding sites in the screen due to the two types of radiation.

We previously found that the shapes of the compressed breasts within a given class differ most near the chest wall. This no doubt contributes to the larger standard deviations in relative exposures that are observed along normals 1 and 17 in Fig. 6. The plots in Fig. 6 also show relatively large standard deviation inside the breast in the nipple region (e.g. along normal # 9). Large variations in relative exposure for breasts within a given class makes it unlikely that a single filter will adequately equalize all of the images, i.e., some will be much better equalized than others. We plan to perform studies to quantify the amount of exposure variation that is acceptable. It is possible that additional filters will be needed to accommodate exposure variations that are greater than a certain value.

(c) Automated edge detection

We have developed a fully automated breast border detection program. A mammogram is digitized with a CCD digitizer and the image file is analyzed with an edge detection and tracking program. The program first eliminates the unexposed area and labels from the image, and then finds the breast border within the area of the film exposed by x-rays. Our algorithm uses a two-pass approach. In the first pass, a simple edge detection routine using the first and second derivatives locates the approximate edge coordinates for each line. These coordinates are used as a tracking path for the accurate edge detection routine. In the second pass, the program first estimates the mean and standard deviation of the background level just outside the breast border paralleling the simple edge coordinates. These local statistics are used to determine an adaptive threshold criterion to find the exact edge location. Starting from about the nipple of the breast image and tracking in both directions, the program searches for an edge point that passes the second derivative and threshold criterion within a range of pixels centered on the approximate edge. For subsequent edge points, the edge tracking routine uses previously found edge points to adjust the search direction and the approximate location of the breast border. The
edge tracking routine searches in both directions starting from the nipple of the breast and continues to the edge of the image.

This model-guided search method reduces the likelihood that the edge tracking routine will be misled by noisy background fluctuations. At present, the edge tracking routine can accurately detect the breast borders in over 97% of the images digitized for the breast shape analysis. We have so far collected 123 cases of randomly selected patient mammograms. Most of the cases consist of four views, the cranio-caudal (CC) view and the mediolateral (MLO) view of the left and right breasts. There have been a total of approximately 464 mammograms digitized for this project, about half of which are CC views, and half are MLO views. An example of a case with the four views is shown in Fig. 7. It can be seen that the edge tracking program follows the breast borders on the displayed images very closely. A comparison study of the detected breast borders with a manually traced 'true' breast border is underway to quantify the accuracy of the edge detection algorithm.

Figure 1. Average breast border for CC view mammograms: (a) group1 training result, (b) group 1, test result. The error bars represent ± one standard deviation of the distribution of breast borders within the group.

Figure 2. Average breast border for LAT view mammograms: (a) group1 training result, (b) group 1, test result. The error bars represent ± one standard deviation of the distribution of breast borders within the group.
Figure 3. Digitized mammogram showing the average breast border and 17 normals to that border for a group of mammograms within a given class.
Figure 4. Plots of mean exposure values as a function of position for normals number 1, 5, 9, 13, and 17. These are represented by m1, m5, m9, etc. in the figure. Normal number 1 is at the top of the mammogram, normal number 9 is about at the location of the nipple, and normal number 17 is at the bottom of the mammogram (See Fig. 3).

Figure 5. 3-D plot of relative mean exposure as a function of position for the 17 normals.
Figure 6. (A-F) Mean relative exposure along normals ± 1 standard deviation for normals 1, 5, 9, 13, and 17. Positions 1 to 10 along the normal (see abscissa) are outside the breast, and positions 12-21 are inside the breast.

(See Figure 7 in the following pages)

Figure 7. (a) An example of the breast borders found by our automated edge tracking program in the four mammograms (Left: CC, MLO; Right: CC, MLO) of the same patient. (b) The four mammograms used in Fig. 7 (a) shown without and with the breast borders superimposed.
Mammography Tracking Program

R-CC

R-MLO

L-CC

L-MLO
(7) Conclusion

The studies performed in the first year of the funding period confirmed that designing an x-ray equalization filter device for nearly patient-specific equalization in mammographic imaging is a viable approach. We have conducted several studies in order to design the filter device: (a) testing of the breast shape classification method, (b) analysis of the exposure profile of breast images in order to design the compensating filter, (c) development of a fully automated breast image segmentation and edge detection program on a personal computer (PC) to prepare for on-line analysis of patient breast shape. The results of these studies will be used for the implementation of the filter device.

Tasks to be performed next year

1. Develop computerized edge detector and breast shape classification scheme --

   We will perform a study to determine the accuracy of the newly developed compressed breast border detector. We will manually trace the borders on the digitized mammograms which will be edge- and contrast-enhanced and displayed on a workstation. These borders will be compared with those detected automatically with the new algorithm. The comparison will be in terms of a root-mean-square distance between corresponding borders. A total of 500 images will be compared.

   We will continue to expand our database of digitized mammograms for analysis of the compressed breast shapes and exposure profiles. The adequacy of using the two-term polynomial \( y = ax^2 + bx^3 \) for classification of this larger set of images will be examined. Alternative shape descriptors and classification schemes will be investigated.

2. Design and build exposure equalization filters --

   The Monte Carlo code described above will be completed, and scatter-to-primary ratios at various sites will be determined. The exposures computed from digitized mammograms using sensitometric curves of the films will be corrected to eliminate the scatter component. The exposure profiles of the primary radiation recorded on the detector will be used to design the x-ray equalizing filters. Simulation studies will be performed to determine the filter profiles with different materials.

3. Design and build filter positioning apparatus --

   We will study the proper location of mounting the filter positioning apparatus on commonly used mammographic systems. Preliminary design of the filter positioning device will begin.

4. Evaluate effects of equalization on image quality and patient dose by Monte Carlo simulation studies and optimize imaging techniques for dense breasts--

   We plan to use the MCNP4 Monte Carlo program, incorporate the mammographic imaging geometry, breast tissue equivalent phantom, and detector response, and investigate the dependence of image quality and patient dose on imaging conditions and breast thickness and density. The effects of equalization on dense breasts will be analyzed.

5. Study the feasibility of very high-contrast technique --
The Monte Carlo simulation will be used to study the feasibility of very high-contrast technique using low kVp and/or high contrast films.

6. Design and build custom phantoms --

The exposure profiles of the primary radiation penetrating the breasts will be used to estimate average breast densities and thickness profiles for different classes of breast shapes. We will begin designing of custom breast phantoms based on the estimated average breast densities and thickness profiles.

(8) References


(9) Appendix

Publications as a result of this grant
