AD_____

GRANT NUMBER: 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: August 1996

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1996	Annual (11 Jul 95 - 10 Jul 96)

4. TITLE AND SUBTITLE
Improved Mammographic Technique for Breast Cancer Diagnosis

5. FUNDING NUMBERS
DAMD17-94-J-4292

6. AUTHOR(S)
Heang-Ping Chan, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, Michigan 48109-1274

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200)*

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.

We have performed the following studies in the second year of the funding period: (1) collected a large data base for analysis of breast shapes and exposure profiles, (2) evaluated the accuracy of the automated breast border detection program, (3) evaluated the adequacy of two polynomial models for fitting the breast shapes and classification, (4) developed a clustering procedure for classifying the breast shapes into a small number of classes, (5) developed an automated breast shape classification program to prepare for on-line analysis of patient breast shape, (6) developed a Monte Carlo simulation model for mammographic imaging (7) evaluated experimentally and by Monte Carlo simulation the scatter fraction on the imaging plane and studied the primary exposure profiles in a given breast shape class. The results of these studies will be useful for the implementation of the filter device as planned in the future years.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer Mammography, exposure equalization, dynamic range compression breast cancer detection	25 16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	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.

APC 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.

____ 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).

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

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

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

____ 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 Ping 8/9/96

PI - Signature Date

(4) Table of Contents

(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 proposed a practical and cost-effective exposure equalization method for reducing the dynmaic 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 second year (7/11/95-7/10/96) of this grant, we have performed the following studies:

(a) Evaluation of Automatic Border Tracking Program

We developed an automated breast border tracking program and began to collect a data base of mammograms for analysis of breast shapes in the first year. We continued to collect the data base in the second year. Mammograms were randomly selected from patient files in the Division of Breast Imaging. Most of the selected cases consisted of four views: the cranio-caudal (CC) view and the mediolateral (MLO) view of the left and right breasts. The selected mammograms were digitized with a 21-micron resolution CCD film digitizer. To reduce storage space and image processing time, the resolution of the digitized images were reduced to 1 mm pixel size before storage and analysis. A total of 1004 MLO and CC view mammograms have been collected in the data base.

A study was performed to determine the accuracy of the computer routine that was developed to trace the outer contours of the breasts in digitized mammograms. Hand traced borders were compared to those that were automatically tracked. To trace the borders by hand, the mammograms were first processed with a Sobel filter to enhance the edges. Next, a rough border was drawn using a commercially available program, OSIRIS. Then, the hand-traced borders were meticulously adjusted to best match those observed visually. This was accomplished with the OSIRIS program. Multiple image magnification levels and gray-scale enhancement settings (window level and width values) were applied in an iterative operation in which the many border position adjustment points or curve "handles" were individually moved to fine-tune the position and shape of the hand-drawn border. A computer program, CONTOUR_ERROR, was written to measure the mean and root mean square distances between the hand traced and automatically traced borders along normals to the hand traced borders. In this program, positive distances corresponded to the hand-drawn border being inside the automatically traced border and negative distances corresponded to the hand-drawn border being outside the automatically traced border. Histograms of the frequencies of occurrence of distances in bins of 1 mm width (e.g. -2.5 to -1.5, -1.5 to -0.5, -0.5 to +0.5, 0.5 to 1.5, etc.) were also computed. The program was applied to 3 sets of images: (1) a "Test" set consisting of 3 images for which the vast majority of the automatically traced border points had very high confidence values, and 3 for which the confidence values were less; (2) a "Difficult" set consisting of 10 images for which many of the automatically traced border points had very low confidence values, and (3) a "Random" set of 10 images selected randomly from the data base. The results are summarized in Tables 1-3 below.

Table 1. RMS and mean distances between hand-traced and automatically traced breast borders: TEST SET

	RMS Distance (mm)	Mean Distance (mm)
	1.20	0.73
	1.09	0.72
	1.24	0.81
	1.04	0.75
	0.98	0.77
	0.88	0.48
Mean for Group:	1.07	0.71

Table 2. RMS and mean distances between hand-traced and automatically traced breast borders: DIFFICULT SET

	RMS Distance (mm)	Mean Distance (mm)
	5.98	4.58
	3.80	-2.74
	4.59	-3.49
	3.25	-2.10
	4.07	-1.44
	3.21	-2.11
	6.07	-4.37
	2.93	-1.28
	5.14	-3.40
	2.77	-1.77
Mean for Group:	4.18	-1.81

Table 3. RMS and mean distances between hand-traced and automatically traced breast borders: RANDOM SET

	RMS Distance (mm)	Mean Distance (mm)
	0.69	0.39
	0.83	0.59
	1.03	0.83
	1.16	0.98
	3.88	-0.73
	1.18	1.04
	1.50	0.89
	1.50	0.96
	1.04	0.73
	1.40	4.52
Mean for Group:	1.42	0.61

In general, the above results demonstrate excellent agreement between the hand-drawn and automatically determined borders. The mean distances for the Test and Random groups are less than 1 pixel (1 pixel = 1 mm), and even the Difficult borders agreed to within a mean distance of less than 2 pixels.

A review of the histograms of the frequencies of occurrence of the distances between the borders in the Test and Random sets shows that for nearly all of these borders, 50-70% of the distances fall in the bin centered at +1 mm. This indicates that, on the average, the automatically traced borders are 1 mm (= 1 pixel) outside the hand-drawn borders. It is most probably a systematic error due to differences in the criteria for defining the edge for the two methods. The automatic tracking method takes a gradient from outside the breast to inside the breast and defines the edge as the outermost point of a set of 4 contiguous points characterized by increasing or 0 gradients (e.g., pixel value (point 1) $\geq$ pixel value (point 2) $\geq$ pixel value (point 3) $\geq$ pixel value (point 4), where point 1 is just outside the breast and point 4 is well inside the breast.) On the other hand the border points that were selected by eye tended to exhibit the maximum gradient which in general corresponded to point 2 of the automatic tracking method. An example of the automatic and hand-drawn borders is displayed in Figure 1, below.

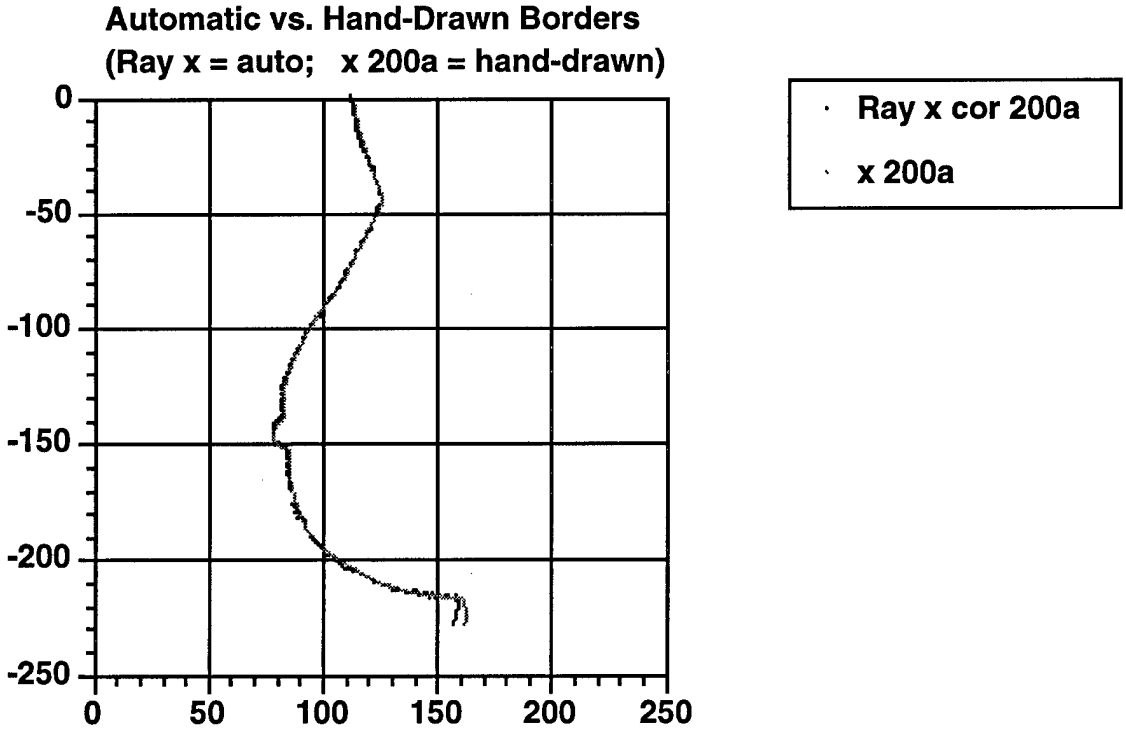

Figure 1. Example of automatic vs. hand-traced compressed breast borders. The hand-traced border is indicated by the lighter gray curve.

The results of the above study as well as a description of the automatic tracking procedure was presented at the 1996 annual meeting of the American Association of Physicists in Medicine (Medical Physics 1996;23:1107-1108).

This study demonstrates that the automated border detection program can accurately track the breast borders. The breast borders will be used for breast shape analysis and filter design as described below. When the x ray equalization filters are incorporated in a mammography system, the border detection program will also be needed for on-line detection of the patient breast shape so that an appropriate filter can be chosen for the patient. In that case, the border will be traced from a visible light image of the breast as recorded by a TV camera interfaced to a PC.

(b) Compressed Breast Shape Classification

In the progress report last year, we described the preliminary study of fitting 75 CC and 105 MLO view compressed breast borders with the equation $y = ax^2 + bx^3$ and the success of classifying the borders using the "a" and "b" coefficients of the fitted curves. A more extensive investigation using the newly collected large data base was performed this year. Some of the distinguishing features of this investigation included:

(1) The evaluation of a much larger set of images - 470 CC-view and 484 MLO view.
(2) The employment of smoothed (run-length averaged) borders for the fits rather than the raw data of the detected borders.
(3) The fitting of a greater portion of the borders. ~75-100% of the automatically determined borders were fit; whereas, previously only the anterior portion (~50%) was fit.

(4) The comparison of both fitting the borders and classifying the borders with $y = ax^2 + bx^3$ vs. $y = ax^2 + bx^3 + cx^4$.

(5) The use of a more standard and efficient clustering method for classification of the borders. Previously the entire "a" -"b" space was divided into 4, 6, 8, or 10 equally dimensioned boxes to define the clusters. In the present study, the K-means clustering method in the SPSS statistical package [ref. 23] which employs an iterative nearest neighbor centroid sorting technique was employed.

(6) The evaluation of a "hybrid" clustering method in which the a,b parameters were employed (c set to 0) when the fit to a given border using $y = ax^2 + bx^3$ was better than a threshold value, and the a,b, and c-parameters of the fit of the border to $y = ax^2 + bx^3 + cx^4$ was employed otherwise.

Considerable time and effort were expended in making the fitting program more efficient. Part of this involved developing a better method for defining the starting axes for the fits. Previously, a very simple method was employed in which the origin was selected as the minimum y-value of the automatically traced border (the orientation of the automatically traced borders was such that the nipple was towards the bottom of the image) and the starting x and y axes were perfectly horizontal and vertical relative to the image frame. A variety of alternative methods were tried. The one that was found to work best defined the y-axis as the linear fit to the midpoints of a set of lines drawn between selected pairs of points from the left to the right sides of the smoothed border. The name of the computer program that was developed for this purpose was STARTAXIS2. The selected pairs of points in this routine were defined as follows. The derivative (dy/dx) of the first 25 points on the left side of the border (starting from the chest wall) was computed, and the starting point "(x1, ymax_left)" was chosen as the point at which that derivative was a minimum (e.g. most negative). Next, the derivative (dy/dx) was computed starting 25 points from the right end of the contour. The ending point "(x2, ymax_right)" was chosen as the point at which that derivative was a maximum (i.e. most positive). The number of points between the starting point and the point on the border at which the y-value was a minimum (x3, y_min) and the number of points between the ending point and (x3, y_min) were then computed. The smaller of those numbers of points was divided into 5 equal increments (inc). The midpoint of the line drawn between the (starting+inc) point on the left and the (ending-inc) on the right was computed as were the midpoints of the lines between the (starting+n*inc) point on the left and the (ending-n*inc) point on the right , where n=2,4. Finally, a line was least square fit through the above-mentioned midpoints. This line represented the y-axis. An example of one of the y-axes that was determined using this method is displayed in Figure 2, below. The origin was selected as the point at which the y-axis intersected the border (near y-min), and the x-axis was the line perpendicular to the y-axis passing through the origin. This method was found to work very well for both CC and MLO views.

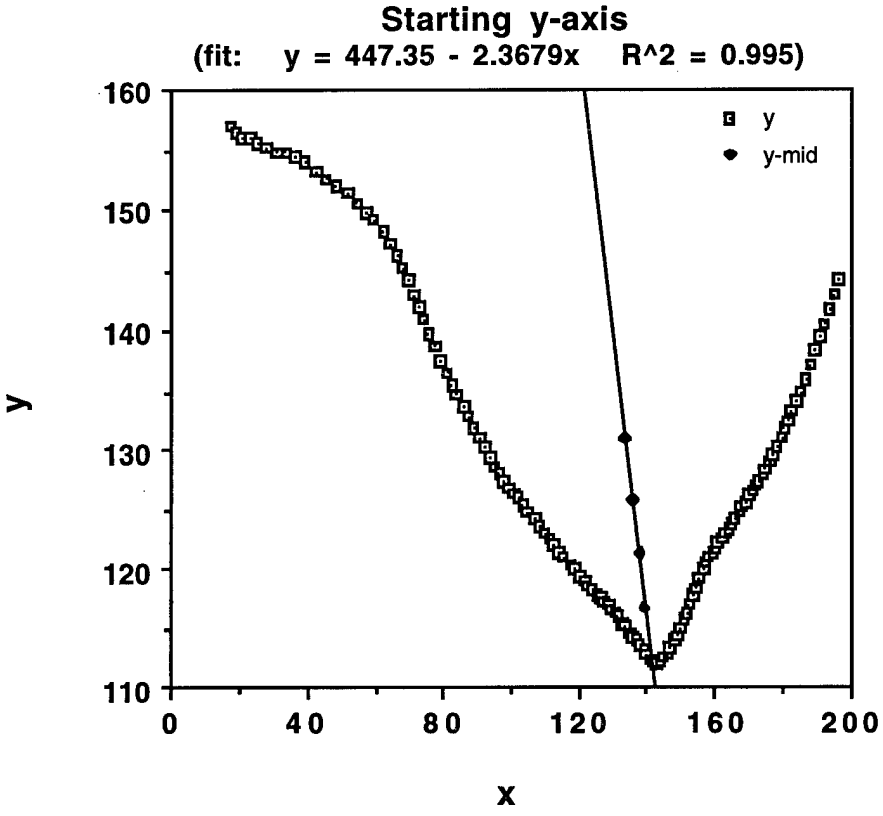

Figure 2. Example of y-axis computed using subroutine STARTAXIS2. The border is translated and rotated about this axis to find the best fit to the equations $y = ax^2 + bx^3$ and $y = ax^2 + bx^3 + cx^4$.

The computer program that translates and rotates the breast borders about the axes (or equivalently rotates and translates the axes about the borders) was extensively rewritten to improve its efficiency. In particular, use of coarse increments followed by fine increments in the x and y translations and the rotations resulted in a tremendous reduction in the amount of time required to find the best fit. Other changes made in the ordering and style of the computations (e.g., use x*x*x instead of x**3) had a substantial effect. Overall, the revised program is 100 times faster than its predecessor. It now takes about 1/2 second to perform the fitting vs. about 20 seconds previously. The 1/2 second time is for both the fit to the equation $y = ax^2 + bx^3$ and for the fit to the new equation $y = ax^2 + bx^3 + cx^4$ which was employed for comparison purposes The 1/2 second time was achieved using coarse x-translation ranges of -40 mm to +40 mm in steps of 4 mm, y-translation ranges of -20 mm to +20 mm with steps of 4 mm, and rotation ranges of -40 degrees to +40 degrees with steps of 4 degrees. The subsequent fine x- and y-translation ranges were $\pm$ 4 mm about the best translations in steps of 1 mm, and the subsequent fine rotation range was $\pm$ 4 degrees in steps of 1 degree. All computations were performed on a Digital Equipment Corporation AlphaStation. The validity of the above chosen ranges is confirmed by the fact that the fits within these ranges had root-mean-square (RMS) errors of about 2 mm or less and only in very rare instances did the best fit occur at the limit of translation or rotation, and in the majority of those cases, the fit errors at those extrema were less than 2.5 mm, which is considered a very good fit. Some of the statistics regarding the fits are listed in Table 4, below.

Table 4. Results of Fitting the CC- and MLO- view borders with the equations $y = ax^2 + bx^3$ and $y = ax^2 + bx^3 + cx^4$

No. of Images	View	Type of fit	Mean RMS error between fit and run-length averaged automatically tracked border (mm)	No. of best fits that occur at one or more extrema of the translation and rotation ranges (RMS errors (mm) in parentheses)
470	CC	a,b	2.18	0
470	CC	a,b,c	1.07	3 (0.9,1.0,5.9)
484	MLO	a,b	1.53	1 (2.3)
484	MLO	a,b,c	1.07	4 (0.5,0.7,1.4,2.3)

As demonstrated by the data in the fourth column of Table 4, use of 3-parameters (a,b,c) reduced the fitting errors by about 30-50%. An example of a fit for which the RMS error is equal to the mean value for all MLO views using the a and b parameters is displayed in Figure 3, below.

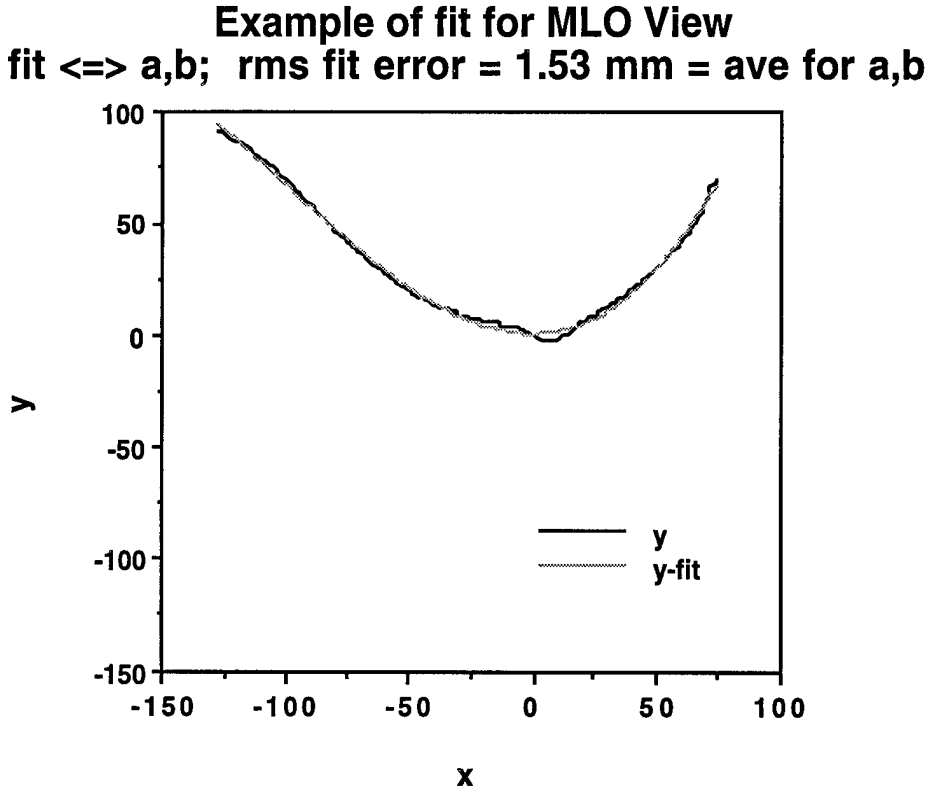

Figure 3. Example of fit for MLO view in which the fitted border (lighter curve) has an RMS error equal to the mean value for all MLO views obtained using the a and b-parameter fitting routine.

Breast Shape Classification

All a,b and a,b,c values for each CC-view border were introduced into the K-Means Cluster Analysis Routine of the SPSS statistical analysis package. This routine uses a centroid sorting method to determine cluster (or class) membership. Once the classification of each border was determined, the mean a and b values or mean a,b, and c-values for the borders within each class were determined. The border defined by the equation using the mean coefficients defined the "filter." A computer routine was written to translate and rotate each of the borders within the class to best match the individual borders with the "filter." This routine utilizes coarse and fine increments for translation and rotation similar to the increments employed in the original fitting routine. The translation and rotation values determined with the original fitting routine were employed as starting values for the matching routine. The mean and standard deviations of the RMS distances between the individual borders and the filters were computed for a quantitative assessment of the success of the border classification. In addition, the newly rotated and translated borders were fit again using the equations $y = ax^2 + bx^3$ and $y = ax^2 + bx^3 + cx^4$. The new a's and b's and a's, b's, and c's were plotted to provide a visual evaluation of the success of classification. The entire procedure was repeated for the MLO borders. Results are listed in Tables 5 and 6.

Table 5. Mean RMS distances (mm) between individual borders and "filters" for **CC** views. (Standard deviations for each distribution are noted in parentheses) . Results in each row are ordered from the smallest mean RMS distance to the largest. n= number of borders in a particular cluster.

(A) a,b fit

	Cluster									
Total no. of clusters	1	2	3	4	5	6	7	8	9	10
4	2.0 (1.1) n=93	2.4 (1.1) n=235	3.5 (4.0) n=9	3.6 (2.0) n=133						
6	2.1 (1.1) n=131	2.5 (1.2) n=231	3.1 (3.3) n=5	3.7 (2.7) n=4	3.7 (2.1) n=95	3.9 (6.1) n=4				
8	1.7 (0.9) n=52	2.0 (0.9) n=119	2.3 (1.2) n=144	3.0 (1.7) n=93	3.1 (3.3) n=5	3.7 (2.7) n=4	3.8 (2.3) n=49	3.9 (6.1) n=4		
10	0.9 (0) n=1	1.7 (0.9) n=52	2.0 (0.9) n=113	2.3 (1.2) n=145	2.8 (1.5) n=89	3.2 (3.1) n=4	3.4 (2.9) n=2	3.8 (2.9) n=2	3.9 (2.3) n=58	3.9 (6.1) n=4

(B) a,b,c fit

	Cluster									
Total no. of clusters	1	2	3	4	5	6	7	8	9	10
4	2.7 (1.8) n=205	2.9 (1.9) n=60	3.2 (2.0) n=199	3.6 (2.2) n=6						
6	1.0 (0.0) n=1	2.4 (1.4) n=34	2.7 (1.7) n=136	2.9 (2.2) n=5	3.0 (2.0) n=204	3.6 (2.4) n=90				
8	1.0 (0.0) n=1	1.3 (0.0) n=1	1.8 (1.0) n=4	2.4 (1.5) n=28	2.7 (1.6) n=109	2.8 (2.0) n=130	3.0 (1.8) n=145	4.0 (2.7) n=52		
10	0.7 (0.0) n=1	1.0 (0.0) n=1	1.3 (0.0) n=1	2.1 (1.0) n=3	2.4 (1.5) n=28	2.7 (1.6) n=103	2.8 (2.0) n=107	2.9 (1.7) n=122	3.1 (2.0) n=72	4.2 (2.8) n=32

Table 6. Mean RMS distances (mm) between individual borders and "filters" for **MLO** views. (Standard deviations for each distribution are noted in parentheses) . Results in each row are ordered from the smallest mean RMS distance to the largest. n= number of borders in a particular cluster.

(A) a,b fit

	Cluster									
Total no. of clusters	1	2	3	4	5	6	7	8	9	10
4	1.7 (1.7) n=107	1.9 (0.8 n=219	2.9 (1.4) n=134	4.1 (1.9) n=24						
6	1.6 (0.7) n=109	1.9 (0.8) n=179	2.1 (0.7) n=18	2.5 (1.3) n=129	3.3 (3.1) n=3	3.7 (1.9) n=46				
8	0.6 (0.0) n=1	0.8 (0.0) n=1	1.6 (0.7) n=142	1.8 (0.8) n=64	2.0 (0.8) n=165	3.0 (1.4) n=87	4.0 (2.0) n=22	4.0 (3.2) n=2		
10	0.6 (0) n=1	0.8 (0.0) n=1	1.5 (0.6) n=81	1.5 (0.0) n=1	1.7 (0.7) n=117	1.9 (0.8) n=150	2.1 (0.7) n=17	2.8 (1.4) n=71	3.3 (1.6) n=35	4.8 (2.3) n=10

(B) a,b,c fit

	Cluster									
Total no. of clusters	1	2	3	4	5	6	7	8	9	10
4	2.9 (2.2) n=229	2.9 (2.2) n=123	4.0 (2.5) n=130	6.4 (1.6) n=2						
6	2.3 (2.1) n=140	2.9 (1.8) n=71	3.3 (2.4) n=147	3.7 (2.3) n=111	4.8 (3.3) n=13	6.4 (1.6) n=2				
6 Hybrid*	1.5 (0.3) n=2	2.1 (1.5) n=138	2.4 (2.4) n=152	2.6 (1.1) n=65	3.3 (3.0) n=106	5.5 (4.8) n=21				
8	0.3 (0.0) n=1	1.3 (0.0) n=1	2.6 (2.0) n=127	2.6 (1.4) n=53	2.7 (2.7 n=76	3.3 (2.3) n=115	3.7 (2.4) n=98	4.8 (3.3) n=13		
10	1.3 (0.0) n=1	2.4 (2.4) n=78	2.8 (2.0) n=114	2.8 (1.0) n=13	2.8 (1.5) n=28	3.1 (0.0) n=1	3.2 (2.0) n=76	3.4 (2.4) n=97	3.5 (3.2) n=7	4.9 (3.4) n=41

*6 Hybrid = use a,b coefficients (c=0) when original fit error ≤ 3.0 mm, use a,b,c coefficients when original fit error > 3.0 mm.

The above results indicate that although the original fits are better using three parameters (a,b, and c) instead of two (a and b), there is either only slight improvement (for the CC case) or no improvement (for the MLO case) in the clustering success as measured by the mean RMS

distances between individual borders within a class and the mean border or "filter". In particular, one will notice a small improvement in these RMS distances for the CC views with the a,b clustering vs. the a,b,c clustering for the 6 and 8 total number of clusters; however, for the same total number of clusters and the MLO views, the a,b,c mean RMS distances are in general worse than their counterparts using a and b. The hybrid idea of using a and b values with c set to zero when the original fit to the automatically traced border is less than or equal to a threshold value and using a, b, and c values otherwise does improve the clustering as shown in Table 6(B); however, for the particular case shown in this table, clustering using only the a and b parameters of the fit equation $y = ax^2 + bx^3$ yielded superior clustering.

Plots depicting the clustering success of the clustering for the MLO views are shown in Figures 4 and 5, below. These plots depict a and b or a, b, and c values of the second stage fits to the borders after they were translated and rotated to best match the cluster "filter". Results for 4 and 8 total clusters are shown. Similar results were obtained for the CC-view.

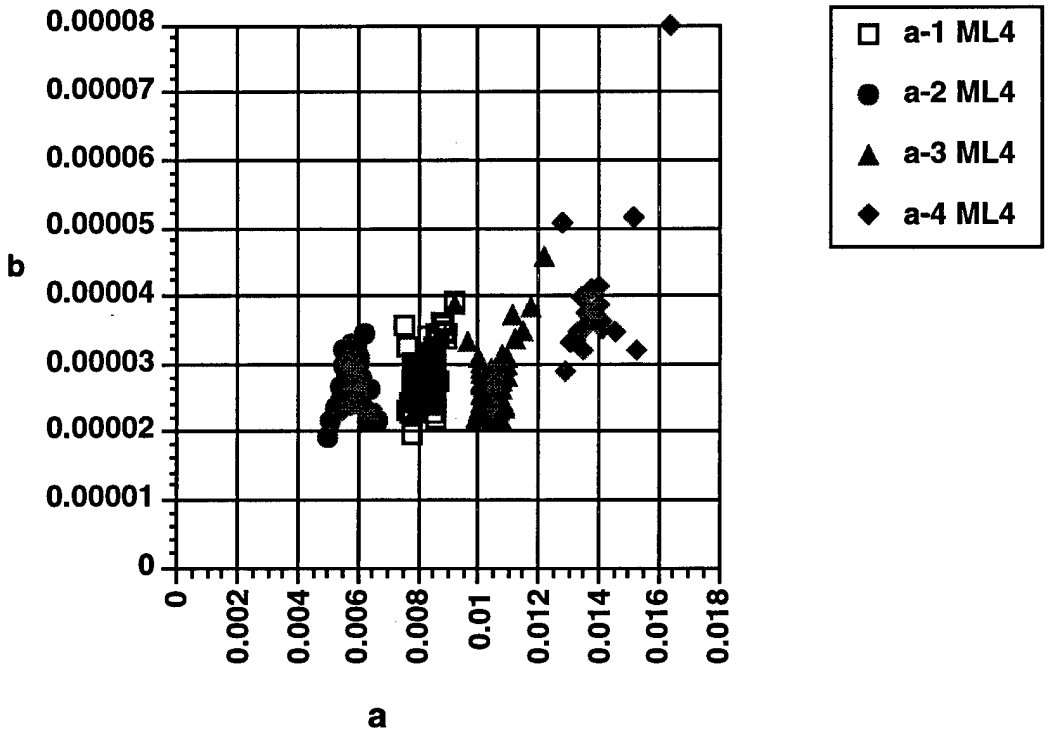

Figure 4(A). 4-group clustering using a and b parameters: 484 MLO views.

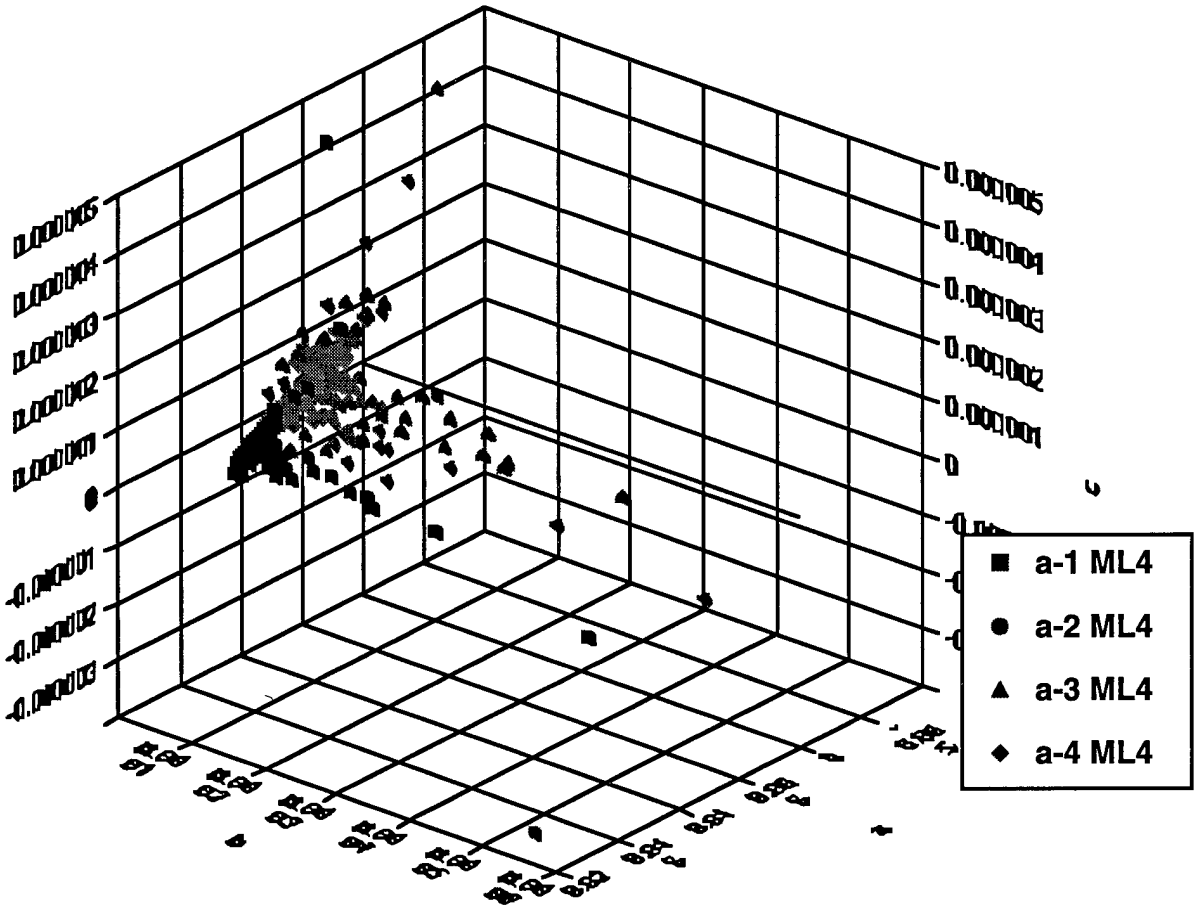

Figure 4(B) 4-group clustering using a b and c parameters: 484 MLO views.

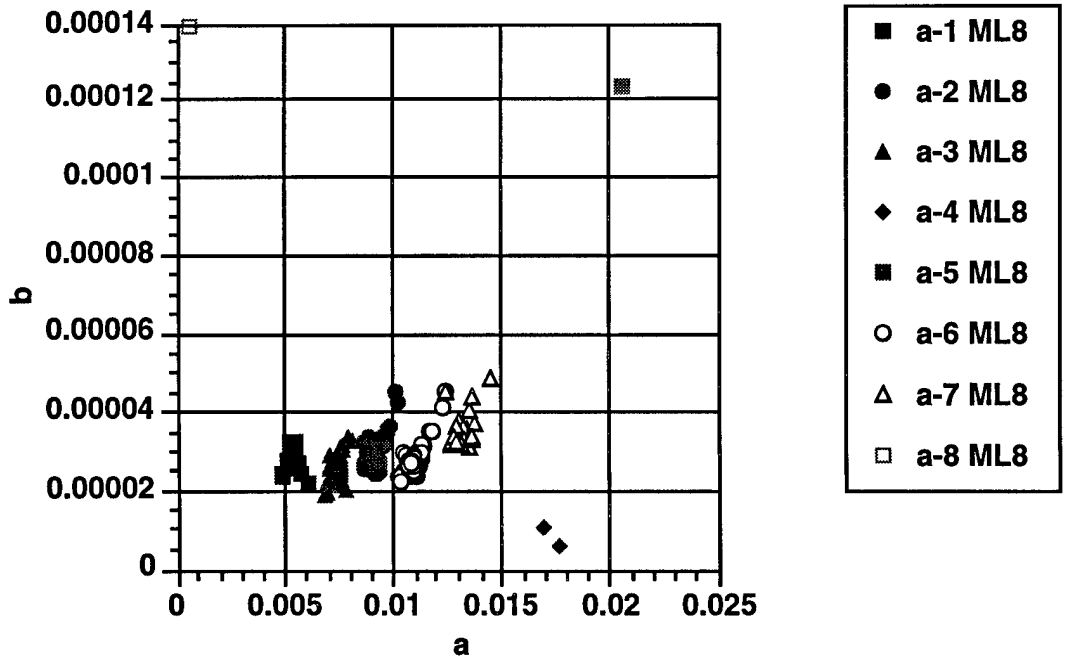

Figure 5(A). 8-group clustering using a and b parameters: 484 MLO views.

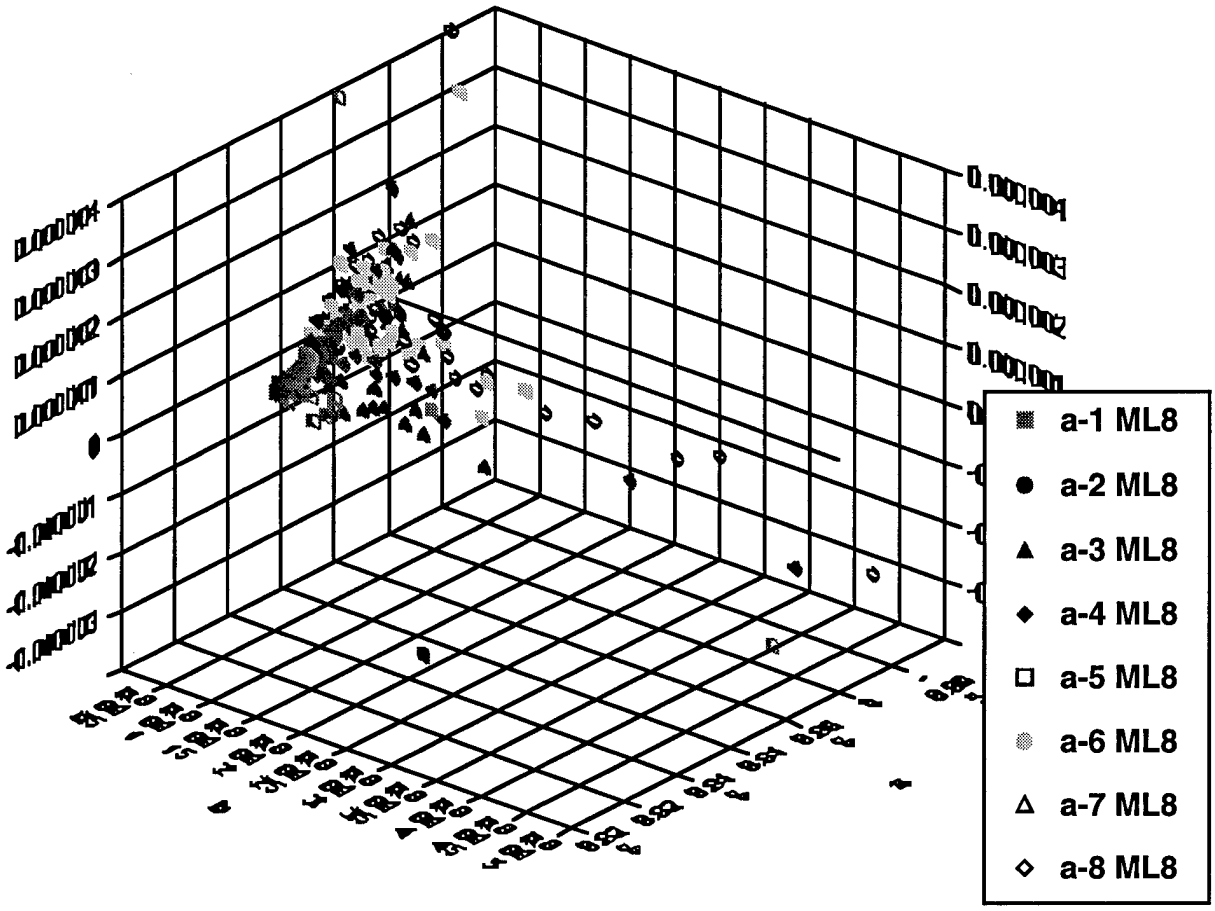

Figure 5(B). 8-group clustering using a b and c parameters: 484 MLO views.

The results of the classification study discussed above were presented at the 1996 annual meeting of the American Association of Physicists in Medicine (Medical Physics 1996;23:1108). Future work in this area will involve evaluation of other clustering methods. For example, we will try a method in which we will select the two best clusters (as determined from the mean and standard deviations of the distances between the individual borders and the "filters") from a 4 cluster analysis, eliminate the corresponding borders and perform a new 2 to 4 cluster analysis.

The significance of this study is that, using a large data base of about 500 mammograms in each view, the results again 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. This is the basis of our approach to exposure equalization in mammographic imaging.

(c) Estimation of Primary X-ray Exposure Distribution in Breast Images for the Design of X-ray Exposure Equalization Filters

The equalization filters reduce the primary x-ray exposure incident upon the breast. The design of the filter requires knowledge of the primary exposure to the detector, which can be

derived from the measured total (primary plus scatter) exposure and the scatter-to-primary ratio. We performed a study to estimate the primary exposure distribution. A set of randomly selected mammograms was digitized and the breast borders detected by the automated edge tracking program discussed above. The breast shapes were classified into a small number of groups. Within each group, the average exposure profiles along radial directions of the breast images were derived from the digital values by using the calibration curve of the film digitizer and the x-ray sensitometric (H and D) curve of the screen-film system. The sensitometric curve was measured using a bootstrap technique in which various thicknesses of Lucite attenuator and various focus-to-film distances were used at the same mAs so as to avoid errors due to reciprocity law failure. To estimate the scatter fraction profile along the radial directions, we used a Monte Carlo code (the MCNP code developed at the Los Alamos National Laboratory) to simulate photon transport through breast phantoms of different thicknesses, shapes, and compositions. An anti-scatter grid was modeled in the Monte Carlo calculation and its effect on the scatter fraction profile could be evaluated. The validity of the Monte Carlo modeling was verified by experimental measurement along a radial direction of a CIRS breast phantom using a beam stop technique. The average primary profiles of the breast images were then derived from the profiles of total exposure and the scatter fraction using the relationship:

Primary Exposure = Total Exposure X (1 - Scatter Fraction).

We performed a preliminary evaluation of the primary exposure profiles in a class of breast shapes. The variation of primary exposure profiles of the breast images in the class was examined. It was found that the variation of the primary exposure profiles within a class was quite large. It was likely that up to three filters of different degrees of equalization would be necessary for the dense, mixed dense and fatty, and fatty breasts or for the thick, medium, and thin breasts in the same breast shape class. We will conduct further studies to determine the acceptable range of variation in the primary exposure profiles for each filter subclass. Once the filter subclass criteria are set and the breast images are grouped into the subclasses, the average primary exposure profile of the breast images in a given filter subclass will be estimated by averaging the primary exposure profiles obtained from the individual mammograms in that subclass. The thickness profile of a filter for this subclass can then be derived for a given filter material.

Some of the results of this study are displayed in Figures 6-11, below.

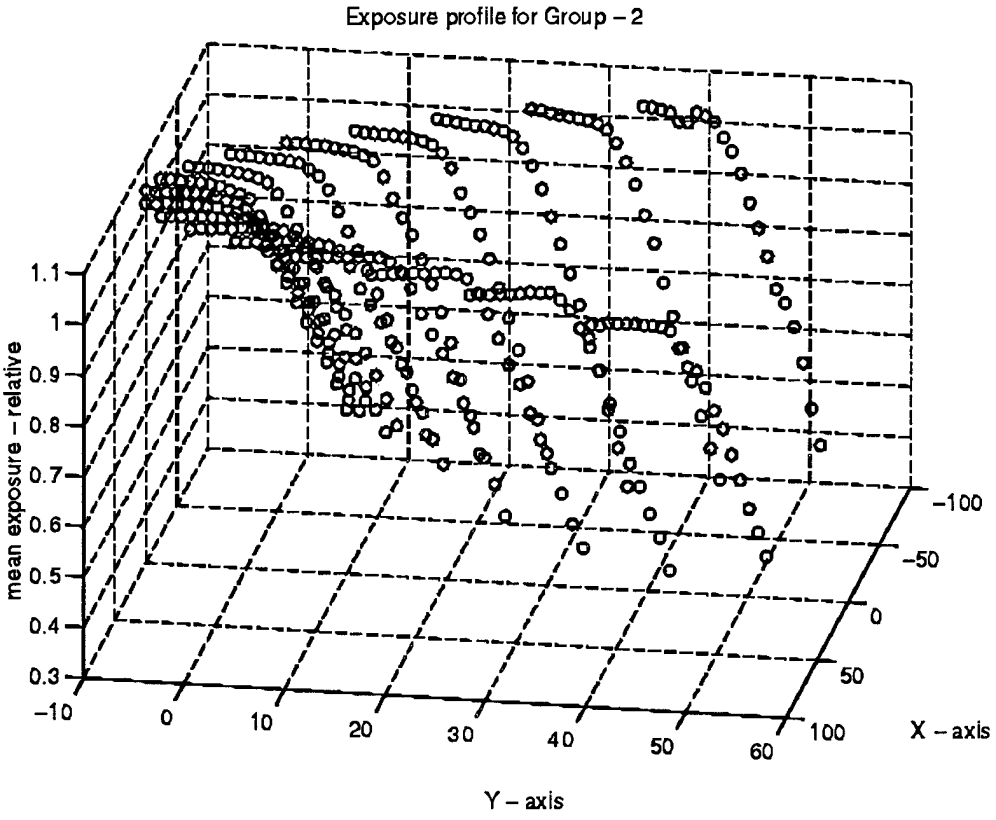

Figure 6. The mammograms are digitized and the breast boundary is detected using the edge tracking program. The breast shapes are classified into a small number of groups. Shown here is the average exposure profile (scattered + primary) of one of the groups.

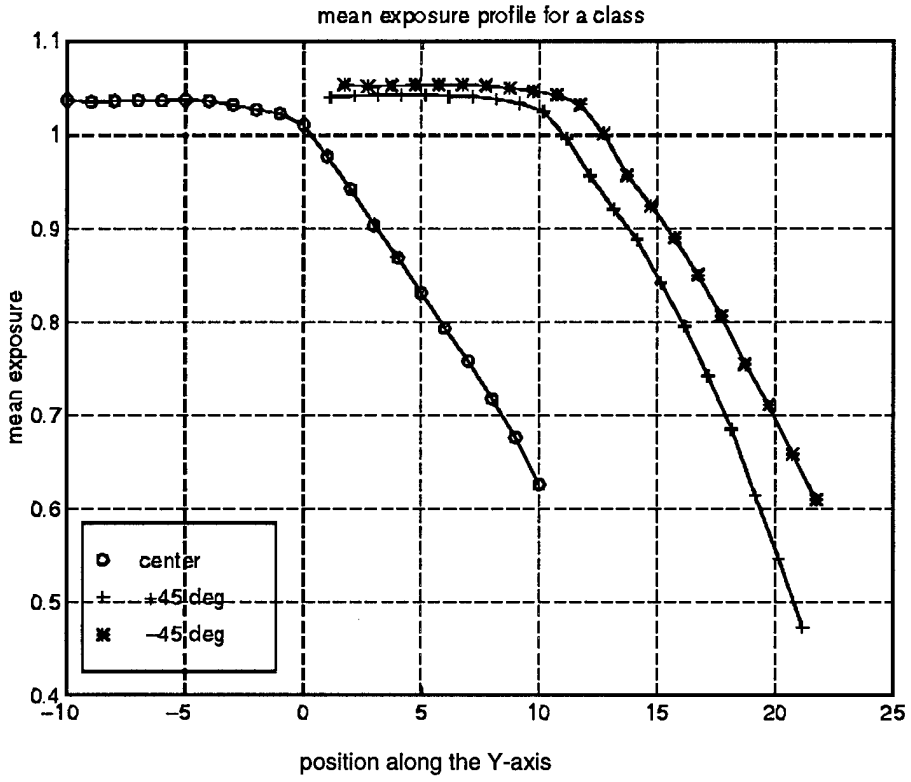

Figure 7. The exposure profile, shown in Figure 6, along the radial and selected directions.

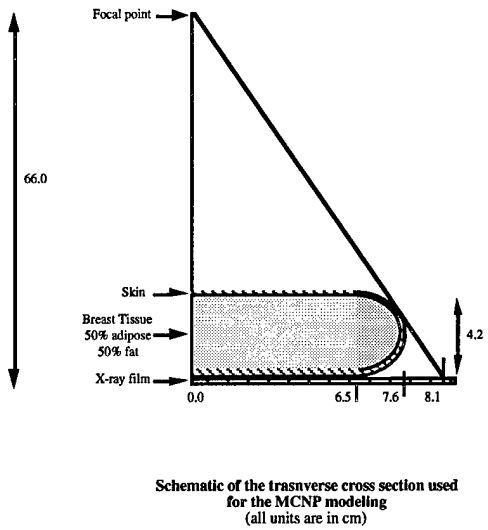

Figure 8. Schematic of the model used for Monte Carlo simulation.

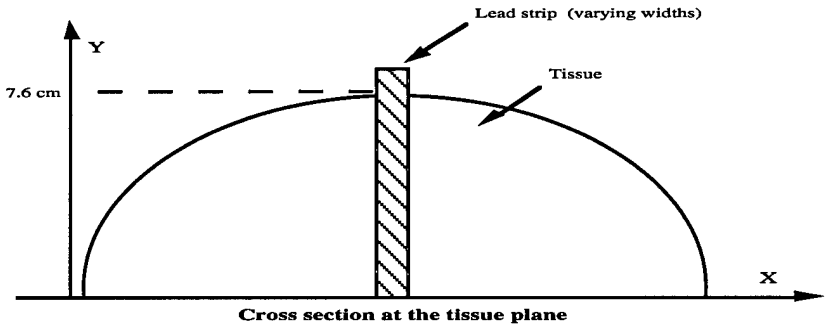

Figure 9. Cross section of the model used, showing the detector locations and the beam stop.

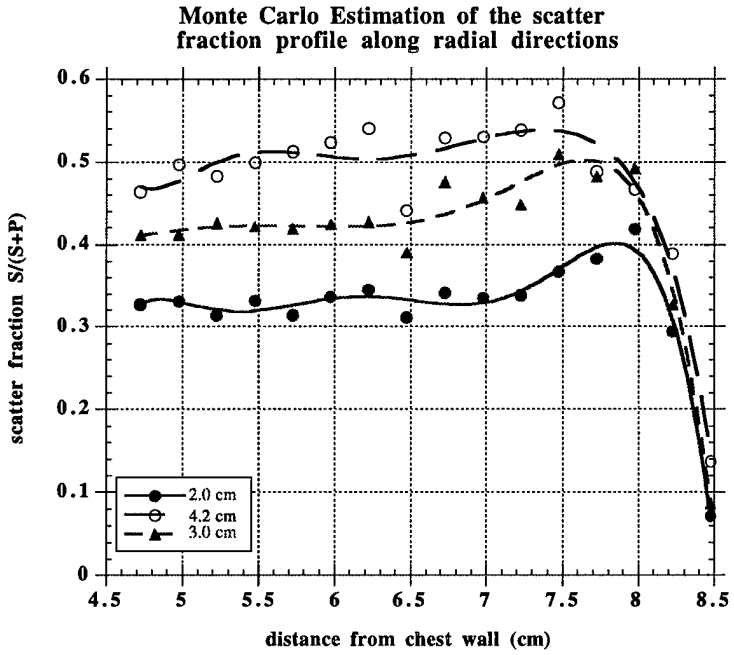

Figure 10. Scatter fractions calculated by Monte Carlo simulation for various tissue thicknesses, imaged at 28 kVp, without the anti-scatter grid.

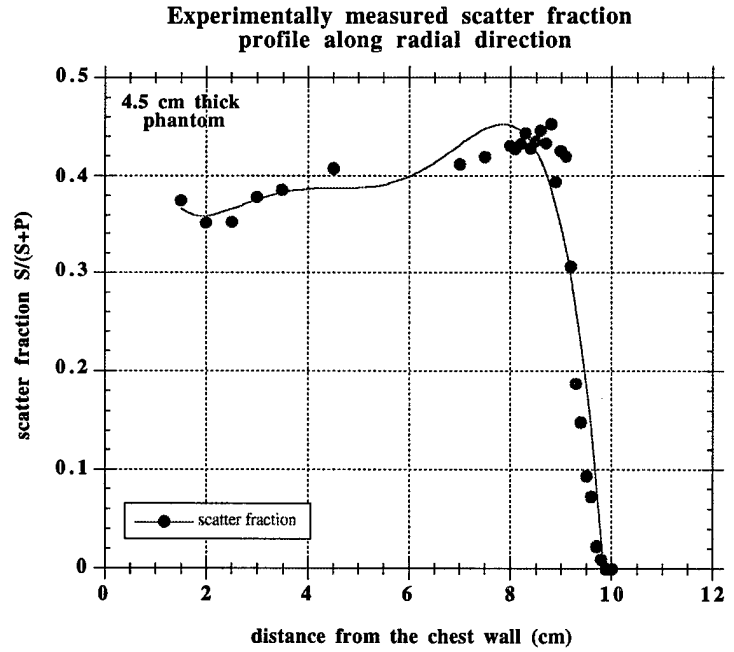

Figure 11. Experimentally measured scatter fractions for a 4.5-cm-thick CIRS phantom, imaged at 28 kVp, without the anti-scatter grid.

(d) Difficulties encountered this year

Although we have made good progress in some areas of the project, we have not accomplished all of the planned studies in this year. One major problem was that our postdoctoral research fellow, whom we recruited to train and assist in this project at the beginning of the second year, was offered an industrial job and left in the middle of the year. Because of the loss of the postdoctoral fellow, some studies including the analysis of the primary exposure profiles and filter design, and evaluation of the effect of equalization on image quality and patient dose by Monte Carlo simulation studies could not proceed as planned. We had begun to recruit a new postdoctoral fellow as soon as we were informed of the resignation. After an extensive search including several interviews, we have found an appropriate person for the position. The new postdoctoral fellow will start in the middle of August. We plan to resume some of the subprojects as soon as the new postdoctoral fellow is trained. Because of this unexpected half-year delay, we may have to request a no-cost extension at the end of the fourth year.

Another problem is that the laboratory space for the experimental mammography system is not ready so that we have not purchased the mammography system planned in the second year of the budget. The University of Michigan Hospital is constructing a Cancer Geriatric Center (CGC) which was originally scheduled to finish in the middle of 1996. We have acquired laboratory space in the new center for installation of the mammography system. It will be an ideal location for a preclinical trial because the space is within the Division of Breast Imaging where patients will visit for routine breast imaging and examinations. The CGC is now expected to finish at the beginning of 1997. The laboratory for the mammography system therefore will not be ready before that time. It is possible to find an off-site location on campus to accommodate the experimental mammography system temporarily. However, we have decided to hold off the purchase of the system until the permanent laboratory is available because of the additional cost involved in installing and moving the x-ray system from one location to another. There is also the risk of damaging the system during the move. This delay in purchasing the mammography system, however, will not delay the progress of our project at this stage. We have access to the clinical mammography systems in the Radiology department in the evenings and weekends. We have been using the clinical systems for experiments such as the measurement of sensitometric curves and scatter fractions, and testing of preliminary filters. We believe that when our filter device and video camera are ready to be installed on the experimental mammography system, the laboratory will be ready for our use.

(7) Conclusion

We have performed the following studies in the second year of the funding period: (1) collected a large data base for analysis of breast shapes and exposure profiles, (2) evaluated the accuracy of the automated breast border detection program by comparing the automatically detected breast borders with manually traced breast borders, (3) evaluated the adequacy of two polynomial models for fitting the breast shapes and classification, (4) developed a clustering procedure for classifying the breast shapes into a small number of classes, (5) developed an automated breast shape classification program to prepare for on-line analysis of patient breast shape. (6) developed a mammographic imaging model for Monte Carlo simulation of photon transport in the imaging system, and (7) evaluated the scatter fraction on the imaging plane experimentally and by Monte Carlo simulation, and studied the primary exposure profiles in a given breast shape class and investigated criteria for subgrouping. The results of these studies will be useful for the implementation of the filter device.

(8) References

1. Shapiro S, Venet W, Strax P, Venet L, Roeser R: Ten-to-fourteen-year effect of screening on breast cancer mortality. JNCI 69:349-355, 1982.
2. Lester RG: The contributions of radiology to the diagnosis, management, and cure of breast cancer. Radiology 151:1-7, 1984.
3. American Cancer Society: Mammography guidelines 1983: Background statements and update of cancer-related checkup guidelines for breast cancer detection in asymptomatic women age 40 to 49. CA Cancer J Clin 33:255, 1983.
4. Moskowitz M: Breast cancer: Age-specific growth rates and screening strategies. Radiology 161:37-41, 1986.
5. Verbeek ALM, Hendriks JHCL, Holland R, Mravunac M, Sturmans F, Day NE: Reduction of breast cancer mortality through mass screening with modern mammography: First results of the Nijmegen project, 1971-1981, Lancet i:1222-1226, 1984.
6. Moskowitz M: Benefit and Risk. In: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging. 2nd edition. Eds. Bassett LW, Gold RH. Grune and Stratton, NY, 1987.
7. Baker LH: Breast Cancer Detection Demonstration Project: A Five-Year Summary Report. CA Cancer J Clin 32:194-225, 1982.
8. American Cancer Society: Cancer Facts & Figures - 1987. Breast Cancer, p.10, 1987.
9. Baines CJ, Miller AB, Wall C, et al: Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers. Radiology 160:295-298, 1986.
10. Pollei SR, Mettler FA, Bartow SA, Moradian G, Moskowitz M: Occult breast cancer: Prevalence and radiographic detectability. Radiology 163:459-462, 1987.
11. Haug PJ, Tocino IM, Clayton PD, Bain TL: Automated management of screening and diagnostic mammography. Radiology 164:747-752, 1987.
12. D'Agincourt L: Technique is everything when breast is dense. Diagnostic Imaging, September: 57-61, 1993.
13. Wallis MG, Walsh MT, Lee JR: A review of false negative mammography in a symptomatic papulation. Clinical Radiology 44: 13-15, 1991.
14. Sickles EA: Mammographic features of "early" breast cancer. AJR 143:461-464, 1984.
15. Sickles EA: Mammographic features of 300 consecutive nonpalpable breast cancers. AJR 146:661-663, 1986.
16. Logan WW, Janus JA: Screen/film mammography. In: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging. 2nd edition. Eds. Bassett LW, Gold RH. Grune & Stratton, NY, 1987.
17. Nishikawa RM, Mawdsley GE, Fenster A, Yaffe MJ: Scanned-projection digital mammography. Med Phys 14:717-727, 1987.
18. Nishikawa RM, Yaffe MJ: An investigation of digital mammographic imaging. Proc SPIE 419:192-200, 1983.
19. Bunch PC, Huff KE, Van Metter R: Analysis of the detective quantum efficiency of a radiographic screen/film system. J Opt Soc Am A 4:902-909, 1987.
20. Blackwell HR: Contrast thresholds of the human eye. J Opt Soc Am 36:624-643, 1946.
21. Baxter B, Ravindra H, Normann RA: Changes in lesion detectability caused by light adaptation in retinal photo-receptors. Invest Radiol 17:394-401, 1982.
22. Snyder HL: Chapter 3: The Visual System: Capabilities and Limitations. In: *Flat-Panel Display and CRTs*. Ed. Tannas LE Jr, Van Nostrand Reinhold, New York, 1985.
23. SPSS: *SPSS for Windows Release 6 Professional Statistics* (Chicago, IL: SPSS inc.)

(9) Appendix

Publications in current year as a result of this grant

1. Liu B, Goodsitt MM, Chan HP. Normalized average glandular dose in magnification mammography. Radiology. 1995; 197:27-32.

2. Goodsitt MM, Chan HP, Liu B, Guru S, Morton R. Compressed breast shape classification for the design of mammography equalization filters. Presented at the 38th Annual Meeting of the American Association of Physicists in Medicine. Philadelphia, Pennsylvania, MA, July 21-25, 1996. Medical Physics 1996; 23: 1108 (Abstract).

3. Morton AR, Chan HP, Goodsitt MM. Automated model-guided breast segmentation algorithm. Presented at the 38th Annual Meeting of the American Association of Physicists in Medicine. Philadelphia, Pennsylvania, MA, July 21-25, 1996. Medical Physics 1996; 23: 1107-1108 (Abstract).