AD

AD-A279 866

GRANT NO: DAMD17-93-J-3007

TITLE: IMPLEMENTATION OF COMPUTER ASSISTED BREAST CANCER DIAGNOSIS

PRINCIPAL INVESTIGATOR: Shih-Chung B. Lo, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University
37th & O Streets, NW
Washington, DC 20057

REPORT DATE: January 3, 1994

DTIC
ELECTE
JUN 01 1994
S B D

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research, Development, Acquisition, and Logistics Command, (Provisional), Fort Detrick, Frederick, 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.

94-16228

94 5 31 064

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	3 January 1994	Annual Report (12/1/92 - 11/30/93

4. TITLE AND SUBTITLE
Implementation of Computer Assisted Breast Cancer Diagnosis

5. FUNDING NUMBERS
Grant No. DAMD17-93-J-3007

6. AUTHOR(S)
Shih-Chung B. Lo, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Georgetown University
37th & O Streets, NW
Washington, DC 20057

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research, Development, Acquisition, and Logistics Command (Provisional)
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 words)*

This project aims at the implementation of a computer-aided diagnosis system for the detection of microcalcifications on mammograms based on the algorithms developed by the principal investigator and others. In addition, proposed research includes: (1) algorithm improvement for the detection of microcalcifications, (2) mammographic image compression and its impact on computer-aided diagnosis (CADx), and (3) computer-aided classification of benign and malignant masses on mammograms.

In the past year, we have developed several algorithms and have studied part of the proposed research: (a) development of filtering techniques with wavelet transform to reduce mammographic structures other than microcalcifications, (b) performance of preliminary study in the detection of microcalcifications, (c) performance of mammographic compression studies using split gray values and full-frame DCT techniques, (d) evaluation of the impact of the compression with respect to various degrees of data compression, and (f) partial implementation of CADx system in a DEC Alpha workstation (user interface and some image functions).

14. SUBJECT TERMS
Breast Cancer, Diagnosis, Computer

15. NUMBER OF PAGES

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.

____ 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).

____ 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.

Accession For

NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

[signature] 1/31/94

PI - Signature Date

The First Annual Report for Project, Entitled:

Implementation of Computer Assisted Breast Cancer Diagnosis

(US Army Grant No. DAMD17-93-J-3007)

TABLE OF CONTENTS

The First Annual Report for Project, Titled:

Implementation of Computer Assisted Breast Cancer Diagnosis

(US Army Grant No. DAMD17-93-J-3007)

1. Introduction

Many investigators have attempted to analyze mammographic abnormalities. Recently, several investigators have proposed various methods for the automatic detection of microcalcifications and masses on mammograms. Vast improvements in accuracy have been made since the initial attempt [Chan 1987; 1988] to apply the computer algorithms for the detection of microcalcifications. We believe that it is important to implement the program into a high speed workstation and conduct a large scale preclinical trial in order to evaluate its clinical practicability and limitations. The false-positive rate for the detection of masses is still very high. On the other hand, we believe that the creation of a computer program to analyze features of suspected masses will give rise to a more useful and fundamental approach to computer-aided diagnosis. We propose to use an artificial neural network to classify malignant and benign masses.

With the use of high-resolution digital mammograms, data compression is an important means to facilitate the mammographic image transmission and storage. We have studied some characteristics of the mammograms using gray value splitting and full-frame discrete cosine transform (DCT) methods. Effects of applying the data compression to the proposed computer aided diagnosis (CADx) scheme in the detection of microcalcifications were also tested in our preliminary evaluation.

2. Refinement of Image Processing Prior to the Detection of Microcalcifications

At Georgetown, we have developed a band-passed filter based on the wavelet transform domain. Since the wavelet transform can decompose both frequency and local spatial information into its transform domain, some breast tissue structures (e.g., vessels and ducts, etc.) can be easily extracted in the transform domain.

In the wavelet transform, many line structures were extracted according to three different high frequency regions, namely: horizontal, vertical, and diagonal. In these regions, one can easily detect the line and band structures using modified Hough Transform. Removed lines and bands are compensated by relaxation algorithms in the original image. Once these background structures are reduced, microcalcifications can be extracted much more accurately by the following CADx detection procedure.

3. Refinement of CADx Algorithm in the Detection of Microcalcifications

We also spent a great deal of effort towards the improvement of the original CADx program. The microcalcification searching algorithm, previously bottlenecked , has been greatly improved. The new program uses "the chain algorithm" to search for the boundary of each island based on a given threshold. The suspected microcalcifications, which are "islands" in a large image, are tested by histogram thresholding and root-mean-square variation methods.

4. Mammographic Image Compression.

Based on the splitting method which was employed to reduce edge effects and to obtain maximum compression efficiency [Lo 1991], we have refined our compression method using alternate value contour coding and full-frame entropy encoding. Trade-off studies between irrerversible and error-free compression were also tested on selected chest and mammographic images [Lo 1993]. We have tested 15 digitized mammograms based on a refined compression scheme. The results are reported below:

4.1. Error-Free Compression for Images Containing Most Significant Values

We have developed an efficient compression method called "alternate value contour coding" for the step-type image. The most significant value images containing the 3 most significant bits (3MSB) of digital radiographs belong to this type of image. We have tested the newly developed compression method on chest radiographs. This method performed much better for error-free compression than DPCM/run-zero/arithmetic coding proposed earlier due to the method of turning the entire 2-D image data into a l-D edge tracking sequence. In addition, contouring for the adjacent values is ignored and the image data is fully recoverable. The only drawback of the alternate value contour coding is that the algorithm is somewhat complicated and demands error checking procedures to ensure the error-free requirement is fulfilled.

4.2. Compression for Images Containing Least Significant Value

Based on full-frame entropy encoding (FFEC), the remapped 9 least significant value (R9LSB) images were decomposed by 2-D full-frame DCT followed by a quantization procedure and an entropy coding (arithmetic encoding) as indicated in our proposal. Preliminary results obtained from the studies using the proposed methods are:

4.2.1. Quantization - we have evaluated the density distribution of DCT coefficients for several chest R9LSV images. We found that the distribution density of R9LSV image can be modelled by a Gaussian function. However, the standard deviation of the Gaussian distribution tends to be large which makes non-linear quantization [Max 1960, Modestino 1985] less useful. The reason for a broad Gaussian distribution is that the low bit data contains a much lower signal to noise ratio. It is relatively difficult to quantize and to encode noise dominated images both in the spatial and in the frequency domain. Our initial results indicated that the advantage of using a non-linear quantizer over a linear quantizer is very small (about 5-8%) for R9LSV images.

4.2.2 Noise Evaluation and Coding - As far as RLSV images are concerned, the FFEC [Lo 1991-SPIE] is the primary algorithm. We found that digitized radiographs contained not only white noise but also system structure noises (e.g., system electronic and mechanical noises and dust on the computed radiographic plate or lens). Although the splitting method has partially solved problems causing by sharp edges, spots and shallow lines are the main structures for encoding. We therefore spent some time overseeing image quality and consulting with vendors to adjust our computed radiographic and laser film digitizer systems. However, we did not succeed in overcoming all the structure noises, particularly in film digitizers. We have evaluated images using a step wedge and found that a minimum of 4 of the least significant bits out of 12-bit values are noise. These results are confirmed by both signal to noise (S/N) and covariance studies with single displacement. Among the sampled gray spectra, the maximum S/N is 120. By removing up to 4 least significant bit data through round-off, the maximum covariance is less than 0.07. However, the test of covariance is drastically increased to 0.2 with the 5 least significant bit data. These results indicated that the image data contained only about 8-bit information. Based on these noise characteristics, we can limit our frequency quantization corresponding to gray value variance (e.g., 15).

4.3. Image Compression in Digital Mammography and Its Effects on Computerized Detection of Subtle Microcalcifications

Our previous receiver operating characteristic (ROC) study indicated that the detection accuracy of microcalcifications by radiologists is significantly reduced if mammograms are digitized at 0.1mm. Our recent study also showed that detection accuracy by computer decreases as the pixel size increases from 0.035mm. Clearly, the digitization of mammograms requires very large matrix sizes. Efficient compression techniques will be needed to facilitate communication and archiving of digital mammograms.

In this study, mammograms were digitized with a laser scanner at a pixel size of 0.035 mm and 12 bits. We studied two compression techniques: (a) full frame DCT coding with entropy coding and splitting of bits, and (b) Laplacian pyramid hierarchical coding (LPHC) with linear requantization. The effectiveness of the techniques is compared in terms of the bit rate, the mean-square-error, the visual quality of reconstructed and error images, and the detection of microcalcifications by computer.

With LPHC, significant degradation of detection accuracy was observed when the compression ratio was greater than 3.6:1. The DCT technique provided a higher compression efficiency at comparable detection accuracy. A compression ratio of 9.6:1 was achieved without significant degradation in the detection of microcalcifications. Furthermore, it was found that the mean-square error was not a good indicator for the evaluation of information loss due to image compression.

In summary, our study showed that there is a trade-off between reconstructed image quality and compression efficiency. Further investigation is needed for selection of optimal compression technique for digital mammography.

5. Implementation of CADx for the Detection of Clustered Microcalcifications

We have started to implement the CADx program into a DEC Alpha workstation which is currently the fastest workstation on the market. The basic user interface is nearly complete. However, it requires some final modifications. The user interface can select a mammogram and display it on the workstation. Several basic image functions are also available: (1) "window and level" for the adjustment of the brightness and contrast, (2) pan, and (3) a cursor box for the user to select the area of interest.

6. Contractual (SOW) Issues

We have not completed our research and implementation of CADx workstation. At this point, we have contacted Dr. R.V. Shah, Chief Breast Radiologist, at Brook Army Medical Center and Dr. Don Smith, attendant breast radiologist, at Madigan Army Medical Center. They have agreed to provide their proven cases associated with mammographic microcalcifications for inclusion in our test database [Private Communication]. We will provide our software for the evaluation at Army Hosipitals after the end of this project.

7. Conclusion of the Annual Report

At this point, we have refined our CADx algorithms in image preprocessing, detection effectiveness, and computer speed. We have also done studies in mammographic image compression. More studies on the impact of compression on the CADx are in progress. Database collection is ongoing and will continue up to the final stage of this project. Several basic functions and user interface have been implemented in the workstation.

During the next year, we are going to convert and test some of our FORTRAN codes to C computer language. It will take great engineering efforts to merge our newly developed algorithm in C and useful old codes developed by Dr. Chan and her colleagues. We will spend most of our research time in the evaluation of the effect of digitization on accuracy of CADx using the proposed computer scheme.

References

Chan HP, Doi K, Galhotra S, Vybomy CJ, MacMahon H, Jokich PM: Image Feature Analysis and Computer-Aided Diagnosis in Digital Radiography. 1. Automated Detection of Microcalcifications in Mammography. Med Phys 1987;14:538.

Chan HP, Doi K, Vybomy CJ, Lam KL, Schmidt RA: Computer-aided Detection of Microcalcifications in Mammograms: Methodology and preliminary clinical study. Invest Radiol 1988;23:664.

Lo, SC, Shen, E, Mun, SK, and Chen, J: A Method for Splitting Digital Value in Radiological Image Compression. Medical Physics Journal, 18(5), 1991, pp. 939-946.

Lo, SC, Krasner, BH, Mun, SK, and Horii, SC: Full-Frame Entropy Encoding for Radiological Image Compression. SPIE Proc. Med. Imaging V, Vol. 1444, 1991. pp. 265-277.

Lo, SC, Krasner, BH, and Mun, SK: Gain of Irreversible over Error-Free Compression in Radiological Imaging. SPIE Proc. Med. Imag. VII, vol. 1897, 1993 , pp. 292-297.

Max, J: Quantization for Minimum Distortion. IEEE Trans. Inf. Th., IT-16, 1960, pp. 7-12.

Modestino, JW, Farvardin, N, and Ogrinc, MA: "Performance of Block Cosine Image Coding with Adaptive Quantization," IEEE Trans. on Comm. Vol. COM-33, No. 3, March 1985, pp. 210-217.

Presentations and Publications:

1. Chan HP, Niklason LT, Lo SC, Ikeda DM, Helvie MA, Adler DD, Cheng SNC: Image *Compression in* Digital Mammography and Its Effects on Computerized Detection of Subtle microcalcifications. 78th Scientific Assembly and Annual Meeting of the Radiological Society of North America, November 29 - December 4, 1992. Radiology 1992; 185(P): 254.

2. Chan HP, Lo SCB, Helvie MA, Goodsitt MM, Cheng SNC, Adler DD. Recognition of mammographic microcalcifications with artificial neural network. Accepted for presentation at the 79th Scientific Assembly and Annual Meeting of the Radiological Society of North America. November 28-December 3, 1993, Chicago, IL. Radiology 1993; 189(P): 318.

3. Lo S-C B, Chan HP, Freedman MT, Mun SK. Detection of Clustered Microcalcifications Using Two-dimensional Convolution Neural Network. 35th Annual Meeting of the American Association of Physicists in Medicine. Washington, D.C., August 8-12, 1993. Medical Physics 1993; 20: 881.

4. Lo S-C B., Y.Wu, A. Hasegawa, M.T. Freedman,a nd S.K. Mun: Fuzzy Neural Network and Fuzzy System Modeling for General Medical Image Pattern Recognition. 79th Scientific Assembly and Annual Meeting of the Radiological Society of North America, November 28 - December 3, 1993. Radiology 1993; 189(P): 218.

5. Lo, SC, Krasner, BH, and Mun, SK: Gain of Irreversible over Error-Free Compression in Radiological Imaging. SPIE Proc. Med. Imag. VII, vol. 1897, 1993 , pp. 292-297.

6. Petrosian A, Chan HP, Helvie MA, Goodsitt MM, Adler DD. Computer-aided Diagnosis in Mammography: Detection of masses by texture analysis. 35th *Annual Meeting of the* American Association of Physicists in Medicine. Washington, D.C., August 8-12, 1993. Medical Physics 1993; 20: 880.

7. Cheng SNC, Chan HP, Helvie MA, Goodsitt MM, Adler DD, St. Clair DC. Classification of Mass and Mon-mass Regions on Mammograms Using Artificial Neural Network. Medical Physics (submitted).

8. Petrosian A, Chan HP, Helvie MA, Goodsitt MM, Adler DD. Computer-Aided Diagnosis in Mammography: Classification of Masses and Normal Tissue by Texture Analysis. Physics in Medicine and Biology (submitted).