

AD \_\_\_\_\_

GRANT NO: DAMD17-94-J-4196

TITLE: Digital Mammographic Image Compression

PRINCIPAL INVESTIGATOR(S): Jun Wang  
Doctor H. K. Huang



CONTRACTING ORGANIZATION: University of California, San Francisco  
San Francisco, California 94143-0962

REPORT DATE: July 15, 1995

TYPE OF REPORT: Annual

19951018 169

PREPARED FOR: Commander  
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.

DTIC QUALITY INSPECTED 5

# 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<br>15 Jul 95                 | 3. REPORT TYPE AND DATES COVERED<br>Annual 15 Jul 94 - 14 Jul 95 |  |
| 4. TITLE AND SUBTITLE<br><br>Digital Mammographic Image Compression                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                             |                                             | 5. FUNDING NUMBERS<br><br>DAMD17-94-J-4196                       |  |
| 6. AUTHOR(S)<br><br>Ms. Jun Wang<br>Dr. H. K. Huang                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                             |                                             |                                                                  |  |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>University of California, San Francisco<br>San Francisco, California 94143-0962                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                             |                                             | 8. PERFORMING ORGANIZATION<br>REPORT NUMBER                      |  |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>U.S. Army Medical Research and Materiel Command<br>Fort Detrick, Maryland 21702-5012                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                             |                                             | 10. SPONSORING/MONITORING<br>AGENCY REPORT NUMBER                |  |
| 11. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                             |                                             |                                                                  |  |
| 12a. DISTRIBUTION / AVAILABILITY STATEMENT<br><br>Approved for public release; distribution unlimited                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                             |                                             | 12b. DISTRIBUTION CODE                                           |  |
| 13. ABSTRACT (Maximum 200 words)<br><br>This research investigated both lossless and lossy compression methods for digital mammography. A structure lossless compression algorithm was developed for digital mammograms. The algorithm utilizes the unique shape and image characteristics of mammograms. The algorithm first segments a breast image from its background. An averaging predict coding was then used to compress the segmented image. The average bit rate compression ratio achieved is about 3.4 for 100 mammograms. The results also show that high resolution images can be compressed more than low resolution images.<br><br>In the lossy compression, a wavelet compression scheme was developed. The preliminary results show that the wavelet transform is able to concentrate pixel energy into a very small area. This characteristic is promising for obtaining high quality compressed images at reasonable compression ratios. |                                             |                                             |                                                                  |  |
| 14. SUBJECT TERMS<br>Digital mammograms, Structure lossless compression,<br>Segmentation, Predict coding, Wavelet transformation.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                             |                                             | 15. NUMBER OF PAGES<br>23                                        |  |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                             |                                             | 16. PRICE CODE                                                   |  |
| 17. SECURITY CLASSIFICATION<br>Unclassified                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 18. SECURITY CLASSIFICATION<br>Unclassified | 19. SECURITY CLASSIFICATION<br>Unclassified | 20. LIMITATION OF ABSTRACT<br>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.

Wayne  
PI - Signature

8/9/95  
Date

## Table of Contents

|                                                           |    |
|-----------------------------------------------------------|----|
| 1. Introduction .....                                     | 1  |
| 1.1 Digital Mammography .....                             | 1  |
| 1.2 Image Compression.....                                | 2  |
| 1.3 The Goal Of The Present Work.....                     | 3  |
| 1.4 The Approach.....                                     | 3  |
| 2. Methods.....                                           | 4  |
| 2.1 Digitization .....                                    | 4  |
| 2.2 Orientation and Segmentation.....                     | 4  |
| 2.3 Prediction Compression .....                          | 7  |
| 2.4 Lossy Compression Method With Wavelet Transform ..... | 8  |
| 2.4.1. Wavelet Transform.....                             | 8  |
| 2.4.2. Quantization and Entropy Coding .....              | 10 |
| 3. Results .....                                          | 12 |
| 3.1 Structure Lossless Compression .....                  | 12 |
| 3.2 Wavelet Compression.....                              | 13 |
| 4. Conclusions .....                                      | 15 |
| 5. References.....                                        | 16 |
| Appendix A: Bibliography.....                             | 18 |
| Appendix B: Personnel Receiving Pay.....                  | 19 |

|                      |                                     |
|----------------------|-------------------------------------|
| Accession For        |                                     |
| NTIS                 | <input checked="" type="checkbox"/> |
| CRA&I                | <input type="checkbox"/>            |
| DTIC                 | <input type="checkbox"/>            |
| TAB                  | <input type="checkbox"/>            |
| Unannounced          | <input type="checkbox"/>            |
| Justification .....  |                                     |
| By .....             |                                     |
| Distribution / ..... |                                     |
| Availability Codes   |                                     |
| Dist                 | Avail and/or Special                |
| A-1                  |                                     |

# **Digital Mammographic Image Compression: Structure Lossless Mammogram Compression and Preliminary Lossy Compression By Wavelet Transformation**

Jun Wang

Dr. H.K. Huang

## **1. Introduction**

### **1.1 Digital Mammography**

Breast cancer is the second leading cause of death from cancer among women in the United States <sup>1</sup>. One in every nine women will have breast cancer at some point in her life time. Studies show that early breast cancer detection can reduce the mortality <sup>2</sup>. Routine screen-film X-ray screening has been recognized as the best means to detect early cancer at the present <sup>3,4,5</sup>. Although conventional screen-film mammography has been shown to be very effective, digital mammography has emerged as a new technology that will overcome some of the limitations of the screen-film mammography and offers other additional benefits <sup>6,7,8</sup>.

Digital mammography offers the following advantages. First, screen-film mammographic systems produce poor contrast images for dense breasts because the glandular tissues and lesions are essentially the same density. Digital detectors have wider dynamic range and images can be post-processed to offer even more contrast. This makes direct digital mammography an attractive way to image dense breasts. Second, digital image processing techniques allow the change of the image contrast, enhance lesion detection, and optimize image display. Such techniques thus provide better image quality to support higher diagnostic accuracy. Third, computer-aided diagnosis (CAD) involves a computerized analysis of radiographs that can be used as a "second opinion" by the radiologist. For mammography, CAD offers many practical schemes to detect the masses and clustered microcalcifications on mammograms <sup>9</sup>. Fourth, telemammography provides a fast way to access experts over long distance for consultation <sup>10</sup>.

There are two ways to generate a digital mammogram: a) secondary digitization, in which conventional films are digitized; and b) acquisition of primary digital images. Previous studies showed that at least 100  $\mu\text{m}$  sampling distance is required in order to achieve an adequate resolution for digital mammography<sup>11,12</sup>. Digitization creates 9 Mbytes of data at the resolution of 100 $\mu\text{m}$  for a 7'X9' film, and 36 Mbytes of data at a resolution of 50  $\mu\text{m}$ . Storage cost and the transmission cost of such large images will be a problem as more digital mammograms are created. Image compression reduces the data storage while preserving the useful information. It provides a solution to handling such large digital images efficiently.

## 1.2 Image Compression

Image compression is a technique that reduces the number of bytes needed to store an image. As a result it also reduces the transmission time of the image. The goal of image compression is to use a minimum number of bytes to store an image while preserving the original image quality. There are two types of image compression: lossy and lossless. Lossless compression can achieve a compression ratio of around 2 or 3 for medical images, and will allow exact reconstruction of the original image. Lossy compression can achieve much higher compression ratios, although the original image can not be exactly reconstructed.

Various lossless and lossy image compression techniques have been developed in the past years<sup>13-17</sup>. Most of them, however, are for engineering applications, and do not consider the stringent requirements of image size and quality in medicine. Although several investigators have studied compression techniques for medical images<sup>18-22</sup>, few such techniques concern mammograms. Kuduvalli investigated performance of several reversible image compression techniques for high resolution teleradiology<sup>23</sup>. The compression schemes implemented are: Huffman coding, Arithmetic Coding, Lempel-Ziv coding, two-dimensional linear predictive coding, and transform coding. These coding techniques are applied to a few mammograms and chest radiographs digitized up to 4k X 4k, 10 bits/pixel. The results show that multichannel linear prediction gave the best results at an average pixel rate of 2.83 bits/pixel for five mammograms.

Benoit-Cattin applied a wavelet transform to mammograms<sup>24</sup>. Then JPEG was applied to code only the low resolution subimage. They found that the hybrid method achieves a lower bit rate than solely the JPEG method at the same SNR. Lucier evaluated a Harr wavelet transform coding on mammograms<sup>25</sup>. Images with clustered microcalcifications were compressed at two different ratios, and the image quality was evaluated. The results show that wavelet compression could be used to achieve a high compression rate of mammograms without losing small details such as microcalcifications.

Mammograms have unique characteristics that previous work did not consider: a) breast images have a regular round shape, the background area is relatively large in a rectangular film; b) breasts contain mostly soft tissue, so image contrast is less compared with other types of images. In this work, we develop a mammogram compression method that takes advantage of these characteristics.

### **1.3 The Goal Of The Present Work**

The purposes of this research are:

- 1) Develop a structure lossless compression technique for mammograms by incorporating the breast imaging shape and contrast characteristics. The algorithm developed combines segmentation with a lossless compression scheme to achieve higher compression ratio than those of other types of lossless compression methods.
- 2) Implement a wavelet lossy compression method for mammograms.

### **1.4 The Approach**

Structure lossless compression has three steps. The first step identifies the orientation of the image, i.e. which side of the image that the chest wall is located. The second step segments the breast image from its background. The third step compresses the portion of the image that contains the breast tissue with the lossless compression technique.

## 2. Method

### 2.1 Digitization

We selected 14 cases from the UCSF mammography teaching file. A total of 50 films were digitized at the resolution of 100  $\mu\text{m}$  and 200  $\mu\text{m}$ . A resolution of 200  $\mu\text{m}$  may not be good enough for clinical purposes. It is so chosen simply because images digitized at this resolution have small sizes and are convenient for algorithm development.

A sheet of film contains an image of a breast, a label indicating screening position, and possibly patient information. Film size is normally 7 X 9 inches or 10 X 12 inches. The size of a breast image in a film may vary a great deal. An image may only occupy a small portion of the film. We only digitized a rectangular area that contains the breast image rather than whole film. The patient information and position information, which are located on the edge of the film, are often not digitized. This information can be input as a text file from the image header.

The average image size is 1.3 Mbytes at 200  $\mu\text{m}$ , and 6 Mbytes at 100  $\mu\text{m}$ .

### 2.2 Orientation and Segmentation

The first step of the compression process is to find the orientation of the image, i.e. which side of the image is the chest wall. Mammograms normally have a Cranial-Caudal (CC) view and a Medial Lateral Oblique (MLO) view shown in Figure 1. The background is normally dark and the portion with the breast tissue is normally gray. The chest wall can be either on the left side or the right side of the image, depending on the way that the film is digitized.

To determine the orientation of the image, ten reference points along the left and the right edges of the image were selected as shown in Figure 2. The x positions of the left and right reference points are located on 5% and 95% of total width of the image, respectively. The y positions of the five reference points are located on 10% to 90% of total length at a 20% increment. The average gray level of each reference point is calculated by averaging the surrounding 3 X 3 pixels. If a reference point is located inside the breast, the gray level is higher than if a reference point is

located in the background. Comparing the gray level of the reference points on the left and right side, we can determine the orientation of the breast image and the background gray level.

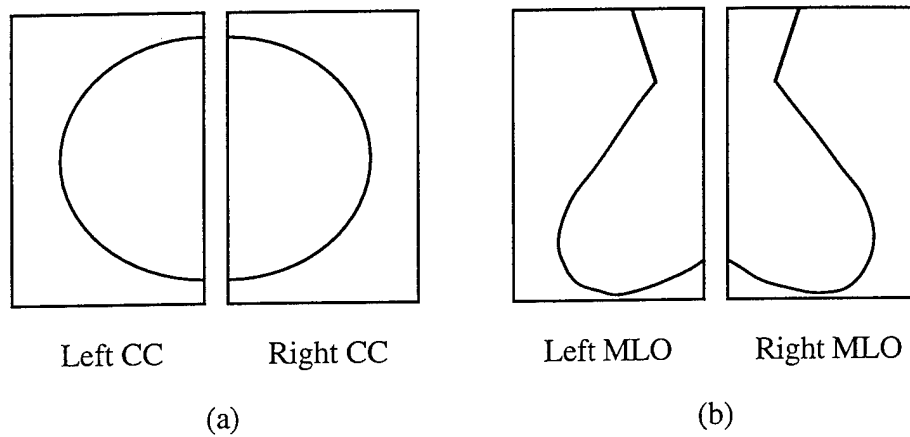


Figure 1. Two different views of mammograms (a) CC view (b) MLO view

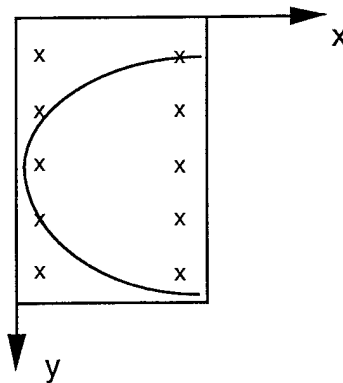


Figure 2. The positions of the reference points

The next step of the compression method is to find the boundary of the breast tissue and segment the breast from the rest of the background. We first analyze the gray level distribution of a mammogram. Figure 3 shows a profile of a horizontal scan line of an right CC image. The left side is the chest wall and the right side is the background. The gray level of the image is around 2000-3500 and drops to a background of around 200-300. Near the edge of the breast, the gray scale decreases gradually.

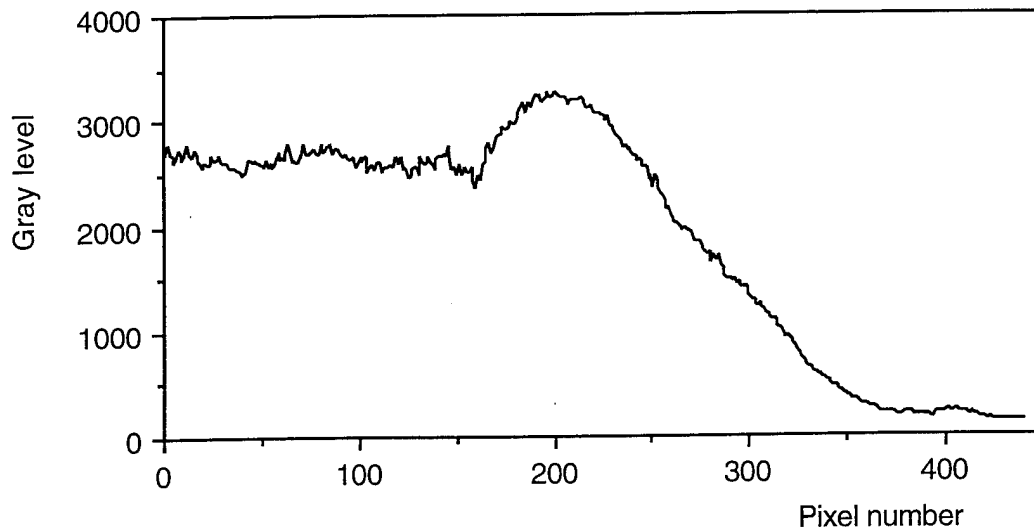


Figure 3. One scanning line of an Right CC mammographic image

We use a thresholding method to find the boundary. The edge of the breast is not clearly defined, because of background noise. The first step is to find an iso-gray level line near the edge of the breast. We setup a thresholding level about  $n$  times the background gray level. The ratio  $n$  is adaptively determined by the ratio of the average image gray level and the average background gray level for each image. The image is scanned line by line to find the location that reaches the thresholding level.

The first scanning line is started from one third of the height of the image from the top instead of starting at the top of the image in order to avoid the unpredicted behavior at the beginning of the image. After finding the first thresholding location, the scanning continues to the top and to the bottom of the image. Since the breast normally has a smooth boundary, only neighboring pixels of the previous thresholding position are scanned in each line.

Since the threshold position is not the true boundary, the second step is to expand the threshold position further to the background region until the true boundary is reached to include all the breast tissue. Figure 4 shows the iso-gray level line on a CC view image. The bold line is the true boundary of the breast and the thin line is the threshold position of the image based on the thresholding. The size of the expanding zoom  $\delta$  is determined by the following procedure. Ten

equally spaced horizontal scan lines are selected. Each scan line is smoothed and the distance between the threshold point and the edge of the image is determined. The average distance is determined from the 10 scan lines. The expending zoom  $\delta$  which equals to the average distance is added to each thresholding position.

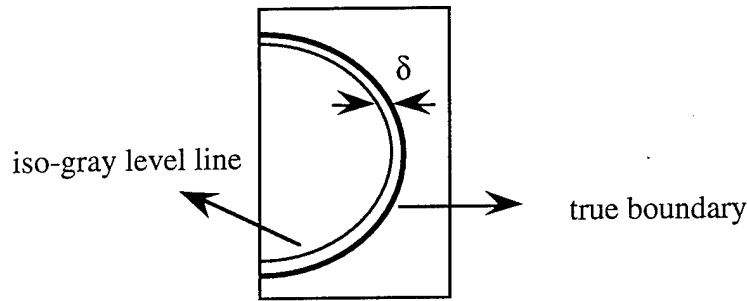


Figure 4. The thresholding of a CC view mammogram

It is worth mentioning here that mammograms normally have labels on films. It is most likely located in the background. When we segment breasts from the background, the labels are also removed. In a digital image, we can overlay text on the image. As long as the image header contains the information about the view, it does not affect the identification of the image.

### 2.3 Prediction Compression

The next step is to compress the image within the boundary using prediction lossless compression. X-ray images are formed due to the different attenuation coefficients of tissue. Breasts are mainly composed of adipose and fibroglandular tissues. Their attenuation coefficients differ slightly on the range of the diagnostic x-ray energies. Therefore, mammograms have less sharp edges and small contrast in comparison with other types of medical images as breasts contain primarily soft tissue. Due to lack of edges, we select prediction as an average of the previous pixels. The predicted pixel can be written as,

$$\hat{f}(x, y) = \sum_{i, j > 0} f(x - i, y - j) / n \quad (1)$$

where  $\hat{f}(x,y)$  is the predicted pixel,  $f(x,y)$  is the original pixel,  $i, j$  are integers, and  $n$  is the number of the pixel summed. The prediction error can be expressed as,

$$e(x,y) = f(x,y) - \hat{f}(x,y) \quad (2)$$

$e(x,y)$  is the error. After the prediction, the error image has less entropy and therefore more efficient for the entropy coding. Huffman coding is selected for the purpose of the entropy coding. Section 3.1 shows the results.

## 2.4 Lossy Compression Method With Wavelet Transform

We are also investigating a lossy compression scheme using wavelet transform. The wavelet theory has been developed for the past few years and it has been successfully applied to image compression<sup>26-30</sup>. A wavelet compression method normally has three steps: wavelet transformation, quantization, and entropy coding.

The basic idea of wavelet transformation is to represent any arbitrary function as a superposition of a wavelet basis<sup>31</sup>. A wavelet transform decomposes a signal into a series of smooth signals and their associated detailed signals at different resolution levels. At each level, the smooth signal and associated detailed signal have all the information necessary to reconstruct the smooth signal at the next higher resolution level. The transformed signal has both spatial and frequency information from the original signal and it provides a good representation for coding.

The quantization step reduces the data precision by coarsely quantizing the transformed data. In the last step, entropy coding such as Huffman or arithmetic coding can be used to reduce the coding redundancy.

### 2.4.1. Wavelet Transform

The wavelet transform can be implemented by a two-channel perfect reconstruction (PR) filter bank<sup>31</sup>. A filter bank is a set of filters, which are connected by sampling operators. Figure 5 shows an example of a two-channel filter bank applied to a one-dimensional signal.  $x(n)$  is an input signal.  $H_0$  and  $H_1$  are analysis filters and  $G_0$  and  $G_1$  are synthesis filters.  $H_0$  is a lowpass

filter and  $H_1$  is a highpass filter. If the output  $\hat{x}(n) = x(n-l)$ , where  $l$  is a delay, then the two-channel filter bank is called a perfect reconstruction (PR) filter bank.

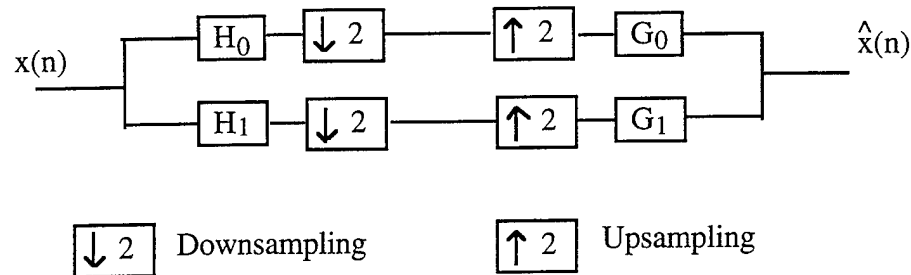


Figure 5. The perfect reconstruction filter

The wavelet transform of a signal can be obtained by repeatedly applying a PR filter bank to the signal in a pyramidal scheme. In a two-dimensional case, a two level wavelet decomposition process is shown in Figure 6:

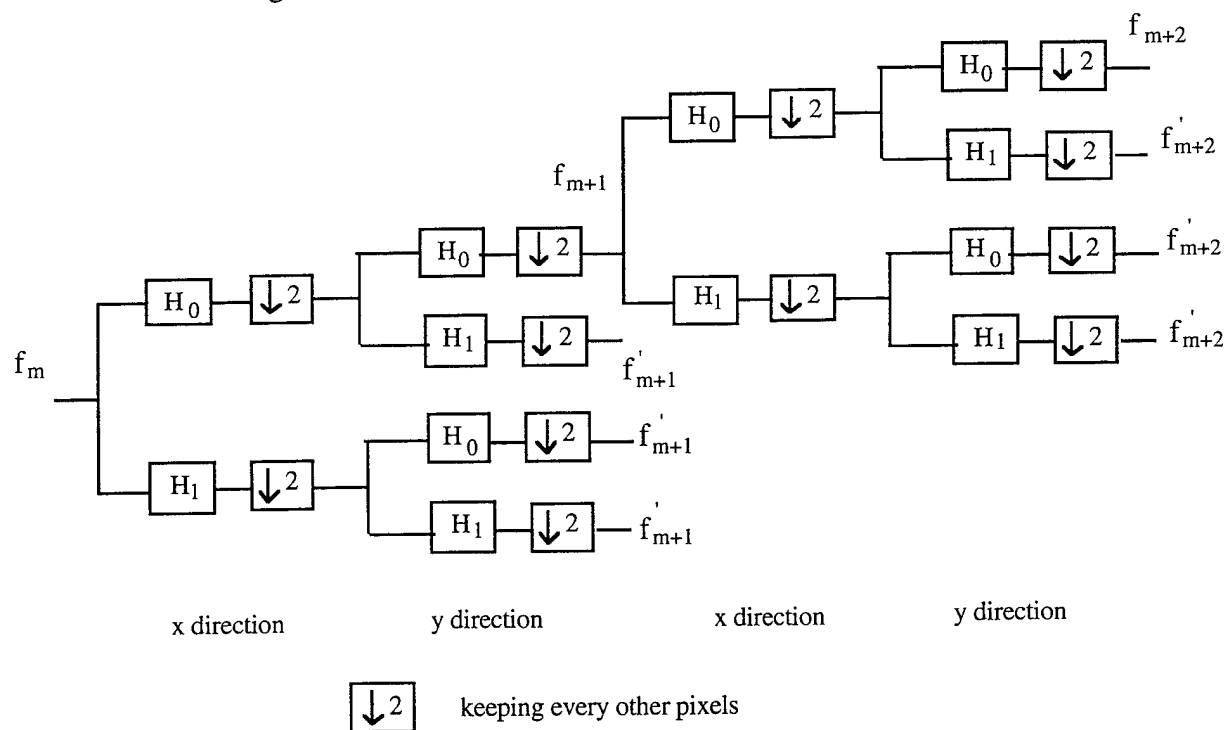


Figure 6. Two dimensional wavelet decomposition

We first convolve each line in the x-direction with filters  $H_0$  and  $H_1$  respectively, followed by subsampling every other pixel. We then convolve the resulting signals with  $H_0$  and  $H_1$  in the y-

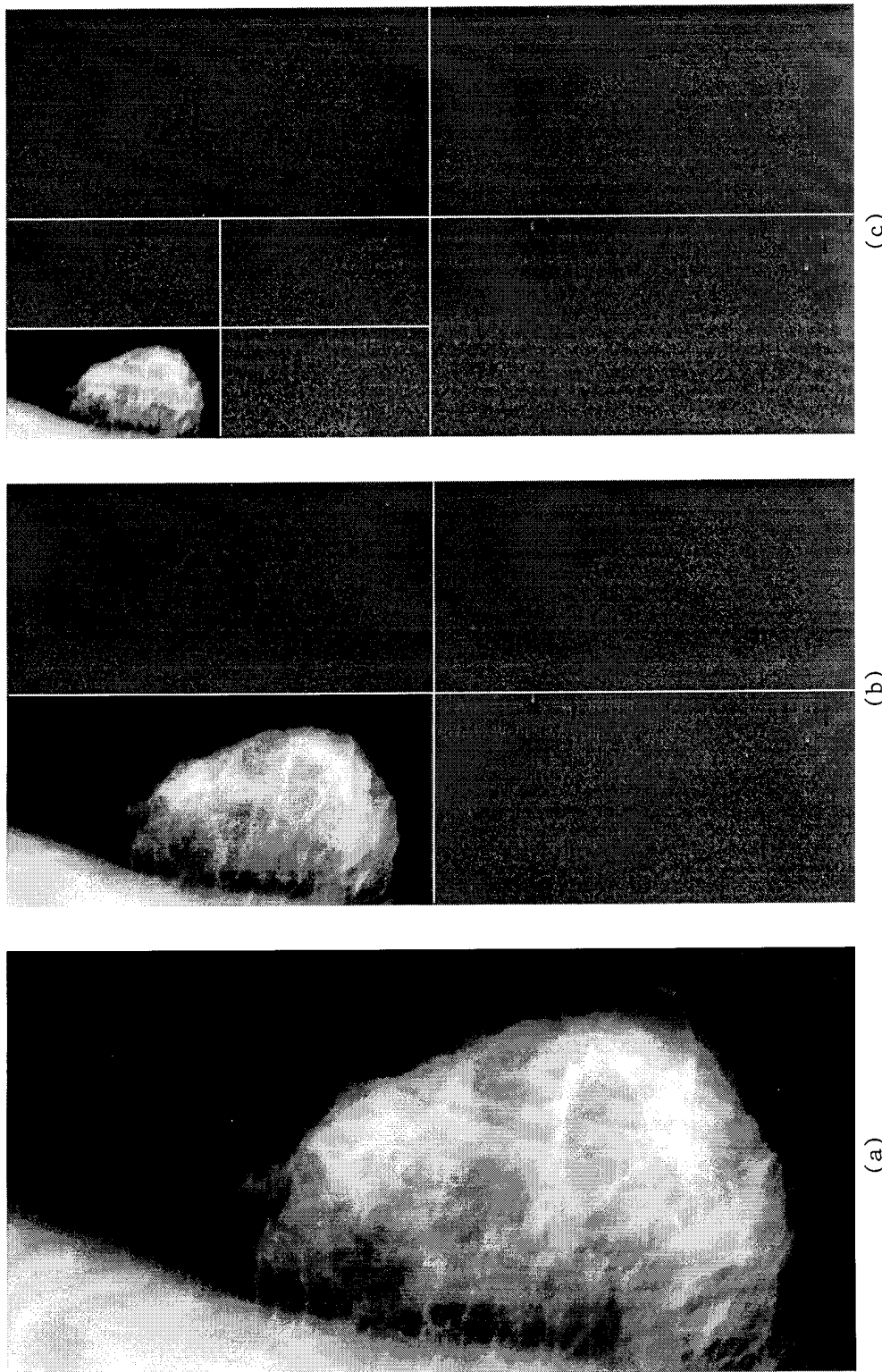
direction, followed by subsampling. Because  $H_0$  is a lowpass filter, and  $H_1$  is a highpass filter the resulting signals are a smooth signal  $f_{m+1}$ , and detailed signals  $f'_{m+1}$ , respectively. The smooth signal  $f_{m+1}$  is again convoluted with filters  $H_0$  and  $H_1$  respectively in x and y directions, sampled at every other pixel, creating  $f_{m+2}$  and  $f'_{m+2}$ s. The same process can be continued until the desired level is reached. The result of the two level decomposition contains  $f_{m+2}$ ,  $f'_{m+2}$ s and  $f'_{m+1}$ s.  $f_{m+2}$  is the smooth signal and  $f'_{m+2}$ s, and  $f'_{m+1}$ s are detailed signals in the different frequency bands.

Figure 7 shows one mammogram and its one level and two levels of the wavelet decomposition. Figure 7(a) is the original image. After the first level wavelet decomposition (Figure 7(b)) the image contains four components. The upper left corner is a smooth image, the rest of the corner images are detailed images. If the image has sharp edges on the x or y direction, the upper right corner will contain the edge in the y direction, the lower left corner will contain the edge in the x direction, and the lower right will contain the edge in the diagonal direction. Since the mammogram does not contain sharp edges. The detailed images do not show the visible edges. The upper right corner image can be further decomposed similarly into four more components at the next level of resolution (Figure 7(c)).

In general, the low frequency component of the wavelet transform is about  $1/2^{2M}$  of the original image size, but contains more than 99 % of the total energy, where  $M$  is the level of the decomposition. The high frequency components are separated into different resolution levels. At a particular resolution level, each block contains the high frequency information in certain directions. Blocks at different levels contain similar structure but different frequency band information.

#### 2.4.2. Quantization and Entropy Coding

The second step of the 3D wavelet compression is quantization. The purpose of quantization is to reduce data entropy by compromising the precision of the data. Reducing entropy allows more compression. The quantization step maps a large number of input values into a smaller set of output values. The original data cannot be recovered exactly after quantization. It is therefore very



(a) (b) (c)

Figure 7. A mammogram and its wavelet transformation. (a) Original image. (b) One level decomposition. (c) Two level decomposition.

important to design a quantization strategy which selectively quantizes the wavelet coefficients and preserves the image quality. A simple uniform quantization is implemented in this study.

In the third step, run-length coding followed by Huffman coding is applied to the quantized data. Run-length coding is effective when there is more than one pixel with the same gray level in a sequence. This method uses two integers to represent a sequence of the same gray level. The first integer represents the length of the sequence, and the second integer represents the gray level of the sequence. The longer is a sequence, the more efficient the run-length coding. Since thresholding of the high resolution components results in a large number of zeroes, run-length can be expected to significantly reduce the data.

### 3. Result

#### 3.1 Structure Lossless Compression

The structure lossless compression method was applied to all 100 digitized mammograms at a resolution of 100  $\mu\text{m}$  and 200  $\mu\text{m}$ . The segmentation algorithm correctly segmented breast images from their backgrounds for each of the 100 mammograms. We applied five different predictions. The predictions are one of the previous neighboring pixels  $a$ ,  $b$ ,  $c$ , or their combinations  $(a + c) / 2$ ,  $(a + b + c) / 2$  as shown in Figure 8.

|   |   |   |
|---|---|---|
|   | b | c |
|   | x | x |
| a | x | 0 |

Figure 8. The distribution of the prediction pixels

Table 1 summarizes the results. The compression ratios are higher at 100  $\mu\text{m}$  than those at 200  $\mu\text{m}$  for all the predictors. Column 5 gives the best compression ratios. Comparing the first three columns, predictor  $a$  gives the best results. This may be the partially caused by digitization

process. This is because the digitizer scans an image line by line, an immediate previous pixel of the same scanning line has better correlation with the current pixel.

| Predictions       | a    | b    | c    | $(a+c)/2$ | $(a+b+c)/2$ |
|-------------------|------|------|------|-----------|-------------|
| 200 $\mu\text{m}$ | 3.17 | 3.07 | 3.01 | 3.24      | 3.19        |
| 100 $\mu\text{m}$ | 3.34 | 3.24 | 3.18 | 3.42      | 3.37        |

Table 1. Summary the average compression ratios for different predictors

### 3.2 Wavelet Compression

We also applied the wavelet compression to a mammogram at different compression ratios. The energy distribution of the wavelet transformed image was calculated. The total energy of a digital image is defined as the summation of the square of the pixels. Table 2 lists the energy distributions at the different levels of decomposition.

|         | Low Resolution | High Resolution |         |         |         |
|---------|----------------|-----------------|---------|---------|---------|
|         |                | level 1         | level 2 | level 3 | level 4 |
| 1 level | 99.9844        | 0.0156          |         |         |         |
| 2 level | 99.9750        | 0.0155          | 0.0095  |         |         |
| 3 level | 99.9629        | 0.0154          | 0.0093  | 0.0124  |         |
| 4 level | 99.9343        | 0.0151          | 0.0092  | 0.0122  | 0.0292  |

Table 2. The energy distribution of the wavelet transformed image.

The first column shows the number level of the decomposition. The second column shows the percentage energy in the low resolution component in that level. The rest columns show the percentage energy at different levels of the decomposition. The energy in the low resolution area is more than 99.9 percent for all level of the decomposition and the total energy in the high resolution

areas is less than 0.06 percent. Therefore, selectively quantizing the high resolution components will only introduce small errors.

The decompressed image quality is measured by the Peak Signal to Noise Ratio (PSNR) defined as:

$$PSNR = 20 \log \frac{f_{\max}}{\frac{\{\sum [f(x,y) - f_c(x,y)]^2\}^{1/2}}{N}} \quad (3)$$

where  $f_{\max}$  is the maximum gray level of the image set,  $N$  is total number of pixels,  $f(x,y)$  is the original image, and  $f_c(x,y)$  is the decompressed image. The denominator is the Root Mean Square Error of the decompressed image. The larger the PSNR, the better the decompressed image quality is.

The compression performance is measured by Compression Ratio (CR) which is defined as

$$CR = \frac{\text{Original bits / pixel}}{\text{Compressed bits / pixel}} \quad (4)$$

Figure 9 shows compression ratio versus PSNR at different wavelet decomposition levels. The level of decomposition ranges from 1 to 4. At each level of decomposition, a range of compression ratios were obtained by varying the quantization numbers. The PSNR is used to measure the decompressed image quality. At the same PSNR, the compression ratio increases as the level of the wavelet decomposition increases. This is because at higher levels the image energy is concentrated in a smaller low resolution area, and the high resolution components are separated into different frequency bands, it is therefore convenient to quantize the high resolution components. The compression performance does not increase when the level of the wavelet decomposition is higher than four.

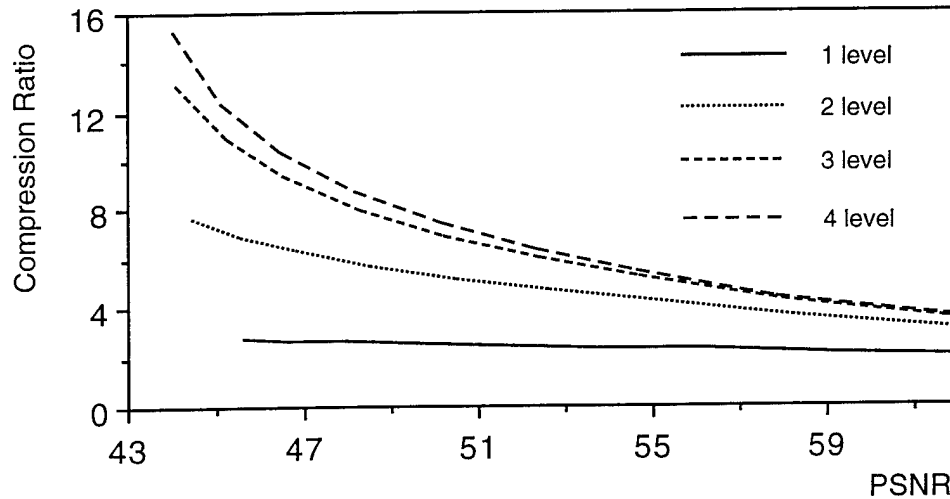


Figure 9. The compression ratio versus PSNR at different wavelet decomposition level

## 4. Conclusions

During the past year, we investigated both lossless and lossy compression methods. In the lossless compression, we developed a structure lossless compression method for mammogram image compression. The algorithm utilizes the unique shape characteristics and image characteristics of mammograms. The combination of the segmentation and the prediction coding enable us to achieve high compression ratios and without losing any useful information.

In the lossy compression, we developed a wavelet method which formulates the background for improving image quality at high compression ratios for mammograms. Currently, we implemented a limited number of wavelet filters for the transformation and utilized simple uniform quantization. In the next stage, we will study the wavelet filter characteristics and select the best filter for mammogram compression. Uniform quantization will be replaced by an adaptive quantization method based on mammogram characteristics.

PSNR and compression ratio have been used for assessing image quality and compression performance. However, PSNR cannot represent the human visual system. In the next stage, we will propose quantitative parameters for better assessing image quality.

## 5. References

1. American Cancer Society. Cancer Facts and Figures, 1992 Atlanta: American Cancer Society.
2. Silverberg E, Lubera JA. Cancer Statistics CA-A Cancer. *Journal for Clinicians* , Vol. 39(1) 3-20.
3. Crisp WJ, Higgs MJ. Screening for Breast Cancer Detects Tumors at an Earlier Biological Stage. *Br. J. Surg.* Vol. 80(7), 863-865, 1993.
4. Seidman H, Gelb SK, Silverberg E, Laverda N, Lubera JA, Survival Experience in Breast Cancer Detection Demonstration Project. *CA Cancer J. Clin.* , Vol. 37, pp. 258-90, 1987.
5. Noriaki Ohuchi, Koichi Yoshido. Improved Detection Rate of Early Breast Cancer in Mass Screening Combined with Mammography. *Jap. J. Cancer Res.* Vol. 84(7), pp. 807-812, 1993.
6. Yaffe, MJ. Digital Mammography, 1993.
7. Shtern F. Digital Mammography And Related Technologies: A Perspective From The National Cancer Institute. *Radiology* . Vol. 183, pp. 629-630, 1992.
8. Anderson LF. Large-Scale Effort Tests Digital Mammography's Potential. *Journal Of The National Cancer Institute.* Vol. 86(8), pp.580-2, 1994.
9. Wu Y, Doi K, Giger ML, Nishikawa M. Computerized Detection of Clusters Microcalcification in digital Mammograms: Applications of artificial Neural networks. *Med. Phys.* Vol. 19(3), pp.555-560, 1992.
10. Barnes CT, Morin RL, Staab EV Inforad: Computer for Clinical Practice and Education in Radiology Teleradiology: Fundamental Considerations and Clinical Application. *Radiographics.* Vol. 13(3), pp. 673-81, 1993.
11. Karssemeijer N, Frieling JTM, Hendriks, JHCL, Spatial Resolution in Digital Mammography. *Investigative Radiology*, Vol. 28(5), pp. 413-419.
12. Nawano S, Evaluation of Digital Mammography in Diagnosis of Breast Cancer. *Journal of Digital Imaging*, Vol. 8 (1) Suppl 1, pp. 67-69, 1995.
13. Huffman DA, A Method for the Construction of Minimum-redundancy Codes. *Proceedings of the IRE*, Vol. 40(10) pp. 108-1101, 1952.
14. Margos PA, Schafer RD and Mersereau RM, Two-Dimensional Linear Prediction and Its Application to Adaptive Predictive Coding of Images. *IEEE Trans. on Acoustics, Speech, and Signal Processing.* Vol. ASSP-32(6), pp.1213-1229, 1984.
15. Wintz PA, Transform Picture Coding. *Proceedings of IEEE.* Vol. 60(7). pp. 809-820, 1972.
16. Woods JW and O'Neil SD, Subband Coding of Images. *IEEE Trans. on Acoustics, Speech, and Signal Processing.* Vol. ASSP-34(5), pp.1278-1288, 1986.

17. Ramamurthi B, and Gersho A, Classified Vector Quantization of Images. *IEEE Transactions on Communications*. Vol. COM-34(11), pp. 1105-1115, 1986.
18. Lo SC, Huang HK. Radiological Image Compression: Full-Frame Bit-Allocation Technique. *Radiology*, Vol.155(3), pp. 811-817,1985.
19. Chan KK, Lou SL, Huang HK. Radiological Image Compression Using Full-Frame Cosine Transform with Adaptive Bit-allocation. *Comp Med Imag and Graphics*. Vol. 13(2), pp. 153-159, 1989.
20. Lee H, Frank, MS, Rowberg AH, Choi HS. A New Method for Computer Tomography Image Compression Using Adjacent Slice Data. *Investigative Radiology*. Vol. 28(8), pp 678-685, 1993.
21. Koo JI, Lee HS, Kim Y. Application of 2-D and 3-D Compression Algorithms to Ultrasound Images. *SPIE* Vol. 1653, pp.434-439, 1992.
22. Roos P., Viergever M.A., Van Dijke M.C.A. and Peters J.H., Reversible intraframe Compression of Medical Image. *IEEE Trans. on Medical Image*, Vol. 7(4), pp. 328-336, 1988.
23. Kuduvalli, G.R.; Rangayyan, R.M. Error-Free Transform Coding By Maximum-Error-Limited Quantization of Transform Coefficients. *Proceedings of the SPIE*, Vol.1818, pt.3:1458-61, 1992.
24. Brettle DS; Ward SC; Parkin GJ; Cowen AR; Sumsion HJ. A Clinical Comparison Between Conventional And Digital Mammography Utilizing Computed Radiography. *British Journal of Radiology*, Vol.67(797), pp 464-8,1994.
25. Lucier BJ, Kallergi M, Qian W, et al. Wavelet Compression and Segmentation of Digital Mammograms. *Journal of Digital Imaging*, Vol. 7(1) pp 27-38, 1994.
26. Daubechies I, "Orthonormal Bases Of Compactly Supported Wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp909-996, 1988.
27. Mallat SG, A Theory For Multiresolution Signal Decomposition: The Wavelet Representation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*," Vol. 11(7), pp674-693, 1989.
28. Antonini M, Barlaud M, Mathieu P, and Daubechies I., Image Coding Using Wavelet Transform, *IEEE Trans. on Image Processing*, Vol.1(2), 1992.
29. Bradley JN, Brislawn M, The Wavelet/Scalar Quantization Compression Standard For Digital Fingerprint Images, *Proc. IEEE ISCA-94*, London.
30. Wang J, and Huang HK, Three-Dimensional Medical Image Compression Using A Wavelet Transform With Parallel Computing, *SPIE Medical Imaging 95*, Vol. 2431, p162-172.
31. Vetterli M and Herley C., Wavelets And Filter Banks: Theory And Design, *IEEE Trans. on Signal Processing*, Vol. 40, pp2207-2232, 1992.

## Appendix A: Bibliography

1. Jun Wang and H.K. Huang, Real-Time Structure-Lossless Mammographic Image Compression. RSNA'95 (Abstract).
2. Jun Wang and H.K. Huang, Three-Dimensional Medical Image Compression Using A Wavelet Transform With Parallel Computing. *SPIE Medical Imaging 95*, Vol. 2431, p162-172.
3. Jun Wang and H.K. Huang, 3D Medical Image Compression by Using 3D Wavelet Transformation. Submit to IEEE Trans. on Medical Imaging.

## **Appendix B: Personnel Receiving Pay**

Ms Jun Wang received total of \$19,000 from July 15, 1994 to July 14, 1995, including \$5,000 tuition and fees.