Enhancement and analysis of digital mammograms using fuzzy models

Sutton, Melanie, Bezdek, James


Event: 26th AIPR Workshop: Exploiting New Image Sources and Sensors, 1997, Washington, DC, United States
Enhancement and analysis of digital mammograms using fuzzy models

Melanie A. Sutton
Jim Bezdek

Department of Computer Science
University of West Florida, Pensacola, FL 32514
msutton@argo.cs.uwf.edu; jbezdek@ai.uwf.edu

ABSTRACT

This paper describes our work in enhancing and analyzing digital mammograms from the Digital Database for Screening Mammography (DDSM). The DDSM will ultimately contain 3000 cases and provides a unique opportunity for researchers from around the world to compare results on a large, diverse data set. However, the size of the database and images within it require careful consideration of memory limitation issues, display device constraints, etc. We address research problems connected with the modification and application of existing fuzzy modeling approaches to this digital mammography domain. Segmentation and edge detection are used as benchmark applications for the comparisons we make.

Key Words: DDSM, fuzzy models, digital mammography, computer assisted diagnosis, breast cancer.

1. PROBLEM DEFINITION

"There is no radiologist, regardless of skill and expertise who has not failed to see a significant abnormality on a mammogram that is visible in retrospect. There is a psycho visual threshold for all involved in observational endeavors that cannot be completely eliminated. Since, at any one time, different observers tend to overlook different findings, multiple observers are less prone to oversight than a single observer. In order to reduce the error rate for the detection of breast cancers, radiologists have developed double reading systems in which more than one trained radiologist reviews the same mammogram. This has been shown to reduce the error rate by 5-15%. One major benefit from CAD [computer assisted diagnosis] is the development of a second reader system that is tireless, consistent, and cannot be distracted1."

The new Digital Database for Screening Mammography (DDSM), developed with support from the U.S. Army Medical Research and Material Command (USAMRMC), offers exciting opportunities for researchers to study and make more thorough comparisons on a large, comprehensive database with ground truth information2. The database currently contains 235 normal 42 micron cases and a set of 56 normal and 61 abnormal (in terms of masses) 100 micron cases (available as of October 1997). The DDSM project will ultimately build a database of 3000 cases.

We need automated techniques to handle this large number of images. Careful consideration must also be given to the way the images will be stored and processed internally. For example, images from this database are 16-bit intensity images which can be as large as 6900x4100 pixels, requiring approximately 56.58 Mbytes of space, if read in their entirety. One image of this size presents the user with time and storage complexity that eclipses most current software for viewing, processing, and interpreting structural details to the degree required by practicing radiologists.

In a clinical environment, radiologists use light boxes to view films for evidence of breast cancer. Potential indicators include the following:

- **microcalcifications** small, bright areas
- **masses** lesions with irregular or spiculated boundaries
- **architectural distortions** indicated by skin protrusions, focally thick areas, etc.
- **asymmetry** when comparing left and right breasts
In order for CAD researchers to contribute to this effort, these films must be digitized, and an appropriate set of algorithms must be developed to aid the diagnostic process. While a number of automated approaches have been tried in this area, none has evolved as clearly superior. Our interest in this area is motivated by past success with using fuzzy models in other medical imaging domains (e.g., detection of brain tumors with MRI) and a desire to test their generalization capabilities to the digital mammography domain.

2. DATABASE SELECTION

In order to develop and systematically test an approach, researchers must first clarify which aspect of the problem (in terms of breast cancer indicator) will be analyzed, and then find an appropriate database of images which can support the endeavor. Fortunately, numerous databases have evolved as researchers have initiated programs to study automation techniques, as described in Table 1.

Table 1. Some databases for digital mammography research

<table>
<thead>
<tr>
<th></th>
<th>LLNL/UCSF</th>
<th>U. of Chicago</th>
<th>MIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>spatial resolution</td>
<td>35 microns</td>
<td>50 microns</td>
<td>50 microns</td>
</tr>
<tr>
<td>digitizer</td>
<td>Dupont NDT</td>
<td>Lumisys scanner</td>
<td>Joyce-Loebl, SCANDIG-3</td>
</tr>
<tr>
<td>greyscale response</td>
<td>linear to transmitted intensity, adjusted for each image</td>
<td>N/A</td>
<td>linear in range 0-3.2 OD</td>
</tr>
<tr>
<td>number of patients</td>
<td>50 (198 images)</td>
<td>N/A (1000 cases expected)</td>
<td>N/A (322 images)</td>
</tr>
<tr>
<td>type of views</td>
<td>MLO, CC</td>
<td>MLO, CC</td>
<td>MLO only</td>
</tr>
<tr>
<td>image categories</td>
<td>normals, calcifications</td>
<td>normals, calcifications, masses, architectural distortion, asymmetry</td>
<td>normals, calcifications, masses, architectural distortion, asymmetry</td>
</tr>
<tr>
<td>reference #</td>
<td>23</td>
<td>24</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Nijmegen</th>
<th>Washington U.</th>
<th>DDSM - A</th>
<th>DDSM - B</th>
</tr>
</thead>
<tbody>
<tr>
<td>spatial resolution</td>
<td>100 microns</td>
<td>N/A</td>
<td>42 microns</td>
<td>100 microns</td>
</tr>
<tr>
<td>digitizer</td>
<td>Eikonix 1412 CCD</td>
<td>LORAD CCD-based</td>
<td>DBA scanner</td>
<td>Lumisys 200 scanner</td>
</tr>
<tr>
<td>greyscale response</td>
<td>grey level 4095 corresponds to 0.18 OD</td>
<td>N/A</td>
<td>logarithmic in range 0-3.6 OD</td>
<td>linear in range 0-3.6 OD, with minimal grey level 495</td>
</tr>
<tr>
<td>number of patients</td>
<td>21 (40 images)</td>
<td>N/A (30 images)</td>
<td>235 (3000 cases expected)</td>
<td>117</td>
</tr>
<tr>
<td>type of views</td>
<td>MLO, CC</td>
<td>core biopsy</td>
<td>MLO, CC</td>
<td>MLO, CC</td>
</tr>
<tr>
<td>image categories</td>
<td>calcifications</td>
<td>lesions</td>
<td>normals (currently)</td>
<td>masses (currently)</td>
</tr>
<tr>
<td>reference #</td>
<td>26</td>
<td>27</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

However, the diversity of databases and quality control mechanisms in place complicate the selection of an appropriate database. Researchers must decide if the spatial and intensity resolution of the database are sufficient for locating a particular abnormality. For example, one investigation states that microcalcifications can be as small as 100 microns/pixel. If more than one database is required (to increase the sample size or test the generalization capabilities of the approach), the databases must also supply digitization information so the data can be normalized prior to comparison. Finally, the necessary hardware components and/or compression techniques must be available or developed to process the images efficiently.

The DDSM database will be one of the largest digital mammography databases ever developed, and each image will have completely annotated ground truth information. The developers of this database...
have carefully studied the effects of film digitization repeatability and checked the intensity values in the resulting images against optical density standards. The quality and size of the database is extremely important if we hope to generalize our techniques to the large volume of data encountered in real screening programs. Statistics gathered by Sacred Heart Women's Hospital (Pensacola, FL) in 1996 indicate that screening programs at regional health care facilities typically involve 10,000-12,000 mammograms per year with cancer rates of approximately 0.9 percent. The DDSM database, with a goal of creating 3000 cases, comes closer than any other database to this case load.

The University of West Florida is fortunate in that two Pensacola radiologists are supporting (part of) the DDSM effort by providing films for the database, and thus, we have access to (some of) the original films used in the digitization process. This is an important aspect of CAD evaluation, since there are three ways (see Figure 1) that radiologists can be involved in performance analysis, but only one of the three types of comparison is available without access to the original films.

### 3. VERIFICATION AND VALIDATION STRATEGIES

#### 3.1 Evaluation by clinician comparison

Usually, clinicians are provided a set of images and asked to evaluate one or more "properties" of the enhancement process, such as: faithful replication of film information, utility of enhancement (edges, regions, etc.), and implications of region of interest prompts.

As shown in Figure 1, radiologists can be involved in performance evaluation in three ways:

- **C1.** original films versus digitized images
- **C2.** original digitized images versus enhanced digitized images
- **C3.** original films versus enhanced digitized images

When radiologists examine mammogram films, a light box display unit is used, and the entire mammogram suite can be viewed. However, when a digital version of the film is viewed on a computer monitor, only a fraction of the digitized film can be viewed at one time, often with 256 shades of grey (assuming an 8 bit display). When printed, these digitized images must be further cropped or compressed to fit the paper constraints of the printer. In this case, an average 600 dpi (dots per inch) laser printer can resolve only 122 grey levels. Research indicates that humans themselves can only resolve 32 grey levels.
Since radiologists are most comfortable with films while non-clinicians such as computer scientists are most comfortable with computer monitors and printouts of digitized images, some compromise is involved when evaluating the performance of the system. Having access to all three of the data types displayed in Figure 1 (original films, digitized images and processed images) maximizes flexibility when determining the nature of this compromise, while still ensuring the development of a sound and repeatable evaluation methodology.

3.2 Evaluation by computational comparison

In addition to clinician involvement in performance evaluation, digital mammographic databases (including DDSM) contain ground truth information which includes American College of Radiology (ACR) ratings, size and locations of lesions, etc. When such information is available, performance analysis can be also be based on various pattern recognition methods such as:

- **PR1.** labeled test data which are used to estimate error rates
- **PR2.** correlation coefficients between computed and ground truth labels
- **PR3.** receiver operating characteristic (ROC) analysis

There are several ways to use these measures. For example, algorithm X can be used with training data to find a classifier, and then the labels the classifier generates on test data can be compared to ground truth labels to assess the quality of the algorithm. Or, any of these measures can be used to compare the relative quality of several algorithms X, Y ... Z on common data sets that possess ground truth. The general situation for assessment by pattern recognition techniques is summarized in Figure 2.

![Figure 2. Supervised pattern recognition techniques for algorithm evaluation](image)

Evaluation with or without clinician involvement provides insights into the success or utility of the proposed technique. However, we feel that clinician involvement is vital to developing a generalizable, non-database specific, repeatable methodology. Because the film to digitization to processing to clinician loop shown in Figure 1 is available to us in Pensacola, with feedback from practicing radiologists active in a breast cancer screening program that is supplying films for the DDSM project, we will be able to use methods C1-C3 as well as PR1-PR3 in the evaluation of proposed fuzzy models.

4. PREVIOUS WORK IN MAMMOGRAPHY WITH FUZZY MODELS

The application of fuzzy models to analysis of digital mammography is quite limited, and has primarily considered only a small set of images. Brzakovic, et al. studied the use of fuzzy pyramid linking for the detection of microcalcifications and nodules\textsuperscript{10}. Lo, et al. also focused on the detection of clustered microcalcifications using fuzzy classification modeling\textsuperscript{11}. Li, et al. examined the use of fuzzy membership functions for Markov random fields for tumor detection in mammogram images\textsuperscript{12}. Sameti, et al. studied mammogram partitioning based on intensity patterns and fuzzy membership functions\textsuperscript{13}.
Petrick is currently studying the detection of mammographic masses using adaptive enhancement and fuzzy classification\(^{14}\). Wang and Karayiannis used fuzzy learning vector quantization and other techniques to compress digital mammograms that contain microcalcifications\(^{15}\).

Fuzzy modeling techniques have also been used in other medical imaging applications\(^{16-18}\). For example, the unsupervised fuzzy c-means (FCM) clustering algorithm often provides better segmentation of magnetic resonance images (MRIs) of the brain than supervised learning techniques such as statistical mixture decomposition, k-nearest neighbor rules, and feedforward neural networks\(^ {17-19}\). Evaluation of algorithmic results in\(^ {17}\) was done by a panel of radiologists who independently compared raw MRI data to the outputs of the algorithm (this is an example of evaluation technique C2). FCM has also been used more successfully than the k-nearest neighbor rule and commercially available techniques for tumor volume estimation\(^ {20}\).

5. OVERVIEW OF THE UWF SYSTEM

In our system an input DDSM image is processed in a sequence of steps to enhance breast substructure. Specifically, FCM is used to segment the breast area into homogeneous tissue regions. An edge detector based on the Takagi-Sugeno (TS) fuzzy input-output model is then used to detect and enhance breast substructure. A general diagram of the flow of control and representative processed/enhanced images is shown in Figure 3, where many of the specific details involved when analyzing these images are omitted. These images are very large and their size heavily impacts system resources. The following paragraphs clarify various steps taken to limit this impact while preserving the diagnostic utility of processed outputs.

**Input.** DDSM images are acquired with the characteristics described in Table 1.

**Database Normalization.** In order to make useful comparisons across databases, images must be normalized to some common space, based on digitization information such as the grey level response characteristics detailed in Table 1. For demonstration purposes, examples shown in this paper have not been normalized.

**Noise Removal.** Median filtering with a 5x5 window is applied to remove noise introduced in the digitization process.

**Reduction/Compression.** Images from this database can be as large as 6900x4100, requiring approximately 56.58 Mbytes of space when intensities are 2 byte integers. However, if subsequent processing requires floating point numbers, the storage required for one temporary image can reach 226.32 Mbytes. A number of researchers use automatic techniques for reducing the image by cropping it to a more reasonable size. For example, the background (air area) may be cropped from the breast area prior to pixel level analysis. The image can also be compressed by other methods such as local averaging of the original image over nxn windows, which compresses the image to \((1/n)^2\) its original size. While reduction by cropping and compression by local averaging retain grey level resolution (i.e. final images, although smaller, are at the same grey level resolution as the input image, 16 bit), the latter method will reduce the spatial resolution of the image. For example, using a 2x2 local averaging window applied to an original 42 micron image results in an 84 micron image. Local averaging using a 6x6 window is used for the results provided in this paper.

**Conditioning (histogram modification).** Histogram modification can be used to modify the grey level distribution for subsequent processing. One common justification for this is to condition the image by enhancing specific intensity value ranges. For example, this method can be used to enhance the distinction between pixel intensity values such as 0 and 1. Histogram modification can also be applied to trim the data requirements from 16 bit to 8 bit. For example, a 16 bit image can be compressed to an 8 bit image by bit shifting from the range \([0,...,65535]\) to \([0,...,255]\), or by applying some function to the input intensity (e.g., square root) which converts the input intensity value to an 8 bit value. These methods do not alter the spatial resolution of the image (i.e. the final image is still 42 micron), but may improve the contrast quality of relatively dark mammographic images with a courser grey level resolution. We condition images with dynamic scaling\(^ {5}\) to non-linearly scale pixels close to 0.
Feature Extraction. Features extracted from the images are used to provide support for "significant" areas of breast substructure. Preliminary analysis using the Sobel, range, and standard deviation features extracted from 3x3 windows applied to images from the MIAS database indicate that the non-directional properties of the range and standard deviation features can be more useful for extracting edges in mammographic images. These features tend to brighten the image (improving contrast) and thicken edges (indicating texture). Range and standard deviation features enjoy three other advantages; they can be used for rectangular (m x n) windows, their formulae are fixed functions of n and m, and they have well known statistical distributions.
Normalization. To judge the effectiveness of different types of extracted features, images are normalized to a common range so that corresponding blended features can be compared in an unbiased manner. We normalized feature images to real numbers in the range \([0, ..., 4]\) for this study.

Blending. The extracted features can be combined in many ways besides the traditional Manhattan and Euclidean norms for enhancement of various breast structures. We have examined the use of both the generalized logistic function and 4-rule Takagi-Sugeno (TS4) fuzzy systems for blending to enhance edges within the breast substructure. Initial experimentation indicates that the TS4 model is particularly useful for strengthening weak edges in the breast substructure without emphasizing strong ones.

Image Scaling (histogram modification). For printing, visual display purposes, and to make accurate comparisons of the outputs of different blending strategies, all blended feature edge images are scaled to the range \([0, ..., 255]\).

Output. The final output image is an unthresholded edge map representing significant areas of breast substructure. This image is displayed on a computer monitor with up to 256 grey levels and can be printed on a 600 dpi laser printer (as used for this paper) with up to 122 grey levels. At this point, the subjective quality of the image is rated by the authors. Ultimately, a set of radiologists will be consulted for a professional interpretation.

6. EXPERIMENTAL RESULTS

6.1. Edge detection

To understand the use of the Takagi-Sugeno fuzzy system\(^{22}\) for blending edge features in the digital mammography domain we provide a brief description of the multiple-input single-output (MISO) case. In step 0, \(\mathbf{x} = (x_1, x_2, ..., x_p)^T\) in \(\mathbb{R}^p\) is an input vector. In our application, \(\mathbf{x}\) is a vector of edge features extracted from the conditioned image. For example, \(\mathbf{x}\) might be the 2-D vector whose coordinates are the horizontal and vertical components of the Sobel estimates of the digital gradients in each 3x3 window in the image. Step 0 identifies the numerical range \(D_k\) associated with linguistic variable \(l_k\) that provides a semantic description of \(r\) subdomains of \(D_k\). \(r\) is the granularity of \(l_k\). The j-th subdomain of \(l_k\) represents a linguistic value, \(l_{kj}\), which is in turn represented by a fuzzy membership function \(m_{kj}: D_k \rightarrow [0, 1]\). \(m_{kj}: 1 \leq j \leq r\) is called the linguistic termset associated with variable \(k\), \(1 \leq k \leq p\). Step 0 is often referred to as fuzzification of the input domain.

Step 0 comprises the reasoning mechanism. The left hand side (LHS) and right hand side (RHS) of the rule base (RB) are composed of M rules \([R_q]\) that operate on \(\mathbf{x}\). \(\alpha_q(\mathbf{x}) = \cap(m^q(\mathbf{x}))\) is called the firing strength of rule \(q\) : \(1 \leq q \leq M\). \(\cap\) is any \(\cap\)-norm (intersection) operator on \([0, 1] \times [0, 1]\). Different membership functions among the \(m_{kj}\) will be used as \(\mathbf{x}\) runs through its domain. Here is rule \(q\):

\[
R_q: \alpha_q(\mathbf{x}) = \cap(m^q(\mathbf{x})) = \cap(m_{k1}(x_1), ..., m_{ktk}(x_k), ..., m_{kp}(x_p)) \Rightarrow u_q(\mathbf{x}) ; 1 \leq q \leq M
\]

The output functions \(\{u_q: \mathbb{R}^p \mapsto \mathbb{R}\}\) comprise the RHS of the rule base. Each \(u_q\) has a functional form (e.g., linear, affine, quadratic, trigonometric, transcendental, power, etc.) specified by the user. Step 0 produces the numerical output \(u(\mathbf{x})\) as

\[
u(\mathbf{x}) = \frac{\sum_{k=1}^{M} \alpha_k(\mathbf{x})u_k(\mathbf{x})}{\sum_{j=1}^{M} \alpha_j(\mathbf{x})}
\]
We call the particular function used for blending edge features which is based on the 4-rule TS model given below $b_{TS4}$; more specifically, in this paper $b_{TS4}: \mathbb{R}^2 \mapsto \mathbb{R}$. Membership functions for "L=low" and "H=high" (on $x$ in [0, 4]) for $b_{TS4}$ are shown in Figure 4. The feature $x$ might be, for example, the range or standard deviation of the intensities in the window, or a Sobel gradient estimate. $b_{TS4}$ uses the same two membership functions along the y axis for feature input vector $x = (x, y)$ in $\mathbb{R}^2$.

![Figure 4. Membership functions for $b_{TS4}$](image)

Let $\tau, \chi, \gamma, \omega \in \mathbb{R}^+$ and define four specific rules for equation (1) as:

- R.1. If $x = L$ and $y = L \Rightarrow u_1(x) = x^\tau + y^\tau$
- R.2. If $x = L$ and $y = H \Rightarrow u_2(x) = \chi$
- R.3. If $x = H$ and $y = L \Rightarrow u_3(x) = \gamma$
- R.4. If $x = H$ and $y = H \Rightarrow u_4(x) = \omega$

Choose the $\tau$ norm as the product of its arguments, $\tau(a,b) = ab$. With the membership functions shown in Figure 4, rules (3 a-3 d) can be written explicitly:

$$b_{TS4}(x; \tau, \chi, \gamma, \omega) = m(x)m(y)[x^\tau + y^\tau + \omega - \chi - \gamma] + m(x)[\chi - \omega] + m(y)[\gamma - \omega] + \omega,$$

$$m(z) = \max\{0, 1 - |z|/4\}, \quad z = x \text{ or } y \in [0, 4]$$

Consult reference 5 for a detailed analysis of characteristics of input-output surfaces that are generated by (4). In brief, $\chi, \gamma, \omega$ pin the 3 non-zero corners of the output surface at the heights given by these values. The output function for rule 1, $u_1(x) = x^\tau + y^\tau$, controls $b_{TS4}$ in the neighborhood of the origin of feature space, and is the most important rule for variation of TS edge images.

Edge detection based on $b_{TS4}$ was used to enhance the substructure of the breast areas of the two DDSM input images shown in the left views of Figure 5. The normal image on the top is case A_0329_1.RIGHT_CC in DDSM, scanned with the OBA scanner. The image on the bottom is case B_9023_1.LEFT_MLO in DDSM, part of a special volume of abnormal images acquired using the Lumisys scanner. Database ground truth is seen in the lower left view as the set of white pixels that delineate the very large lesion. The features used in all four processed views were the absolute values of the standard Sobel features, and we fixed $\chi = \gamma = 2$ and $\omega = 3$ in (4). The center views for both input images show the edge images made by applying $b_{TS4}$ to the input images with $\tau = 0.25$, while the right views are made by changing this one parameter of the TS4 blending function to $\tau = 2.25$. This shows that different levels of structural detail can be accomplished by changing a single parameter of the TS4 edge detector. Although the images were not normalized to account for scanner differences, the effect of variation in $\tau$ is similar in the upper and lower sets of images.
Figure 5. Effect of $\tau$ on images for normal (upper) and abnormal (lower) patients

Note the level of enhancement and substructure differentiation in Figure 5. The lower value of $\tau$ enhances the skin line (the nipple in the upper central view, for example, which cannot be seen in the input image), while the higher value enhances the internal breast area. $b_{TS4}$ can be used to narrow the search for breast cancer indicators. For example, for masses, the characterization of the margins can provide an indication of malignancy and certain settings of $\tau$ seem to help detect such margins. In addition, malignancy can be indicated by edge distance and intensity variation measures around the lesion. We are studying optimization of $b_{TS4}$ for these purposes. Further, we are studying the possibility of practicing radiologists varying $b_{TS4}$ parameters (especially $\tau$) on-line to select the best (most useful) output(s) for a particular image.
6.2. Segmentation of breast area

FCM is well known for its ability to provide reliable segmentations of many kinds of images\(^3\), but has not, to our knowledge, been used in the mammography domain. We applied FCM clustering to the normal and abnormal input images shown in the left views of Figure 5. For these runs the input variable was just pixel intensity, the number of clusters was prespecified as \(c = 5\), the FCM weighting exponent was fixed as \(m = 2\), and the FCM objective function used the Euclidean Norm. Figure 6 shows artificially colored segmentations of these input images made by hardening the terminal partitions found by FCM. These results show the influence of the digitization process and the importance of normalizing the images to account for scanner differences prior to comparison of outputs. The image on the left was scanned on a DBA scanner (logarithmic response to OD) while the image on the right was scanned on a Lumisys scanner (linear response to OD). Changes in pixel intensity values do not have the same meaning in these two scanners.

![Figure 6. FCM clustering (c = 5) of images from normal and abnormal patients](image)

7. CONCLUSIONS

The impact of automated computer interpretation systems which preprocess mammogram images is well established\(^28\). Image processing steps such as segmentation and edge detection lead to image enhancements that will significantly ease the workload of practicing radiologists. The underlying premise of the proposed work is that fuzzy models have been successfully applied to a wide variety of other medical applications (e.g., brain tumor detection and volume estimation from magnetic resonance imagery) with excellent results. However, each data domain and application has unique peculiarities which necessitate adjustments to or even replacement of existing methods. Our first experiments indicate the potential success of applying fuzzy models to images from the DDSM database. As this database is expanded to reach its goal of 3000 cases, we will continue to test the generalization capabilities of our approach on additional normal and abnormal cases.

There are a number of additional avenues we are pursuing in this project. First, we are investigating the integration of multiple cues for substructure / skin line extraction. To reduce the set of plausible significant edges, outputs from multiple fuzzy models can be integrated to essentially cast a weighted vote for "significant" areas within the breast substructure. In this area we are investigating the coordinated use of FCM and TS4 to locate the air-skin interface and guide region segmentation of the breast area. Two major problems that must be solved when using any clustering algorithm for segmentation are how to choose \(c\), the "best" number of clusters; and how to assign physical labels to the pixels at the termination of processing. We plan to study both problems.
In addition to affecting the quality of enhancement (contrast and structural detail) in the final image, the modification of resolution has important consequences for the time and space complexity of subsequent processing stages. We are examining the impact of varying both spatial and grey level resolution on system resource use and the quality of outputs from the TS4 and FCM algorithms.

We are also in the process of developing and testing instruments (rating scales) which two consulting radiologists will use independently to evaluate outputs of the system. Each radiologist will be asked to judge the enhancement quality of the breast skin line and breast substructure, the sufficiency of the detail level within and outside these areas, and the level of differentiation between normal and abnormal tissues. The classification of suspicious areas in each image will be compared quantitatively to ground truth patient information, in terms of the number of false positives and false negatives.

To factor in the influence of the film to digitization process, we will present the results of our system to the radiologists for their comparisons as in Figure 1 using three devices: (1) printouts only; (2) computer monitor displays only and (3) original films with printouts and/or computer monitor displays. We are also developing test cases where the radiologist is allowed to vary system parameters such as \( \tau \) in TS4 on-line.

8. ACKNOWLEDGMENTS

We would like to thank Suzanne Wooldridge, M.D. and Patricia Griffith, M.D. of the Ann L. Baroco Center for Women’s Health at Sacred Heart Women’s Hospital for their participation in this research. In addition, we would like to express our gratitude to the University of South Florida for providing access to and assistance with the DDSM and associated software.

9. REFERENCES


25. Mammographic Image Analysis Society (MIAS), Digital Mammogram Database. Published electronically, 1994. Electronic contact addresses as of October 1997 are mias@svl.smb.man.ac.uk for E-mail and http://s20c.smb.man.ac.uk/services/MIAS/MIAScom.html for World Wide Web.

26. Nijmegen Digital Mammogram Database. Electronic contact addresses as of October 1997 are nico@mbfys.kun.nl for E-mail and http://marathon.csee.usf.edu/Mammography/Nijmegen/README.1 for World Wide Web.
