

SAND--96-8498C  
CONF-960219--1

# Detecting Circumscribed Lesions with the Hough Transform

RECEIVED  
MAR 01 1996  
OSTI

Bennett R. Groshong

W. Philip Kegelmeyer, Jr.

Sandia National Laboratories, Livermore, CA 94551-0969

January 11, 1996

## ABSTRACT

We have designed and implemented a circumscribed lesion detection algorithm, based on the Hough Transform, which will detect zero or more approximately circular structures in a mammogram over a range of radii from a few pixels to nearly the size of the breast. We address the geometrical behavior of peaks in Hough parameter space  $\{x, y, r\}$  for both the true radius of a circular structure in the image ( $r = r_0$ ), and for the parameter  $r$  as it passes through this radius. In addition, we evaluate peaks in Hough parameter space by re-analyzing the underlying mammogram in the vicinity of the circular disk indicated by the peak. Discs suggested by the resulting peaks are accumulated in a feature image, scaled by a measure of their quality. These results are then rectified with respect to image contrast extremes and average value. The result is a feature with a continuously scaled pixel level output which suggests the likelihood that a pixel is located inside a circular structure, irrespective of the radius of the structure and overall mammogram contrast. These features are evaluated fast qualitative and quantitative performance metrics which permit circumscribed lesion detection features to be initially evaluated without a full end-to-end classification experiment.

## 1 INTRODUCTION

This paper reports on progress in a larger effort to develop an integrated approach for the detection of abnormalities in screening mammograms. The end goal is a computer tool which may be used as a second reader of mammograms, providing a second opinion to a screening radiologist, pointing out to them areas worthy of their focused attention.

There has been a fair amount of work devoted to this general purpose.<sup>1</sup> However, there are many pertinent mammographic signs of breast cancer (spiculated lesions, circumscribed lesions, calcification clusters, asymmetry, and others), but little published work to date on the integration of the detection of *all* such signs. Further, a common approach in much of this work is to put most of the algorithmic effort into the low-level detection of the abnormalities, and then to use heuristics to interpret the features.

Therefore, a method which naturally integrates disparate cancer signs and which employs a rigorous pattern recognition methodology would be of value. Per-pixel feature extraction processed by binary decision trees<sup>2</sup> (BDTs) to create probability image outputs is such a method, and one which we explore here, reporting on the initial design of features pertinent to circumscribed lesion detection.

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, 175 Oak Ridge Turnpike, Oak Ridge, TN 37831; prices available at (615) 576-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161; phone orders accepted at (703) 487-4650.

#### **DISCLAIMER**

**Portions of this document may be illegible  
in electronic image products. Images are  
produced from the best available original  
document.**

## 1.1 A Review Of Related Work

The earliest work in circumscribed lesion analysis by computer was concerned with classification, rather than detection.<sup>3</sup> Classification of lesions (in the hopes of reducing the rate of false biopsies) remains a popular area of interest.<sup>4,5</sup>

*Detection* of such lesions is the focus of the current paper, however, and there has been a wide variety of approaches to this problem as well.<sup>6-12,5</sup> Most of this work has been “region-oriented”. That is, the basic organization of the method is to generate regions of interest by one technique or another, and then analyze features of those regions to decide whether or not to report them as potential lesions.

One investigation,<sup>11</sup> though avowedly region-oriented, is related to the current work in that they also use BDTs with probabilistic outputs to analyze the features generated by each candidate region. Another approach<sup>5</sup> begins in an essentially pixel-oriented fashion, using correlation analysis at every pixel to detect candidate matches against a matched filter, but then groups these candidates into regions and proceeds to do all subsequent analysis on a region basis.

There has been, however, at least one other effort taking a purely pixel-oriented approach, inspired by the close theoretical connections between image compression and image classification.<sup>12</sup> In this case, though, the pixel “features” were simple 2 by 2 blocks of pixels, and so were not designed to specifically respond to either lesions or normal pixels. As a result, the false alarm rate was very high; 30% of the normal pixels in a typical result image were reported as being lesion pixels.

## 1.2 A Review of the Dense Feature Map (DFM) Approach

We have previously described a method for the automatic detection of spiculated lesions in digitized mammograms,<sup>13</sup> reported on an observer study employing this process in a second reader scenario,<sup>14</sup> and have extended the method to detection of calcification clusters.<sup>15</sup>

The technique extracts image features at *every* pixel from a set of training images, pairs them with “truth” images to grow binary decision trees, and uses those trees to label every pixel of new mammograms with the probability of their being located on an abnormality. The effect, as illustrated in Figure 1, is to expand each mammogram into a stack of feature images which is then compressed back into a single “probability of suspiciousness” image, with values on [0, 1], by the binary decision tree classifier. After each pixel is independently labeled, a spatial filtering step is applied to extract a spatial consensus on the presence or absence of a lesion or calcification cluster.

Note that this high-level description of the method for detection of breast abnormalities does not depend on the precise nature of the abnormality of interest. In previous work, we have selected features which respond well to the image signature of spiculated lesions or calcifications. As this is a binary (suspicious/nonsuspicious) detection problem, features which correlate with abnormalities *and* features which correlate with normal tissue are useful. Therefore we also use general texture characterization features<sup>14,16</sup> whose purpose is to model the broad range of *normal* breast tissue patterns, and so suppress false alarms.

In the current paper, we report on the development of pixel-level features that will respond to circumscribed lesions, permitting the same overall methodology to apply to their detection as well.

## 1.3 The Circumscribed Lesion Model

We model a circumscribed lesion as a roughly circular shape with a well-defined edge around much, but not necessarily all, of its circumference. We also expect the interior of the lesion to be either darker (radiolucent) or lighter (radiopaque) than the surrounding tissue.<sup>17</sup> Finally, circumscribed lesions can vary dramatically in size, with radii from 1 to 30 millimeters.

Further, the context in which the circumscribed lesions are to be detected is one containing a great deal of structured

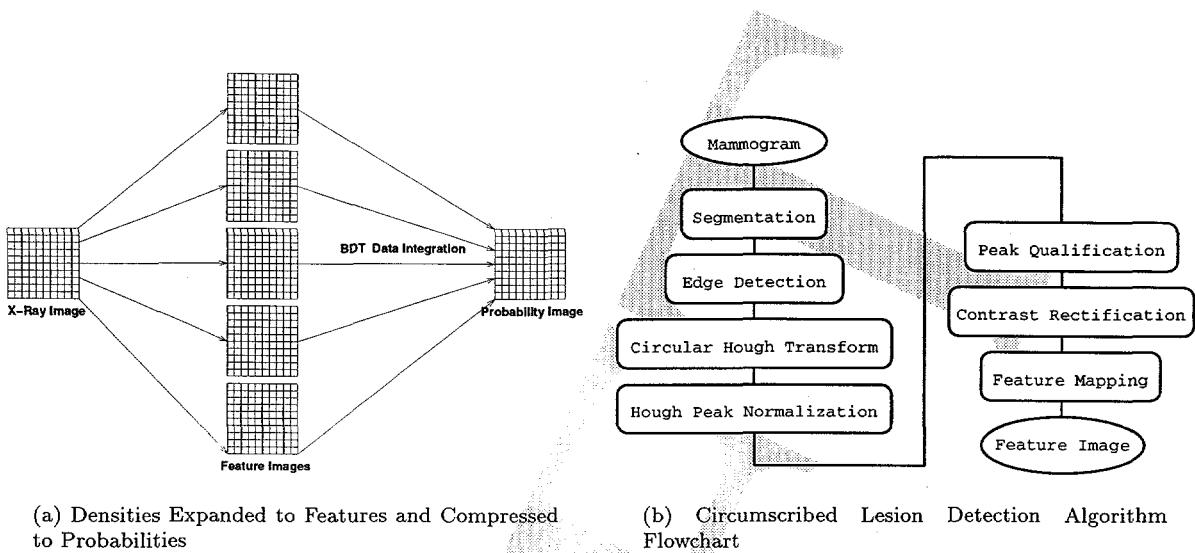


Figure 1: Classification and Detection Algorithms

noise (due to the complex appearance of normal parenchyma) and wide ranges in brightness (due to the possible presence of fatty and/or dense tissue).

Accordingly, the situation calls for an detection approach which is successful despite partial or distorted shapes, robust in the face of structured noise, and capable of generalizing to a wide range of scales. All of these considerations motivated the development of a method based on the circle Hough Transform, as will be discussed in detail in Section 4. First, however, we will describe other approaches to the problem, our own test data, and methods for evaluating the usefulness of a given circumscribed lesion feature.

## 2 TEST DATA

### 2.1 Data Description

The data analyzed in this investigation is a subset of version 1.2 of the digital mammographic database created and distributed by the Mammographic Image Analysis Society (MIAS)<sup>1</sup>. These are single-view mammograms scanned with the Joyce-Loebl microdensitometer SCANDIG-3, which has a linear response in the optical density range 0–3.2. The resulting pixels have a spatial resolution of 50 microns and a grey-scale resolution of 8 bits.

The MIAS database characterizes each image according to the class of abnormalities present. There are 23 images that contain circumscribed lesions. One of these, however (mdb0591s), has no associated groundtruth information, and so it was impossible to objectively determine its location or extent; accordingly, it was not included in our analysis. This results in 22 images containing 24 lesions, as two of the images contain two lesions apiece.

To match the 22 circumscribed lesion images, 22 entirely normal images were selected at random from the set of 204 normal images in the MIAS database, resulting in a data subset containing 44 images. The names of the images selected are listed in Table 1.

<sup>1</sup>See the WWW site <http://skye.icr.ac.uk/miasdb/miasdb.html> for more information.

Lesion Images	001lm	002rl	005ll	010rm	012rl	015lm	017ls	019ll	021ll	023ll	025ll
	028rl	069ll	080rm	091lm	132rx	141lx	142rx	244rm	270rm	290rs	315ll
Normal Images	036rs	041ll	047lm	066rm	068rl	085lm	094rm	098rl	114rs	122rl	138rl
	143lx	154rx	173ll	180rs	229ll	230rl	235ll	279ll	296rl	297ll	319ll

Table 1: Selected Subset of the MIAS Database

## 2.2 Groundtruth Data

Version 1.2 of the MIAS database includes information (“groundtruth”) as to the location and extent of each abnormality in the database. This information is encoded as circles which indicate the approximate center and radius of the abnormality.

As circumscribed lesions are rarely perfectly circular, and as the MIAS database had a policy of erring on the side of making the groundtruth circles completely inclusive rather than too small, the groundtruth regions often contain a substantial amount of normal tissue. Therefore, for the subset of lesions in which the border was both clearly visible and substantially different from the original groundtruth circle, the truth regions were edited to more closely reflect the actual lesion shape.

These edited truth regions are used when training classifiers and when applying the performance metrics discussed in Section 3. However, the original truth regions are and will be used whenever end-to-end detection rates are computed, in order to permit accurate comparison with results from other institutions making use of the same data.

## 2.3 Pre-Processing

At 50 microns, the original data is at a resolution finer than practical for initial algorithm investigation. Accordingly, it was subsampled by a factor of 16, using Gaussian/sinc interpolation,<sup>18</sup> to create data with a pixel resolution of 800 microns. At this spatial resolution, the circumscribed lesions in the MIAS database have a radius that varies from 4.5 to 49 pixels. (Determining whether higher resolution processing is strictly necessary is one of the future aims of this work.)

After subsampling, the images were edited by hand to remove artifacts from the background (non-tissue) regions. They were segmented by a process that thresholds the image at a level dynamically chosen to be just higher than the background histogram peak, and then finds and retains the largest resulting 4-connected object. The rest of the objects are merged into the background, and the background value is set everywhere to zero. The result is an image in which the breast region is the only region that contains non-zero values.

## 3 PERFORMANCE ANALYSIS METHODS

The end goal of feature development in the DFM approach is to come up with features that will suggest whether a given pixel is, or is not, located within a circumscribed lesion. As illustrated in Figure 1, these features are analyzed by a binary decision tree (BDT) process, first to grow a tree on pixels with known classifications, and then to classify new, unknown mammogram pixels.

The BDT is very fast when classifying new feature vectors, as in application it is simply a small number of threshold comparisons. So it is well suited for the classification of large mammogram images. The BDT’s matching disadvantage, however, is that it is very time-consuming to generate the classification tree in the first place. It is not an iterative process, so it completes within a predictable amount of time, but for any training set of reasonable size it can be very slow.

In practice, with a finished and polished set of pixel-level features, this is not a limitation, as the tree needs to be grown only once. But in the context of feature discovery and exploration, the full BDT process has too lengthy a cycle time to permit its application to every set of feature variations.

Accordingly we have implemented two intermediate feature performance measures, one qualitative and one quantitative, to aid in the fast evaluation of candidate features. Both hinge on the existence of groundtruth information (see Section 2.2), as both are based on the *degree of separation* of a feature between the lesion areas and the background areas.

That is, a frequent approach to classification of features is simple thresholding of a single feature. This generally does not suffice, as in a sufficiently rich feature set there will be *combinations* of features who can have values indicative of the target class, where any single feature would not suffice. It is precisely this fact that gives the BDT approach its utility. Still, if a feature *does* well separate the target classes (in this case, lesion pixels and normal tissue pixels), then it will perform well for any classifier. In other words, feature separation is a sufficient but not necessary criteria for a good feature. Accordingly, both of the fast evaluation methods to be described are based on feature separation.

### 3.1 Separation Plots

One fast, though crude, approach is to determine the mean and standard deviation of the feature over the lesion area and over the background. If the spread of the means is large with respect to their standard deviations, then the feature well separates the two classes.

In the current context, the same classifier is going to be applied to all images. Thus it is important to note not only mean values and spread within a given image, but also how the values change from image to image. If, for instance, the low values in the lesion area, as computed across the entire image set, markedly overlap the high values of the normal area, then this feature will confuse a classifier, even if there is never any overlap in any given image.

These behaviors can be examined qualitatively with separation plots. The source data is the mean and standard deviation of the normal tissue and the lesion tissue (when present), computed across all 44 images in the current data set. These values are plotted on a single graph, one curve following the lesion means and one following the normal tissue means, with the error bars indicating the magnitude of the standard deviation in each image. An example is in Figure 5. A promising feature would be one where the lowest point on the lesion curve is well-separated from the highest point on the normal tissue curve, regardless of the images in which they occur.

### 3.2 t-statistic Analysis

The separation plot is a qualitative means of assessing the degree of separation. A more quantitative approach is to consider the lesion and normal tissue feature values to be sets of samples from two distributions, and to ask whether they are *different* distributions. If so, then it should be possible to construct a separation test (via an automatic process, in the case of the BDT), and so this feature is likely to be a good one.

A standard approach to this question in the context of Gaussian distributions is the Two-Sample Student  $t$  test. Though there is no a priori reason to expect the circumscribed lesion features to obey a Gaussian distribution, nonetheless the  $t$  test is appropriate, as it is particularly robust even when the underlying distributions are not Gaussian, and becomes more so with large sample sizes,<sup>2</sup> which we certainly have here. Accordingly, we have implemented this test, using the groundtruth data to define the regions of the lesion and normal tissue populations. (The non-tissue part of the mammogram is not considered, either here or in the separation plots above.)

The size of the  $t$  statistic is used to determine whether the claim that the distributions are the same can be rejected at some level of confidence. The magnitude of the  $t$  statistic also roughly indicates the degree of separation, and so may be used as a quick quantitative measure of the usefulness of a candidate feature, across an image or across a set of images.

## 4 LESION DETECTION METHODS

Our method for detecting circumscribed lesions in mammograms is built around the circular Hough<sup>19</sup> transform. The algorithm we use is composed of a series of steps as outlined in Figure 1b. The input is a mammogram as described in Section 2, and the end result is a feature image that correlates with the likelihood of a circumscribed lesion at each pixel location in the input mammogram. We segment the input mammogram as described in Section 2.3, detect intensity edges, perform the circular Hough transform, filter the Hough accumulator to normalize with respect to radius, qualify peaks in the accumulator using the contrast in the mammogram, and finally map the qualified peaks into a feature image. We describe these algorithm steps in more detail in the following Sections (4.1–4.6)

### 4.1 Edge Detection

Our objective for edge detection is a sensitive edge indicator, combined with accurate normal directions, as these are required by the HT. We use a simple algorithm based on that of Canny.<sup>20</sup> We filter the input image with a 2-d symmetrical Gaussian ( $\sigma_x = \sigma_y = 1.0$ ), then estimate the image gradient using first differences taken in the image  $X$  and  $Y$  directions, then compute the magnitude and phase angle of the gradient, and finally suppress all points but those which are a local maxima in the direction of the gradient.

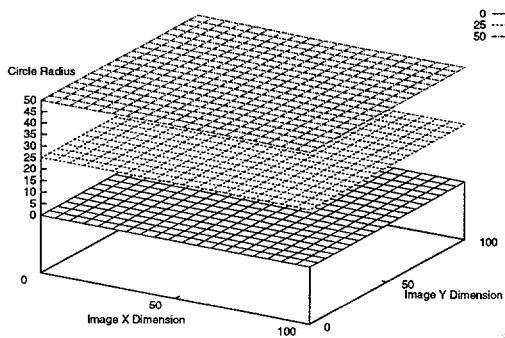
We experimented with a larger Gaussian ( $\sigma_x = \sigma_y = 2.0$ ), which increased the accuracy of the gradient phase angle dramatically for smaller radii circles, but failed to improve end-to-end detection performance noticeably. We also experimented with a directional Canny<sup>20</sup> using 8 oriented filters ( $\sigma_x = 1.0, \sigma_y = 3.0$ ) with similar results. We attribute this lack of improved detection performance to the quantity of background structure cluttering in the mammograms.

### 4.2 Circular Hough Transform

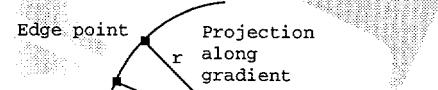
We chose the Circular Hough<sup>21–23</sup> transform as a detection algorithm for its performance over a wide radius range in the face of noise and background clutter, both prevalent features in our mammogram data set. We chose to parameterize the Hough accumulator for circles for two reasons: first, circles are a reasonable first order approximation for circumscribed lesions, and second, to limit the dimensionality of the search for peaks in the accumulator. Figure 2a schematically illustrates the circular Hough transform accumulator, which is parameterized at 1 unit in  $x$ ,  $y$ , and  $r$ , where  $\{x, y\}$  indicate the origin of a circle and  $r$  indicates its radius. Edge points are projected in the direction normal to the rising edge as shown in Figure 2b. Figure 2c illustrates the projection of a gradient local maxima at  $\{40, 20\}$  with direction  $1, 0$  through the Hough accumulator  $\{x, y, r\}$ , incrementing each point along the line  $\{40 + r, 20, r\}$ . Similarly, a maxima at  $\{60, 20\}$  with direction  $\{-1, 0\}$  increments each point in the accumulator along the line  $\{60 - r, 20, r\}$ . Edge points arranged in a circle will thus result in a cluster of values about the point  $50, 20, 10$  in the Hough accumulator where the projected lines converge. The result is a volume of points with higher densities indicating the possibility of circular structure in the original image, as seen in Figure 2d.

### 4.3 Hough Peak Normalization

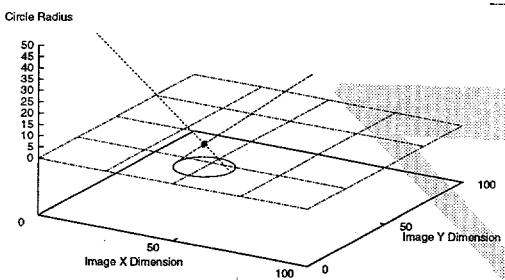
We desire an indication of circular structure in the input mammogram which is independent of radius, as we will be mapping peaks in the 3-d Hough accumulator down to the 2-d pixel level feature image (Section 4.6) to indicate the significance of a detected lesion, irrespective of its radius. Unfortunately, the height and density of peaks in the raw Hough accumulator are affected by the number of edge points contributing to a peak and by the accuracy of the edge location and gradient normal. The total number of edge points contributing to an accumulator peak for a circle is  $\propto 2\pi r$ , but the spatial distribution of the edge points projected along the gradient normal is affected in a complex, deterministic way by the pixel sampling of the circle border, which is both a function of circle radius and location. This results in peak heights



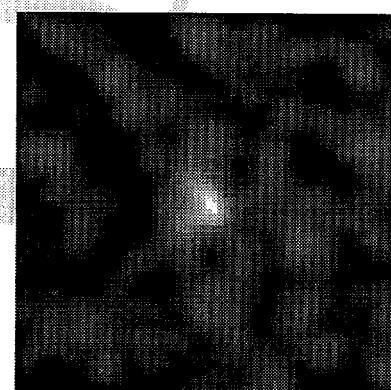
(a) Schematic of Hough accumulator



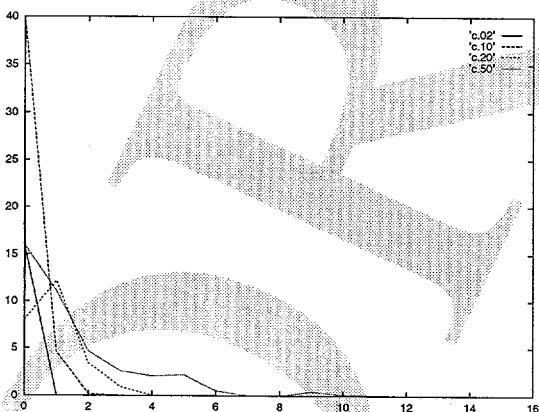
(b) Edge Gradient Projection (XY Plane View)



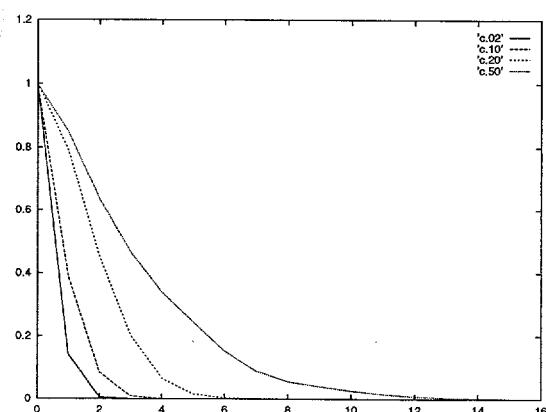
(c) Gradient Projection for circle  $\{50, 20, 10\}$



(d) Smoothed Hough accumulator  
lesion subimage



(e) Raw accumulator radial profile



(f) Normalized accumulator radial profile

Figure 2: Circular Hough Transform Accumulator

and densities that vary considerably with respect to radius, even for images of perfectly circular structures. Figure ??a is a plot of peak height versus radial distance from the peak center for circular disks of radius 2, 10, 20, and 50. We compensate for this behavior by normalizing the Hough accumulator with respect to  $r$  using an empirically designed exponential filter matched to the peak shapes created by circular disks. The filtering operation is the convolution

$$g(x, y, r) = f(x, y, r) * h(x, y, r) \quad (1)$$

of the raw Hough accumulator  $f(x, y, r)$  with the exponential normalizing filter

$$h(x, y, r) = c(r) \exp \left( -\frac{(x^2 + y^2)^{a(r)/2}}{b(r)} \right) \quad (2)$$

resulting in normalized Hough accumulator  $g(x, y, r)$ . The terms  $a(r)$ ,  $b(r)$ , and  $c(r)$  are weights as a function of the circle radius  $r$ , and were empirically determined for each  $r$  by summing the Hough accumulator response for a large number of circular disks displaced uniformly over the interval  $x_0 \pm 0.5, y_0 \pm 0.5$ . Figures 2e-f illustrate the effect of normalizing the hough accumulator for circles of varying radius.

#### 4.4 Hough Peak Qualification

Peaks in the Hough transform accumulator described above indicate circular structures in the mammogram. Unfortunately, many Hough peaks of similar height are created, due to both interfering structure and to lesion deviation from a circle. The Hough transform only measures the outline of a structure, as indicated by edge points on its periphery. Our approach to improving this performance is to measure the image contrast between the interior and exterior of the circle described by a peak in Hough space by computing the mean value of the interior of the circular region and the mean of an annulus exterior to the circle as shown in Figure 3a. We then use this contrast to scale the Hough peak in a multiplicative fashion, resulting in a measure of the region contrast, as well as the support for a circle.

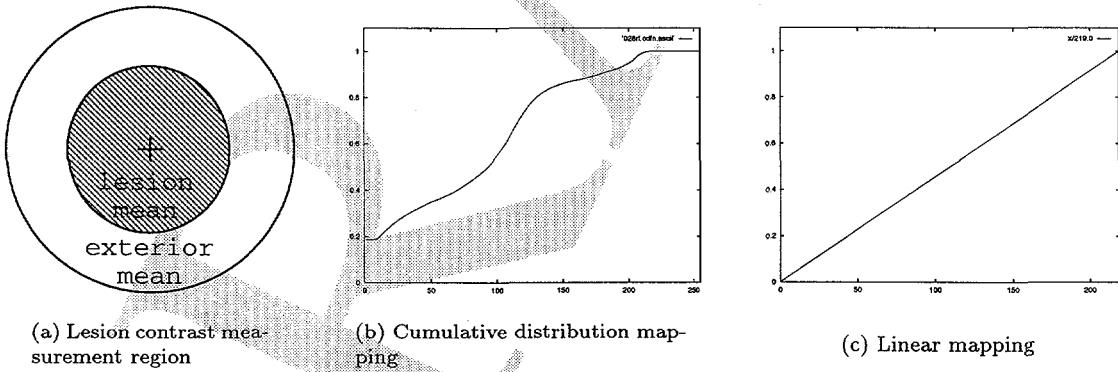


Figure 3: Lesion contrast schematic and image mapping functions

#### 4.5 Contrast Rectification

We address image to image variability in contrast by mapping the image brightness range  $[\min - \max]$  to  $[0 - 1]$  through either a linear or cumulative distribution function, as illustrated in Figure 3b and c. This results in a volume of points  $\{x, y, r\}$  describing the relative likelihood of a dense circular structure with radius  $r$  at a point  $\{x, y\}$ .

## 4.6 Feature Mapping

The last step in our algorithm is to create a feature image indicating the relative likelihood of a circumscribed lesion at a pixel level by mapping points in the qualified, rectified Hough image in  $x, y, r$  space onto disks of radius  $r$  on the image plane at  $x, y$ . We experimented with two different methods of collapsing the vector of  $r$  possible Hough space values onto the feature image plane:

**Peak disk:** Setting each feature image pixel to the maximum likelihood encountered over  $r$  at the point  $x, y$ .

**Sum disk:** Setting each feature image pixel to the sum of the likelihoods encountered over  $r$  at the point  $x, y$ .

Figure 4a-c illustrates this mapping for the “028rl” image. Figure 4 illustrates the “Peak disk” and “Sum disk” features (Section 4.6). for selected mammograms, as well as the original mammograms. Figures 5a and b are plots of the “Peak disk” and “Sum disk” features without contrast rectification (Section 4.5) for all 44 mammograms in the data set. Figures 5c and d illustrate the linear map contrast rectification over the set and Figures 5e and f illustrate the cumulative distribution map contrast rectification.

## 5 CONCLUSIONS

Our intent is to extend the dense feature map method to the detection of circumscribed lesions in screening mammograms. In prior work we have established the pattern recognition infrastructure and the means for modeling normal breast tissues (via the Laws texture features).

The current effort, then, has been devoted to remaining piece of the puzzle, that is, the development of pixel-level features which suggest, for each pixel, the likelihood that that particular pixel is located within a circumscribed lesion. This effort is not complete; at the least, a full end-to-end detection performance evaluation, represented as a free-response receiver-operator characteristic (FROC) curve would be required to assess the merit of these features. We have demonstrated, however, the value of these features as assessed by faster (though less sensitive) qualitative and quantitative measures.

Future effort will focus on pre-processing to reduce the structural noise in mammograms and to improve the assessment of contrast, analysis of the Hough space to sharpen the distinction between true and false peaks, and full FROC performance evaluation.

## 6 REFERENCES

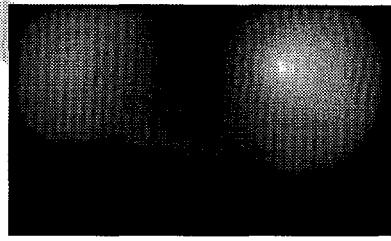
- [1] Carl J. Vyborny and Maryelle L. Giger. Computer vision and artificial intelligence in mammography. *AJR*, 162:699–708, 1994.
- [2] Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. *Classification and Regression Trees*. Wadsworth International Group, Belmont, California, 1984.
- [3] Laurens V. Ackerman and Earl E. Gose. Breast lesion classification by computer and xeroradiograph. *Cancer*, 30:1025–1035, 1972.
- [4] Datong Wei, Heang-Ping Chan, Mark A. Helvie, Berkman Sahiner, Nicholas Petrick, Dorit D. Adler, and Mitchell M. Goodsitt. Classification of mass and normal breast tissue on digital mammograms: Multiresolution texture analysis. *Medical Physics*, 22(9):1501–1513, September 1995.
- [5] Shun Leung Ng and Walter F. Bischof. Automated detection and classification of breast tumors. *Computers and Biomedical Research*, 25:218–237, 1992.



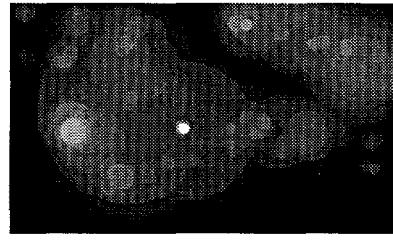
(a) Peak disk feature for "028rl"



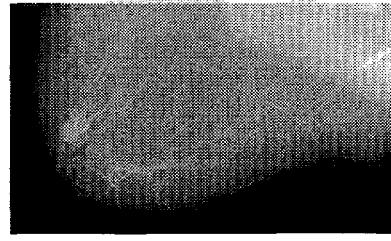
(b) Segmented mammogram "028rl"



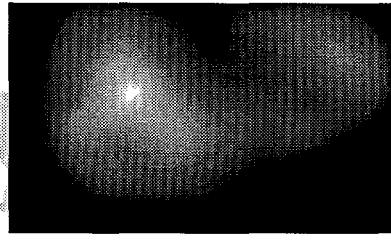
(c) Summed disk feature for "028rl"



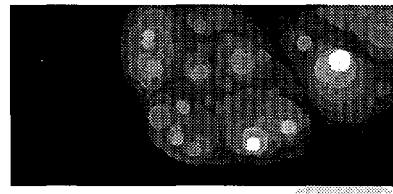
(d) Peak disk feature for "005ll" image



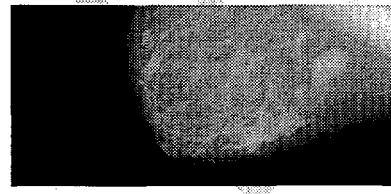
(e) Original "005ll" mammogram



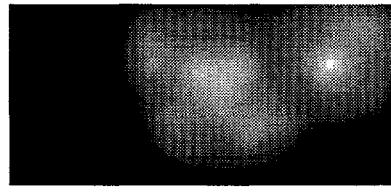
(f) Summed disk feature for "005ll" image



(g) Peak disk feature for "015lm" image



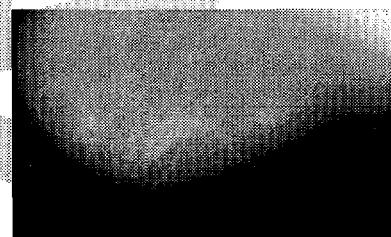
(h) Original "015lm" mammogram



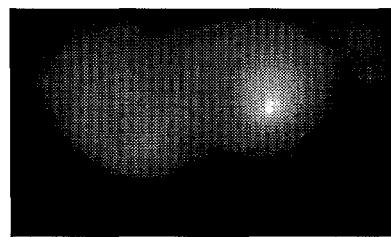
(i) Summed disk feature for "015lm" image



(j) Peak disk feature for "023ll" image

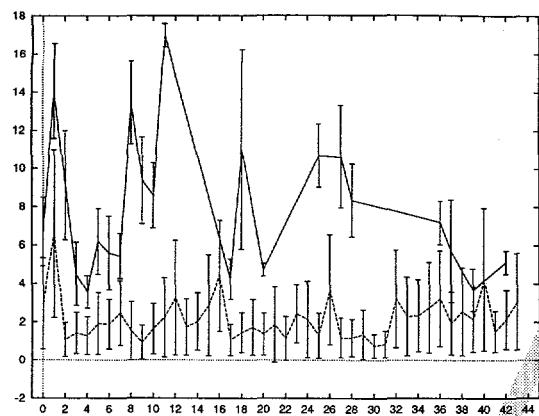


(k) Original "023ll" mammogram

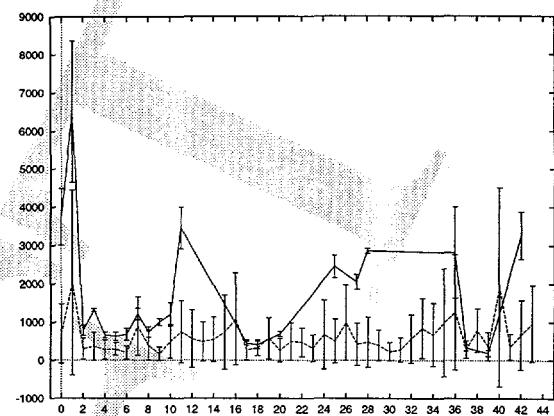


(l) Summed disk feature for "023ll" image

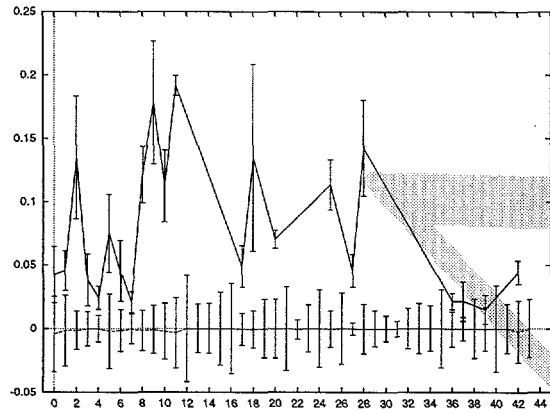
Figure 4: Peak, original, and summed disk feature images for selected mammograms



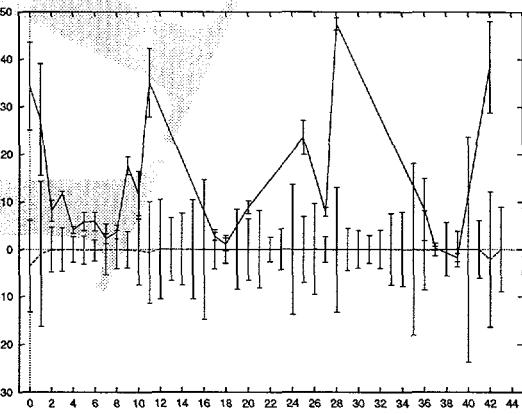
(a) Original peak disk feature plot



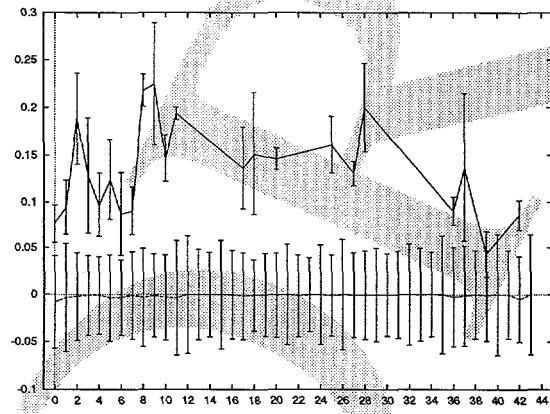
(b) Original summed disk feature plot



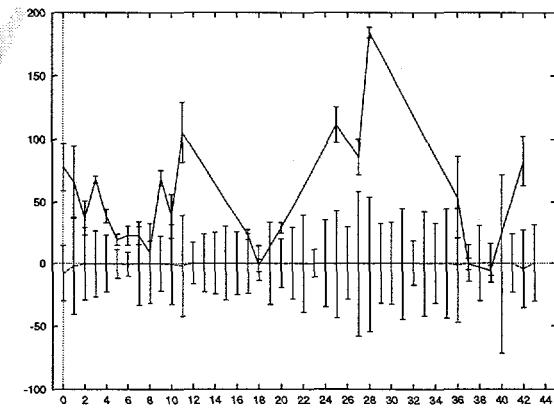
(c) Linear map, peak disk feature plot



(d) Linear map, summed disk feature plot



(e) Cdf map, peak disk feature plot



(f) Cdf map, summed disk feature plot

Figure 5: Plots of features across the 44 mammogram data set

- [6] D. Brzakovic, X.M. Luo, and P. Brzakovic. An approach to automated detection of tumors in mammograms. *IEEE Transactions on Medical Imaging*, 9(3):233–241, September 1990.
- [7] Shuk-Mei Lai, Xiaobo Li, and Walter F. Bischof. On techniques for detecting circumscribed masses in mammograms. *IEEE Transactions on Medical Imaging*, 8(4):377–386, December 1989.
- [8] Maryellen L. Giger, Fang-Fang Yin, Kunio Doi, Charles E. Metz, Robert A. Schmidt, and Carl J. Vyborny. Investigation of methods for computerized detection and analysis of mammographic masses. In *Proceedings of the SPIE Conference on Medical Imaging IV: Image Processing*, volume 1233, pages 183–184. SPIE, 1990.
- [9] Maryellen L. Giger, P. Lu, Z. Huo, U. Bick, C.J. Vyborny, R.A. Schmidt, W. Zhang, C.E. Metz, D. Wolverton, R.M. Nishikawa, W. Zouras, and K. Doi. Cad in digital mammography: Computerized detection and classification of masses. In Alastair G. Gale, Sue M. Astley, David Dance, and Alistair Y. Cairns, editors, *Digital Mammography: Proceedings of the 2nd International Workshop On Digital Mammography*, number 1069 in International Congress Series, pages 281–287, Amsterdam, The Netherlands, 1994. Elsevier Science B.V.
- [10] Fang-Fang Yin, Maryellen L. Giger, Kunio Doi, Carl J. Vyborny, and Robert A. Schmidt. Computerized detection of masses in digital mammograms: Automated alignment of breast masses and its effect on bilateral-subtraction technique. *Medical Physics*, 21(3):445–452, March 1994.
- [11] H.D. Li, M. Kallergi, L.P. Clarke, V.K. Jain, and R.A. Clark. Markov random field for tumor detection in digital mammography. *Transactions on Medical Imaging*, 14(3):565–576, September 1995.
- [12] Cheryl L. Nash, Keren O. Perlmutter, and Robert M. Gray. Evaluation of bayes risk weighted vector quantization with posterior estimation in the detection of lesions in digitized mammograms. In *Twenty-Eighth Asilomar Conference on Signals, Systems, and Computers*, volume 1, pages 716–720. IEEE, IEEE Computer Society Press, October–November 1994.
- [13] W. Philip Kegelmeyer, Jr. Evaluation of stellate lesion detection in a standard mammogram data set. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(6):1477–1492, December 1993.
- [14] W. Philip Kegelmeyer, Jr., Joe M. Pruneda, Philip D. Bourland, Argye Hillis, Mark W. Riggs, and Michael Nipper. Computer-aided mammographic screening for spiculated lesions. *Radiology*, 191(2):331–337, May 1994.
- [15] W. Philip Kegelmeyer, Jr. and Mark C. Allmen. Dense feature maps for detection of calcifications. In Alastair G. Gale, Sue M. Astley, David R. Dance, and Alistair Y. Cairns, editors, *Digital Mammography: Proceedings of the Second International Workshop on Digital Mammography, York, England*, pages 3–12. Elsevier Science B.V., 1994.
- [16] K. I. Laws. *Textured Image Segmentation*. PhD thesis, University of Southern California, 1980.
- [17] Laszlo Taber and Peter B. Dean. *Teaching Atlas of Mammography*. Georg Thieme Verlag, Thieme Inc., Stuttgart, New York, 2nd revised edition, 1985.
- [18] E. Karabassis and M. E. Spetsakis. An analysis of image interpolation, differentiation, and reduction using local polynomial fits. *Graphical Models and Image Processing*, 57(1):183–195, May 1995.
- [19] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, Jan 1972.
- [20] John F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–697, November 1986.
- [21] C. Kimme, D. Ballard, and J. Sklansky. Finding circles by an array of accumulators. *Communications of the ACM*, 18(2):120–122, 1975.
- [22] D. C. W. Pao, Hon F. Li, and R. Jayakumar. Shapes recognition using the straight line hough transform: Theory and generalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1076–1089, Nov 1992.
- [23] V. F. Leavers. Which hough transform? *Computer Vision, Graphics, and Image Processing: Image Understanding*, 58(2):250–264, Sep 1993.