Improvement of computerized mass detection on mammograms: Fusion of two-view information

Sophie Paquerault,[a] Nicholas Petrick, Heang-Ping Chan, Berkman Sahiner, and Mark A. Helvie
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-0030

(Received 27 June 2001; accepted for publication 19 November 2001; published 25 January 2002)

Recent clinical studies have proved that computer-aided diagnosis (CAD) systems are helpful for improving lesion detection by radiologists in mammography. However, these systems would be more useful if the false-positive rate is reduced. Current CAD systems generally detect and characterize suspicious abnormal structures in individual mammographic images. Clinical experiences by radiologists indicate that screening with two mammographic views improves the detection accuracy of abnormalities in the breast. It is expected that the fusion of information from different mammographic views will improve the performance of CAD systems. We are developing a two-view matching method that utilizes the geometric locations, and morphological and textural features to correlate objects detected in two different views using a prescreening program. First, a geometrical model is used to predict the search region for an object in a second view from its location in the first view. The distance between the object and the nipple is used to define the search area. After pairing the objects in two views, textural and morphological characteristics of the paired objects are merged and similarity measures are defined. Linear discriminant analysis is then employed to classify each object pair as a true or false mass pair. The resulting object correspondence score is combined with its one-view detection score using a fusion scheme. The fusion information was found to improve the lesion detectability and reduce the number of FPs. In a preliminary study, we used a data set of 169 pairs of cranio-caudal (CC) and mediolateral oblique (MLO) view mammograms. For the detection of malignant masses on current mammograms, the film-based detection sensitivity was found to improve from 62% with a one-view detection scheme to 73% with the new two-view scheme, at a false-positive rate of 1 FP/image. The corresponding cased-based detection sensitivity improved from 77% to 91%. [DOI: 10.1118/1.1446098]

Key words: computer-aided diagnosis, mammography, mass detection, classification, fusion of information

I. INTRODUCTION

X-ray mammography is the only proven diagnostic technique for detecting breast cancer in its early stages.[1,2] In mammographic screening, a cranio-caudal (CC) and a mediolateral oblique (MLO) view are routinely taken for each breast. The two views not only allow most of the breast tissue to be imaged but also improve the chance that a lesion will be seen in at least one of the views. Radiologists analyze the different mammographic views to detect calcifications and masses that may be a sign of breast cancer and to decide whether to call the patient back for further diagnostic evaluations. They also use the two views to reduce false positives such as overlapping dense tissue in one view that mimics masses. Their interpretation integrates complex criteria of human vision and intelligence, including morphology, texture, and geometric location of any suspicious structures of the imaged breast, combining information from different views, checking differences between the two breasts, and looking for changes between the prior and current mammograms when available. Clinical studies indicate that lesion detectability in two-view mammograms is more accurate than when only one view is available.[3–5]

It has also been shown that independent double reading by two radiologists significantly increases the sensitivity of mammographic screening.[6,7] However, the increased cost and workload to the radiologists make double reading impractical in most screening situations. To provide a second opinion to the radiologists, computer-aided diagnosis (CAD) systems have been developed using computer vision and pattern recognition techniques to automatically detect and characterize abnormal lesions on mammograms. Although it has been reported that these systems are useful in reducing the error rate in mammographic screening,[8–10] the detection sensitivity of these systems needs to be improved and the false-positive (FP) rate reduced to provide maximum benefit to the radiologist and the patient. CAD algorithms reported in the literature so far use one-view information for the detection of lesions even though the accuracy may be scored and reported using two views. Yin *et al.*[11] used bilateral subtraction in a prescreening step of a mass detection program to locate mass candidates, but the subsequent image analysis was performed based only on a single view. Recently, Hadjiiski *et al.*[12–14] have developed an interval change analysis of masses on current and prior mammograms and found that the classifi-

cation accuracy of malignant and benign masses can be improved significantly in comparison to single image classification. These studies demonstrated the potential of using multiple image information for CAD. However, current CAD algorithms have not utilized one of the most important pieces of information available in a mammographic examination—the correlation of computer-detected lesions between the two standard views. This is a very difficult problem for computer vision because the breast is elastic and deformable. The overlapping tissue and the relative position of the breast structures are generally different even when the breast is compressed in the same view two different times. The change in geometry for an elastic object and lack of invariant ''landmarks'' make it difficult, if not impossible, to correctly register two breast images in the same view by any established image warping technique or by using an analytic model to predict corresponding object locations in the different views of the same breast.

Few studies have been conducted on how to find the relationship between structures in different mammographic views. Kita *et al.*[15] proposed a breast deformation model for compressed breasts and used the model for finding corresponding points in two different views. They demonstrated with a data set of 24 cases (a total of 37 lesions) that this method allowed the prediction of location in a second view within a band of pixels ± 27 mm from an epipolar line. However, assumptions on the parameters and the deformation of a compressed breast had to be made and the robustness of the model has yet to be validated. More practical approaches, which do not depend on a large number of assumptions, may be preferable. Good *et al.* and Chang *et al.* recently reported a preliminary attempt of matching computer-detected objects in two views.[16,17] They demonstrated the feasibility of identifying corresponding objects ($A_z = 0.82$) in the two views by exhaustive pairing of the detected objects and feature classification. None of these studies attempted to use the two-view correspondence information to improve lesion detection or classification.

During mammographic interpretation, if a suspicious breast mass is found in one view, the radiologist will attempt to find the same object in the other available views in order to identify the object as a true or a false mass. Radiologists commonly consider the distance from the nipple to the center of the suspicious lesion in one view and then search the corresponding object in the second view in an annular region at about the same radial distance from the nipple. Based on this approach, we previously developed a regional registration technique to identify corresponding lesion locations on current and prior mammograms of the same view.[13,18] We have also designed geometric models that can localize corresponding lesions within a search region when two-view or three-view mammograms are available for lesion localization.[19] With the geometric information, the computer searches for a corresponding lesion in the other view within a limited search region. The object of interest can then be matched with possible corresponding objects in the search region using the similarity of feature measures. We have found that the geometric constraints improved the chance of correctly matching lesions in current and prior mammograms for the classification of malignant and benign masses.[14] In this study, we explore the use of the regional registration technique as a basis to correlate lesions in two views. The correspondence information is used to reduce false detections produced by our one-view CAD algorithm.[20] The detection accuracy of the two-view scheme was evaluated and compared to our current one-view CAD scheme using free response receiver operating characteristic (FROC) analysis.

II. MATERIALS AND METHODS

Our approach to improving the accuracy of the mass detection is to merge information from corresponding segmented structures in the two standard views of the same breast.[20] We first assume that a true mass will have a higher chance of being detected in both views. Likewise, we assume that the objects corresponding to the same mass detected in the two different views (a TP–TP pair) will be more similar in their feature measures than a mass object compared to normal tissue (a TP–FP pair), or two false-positives (an FP–FP pair). Object matching is performed in two stages. First, all possible pairing of the detected objects on the two views are determined, taking into account geometric constraints. Second, features are extracted from each object, similarity measures for the features pairs are derived, and a classifier is trained to classify true pairs (TP–TP pairs) from false pairs (TP–FP, FP–TP, or FP–FP pairs) using the similarity measures. The two stages are detailed below. The data sets used in the development and evaluation of this approach are described next.

A. Image acquisition and data set

Two data sets of two-view mammograms were collected and separately used to train and test the geometric models and our proposed two-stage information fusion technique. These mammograms were selected from patient files in the Breast Imaging Division at the University of Michigan.

For the geometric modeling of object location on two views, the database consisted of 116 cases with masses, large benign calcifications, or clustered microcalcifications identifiable on both views of the same breast. The mammograms were digitized with a LUMISYS 85 film scanner with a pixel size of 50 μm and 12-bit gray levels. The gray levels were calibrated to be linearly proportional to optical density in the 0.1 to 4.0 O.D. range. The images were reduced to a pixel resolution of 800 μm$\times$800 μm by averaging 16$\times$16 neighboring pixels and down-sampling. For each case, the two standard mammographic views were available. A total of 177 objects were manually selected and marked by an expert radiologist on each of these two views. The nipple location was also identified for each breast image. The radial distance of the selected objects was calculated and the prediction model of an object location in one view from its location in the other view was estimated, as described above.

For the evaluation of the two-view mass detection scheme, a data set of 169 pairs of mammograms containing masses on both the CC and MLO views was used. The mam-

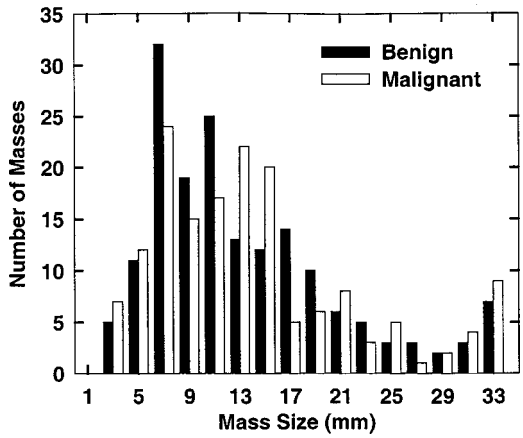

FIG. 1. Histograms of the size (the longest dimension) of the benign and malignant masses contained in the data set of 338 one-view mammograms and rated by a MSQA-radiologist. Eight masses in the prior mammograms of the data set did not receive a rating because the radiologist could not delineate the mass even in retrospect, although a focal density could be seen.

mograms were obtained from 117 patients, of which 128 pairs were current mammograms (defined as mammograms from the exam before biopsy) and 41 pairs were from exams 1 to 4 years prior to biopsy. 58 of the 128 current and 26 of the 41 prior image pairs contained a malignant mass. The 338 mammograms were also digitized with the LUMISYS 85 film scanner. The true mass locations on both views were identified and rated by a radiologist approved by the Mammography Quality Standards Act (MQSA). The histograms of the size (longest dimension) and the subtlety rating of the benign and malignant masses contained in this data set are shown in Figs. 1 and 2, respectively. The subtlety of the masses was estimated subjectively on a 10-point scale by the experienced radiologist relative to the masses encountered in clinical practice.

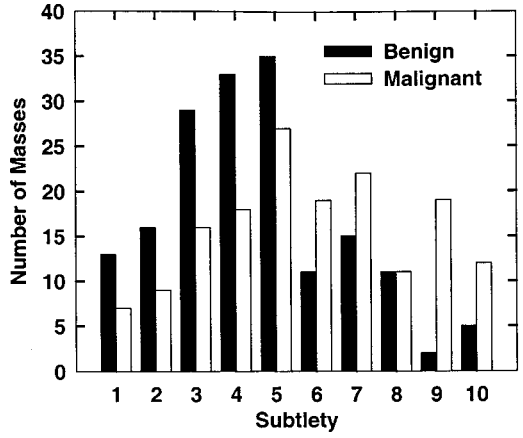

FIG. 2. Histograms of the subtlety (1=most obvious, 10=subtlest) of the benign and malignant masses contained in the data set of 338 one-view mammograms and rated by a MSQA-radiologist. Eight masses in the prior mammograms of the data set did not receive a rating because the radiologist could not delineate the mass even in retrospect, although a focal density could be seen.

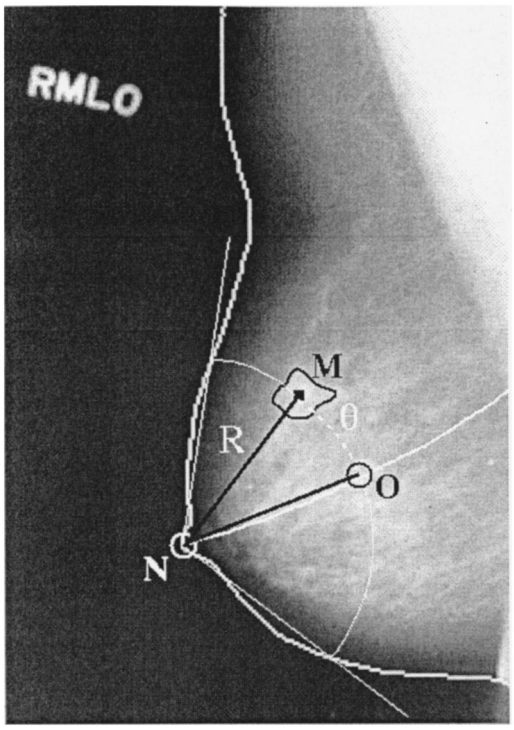

FIG. 3. An example of the coordinate system used to localize an object in a mammographic view. An automatic boundary tracking process is used to segment the breast. The nipple location was identified by a MQSA-approved radiologist. The distance of the object from the nipple location is defined by $R=\|\overrightarrow{NM}\|$. The angle of the mass from the midline of the breast is defined by the angle between the vectors $\overrightarrow{NM}$ and $\overrightarrow{NO}$.

B. Geometrical modeling

We will first describe the geometric models that we developed for predicting the location of an object in the MLO view from that in the CC view or *vice versa*. For the purpose of studying the geometric relationship between the locations of an object imaged on the two mammographic views, any identifiable objects can be used. We therefore chose two-view mammograms that contained masses, microcalcification clusters, and large benign calcifications identifiable on both views. This data set was different from that used for mass detection to be described below. The locations of the corresponding objects on the two views and the nipple locations were identified on the mammograms by the MQSA-approved radiologist. For a large object such as a mass or a microcalcification cluster, the manually identified ''centroid'' was taken as its location. A breast boundary tracking program was used to segment the breast area from the mammogram.[21,22] Using the nipple location as the origin, concentric circles were drawn, each of which intersected the breast boundary at two points and defined an arc. The locus of the mid-points of these arcs was considered to be the breast midline. The breast length was defined as the distance from the nipple to the point where the midline intersected the chest wall. From these parameters, the polar coordinates (R_x, θ_x) with $x=C$ (CC view), or M (MLO view), as shown in Fig. 3, were defined, where R_x was the distance from the nipple to the object center and θ_x, the angle between R_x and

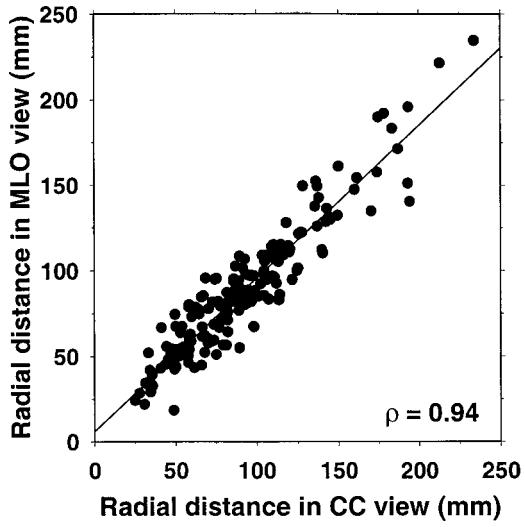

FIG. 4. The CC view versus the MLO view of the radial distances of the identified objects from the nipple location.

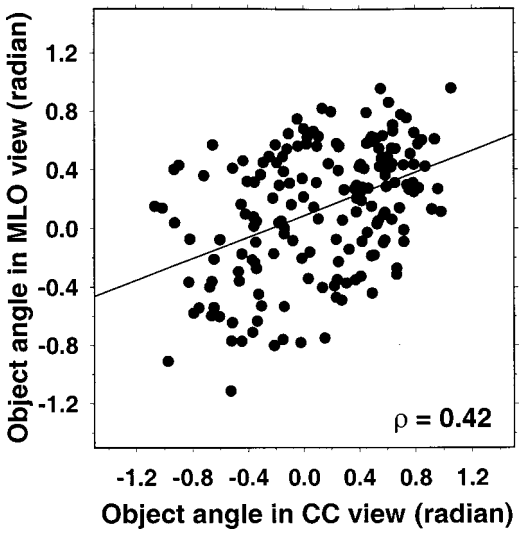

FIG. 5. The CC view versus the MLO view of the angular coordinates of the identified objects from the breast midline.

the line from the nipple to the mid-point of the arc intersecting the object. We investigated the relationship between the coordinate of the object on one view and that on the other view in this coordinate system.

Scatter plots of the radial distance and the angle of the radiologist-identified objects on the two views in the data set are shown in Fig. 4 and Fig. 5, respectively. It can be seen that there is a high correlation (correlation coefficient $=0.94$) of the radial distances of the corresponding objects in the two views. However, the angular coordinates in the two views are basically uncorrelated (correlation coefficient $=0.42$). We therefore chose a linear model for predicting the radial distance of an object in a second view from that in the first view:

$$R_y = a_r \cdot R_x + b_r . \quad (1)$$

Because of the variability of the breast tissue caused by compression, the predicted location for an individual case could deviate from its ''true'' location, as determined by the radiologist, by a wide range. Therefore, we estimated a global model using a set of training cases with radiologist-identified object locations on both views. The model coefficients were obtained by minimizing the mean square error between the true and the predicted coordinates in the second view. The error in this estimation was then used to define an annular search region, which had a center at a radial distance R_y from the nipple as predicted by the model, and a width of $\pm \Delta R$ as estimated from the localization errors observed in the training set. This search region avoids using the entire area of the breast and eliminates many inappropriate pairings between detected objects on the CC view and the MLO view in the second stage, discussed in Sec. II D.

We randomly divided the available data set into a training set and a test set in a 3:1 ratio. The training set was used for the estimation of the model coefficients and the search region width. The test set was used for evaluating the prediction accuracy of the model. Four nonoverlapping partitions separating the database into training and test sets were considered. The model performance was then obtained by combining the results of the four test sets.

The geometrical analysis is then used for pairing objects detected on the two views of the same breast in the prescreening stage of our mass detection program as detailed below.

C. One-view analysis

The one-view approach is used to identify potential breast masses among the suspicious objects. The one-view prescreening used in this study is similar to that discussed previously.[23–25] The only difference is that the false positive (FP) reduction step was modified such that a slightly different object overlap criterion was employed. The block diagram for the one-view mass detection scheme is shown in Fig. 6. A density-weighted contrast-enhancement (DWCE) filter is first applied to each digitized mammogram. The DWCE filter enhances mammographic structures in the breast image. Following this preprocessing filtering, edge detection is employed to refine the borders of the detected regions. *K*-means clustering is then applied to a 25 mm $\times$25 mm, background-corrected region of interest centered on each initially detected object to improve the object border. This segmentation process extracts a large number of objects, including masses and normal breast structures. In order to reduce the number of nonmass objects, different FP reduction stages based on morphological features, overlap of the detected regions, and texture features were designed and trained using an independent set of mammograms in a previous study.[25,26] It was found that 11 morphological features composed of shape descriptors and 15 spatial gray level dependence (SGLD) texture features extracted for each object were useful for FP reduction.[27,28] In this study, rule-based classification using the 11 morphological features reduced the average number of objects from 37 to about 29 per image and lowered the TP detection sensitivity from 91.1% to

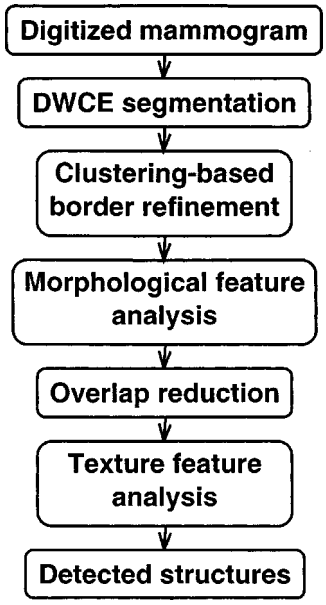

FIG. 6. A schematic diagram for the current one-view prescreening detection algorithm.

87.9% at this stage. The 15 texture features were then used as the input variables for a linear discriminant analysis (LDA) classifier. A texture score for each object was obtained from the classifier. Overlap reduction was then applied using these texture scores as discussed below.

During object segmentation, the border of an object is obtained by *K*-means clustering in a fixed sized region centered on a ''seed'' object. If the seeds from two objects are close to each other, the two segmented objects can overlap each other. This occurs when the two detected objects are neighboring structures that overlap in the mammographic view or they may be part of a large single structure that was initially detected in multiple pieces. An overlap criterion based on the texture scores is imposed to select one of the two overlapping objects as a mass candidate. In this study, we used the shape of the segmented objects to estimate the overlapping area between the two neighboring objects on the mammogram. An overlap fraction was defined as

$$\text{Overlap} = \frac{O_1 \cap O_2}{O_1 \cup O_2}, \tag{2}$$

where O_1 and O_2 are the segmented areas of the overlapping objects. A threshold on the overlap fraction was chosen such that if the overlap fraction of two objects exceeded the threshold, the object with the higher texture score (i.e., more likely to be a mass candidate) was kept and the other was discarded as an FP. The sensitivity and the specificity of differentiating true and false masses depend on the selection of the overlap threshold. We chose an overlap threshold of 15% which led to an average of 15 objects per image at a detection sensitivity of about 85%. As shown later in the Results section, the overall detection accuracy was relatively

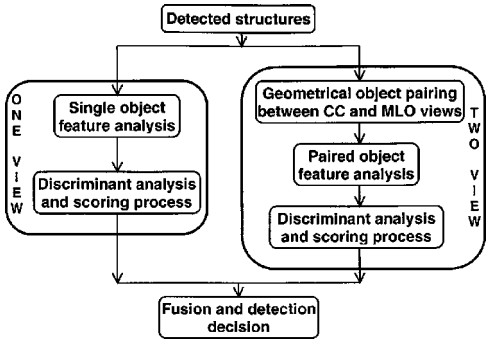

FIG. 7. A schematic diagram for the proposed two-view fusion scheme.

independent of the FP rate in this intermediate stage so that the selection of the 15% overlap threshold was not a critical factor.

After overlap reduction, our current one-view algorithm employed a final stage of FP reduction based on the texture scores, as illustrated in the block diagram in Fig. 7. A decision threshold was applied to the texture scores such that objects with scores lower than the threshold were excluded as FPs. In addition, another criterion was imposed so that no more than three objects were kept on each image. By comparing the retained objects with the true mass locations on each image for a range of decision thresholds, an FROC curve characterizing the sensitivity as a function of the number FPs per image could be generated.[29]

D. Two-view analysis

The block diagram in Fig. 7 illustrates our two-view mass detection scheme and its relationship to our current one-view approach. The detection algorithm described above was used as a prescreening stage in our two-view fusion approach. The only difference was that the operating threshold that limits the maximum number of objects on an image was relaxed to increase sensitivity while retaining a larger number of FPs. The remaining objects after this threshold will be still referred to as the prescreening objects in the following discussions. To investigate the dependence of the overall detection accuracy of our two-view detection scheme on the initial number of prescreening objects, three different decision thresholds were selected to obtain a maximum of either 5, 10, or 15 objects per image.

To further perform the two-view information fusion analysis, an expanded set of morphological features was extracted from each prescreening object. These morphological features included the 11 shape descriptors discussed previously, and 13 new contrast measures[30] and 7 new shape features. In order to evaluate the new method, we randomly divided the available cases into a training and a test set using a 3:1 training/test ratio. The training set was used to select a subset of useful morphological features using stepwise feature selection and to estimate the coefficients of an LDA classifier. To reduce biases in the classifier, 50 random 3:1 partitions of the cases were employed. A morphological score was obtained for each individual object by averaging

the test score of the object obtained from the different partitions. The morphological score was then combined with the one-view texture score by averaging the two scores. A single combined score thus characterized each prescreening object. This one-view score was further fused with the discriminant score obtained by the two-view scheme, as described below.

The prescreening objects were analyzed by the two-view method shown in the right branch of the diagram in Fig. 7. All possible pairing between the prescreening objects in the two views of the same breast was determined using the distance from the nipple to the centroid of each object and the geometrical model described above. Since the location of a given object detected in one view cannot be uniquely identified in the other view, as described in Sec. II B, an object was initially paired with all objects with centroids located within its defined annular region in the other view. The geometric constraints reduced the number of object pairs that needed to be classified as true or false correspondences in the subsequent steps. A true pair (TP–TP) was defined as the correspondence between the same true masses on the two mammographic views, and a false pair is defined as any other object pairing (TP–FP, FP–TP, and FP–FP). For each object pair, the set of 15 texture and 31 morphological features described above were used to form similarity measures. In this preliminary study, two simple measures, the absolute difference and the mean, were used. A total of 30 texture measures and 62 morphological measures were thus obtained for each object pair. The absolute difference between the nipple-to-object distances in the CC and MLO views was also included in both the texture and morphological feature sets as a feature for differentiating true from false object pairs. Two separate LDA classifiers with stepwise feature selection were trained to classify the true and false pairs using the similarity features in the morphological and texture feature spaces, respectively.

For training the classifiers, the data set was randomly divided into a training set and a test set again using a 3:1 training/test ratio. Fifty random 3:1 partitions of the cases were used to reduce bias. Individual morphological and texture scores were obtained for each object pair by averaging the test scores of each object pair obtained from the different partitionings. The two classification scores were then averaged to obtain one''correspondence'' score for each object pair. This approach of merging scores from different classifiers trained in parallel for the same task into a single score for further discrimination is similar to our previous work that used neural networks in morphological and texture feature spaces.[31] The correspondence score along with the one-view prescreening score were used in the following fusion step.

E. Fusion analysis

The fusion of the one-view prescreening scores with the two-view correspondence scores was the final step in our two-view detection scheme. In this study, we designed a fusion scheme that combines ranking and averaging of the prescreening and correspondence scores. We first ranked all prescreening object scores within a given film from the largest to the smallest. The correspondence scores were ranked in a similar way. These two new rank scores were then merged into a single score for each object in each view. Since an object could have more than one correspondence score, its two-view correspondence score was taken to be the maximum correspondence score among all object pairs in which this object was a member. There can be many variations for the fusion step.[32,33] In this preliminary study, the final discriminant score for an object was obtained by averaging its two-view correspondence score rank with its one-view prescreening score rank.

The FROC performance curve for the two-view analysis was generated by varying the decision threshold on the final discriminant score for each object and determining the sensitivity and FP per image at each threshold. We compared the FROC performance curves obtained by the two-view scheme when starting with 5, 10, and 15 prescreening objects per image and that obtained with the one-view detection scheme.

III. RESULTS

A. Geometrical modeling

In the geometrical analysis experiments, we first estimated a prediction model of the radial distance of an object in a second view from its radial distance in the first view using the training set. The model was then used to predict object location from one view to the other for the independent test cases. Since the model did not provide an exact solution, a search region, $R \pm \Delta R$, where R was the predicted radial distance and ΔR the half width of an annular region, was defined. The percentage of the true object centroids enclosed within the search region was measured as a function of the size of $2\,\Delta R$. Figure 8 shows the prediction accuracy as a function of $2\,\Delta R$ for estimating the object radial distance in the MLO view from that in the CC view. Figure 9 shows the corresponding results for predicting the object radial distance in the CC view from that in the MLO view. The training and test curves almost overlap in each case. The

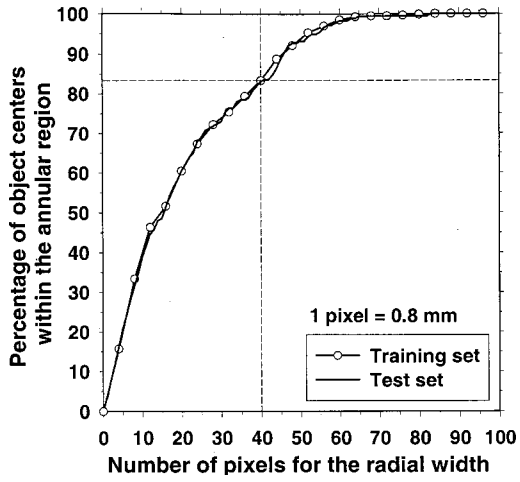

FIG. 8. The prediction of the center of an object in the MLO view from its location in the CC view. Training and test performances are given as a function of the radial width of the annular search region.

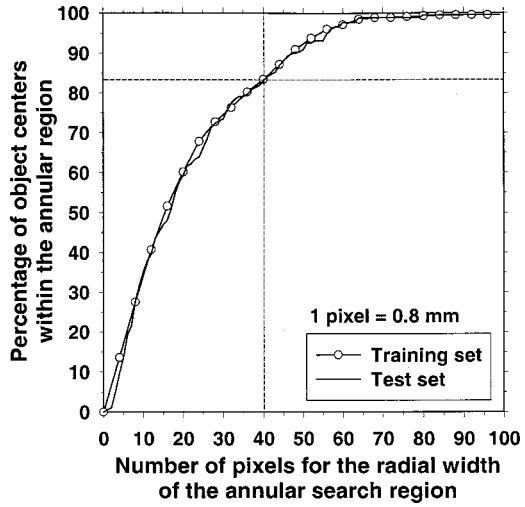

FIG. 9. The prediction of the center of an object in the CC view from its location in the MLO view. Training and test performances are given as a function of the radial width of the annular search region.

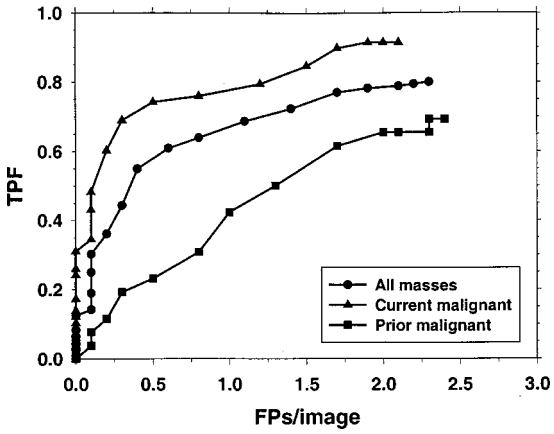

FIG. 11. Case-based performances of the current one-view mass detection algorithm applied to the data set of 169 pairs of mammograms. The FROC curves are plotted for the detection of all malignant and benign masses, and of the malignant masses on the current and the prior mammograms. Higher sensitivity was obtained for the detection of malignant masses on current mammograms.

difference in the accuracy between searching the object centers in the CC or MLO views is small. About 83% of the object centers are within the search region when the radial width of the search region is about 40 pixels (32 mm) for either the CC view or the MLO view. These results indicate that the search region, although large, is much smaller than the entire area of the breast. The limited search region size reduces the number of object pairs to be analyzed in the two-view detection scheme. To avoid missing any pairs of true masses in the two-view scheme, we chose to set the radial width of the annular search region to about 80 pixels. This led to a larger number of false pairs, but it was substantially less than that if the entire breast area was considered.

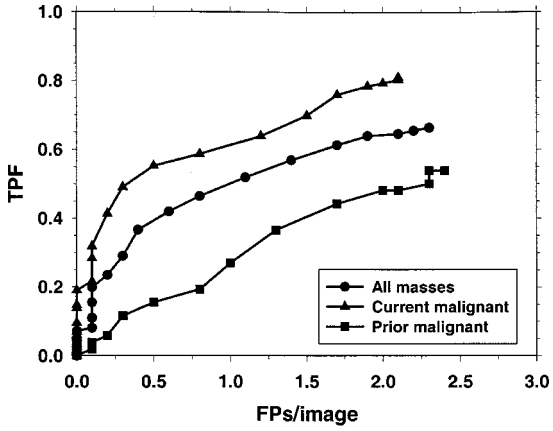

FIG. 10. Film-based performances of the current one-view mass detection algorithm applied to the data set of 338 one-view (169 pairs) mammograms. The FROC curves are plotted for the detection of all malignant and benign masses, and of the malignant masses on the current and the prior mammograms. Higher sensitivity was obtained for the detection of malignant masses on current mammograms.

B. One-view analysis

The FROC curve obtained from our current one-view mass detection algorithm[25] applied to the data set of 338 images is shown in Fig. 10. The FROC curves for the detection of the malignant masses on the current and prior mammograms are also plotted for comparison.

In clinical application, if the mass is detected on one-view by the computer and the radiologist is alerted to the mass, the radiologist will likely find the mass on the other view, if it is visible, even if the CAD algorithm misses it on the other view. Some researchers therefore consider a true-positive as the detection of the mass on one or two views of the breast.[10] We refer to this as case-based analysis. In this situation, the total number of masses or cases in this study was 169. For comparison purposes, we plot the case-based FROC curves for all masses, malignant masses on current mammograms, and malignant masses on prior mammograms in Fig. 11.

C. Fusion analysis

Three different decision thresholds that retained a maximum of 5, 10, and 15 objects per image after the one-view prescreening stage were used to select mass candidates as inputs to the two-view detection scheme. Table I summarizes the characteristics of these three object sets. The average number of prescreening objects per image was smaller than

TABLE I. Characteristics of the 3 sets of objects to be input to the two-view scheme. The objects were obtained by applying a detection threshold at the prescreening stage to extract a maximum of 5, 10, and 15 objects per image.

Prescreening threshold objs/image	Avg. objs/image	Sensitivity film-based (%)	Sensitivity case-based (%)	No. of pairs/case
5	4.9	72.7	85.2	14.2
10	9.4	79.8	89.3	49.4
15	12.6	83.4	92.3	85.9

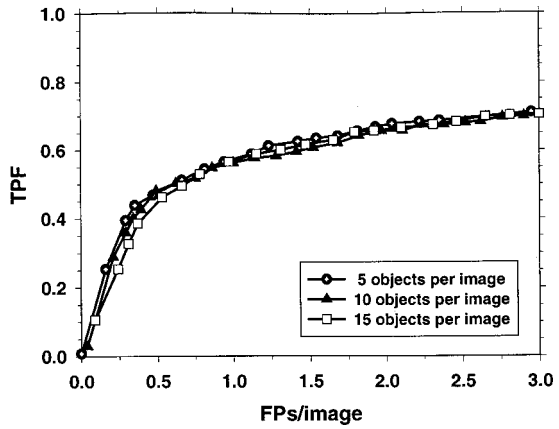

FIG. 12. Film-based performances of the proposed two-view detection scheme for all masses. Three initial conditions depending on the maximum number of retained objects per image (5, 10, and 15 objects per image) at the prescreening stage were evaluated.

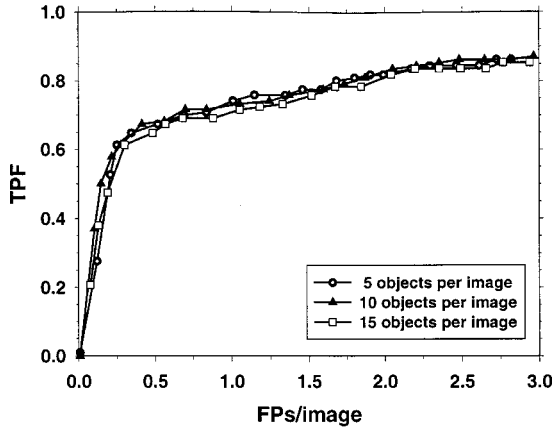

FIG. 13. Film-based performances of the proposed two-view detection scheme applied to the current malignant masses. Three initial conditions depending on the maximum number of retained objects per image (5, 10, and 15 objects per image) at the prescreening stage were evaluated.

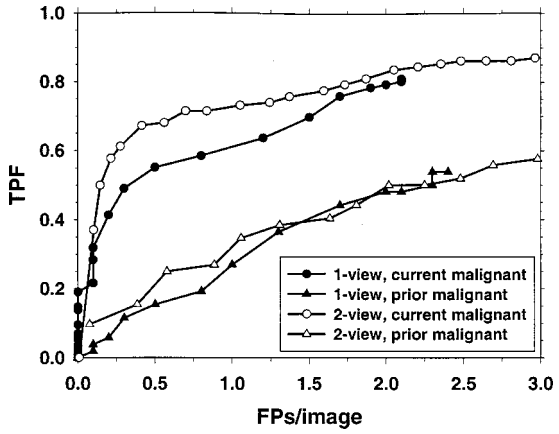

FIG. 14. A comparison of the film-based performance of the one-view and two-view detection methods for the detection of malignant masses on current mammograms and prior mammograms.

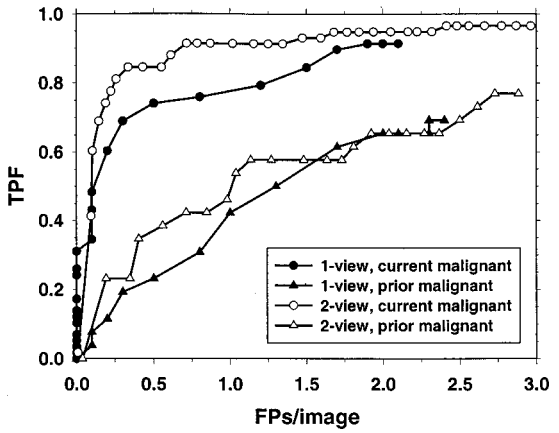

FIG. 15. A comparison of the case-based performance of the one-view and two-view detection methods for the detection of malignant masses on current mammograms and prior mammograms.

the maximum number allowed per image because the total number of objects in some images was smaller than the maximum number.

The FROC curves for the detection of malignant and benign masses on each image, using our two-view fusion technique, are shown in Fig. 12. The curves are similar for the three thresholds of 5, 10, 15 prescreening objects per image. This similarity also holds for the FROC curves for the detection of malignant masses, as illustrated in Fig. 13. The improvement in detection by our current two-view fusion method therefore seems to be independent of the operating threshold when the maximum number of objects retained per image in the prescreening stage is between 5 and 15.

Figure 14 compares the film-based FROC curves for the detection of malignant masses by the one-view and two-view fusion methods obtained from the condition of 10 prescreening objects per image. Figure 15 compares the corresponding case-based FROC curves. A comparison of the detection sensitivity at 1 FP image between the one-view and two-view fusion methods is given in Table II for both film-based and case-based detection.

IV. DISCUSSION

In this work, we propose a new technique based on the fusion of one-view and two-view information to improve the performance of mammographic mass detection. The results of our preliminary study show that including correspondence

TABLE II. A comparison of detection sensitivities obtained by the one-view and the two-view fusion schemes for film-based and case-based detection.

	Sensitivity-film-based (1 FPs/image)		Sensitivity-case-based (1 FPs/image)	
Mass type	1-view	2-view	1-view	2-view
All	50%	56%	67%	73%
Current malignant	62%	73%	77%	91%
Prior malignant	27%	33%	42%	52%

information from two mammographic views is an effective technique for reducing FPs. At a case-based detection sensitivity of 75% for all masses, the number of FPs per image was reduced from 1.5 FPs/image using the one-view detection technique to 1.13 FPs/image using the two-view fusion technique. The results also indicate that our proposed method is more effective in reducing FPs in the subset of cases containing malignant masses on current mammograms. At a case-based sensitivity of 85% for malignant masses on current mammograms, the number of FPs per image was reduced from 1.5 FPs/image to 0.5 FPs/image using the two-view fusion technique (Fig. 15). Alternatively, at 1 FPs/image, the two-view algorithm achieved a case-based detection sensitivity of 91% whereas the current one-view scheme had a 77% sensitivity at the same number of FPs/image (Table II).

The two-view correspondence analysis is more useful for mammogram pairs for which the mass is detected on both views in the prescreening stage. The fusion process is designed to both increase the scores for the TPs and reduce the scores for FPs for such cases. For the data set of 169 pairs of mammograms under the condition of 10 prescreening objects per image, the mass was detected on both CC and MLO views in a subset of 120 cases and on only one view in another subset of 32 cases. If we analyzed the subset of cases in which the mass was detected in both views, at 1 FP/image, the case-based detection sensitivity increased from 82.5% for the current one-view algorithm to 93.3% using the two-view fusion technique. However, for the subset of cases in which the mass was detected on only one view at the prescreening stage, the fusion analysis reduced the scores for TPs. At 1 FP/image, the case-based detection sensitivity was reduced from 50% for the current one-view algorithm to 43.7% using the two-view fusion process. Similar trends for the detection results were observed when 5 and 15 objects per image were retained in the prescreening stage.

In this study, we chose the radial width of the annular search region to be 80 pixels for all mammograms. This radial width reduced the search region to only a fraction of the breast area for large breasts but it covered most of the breast area in smaller breasts. Therefore, the advantage of geometric correlation has not been fully utilized in small breasts. One approach to reducing the search region size for small breasts would be to choose the region size as a percentage of the breast area so that the actual width of the annular region will be different for each pair of mammograms. This will lead to a reduction in the number of false object pairs for small breasts. The second approach would be to use a third mammographic view when it is available. As we discussed previously,[19] using the three standard views (CC, MLO, and Lateral) of the breast allow more accurate localization of a lesion to within a small fan-shaped region. This approach would require further adaptation of our two-view scheme to a three-view fusion scheme. Although 3-view mammograms are not generally available for screening, it will be of interest to investigate how 3-view mammograms will improve the detection of malignancy in the breast by the computer.

In this study, we used radiologist-identified nipple locations for the geometric correlation process. In a fully automated mass detection program, this step will have to be automated. We are developing an automated nipple detection program. This detection program could identify the nipple within 1 cm of the true location in 88% of the 311 mammograms in a study set.[22] For the purpose of this study, we did not use automated nipple detection because it will complicate our analysis of the two-view fusion techniques if errors in nipple detection have to be taken into account. We therefore isolated their effects by using manually identified nipple locations. We will continue to improve the automated nipple detection algorithm and incorporate this step into the two-view mass detection scheme in the future.

In this preliminary study, we used two simple similarity measures for the classification of object correspondence. The fusion of the two-view and one-view scores for the individual objects was performed with a relatively simple ranking and averaging methods. These approaches already provided substantial improvement in the detection accuracy, indicating the promise of the two-view method for mass detection and FP reduction. Further studies are being conducted to optimize the various steps in the two-view classification and fusion schemes.

V. CONCLUSION

We are developing a two-view fusion technique to improve computerized mass detection on mammograms. Starting from objects detected in a prescreening stage, we defined all possible pairing based on geometry and then combined morphological and textural characteristics from these paired objects into a correspondence score for each object. A classifier was trained to differentiate the true mass pairs from the false pairs. A final fusion stage combined the two-view object pair information with the one-view object scores. Our preliminary results demonstrate that the proposed two-view scheme can reduce FPs in comparison with our current one-view method. The mass detection sensitivity is also improved by using information from the two-views. Further studies are underway to optimize the prescreening process, the design of the similarity measures, as well as the two-view fusion scheme. When fully developed and integrated into the CAD system, it is expected that our proposed two-view technique will improve upon the current one-view scheme and provide a useful second opinion to radiologists in the detection of breast cancer on mammograms.

ACKNOWLEDGMENTS

This work is supported by USPHS Grant No. CA 48129, USAMRMC Grant No. DAMD 17-96-1-6254, and a Career Development Award (B.S.) DAMD 17-96-1-6012 from the USAMRMC. The content of this publication does not necessarily reflect the position of the government and no official endorsement of any equipment and product of any companies mentioned in the publication should be inferred.

[a]Correspondence: Sophie Paquerault, Ph.D. c/o Heang-Ping Chan, Ph.D., Department of Radiology, University of Michigan, 1500 E. Medical Center Drive, B1F510B, Ann Arbor, Michigan 48109-0030. Telephone: (734) 936-4357; fax: (734) 615-5513; electronic mail: chanhp@umich.edu

[1] L. Tabar *et al.*, "Reduction in mortality from breast cancer after mass screening with mammography," Lancet **1**, 829–832 (1985).

[2] H. C. Zuckerman, "The role of mammography in the diagnosis of breast cancer." in: *Breast Cancer Diagnosis and Treatment*, edited by I. M. Ariel and J. B. Cleary (McGraw-Hill, New York, 1987), pp. 152–172.

[3] L. W. Bassett, D. H. Bunnell, R. Jahanshahi, R. H. Gold, R. D. Arndt, and J. Linsman, "Breast cancer detection: one versus two views," Radiology **165**, 95–97 (1987).

[4] E. Thurfjell, "Mammography screening: One versus two views and independent double reading," Acta Radiol. **35**, 345–350 (1994).

[5] R. G. Blanks, M. G. Wallis, and R. M. Given-Wilson, "Observer variability in cancer detection during routine repeat (incident) mammographic screening in a study of two versus one view mammography," J. Med. Screen **6**, 152–158 (1999).

[6] E. D. C. Anderson, B. B. Muir, J. S. Walsh, and A. E. Kirkpatrick, "The efficacy of double reading mammograms in breast screening," Clin. Radiol. **49**, 248–251 (1994).

[7] E. L. Thurfjell, K. A. Lernevall, and A. A. S. Taube, "Benefit of independent double reading in a population-based mammography screening program," Radiology **191**, 241–244 (1994).

[8] H. P. Chan, K. Doi, C. J. Vyborny, R. A. Schmidt, C. E. Metz, K. L. Lam, T. Ogura, Y. Wu, and H. MacMahon, "Improvement in radiologists' detection of clustered microcalcifications on mammograms. The potential of computer-aided diagnosis," Invest. Radiol. **25**, 1102–1110 (1990).

[9] W. P. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions," Radiology **191**, 331–337 (1994).

[10] L. J. Warren Burhenne, S. A. Wood, C. J. D'Orsi, S. A. Feig, D. B. Kopans, K. F. O'Shaughnessy, E. A. Sickles, L. Tabar, C. J. Vyborny, and R. A. Castellino, "Potential contribution of computer-aided detection to the sensitivity of screening mammography," Radiology **215**, 554–562 (2000).

[11] F. F. Yin, M. L. Giger, K. Doi, C. E. Metz, C. J. Vyborny, and R. A. Schmidt, "Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images," Med. Phys. **18**, 955–963 (1991).

[12] L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. Gurcan, "Analysis of temporal change of mammographic features for computer-aided characterization of malignant and benign masses," Proc. SPIE **4322**, 661–666 (2001).

[13] L. M. Hadjiiski, H. P. Chan, B. Sahiner, N. Petrick, and M. A. Helvie, "Automated registration of breast lesions in temporal pairs of mammograms for interval change analysis—local affine transformation for improved localization," Med. Phys. **28**, 1070–1079 (2001).

[14] L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. N. Gurcan, "Analysis of temporal change of mammographic features: Computer-aided classification of malignant and benign breast masses," Med. Phys. **28**, 2309–2317 (2001).

[15] Y. Kita, R. P. Highnam, and J. M. Brady, "Correspondence between different view breast x rays using curved epipolar lines," Comput. Vis. Image Underst. **83**, 38–56 (2001).

[16] W. F. Good, B. Zheng, Y. H. Chang, Z. H. Wang, G. S. Maitz, and D. Gur, "Multi-image CAD employing features derived from ipsilateral mammographic views," Proc. SPIE **3661**, 474–485 (1999).

[17] Y. H. Chang, W. F. Good, J. H. Sumkin, B. Zheng, and D. Gur, "Computerized localization of breast lesions from two views—An experimental comparison of two methods," Invest. Radiol. **34**, 585–588 (1999).

[18] S. S. Gopal, H. P. Chan, T. E. Wilson, M. A. Helvie, N. Petrick, and B. Sahiner, "A regional registration technique for automated interval change analysis of breast lesions on mammograms," Med. Phys. **26**, 2669–2679 (1999).

[19] S. Paquerault, B. Sahiner, N. Petrick, L. M. Hadjiiski, M. N. Gurcan, C. Zhou, and M. A. Helvie, "Prediction of object location in different views using geometrical models," *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, Canada, 11–14 June 2000, Digital Mammography, Madison, WI (Medical Physics Publishing), pp. 748–755.

[20] S. Paquerault, N. Petrick, H. P. Chan, B. Sahiner, and A. Y. Dolney, "Improvement of mammographic lesion detection by fusion of information from different views," Proc. SPIE **4322**, 1883–1889 (2001).

[21] A. R. Morton, H. P. Chan, and M. M. Goodsitt, "Automated model-guided breast segmentation algorithm," Med. Phys. **23**, 1107–1108 (1996).

[22] C. Zhou, H. P. Chan, N. Petrick, M. M. Goodsitt, C. Paramagul, and L. M. Hadjiiski, "Computerized image analysis: breast segmentation and nipple identification on mammograms," *Proceedings of the Chicago 2000-World Congress on Medical Physics and Biomedical Engineering*, Chicago, Illinois, 23–28 July, 2000, Paper No. TH-Aa201-04.

[23] N. Petrick, H. P. Chan, D. Wei, B. Sahiner, M. A. Helvie, and D. D. Adler, "Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification," Med. Phys. **23**, 1685–1696 (1996).

[24] N. Petrick, H. P. Chan, B. Sahiner, and D. Wei, "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection," IEEE Trans. Med. Imaging **15**, 59–67 (1996).

[25] N. Petrick, H. P. Chan, B. Sahiner, and M. A. Helvie, "Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms," Med. Phys. **26**, 1642–1654 (1999).

[26] N. Petrick, B. Sahiner, H. P. Chan, M. A. Helvie, and S. Paquerault, "Preclinical evaluation of a CAD algorithm for early detection of breast cancer," *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, Canada, 11–14 June 2000, Digital Mammography, Madison, WI (Medical Physics Publishing), pp. 328–333.

[27] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," IEEE Trans. Syst. Man Cybern. **SMC-3**, 610–621 (1973).

[28] R. M. Haralick, "Statistical image texture analysis," in *Handbook of Pattern Recognition and Image Processing* (Academic, New York, 1986).

[29] P. C. Bunch, J. F. Hamilton, G. K. Sanderson, and A. H. Simmons, "A free response approach to the measurement and characterization of radiographic observer performance," J. Appl. Photogr. Eng. **4**, 166–171 (1978).

[30] G. M. te Brake, N. Karssemeijer, and J. H. C. L. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," Phys. Med. Biol. **45**, 2843–2857 (2000).

[31] B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: An artificial neural network with morphological features," *Proc. World Cong. Neural Net. II*, 1995, pp. 876–879.

[32] A. Kandel, *Fuzzy Techniques in Pattern Recognition* (Wiley, New York, 1982).

[33] B. Sahiner, N. Petrick, H. P. Chan, S. Paquerault, M. A. Helvie, and L. M. Hadjiiski, "Recognition of lesion correspondence on two mammographic views—A new method of false-positive reduction for computerized mass detection," Proc. SPIE **4322**, 649–655 (2001).