Computerized image analysis: Texture-field orientation method for pectoral muscle identification on MLO-view mammograms

Chuan Zhou,[a] Jun Wei, Heang-Ping Chan, Chintana Paramagul, Lubomir M. Hadjiiski, and Berkman Sahiner
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-5842

Julie A. Douglas
Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109-5618

(Received 10 December 2009; revised 11 March 2010; accepted for publication 22 March 2010; published 28 April 2010)

Purpose: To develop a new texture-field orientation (TFO) method that combines *a priori* knowledge, local and global information for the automated identification of pectoral muscle on mammograms.
Methods: The authors designed a gradient-based directional kernel (GDK) filter to enhance the linear texture structures, and a gradient-based texture analysis to extract a texture orientation image that represented the dominant texture orientation at each pixel. The texture orientation image was enhanced by a second GDK filter for ridge point extraction. The extracted ridge points were validated and the ridges that were less likely to lie on the pectoral boundary were removed automatically. A shortest-path finding method was used to generate a probability image that represented the likelihood that each remaining ridge point lay on the true pectoral boundary. Finally, the pectoral boundary was tracked by searching for the ridge points with the highest probability lying on the pectoral boundary. A data set of 130 MLO-view digitized film mammograms (DFMs) from 65 patients was used to train the TFO algorithm. An independent data set of 637 MLO-view DFMs from 562 patients was used to evaluate its performance. Another independent data set of 92 MLO-view full field digital mammograms (FFDMs) from 92 patients was used to assess the adaptability of the TFO algorithm to FFDMs. The pectoral boundary detection accuracy of the TFO method was quantified by comparison with an experienced radiologist's manually drawn pectoral boundary using three performance metrics: The percent overlap area (POA), the Hausdorff distance (Hdist), and the average distance (AvgDist).
Results: The mean and standard deviation of POA, Hdist, and AvgDist were $95.0 \pm 3.6\%$, 3.45 ± 2.16 mm, and 1.12 ± 0.82 mm, respectively. For the POA measure, 91.5%, 97.3%, and 98.9% of the computer detected pectoral muscles had POA larger than 90%, 85%, and 80%, respectively. For the distance measures, 85.4% and 98.0% of the computer detected pectoral boundaries had Hdist within 5 and 10 mm, respectively, and 99.4% of computer detected pectoral muscle boundaries had AvgDist within 5 mm from the radiologist's manually drawn boundaries.
Conclusions: The pectoral muscle on DFMs can be detected accurately by the automated TFO method. The preliminary study of applying the same pectoral muscle identification algorithm to FFDMs without retraining demonstrates that the TFO method is reasonably robust against the differences in the image properties between the digitized and digital mammograms. [DOI: 10.1118/1.3395576]

Key words: computer-aided detection, mammography, pectoral muscle trimming, texture orientation field analysis

I. INTRODUCTION

Computerized analysis of mammograms can assist radiologists in detection and characterization of breast lesions, and estimation of breast density that may be used for prediction of breast cancer risk. We have previously developed a computerized system, mammographic density estimator (MDEST), to estimate breast density automatically on digitized film mammograms (DFMs).[1] For each mammogram, the breast region was first segmented by breast boundary detection and, for the mediolateral oblique (MLO) view, with additional pectoral muscle trimming. A gray level threshold was then automatically determined to segment the dense tissue from the breast region. The breast density was estimated as the percentage of the segmented dense area relative to the breast area. Our previous study[1] indicated that the computer-estimated mammographic breast density correlated closely with the "reference standard" obtained by averaging five experienced radiologists' manual segmentations and the average bias was much less than that of the radiologists' visual estimation. However, density estimation on MLO-view mammograms was not as accurate as that on CC-view mammograms. For example, because the pectoral region has higher brightness than the fatty regions in the breast, if the

pectoral muscle is not completely excluded, the residual pectoral region will be counted as dense area and will cause an overestimation of the breast density. Conversely, if the trimmed pectoral region inadvertently includes some of the dense area adjacent to the pectoral region, the breast density may be underestimated. Accurate segmentation of the pectoral muscle on MLO-view mammograms can reduce the bias in mammographic density estimation and improve the performance of our MDEST method.

Automatic identification of the pectoral muscle is also an essential step in other mammographic image analysis applications. For example, it can enable region-specific processing in lesion detection programs to reduce false negatives. False positives can be reduced if the detected objects in the pectoral muscle area can be selectively suppressed. The pectoral muscle may also be used as a reference in image registration algorithms for multiple-view analysis of mammograms.[2–4]

A few studies have been presented to identify the pectoral muscle to date. In the study of the classification of parenchymal patterns, Karssemeijer *et al.*[5] applied Hough transform and thresholding method to detect the pectoral muscle, based on the assumption that the boundary of the pectoral muscle is approximately a straight line at an angle between 45° and 90° in the oblique view. Although 615 oblique mammograms were used as a test set for the evaluation of parenchymal pattern classification, the performance of the pectoral muscle detection was not reported in the study. Ferrari *et al.*[6] implemented and modified the Hough transform method by Karssemeijer *et al.*[5] To overcome the limitation of the straight line representation of the pectoral muscle, a Gabor filter-based edge enhancement method was used to improve their initial pectoral muscle detection technique. The pectoral muscle edges were manually drawn by the author, checked by a radiologist, and used as reference regions for performance evaluation. The pixels identified by computer but outside the reference region were defined as false positive (FP) pixels, and the pixels within the reference region but missed by computer were defined as false negative (FN) pixels. FP and FN were normalized with the reference region size. The Gabor filter-based method resulted in average FP and FN rates of 0.58% and 5.77%, respectively, for 84 MLO testing mammograms. Ma *et al.*[7] used two image segmentation methods based on graph theory in conjunction with active contour to segment pectoral muscle in screening mammogram. One method was based on adaptive pyramids (APs) and the other was based on minimum spanning trees (MSTs). Using the same data set of 84 MLO test mammograms and the reference standard as those in Ferrari's study,[6] the average FP and FN rates of 3.71% and 5.95%, respectively, were obtained by the AP method, and 2.55% and 11.68%, respectively, by the MST method (MST method did not detect the pectoral muscle in two mammograms; the average FP and FN were based on 82 mammograms accordingly). Several other studies first approximated the pectoral muscle edge by a straight line, which was then refined using techniques such as neural network,[8] Radon transform,[9] iterative cliff detection,[10] voting scheme,[11] and region growing.[12] The detection performance in all these studies were evaluated with small data sets (less than 100 mammograms), except for the studies by Kwok *et al.*[10] (322 mammograms), Raba *et al.*[12] (300 mammograms), and Kinoshita *et al.*[9] (540 mammograms). However, the performance of the iterative cliff detection method by Kwok *et al.*[10] was only visually assessed by radiologists using a five-point assessment scale, who rated 83.9% of the segmentation as adequate or better, and no pixelwise quantitative evaluation was performed. The performance of the region growing method by Raba *et al.*[12] was also visually inspected by a radiologist who classified 98% and 86% of the 320 images as "near accurate" and "good," respectively. The Radon transform based method by Kinoshita *et al.*[9] achieved accurate pectoral muscle segmentation in 28.9% (156/540) and acceptable in 40.7% (220/540), but unacceptable in 30.4% (164/540) of the mammograms compared to a radiologist's manual segmentation.

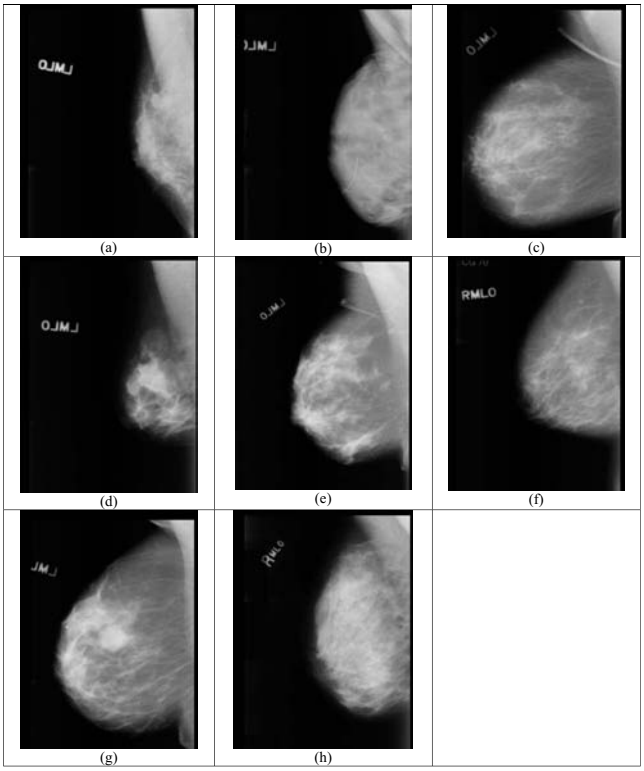


FIG. 1. Examples of different patterns of pectoral muscle on the mammograms that are challenging for automated pectoral muscle detection. [(a) and (b) Dense breast containing large area of fibroglandular tissue overlapping the pectoral muscle; [(b), (c), (e), and (g)] skin fold located in the upper region of the breast mimicking pectoral muscle boundary; [(b), (d), (e), and (h)] fuzzy pectoral edges in the lower region of the pectoral muscle; and [(f) and (g)] small pectoral muscle.

The pectoral muscle boundary on a mammogram can be complicated, especially for the improperly positioned views and when dense glandular tissue overlaps with the pectoral muscle region. Figure 1 shows examples of four typical pectoral muscle patterns observed in our studies that are challenging for automated detection: (1) Dense breast containing large area of dense fibroglandular tissue overlapping the pectoral muscle [Figs. 1(a) and 1(b)]; (2) skin fold located in the

upper region of the breast mimicking pectoral muscle boundary [Figs. 1(b), 1(c), 1(e), and 1(g)]; (3) fuzzy pectoral edges in the lower region of the pectoral muscle [Figs. 1(b), 1(d), 1(e), and 1(h)]; and (4) small pectoral muscle region [Figs. 1(f) and 1(g)]. Note that the small bright triangle region in the upper-right corner of Fig. 1(h) is not the pectoral muscle, but may be detected mistakenly by the computer because of its strong and straight edge and location. The complicated and varied patterns seen in these examples illustrate the difficulties for computerized pectoral muscle detection. A large number of studies on CAD methods for mammography have been reported in the literature; however, few studies evaluated the performance of a specific automated pectoral muscle identification method using a large data set in comparison to radiologist's manual segmentation as a reference standard. It is not clear whether the published methods are robust against the difficult situations demonstrated in Fig. 1. Although the difficult cases may only constitute a small fraction of all cases, they are major hurdles to the full automation of advanced CAD techniques.

From our experiences in analysis of mammograms and a previous study on pectoral muscle trimming,[1–4] we have observed that the texture orientation and the local and global features in the pectoral region may be useful for detection of pectoral muscle boundary. In this study, we developed a new texture-field orientation (TFO) method that utilized two gradient-based directional kernel (GDK) filters to enhance pectoral edges. The first GDK filter was used to enhance the linear texture structures, and then a gradient-based texture analysis was designed to extract a texture orientation image that represented the dominant texture orientation at each pixel. The texture orientation image was enhanced by a second GDK filter for ridge point extraction. After validation of the extracted ridge points, a shortest-path finding method was used to estimate the probability of each ridge point lying on the true pectoral boundary. Finally, the pectoral boundary was tracked by searching for and connecting the ridge points with higher probability. The accuracy of the pectoral boundary detection was evaluated on a large independent data set by comparison with an experienced breast radiologist's manually drawn pectoral boundary. The robustness of the algorithm was further evaluated on an independent data set of full field digital mammograms (FFDMs).

II. MATERIALS AND METHODS

II.A. Data sets

A data set of 260 four-view screen-film mammograms from 65 patients was used to evaluate the performance of an automated breast density segmentation method in our previous study.[1] In this study, 130 MLO-view mammograms from the above data set were used as a training set for the development of pectoral muscle identification method. An independent data set of 637 MLO-view mammograms from 562 patients was used as a test set. Among the 637 mammograms, 531 mammograms from 463 patients were randomly selected from the patient files of an ongoing NIH supported and Institutional Review Board (IRB) approved genetic study of breast density in women from the Old Order Amish population of Lancaster County, Pennsylvania.[13] Another 106 mammograms from 99 patients were collected from the patient files in the Radiology Department at the University of Michigan (UM) hospital with IRB approval. The mammograms were digitized with a LUMISYS 85 laser film scanner with a pixel size of 50×50 μm^2 and 4096 gray levels. The gray levels are linearly proportional to optical densities (O.D.s) from 0.1 to greater than 3 O.D. units. To reduce processing time and noise, the full resolution digitized mammograms were first smoothed with a 16×16 box filter and subsampled by a factor of 16, resulting in a pixel size of 800×800 μm^2 with approximately 225×300 pixels in an image.

A data set consisting of FFDMs of 92 patients with 92 MLO-view images collected from the patient files in the Radiology Department at the UM hospital was used as an independent test set to evaluate the feasibility of applying our pectoral muscle identification method to FFDMs without retraining. The FFDM was acquired with a GE Senographe 2000D system. To avoid the changes in image quality in the "For Presentation" images processed by the manufacturer's proprietary image enhancement methods that may change over time, we generally use the raw images as input to our computerized image analysis algorithms. A logarithmic transform was applied to the raw FFDM to convert the linear relationship of x-ray intensity vs pixel value to a logarithmic relationship and 12-bit gray levels.[14] The image after logarithmic transform was smoothed with an 8×8 box filter and subsampled by a factor of 8 to a pixel size of 800×800 μm^2 to match the same input resolution as that of the DFMs.

For each MLO-view DFM and FFDM, an experienced Mammography Quality Standards Act (MQSA) radiologist visually inspected the image displayed on a monitor with a graphical user interface and used the windowing function to enhance the visibility of the pectoral muscle boundary. The pectoral muscle boundary was manually drawn by the radiologist using the cursor, which was then used as the reference standard for the evaluation of the computer performance. The radiologist also rated the visibility of the pectoral muscle on a four-point scale, and the breast density in terms of breast imaging reporting and data system (BI-RADS) category[15] for each MLO-view mammogram. Among the 637 MLO-view DFMs that were used as an independent test set, 60.3%, 23.1%, 10.5%, and 6.1% of these 637 mammograms were rated as one of four categories from high to low visibility of pectoral boundaries, respectively. For the density classification, 21.0%, 52.0%, 22.0%, and 5.0% of the 637 MLO-view DFMs were rated as BI-RADS categories 1 to 4, respectively.

II.B. Methods

II.B.1. Enhancement in pectoral muscle edges

With the assumption that there exists a dominant orientation at each pixel within a texture pattern, an "orientation image" can be computed from the gray level mammogram

using least mean squares estimation based on the optimal solution of Rao.[16] Let $g_x(u,v)$ and $g_y(u,v)$ represent the horizontal and vertical gradients at pixel (u,v) in the image. The gradient magnitude is computed as $G_{u,v}=\sqrt{g_x^2(u,v)+g_y^2(u,v)}$, and the gradient orientation is computed as $\theta_{u,v}=\arctan(g_y(u,v)/g_x(u,v))$. The sum-of-squares S in an $M\times N$ local neighborhood centered at pixel (i,j) can be computed as

$$S=\sum_{u=1}^{M}\sum_{v=1}^{N} G_{u,v}^2\cos^2(\theta_{u,v}-\phi(i,j)), \tag{1}$$

where S is the sum of the squared gradient magnitudes projected along a direction $\phi(i,j)$ in this neighborhood. $\phi(i,j)$ is referred to as the dominant orientation when S is the maximum. The maximum of S with respect to $\phi(i,j)$ can be found by solving the equation $(dS/d\phi(i,j))=0$, the left-hand side of which is given by

$$\frac{dS}{d\phi(i,j)}=2\sum_{u=1}^{M}\sum_{v=1}^{N} G_{u,v}^2\cos(\theta_{u,v}-\phi(i,j))\sin(\theta_{u,v}-\phi(i,j)). \tag{2}$$

Thus, the dominant orientation $\phi(i,j)$ can be estimated as

$$\phi(i,j)=\frac{1}{2}\tan^{-1}\left(\frac{\Sigma_{u=1}^{M}\Sigma_{v=1}^{N}G_{u,v}^2\sin 2\theta_{u,v}}{\Sigma_{u=1}^{M}\Sigma_{v=1}^{N}G_{u,v}^2\cos 2\theta_{u,v}}\right)=\frac{1}{2}\tan^{-1}\left(\frac{\Sigma_{u=1}^{M}\Sigma_{v=1}^{N}2g_x(u,v)g_y(u,v)}{\Sigma_{u=1}^{M}\Sigma_{v=1}^{N}(g_x^2(u,v)-g_y^2(u,v))}\right). \tag{3}$$

In our study, the dominant orientation $\phi(i,j)$ was estimated in a $5\times 5(M=N=5)$ local region. The size of the local region was chosen by experimenting with the training set. If the size is too small, the orientation of the texture will not converge and the estimated dominant orientation is very noisy. If the size is too large, the local region may contain more than one major linear texture structures with different orientations, which can diffuse the dominant orientation of the textures depicting the pectoral muscle boundary. An appropriate size will be more effective in enhancing the linear texture structures. However, given the varied textures on individual mammograms, the choice of the 5×5 kernel size is only a compromise among the population as estimated from the training set, rather than the optimal for each type of pectoral muscle textures.

The MLO view is preferred over a lateral 90° projection in screening mammography because more of the breast tissue in the upper outer quadrant of the breast and the axilla can be imaged. On the MLO view, the pectoral muscle is depicted obliquely from above and down to the level of the nipple or further down. The medial (middle) portion of the breast should be prominent in the MLO view. The orientation of the digitized image can be automatically determined by the curvature of the breast boundary. To facilitate image processing, all mammograms (regardless of laterality) were first oriented such that the chest wall was on the right side. We could then assume that the direction of the pectoral muscle boundary ran approximately from top left to bottom right with greater than 45° angulation. We developed a GDK filter to enhance the linear texture structures on the mammogram at approximate 45° from top left to bottom right. The filter kernel is defined as:

$$h_1=\begin{bmatrix} a & a & a & \dots & a\\ b & a & a & \dots & a\\ b & b & a & \dots & a\\ . & & & & .\\ . & & & & .\\ b & b & \dots & b & a\end{bmatrix}\quad h_2=\begin{bmatrix} a & \dots & a & a & a\\ a & \dots & a & a & b\\ a & \dots & a & b & b\\ . & & & & .\\ . & & & & .\\ a & b & \dots & b & b\end{bmatrix}, \tag{4}$$

where, $a=1$ and $b=-1$. The size of the kernel h_1 and h_2 was chosen to be 11×11 pixels (8.8×8.8 mm^2), which is about two times of the local region for estimation of the dominant texture orientation, described above.

The gradient at pixel (u,v) is estimated by convolving the image f with the kernels defined in Eq. (4) as follows:

$$g_x(u,v)=f(u,v)*h_1$$

$$g_y(u,v)=f(u,v)*h_2. \tag{5}$$

Combining Eqs. (3) and (5), the dominant orientation $\phi(i,j)$ can be used to enhance the texture pattern around the pectoral muscle, where strong gray level changes occur along the direction normal to the pectoral muscle edge.

Due to the presence of noise, the estimated local texture orientation may not always be correct. The orientation image is then smoothed using an edge preserving mean shift algorithm[17] that iteratively shifts each pixel to the average of the pixels in its neighborhood.

The texture patterns with dominant texture orientations directing from top left to bottom right, which are more likely to be the pectoral edges, are enhanced by applying a second GDK filter to the smoothed orientation image. The kernel of the second GDK is defined as

$$h=\begin{bmatrix} b_1 & \dots & b_4 & a_1 & \dots & a_5\\ . & & . & . & & .\\ . & & . & . & & .\\ b_1 & \dots & b_4 & a_1 & \dots & a_5\end{bmatrix}, \tag{6}$$

where $a_m=1$ $(m=1,\dots,5)$, $b_n=-1$ $(n=1,\dots,4)$, and the size of the kernel h was chosen to be 9×9 $(=7.2\times 7.2$ mm$^2)$. After the first GDK filtering followed by mean shift smoothing, the pixels located at the pectoral boundary with sharp edges form band structures along the pectoral boundary. By experimenting with the kernel size of the second GDK filter on the training set, it was observed the band structures along the pectoral boundary could be enhanced to be ridges with kernel sizes in the range of eight to ten pixels, while most other structures, including the band structures not matched with the filter kernel, would be suppressed. The kernel size was therefore chosen to be 9×9 pixels. In this step, a GDK filter that enhanced the horizontal gradient was found to be sufficient because most of the pectoral boundary had an orientation greater than 45°.

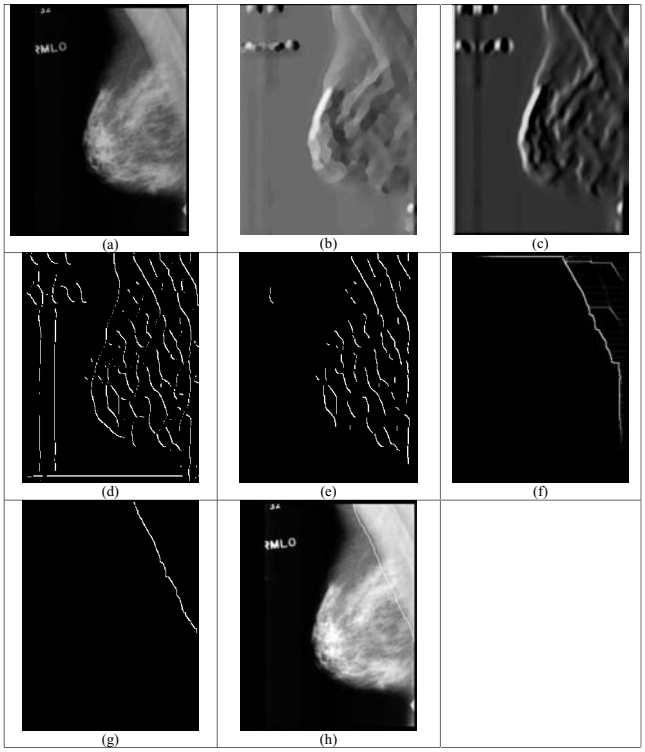


FIG. 2. Example of images of pectoral boundary enhancement and edge tracking in the intermediate steps of the new pectoral muscle detection method based on texture-field orientation. (a) Original image; (b) texture orientation image after first GDK filter and texture-flow analysis; (c) ridge image enhanced by the second GDK filter; (d) detected ridges; (e) validated ridges; (f) probability p_F, the pixel intensity represented the probability of having the shortest paths passing through the pixel; (g) tracked pectoral edges on image (f); and (h) the final identified pectoral boundary after running box smoothing along the edge.

Figure 2(b) shows an example of the smoothed orientation image. The texture along the direction of the pectoral muscle edge is enhanced to form band structures, while the glandular tissues overlapping with the pectoral muscle in the approximately normal direction are suppressed. Figure 2(c) shows the image enhanced by the second GDK filter in which the band structures that match the kernel are enhanced as ridges. Candidate edges of the pectoral muscle are detected on the enhanced orientation image using a ridge detection algorithm. The ridges are extracted by searching for the local maximum along the approximately 45° direction from top left to bottom right. Figure 2(d) shows the ridges extracted from the ridge image shown in Fig. 2(c).

II.B.2. Identification of pectoral muscle boundary

II.B.2.a. Validation of the ridge points. After the ridges are detected as candidate pectoral muscle edge points, a ridge validation process is used to remove the ridges that are less likely to lie along the pectoral boundary by the following knowledge-based decision criteria: (1) For a single segment connected by ridge points, a straight line is fitted to determine the line direction of this single segment. If the difference between the line direction and the presumed pectoral orientation is large than 45°, then the ridge segment will be eliminated. (2) A single segment containing a small number of points (less than ten) is removed without line fitting. (3) For each ridge segment validated by (1), two bands are formed parallel to the ridge segment on its left and right with the segment as the centerline. The average gray level within each band is calculated. Because the pectoral region is assumed to have higher brightness, if the average gray level in the left band is higher than that in the right band, then the current ridge segment will be removed as noise. Figure 2(e) shows a number of ridge points that were validated to be along the pectoral orientation.

II.B.2.b. Tracking of the pectoral muscle boundary. To track the true pectoral muscle boundary points, a seed point has to be determined. Because most of the pectoral muscle boundaries can be fitted to a straight line or a second-order curve, we can consider the pectoral boundary to be the shortest path traveling through validated ridge points from top left to bottom right.

The image containing ridge points can be considered a graph with ridge points as vertices. The algorithm Dijkstra[18] is a greedy algorithm that can solve the problem of finding the shortest path from a single vertex to all other vertices in a weighted graph. In graph theory, a graph is a set of objects called vertices connected by links called edges. Let V denote a set of all vertices in the weighted directed graph G. An ordered pair of vertices α and β is connected by an edge from α to β, and the weight $w(\alpha,\beta)$ of the edge connecting vertices α and β is the non-negative cost of moving from vertex α to β. Starting from vertex s and initializing all vertices in the graph to be not-yet-visited vertices, Dijkstra's algorithm repeatedly visits the closest not-yet-visited vertices and selects the vertex with the shortest-path cost to the starting vertex s, adding it to the set of vertices already visited. The weights of all vertices adjacent to the currently selected vertex are then updated, which is commonly referred to as "relaxation" of the edges between the vertices. The algorithm ultimately connects all vertices reachable from the starting point with the shortest path.

In our study, the image containing detected ridges constitutes a directed graph, where the ridge points are graph vertices and the oriented gray level gradients of ridge points represent edges of the graph. From a starting vertex s in the image, a map of shortest paths from all vertices to s is calculated using Dijkstra's algorithm. With the assumption that the pectoral edge intercepts the top and the right of the image, as shown in Fig. 3, the shortest-path searching method is applied to the ridge image at multiple starting points along the top margin of the image to extract the ridge paths ending at multiple points along the right margin of the image.

Let S be a set of starting vertices $S=\{s_1,\ldots,s_m\}$, and T a set of destination vertices $T=\{t_1,\ldots,t_n\}$. Then $K=m\times n$ shortest paths from S to T can be found by using Dijkstra's algorithm. Let $P_k(x,y|s_i,t_j)$ be a binary index such that $P_k(x,y|s_i,t_j)=1$ if point (x,y) is on the shortest path k from s_i to t_j, otherwise, $P_k(x,y|s_i,t_j)=0$. An image $F(x,y)$ can

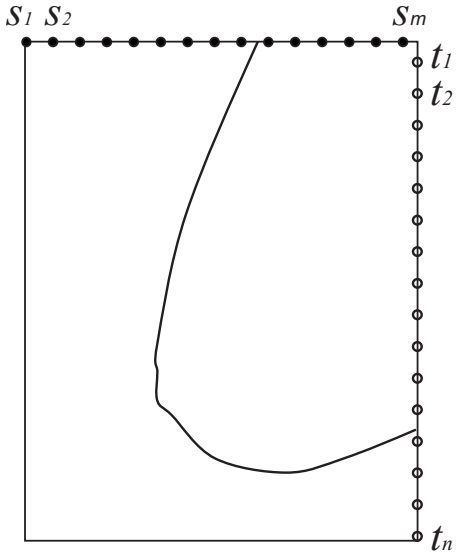


FIG. 3. Illustration of the shortest-path search with different starting (solid dots) and ending points (open dots).

then be generated to count the number of paths passing through a point (x,y) by summing over all of the paths as follows:

$$F(x,y)=\sum_{k=1}^{K} P_k(x,y|s_i,t_j), \quad i=1,\ldots m; \quad j=1,\ldots,n.$$

Therefore, the probability of a ridge point (x,y) having K shortest paths passing through it is given by

$$p_F(x,y)=F(x,y)/K.$$

Figure 2(f) shows an example of a probability image. In the probability image, the higher gray level of a point, the higher probability of this point being tracked as a true pectoral edge because more shortest paths pass through this pixel.

The seed point for tracking the pectoral boundary is then determined at the point having a maximum probability value of p_F. Note that the number of points having a maximum p_F can be greater than one because an equal number of paths can pass through different ridge points. If more than one point with the maximum p_F are found, the cluster containing the maximum number of ridge points is determined by a connected component analysis. The seed point is then identified as the center of the cluster. Note that the seed point may not be one of the points at the top or right edge. It can be located within the breast region.

Once the seed point is determined, the tracking of the pectoral boundary is performed upward and downward. Starting from the seed point, the pectoral boundary point is found as the local maximum probability of p_F by a line-by-line search within a small window (window size=16 mm or 20×20 pixels). A box smoothing filter is finally used to smooth the tracked points to generate the final pectoral muscle boundary. Figures 2(g) and 2(h) show the initial tracked pectoral edges and the final identified pectoral boundary after box smoothing.

II.B.3. Performance evaluation

The pectoral muscle boundary manually drawn by an experienced MQSA radiologist for each MLO-view mammogram was used as the reference standard for the evaluation of the performance of our automated pectoral muscle detection method. Let $C=\{c_1,c_2,\cdots,c_p\}$ be computer-identified pectoral boundary that contains p singly connected points, and $R=\{r_1,r_2,\cdots,r_q\}$ be radiologist's manually drawn pectoral muscle boundary that contains q singly connected points. The Euclidean distance between a computer-identified pectoral point c_i and a reference standard point r_j is $\mathrm{Dist}(c_i,r_j)$, or equivalently, $\mathrm{Dist}(r_j,c_i)$. For each MLO-view mammogram, the accuracy of pectoral boundary detection is evaluated by three performance metrics

(1) Percent overlap area (POA)

$$\mathrm{POA}(C,R)=\frac{A_C\cap A_R}{A_C\cup A_R},$$

where A_C and A_R are the computer detected pectoral muscle area and the reference standard pectoral muscle area enclosed by the boundaries C and R, respectively.

(2) Hausdorff distance (Hdist)

$$\mathrm{Hdist}=\max\{\max_{i\in(1,\ldots,p)}\{\min_{j\in(1,\ldots,q)}\{\mathrm{Dist}(c_i,r_j)\}\},\ \max_{j\in(1,\ldots,q)}\{\min_{i\in(1,\ldots,p)}\{\mathrm{Dist}(r_j,c_i)\}\}\}.$$

(3) Average distance (AvgDist)

$$\mathrm{AvgDist}=\frac{1}{2}\left(\frac{1}{p}\sum_{i=1}^{p}\min_{j\in(1,\ldots,q)}\{\mathrm{Dist}(c_i,r_j)\}+\frac{1}{q}\sum_{j=1}^{q}\min_{i\in(1,\ldots,p)}\{\mathrm{Dist}(r_i,c_j)\}\right).$$

The distance measures are calculated in units of mm (1 pixel=0.8 mm).

III. RESULTS

III.A. Accuracy of computerized pectoral muscle detection on digitized film mammograms

Figure 4 shows examples of the computer-identified pectoral boundary corresponding to the mammograms shown in Fig. 1, which are typical pectoral muscle patterns that are challenging for automated detection.

For the test data set of 637 MLO-view mammograms, the mean and standard deviation of the POA are $95.0\pm3.6\%$ (Table I). The mean and standard deviation of the Hdist, and the AvgDist are 3.45 ± 2.16 and 1.12 ± 0.82 mm, respectively. Figure 5(a) shows the cumulative percentage of images having POA greater than a certain value; 91.5% (583/637), 97.3% (620/637), and 98.9% (630/637) of computer detected pectoral muscles had percent overlap area greater than 90%, 85%, and 80%, respectively. For the distance measures [Fig. 5(b)], 85.4% (544/637) and 98.0% (624/637) of computer detected boundaries had Hausdorff distances

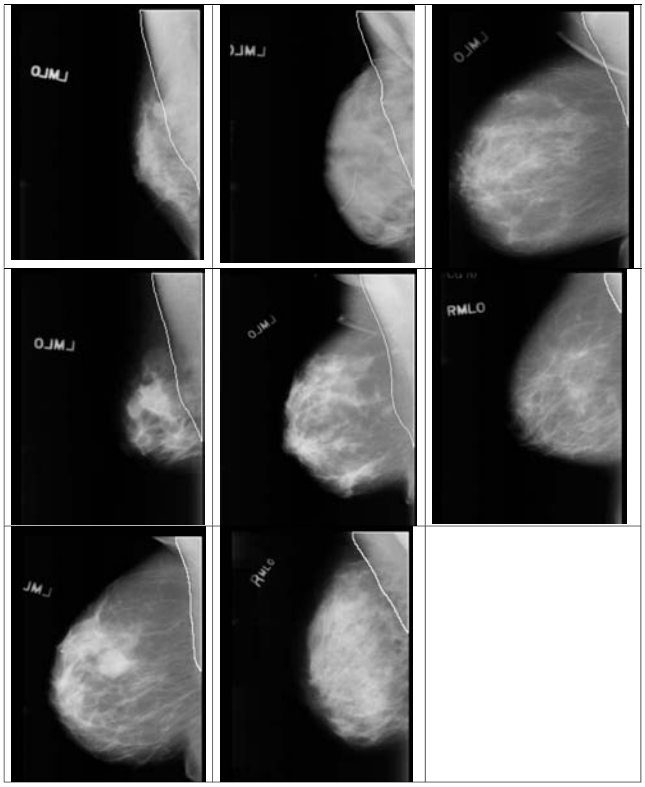
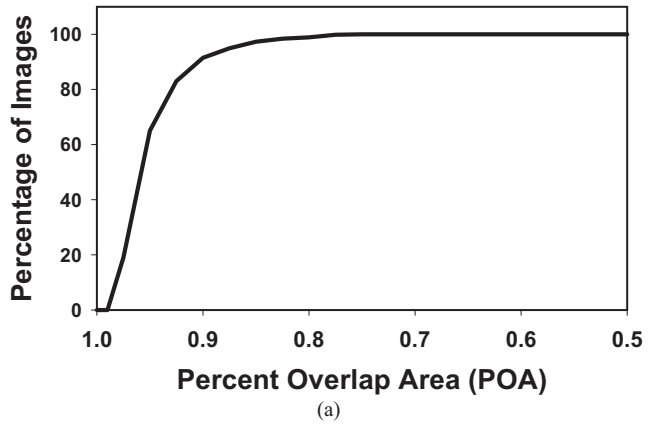
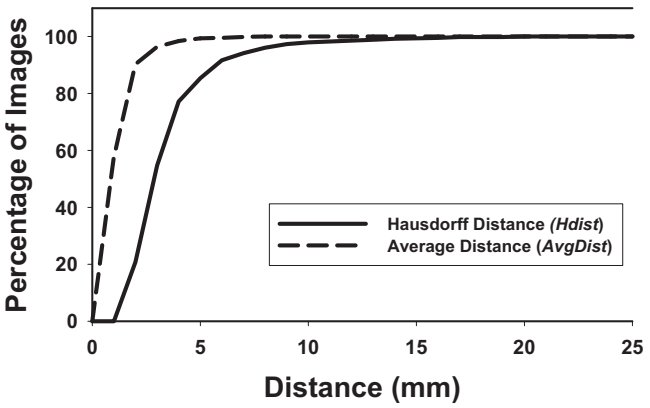


FIG. 4. Examples of pectoral boundary detected by the TFO method superimposed on the mammograms shown in Fig. 1.



(a)



(b)

FIG. 5. Cumulative percentage of images having the performance metrics greater than a certain value for the test data set of 637 MLO-view mammograms. (a) Percent overlap area between TFO-detected pectoral muscle area and reference standard pectoral muscle area and (b) two distance measures between the TFO-detected pectoral boundary and manually drawn pectoral boundary.

within 5 and 10 mm from the reference boundaries, respectively, and 99.4% (633/637) of computer detected pectoral muscle boundaries had average distances within 5 mm from the reference boundaries.

III.B. Accuracy of computerized pectoral muscle detection on full field digital mammograms

Our TFO algorithm was directly applied to the test set of 92 MLO-view FFDMs without retraining, after the raw FFDM images were preprocessed using logarithmic transform and reduced to 800 μm pixel size to match the pixel size of the DFMs used for training and testing of the algorithm. In comparison with radiologist's manually drawn pectoral muscle boundary, the mean and standard deviation of the percent overlap area POA are $89.3 \pm 12.7\%$. The mean and standard deviation of the Hdist and AvgDist are 5.37 ± 4.74 and 2.54 ± 4.04 mm, respectively. For the POA measure, 75.0% (69/92), 87.0% (80/92), and 89.1% (82/92) of computer detected pectoral muscles had percent overlap area greater than 90%, 85%, and 80%, respectively. For the distance measures, 65.2% (60/92) and 91.3% (84/92) of computer detected boundaries had Hausdorff distances within 5 and 10 mm from the reference boundaries, respectively, and 91.3% (84/92) of computer detected pectoral muscle boundaries had average distances within 5 mm from the reference boundaries.

TABLE I. The agreement of computerized pectoral boundary detection with reference to an experienced radiologist's manually drawn boundaries for 637 MLO-view DFMs and 92 MLO-view FFDMs.

		POA (%)	Hdist (mm)	AvgDist (mm)
DFM	Mean	95.0	3.45	1.12
	Standard deviation	3.6	2.16	0.82
FFDM	Mean	89.3	5.37	2.54
	Standard deviation	12.7	4.74	4.04

III.C. Observer variability for identifying pectoral muscle

The subset of 106 MLO-view digitized film mammograms collected at the University of Michigan was used to study the effect of the variability in the reference standard on the performance evaluation. For each mammogram, the pectoral muscle boundary was manually drawn by the same experienced MQSA radiologist twice (R1 and R2) with an interval of about one year. Table II shows the agreement between the two hand-drawn pectoral muscle boundaries by the same radiologist and between the computer detected pectoral muscle boundary and the radiologist's manual segmen-

TABLE II. Effect of radiologist's variability on evaluation of pectoral boundary segmentation. The comparison between an experienced MQSA radiologist's two manual segmentations and the segmentation by the new TFO method for 106 MLO-view mammograms is shown. R1 and R2 denote the first reading and second reading of the same radiologist.

		POA (%)	Hdist (mm)	AvgDist (mm)
R1 vs R2	Mean	92.8	4.44	1.73
	Std Dev	4.85	2.62	1.31
TFO vs R1	Mean	93.3	4.13	1.46
	Std Dev	5.1	2.60	1.18
TFO vs R2	Mean	91.3	4.28	1.97
	Std Dev	6.2	2.96	2.09

tation. For the intraobserver variability evaluation, the results show that the average and the standard deviation of POA, the Hdist, and the AvgDist are $92.8\% \pm 4.85\%$, 4.44 ± 2.62 mm, and 1.73 ± 1.31 mm, respectively. Figure 6 shows the cumulative percentage of images having performance metrics greater than a certain value. For the POA measure, 81.1% (86/106), 95.3% (101/106), and 98.1% (104/106) of the radiologist's two segmentations had POA larger than 90%, 85%, and 80%, respectively. For the distance measures, 72.6% (77/106) and 97.2% (103/106) of the radiologist's two segmentations had Hausdorff distances within 5 and 10 mm, respectively, and 99.1% (105/106) of the two manually drawn boundaries had AvgDist within 5 mm. The differences in the average performance metrics between the computer detected pectoral boundaries and the radiologist's two segmentations are 2.0% for POA and smaller than 0.6 mm for the two distance measures (Table II).

Table III shows the two-tailed *P*-values estimated from the paired *t*-test for the differences in the corresponding performance measures between pairs of the three comparisons. With the radiologist's first reading (R1) as the reference standard, the differences in performance between the radiologist's two hand segmentations and between the computer's segmentation and R1 did not reach statistical significance for any of the three performance metrics ($p > 0.05$). With the radiologist's second reading (R2) as the reference standard, the paired *t*-test shows that the difference was statistically significant for the POA measure but did not reach statistical significance for the two distance measures. Similar results were observed for the differences in the performance measures when the computer segmentation was evaluated using R1 as reference standard or using R2 as the reference standard. Figure 7 shows examples of mammograms that had large variations in radiologist's manually drawn pectoral muscle boundaries. The computer detected boundaries are also shown for comparison.

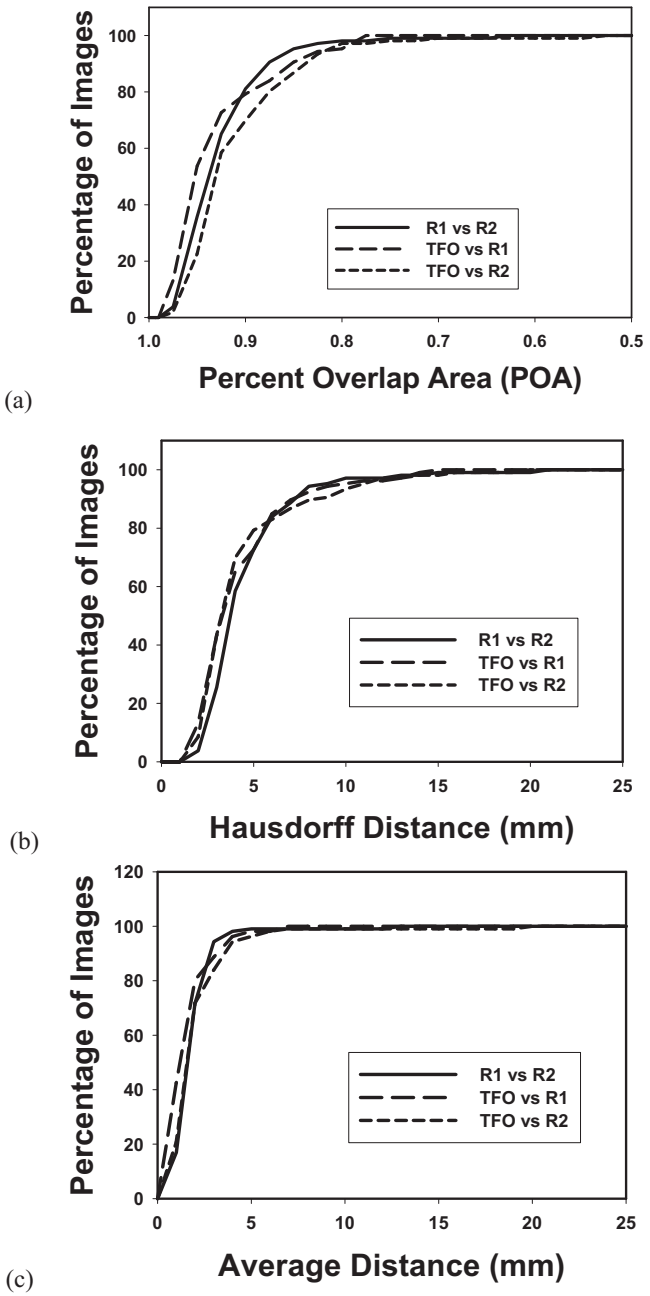


FIG. 6. Cumulative percentage of images having the performance metrics greater than a certain value for the subset of 106 MLO-view mammograms: (a) Percent overlap area, (b) Hausdorff distance, and (c) average distance comparing the pectoral boundaries between two segmentations (R1 and R2) by the same radiologist, between TFO segmentation and R1, and between TFO segmentation and R2.

IV. DISCUSSION

One of the advantages of our TFO method is the utilization of the combined *a priori* knowledge, local and global information for pectoral muscle detection. The *a priori* knowledge on pectoral muscle is its approximate direction and relatively high gray level intensity. The local information at a pixel is represented by the high gradient in a direction approximately normal to the pectoral boundary, while the global information is represented by the relationship between the potential pectoral muscle boundary points.

We designed two new GDK filters to extract local and global information. The first GDK filter with 11×11 pixels (8.8×8.8 mm^2) kernel was designed to enhance edges locally along the assumed direction of pectoral muscle for the calculation of the texture orientation image,

TABLE III. The two-tailed *P*-values estimated from the paired *t*-test on the differences in the three performance measures between pairs of three comparisons: The pectoral boundaries between two segmentations (R1 and R2) by the same radiologist, between TFO segmentation and R1, and between TFO segmentation and R2.

	POA	Hdist	AvgDist
R1 and R2 vs TFO and R1	0.320	0.258	0.060
R1 and R2 vs TFO and R2	0.0008	0.579	0.064
TFO and R1 vs TFO and R2	0.0001	0.448	0.065

which depicts global information of the location and the direction of the pectoral muscle. The smoothed texture orientation image was further enhanced to be a ridge image by the second GDK filter in a 9×9 pixels (7.2×7.2 mm^2) local region. The two GDK filters were designed to enhance the edges that exhibited line structures in the regions with a certain size. The sizes of the filters were determined experimentally using the training set and the selected sizes were fixed for all test images. In this study, the pixel size of the DFMs and FFDMs was reduced to 0.8×0.8 mm^2 for identification of the pectoral muscles both by the radiologists and by the computer. This low resolution was chosen in order to increase the computational efficiency and reduce noise. If high resolution images are used, the sizes of the enhancement filters can be scaled proportionally but the effect of noise will have to be reduced appropriately.

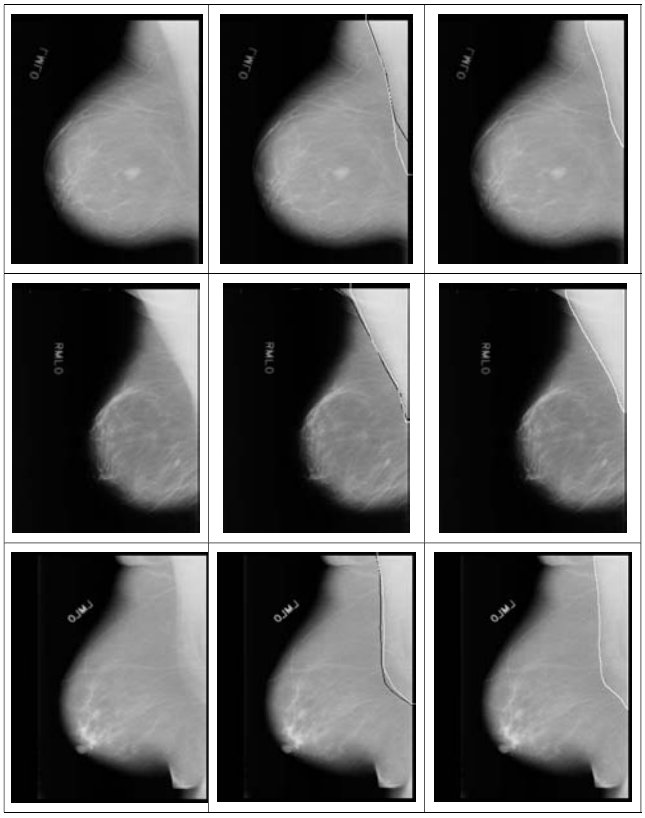


FIG. 7. Examples of radiologist's intraobserver variability and TFO-detected pectoral boundary. First column: Original images; second column: Radiologist's first (black) and second (white) reading; third column: TFO-detected boundaries.

Because a lot of line structures are extracted from the ridge image after edge enhancement, we developed a shortest-path search method to further utilize the global relationship among the ridge points. For a pair of starting and ending points, the shortest-path search method identifies a path along the direction derived from the starting and ending points. The probability image of the ridge points, in which the value of a ridge point pixel represents the probability of being a point on the shortest paths, therefore contains the global information indicating the relationship of this ridge point to other ridge points and pectoral boundary. The importance of combining global and local information is demonstrated by the example shown in Fig. 2, in which the pectoral muscle was obscured by dense tissue and gaps and noisy ridge segments occurred in the region of the true pectoral muscle boundary [Fig. 2(e)]. Tracking of the pectoral edges might fail if it was performed only locally by searching for strong edges. The tracking of pectoral edges was facilitated by searching for the ridge points that are most likely to be a true pectoral edge as indicated by the high probability values derived from the shortest-path method.

From the study of intraobserver variability as described in Tables II and III, it can be seen that the agreement between the two readings by the same radiologist is similar to the agreement between the computer and each of the two readings, in terms of the three performance metrics. As shown in Fig. 7, one of the major reasons for disagreement between radiologist's two readings on a given image is the weak pectoral edges (shown in the first and the third rows of images). For the pectoral muscle that did not depict a clear edge, the radiologist judged subjectively where the edge was. For example, our intraobserver study indicated that subjective identification of the pectoral edges by the same radiologist two different times can deviate as much as 20.8 mm in terms of Hausdorff distance measurement. Using the TFO method, the largest deviations in Hausdorff distance were 14.4 and 19.0 mm, respectively, when radiologist's first and second reading were used as reference standard. These results indicate that the TFO method can detect pectoral muscle with errors within the intraobserver variability of an experienced radiologist.

FFDMs are increasingly used in screening and diagnosis. Image analysis algorithms developed for DFMs have to be adapted to FFDMs. We conducted a preliminary study to evaluate whether our TFO algorithm trained with DFMs can be generalized to FFDMs. The TFO algorithm was applied directly to the FFDMs without any retraining. The results demonstrate that the algorithm could process all FFDM images and find the location of the pectoral muscles. Although the TFO algorithm did not track the pectoral muscle boundary in FFDM as closely as that in DFMs, it self-adapted to the potential differences in gray level and noise characteristics between the two modalities. Thus, as expected, the same method can be used for segmenting the pectoral muscle automatically on FFDM and DFM. Still, minor adjustments of

TABLE IV. Comparison with reported studies using performance metrics: Mean and standard deviation of FP and FN rates, the Hausdorff distance, and the percentage of images with the conditions set in the reported studies, where the FP and FN rates are considered to be accurate (<5%), acceptable (5% to 10%), and unacceptable (>10%). The study using the Radon method used 15% as the upper threshold between acceptable and unacceptable. The data not provided in the publications are marked as NA.

Method	Hough	Gabor	AP	MST	Radon	TFO
Number of images in test set	84	84	84	82	540	637
FP (mean ± std dev)	1.98 ± 6.09	0.58 ± 4.11	3.71	2.55	8.99 ± 38.72	2.33 ± 3.10
FN (mean ± std dev)	25.19 ± 19.14	5.77 ± 4.83	5.95	11.68	9.13 ± 11.87	2.88 ± 3.19
Percentage of images with FP < 5% and FN < 5%	11.90%	53.57%	59.52%	48.78%	28.89%	74.69%
min(FP,FN) < 5% and 5% < max(FP,FN) < 10%	NA	NA	21.43%	24.39%	NA	17.93%
min(FP,FN) < 5% and max(FP,FN) > 10%	NA	NA	13.10%	21.95%	NA	6.76%
5% < FP < 10% and 5% < FN < 10%	9.52%	26.19%	0.00%	0.00%	NA	0.31%
5% < min(FP,FN) < 10% and max(FP,FN) > 10%	NA	NA	0.00%	1.22%	NA	0.31%
FP > 10% and FN > 10%	78.57%	20.24%	5.95%	3.66%	NA	0.00%
5% < FP < 15% and 5% < FN < 15%	NA	NA	NA	NA	40.74%	0.63%
FP > 15% and FN > 15%	NA	NA	NA	NA	30.37%	0.00%
Hausdorff distance (mm)	7.08 ± 5.26	3.84 ± 1.73	NA	NA	12.45 ± 22.96	3.45 ± 2.16

the processing parameters may be beneficial in improving the accuracy for segmentation in FFDMs to the same level as in DFMs and will be pursued in future studies.

Table IV shows the comparison between our TFO method and five reported methods for the performance metrics provided in their publications: The mean and standard deviation of FP and FN rates and the Hausdorff distance in comparison to radiologist's manually segmented pectoral muscle region. The five published methods are based on Hough transform,[6] Gabor filter,[6,14] AP and MSTs,[7] and Radon transform,[9] respectively. In the study using Hough transform and Gabor filter, pectoral muscle segmentation was considered to be accurate when both FP and FN rates were less than 5%, acceptable when 5% < FP and FN < 10%, and unacceptable if both FP and FN > 10%. In the study using the Radon method, 15% was considered to be the upper threshold between acceptable and unacceptable. However, the above counting criteria that were chosen in these three studies were not able to include all situations observed in studies using the AP and MST methods[7] and our TFO method. The study using the AP and MST methods[7] designed a complete set of conditions that included all situations in their studies. To compare the performance to the reported studies, the performance metrics under different conditions were calculated accordingly for our method. This comparison shows that our TFO method achieved higher performance than the previous methods. However, it is known that algorithm performance may depend on the data set and there may be larger variances in results estimated with smaller data sets. Comparison of the performances among different methods can be more definitive only if they are evaluated using a large common data set with common reference standards.

In our previous study of automated estimation of mammographic density using the MDEST method,[1] the breast region was extracted by breast boundary detection and the pectoral muscle in an MLO-view mammogram was excluded by a gradient-based pectoral muscle trimming (GPMT) method. To evaluate the effect on breast density estimation by the new TFO method, the performance of MDEST using the TFO algorithm for pectoral muscle detection was compared with that using the GPMT method. In the data set containing the 531 MLO-view DFMs from the genetic study of breast density in the Amish population,[13] an experienced radiologist provided the reference standard of percent breast density by manual thresholding of the dense area on each mammogram using a computer interface. The MDEST with the TFO method and with the GPMT method for pectoral muscle segmentation were applied to the 531 DFMs and the percent breast density was computed using each method. The accuracy of the breast density estimation by our MDEST method was influenced not only by the pectoral muscle trimming, but also by other factors, such as the breast boundary detection, the automated classification of breast density patterns, and the determination of the gray level threshold in each pattern. To compare the performance of the TFO method to that of the GPMT method, we selected the subset of mammograms for which the breast density estimation by MDEST was mainly affected by the pectoral muscle detection methods so that the pectoral muscle trimming error would not be masked by the other errors. For the 262 mammograms that the difference between the pectoral areas segmented by TFO and GPMT was larger than 200 pixels, the difference in the percent breast density between the two methods was statistically significant ($p < 0.05$), and the correlation between the MDEST-estimated percent density and the radiologist's reference standard was improved from 0.87 to 0.89 when the TFO method was used with MDEST in place of the GPMT method. The comparison demonstrated that the improved pectoral muscle identification using the TFO method can improve the accuracy of breast density estimation.

The major limitation of our study was that the reference standard was provided by only one radiologist such that there was a lack of interobserver variability analysis. We will continue to expand the data set and conduct interobserver variability analysis in future studies.

V. CONCLUSION

Accurate identification of the pectoral muscle on MLO-view mammograms is challenging due to the complicated patterns of pectoral muscle, improperly positioned mammograms, variation in image quality, and dense glandular tissues overlapping with the pectoral muscle region. In this work, we developed an automated pectoral muscle boundary enhancement and tracking method based on texture-field orientation that utilizes a combination of *a priori* knowledge, local and global image information. The results demonstrated that the pectoral muscle can be identified accurately by our TFO method compared to an experienced radiologist's manual segmentation as reference standard. The study of intraobserver variability demonstrated that the average deviation of the boundary detected by the TFO method from the radiologist's manually drawn boundary had approximately the same magnitude as the intraobserver variation in terms of three performance metrics. The TFO method had higher accuracy than the published methods in terms of the false positive and false negative rates in the segmented region and the Hausdorff distance measure. The preliminary study of testing the TFO algorithm on FFDMs demonstrated that the algorithm could achieve reasonable performance on FFDMs without retraining, indicating its adaptability to the new mammographic modality. An automated pectoral muscle identification method will provide a foundation for many mammographic image analysis tasks in CAD applications.

ACKNOWLEDGMENTS

This work was supported by USPHS Grant Nos. RO1 CA95153 and R33 CA120234 (to H.P.C.). The genetic study of breast density in the Amish population is supported by USPHS Grant No. RO1 CA122844 (to J.A.D.). The content of this paper does not necessarily reflect the position of the government and no official endorsement of any equipment and product of any companies mentioned should be inferred.

[a]Author to whom correspondence should be addressed. Electronic mail: chuan@umich.edu; Telephone: 734-647-8554; Fax: 734-615-5513.

[1]C. Zhou, H. P. Chan, N. Petrick, M. A. Helvie, M. M. Goodsitt, B. Sahiner, and L. M. Hadjiiski, "Computerized image analysis: Estimation of breast density on mammograms," Med. Phys. **28**, 1056–1069 (2001).
[2]S. Paquerault, N. Petrick, H. P. Chan, B. Sahiner, and M. A. Helvie, "Improvement of computerized mass detection on mammograms: Fusion of two-view information," Med. Phys. **29**, 238–247 (2002).
[3]B. Sahiner, H.-P. Chan, L. M. Hadjiiski, M. A. Helvie, C. Paramagul, J. Ge, J. Wei, and C. Zhou, "Joint two-view information for computerized detection of microcalcifications on mammograms," Med. Phys. **33**, 2574–2585 (2006).
[4]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).
[5]N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," Phys. Med. Biol. **43**, 365–378 (1998).
[6]R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic identification of the pectoral muscle in mammograms," IEEE Trans. Med. Imaging **23**, 232–245 (2004).
[7]F. Ma, M. Bajger, J. P. Slavotinek, and M. J. Bottema, "Two graph theory based methods for identifying the pectoral muscle in mammograms," Pattern Recogn. **40**, 2592–2602 (2007).
[8]J. Suckling, D. R. Dance, E. Moskovic, D. J. Lewis, and S. G. Blacker, "Segmentation of mammograms using multiple linked self-organizing neural networks," Med. Phys. **22**, 145–152 (1995).
[9]S. K. Kinoshita, P. M. Azevedo-Marques, R. R. Pereira, Jr., J. A. H. Rodrigues, and R. M. Rangayyan, "Radon-domain detection of the nipple and the pectoral muscle in mammograms," J. Digit Imaging **21**, 37–49 (2008).
[10]S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," IEEE Trans. Med. Imaging **23**, 1129–1140 (2004).
[11]S. R. Aylward, B. M. Hemminger, and E. D. Pisano, "Mixture modeling for digital mammogram display and analysis," in *Digital Mammography*, Computational Imaging and Vision Series Vol. 13, edited by N. Karssemeijer *et al.* (Kluwer Academic Publishers, Dordrecht, 1998), pp. 305–312.
[12]D. Raba, A. Oliver, J. Marti, M. Peracaula, and J. Espunya, "Breast segmentation with pectoral muscle suppression on digital mammograms," *Pattern Recognition and Image Analysis*, Lecture Notes in Computer Science Vol. 3523 (Springer-Verlag, Berlin Heidelberg, 2005), pp. 471–478.
[13]J. A. Douglas, M.-H. Roy-Gagnon, C. Zhou, B. D. Mitchell, A. R. Shuldiner, H.-P. Chan, and M. A. Helvie, "Mammographic breast density—Evidence for genetic correlations with established breast cancer risk factors," Cancer Epidemiol. Biomarkers Prev. **17**, 3509–3516 (2008).
[14]A. E. Burgess, "On the noise variance of a digital mammography system," Med. Phys. **31**, 1987–1995 (2004).
[15]American College Of Radiology, *Breast Imaging—Reporting and Data System (BI_RADS)*, 3rd ed. (American College of Radiology, Reston, 1998).
[16]A. R. Rao and B. G. Schunck, "Computing oriented texture fields," Comput. Vis. Graph. Image Process. (CVGIP): Graphical Models Image Processing **53**, 157–185 (1991).
[17]D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell. **24**, 603–619 (2002).
[18]E. W. Dijkstra, "A note on two problems in connexion with graphs," Numer. Math. **1**, 269–271 (1959).