Robust Automatic Pectoral Muscle Segmentation from Mammograms Using Texture Gradient and Euclidean Distance Regression

Vibha Bafna Bora¹ · Ashwin G. Kothari² · Avinash G. Keskar²

Published online: 11 August 2015
© Society for Imaging Informatics in Medicine 2015

Abstract In computer-aided diagnosis (CAD) of mediolateral oblique (MLO) view of mammogram, the accuracy of tissue segmentation highly depends on the exclusion of pectoral muscle. Robust methods for such exclusions are essential as the normal presence of pectoral muscle can bias the decision of CAD. In this paper, a novel texture gradient-based approach for automatic segmentation of pectoral muscle is proposed. The pectoral edge is initially approximated to a straight line by applying Hough transform on Probable Texture Gradient (PTG) map of the mammogram followed by block averaging with the aid of approximated line. Furthermore, a smooth pectoral muscle curve is achieved with proposed Euclidean Distance Regression (EDR) technique and polynomial modeling. The algorithm is robust to texture and overlapping fibro glandular tissues. The method is validated with 340 MLO views from three databases—including 200 randomly selected scanned film images from miniMIAS, 100 computed radiography images and 40 full-field digital mammogram images. Qualitatively, 96.75 % of the pectoral muscles are segmented with an acceptable pectoral score index. The proposed method not only outperforms state-of-the-art approaches but also accurately quantifies the pectoral edge. Thus, its high accuracy and

relatively quick processing time clearly justify its suitability for CAD.

Keywords Computer aided diagnosis · Hough transform · Euclidean distance regression · Mammograms · Pectoral muscle detection · Texture gradient

Introduction

In CAD system for mammography, three anatomical landmarks, i.e., pectoral muscle, breast border, and nipple have to be first extracted automatically before the mammogram is to be analyzed for cancer detection [1]. In a high-quality mammogram, the pectoral muscle should be seen at the level of the nipple or below [2]. Consequently, the pectoral muscle always appears as a high-intensity triangular region in the upper portion of the mediolateral oblique (MLO) view of mammogram.

Accurate segmentation of the pectoral muscle in MLO view is very useful in many areas of automated mammographic analysis. In an automated breast tissue density quantification method, the pectoral muscle must be excluded completely. If this step is not followed, fatty tissues will cause an overestimation of the breast density [3]. On the other hand, if the triangular pectoral region inadvertently includes some of the dense glandular area adjacent to the pectoral region, the mammogram density may be underestimated. Moreover, as mentioned by Gupta et al. [4], due to the similarity between pectoral muscles and the mammographic glandular parenchyma, a significantly high number of false positives can occur in mass detection procedures.

The importance of pectoral muscle segmentation makes it a great source of encouragement for researchers working on CAD. In the recent past, several works were proposed on automatic segmentation of pectoral muscle. Karssemeijer

✉ Vibha Bafna Bora
vibha.bora@raisoni.net

Ashwin G. Kothari
ashwinkothari@ece.vnit.ac.in

Avinash G. Keskar
agkeskar@ece.vnit.ac.in

¹ Department of Electronics and Telecommunication Engineering, G. H. Rasoni College of Engineering, CRPF Gate No. 3 Hingna Road, Nagpur 440016, Maharashtra, India

² Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, South Ambazari Road, Nagpur 440010, Maharashtra, India

et al. [5] applied Hough transform and thresholding method to detect the pectoral muscle with an accuracy of 67 %. Aylward et al. [6] first approximated the pectoral muscle edge by a straight line. This was later refined using a voting scheme. Ferrari et al. [7] implemented and modified the Hough transform, and segmented 21.42 % of pectoral muscle with false positive (FP) and false negative (FN) <10 %. Kinoshita et al. [8] applied Radon transformation to segment muscle with an accuracy of 69.6 % with FP and FN <15 %. However, all these methods consider the muscle to be straight and exhibit a lack in considering its actual curved nature. The obvious outcome of this lacuna is the moderate accuracy of these methods.

Further, Ferrari et al. [9] implemented an efficient detection algorithm based on Gabor wavelet to obtain a smooth pectoral edge with an accuracy of 79.76 % for FP and FN <10 %. However, this method was very time-consuming. Kwok et al. [10] used an iterative threshold method to detect the pectoral muscle approximately and then applied an iterative cliff algorithm with an accuracy of 83.9 % with no quantitative evaluation. Also, Kwok's algorithm cannot segment highly dense and textured muscle. Raba et al. [11] combined an adaptive histogram approach and a selective region growing method to obtain the pectoral muscle edge. However, in mammograms where the fibro glandular tissues appear near the pectoral muscle, this method may result in over-segmentation. Ma et al. [12] applied two graph-based detection methods to segment the muscle where 80 % images had average error less than 2 mm. Zhou et al. [13] realized that the texture-field orientation (TFO) method performs quantitative assessment with high accuracy but it shows a high false-positive region near textured muscle in mammograms of type BI-RADS-4. This region is important for the analysis of lymph node detection, density estimation, etc. Chakraborty et al. [14] applied an average gradient and shape-based feature method to delineate muscle with an accuracy of 71 % for FP and FN <10 %. However, this method failed to detect small and low contrast pectoral muscle and, thus, left room for improvement in accuracy.

The aim of the proposed work is to develop a robust technique for the detection of pectoral muscle on MLO view which must consider curvature as well as the overlapping of muscle with dense tissues and variable textures accurately for any shape and size. The technique should be applicable to any kind of mammographic image. In the proposed method, Hough transform is used to obtain straight line approximation on a probable texture gradient map and Euclidean distance regression is applied to refine the detection results. The final segmented pectoral muscle region is assessed by two radiologists and compared with two state-of-the-art methods for pectoral muscle detection. A total of 340 MLO mammograms from screen film, digitized, and digital databases are used in this study.

Materials

The methodology presented in this work is applied on three types of database:

1. Scanned film images

The images from miniMIAS database [15] originate from a film-screen mammographic imaging process in the United Kingdom National Breast Screening Program. These images have a spatial resolution of $1,024 \times 1,024$, pixel resolution of $200 \mu\text{m}$ and a gray value resolution of 8 bit. The experiment is conducted on 200 MLO mammograms from miniMIAS consisting of simple, curved, and complex muscles.

2. Computed radiography images (CR):

The proposed algorithm is applied to 100 MLO mammograms collected from Central India Cancer Research Institute (CICRI) Hospital, Nagpur, India. The images have been digitized using scanner GE Senographe 600T. These images have a spatial resolution of $2,370 \times 1,770$, pixel resolution of $100 \mu\text{m}$ and a gray value resolution of 12 bit. All images are

Fig. 1 Flow of the proposed method for pectoral muscle detection

Fig. 2 Oriented cropped mammogram with pectoral muscle in the *top left corner*. PQ is the *straight line* approximation to the pectoral edge. $P(P_w, 1)$ and $Q(1, P_h)$ are the coordinates of top edge and bottom edge of the pectoral muscle, respectively. M is the width and N is the height of the mammogram

down sampled by a factor of 2, without loss of significant information to $1,185 \times 885$ pixels, with an effective resolution of 200- μm pixel and held at 8-bit gray level scale images in order to simplify calculations.

3. Full-field digital mammograms images (FFDM):

The database consists of 40 digital MLO mammograms acquired with Hologic Selenia Dimensions Digital Tomosynthesis Mammography unit at Star Imaging and Research Centre, Pune, India. These images have a spatial resolution of $3,328 \times 4,096$, pixel resolution of 70 μm , and a gray value resolution of 12 bit. The images were down sampled by a factor of 4 which transforms the size to $832 \times 1,024$ and pixel resolution to 280 μm .

Method

The flow of the proposed method for pectoral muscle detection for CAD is as shown in Fig. 1. The *preprocessing* steps are: *median filtering*, *artifacts removal* like labels, tags, etc., *orientation correction*, and *contrast enhancement*. Using morphological operation and connected component algorithm, artifacts were removed. As in mammography, there is no orientation standard and known orientation makes detection of muscle somewhat easier than unknown position [16], so muscles are correctly oriented with the help of pattern-based Radon transform [17]. As a result, each mammogram has chest wall on the left side and pectoral muscle in the upper left corner as shown in Fig. 2. The breast background is also cropped in order to reduce processing time and image file size without losing anatomical information. Furthermore, to improve image contrast and accuracy of muscle segmentation, enhancement [18, 19] is done.

Probable Texture Gradient Map and Straight Line Approximation

Shannon entropy [19] is used to characterize the texture of muscle. Because of the sharp intensity variation of the pectoral edge, the muscle edge pixels indicate high entropy. Instead of calculating global entropy, entropy in the 9×9 -pixel neighborhood is calculated to extract the texture information locally. Furthermore, in properly oriented mammograms the pectoral muscle edge is oriented in the first quadrant and in the direction perpendicular to the gradient at each pixel location. With the help of the two above-mentioned features and intensity, a probable texture gradient (PTG) map is plotted as the collection of all pixels that have entropy above the threshold entropy,

Fig. 3 Straight line approximation: **a** preprocessed original mammogram mdb061 from miniMIAS database; **b** enhanced mammogram; **c** probable texture gradient map (PTG); **d** Initial ROI and straight line estimation using Hough Transform (*white line*); **e** Straight line overlapped on mammogram

Fig. 4 Non-linear regression: **a** polynomial model for nonlinear regression of y vs. x for pectoral edge data. **b** 95 % Confidence band; **c** 95 % prediction band

orientation between 30° and 80° , and intensities greater than 50 % of the maximum intensity.

The Hough Transform [19] can detect straight lines in an image. However, as dense tissues overlap the pectoral edge, accurate straight line approximation also becomes challenging. Given this, instead of applying the Hough transform to an intensity map, it is applied to a PTG map to obtain accurate straight approximated muscle. Out of n pixels from a PTG map, the Hough transform finds sets of pixels from the accumulator cells that form the strongest straight line as the approximated pectoral edge. Let the coordinates of the straight pectoral edge be $P(P_w, 1)$ and $Q(1, P_h)$, where P_w is the width and P_h is the height of the pectoral muscle as shown in Fig. 2.

Steps for straight line approximation of the muscle using Hough transform on PTG map are as follows:

1. Consider oriented mammogram $I(x, y)$ (Fig. 3a).
2. Improve contrast of the oriented mammogram (Fig. 3b).
3. Calculate entropy $H(x, y)$ in the 9×9 -pixel neighborhood.
4. Apply Sobel operator to obtain gradient orientation angle $\alpha(x, y)$.

5. Obtain probable texture gradient (PTG) map of the mammogram.

$$PTG(x, y) = \left\{ 1 \mid H(x, y) \geq H_T, 30^\circ \leq \beta(x, y) \leq 80^\circ, I(x, y) \geq 0.5 I_{\max} \right\} \quad (1)$$

where H_T the entropy threshold is equal to H_{mean} ; the mean value of the local entropy for miniMIAS images or is equal to H_{UQ} ; upper quartile of the local entropy for CR and FFDM images. $I_{\max}$, the maximum intensity of the image. β , the pectoral edge orientation perpendicular to gradient given as $\beta(x, y) = \alpha(x, y) - 90^\circ$ where α is the angle of the gradient (Fig. 3c).

6. Consider the initial Region of interest (ROI) from PTG for straight line estimation.

$$ROI = \{(x, y) : 0 \leq x \leq M \text{ and } 0 \leq y \leq 3N/4\} \quad (2)$$

where M and N are width and height of the mammogram, respectively (Fig. 3d).

7. Apply Hough transform to obtain strongest straight line as approximated muscle edge (white line) (Fig. 3d, e).
8. If the coordinates of the end point of the strongest line contain first row and first column, then the strongest line is the approximated straight pectoral edge. If not, change the ROI by 20 % and repeat steps 6–8.

Fig. 5 Euclidean distance regression on pectoral muscle

Region of Interest for Pectoral Muscle and Block Averaging

After a straight line approximation, the Region of Interest for Pectoral Muscle (ROIPM) is obtained by shifting the approximated straight pectoral edge left at $(P_w - N_p, 1)$ to $(1, P_h - N_p)$ and right at $(P_w + N_p, 1)$ to $(1, P_h + N_p)$ on a PTG map. This approximation allows for accurate prediction of the pectoral edge and can remove distant outliers, where N_p is the

Fig. 6 Pectoral muscle segmentation using Euclidean distance regression: **a** Region of interest for pectoral muscle (ROIPM); **b** 5×5 Block averaging on ROIPM; **c** Euclidean distance regression applied to ROIPM (dark gray); **d** Zoomed view of white rectangular area shown in part c; **e** Pectoral edge smoothed using cubic polynomial

horizontal distance in terms of number of pixels. Furthermore, before muscle refining, $m \times m$ block averaging is done on ROIPM to obtain important clustered pixels. Blocks with a mean value of 0.51 or higher are considered as 1 while others 0. This step removes those pixels of the ROIPM that are least important or scattered.

Euclidean Distance Regression

Euclidean distance regression (EDR) is designed to refine the straight line approximation into a curve that can demarcate any type of pectoral edge more accurately. A statistical tool such as regression [20] is used to derive the relationship among variables. The regression function is given by $y=f(x_1, x_2, \dots, x_n)$, where x is the horizontal and y the vertical coordinates. Given the pectoral edge points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ the best fit $y = a_0 + a_1x + \dots + a_mx^m$ can be obtained as shown in Fig. 4a where $(m \leq n-2)$.

The regression plot of the best fit curve can include the 95 % confidence band or 95 % prediction band. The confidence band indicates that a 95 % true best fit curve lies within the

confidence band as shown in Fig. 4b. The prediction band indicates the scattering of the expected pixel of pectoral edge data and accounts for uncertainty in the curve itself as shown in Fig. 4c.

To obtain only those pixels that are more likely to lie on pectoral edge, EDR is applied. The initial value $(P_w, 1)$ will choose that pixel from the next non-zero row of block averaged ROIPM, which is at a minimum Euclidean distance until $Q(1, P_h)$ is reached as shown by solid line in Fig. 5. Thus the pixels of ROIPM are iterated for minimum Euclidean distance to obtain a set of pixels belonging to the confidence band P_c and is given by:

$$P_C = \bigcup_{j=1}^N \left\{ \min D_i(x_i, y_j) \mid i \in M \right\} \quad (3)$$

where D is the Euclidean distance between two pixels and given as $D_i(x_i, y_j) = \sqrt{(x_0 - x_i)^2 + (y_0 - y_j)^2}$ where $i \in M$, $j \in N$ & $(x_0, y_0) = (P_w, 1)$ is the initial value.

The curve obtained from EDR is ragged, and is smoothened using cubic polynomial fitting function with four coefficients. Cubic polynomial function has been

Fig. 7 Block averaging on ROIPM: **a** Preprocessed mammogram mdb002; **b** EDR curve on ROIPM without block averaging; **c** Pectoral muscle detected with false negative of 9.65 %; **d** EDR curve on ROIPM with block averaging; **e** Accurate pectoral muscle detected

chosen because of the pectoral edge shape which is usually curved, and is represented by cubic function, instead of linear function.

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \quad (4)$$

Steps of EDR for accurate pectoral muscle detection are as follows:

1. Extract ROIPM from PTG map using approximated straight line (Fig. 6a).
2. Calculate $m \times m$ block averaging (BAG) on ROIPM (Fig. 6b).

$$\text{BAG}[x : x + (m-1), y : y + (m-1)]$$

$$= \frac{\sum_{s=0}^{(m-1)} \sum_{t=0}^{(m-1)} \text{PTG}(x+s, y+t)}{m^2} \quad (5)$$

where

$$x = 1, m+1, 2m+1, \dots, M-(m-1). \text{ and } y = 1, m+1, 2m+1, \dots, N-(m-1)$$

3. Apply thresholding to block averaged image BAG as discussed previously.
4. Apply Euclidean distance regression (Eq. 3) on ROIPM with initial value as $(P_w, 1)$ and iterate till the point $Q(1, P_h)$ is reached giving ragged boundary (Fig. 6c, d).
5. Apply cubic polynomial modeling (Eq. 4) on ragged muscle to obtain smooth curve (Fig. 6e).

Need of Block Averaging on ROIPM

ROIPM is the region near the straight approximated pectoral muscle with pixels that have relatively high value of intensity, and entropy and gradient angle lying between 30° and 80° . EDR initializes with $P(P_w, 1)$ and chooses successive pixels with minimum Euclidean distance on ROIPM. When block averaging is not done, false-positive pixels are collected by EDR as shown in Fig. 7b. It gives inaccurate edge with FN of 9.65 % as shown in Fig. 7c. However, when block averaging is applied, false-positive pixels that are not significant are removed before applying EDR (Fig. 7d). This gives a more accurate pectoral edge as shown in Fig. 7e compare to EDR without block averaging (Fig. 7b, c).

Trimming of Sharp Convex Muscles

Due to sharp curvatures in some mammograms like mdb233, the approximated straight line covers false-positive region near lower left edge of the muscle as shown in Fig. 8a (white line). Furthermore, when EDR is applied (gray line) to the ROIPM, a few outliers near the lower left edge of the pectoral muscle bias the pectoral muscle shape. Hence, EDR needs trimming at lower left edge. Therefore, the algorithm has to be converged for the pixel selected first; close to the left chest wall as shown in Fig. 8b, c.

Choice of Parameters

An optimum size of N_p is chosen to cover pectoral edge variation for highly concave- or convex-shaped muscle. If a small N_p value is chosen, ROIPM will miss pixels that are important for accurate delineation of the pectoral edge. In contrast, if a large value of N_p is chosen, unwanted pixels in the vicinity may get detected to form an inaccurate shape of the muscle. The empirical value of N_p is chosen by varying it for extreme cases of muscle curvature, and corresponding FP and FN are calculated. The N_p that gives the smallest FP and FN is calculated for extreme cases and selected as optimized $N_p = 70$ pixels.

For block averaging, when the block size is large, e.g., 7×7 , 9×9 , and 15×15 , details of the muscles are lost, thus leading to high FN. Conversely, when the block size is smaller, e.g., 3×3 pixels or 5×5 pixels, muscle details are retained properly. For block averaging in this paper, the value of $m \times m$ is chosen as 5×5 .

Fig. 8 Trimming applied to convex muscle; mdb233. **a** Approximated straight line (white) on PTG map and EDR curve (dark gray) with false-positive pixels in white square; **b** Trimming applied on lower left edge in white square; **c** Accurate pectoral muscle detected

Table 1 Pectoral score index

Pectoral Score Index (PSI)	Maximum tolerable deviation allowed to radiologist R1		Maximum tolerable deviation allowed to radiologist R2		Mean score of maximum tolerable deviation (R1 and R2)		Reference derived from mean score
	FP (%)	FN (%)	FP (%)	FN (%)	FP (%)	FN (%)	
Accurate (1)	6.50	6.35	3.40	3.75	4.95	5.05	(FP and FN) ≤ 5 %.
Adequate (2)	11.25	11.2	9.6	9.5	10.42	10.35	5 % < (FP and FN) ≤ 10 %.
Moderate (3)	16.95	16	13.15	13	15.05	14.5	10 % < (FP and FN) ≤ 15 %.
Inaccurate (4)	–	–	–	–	–	–	(FP and FN) > 15 %.

Performance Metrics

Two radiologists visually inspect the enhanced mammograms to draw the standard pectoral muscle using a cursor for the evaluation of the performance of proposed automated pectoral muscle detection method. Let $P=(p_1, p_2, \dots, p_x)$ be automated pectoral muscle boundary pixels, and $Q_1=(q_{11}, q_{12}, \dots, q_{1y})$ and $Q_2=(q_{21}, q_{22}, \dots, q_{2y})$ be each radiologist's manually drawn standard pectoral muscle boundary. For each MLO view mammogram, the accuracy of algorithm is evaluated by three quantitative performance metrics and the two radiologists' proposed Pectoral Score Index (PSI).

1. *Hausdorff Distance (HD)*: The Hausdorff distance is defined as

$$HD(P, Q) = \max(h(P, Q), h(Q, P)) \quad (6)$$

where $h(P, Q) = \max_{p \in P} \min_{q \in Q} \|p - q\|$

$\|\cdot\|$ is the Euclidean distance and where, $P \in$ acquired pectoral muscle region, $Q \in$ standard pectoral muscle region. The average of the HD with respect to two radiologists is taken.

2. *Pixel percentage of false positive (FP)*: A pixel inside the acquired pectoral edge but outside the reference region.

Mathematically, pixel percentage of false positive is given as,

$$\text{Pixel percentage of FP} = \frac{|P \cup Q| - |Q|}{|Q|} \times 100\% \quad (7)$$

3. *Pixel Percentage of False Negative (FN)*:

A pixel outside the acquired pectoral edge but inside the reference region. Mathematically, pixel percentage of false negative is given as,

$$\text{Pixel percentage of FN} = \frac{|P \cup Q| - |P|}{|Q|} \times 100\% \quad (8)$$

4. *Pectoral Score Index (PSI)*:

The pectoral muscle is evaluated using four level pectoral score index recommended by two radiologists after obtaining segmentation results. Deviation, in the form of FP and FN, is mathematically assessed based on the result given by both radiologists. Mean score of the maximum tolerable deviation allowed by radiologists in terms of FP and FN is calculated and expressed as the Pectoral score index given in Table 1. Radiologists classify the observed images into four categories on a scale of pectoral score index: accurate, adequate, moderate and inadequate. Accurate (1) means the segmentation is almost perfect with a straight line or curve overlapping the pectoral muscle

Fig. 9 Pectoral muscle segmentation results for miniMIAS images. **a** mdb123, **b** mdb010, **c** mdb053, **d** mdb025, **e** mdb277

Fig. 10 Pectoral muscle segmentation results for CR images (a–e)

edge correctly. All such images have (FP and FN) < 5 %. Adequate (2) means the straight line or curve does not precisely overlap the pectoral muscle but is sufficiently accurate, and can be used for diagnosis. Such images have $5\% \leq (\text{FP and FN}) < 10\%$. Moderate (3) means some breast tissue is excluded or some non-breast tissue is included in the muscle segmented result. Images show $10\% \leq (\text{FP and FN}) < 15\%$. Inaccurate (4) means the straight line or curve is not present or completely inaccurate; images have (FP and FN) > 15 %.

The results of the proposed method are compared with two well-known pectoral muscle detection algorithms proposed by Kwok et al. [10] and Kinoshita et al. [8] on the same dataset of images and platform. Tables 2, 3, and 4 summarize the comparative analysis of all these methods for miniMIAS (200 images), CR (100 images) and FFDM (40 images), respectively. The table shows mean and standard deviation for all three metrics calculated between the automatically detected and manually drawn pectoral muscle with number of images marked as accurate, adequate, moderate and inaccurate with reference to PSI. In the present study, pectoral muscle segmentation is considered to be acceptable when (FP and FN) $\leq 10\%$.

Results and Comparative Analysis

The proposed algorithm is applied to three different types of database. To evaluate the efficiency of the method, performance evaluation metrics mentioned earlier are used. Figures 9, 10, and 11 demonstrate the results obtained automatically from the proposed method. As is evident, the proposed algorithm is able to detect muscles which are curved and completely overlapped by dense tissues and texture.

Algorithm Processing Time

The algorithm is implemented on MATLAB ver. R2013b. The computer used is a Dell PC with Intel core 2 Duo processor T6670 clocked at 2.2 GHZ with 2 GB RAM. The average processing time required for preprocessing step is 1.85 s, Hough texture based straight line estimation

Fig. 11 Pectoral muscle segmentation results for FFDM images (a–e)

Table 2 Mean and standard deviation values of the Hausdorff distance, pixel % of False-Positive (FP) and pixel % of False-Negative (FN) and number of images as accurate, adequate, moderate and inaccurate for miniMIAS images

Metrics and PSI	The present study	Kwok et al. [10]	Kinoshita et al. [8]
Hausdorff distance (HD in mm)	3.72±1.67	11.06±17.45	23.56±26.45
FP (%)	1.56±3.12	7.17±6.79	18.34±8.38
FN (%)	2.83±3.17	9.13±7.03	16.13±13.54
Accurate (1) (FP and FN)≤5 %	142 (71.00 %)	137 (68.50 %)	80 (40.00 %)
Adequate (2) 5 %<(FP and FN)≤10 %	51 (25.50 %)	30 (15.00 %)	41 (20.50 %)
Moderate (3) 10 %<(FP and FN)≤15 %	5 (2.50 %)	22 (11.00 %)	19 (9.50 %)
Inaccurate (4) (FP and FN)>15 %	2 (1.00 %)	11 (5.5 %)	60 (30.00 %)
Accuracy (FP and FN≤10 %)	96.50 %	83.5 %	60.50 %

is 2.721 s and Euclidean distance regression with polynomial modeling for curve fitting is 2.089 s. Computational time for valid straight line estimation is longer than EDR. As for valid straight line estimation, if needed, ROI needs to be tuned in iterations. Thus the proposed method requires an average processing time of 4.81 s for muscle detection. The processing time required by Kwok method and Kinoshita method is 6.012 and 3.526 s, respectively. While the Kinoshita method requires less processing time, it is not efficient because it only detects straight edges without taking tissue occlusion into account.

Limitations

The proposed method sometimes fails in the detection of muscle that is small in size and low in contrast. Additionally, it must be noted that the accuracy of the algorithm depends on the accuracy of the straight line approximation of the pectoral edge and ROIPM. The proposed method will prove to be more accurate if the accuracy of the above steps increases.

Conclusion

The strength of the proposed method is that it can detect muscle even when completely obscured by dense tissues. In addition, another advantage is its robustness in dealing with multiple textures and artifacts present on the muscle and its edge. Experimentation also shows that preliminary extraction of ROIPM reduces the chance of false detection of muscle. Furthermore, it can be seen from Tables 2, 3, and 4 that for all types of database, the proposed method outperforms Kwok [10] and Kinoshita [8] method in terms of HD (in millimeter), FP (in percent) and FN (in percent). In Kwok's method, because of extensive intensity overlap between pectoral and non-pectoral regions, the straight line estimation technique is not suitable as it is based on iterative threshold selection that affects curve segmentation [21]. In contrast, in the proposed work the Hough transform applied on a PTG map gives accurate straight line estimations even though muscle may be overlapped by dense tissues. The Kinoshita method demonstrates better

Table 3 Mean and standard deviation values of the Hausdorff distance, pixel % of false-positive (FP), and pixel % of false-negative (FN) and number of images as accurate, adequate, moderate and inaccurate for CR images

Metrics and PSI	The present study	Kwok et.al [10]	Kinoshita et.al [8]
Hausdorff distance (HD in mm)	3.14±2.46	8.56±9.45	21.06±16.58
FP (%)	1.52±3.97	6.87±5.72	13.63±7.97
FN (%)	2.28±3.05	8.13±5.66	15.83±14.50
Accurate (1) (FP and FN)≤5 %	72 (72.00 %)	57 (57.00 %)	28 (28.00 %)
Adequate (2) 5 %<(FP and FN)≤10 %	25 (25.00 %)	28 (28.00 %)	35 (35.00 %)
Moderate (3) 10 %<(FP and FN)≤15 %	2 (2.00 %)	9 (9.00 %)	12 (12.00 %)
Inaccurate (4) (FP and FN) >15 %	1 (1.00 %)	6 (6.00 %)	25 (25.00 %)
Accuracy (FP and FN≤10 %)	97.00 %	85.00 %	63.00 %

Table 4 Mean and standard deviation values of the Hausdorff distance, pixel % of false-positive (FP) and pixel % of false-negative (FN) and number of images as accurate, adequate, moderate, and inaccurate for FFDM images

Metrics and PSI	The present study	Kwok et.al [10]	Kinoshita et.al [8]
Hausdorff distance (HD in mm)	2.82±2.18	7.56±8.66	18.06±15.75
FP (%)	1.04±3.35	6.12±4.11	12.63±9.97
FN (%)	2.23±4.27	7.26±9.66	13.67±17.50
Accurate (1) (FP and FN)≤5 %	28 (70.00 %)	25 (62.50 %)	18 (45.00 %)
Adequate (2) 5 %<(FP and FN)≤10 %	11 (27.50 %)	10 (25.00 %)	10 (25.00 %)
Moderate (3) 10 %<(FP and FN)≤15 %	1 (2.50 %)	3 (7.50 %)	5 (12.50 %)
Inaccurate (4) (FP and FN) >15 %	0 (0.00 %)	2 (5.00 %)	7 (17.50 %)
Accuracy (FP and FN≤10 %)	97.50 %	87.50 %	70.00 %

performance when images have straight-line edges and dissatisfactory performance with other types of edges of the pectoral muscle region [8]. However, in the present study EDR helps to detect curved edges. Hence, the proposed method automatically mimics the decision making capability of a radiologist.

A novel texture gradient-based method for detection of the pectoral muscle has been presented in this work. The application of the Hough transform on a probable texture gradient map instead of intensity map, ROIPM extraction, block averaging and EDR with cubic modeling are used to detect pectoral muscle accurately. Furthermore, the algorithm relies on texture, gradient and intensity to detect pectoral muscle, thus overcoming the limitations of state-of-the-art methods. The proposed algorithm is tested on 340 mammograms from three databases, and in the case of 96.75 % of mammograms, pectoral muscle is segmented with an acceptable pectoral score index. With its high accuracy, robustness, and fast processing time, this method has the potential to be an efficient means of automatic pectoral muscle segmentation of MLO mammograms in CAD.

This method shows scope for future research. To overcome the limitations discussed previously, the present method may be added to a region-based contrast enhancement technique [22], and automatic assessment of mammographic positioning [21]. Moreover, when an image is properly exposed at optimal contrast, the pectoral muscle can be differentiated from other glandular tissue with relative ease.

In addition, while drawing a PTG map additional texture features could improve the accuracy of pectoral segmentation. Furthermore, polynomial modeling can be made adaptive with respect to variations in parameters like the sum of squares due to error (SSE) to reduce FPs and FNs in the detection of muscle.

Acknowledgments The authors would like to thank Dr. Varsha S. Hardas, Consultant Radiologist of Star Imaging & Research Centre, Pune, India, for providing FFDM images and assessing the segmentation results. Thanks is also given to Dr Suchitra Mehta, Central India Cancer Research Institute (CICRI), Nagpur, India, for providing CR Mammograms and also for assessing segmentation results. We also acknowledge the financial support provided by Centre of Excellence on Combedded System: Hybridization of Communication and Embedded Systems under TEQIP 1.2.1 at VNIT, Nagpur.

References

1. Ferrari R, Rangayyan R, Borges R, Frere A: Segmentation of the fibro-glandular disc in mammograms using Gaussian mixture modeling. *Med Biol Eng Comput* 42(3):378–387, 2004
2. Eklund GW, Cardenosa G, Parsons W: Assessing adequacy of mammographic image quality. *Radiology* 190(2):297–307, 1994
3. Saha PK, Udupa JK, Conant EF, Chakraborty DP, Sullivan D: Breast tissue density quantification via digitized mammograms. *IEEE Trans Med Imaging* 20:792–803, 2001
4. Gupta R, Undrill PE: The use of texture analysis to delineate suspicious masses in mammography. *Phys Med Biol* 40:835–855, 1995
5. Karssemeijer N: Automated classification of parenchymal patterns in mammograms. *Phys Med Biol* 43:365–378, 1998
6. Aylward SR, Hemminger BM, Pisano ED: In: Karssemeijer N, et al Eds. *Mixture modeling for digital mammogram display and analysis*. Digital Mammography, Comput Imag and Vision Series, vol. 13. Kluwer Academic Publishers, Dordrecht, 1998, pp 305–312
7. Ferrari RJ, Rangayyan RM, Desautels JEL, Frere AF: Segmentation of mammograms: Identification of the skin—air boundary, pectoral muscle, and fibro-glandular disc. In: Yaffe MJ, Ed. *IWDM 2000. Proc. 5th Int Workshop Digital Mammography*, Toronto, ON, Canada, 573–579, 2001
8. Kinoshita SK, Azevedo Marques PM, Pereira Jr, RR, Rodrigues JA, Rangayyan RM: Radon-domain detection of the nipple and the pectoral muscle in mammograms. *J Digit Imaging* 21(1):37–49, 2008
9. Ferrari RJ, Rangayyan RM, Desautels JEL, Borges RA, Frere AF: Automatic identification of the pectoral muscle in mammograms. *IEEE Trans Med Imaging* 23(2):232–245, 2004

10. Kwok SM, Chandrasekhar R, Attikiouzel Y, Rickard MT: Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. *IEEE Trans Med Imaging* 23(9):1129–1140, 2004
11. Raba D, Oliver A, Marti J, Peracaula M, Espunya J: Breast segmentation with pectoral muscle suppression on digital mammograms. *Pattern Recog and Image Anal, Lecture Notes in Computer Science*, vol. 3523. Springer, Berlin Heidelberg, 2005, pp 471–478
12. Ma F, Bajger M, Slavotinek JP, Bottema MJ: Two graph theory based methods for identifying the pectoral muscle in mammograms. *Pattern Recogn* 40(9):2592–2602, 2007
13. Zhou C, Wei J, Chan H, Paramagul C, Hadjiiski L, Sahiner B, Douglas J: Computerized image analysis: texture-field orientation method for pectoral muscle identification on MLO-view mammograms. *Med Phys* 37(5):2289–2299, 2010
14. Chakraborty J, Mukhopadhyay S, Singla V, Khandelwal N, Bhattacharyya P: Computerized image analysis: texture-field orientation method for pectoral muscle identification on MLO-view mammograms. *J Digit Imaging* 25:387–399, 2012
15. Suckling J, Parker J, Dance D, Astley S, Hutt I, Boggis C, Ricketts I, Stamatakis E, Cerneaz N, Kok S, et al: The mammographic image analysis society digital mammogram database, *Exerpta Medical. Int Congr Ser* 1069:375–378, 1994
16. Mustra M, Grgic M: Robust automatic breast and pectoral muscle segmentation from scanned mammograms. *Signal Process* 93(10): 2817–2827, 2013
17. Jain AK: *Fundamentals of Digital Image Processing* Prentice Hall Information and System Sciences Series. Prentice Hall, NJ, USA, 1989
18. Zuiderveld K: Contrast limited adaptive histogram equalization. *Graphic gems IV*. Academic Press Professional, San Diego, 1994, pp 474–485
19. Gonzalez RC, Woods RE: *Digital image processing*, 3rd edition. Pearson Prentice Hall, New Jersey, 2008
20. Motulsky HJ, Christopoulos A: *Fitting models to biological data using linear and nonlinear regression. A practical Guide to Curve Fitting*. GraphPad Software Inc., San Diego CA, www.graphpad.com, 2003
21. Kwok SM: *Attribute-driven segmentation and analysis of mammograms*. Ph.D. dissertation: University of Western Australia-School of Electrical, Electronic and Computer Engineering, and Centre for Intelligent Information Processing Systems. Ch 4-5, 2004
22. Morrow WM, Paranjape RB, Rangayyan R, Desautels JEL: Region-based contrast enhancement of mammograms. *IEEE Trans Med Imaging* 11(3):392–406, 1992