

An Overview of Pectoral Muscle Extraction Algorithms Applied to Digital Mammograms

Suhas Sapate and Sanjay Talbar

Abstract Substantial numbers of patients are reaching to a progressive breast cancer stage due to increase in the false negatives coming out of cumbersome and tedious job of continuously observing the mammograms in fatigue. Hence, the early detection of cancer with more accuracy is highly expected to reduce the death rate. Computer Aided Detection (CADe) can help radiologists in providing a second opinion increasing the overall accuracy of detection. Pectoral muscle is a predominant density area in most mammograms and may bias the results. Its extraction can increase accuracy and efficiency of cancer detection. This work is intended to provide the researchers a systematic and comprehensive overview of different techniques of pectoral muscle extraction which are categorized into groups based on intensity, region, gradient, transform, probability and polynomial, active contour, graph theory, and soft computing approaches. The performance of all these methods is summarized in tabular form for comparison purpose. The accuracy, efficiency and computational complexities of some selected methods are discussed in view of deciding a best approach in each of the categories.

Keywords Breast cancer · Mammography · Computer aided detection · Pectoral muscle extraction · Segmentation

S. Sapate (✉)

Department of Computer Science and Engineering, SGGS IE&T, Nanded, India
e-mail: sapatesuhas@sggs.ac.in

S. Talbar

Department of Electronics and Telecommunications, SGGS IE&T, Nanded, India
e-mail: sntalbar@sggs.ac.in

1 Introduction

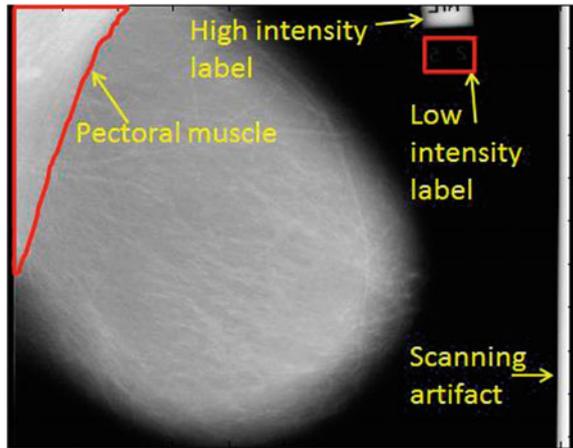
1.1 Overview of Mammography

Breast cancer is the most common cancer and is the second most common cause of cancer deaths in women. Breast cancer incidences worldwide are increasing over the years with more than 1 million new cases reported each year. The chances of success are more if further treatment and therapeutic actions are taken in the early stages of the breast cancer. Thus, early detection plays an important role for improving breast cancer prognosis [1–3].

A mammogram is an X-ray image of the human breast. A careful observation of this image can allow us to identify and evaluate indicators of abnormalities at early stage in the human breast. Screening mammograms are useful in finding likelihood of cancer in patients without any external symptom, whereas patients with some abnormal symptoms or lumps in breast undergo diagnostic mammography. Mammographic images are generated by passing low dose X-ray across each breast. This produces a picture which highlights the soft tissues, dense tissues, pectoral muscle, and fibro-glandular region etc. Expert radiologists can read these mammograms to find out the abnormalities, if are there, in the breast. Any change in two or more mammograms taken over a period, say a year or two, may signify cancer in its early stage. A mammogram can depict changes in the breast up to a year or two before any symptoms observed by patient or physician [1, 4]. If the significant changes are confirmed as early stage cancer, further extensive treatments can be avoided and probability of breast conservation can be improved. Modern mammography machines are with low radiation doses, 0.4 mSv, of X-ray and produces high-quality digital images with 2 views of each breast [5]. In mass screening programs, mammography is the most effective, more popular, cheaper and hence commonly used imaging modality for breast than Magnetic Resonance Imaging (MRI), Nuclear Imaging and Ultrasound [6]. The **mediolateral oblique** (MLO-taken at around 30–70° angle) **view** and **craniocaudal** (CC-top to down) **view**, are two standard mammographic projections used for screening mammography. In MLO view, maximum portion of the breast, including pectoral muscle, is exposed. It is always better to expose maximum portion of pectoral muscle in MLO view to guarantee that each and every part of the breast is covered neatly. Hence, it is the most important projection. Thus, the pectoral muscle on the MLO view is a vital component in confirming correct patient positioning which results in a accurate mammogram of adequately good quality. This is very important to minimize the number of false positives (FP), false negatives (FN) and improve the sensitivity of the mammographic images [7].

As shown in the Fig. 1, the mammographic image consists of various parts apart from the region of interest required for automatic detection of abnormalities. These parts include low and high intensity labels, scanning artifacts etc. all in the background. Pectoral muscle located on top left (left MLO view) or top right (right MLO view) occupies major portion of the breast. The labels, scanning artifacts and

Fig. 1 A typical left MLO view mammographic image



pectoral muscle may increase the computational complexity of the detection process and also cause the reduction in detection accuracy. Hence to remove all these unnecessary parts from the breast region in the mammogram is a vital preprocessing task in CADE system of the breast cancer.

Segmentation of mammographic image into its representative anatomically distinct regions such as background (the non-breast area), pectoral muscle, a nipple, fibro-glandular region (parenchyma), and adipose region etc., is very crucial. It is the first preprocessing step in Computer Aided Diagnosis (CADx) of breast cancer. The different methods available for automatic extraction of pectoral muscle have been categorized as shown in the Fig. 2.

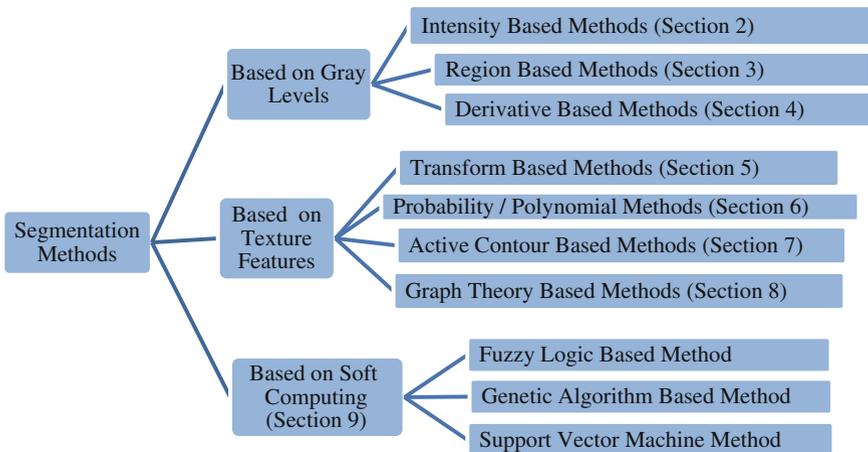


Fig. 2 Mammogram segmentation methods

Table 1 Performance measurement parameters

Performance parameter	Formula to calculate the parameter
Sensitivity	TP cases/total positive cases
Specificity	TN cases/total negatives cases
Precision	TP cases/(TP cases + FP cases)
Accuracy	(TP cases + TN cases)/total samples

The performance, indicating the degree of correctness of the segmentation results and their respective ground truth, is evaluated based on various parameters. It can be assessed subjectively by a expert radiologist by ranking the results or objectively by comparing the results with the ground truth using different metrics. The most widely used interpretation is the confusion matrix consisting of true positives (TP), true negatives (TN), false positive (FP), false negatives (FN). The performance of many methods is measured in terms of: specificity, sensitivity, precision, accuracy rates which are defined Table 1.

Similarly, the metrics used for error evaluation includes average error, Hausdorff distance, Absolute Error Distance etc. are also used in some techniques. Another metric used is ‘Receiver Operating Characteristic’ curve determined by true positives and false negatives results of a given experiment.

1.2 Significance

The accuracy of the automatic detection of breast cancer using CADe systems may be improved by separating region of interest in mammographic images. The presence of labels, noise, artifacts and majorly the pectoral muscle in the breast region may affect the performance and accuracy of the CADe system. Removing these parts out from the mammogram can increase computational complexity of the CADe systems. The presence of the pectoral muscle occupying a predominant region in the MLO view mammogram of breast, as shown in Fig. 1, may affect the results of cancer detecting process very badly. Pectoral muscle extraction is essential to provide effective results in the preprocessing step in CADe of breast cancer. An automatic pectoral muscle extraction plays a vital role in reducing the computational complexity and the errors of CADe systems. The further image analysis for breast cancer detection may become easier in the cancer detection process. The pectoral muscle extraction can also be useful in

- image registration for analyzing abnormality as like bilateral symmetry;
- automatic breast tissue density quantification;
- 3-D reconstructions from multiple mammographic views;
- mammogram-pair registration and comparison etc.

The following are the common preprocessing tasks [8] performed on the input mammographic image.

- Improving the quality of the input image by enhancing its contrast
- Finding out RoI by delineating the breast border in a simple, effective way
- Pectoral muscle is then extracted using a particular segmentation technique

1.3 Challenges

In most of the mammographic images, pectoral muscle detection still remains a challenging task. The major challenges of Pectoral Muscle Extraction [9] are due to its

- unclear and superimposed boundaries due to overlapping features;
- total absence in some cases;
- varying position, size, shape and texture from image to image;
- textural information similar to that of breast tissue, in most of cases;
- concave or convex border with its appearance varying in every other mammogram;
- border which cannot be modeled with any common geometrical or mathematical representation;

Thus, to devise a solution that extract a pectoral muscle accurately and efficiently over a wide variety of mammographic images, possibly from different databases [10, 11], is really a great challenge.

1.4 Motivation

The very solution to the pectoral muscle extraction problem lies in the domain of image segmentation. A variety of techniques are available to solve the basic segmentation problem. However to apply the technique commonly on a variety of images, one has to modify the fundamental segmentation algorithm or support it with some other supplementary methods. A few soft computing and other supporting techniques are also available to modify the basic segmentation algorithm so that the desired result of pectoral muscle extraction is achieved efficiently with sufficient accuracy over a different set of mammographic images.

1.5 Hypothesis

The different methods available for automatic extraction of pectoral muscle have been summarized under different categories in this chapter. The reason behind this detailed overview of automatic pectoral muscle extraction methods is to understand

the merits, demerits, limitations, problems and challenges encountered in each and every method while applying them over a different set of mammographic images. The performance of all the methods under similar category is enlisted for comparison purpose. This book chapter is intended to provide the researchers a systematic and comprehensive overview of different techniques of pectoral muscle extraction from digital mammograms.

1.6 Contributions

One of the intentions behind the work is to bring all the methods applied for pectoral muscle extraction, in a single chapter, so that the researchers get the consolidated information required and the further directions to devise a new simpler approach with even better accuracy, perhaps by combining some good concept proposed in some or the other algorithms enlisted here.

1.7 Organization of Chapter

The rest of this chapter is organized as follows. The Sect. 2 covers all the Intensity and histogram based methods. Region based approaches are discussed in Sect. 3. Section 4 describes all the gradient based approaches. Wavelet based approaches are presented in Sect. 5. Section 6 consists of the probability and polynomial based approaches. Section 7 includes active contour based approaches. Section 8 outlines graph theory based methods and Sect. 9 incorporates the soft computing methods.

2 Intensity Based Approaches

In the intensity based approaches, it is considered that the pectoral muscle area in the mammogram is dense and with high intensity compared to its surrounding tissues. These approaches try to find out change in the intensity levels of the pectoral muscle area and its adjacent parenchymal region. Rise and fall in the intensity levels all over the pectoral muscle plays a vital role in delineating the pectoral muscle border with better accuracy. Though, finding the exact pectoral muscle border in some cases is highly difficult; especially, with overlapping of surrounding tissues. From the literature surveyed, the different solutions based on either intensity, histogram, morphology or their combination with varying rates of success are discussed as given below.

A pectoral muscle extraction using histogram thresholding is proposed by Thangavel and Karnan [12] in a very efficient manner. The global optimum

threshold value is selected first and then the intensities less than this threshold are assigned with zero, whereas the remaining intensities are assigned with one. Morphological operators such as erosion and dilation are then applied for preserving details nearby the pectoral muscle region. This result is then converted to a binary image from which upper left region of white pixels represent a pectoral muscle region of the mammogram image. This proposed algorithm is very simple, easy to implement and yet with good performance. The experimental setup, results, image dataset used etc. are not discussed and the accuracy of the method is also not calculated in the paper.

An automatic method based on interesting properties of watershed transformation was explored by Camilus et al. [13]. In this approach, application of Watershed Transform on gradient images leads to a watershed line matching to the pectoral muscle border which in turn allows an efficient extraction of the pectoral edge. The problem of over-segmentation of the pectoral muscle region is resolved by applying merging algorithm which combines the suitable catchment basins to extract pectoral muscle with better accuracy. This method is validated by performing an experiment on 84 mammographic images from MIAS database which reveals a mean FP to be 0.85 % whereas mean FN is 4.88 %. The cases with FP and FN greater than 0.10 are almost zero, which indicates a good accuracy. The overall performance is claimed to be better than other techniques in this domain. The performance of this simple method is very accurate and efficient. The result is not validated with variety of images over multiple datasets.

A fully automatic breast region segmentation algorithm based on multilevel Otsu [14], gradient estimation and linear regression is presented by Kamila and Justyna [15]. After morphological preprocessing, a fast algorithm for multilevel thresholding classifies pixels in the multiple classes based on a number of gray levels. This separates the region of low intensity background from that of the breast skin-air interface in the image. Applying gradient on this image produces a rough pectoral muscle border which is smoothed by using a linear regression. This linear regression leads to finding the exact border of the pectoral muscle. The algorithm when tested on 300 MIAS database images showed an accuracy of 95–97 % which is quite high in comparison with existing methods. The efficiency of this algorithm measured in terms of total percentage error was found to be 98–99 %. The major success of this method lies in elimination of wrong detection. However, the method is not tested on variety of images from different datasets.

Liu et al. [16] proposed an accurate extraction of pectoral muscle border efficiently. The algorithm works on the basis of position related features of pectoral muscle in the breast area. The method makes repetitive use of the Otsu thresholding along with the morphology based operators to line out a rough edge of the pectoral muscle. This rough approximate edge then passes through a ‘multiple regression analysis’ to give out a refined pectoral muscle edge accurately. When tested on 150 MIAS database images, this algorithm gives almost the same results as that of the expert radiologists over a wide range of mammograms with varying appearances. It is also observed that the algorithm is effective even when the pectoral muscle edge

is obscured by overlapping breast tissue or other artifacts. The performance of this algorithm is validated over the different error metrics such as mean error (1.7188), misclassification error (0.0083), extraction error rate (0.00134), modified Hausdorff distance (0.08702) and average error is quite less. However, the repetitive use of thresholding makes this algorithm computationally intensive.

Duarte et al. [17] presented an automatic method, based on morphological filters, to estimate the pectoral muscle in mammograms. Morphological filters improve the image contrast between breast contour and background, also, between the pectoral muscle and the breast parenchyma. Original image gray level is first reduced from 256 to 9 in a heuristic way. By considering pectoral muscle as one of the densest region of the image, it is segmented by using the seventh and sixth gray-levels as thresholds, which are negated producing images N7 and N6, as shown in [17]. A morphological opening (disc-shaped SE with a 21-pixels diameter) is applied to N7, intending to exclude the smaller bright pixels that are out of the pectoral muscle region. An inferior reconstruction (disc-shaped SE with an 11-pixels diameter) is also applied but to the resulting image (mark), using N6 as its mask. Then, a morphological closing (again, a disc-shaped SE with an 11-pixels diameter) is applied to fill gaps in the reconstructed image contour. The gradient of the image obtained in the previous procedure is determined and a first-order polynomial is adjusted to estimate pectoral muscle. Then, it is tested if this estimated pectoral muscle comes into contact with the upper image edge or any of the lateral edges, as well as, if it does not cross the breast contour. If the above occur, then pectoral muscle is considered the densest region and hence adequately estimated. This method is evaluated by an experienced radiologist. The results of applying such methodology on 154 images (300 dpi, 8 bits) from the DDSM database show acceptable results with 93.6 % accuracy. This morphological operations based method is simple yet effective. However, the method is robust over different sets of images with varying appearances of pectoral muscle.

Burcin et al. [18] presented a novel segmentation algorithm for a pectoral muscle extraction based on Otsu's method in mammograms. The proposed system includes a pectoral muscle extraction on the basis of automatically selected threshold in an unsupervised manner. The process starts with preprocessing operations to remove the artifacts out of the breast border and to enhance the region of interest. A nonparametric, unsupervised extended version of Otsu's segmentation method with $N = 2$ is applied for segmenting the pectoral muscle. Connected component labeling algorithm is used for labeling the segmented regions. A upper of the two largest regions is selected as pectoral muscle. A limit area control mechanism proposed in this method with area value 21000 pixels for 512×512 mammographic images, allows to prevent false segmentation; especially for images with no pectoral muscle. The experimental results on 96 MIAS database images show 93 % accuracy. This method is simple, effective but its experimental results are not validated on different sets of images.

Performance evaluation of intensity based methods for pectoral muscle extraction is tabulated in Table 2.

Table 2 Performance evaluation of intensity based methods

Year authors	Main theme	# Images success%	Advantage/disadvantage
2005 Thangavel and Karnan [12]	Histogram-based thresholding	– –	<ul style="list-style-type: none"> • Easy, efficient • Accuracy not calculated
2011 Camilus et al. [13]	Watershed transformation	84 MIAS FP 0.85	<ul style="list-style-type: none"> • Accurate, efficient • Not robust
2012 Kamila and Justyna [15]	<i>Multilevel Otsu, gradient estimation, linear regression</i>	300 MIAS 95-97	<ul style="list-style-type: none"> • <i>No wrong detection</i> • <i>Not robust</i>
2012 Liu [16]	Iterative Otsu Thresholding and morphological processing	150 MIAS HD-0.087	<ul style="list-style-type: none"> • Accurate, robust, efficient • Computationally intensive
2012 Duarte et al. [17]	Morphological filters, gradient and first order polynomial	154 93.6	<ul style="list-style-type: none"> • Simple, accurate • Not robust
2014 Burcin et al. [18]	Nonparametric, unsupervised extended version of Otsu's method	96 MIAS 93	<ul style="list-style-type: none"> • ccurate, effective, no false segmentation • Not robust

As mentioned in Table 2, a multilevel Otsu's algorithm based method with gradient estimation and linear regression suggested by Kamila and Justyna [15] gives the best performance in this category. This is because of a fast multilevel Otsu's segmentation method that delineates the highlighted pectoral muscle in the preprocessing part and this is accurately marked by the linear regression. There are very rare cases in which computational complexity of the algorithms in terms of speed and time is considered. Further it is revealed that, in a few cases, the density of the pectoral muscle is high and is approximately same as that of the fibro glandular disc or small doubtful masses. Hence, most of the intensity based techniques are not able to discriminate the pectoral muscle from above mentioned dense parts of the breast. Consequently, the performance of all such techniques in these cases is very poor.

3 Region Based Methods

Region is a group of connected pixels with similar properties. Region based segmentation is a technique that allows to determine the regions directly in the given image. Mammographic images can be segmented using initial seed points until some condition or criterion based on distance etc. is satisfied. Region based methods are better than the edge based techniques (Sect. 4) in noisy images where edges are difficult to detect. These methods are simple, fast, leak through weak boundaries. From the literature surveyed, the different solutions provided on the pectoral muscle extraction with the help of region based segmentation techniques with varying rates of success are summarized below.

Raba et al. [19], illustrated an automatic pectoral muscle suppression method using a novel selective region growing algorithm. Initially selected seed point gives a rough approximation of the pectoral muscle region. This rough region is then refined with the help of morphological operations such as opening and closing. With this refinement, pectoral muscle border is highlighted clearly and hence extracted easily. This algorithm when tested on 320 images from MIAS database, showed around 98 % results as “near accurate” out of which 86 % are the good extractions of the pectoral muscle. Moreover, this technique is robust enough to give consistently good results over a wide variety of pectoral muscle appearances among all the mammograms. However, the method is weak in producing correct results when a tissue appears near the pectoral border [9]. The accuracy of the technique can further be improved by taking into account few more shape based and other related features.

Saltanat et al. [20], proposed a different method comprising pixel intensity levels values mapping in an exponential scale followed by a modified thresholding algorithm to line out the pectoral muscle area accurately in an efficient way. A region growing algorithm finds out an approximation of the pectoral muscle area and then verifies the same for exact match with that in the ground truth marked image. If it is not matching exactly, the rough region is adjusted to match with the desired pectoral muscle. This results into a mapped image with brighter regions which is enhanced further to divide it into regions with enhanced contrast. This is followed by specialized thresholding and region growing algorithm with lesser overflow of regions. The method is claimed to be robust over a large number of images with varying size, shape and positions of pectoral muscles appearances. When applied on 322 images of Mammogram Image Analysis Society (MIAS) database, the proposed algorithm gives 84 and 94 % accurate results when evaluated by two radiologists respectively.

A very good effect with simplicity is explored by Nagi et al. [21] through an automated technique for breast segmentation using a seeded region growing algorithm with morphological preprocessing. The process starts with removal of noise in the image using 2-D median filtering. Artifacts are suppressed and background is separated using thresholding and contrast enhancement. A seeded region growing algorithm is then applied to extract the pectoral muscle from the mammogram. A fully automated segmentation leading to accurate breast contour and the better computational performance over a wide range of mammograms with fatty, fatty-glandular and dense-glandular breasts are the two major contributions of the proposed algorithm claimed by the authors. The experimental setup includes two ground truth marked datasets, one is MIAS and the other is UMMC. The proposed method works well on a wide range of mammographic images with varying appearances pectoral muscles and shows good accuracy in pectoral muscle extraction. However, how the initial seed points are selected is not explained at all. The metric of accuracy and the quantified accuracy is not specified.

Nanayakkara et al. [22] proposed a method based on modified Fuzzy C-Means (mFCM) clustering algorithm. The process starts with preprocessing separating out

the region of interest and filtering out unwanted artifacts. A standard FCM is modified to avoid random initialization of cluster seeds and to show better pixel clustering in a speedy way using a block density approach. mFCM makes use of local information to estimate region modes robustly and to classify noisy pixels near the pectoral muscle border. The approximate pectoral muscle boundary obtained is then fitted by using a local maximum average gradient search. The contour obtained thus is smoothed using locally weighted least square fitting mechanism. Performance of the proposed method is tested by using 277 MIAS images with all types of breast tissues and pectoral muscle appearances. The experimental results indicate that the mean FP is 3.35, mean FN is 11.12 and mean Hausdorff distance is 14.83. The performance is also evaluated on some other error metrics and is quite acceptable. The performance of the proposed method is also compared with standard algorithms and it outperforms in terms of parameters such as percent overlap area (POA) and Hausdorff Distance (HD). The method works effectively even in case of pectoral muscle overlapping with breast parenchyma. The experiment is not validated on different sets of images; hence not robust.

The performance evaluation of the above mentioned region based methods for pectoral muscle extraction is presented in Table 3. The solution based on modified Fuzzy C-Means algorithm in [22] is the best among all the methods proposed so far in the region based methods group. This is because this proposed method works accurately well for all images having pectoral muscle overlapping with parenchymal region of breast. However, the computational complexity of the same method is not discussed. It is desired that the researchers should explore region based segmentation further and present a modified version which is simple yet effective on a wide variety of images.

Table 3 Performance evaluation of region based methods

Year authors	Main theme	# Images success %	Advantage/disadvantage
2005 Raba et al. [19]	Region growing and morphological operations	320 86	<ul style="list-style-type: none"> • Robust method, consistent results • Fails when muscle border is near breast tissue
2010 Saltanat et al. [20]	Specialized thresholding algorithm	322 84 and 94	Efficient, robust
2010 Nagi et al. [21]	Seeded region growing and morphological processing	MIAS UMMC	<ul style="list-style-type: none"> • Fully automatic, simple, works on images with varying breast densities • Initial seed points selection is not explained
2013 Klein [22]	<i>Modified fuzzy C-means</i>	277 MIAS FP 3.35 FN 11.12	<ul style="list-style-type: none"> • <i>Accurate, efficient, tested on different metrics</i> • <i>Fails where dense tissue and pectoral muscle overlaps</i>

4 Gradient Based Approaches

Pectoral muscle can be separated from the breast region by a straight or curved line between them. Hence, gradient based line detection methods are becoming a de facto standard for this purpose. A few researchers have proposed gradient based techniques using a straight line estimation to identify the pectoral muscle edge with quiet good accuracy. However, the actual pectoral muscle edge is not always straight; instead it is concave at some places and convex at other places. In the literature, few techniques refine the estimated straight line to fit the actual curved pectoral edge in the mammogram. Based on the available research literature, the different gradient based solutions for pectoral muscle extraction with varying rates of success are explained as given below.

Bezdek et al. [23] described a novel method for pectoral muscle edge and breast border detection effectively in four different stages. First, conditioning which is an important determinant of image quality is either by means of histogram equalization, spatial filtering or contrast enhancement to normalize the image intensities required for linear cumulative histogram. The second stage of feature extraction deals with visualizing more apparent edges, digital butte and canyon. This is achieved by means of Sobel and Prewitt masks followed by the geometric characteristics like range and standard deviation. These parameters lead to an exact separation of flat areas and the edge walls with flat top and steep sides as well as steep-walled valleys. These chosen features are then used in a blending function such as Minkowski norms, generalized logistic function or computational learning model to aggregate the information about the edges and to produce a wide range of edge images. The original “byte images” becomes “float images” after feature extraction and the same are reconverted to “byte images” using ‘dynamic scaling’ functions in the last stage. Once the extracted features match with the proposed blending function, it gives rise to an optimal edge image with full details. A pectoral muscle edge can be easily extracted from this detailed edge image. The overall performance of the algorithm seems to be acceptable for most of the images; however the result analysis with regard to sensitivity, specificity or any other parameter is not carried out in the work undertaken.

Chandrasekhar and Attikiouzel [24] addressed the segmentation of pectoral muscle by modifying the conventional edge detection paradigms to tunable parametric edge detection. The method makes use of four neighborhood based edge features, two directed digital gradients and two statistical descriptors. The pixels in a 3×3 window around a current pixel are “strung out” as a vector ω of dimensions 9, from top to bottom, left to right, in the original neighborhood. The authors have relaxed the constraint that the edge vector component should only be directed digital gradient. Instead, they allowed any combination of edge sensitive features defined as given below.

$$\varphi_1(\omega) = |(\omega_9 + 2\omega_8 + \omega_7) - (\omega_1 + 2\omega_2 + \omega_3)| \quad (1)$$

$$\varphi_v(\omega) = |(\omega_3 + 2\omega_6 + \omega_9) - (\omega_1 + 2\omega_4 + \omega_7)| \quad (2)$$

$$\varphi_r(\omega) = \max_{1 \leq i \leq 9} [\omega_i] - \min_{1 \leq i \leq 9} [\omega_i] \quad (3)$$

$$\varphi_s(\omega) = \sqrt{\left[\frac{1}{9} \sum_{i=1}^9 \omega_i^2 \right] - \left[\frac{1}{9} \sum_{i=1}^9 \omega_i \right]^2} \quad (4)$$

Here φ_h is horizontal Sobel digital gradient, φ_v is vertical Sobel digital gradient, φ_r is range and φ_s is the standard deviation. In order to ensure compatibility between the ranges of the different features, each is normalized so that the range of all four is $[0, 4]$. The algorithm also relaxed the constraint that the function that combines vector components to yield a real scalar magnitude must satisfy the properties of a norm, and instead allowed a generalized logistic function, $b_L(x)$ as a sigmoid blending function to yield a real scalar and defined as given below.

$$b_L(x) = \frac{1}{1 + \exp(-\lambda(x - \beta))} \quad (5)$$

where, λ and β are real positive constants. Thus a modification of the conventional edge detection paradigm gives rise to families of tunable parametric edge detectors, one of which has been used to extract the pectoral edge simply, controllably and reliably from mammograms. When tested on 12 MIAS images with $\lambda = 100$ and $\beta = 0.5$, it gives simple, controllable and reliable segmentation of the edge of the pectoral muscle for 10 images. However, the algorithm fails to yield a binary pectoral edge image alone.

Ferrari et al. [25] discussed a automatic technique of segmenting pectoral muscle edge by means of Hough Transform. The algorithm starts with binarization procedure that automatically identifies the pectoral muscle edge from the selected region of interest (ROI). The limited and bounded ROI minimizes the possibility of other linear structures biasing pectoral muscle edge representation. High frequency noise is then suppressed using Gaussian Filter. Hough transform of the Sobel gradient of ROI is then computed using

$$p = (x - x_o) \cos(\Theta) + (y - y_o) \sin(\Theta) \quad (6)$$

where (x_o, y_o) is the origin of the coordinate system of the image, p indicates the distance and Θ is the angle made by the pixel coordinates under analysis. This method is simple and efficient. However the detailed discussion on the experimental results is not covered.

A fully automatic method for segmenting the pectoral muscle consisting of the muscle edge estimation by straight line and cliff detection is presented by Kwok et al. [26]. The algorithm starts with an iterative thresholding that separates the

pectoral muscle from the parenchymal region. This is followed by a median filtering to remove unwanted noise. A gradient test then eliminates the problematic portions of separating straight line which is then fitted to minimize the least square error. In order to avoid the worst results, the straight line estimation is followed by validation test in an iterative manner till the line is fitted. To refine the muscle edge along this estimated straight line, cliff detection is used. This cliff detection consists of surface smoothing for removal of noise, rough texture etc. and edge detection to find a real shape of the muscle edge. Detecting the cliff is a dynamic process which is carried out until the best curved approximation is determined. Essentially, the intensity drops are identified and the intensity rises are ignored for the better results. This algorithm was tested on MIAS database images and approximately 94 % of images were acceptably segmented. However, this method is weak in detecting exact texture and the vertical pectoral borders especially.

Kwok et al. [27] presented a new adaptive automatic pectoral muscle extraction algorithm in which pectoral muscle edge is roughly identified by a straight line followed by its validation for its location and orientation. The algorithm uses the prior information about position and shape of the muscle edge to approximate the straight line estimation by means of iterative threshold selection to optimize the binarization. Enough care is taken to preserve the average luminance in the binary image. The result which is not always accurate and hence is corrected using cliff detection in a iterative manner to precisely find out the pectoral muscle edge. The algorithm is slightly modified from that of [26] and is designed to identify the intensity cliff nearby the pectoral border. The identified cliff locations are used to remove unwanted locations and to add intermediate values wherever necessary by using two point linear interpolations. This yields an image which is smoothed using average filter to produce a detected curve with some reduction in the sharpness. An iterative refinement then sharpens the edge that separates the pectoral muscle from the parenchymal region to a higher degree of accuracy. The algorithm when applied to MIAS database of 322 images, was found to be robust over a wide range of appearances of the pectoral muscles from all the images. Two expert mammographic radiologists evaluated that the proposed method gives an accuracy of 83.9 %.

Another interesting approach for pectoral muscle extraction is presented by Kwok et al. [28]. The algorithm starts with finding an approximation of the rough straight line along the pectoral muscle edge. Normal's to all the pixels along this rough line directed inwards are calculated to find out the curved portions of the pectoral border. The angles of these normal's vary between 180 to -180 . The value of difference between two consecutive normal's can be negative or zero indicating convex and otherwise concave. Thus overall extraction of the pectoral muscle is acceptably accurate. This method is simple and novel. The experiment performed on 322 MIAS images shows an accuracy of 79.5 %. The method is computationally intensive due to iterative nature.

A novel pectoral muscle edge detection method that overcomes a few drawbacks of the conventional techniques to give high precision results is proposed by Weidong et al. [29]. Firstly, a rough portion of the pectoral border consisting of

various texture features is separated by computing optimal threshold curve and local mean square deviation (MSD) curve. These curves help to find an appropriate threshold with respect to the distributed intensities over the mammographic image. A zonal Hough Transform, which is different than the conventional one, is applied to roughly fit the line along pectoral muscle border. This rough boundary is then refined by using a proposed elastic thread method to fit the actual muscle border which is slightly curved. When tested on 60 MLO view mammograms, the proposed method showed an accuracy of 96 % with a high acceptable precision.

Zhou et al. [30] designed and developed an automated algorithm to identify a pectoral muscle edge based on texture field orientation that utilizes a combination of prior information, local and global image features. The a priori knowledge on this muscle is its approximate direction and high intensities compared to its adjacent region. 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. This is used in this proposed texture-field orientation (TFO) method that utilized two gradient-based directional kernel (GDK) filters: one enhances the linear texture parts followed by extracting a texture orientation of the image on the basis of calculated gradient. This represents the dominant texture orientation at each pixel in the image which is then improved by a second GDK filter for extracting the ridge point. After validation of the extracted ridge points, a shortest-path finding method is applied to prepare the estimation of the probability of each ridge point lying on the actual pectoral border. Thus the ridge points with higher probability are connected to form the pectoral muscle edge. A data set of 130 MLO-view digitized film mammograms (DFMs) from 65 patients, data set of 637 MLO-view DFMs from 562 patients, and data set of 92 MLO view full field digital mammograms (FFDMs) from 92 patients etc. were tested to find out how much adaptive is TFO algorithm. The evaluation showed that 91.3 % of the tested images give out a correct pectoral muscle edge in a acceptable form. Also the technique works well proving its robustness over a wide range of variety of images.

A very simple yet accurate novel method for the detecting the pectoral muscle edge by making use of gradient and shape dependent characteristic traits is highlighted by Chakraborty et al. [31]. The algorithm starts with the pectoral muscle border estimation as a rough line by means of some characteristic traits of the pectoral muscle. This straight line passes through a refinement process to produce a pectoral muscle border more accurately. The method is applied on 200 mammograms (80-MIAS, 80 DR, and 40-CR images) and assessed based upon the false positive (FP), false negative (FN) pixel percentage which was 4.22, 3.93, 18.81 %, and 6.71, 6.28, 5.12 % for selected three databases, respectively. Whereas, mean distance closest point (MDCP) values for the same set of images are 3.34, 3.33, and 10.41 respectively. When compared with two similar techniques for identifying pectoral muscle developed by Ferrari et al. [26] and Kwok et al. [28], proposed technique results are found more accurate. The accuracy of the proposed algorithm still can be improved.

For the detection of pectoral muscle, Molinara et al. [32] presented a new approach based on a preprocessing step useful to normalize the image and highlight

the border separating the pectoral muscle from parenchymal region. This method is based on a preprocessing step that highlights the boundary separating the pectoral muscle from parenchymal region and on the evaluation of the gradient of the image along the x-axis direction. A subsequent step including edge detection and regression via RANSAC algorithm gives rise to a straight line separating pectoral muscle from the parenchymal region. The experiments performed on 55 images from DDSM database, showed that 89.1 % results are acceptable while 10.9 % un-accurate. One of the drawbacks is that this method includes repetitive processes and hence is computationally expensive and slow.

4.1 Proposed Method Using Morphological Operations and RANSAC

A slight modification, in the method suggested in [32], which is based on **RANDOM Sampling Consensus** (RANSAC) algorithm, in terms of the preprocessing for a good quality image followed by a computationally efficient RANSAC algorithm has reflected in acceptable results. In the proposed method, the unwanted noise and artifacts are removed using morphological operations. The upper and lower end points along the pectoral muscle in the top row and left column based on intensity variations is determined. The contrast of the image is then stretched following a binarization using Otsu's graythresh. A sobel operator then used to find out the estimation of edges near pectoral muscle border of the smoothed image. This estimation of the pectoral muscle edge is verified two to three times. The points in between upper and lower end points along approximate pectoral muscle edge are then recorded for RANSAC algorithm.

4.1.1 RANSAC Algorithm

The RANSAC algorithm divides given data into inliers and outliers and yields estimate computed from minimal set of inliers with maximum number of support points. The algorithm used is as given below.

1. Select minimal subset of data points in a random way required to fit a sample model
2. Points within some distance threshold t of model are a consensus set. Size of consensus set is model's support
3. Repeat for N such samples; model with maximum number of points is most robust fit
 - Points within distance t of best model are inliers
 - Fit final model to all inliers.

4.1.2 Experimental Results and Discussion

In order to test the performance of the RANSAC algorithm implemented, 40 images have been selected from mini MIAS database [10] consisting of 322 mammograms. Each of these MLO view mammographic images is having a size of 1024×1024 pixels with 8 bits per pixel. The spatial resolution of each pixel is 200 mm per pixel. Though the images are old and outdated, they are chosen for experimental purpose because they are publicly available. The snapshots of the output images are in Fig. 3. The results of pectoral edge segmentation are evaluated visually by the authors and show promising effects as enlisted in the Table 4.

From the experimental analysis, it is revealed that the algorithm works well in case of strong pectoral muscle borders which are nearly straight. In case of curved edges, the performance of the algorithm is poor and below average. The segmentation is even worse in a few cases. The segmentation in some cases fails due to overlapping of the pectoral muscle in the lower part of the breast tissue as the edge is not at all detectable. The time complexity of RANSAC algorithm is given on the basis of Eq. 7.

$$t = \frac{k}{1-a} (T_m + M_s * N) \quad (7)$$

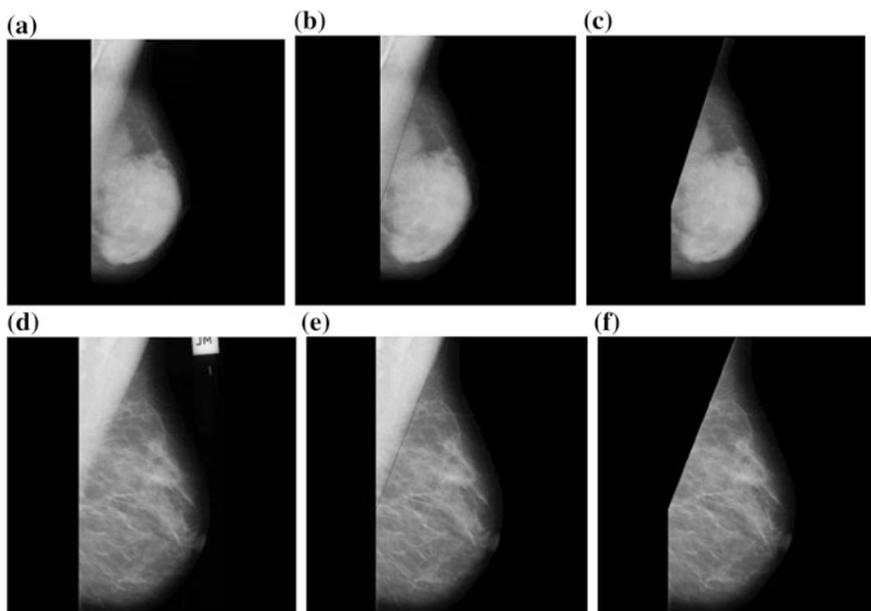


Fig. 3 Experimental results of RANSAC Algorithm. **a** Original image mdb038. **b** Line by RANSAC. **c** Extracted ROI of mdb-038. **d** Original image mdb046. **e** Line by RANSAC. **f** Extracted ROI of mdb-046

Table 4 Experimental results of RANSAC algorithm

Accurate	Acceptable	Unacceptable	Overall accuracy
22	10	8	82 %

Table 5 Computations required for RANSAC algorithm

	Time to compute single model— T_m	Number of samples	Number of models	Total computations
Addition	L	$M - 1$	k	$L(M - 1)(k)$
Subtractions	$L(L - 1)/2$			$L(L - 1)/2$
Multiplications	L			L
Divisions	0			0
Combinations	$L + L/2$	$M - 1$	k	

where, α is probability that a good model is rejected, k is number of samples drawn, N is number of data points, T_m is time to compute a single model, Ms average number of models per sample. k is a function of I, N and h, where I is the number of inliers and h is confidence in the solution. The computations required for implementation of RANSAC algorithm are shown in Table 5.

For N data points, there are $L = N(N - 1)/2$ possible estimated lines altogether as all the points are treated equally. The computational complexity of all the lines is $O(L)$ which is approximately $O(N^2)$, and the time required to select the suitable line fitting the pectoral muscle border is approximately $O(kMN^2)$. Thus, The contribution here lies in the RANSAC algorithm with limited number of iterations (k) and less number of samples (M) to select the best fit. However RANSAC algorithm fails sometimes at producing the correct model with the user-defined probability leading to an inaccurate model output.

Performance evaluation of gradient based methods for pectoral muscle extraction is tabulated in Table 6.

With the aid of Table 6 in Chap. “Electroanatomical Mapping Systems. An Epochal Change in Cardiac Electrophysiology”, the method based on Straight Line Estimation and Cliff Detection presented Kwok et al. [26] gives the best results in terms of 322 number of images. The reason behind the better accuracy of this robust method lies in a special cliff detection mechanism designed to refine the straight line estimate of the pectoral muscle border. This technique succeeds majorly due to its two components, surface smoothing and edge detection. The method presented by Weidong et al. [29] gives 96 % successful results but only on 60 images. No particular method based on gradient that works accurately for identifying the pectoral muscle on a wide range of images with varying positions of pectoral muscle. In majority of the techniques, the solution developed is tested over a specific set of images or a specific problem in given context. There are very rare cases in which computational complexity of the algorithms in terms of speed and time is considered.

Table 6 Performance evaluation of gradient based methods

Year authors	Main theme	# Images success%	Advantages/disadvantages
1998 Bezdek et al. [23]		– –	<ul style="list-style-type: none"> • Simple, effective • No result analysis
2000, Chandrasekhar and Attikiouzel [24]	Tunable parametric edge detection	12 MIAS 93	<ul style="list-style-type: none"> • Easy, efficient • Fails on curved muscles
2000 Ferrari et al. [25]	Hough transform	– –	<ul style="list-style-type: none"> • Simple, effective • Fails on curved muscles
2001 Kwok et al. [26]	<i>Straight line estimation and Cliff detection</i>	322 MIAS 94	<ul style="list-style-type: none"> • <i>Efficient, accurate</i> • <i>Difficult, complex</i>
2004 Kwok et al. [27]	Modified Cliff detection	322 MIAS 83.9	<ul style="list-style-type: none"> • Effective accurate • Computation intensive
2004 Kwok et al. [28]	Straight line, iterative curve method	322 MIAS 79.5	<ul style="list-style-type: none"> • Simple effective • Iterations adds into computational complexity
2007 Weidong et al. [29]	Polyline fitting and elastic thread approach	60 MIAS 96	<ul style="list-style-type: none"> • High precision, effective • Not robust
2010 Zhou et al. [30]	Texture-field orientation method	– 91.3	<ul style="list-style-type: none"> • Robust • Computation intensive
2012 Jayasree [31]	Average gradient and shape based feature	MIAS 4.22FP 6.7FN	<ul style="list-style-type: none"> • Efficient, simple, robust • Fails only on few very complicated images
2013 Molinara et al. [32]	Edge detection, regression via RANSAC	DDSM-55 89.1	<ul style="list-style-type: none"> • Efficient, avg. accuracy • Fails on images with curved pectoral muscles
2015 Proposed method	Morphological operations and RANSAC algorithm	MIAS 40 82	<ul style="list-style-type: none"> • Simple, efficient • Not robust

Hence, there is a tremendous scope to develop several new theories to solve this problem. What is required is a simple method that gives a perfect detection with maximum possible accuracy in terms of sensitivity and specificity in a robust way over a wide range of variety of images from different datasets available.

5 Transform Based Approaches

Texture features are useful and successful in the analyzing and interpreting mild textured mammograms. Texture features and intensity variations can be observed closely by decomposing the complex image to elementary form through wavelet analysis by means of Gabor wavelets, dyadic wavelets, Radon transform, Hough transform etc. These elementary components at different positions and scale helps radiologists to analyze the strong intensity level variations precisely. Wavelet transform analyzes various frequency components at various resolution scales and reveals the spatial as well as frequency components simultaneously from the image.

The original image can be perfectly constructed from these decomposed components. However, only a few researchers have exploited the power of wavelet based analysis to extract the pectoral muscle from the mammographic images. From the literature reviewed, the different ideas presented on transform based segmentation techniques with varying rates of success are described as given below.

Ferrari et al. [33] discussed a ‘Gabor wavelet’ based technique for automatic pectoral muscle edge identification. The algorithm starts with defining a region of interest (ROI) consisting of pectoral muscle in its entirety. This ROI image is then convolved with a specially designed bank of tunable ‘Gabor filters’ which encapsulate maximum information. This convolution enhances the appearances of all the ROI components in terms of their directional gradients, orientation and scale. The ‘Gabor filter’ designed in this method with scale parameter $S = 4$ and $K = 12$ orientations. This set of 48 parameters leads to 48 filters spanning the entire frequency spectrum. Angular bandwidth of each filter is 15° . Assuming the MLO view of optimally positioned breast, the Gabor filters are applied with 12 orientations and 4 scales. Vectors summation of K filtered images separates out phase $\phi(x, y)$ and magnitude $A(x, y)$ images at each pixel location (x, y) , which represent the edge flow vector. Series of nonzero vectors from the opposite directions become candidates for pectoral muscle edge. As optimally positioned MLO view expects the pectoral muscle located within 30° – 80° , the corresponding Gabor filter frequency responses can be oriented at 45° , 60° and 75° in the image domain. Disjoint boundary parts are connected by using iterative linear interpolation method. The longest line with maximum pixels is declared as a pectoral muscle border. This method delineates the pectoral muscle edge accurately with FP rate of 0.58 % and FN rate of 5.77 % from 84 mammographic images of mini MIAS database. Though this method gives accurate results, it is computationally more intensive.

Hough Transform and Radon transform are related to each other. Hough transform can make use of radon function for straight line detection. Hough transform is a special form of a radon transform. Linear features from images with high noise can be extracted using radon transform. Kinoshita et al. [34] presented a novel method for pectoral muscle extraction using radon transform. The preprocessing step includes application of Wiener filter to remove minor artifacts with high contrast and preserves the edge information at the same time. The algorithm proposed starts with finding and edge image using ‘Canny filter’. Radon transform is then applied on this edge image in an appropriate angular interval of 5° to 50° and -5° to -50° for right and left breast respectively. This leads to a number of straight line candidates representing pectoral muscle edge. The longest high gradient straight line candidate is then selected to delineate pectoral muscle edge separating the breast tissue. Localized radon transform used in this algorithm reduces the computational complexity and increases the speed. However, when tested on 540 mammograms, experimental results for 156 images are ‘accurate’ with FP < 5 %, acceptable for 220 images with FN < 15 % whereas 164 images are not accepted. Analysis of the experimented results shows that the algorithm works well for straight line edges while its performance with curved edges is not so good.

Mustra et al. [35] presented a hybrid method for extracting pectoral muscle edge. The algorithm starts with determining a reduced region of interest to understand the breast region orientation with its height multiple of 2^n , usually half of the height of image, and width based on skin-air interface of the breast on top line. Height and width chosen at power of 2 allows proper wavelet decomposition. In order to make edge detection an easy task, it reduces the original image to a 3-bit image. Dyadic wavelet of fourth level decomposes this image into approximate edge images. This approximate edge image undergoes interpolation on the basis of wavelets to prepare the image with same size and brightness as that of the original one. A blurring filters of size 40×40 is then applied for smoother edges. The image is thresholded for spreading the gray intensities evenly over the image. A Sobel filter then finds out the pectoral edge which is approximated then to a straight line separating pectoral muscle and breast tissue. When tested on 40 digital mammograms, the experimental results show ‘good and acceptable’ segmentation on 85 % images. Further analysis reveals that the algorithm works well when there is a high contrast between pectoral muscle and breast tissue. It fails when either the pectoral muscle is small or its contrast is low.

Mencattini et al. [36] presented a method for optimal pectoral muscle extraction using local active contour scheme and Gabor filters in a special combination. As described in [33], original image is initially decomposed using ‘Gabor filters’ and then the magnitude and phase of the image are then calculated. Vecteded summation of 48 ‘Gabor filters’ detect the candidate lines for the pectoral muscle profile, as per the process narrated in [33]. However the candidates selected may mislead increasing the False Negative rate of accuracy. Hence, this method eliminates the false pectoral edge candidates by using different logical conditions as described in [36]. These logical conditions allows to remove false candidate lines and the absent muscle problem is also addressed as well. The experimental results exhibit a very good accuracy up to 90 % on mini MIAS database images.

All the methods discussed above assume that the pectoral muscle can be fitted with a straight line. However, many a times, it is either concave, convex or both. Li et al. [37] presented a homogeneous texture and intensity deviation based method for pectoral muscle segmentation. This method diminishes the limitations of pectoral muscle extraction with a straight line. The process starts with a non-sub-sampled pyramid (NSP) which decomposes the original image into low-frequency and high-frequency sub-bands. The pectoral muscle is represented by means of likelihood maps in texture field calculated through clustering based segmentation and in intensity field calculated using neighbor Mean Square Deviation (MSD) matrix. By combining likelihood maps in a row, initial point on the border of the pectoral muscle is found out first and later other points are obtained by the same process in an iterative manner. The ragged edge obtained this way is further refined with the help of Kalman filter efficiently. The experimental results show an accuracy of 90.06 % on 322 MIAS database images and 92 % on images from DDSM database.

Performance evaluation of transform based methods for pectoral muscle extraction is enlisted in the Table 7. As mentioned in the Table 7, a method

Table 7 Performance evaluation of transform based methods

Year authors	Main theme	#Images success%	Advantages/disadvantages
2004 Ferrari et al. [33]	Tunable Gabor wavelet filters	84 MIAS FP 0.58, FN 5.77	<ul style="list-style-type: none"> • Accurate, efficient • Sensitivity not analyzed
2008 Kinoshita et al. [34]	Weiner filter, radon filters	540 69.62	<ul style="list-style-type: none"> • Fast, less complex • Fails at curved borders
2009 Mustra et al. [35]	Dyadic wavelet decomposition of 4th level	40 85 %	<ul style="list-style-type: none"> • Simple, efficient • Fails if pectoral muscle is small
2011 Mencattini et al. [36]	Gabor filters with logical conditions	90 MIAS 90	<ul style="list-style-type: none"> • Efficient, easy • Not so accurate
2013 Li et al. [37]	<i>Texture and intensity deviation, Kalman filter</i>	322 MIAS, DDSM 90.06, 92	<ul style="list-style-type: none"> • <i>Robust, efficient</i> • <i>Complex</i>

presented by Li [37] is the best with 92 % accuracy on 322 MIAS database images. The best results are possible because of the efficient Kalman filter applied on approximately correct rough estimation of pectoral muscle edge. There are very few methods that identify the pectoral muscle border accurately and efficiently, over different sets of mammograms. Assumption that the pectoral muscle edge can be fitted with straight line is not always true and limits the accuracy of the results. It is revealed that the research published in the domain of pectoral muscle separation based on transform is really low. Hence there is tremendous scope to develop several new theories to solve this problem. What is required is a simple method that gives a perfect detection with maximum possible accuracy in terms of sensitivity and specificity in a robust way over a wide range of variety of images from different datasets available.

6 Probability and Polynomial Based Approaches

The texture, appearance and density of the breast structures can be used to deduce the different statistical parameters for classifying the pixel intensities of digital mammograms. This approach is successfully used by a few researchers to statistically identify the ‘pectoral muscle edge’ in a effective way. From the literature surveyed, the different techniques presented on probability and polynomial based ‘pectoral muscle segmentation’ with varying rates of success are discussed as given below.

Sultana et al. [38] presented a new method with excellent tolerance to noise, for detecting a ‘pectoral muscle’ in ‘mammograms’ by making use of ‘Mean Shift Approach’. Assumption that a straight line can be fitted to a ‘pectoral muscle edge’ fails increasing ‘False positive rate’ which in turn decreases the segmentation

accuracy. This new method smashes out the drawbacks of straight line assumptions and obtains more accurate segmentation results. The process starts with removal of high frequency components in the image that may degrade the segmentation results. 'Region of Interest' consisting of 'pectoral muscle' is selected by using 'Histogram Equalization' followed by thresholding with low value. In the 'mean shift approach', firstly, 'probability density functions' (PDF) is used to estimate the initial points on the edge. To estimate this PDF, the proposed method uses a 'Gaussian kernel' which helps to find out the convergence in a few steps only and forms the cluster of pixels. Approximation of all possible paths in the direction of each point's gradient far from valleys and closer to PDF peak is performed. The process stops after assigning all the pixels a peak value. Thus a labeled map is obtained for each region. The mean value of each region in the map is calculated and the region with mean value bigger than $T = 150$ are registered as selected candidates for the 'pectoral muscle edge'. The selected region fulfilling the local contrast feature is then declared as a 'pectoral muscle edge'. The experimental results show an 84 % TP rate per image and 13 % FP rate per image. The very advantage of this new method is that it is a parameter-less clustering method which doesn't need any priori information about number of clusters and size of each cluster.

A statistical approach using the idea of 'Goodness of Fit' is discussed by Liu et al. [39] for detecting the 'pectoral muscle edge'. This method works on the basis of joint normal distribution applied to determine the probability of a pixel lying along a either high or low intensity region in the image. Based on this decision, a contour is finalized to remove pectoral muscle from breast tissue. The algorithm assumes the mammogram as a set of independent random intensity variables modeled as a normal distribution $N(\mu, \sigma^2)$ where μ is the mean and, σ^2 is the variance. This is $k \times k$ distribution of pixels sharing the same statistical features in the flat regions with strong features. An Anderson Darling (AD) test is applied on this set of pixels to perform a 'Goodness of Fit' test. This AD value is calculated as per the equation given in [39]. A smaller AD value indicates that the pixel belongs to a flat or slow changing (low frequency) component. A larger AD value represents a pixel from the high frequency component or related brighter region in the image. Thus AD value acts as image and edge enhancement measure which is insensitive to the amplitude of intensity variations in the image. Thus when this AD measure is applied on the mammograms, the pectoral muscle with brighter pixels along with its border full of stronger intensity variations is identified very easily. The experimental results on the randomly selected 100 images from MIAS database show that the proposed method gives 'accurate and acceptable' segmentation on 81 images while 'unacceptable' on 19 images. Thus the proposed method works more effectively on 'pectoral muscle extraction'.

Mustra and Grgic [40] discussed a pectoral muscle extraction method that combines conventional pectoral muscle edge identification with the polynomial approximation of curved muscle in six steps. First part includes finding the location where pectoral muscle is situated. This portion is usually 2/3 of the breast height and thus forms a region of interest. Second step is to enhance the contrast using

Contrast Limited Adaptive Histogram Equalization ('CLAHE') algorithm. Third, this is followed by a morphological opening with 3×3 structuring element which eliminates small objects and background noise while preserving the larger objects. Fourth step, a preliminary binary mask is created using previously calculated threshold. The rough pectoral muscle border achieved is then smoothed with the help of cubic polynomial fitting in an iterative manner. In fifth step, from the binary mask, 10 points are selected randomly for polynomial fitting of the muscle boundary. A cubic fitting function is chosen with 4 coefficients as shown in the equation:

$$y = p_1x_3 + p_2x_2 + p_3x + p_4 \quad (8)$$

where y is the horizontal coordinate and x is the vertical coordinate and p_i are the coefficients. In sixth step, a cubic polynomial function has been chosen because of the curved shape of pectoral muscle. An iterative linear fit function which finds correct slope is chosen to avoid wrong choice of points and is defined as

$$y = p_5x + p_6 \quad (9)$$

where y is the horizontal coordinate and x is the vertical coordinates. This proposed method when applied on MIAS database of 322 images showed 91.61 % successful results, 7.45 % acceptable results and 0.93 % unacceptable results.

Oliver et al. [41] presented a different pectoral muscle extraction technique using a supervised single strategy. The process starts by computing the probability density function, A_R for each pixel location (x) which is belonging to either background, pectoral muscle or breast. The method takes the advantage of the fact that usually background is dark, pectoral muscle is bright and breast region is in between bright and dark. There are exceptions as well. The intensity range, I_R , of these regions is determined based on histogram of each of these regions through a training over a set of images. Local binary patterns (LBP) is then used to characterize each pixel based on its texture probability, T_R . The likelihood of the pixel belonging to a particular region is then calculated by multiplying all three probabilities A_R , I_R and T_R . Finally all the pixels are assigned to the region with higher probability. This allows us to extract pectoral muscle easily. The experimental results on 149 MIAS images show a high degree of accuracy. The exact metric of the accuracy and its analysis is not discussed. The method is easy to implement and efficient. Performance evaluation of statistics and probability based methods for pectoral muscle extraction is tabulated in Table 8.

As observed in the Table 8, the method based on edge detection and polynomial estimation, presented by Mustra and Grgic [40] is the best among all methods in this class. The reason behind the success of this method lies in very good rough estimation and the best results with polynomial refinement over estimated pectoral border. It is very clear that the domain of 'Probability and statistics' is not fully explored but there is enough potential as like other application domains.

Table 8 Performance evaluation of statistics and probability based methods

Year authors	Main theme	# Images success%	Advantages/disadvantages
2010 Sultana et al. [38]	Mean-shift segmentation approach	TP 84, FP-13	<ul style="list-style-type: none"> • Accurate, efficient • Computation intensive
2011 Liu et al. [39]	Goodness of fit measure	100 MIAS 81	<ul style="list-style-type: none"> • Simple, effective • Not robust
2013 <i>Mustra and Grgic</i> [40]	<i>Edge detection, polynomial estimation</i>	322 MIAS 92–97	<ul style="list-style-type: none"> • <i>Highly accurate</i> • <i>Little-bit complex</i>
2014 Oliver et al. [41]	Position, intensity, and texture information	149 MIAS High	<ul style="list-style-type: none"> • Fast, reliable • No result analysis

7 Active Contour Based Approaches

‘Active contours’ which are also known as snakes are widely applied in medical image processing for detecting edge or curves, segmenting the image, shape modeling etc. Given the set of contours in the image, the snake tries to minimize the internal and external potential energies of all the possible surrounding neighbors of points along the contours. The internal deformation energy controls the capability of snake for stretching or blending of the contour. The external energy though the local minima attract the snake, The Gaussian smoothing filter defines the local minima which give gradient intensity edge that attracts the snake. Thus the classical snake has capability of extracting smooth edge accurately. However, it cannot deal with images with topological variations. Hence, number of improvements in the classical snake methods is suggested by the researchers in the literature. From the literature reviewed, the different studies presented on ‘active contour’ based segmentation techniques with varying rates of success are discussed as given below.

Wirth and Stapinski [42] suggested a slight modification to the classical ‘active contour method’ to segment the breast region and identify the ‘pectoral muscle edge’. All the initial contour points are identified by applying a dual threshold which is obtained using ‘Uni-modal Thresholding Algorithm’. The edges obtained this way are then enhanced using directed edge enhancing method. The enhanced edges are enlarged by removing noise after applying morphological erosion. A modified snake using a greedy algorithm calculated the energy for all the neighbors of all the pixels along continuity, curvature or gradient in the image. Thus lowest energy pixel is selected and again the energy levels in its neighborhood are calculated. At last the snake stops after defining a contour that represents the pectoral muscle edge. The algorithm when applied on 25 images from MIAS database shows acceptable results.

Ferrari et al. [43] discussed a novel method using adaptive contour model for extracting the breast and pectoral muscle boundary. The algorithm starts with contrast enhancement of the image by applying a logarithmic operation. This results in a significant improvement in the low density regions with fewer details near the

pectoral muscle border and breast border. This is followed by a low distortion binarization using Lloyd-Max algorithm. A morphological opening is then applied to reduce the minute unwanted objects and the noise. This demarks the pectoral muscle border approximately. An adaptive active deformable contour model is then applied on the image by adjusting the internal and external energy controlling parameters at each point. The proposed contour model minimizes the energy by means of a greedy algorithm developed by Williams and Shah (1992). The pectoral muscle segmentation results are evaluated based on the FP rate is 0.41 % and FN rate is 0.58 % for 84 mammographic images from MIAS database.

Though the ‘active contour models’ are useful in accurate extraction of pectoral muscle and other breast regions, the evolution of snake poses several limitations such as (i) sensitivity to initial contour position, quantity of internal parameters, weak edges, noise etc. (ii) an appropriate local minimum may be missed creating problem for convergence of points. (iii) Placing an initial contour closer to expected border (iv) lack of hard constraint regarding specific distance between two pixels.

The approaches discussed below try to eliminate the above mentioned limitations and suggest the modifications in the ‘active contour model’ to optimize the results.

Chaabani et al. [44] illustrated a method for identifying a pectoral muscle using Hough Transform and active contour. The algorithm starts with application of Canny edge detection followed by a Hough Transform in the angle interval between 135° to 165° . A line with the maximum number of pixels belonging to the contour is selected as pectoral muscle edge. This estimated line is further refined using the active contour model by virtue of energy minimizing spline. The algorithm when applied on DDSM database of mammograms showed that the success rate of pectoral muscle extraction was 92.5 % whereas there are 7.5 % images are unaccepted.

Wang et al. [45] presented a novel method for detecting pectoral muscle edge automatically with the help of ‘discrete time Markov Chain’ (DTMC) along with a ‘active contour’ method. Markov chain represents a portion of the object in a random discrete set of current pixel locations over time. The next pixel location is determined by using n-step transition probabilities. This is combined with two properties such as continuity and uncertainty belonging to pectoral muscle region for detecting the approximate border of the pectoral muscle. In the given algorithm, the rows and columns of the image are represented by time and state of the DTMC respectively. Thus DTMC algorithm obtains a rough edge of the pectoral muscle in an iterative manner. The detailed procedure for finding a rough pectoral border is explained in [45]. This rough border is further validated by replacing the false part with a straight line. This coarse pectoral muscle edge is refined by a snake algorithm with a slight modification. The internal energy parameter in the modified snake obtains a smooth pectoral muscle border whereas the external energy stretches the pectoral border as long as possible. The experiment performed on 200 images from DDSM database shows a ‘good’ segmentation on 75 % images and ‘acceptable’ segmentation on 91 % images. Accuracy of the detection can further be improved by developing a method searching the pectoral muscle border on the initial row itself.

The multiphase segmentation model proposed by Vese and Chan combines each phase using ‘level set functions’ for representing 2^n regions. At every stage of contour evolution, the ‘level set function’ is deviated aside from ‘signed distance function’ (SDF). Hence it requires costly re-initialization in each curve evolution.

In a topological analysis of medical images, isocontour mapping is very useful in retrieving meaningful information. Kim et al. [46] developed an algorithm focusing on intensity analysis of mammographic images and generates a adaptive contour map using a modified ‘active contour model’. In this approach, the complex mammographic images are analyzed to extract topographic features from rough to fine scale and are represented in an isocontour map. This isocontour map image causes the reduction in analysis complexity. The algorithm presented here starts with applying a denoising method for reducing interference noise from the image. This image then undergoes two-phase segmentation and two sub-regions are created. This partitioning is achieved by using the Mumford Shah energy functional recursively. A multipass ‘active contour’ that is based on ‘active contour without edges’ (AWCE) proposed by Chan and Vese is used to extract local regional information. In an image with weak edges and intense noise, AWCE model partitions the regions based on energies. The algorithm again partitions one sub-region iteratively by using level set evolution without re-initialization (LSEWR) by minimizing a new energy model. This LSEWR introduces an internal energy term which doesn’t allow the ‘level set function’ to deviate from a ‘signed distance function’ (SDF) in every contour evolution. This segmentation of sub-regions results into a tree-like structure of all the sub partitions forming a map of adaptive contours. This map is then finalized after skipping the isocontours with same energy. Thus the algorithm works very well on mammographic images with weak and blurred edge effectively and also reduces the isocontour maps quantity from 206 to 11.

Looking into the several limitations posed by snakes based methods, Akram et al. [47] proposed a preprocessing algorithm to remove a pectoral muscle edge along with the other unwanted artifacts from the mammograms. This algorithm makes use of a modified ‘active contour method’ proposed by Chan and Vese which is based on the Mumford Shah model. The algorithm in its first part, converts a given image into a binary image using a threshold $T = 15$, and then removes the low and high intensity labels along with scanning artifacts by computing a row and column wise pixel summation method. In its second stage, the pectoral muscle border is traced by using multiphase ‘active contour method’ which is based on Mumford Shah model. The algorithm introduces a new term M^k which allows moving the contour inwards and also computes its stopping point based on the difference between consecutive contours. Thus the contour of the pectoral muscle and other breast regions is derived. In the third part, the pectoral muscle is extracted out using M^k value. The algorithm when tested on few images from mini MIAS database, shows an accuracy of 77.10 % on images with bad preprocessing results while it is 97.84 % on images with accurate preprocessing results. Thus, the accuracy of the technique discussed herein is highly dependent on the preprocessing results and the value of stopping point in the contour model.

Table 9 Performance evaluation of active contour based methods

Year authors	Main theme	# Images success%	Advantages/disadvantages
2003 Wirth and Stapinski [42]	'Unimodal Thresholding Algorithm' and modified contour model	25 MIAS • Good	• Simple, efficient • No result analysis
2004 Ferrari et al. [43]	Adaptive contour model	84 MIAS FP 0.41, FN-0.58	• Efficient, accurate • Not robust
2010 Chaabani et al. [44]	<i>Hough transform and active contour</i>	<i>DDSM</i> 92.5	• <i>Accurate, effective</i> • <i>Not robust</i>
2010 Wang et al. [45]	Time Markov chain and active contour model	200 DDSM 75, 91	• Simple, efficient • Not robust
2013 Kim et al. [46]	Active contour without edges, level set evolution without re-initialization	– –	• Simple • Not robust
2013 Akram et al. [47]	Multiphase active contour method	MIAS 77.10, 97.84	• Accurate with good preprocessing • Not robust

Performance evaluation of active contour based methods for pectoral muscle extraction is tabulated in Table 9. As mentioned in Table 9, the method based on Hough Transform and active contour by Ali Cherif [44], gives the best results among all. The best results in this method are possible due to effective refinement work by the active contour model suggested. As such, there is no particular method that works satisfactorily with better accuracy for the problem of identifying the pectoral muscle, uniformly over a wide variety of mammograms. In majority of the methods, the solution developed is tested over a set of limited images or a specific database images only. There are very rare cases in which computational complexity of the algorithms is considered. Though the researchers are trying their level best to find out an accurate solution, it is revealed from the literature reviewed that the research published in the domain of pectoral muscle separation based on active contour methods is really low. And hence there is tremendous scope to develop several new theories to solve this problem.

8 Graph Based Methods

Image segmentation based on graph theory based methods though computationally intensive can be applied for pectoral muscle edge detection to obtain the expected results. Recently, the appropriate selection of local and global information features along with simplified efficient techniques such as Minimum Spanning Trees and Shortest path have come up with promising results. Based on the research work

studied from the available literature, the different solutions presented on the basis of graph theory for pectoral muscle border identification with varying rates of success are discussed as given below.

Ma et al. [48] presented two methods, one on the basis of adaptive pyramids (AP) and other on minimum spanning tree (MST), for pectoral muscle identification in digital mammograms. The first method implemented in this paper is based on the algorithms suggested by Jolion and Montanvert for building a pyramid graph of vertices (pixels) in the given image. The ‘interest operator’ and ‘two state variables’ allow choosing the surviving vertices while exploiting different characteristics of the image. The two state processes for selecting these two state variables is explained in [48]. Thus a graph pyramid consisting of significant components of the image with non surviving vertex as a root is constructed. The reduction in the level of pyramid is dependent on the changing image information and hence the pyramid is adaptive. The second method based on MST constructs a graph of edges (connecting pixels as vertices) with weights defined by a function based on intensity differences. The algorithm proceeds forming a segment of pixels with minimum internal variation and merging two segments with less internal variations. The implementation of MST based algorithm is computationally intensive. None of these methods give accurate pectoral muscle segmentation; any one can be chosen for further smoothing of the results. An active contour is used to bring the rugged pectoral muscle edge closer to the real one. The internal and external energies represented in [48] produce smoothing and highlighting effects on the pectoral muscle border. The implementation of the methods with the selected 84 mammographic images from mini MIAS database shows moderately acceptable results. The performance of the methods based on the error measure of average distance between actual and computed border is less than 2 mm for 80 % and it is less than 5 mm for 97 % of the selected 84 images. Being a first attempt to identify the pectoral muscle using graph theory based methods; the results are encouraging and open a wide scope for further experiments with different local and global characteristics features of the image.

Camilus et al. [49] proposed a graph cut based method to automatically identify the pectoral muscle edge in digital mammograms in an efficient way. The algorithm starts with careful observation of anatomical features and cropping of the mammogram to a region of interest which completely includes pectoral muscle and thus eliminates the unwanted components while reducing the time complexity. The proposed method achieves the segmentation in three major steps. The first step formulates the weighted graph of edges formed by joining the image pixels as vertices. The dissimilarity in the pixels (usually intensities or Euclidean distance) determines the weight on the edges which are then sorted in non decreasing order of weights. The second step of the algorithm sorts the edges based on their weights and homogeneity of edges. Here the ROI gets divided into different segments based on intra region and inter region dissimilarity factors. The mean of all the edges known as intra-region edge average (IRA) calculated with formula specified in

Eq. (1) in [49] represents the homogeneity of the probable image segment. Similarly, inter region edge mean (IRM), as defined in Eq. (6) of [49], allows merging two closely resembling regions. Selection of proper values of parameters δ_1 and δ_2 for dynamic threshold ultimately leads to a coarse region identified at the top left corner of the ROI. The third step includes the application of Bezier curves to rough pectoral muscle edge. The experiment performed on randomly selected 84 images from MIAS database with ground truth marked by the expert radiologists gives consistent accuracy in terms of FP as 0.64 % and FN as 5.58 %. In most of the tested images, the error rate is very less; especially FP and FN either of which may be less than 0.05 but not both at a time. Thus the results are quite superior to earlier method [48]. The proposed method even works well in case of pectoral muscle border near to the dense tissues and also in case of very small pectoral muscle. However, the results of the method can be improved further by incorporating a few more low level features along with high level features and some more anatomical constraints.

Cardoso et al. [50] presented an algorithm based on a shortest path on a graph to detect the pectoral muscle border automatically. The algorithm assumes that the pectoral muscle, if present, is the change in the intensity levels of the image pixels which ranges from top margin of the image to the left margin. Assuming the origin at the top left corner, the left columns are mapped to the bottom rows due to which the pixels along the pectoral muscle border remains in vertical direction along top to bottom rows with one and only one pixel along each row. A weighted graph of the image is then constructed to find out the optimal vertical path using the cumulative minimum cost C for each pixel using the formula given in [50]. The weight on each edge in the graph is computed with a formula given in [50]. Once the shortest path is constructed, the pectoral muscle edge is finalized. The rows are then transformed back to the Cartesian coordinate system. The contour validation rule is applied to verify if there is no pectoral muscle present in the image. The experiment performed on a set of 50 DDSM images and 100 images from HSJ Portugal, with ground truth marked by expert radiologists, shows the Hausdorff distance of 0.1387 and 0.1426 whereas Mean distance of 0.0545 and 0.0387 respectively. These results are quite good among all the graph based methods for the same task. However, this method may give wrong results in case of multiple strong pectoral muscle borders present in the image.

Performance Evaluation of graph theory based methods for pectoral muscle extraction is enlisted in the Table 10. As seen in the Table 10, the method based on shortest path and support vector machine approach, by Cardoso et al. [50], is the best among all. The better result is possible because of the accurately constructed weighted graph using cumulative minimum cost measure. Further, it is revealed that the crucial tasks in all the graph based methods include constructing the graph, sorting the edges and determining the edge weights in the given image. The different parameters selected to provide either local or global image information plays a vital role in the overall algorithm. The results of some of the recent methods have proved to be really promising but still there is a lot expectation from the accuracy point of view. Hopefully, the researchers will be able to exploit the real power of

Table 10 Performance evaluation of graph theory based methods

Year authors	Main theme	# Images success%	Advantages/disadvantages
2007 Ma et al. [48]	Adaptive pyramids (AP), minimum spanning tree (MST), active contours	84 MIAS 2 mm, Error 80 5 mm, Error 97	<ul style="list-style-type: none"> • Accurate, effective • Computation intensive
2010 Camilus et al. [49]	Graph cut based method, Bezier curve	84 MIAS FP 0.64 FN5.58	<ul style="list-style-type: none"> • Efficient • Complex
2011 Cardoso et al. [50]	<i>Shortest path method</i>	50 DDSM, 100 FSJ HD 0.1387, 0.1426 MD 0.0545 0.0387	<ul style="list-style-type: none"> • Accurate, effective • Computation intensive

graph theory with some other concepts leading to a accurate solution for the pectoral muscle identification efficiently.

9 Soft Computing Methods

Soft computing is a new emerging trend of obtaining precise solution for complicated cases of the problems. The elements of soft computing includes fuzzy logic, genetic algorithm, neural computing and evolutionary computation. Soft computing techniques can be used for wide range of applications including image segmentation. A few important soft computing based methods for pectoral muscle extraction are explored briefly below.

Karnan and Thangavel [51] presented a two step approach to detect a breast border separating a pectoral muscle indirectly using Genetic Algorithm. The breast border identification process in the proposed work starts with binarization of the given mammographic image using local minima of the histogram as the threshold value. The connected components in the binary image are then smashed out using morphological operations. This results into a binary image showing a breast border. Pixels on this border with a neighborhood window of size 3×3 form a binary kernel which represents the population string in the proposed genetic algorithm. Population strings along fitness values which are sum of intensities along border, generates new population using the genetic ‘reproduction’ for crossover. The crossover operator then allows exchanging of bits in the 2×2 window of reproduced kernels. This is followed by a 2 dimensional mutation operation in which a transformation is performed if the kernel matches any one of the 14 windows shown in [52]. The kernels in final population represent the enhanced border points on the breast border which indirectly separates a pectoral muscle in the left top corner of

mammogram. The performance of the algorithm analyzed on 114 images with malignancy from MIAS database shows the accuracy of 90.60 % for detection. Further analysis plotting True Positives versus False Positives shows True Positive Fraction as 0.71 and 0.938 whereas False Positive Fraction 0.2890 and 0.0614 with threshold 50 and 150 respectively.

Domingues et al. [53], proposed a fully automatic yet simpler method to detect the pectoral muscle border using a shortest path and support vector machine approach. The method first finds out the region of interest by removing unwanted labels and artifacts in the background by using an adaptive thresholding approach. The image is then cropped to reduce the area of the breast and the computational complexity subsequently. The two endpoints on the pectoral muscle are detected based on two support vector regression (SVR) models. The end point on the pectoral muscle on the top row is detected using a SVR model which is based on the input features obtained from a 32×32 thumbnail from the upper half of the cropped image. The other end point on the pectoral muscle on the left column is detected using a SVR model which is based on the input features obtained from a 32×32 thumbnail from the lower half of the cropped image. The pectoral muscle border is along the shortest path through edges represented in a graph, in between these two end points. A weighted graph with pixels as nodes and edges connecting neighboring pixels with its magnitude as weight, is searched for a shortest path which demarks the pectoral muscle. When tested, this algorithm shows the Hausdorff distance of 0.0860 and 0.1232 whereas Mean distance of 0.1232 and 0.0340 on 50 images from DDSM database and HSI database respectively. Though the accuracy of the proposed algorithm is low, its simplicity is really very acceptable by different manufacturers for devising a solution.

Aroquiaraj et al. [54], proposed a novel pectoral muscle extraction method which is merely a combination of straight line techniques, Connected Component Labeling algorithm (CCL) and, Fuzzy Logic. The method is validated on 322 images from the Mammographic Image Analysis Society (MIAS) database. The evaluation was done using various parameters such as Mean Absolute Error (MAE), Hausdorff Distance (HD), Probabilistic Rand Index (PRI), Local Consistency Error (LCE) and Tanimoto Coefficient (TC). The combination of fuzzy with straight line algorithm gives more than 95.5 % accuracy which is quite high and acceptable.

Sapate and Talbar [55] discussed a modified 'K-means clustering' [56] for eliminating a pectoral muscle from the breast tissue leading to a substantial accuracy. The algorithm starts with applying a combination of image filters and morphological operations for removing noise, scanning artifacts, low and high intensity labels from the mammographic images along with accentuating some specific features. A modified K-means algorithm presented in this method attempts to improve the original algorithm in both of its major phases i.e. computing cluster centers and assigning pixels to appropriate clusters with $K = 4$. The automatic selection of initial cluster centers improves the accuracy of segmentation in the proposed method. The experimental results show that the accuracy and the computational complexity, both, are improved over the original algorithm.

Table 11 Performance evaluation of soft computing based methods

Year authors	Main theme	# Images success%	Advantages/disadvantages
2007 Karnan [51]	Genetic algorithm	114 MIAS 90.60	<ul style="list-style-type: none"> • Accurate, efficient • Not robust
2010 Domingues et al. [53]	Shortest path and support vector machine approach	50 DDSM HD.0860	<ul style="list-style-type: none"> • Simpler, efficient • Less accurate, not robust
2014 Laurence [54]	<i>Connected component labeling, fuzzy logic, straight line estimation</i>	MIAS322 95.5	<ul style="list-style-type: none"> • Accurate, efficient • Complex, computation intensive
2015 Sapate [55]	Modified K means clustering method	150 MIAS 86	<ul style="list-style-type: none"> • Simple, efficient • Not robust

Experimental results on 130 images from MIAS database show the accuracy of pectoral muscle extraction is 86 %. The method is not robust as its results are not validated with different datasets of mammograms.

Performance evaluation of soft computing methods for pectoral muscle extraction is tabulated in Table 11. With aid of the Table 11, the method by Aroquiaraj et al. [54] combining connected component labeling, fuzzy logic and straight line estimation approaches is the best among all. The reason behind these best results is that the fuzziness of the gray scale mammograms is correctly modeled by this fuzzy based approach. The soft computing based approaches give better performance over the existing traditional techniques for the pectoral muscle extraction. However, very few of the soft computing approaches are explored for extracting the pectoral muscle. Therefore, there is tremendous scope for exploring further the potential of soft computing based other approaches to improve the accuracy of the pectoral muscle extraction problem.

10 Conclusions

The overview of the different techniques covered in this chapter focuses on the efforts made in the direction of solving the pectoral muscle extraction problem in the preprocessing part of the CADe systems for detecting breast cancer in its early stage using digital mammograms. The discussion about all the different methods proposed by researchers in literature reveals that there exists very few methods which give more accurate results on a wide range of images with varying position, shape and size of the pectoral muscle in the mammographic image of the breast. On the other hand, there are very rare cases where the computational complexity of the proposed algorithm has been calculated with a due importance. The performance and accuracy of techniques enlisted may be useful for comparison purpose. Hopefully, this study will be useful for the researchers to find out a better scope to devise a robust yet simple pectoral muscle extraction algorithm with better accuracy

over a wide range of mammograms with varying positions, shapes and intensities of the pectoral muscle regions.

Acknowledgments The authors are thankful to Dr. L.M. Waghmare, Director, SGGS IE&T, Nanded and Dr. M.B. Kokare, Coordinator, Center of Excellence in Signal and Image Processing, for their constant encouragement, great support and 24×7 open access to state of the art laboratory facilities. Authors are thankful to Dr. Ravindra C. Thool for his constant encouragement. Authors are really very grateful to the referees for their valuable suggestions and comments.

References

1. The American College of Radiology.: <http://www.mammographysaveslives.org/Facts>
2. Amin, I.I., Hassaniien, A.E., Kassim, S.K., Hefny, H.A.: Big DNA Methylation data analysis and visualizing in a common form of breast cancer. In: Hassaniien, A.E. et al. (eds.) *Big Data in Complex Systems, Studies in Big Data*, vol. 9. Springer International Publishing Switzerland (2015). doi:10.1007/978-3-319-11056-1_13
3. Hassaniien, A.E., Moftah, H.M., Azar, A.T., Shoman, M.: MRI breast cancer diagnosis hybrid approach using adaptive ant-based segmentation and multilayer perceptron neural networks classifier. *Appl. Soft Comput.* **14**, 62–71 (2014)
4. Moftah, H.M., Azar, A.T., Al-Shammari, E.T., Ghali, N.I., Hassaniien, A.E., Shoman, M.: Adaptive K-means clustering algorithm for MR breast image segmentation. *Neural Comput. Appl.* **24**(7–8), 1917–1928 (2014)
5. American Cancer Society.: <http://www.cancer.org/treatment/understandingyourdiagnosis>
6. BreastCancer.Org.: <http://www.breastcancer.org/symptoms/testing/types/mammograms>
7. Radiopaedia.org.: <http://radiopaedia.org/articles/mediolateral-oblique-view>
8. Bhateja, V., Misra, M., Urooj, S., Lay-Ekuakille, A.: A robust polynomial filtering framework for mammographic image enhancement from biomedical sensors. *IEEE Sens. J.* **13**(11), 4147–4156 (2013)
9. Karthikeyan, G., Rajendra, A., Kuang, C.C., Lim, C.M., K Thomas, A.: Pectoral muscle segmentation: a review. *Elsevier J. Comput. Methods Progr. Biomed.* 48–57 (2013)
10. Suckling, J.: The Mammographic image analysis society digital mammogram database. In: *Exerpta Medica. International Congress Series*, vol. 1069, pp. 375–378 (1994)
11. Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, W.P.: the digital database for screening mammography. In: *Proceedings of the 5th International Workshop on Digital Mammography*, pp. 212–218. Med. Physics Publishing (2001)
12. Thangavel, K., Karnan, M.: Computer aided diagnosis in digital mammograms: detection of micro-calcifications by meta heuristic algorithms. *GVIP J.* **5**(7) (2005)
13. Camilus, K., Govindan, V., Sathidevi, P.: Pectoral muscle identification in mammograms. *J. Appl. Clin. Med. Phys. North America* **12**(3), 215–230 (2011)
14. Otsu, N.: A threshold selection method from gray-level histogram. *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979)
15. Kamila, C., Justyna, W.: Automatic breast-line and pectoral muscle segmentation. *Schedae Inf.* **20**, 195–209 (2012)
16. Liu, C.C., Tsai, C.Y., Liu, J., Yub, C.Y., Yub, S.S.: A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis. *Elsevier J. Comput. Math. Appl.*, 1100–1107 (2012)
17. Duarte, M.A., Alvarenga, A.V., Azevedo, C.M., Infantosi, A.F.C., Pereira, W.C.A.: Estimating the pectoral muscle and the nipple positions in mammographies using morphological filters. In: *XXIII Congresso Brasileiro em Engenharia Biomédica* (2012)

18. Burcin, K., Nabiyevb, V.V., Turhanc, K.: A novel automatic suspicious mass regions identification using Havrda & Charvat entropy and Otsu's N thresholding. *Comput. Methods Progr. Biomed.*, 349–360 (2014)
19. Raba, D., Oliver, A., Joan, M., Marta, P., Joan, E.: Breast segmentation with pectoral muscle suppression on digital mammograms. *Lecture Notes in Computer Science*, pp. 153–158 (2005)
20. Saltanat, N., Hossain, M.A., Alam, M.S.: An efficient pixel value based mapping scheme to delineate pectoral muscle from mammograms. In: *IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, pp. 1510–1517 (2010)
21. Nagi, J., Kareem, S.A., Nagi, F., Ahmed, S.K.: Automated breast profile segmentation for ROI detection using digital mammograms. In: *IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 87–92 (2010)
22. Nanayakkara, R.R., Yapa, Y.P.R.D., Hevawithana, P.B., Wijekoon, P.: Automatic breast boundary segmentation of mammograms. *Int J. Soft Comput. Eng. (IJSCE)*. **5**(1) (2015)
23. Bezdek, J.C., Chandrasekhar, R., Attikiouzel, Y.: A Geometric approach to edge detection. *IEEE Trans. Fuzzy Syst.* **6**(1), 52–75 (1998)
24. Chandrasekhar, R., Attikiouzel, Y.: Segmentation of the pectoral muscle edge on mammograms by tunable parametric edge detection. In: *Australian Research Centre for Medical Engineering (ARCM E)* (2000)
25. Ferrari, R.J., Rangayyan, R.M., Desautels, J.E.L., Frère, A.F.: Segmentation of mammograms: identification of the skin–air boundary, pectoral muscle, and fibro-glandular disc. In: *Proceedings of 5th International Workshop on Digital Mammography*, Toronto, ON, Canada. pp. 573–579 (2000)
26. Kwok, S.M., Chandrasekhar, R. and Attikiouzel, Y.: Automatic pectoral muscle segmentation on mammograms by straight line estimation and Cliff detection. In: *Seventh Australian and New Zealand Intelligent Information Systems Conference*, Perth, Western Australia, pp. 67–72 (2001)
27. Kwok, S.M., Chandrasekhar, R., Attikiouzel, Y., Rickard, M.T.: Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. *IEEE Trans. Med. Imaging* **23**(9), 1129–1139 (2004)
28. Kwok, S.M., Chandrasekhar, R.A., Attikiouzel, Y.: Automatic assessment of mammographic positioning on the mediolateral oblique view. In: *International Conference on Image Processing ICIP '04*, vol. 1, pp. 151–154 (2004)
29. Weidong, X., Lihua, L., Wei, L.A.: Novel pectoral muscle segmentation algorithm based on polyline fitting and elastic thread approaching. In: *The 1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 837–840 (2007)
30. Zhou, C., Wei, J., Chan, H.P., Paramagul, C., Hadjiiski, L.M., Sahiner, B.: Computerized image analysis: texture-field orientation method for pectoral muscle identification on MLO-view mammograms. *Am. Assoc. Med. Phys.*, 2289–2299 (2010)
31. Chakraborty, J., Mukhopadhyay, S., Singla, V., Khandelwal, N., Bhattacharyya, P.: Automatic detection of pectoral muscle using average gradient and shape based feature. *J. Digit. Imaging* **25**(3), 387–399 (2012)
32. Molinara, M., Marrocco, C., Tortorella, F.: Automatic Segmentation of the pectoral muscle in mediolateral oblique mammograms. In: *IEEE Conference on Computer-Based Medical Systems*, pp. 506–509 (2013)
33. Ferrari, R.J., Rangayyan, R.M., Desautels, J.E.L., Borges, R.A., Frere, A.F.: Automatic identification of the pectoral muscle in mammograms. *IEEE Trans. Med. Imaging* **23**(2), 232–245 (2004)
34. Kinoshita, S.K., Azevedo-Marques, P.M., Pereira Jr, R.R., Rodrigues, J.A.H., Rangayyan, R. M.: Radon-domain detection of the nipple and the pectoral muscle in mammograms. *J. Digit. Imaging* **21**(1), 37–49 (2008)
35. Mustra, M., Bozek, J., Grgic, M.: Breast border extraction and pectoral muscle detection using wavelet decomposition. In: *Proceedings of IEEE Conference EUROCON'09*, pp. 1426–1433 (2009)

36. Mencattini, A., Salmeri, M., Casti, P., Pepe, M.L.: Local active contour models and Gabor wavelets for an optimal breast region segmentation. *Int. J. Comput. Assist. Radiol. Surg.* (2012)
37. Li, Y., Chen, H., Yang, Y., Yang, N.: Pectoral muscle segmentation in mammograms based on homogenous texture and intensity deviation. *Elsevier J. Pattern Recogn.* **46**(3), 681–691 (2013)
38. Sultana, A., Ciuc, M., Strungaru, R.: Detection of pectoral muscle in mammograms using a mean-shift segmentation approach. In: *IEEE 8th International Conference on Communications*, pp. 165–168 (2010)
39. Liu, L., Wang, J., Wang, T.: Breast and pectoral muscle contours detection based on goodness of fit measure. In: *IEEE 5th International Conference on Bioinformatics and Biomedical Engineering*, pp. 1–4 (2011)
40. Mustra, M., Grgic, M.: Robust automatic breast and pectoral muscle segmentation from scanned mammograms. *Elsevier J. Signal Process.* **93**(10), 2817–2827 (2013)
41. Oliver, A., Llado, X., Torrent, A., Marti, J.: One-shot segmentation of breast, pectoral muscle, and background in digitized mammograms. In: *IEEE International Conference on Image Processing (ICIP)*, pp. 912–916 (2014)
42. Wirth, M.A., Stapinski, A.: Segmentation of the breast region in mammograms using active contours. In: *Visual Communications and Image Processing, International Society for Optics and Photonics*, pp. 1995–2006 (2003)
43. Ferrari, R.J., Frere, A.F., Rangayyan, R.M., Desautels, J.E.L., Borges, R.A.: Identification of the breast boundary in mammograms using active contour models. *J. Med. Biol. Eng. Comput.* **42**(2), 201–208 (2004)
44. Chaabani, A.C., Boujelben, A., Mahfoudhi, A., Abid, M.: An automatic-pre-processing method for mammographic images. *Int. J. Digit. Content Technol. Appl.* **4**(3) (2010)
45. Wang, L., Zhu, M., Deng, L., Yuan, X.: Automatic pectoral muscle boundary detection in mammograms based on Markov chain and active contour model. *J. Zhejiang Univ.-Sci. C (Computers & Electronics)* **11**(2), 111–118 (2010). ISSN 1869-1951
46. Kim, J.H., Park, B.Y., Akram, F., Hong, B.W., Choi, K.N.: Multipass active contours for an adaptive contour map. *Sensors* **13**(3), 3724–3738 (2013). ISSN 1424-8220
47. Akram, F., Kim, J.H., Whoang, I., Choi, K.N.: A preprocessing algorithm for the CAD system of mammograms using the active contour method. *Appl. Med. Inf.* **32**(2), 1–13 (2013)
48. Ma, F., Bajgera, M., John, P., Slavotinekb, Bottemaa, M.J.: Two graph theory based methods for identifying the pectoral muscle in mammograms. *J. Pattern Recogn.* **40**, 2592–2602 (2007)
49. Camilus, K.S., Govindan, V.K., Sathidevi, P.S.: Computer-aided identification of the pectoral muscle in digitized mammograms. *J. Digit. Imaging* **23**(5), 562–580 (2010)
50. Cardoso, J.S., Domingues, I., Amaral, I., Moreira, I., Passarinho, P., Comba, J.S., Correia, R., and Cardoso, M.J.: Pectoral muscle detection in mammograms based on polar coordinates and the shortest path. In: *Engineering in Medicine and Biology Society, Annual International Conference of the IEEE.*, pp. 4781–4784 (2010)
51. Karnan, M., Thangavel, K.: Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of micro-calcifications. *Comput. Methods Programs Biomed.* **87**(1), 12–20 (2007)
52. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888–905 (2000)
53. Domingues, I., Cardoso, J.S., Amaral, I., Moreira, I., Passarinho, P., Comba, J.S., Correia, R., Cardoso, M.J.: Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs. In: *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3158–3161 (2010)
54. Aroquiaraj, I.L., Thangavel, K.: Pectoral Muscles Suppression in Digital Mammograms using Hybridization of Soft Computing Methods (2014). [arXiv:1401.0870](https://arxiv.org/abs/1401.0870)
55. Sapate, S.G., Talbar, S.N.: Pectoral muscle extraction using modified K-means algorithm for digital mammograms. *J. Med. Phys.* (2016)
56. Hartigan, J.A., Wong, M.A.: A K-means clustering algorithm. *J. R. Stat. Soc. Blackwell Publishing* **28**(1), 100–108 (1979)