Automatic Extraction of Pectoral Muscle in MLO View of Mammograms

C K Feudjio[1]‡, J Klein[2], A Tiedeu[1] and O Colot[2]

[1] LETS, National Advanced School of Engineering, P.O. Box 8390 Yaoundé, Cameroon

[2] LAGIS - UMR CNRS 8219, Université Lille1, France

E-mail: `cyrille.feudjio-kougoum@univ-lille1.fr`

Abstract. Mammogram is the standard modality used for breast cancer screening. Computer Aided Detection (CAD) approaches are helpful for improving breast cancer detection rates when applied to mammograms. However, automated analysis of mammogram often leads to inaccurate results in presence of the pectoral muscle. Therefore, it is necessary to first handle pectoral muscle segmentation separetely before any further analysis of a mammogram. One difficulty to overcome when segmenting out pectoral muscle is its strong overlapping with dense glandular tissue which tampers with its extraction. This paper introduces an automated two step approach for pectoral muscle extraction. The pectoral region is firstly estimated through segmentation by mean of a modified Fuzzy C-Means clustering algorithm. After contour validation, the final boundary is delineated through iterative refinement of edge point using average gradient. The proposed method is quite simple in implementation and yields accurate results. It was tested on a set of images from the MIAS database and yielded results which compared to those of some state-of-the-art approaches, were better.

‡ Also with LAGIS - UMR CNRS 8219, Université Lille1, France and Department of Electrical Engineering, College of Technology, P.O. Box 63 Buéa, Cameroon

1. Introduction

Mammography is the standard method used for breast cancer screening. During screening campaigns, a huge amount of images is collected. These mammograms are to be analyzed by a few radiologists. Computer aided detection (CAD) has therefore been introduced to relieve radiologists workload by providing a first aid opinion and improving detection rates (Reddy and Given-Wilson, 2006). In this case, CAD consists of an automatic analysis of images in order to extract or identify patterns in breast region which can facilitate earlier breast cancer detection.

In most CAD systems, a prior extraction of patterns like breast contour (Feudjio et al., 2012; Karnan and Thangavel, 2007; Liu et al., 2011), nipple (Karnan and Thangavel, 2007) and pectoral muscle (Ferrari et al., 2004; Kwok et al., 2004; Liu et al., 2011) have to be carried out to allow accurate analysis of breast tissues. In this paper, we focus on the pectoral muscle extraction step. Pectoral muscle mainly appears in MLO view mammograms. Roughly speaking, it is assumed to be a triangular region with high grey levels intensities located at the upper left corner of the breast region, provided that the breast is right oriented in the image. In cranio-caudal (CC) view, pectoral muscle is found only in 30%-40% of cases (Eklund et al., 1994). If present, the pectoral muscle covers a little area and does not meaningfully impair automatic analysis of CC view mammograms.

Pectoral muscle is a high density tissue, thereby leading to grey level intensities and texture characteristics similar to masses and microcalcications. It is consequently hard to automatically and jointly characterize fibro-glandular tissues, dense structures in breast region as well as masses and microcalcifications in raw mammogram images. Both texture based methods (Ferrari et al., 2004) and breast tissue density based methods (Karssemeijer, 1998; Saha et al., 2001) for mammogram analysis are impaired by the presence of the pectoral muscle. Reliable and discriminative features cannot be extracted which accounts for both false positives and false negatives. False negative number is all the larger as the overlap region between the pectoral muscle and glandular tissues is a common area for cancer to develop and is particularly checked by radiologists to reduce the rate of missed cancers.

Another interest of extracting pectoral muscle is that its contour is used as landmark for registration (Ma et al., 2007) in mammograms comparison or as three-dimensional axis required for breast reconstruction from multi-views of mammograms (Zhou et al., 2010; Kwok et al., 2004). In view of all these motivations, segmenting out pectoral muscle in MLO view mammograms is a justified preprocessing step to improve computerized aided analysis for breast cancer detection.

1.1. Related works

Several major difficulties to overcome for automatic extraction of pectoral muscle in mammograms are known:
- There is a strong overlap with fibro-glandular tissues that makes it difficult to reliably

characterize the pectoral muscle image region.

- Some skin folding tampering with the pectoral muscle contour can be found.

- X-ray images tend to produce blur edges for pectoral muscles espcially in its lower part.

- There is a high variabilty of region surface ; the most tedious cases are met when the pectoral mucle covers a very small area or is completely missing.

These various and complex factors illustrate how difficult it is to automatically extract pectoral muscle in mammograms (Zhou et al., 2010).

Many approaches dealing with pectoral muscle extraction have been introduced in the literature. Suckling et al. (1995) extracted pectoral muscle by segmenting a mammogram in four main type of components which includes background, pectoral muscle, fibro-glandular region and adipose region using a multiple linked self-organizing neural network. A semi-automatic method introduced by Saha et al. (2001) requires an input from an operator to locate the pectoral muscle and therefore its delineating was then performed automatically. Raba et al. (2005) used region growing method to extract pectoral muscle. However, their results were assessed only visually and rated on a two scale grade (adequate and quite adequate).

Another commonly used approach consists of estimating the boundary between pectoral muscle and mammary tissues as a straight line by the use of Hough transform (Ferrari et al., 2004; Karssemeijer, 1998). The main drawback of this method is that the pectoral muscle edge is not always straight and may sometimes present concavities. To cope with this inconvenient Kwok et al. (2004) firstly estimated the straight line delineating the pectoral muscle using iterative thresholds and then refined this line by cliff detection to be aligned on the pectoral muscle boundary curvature.

Other algorithms developed for extraction of pectoral muscle are based on texture-field orientation (Zhou et al., 2010), wavelets decomposition (Mustra et al., 2009), Gabor wavelets (Ferrari et al., 2004) and non linear filtering (Mirzaalian et al., 2007). In these approaches, the final pectoral muscle contour is obtained by thresholding of the line segments or candidate regions based on shape and size criteria. Chakraborty et al. (2012) introduced a method based on the average of gradient to extract pectoral muscle. They used a weighted average gradient and adaptive band selection to approximate the straight line edge of the pectoral muscle and local gradient to adjust the line to the edge of the pectoral muscle.

Recently, a new image segmentation approach based graph theory has been introduced. This approach relies on the assumption that the pectoral muscle region boundary has its two extremities in the first row of the image and the first non blank column respectively. These extremities are called end-points and the shortest path between two endpoints is searched by minimization of a cost-function. Domingues et al. (2010) estimated the pectoral muscle contour as the shortest path throughout the two end-points in the image gradient. The latter is represented as a weighted graph where nodes are pixels and edges are connecting neighbourhood pixels. Finally the shortest path is found by mean of a Support Vector Machine (SMV) previously trained. Ma et al. (2007)

introduced another graph theory approach based on adaptive pyramids and spanning trees to extract pectoral muscle.

It is noteworthy that most of these works were tested on small datasets and that the accuracy of their algorithms was based on visual inspection. Moreover, only few perfomance metrics are provided. In this work, we report our results compared to a gold standard and attempt to provide a set of performance metrics covering all those chosen in orther studies in oder to make algorithms performance comparison easier.

The analysis of this litterature review shows how complex it is to extract the pectoral muscle and that its boundary cannot be retrieved through a straightforward method. Commonly (Chakraborty et al., 2012; Kwok et al., 2004; Ma et al., 2007), the extraction is done in two main steps. The pectoral region frontier is roughly estimated based on gray level homogeneity or a priori information on the pectoral muscle location. The boundary is then refined by selection of segment candidate or research of the shortest path (Ma et al., 2007; Zhou et al., 2010), or cliff detection using gray level variation (Kwok et al., 2004). In this paper, we also introduce an approach based on these two steps. The pectoral region is firstly estimated using a modified version of the Fuzzy C-Means (FCM) algorithm. Secondly, for each contour pixel of the estimated region, the actual pectoral muscle boundary is sought along the orthogonal direction to the contour by mean of average gradient. In the following subsection, a more detailed outline of our approach is given.

1.2. Proposed method outline

The flowchart of the proposed method is presented in figure 1. Regarding pectoral muscle layouts in mammograms, we make the following reasonable assumption: Pectoral muscle appears at only two different locations in mammograms considering left or right breast. Therefore, a pre-processing is needed so that the pectoral muscle lies at the same location (upper left corner). The symmetry performed here aims at reducing programming complexity. By convention, only images with left oriented breast have to be flipped sideways. After this step, a region of interest (ROI) is defined on the basis that the pectoral muscle lies at upper left corner. This ROI is obtained using a heuristic based on a priori knowledge on women anatomy and mammography protocol. Defining a ROI allows to reduce the area of research and the computation time while ensuring better segmentation results.

Once the ROI is obtained, it can be segmented into regions. The only pieces of information available for segmentation are textures and gray levels intensities (Zhou et al., 2010). However, as mentionned before, texture and gray level features are not very discriminative. The feature distributions of the regions are overlapping. The FCM segmentation algorithm was introduced to deal with such data. In this work, we estimated the pectoral muscle with a modified version of Fuzzy C-Means (mFCM). The standard FCM algorithm has been modified so that it could advoid random initialization of clusters to speed up the clustering. The mFCM also uses local information in
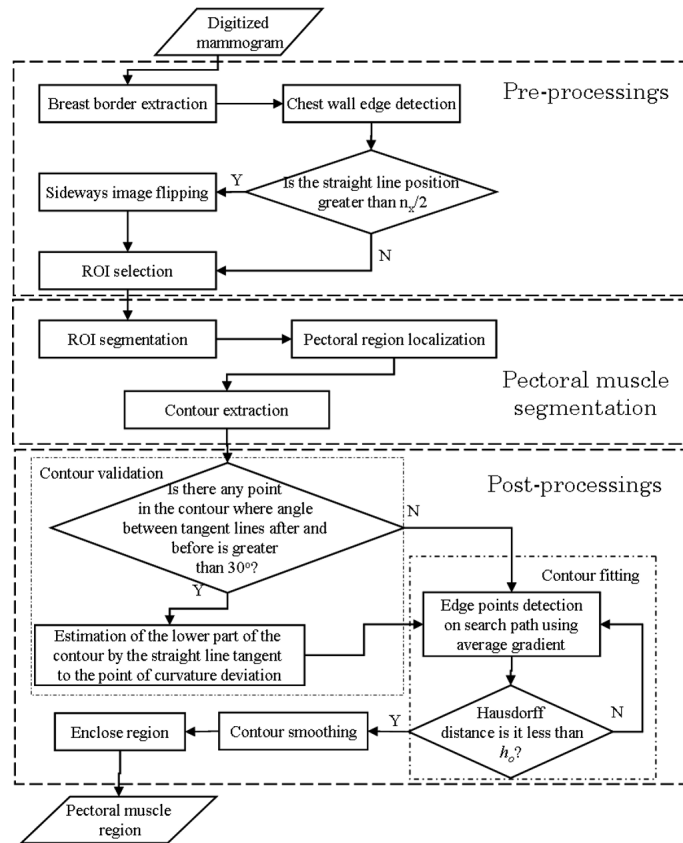
Figure 1. Flowchart of the proposed method for pectoral muscle detection

membership function to better classify noisy pixel particularly those located on the vicinity of the pectoral edge.

The segmentation step provides us with an approximation of the pectoral muscle boundary. In some cases, the lower part of the pectoral muscle completely overlaps with glandular tissue leading to an inaccurate pectoral region estimation. The pectoral region is overestimated and its contour deviates from the true boundary towards breast tissue. To cope with this difficulty a post-processing step for contour refinement is needed. Angle curvatures are computed at each point of the contour. Starting from image top contour point, each next contour point is processed as follows: if the contour point curvature is abnormally high, the following contour point are replaced using a tangent line computed from the nearest non-deviating contour points. Finally, the contour is iteratively refined by using average gradient to accurately estimate the points of the pectoral edge.

2. Materials and methods

2.1. Materials

2.1.1. Dataset. The proposed method was tested on 277 images selected from the mini-MIAS database (Suckling et al., 1994). This open database was chosen because it is the most widely used in the literature dealing with pectoral muscle extraction. It therefore allows an easy comparison of our results with those of other methods. All images in this database are MLO views mammograms digitized at 200 μm and 8-depth resolution. The images are 1024x1024 pixels size and were further subsampled to 512x512. The images not presenting the pectoral muscle as well as those having a band tape or very poor contrast quality were not included to the study. Following this pre-selection policy, only 45 images out of 322 were discarded. The selected images were acquired from 157 patients whereas the 83 images used for inter-observer variability study were a subset of the selected images.

2.1.2. Radiologists' expertise as reference standard. The coordinates of the radiologist's drawn boundaries of pectoral muscle of images used in (Ferrari et al., 2004) as well as coordinates of pectoral muscle obtained by theirs algorithms were kindly provided by Rangayyan so that our results could be compared to the same standard. To extend the experiment to the whole database, another radiologist manually drew the pectoral muscle contour of all mammograms in the dataset. His expert work was used as reference standard for evaluation of the computed performances of pectoral muscle extraction. An inter-obsever variabilty was conducted to evaluate the agrement between the two hand-drawn pectoral muscle boundaries for the set of images used in (Ferrari et al., 2004). This inter-observer varibility shows a strong correlation between the two hand-drawn as described in section 3.2. The performances of the proposed method is assessed on several criteria as defined in the next section.

2.1.3. Performance metrics The accuracy of the pectoral muscle detection is evaluated by the following performance metrics:
(1) false positive (FP) and false negative (FN) rates, which evaluate pixel assignment errors. A FP pixel is one assigned by the algorithm as belonging to the pectoral muscle but assigned by the radiologist outside of the pectoral muscle. A FN pixel is one assigned outside of the pectoral muscle by the algorithm but assigned inside by the radiologist. The FP and FN rates for an image are computed as follows:

$$FP = \frac{\mid A_C \cup A_R \mid - \mid A_R \mid}{\mid A_R \mid}$$

$$FN = \frac{\mid A_C \cup A_R \mid - \mid A_C \mid}{\mid A_R \mid}$$

where A_C and A_R are the areas of the pectoral muscle regions obtained by the algorithm and the radiologist respectively.

(2) Percent Overlap Area (POA), which expresses the accuracy of the region delineated by the algorithm to the reference one drawn by the radiologist. It is defined as:

$$POA = \frac{A_C \cap A_R}{A_C \cup A_R}$$

(3) The Hausdorff distance (H_{dist}), which is used to estimate the discrepancy and the dissimilarity of the detected boundary by the algorithm with the one drawn by the radiologist. It is defined as:

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

where c_i and r_j are contour points obtained by the algorithm and the radiologist respectively, $Dist(c_i, r_j)$ is the Euclidean distance between points c_i and r_j calculated in pixel units (1pixel=0.4mm).

These metrics are widely used thereby allowing a fair comparison with related works. There are also complementary and can thus catch every aspects of image segmentation performances.

2.2. Methods

2.2.1. Pre-processings. In order to make the segmentation task easier, a ROI is extracted in the upper left corner of the image as illustrated in figure 2. But, as the pectoral muscle does not always lie at that location, breast orientation needs first to be retrieved. Consequently, our algorithm starts with a pre-processing step which aims at producing mammogram images with identical region layouts. The sequel of this section discusses breast orientation detection in mammograms, ROI selection and ROI characteristics analysis.

Breast orientation detection. Breast orientation is determined through the following steps:

- breast contour detection: we use the approach described in (Feudjio et al., 2012). Note that the main pattern for breast orientation detection is the breast shape. So, even using a rough segmentation technique such as thresholding the image with its mean value enables to retrieve the breast shape.

- chest wall position estimation: the chest wall is in first approximation a straight line. Its position is thus determined by applying a Radon transform to the breast contour image. In figure 2, the chest wall corresponds to segment OB.

- breast orientation retrieval: if the chest wall edge is located in the first right half of the image, the breast is oriented right otherwise it is oriented left.

Left-oriented breast mammograms are then vertically mirrored thereby allowing to process images with identical region layouts.
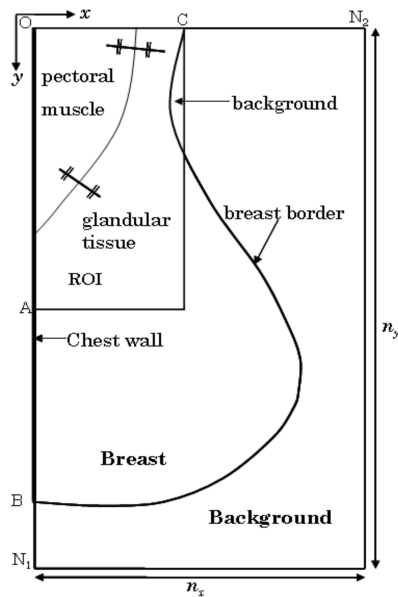
Figure 2. Mammogram with a right orientation and a pectoral muscle at the upper left corner of the image. The coordinate axis are directed as shown with the origin at top left. N_1 and N_2 are respectively the number of rows and columns in the image. ROI is selected as a rectangular window of size $OAxOC$, where OA is half height of the image and C is the upper endpoint of the breast contour.

ROI selection. As the pectoral muscle is roughly triangular, the ROI is therefore selected as a rectangular region starting at the origin O. Its height is taken equal to OA where A is the half height of the image. This allows to reduce the amount of dense glandular tissue in the ROI in order to guarantee good segmentation results. On the other hand, the ROI width should be large enough to fully contain the pectoral muscle. A simple way to ensure this is to take the ROI width equal to OC where C is the top endpoint of the breast border (see figure 2). Mammograms are now ready to be processed for pectoral muscle segmentation.

2.2.2. Pectoral muscle segmentation. Ideally, the ROI should entirely contain the pectoral muscle while rejecting as much as possible other regions in order to ensure good extraction results. However, in practice such an hypothesis is not always met since there is a high variability in shape and size from one patient pectoral muscle to another. The concavity of the breast border in the neighbourhood of point C (see figure 2) implies that the ROI comprises a small part of the background. Therefore, the ROI selected mainly consists of three regions: pectoral muscle, breast glandular tissue, and background.

The choice of the segmentation method applied in our approach is justified from an analysis of characteristics of regions contained in the ROI. Figure 3(a) shows the ROI extracted in a mammogram where the pectoral muscle has been delineated from breast

(a) (b)
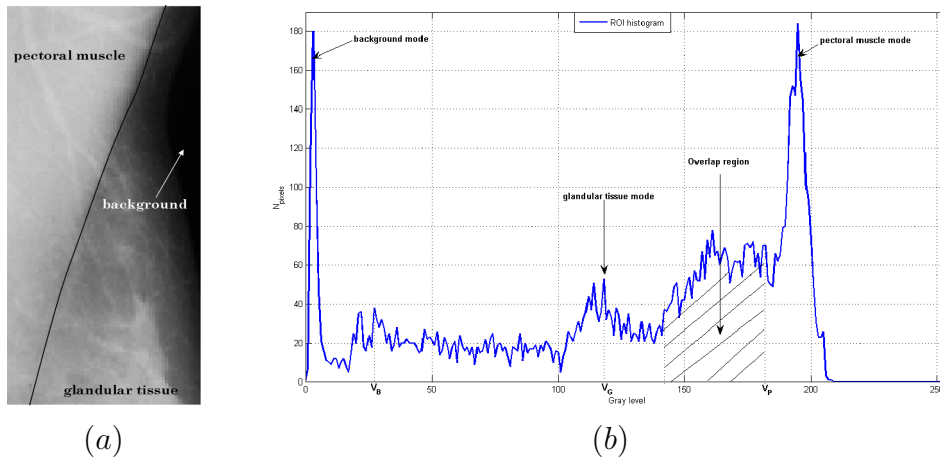
Figure 3. ROI extracted from a mammogram (mdb018). (a) ROI with pectoral muscle, breast tissue and non breast tissue. Pectoral muscle boundary was traced out manually. The lower portion of the ROI presents an overlap of pectoral muscle and breast tissue. (b) Intensity histogram of the ROI. V_B, V_G and V_P are modes of the following regions background, breast tissue and pectoral muscle respectively.

tissue manually. In many MLO mammograms, the lower portion of the pectoral muscle is spatially superimposed on some glandular tissue. The impact of the overlap between glandular tissue and pectoral muscle can easily be observed on the ROI histogram (see figure 3(b)). There is no clear cut separation between the modes of the histogram. Thus, it is not always possible to find thresholds that can completely separate the pectoral muscle from other tissues (Kwok et al., 2004). In such cases, FCM algorithm which is an unsupervised classification method can be used to optimally cluster overlapping data. The next subsections address the FCM algorithm and show how we modified it to take into account spatial information.
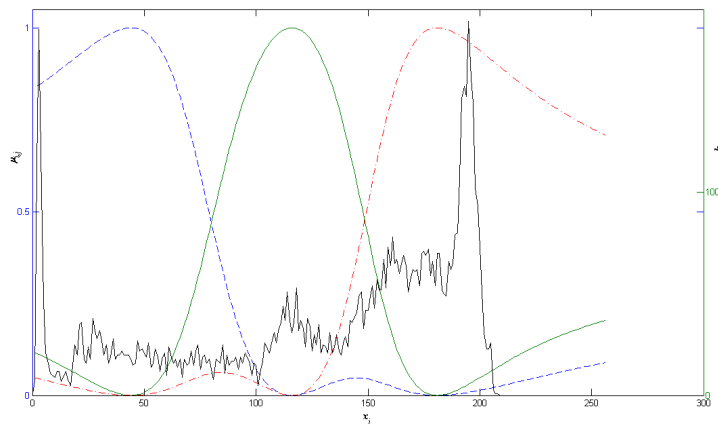
Fuzzy C-Means algorithm. The FCM is an unsupervised data labelling algorithm commonly used in image processing for segmentation tasks. This is an appropriate method for clustering overlapping data. Let us denote the gray level intensity of the jth pixel in the ROI by x_j. For each image region in the ROI, one membership function is defined with respect to x_j. The value of the membership function $\mu_{i,j}$ depicts the possibility for a pixel to belong to the ith region given its grey level intensity x_j. Membership functions are given by:

$$\mu_{i,j} = \left(\sum_{k=1}^{c} \left(\frac{d(x_j, v_i)}{d(x_j, v_k)} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{1}$$

with $m > 1$ a parameter of fuzzification control and where v_i is the mode of the ith image region:

$$v_i = \frac{\sum_{j=1}^{n} \mu_{i,j}^m x_j}{\sum_{j=1}^{n} \mu_{i,j}^m}. \tag{2}$$

The membership functions only depend on the gray value level and can thus be superimposed with the ROI histrogram (see figure 4).

Figure 4. Membership functions of regions: background (- - - -), glandular tissue (——) and pectoral muscle (— · —) superimposed on the ROI histogram of figure 3(b).

The FCM algorithm main principle consists in minimizing the inter-class distance through an objective function J_m defined by:

$$J_m(\mu, v) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{i,j}^m d^2(x_j, v_i) \tag{3}$$

where n is the number of pixels in the ROI and c the number of image regions to segment. d is the Euclidean distance from a pixel intensity x_j to the region mode v_i : $d(x_j, v_i) = \|(x_j - v_i)\|$. Moreover, we have the following constraint : $\forall j, \sum_{i=1}^{c} \mu_{i,j} = 1$. As can be observed from equations (1) and (2), one needs the region modes to compute membership functions and conversely, therefore the FCM algorithm is an iterative algorithm which starts by randomly initializing the modes and evaluating the membership function. At each iteration, the membership function and the clusters center are updated. An optimal partition of clusters is obtained when from one iteration to another the euclidean distance between the previous and the current cluster center is less than a convergence parameter noted ϵ. Segmentation of the image is then performed by assigning each pixel to the region with the highest membership function value .
Two major weaknesses of the classical FCM algorithm are random initialization of region modes which penalizes fast convergence of the algorithm and non-use of spatial information leading to clustering two pixels of same gray level to the same region no matter their spatial localization. Furthermore as the FCM algorithm is gradient descendent, random initialization of cluster may lead the algorithm to converge towards the local minimum.

Modified FCM. As explained in the previous subsections, the FCM algorithm often misclassifies some pixels, we therefore introduce a modified FCM algorithm that allows better pixel classification results. It has been shown that integrating local information

when updating pixel membership functions significantly improves the segmentation performances in MRI brain images (Wang and Wang, 2008). The neighborhood information is integrated in the clustering process by computing a weighted version of the membership function defined as followed:

$$\tilde{\mu}_{i,j} = \frac{\mu_{i,j}^m S_{i,j}^n}{\sum_{k=1}^c \mu_{k,j}^m S_{k,j}^n} \tag{4}$$

where $S_{i,j} = \sum_{k \in N(x_j)}^c \mu_{i,k}$ is the spatial function computed in a squared window N centered on the pixel j and m, n are weighting parameters. Consequently, the membership functions are now not only depending on pixel gray level intensity but also on neighboring pixels and they can no longer be superimposed on the ROI histogram. Figure 5 shows the effects of integrating local information on segmentation.
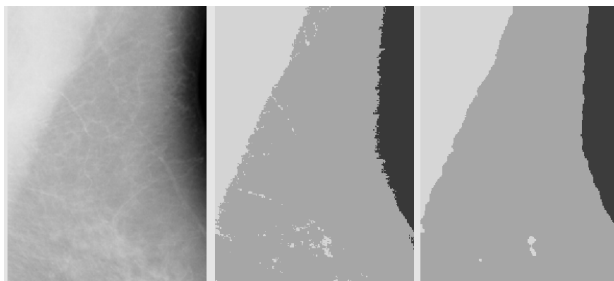
Figure 5. Segmentation of the ROI into 3 regions: pectoral muscle, breast tissue and background. On the left column the initial image, the two next columns are results of segmentation using FCM algorithm without and with local information respectively.

One can see that some misclassified pixels in the ROI have been well classified thanks to neighborhood information. This modified FCM algorithm can still be criticized based on the following grounds: increased computation load and dependance on clusters initialization. Similar conclusions are drawn in (Wang and Wang, 2008) and the FCM algorithm needs further improvements to prevent from bad segmentation results.

It has been shown that neatly estimating region modes significantly reduces the number of iterations while enhancing accuracy in the results. One appealing method of estimating region modes is the block density approach (Guo et al., 2009). In the latter, the image histogram h_j is subdivided into c blocks B_i of equal size, one block for each region involved in the segmentation. If the image has 256 gray levels, then $B_i = \left\{(i-1) \times \frac{256}{c}, ..., i \times \frac{256}{c} - 1\right\}$. A random variable X is defined to represent the probability for a pixel to have x_j as gray value level. The distribution of variable X corresponds to gray level occurrence in the image, *i.e.* the histogram: $P(X = x_j) = h_j$. The initial mode of the first region denoted by $v_1^{(0)}$ is computed as the highest probability level belonging to the first block and its probability is denoted by h_{v_1}. The subsequent region modes $v_2^{(0)}, ..., v_c^{(0)}$ are then computed by finding the gray level of the block that maximizes the following expression:

$$v_i^0 = \arg\max_{j \in B_i} h_j \times D(x_j, v_{i-1}^0) \tag{5}$$

with $D(x_j, v_{i-1}^0)$ the distance between two levels which is given by:

$$D(x_j, v_{i-1}^0) = \left((x_j - v_{i-1}^0)^2 + (h_j - h_{v_{i-1}})^2\right)^{1/2} \tag{6}$$

The membership function are initialized at robust regions modes and then updated by using local information. Here is a outline of the algorithm:

Algorithm 1 Modified FCM (mFCM)

Require: $c = 3$, $m = 2$, $n = 3$, $\epsilon = 0.01$, n_w=5x5

$\boldsymbol{v}^{(0)} = \left\{v_1^{(0)}, ..., v_c^{(0)}\right\} \leftarrow$ Estimate region modes using sample block density with equation (5)

for i from 1 to c **do**

  $\mu_{i,j}^{(k+1)} \leftarrow$ Evaluate membership function with equation (1)

end for

while $\|\boldsymbol{v}^{(k+1)} - \boldsymbol{v}^{(k)}\| > \epsilon$ **do**

  for i from 1 to c **do**

    $\tilde{\mu}_{i,j}^{(k+1)} \leftarrow$ Update weighted membership function with equation (4)

  end for

  for i from 1 to c **do**

    $v_i^{(k+1)} \leftarrow$ Update regions mode with equation (2)

  end for

  $k \leftarrow k + 1$

end while

2.2.3. Post-processings for pectoral muscle contour refinement

Incorrect pectoral muscle region detection and correction.   Once the mFCM algorithm has converged, the segmentation is performed by assigning each pixel to a region where its membership function is the highest. The pectoral region is then identified as the area with highest gray level and located upper left corner of the image.

The mFCM segmentation results are satisfying in most cases but a post-processing is needed to detect and correct wrongly segmented images. Figure 6 shows the segmentation results in a rather difficult case. The slight variation of gray level intensities along the pectoral muscle edge leads to an inaccurate contour estimation. In addition, the strong overlapping between glandular tissues and pectoral muscle in its lower part results in over-estimated pectoral muscle region. In this latter case, the contour of the pectoral muscle deduced from the segmentation deviates inward-breast direction with a concave curvature.

A validation process of the estimated pectoral region is carried out to detect and remove the overestimated area. This is done by computing the angle of curvature to find contour points with the highest deviation. The angle of curvature θ_j of the jth contour point is computed as angle between the tangent line of preceding contour points and
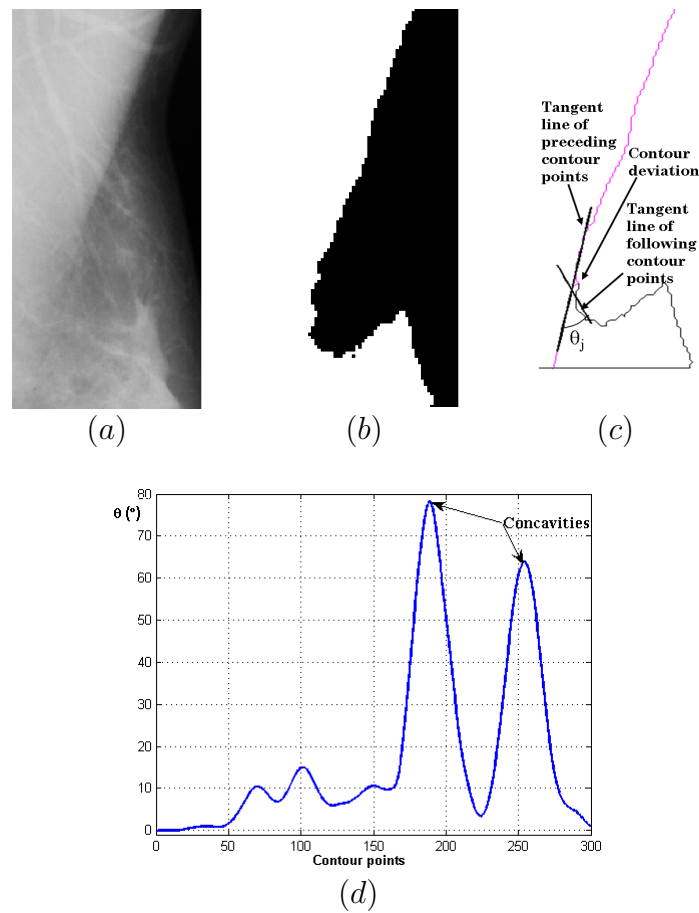
Figure 6. Pectoral muscle region validation: (a) Initial image, (b) Estimated pectoral muscle region from segmentation, (c) Extracted contour with tangent lines (bold black segments) at the deviation point of the contour and the final estimated contour with the lower part fitted to straight line after the point of deviation, (d) Angle curvature of the extracted contour. The peaks indicate the concavities due to contour deviation.

the tangent line of following contour points at each point of the contour (see figure 6). The angle between these two tangent lines is given by the formula:

$$\theta_j = tan^{-1}\left(\frac{a_1 - a_2}{1 + a_1 a_2}\right) \tag{7}$$

where a_1 and a_2 are the respective slopes of the straight tangent lines.

It can be seen from figure 6 that the estimated pectoral muscle contour is very noisy (quick variation of contour pixels' position). Consequently it is necessary to compute the tangent lines on a rather wide range of contour points to smooth the absolute value of the angle of curvature curve. We obtained reliable information on contour for smoothed angle curvature computed on 25 pixels (1cm) tangent lines.

Abnormal deviations in the contour are found in peak areas in the angle of curvature. Conversely if there is no abnormal deviation in the contour, the angle of curvature will remain quite flat. Figure 6(d) shows the angle of curvature after smoothing in the case

where the pectoral region was overestimated due to glandular tissue overlapping. The stronger the peak, the more likey it is to have a wrong contour deviation at that point. Based on this observation, a threshold was established at 30° to detect contour areas where an abnormal deviation occurs. When the angle of curvature is greater than this threshold, the rest of the contour is extrapolated from the tangent straight line of points preceding the deviation. This approximation allows to cope with the cases of pectoral region overestimation especially where there is an overlap between pectoral muscle and glandular tissue. In cases where the pectoral muscle is not entirely contained in the ROI, the rest of its contour in the image is estimated by a straight line on the basis that pectoral muscle is in first approximation triangular. In the subsequent steps, the images are processed in their full height.

Now, the contour obtained from the estimated pectoral region roughly corresponds to the true pectoral boundary but does not perfectly fit the actual pectoral muscle contour. To improve the accuracy of the algorithm, a refinement of the extracted contour is carried out. The next subsection explains the strategy carried out to refine the contour in order to improve the accuracy of the pectoral muscle extraction. It is made of two sub-steps: pectoral muscle boundary fitting and pectoral muscle contour smoothing.

(a)   (b)   (c)
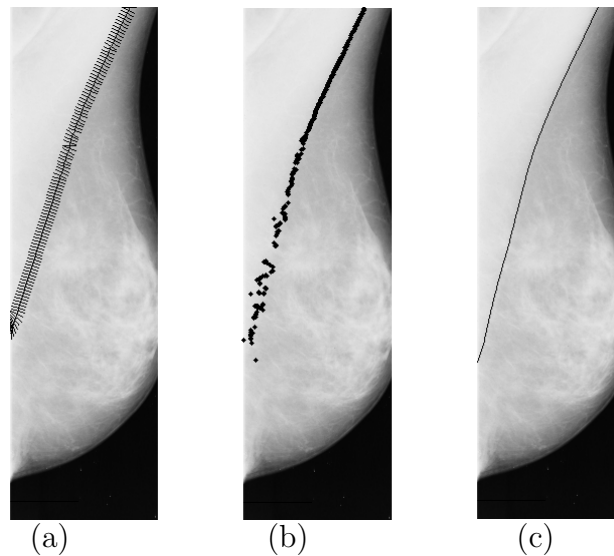
Figure 7. Pectoral muscle contour refinement. (a) Validated contour with search paths drawn every three contour points, (b) True pectoral edge points detected after boundary fitting, (c) Final pectoral muscle contour obtained after smoothing the true pectoral edge points.

Pectoral muscle boundary fitting. The pectoral muscle contour fitting approach developed in this work is based on the idea introduced in (Kwok et al., 2004). This approach relies on the definition of a search path for each point of the estimated pectoral contour. A search path is defined as the orthogonal line segment centered on its corresponding contour point. The search path slope is obtained straightforwardly

from the previously computed angle curvature by adding 90°. The length of the search path denoted L was fixed at 15 pixels (6mm). At some points of contour, search paths exceed the ROI. In these cases, they are rotated accordingly so that they lie within the breast area (see figure 7(a)). Having obtained these paths, we make the following reasonable assumption: the correct pectoral muscle contour point belongs to the search path, one just has to find out which one it is. In (Kwok et al., 2004), the intensity profile along the search path is modelled as a sigmoid function. The pectoral muscle edge point is supposed to lie at the inflection point of the sigmoid. However, it is hard to model the intensity profile to a sigmoid in the area where the pectoral muscle overlaps with glandular tissue because there is not a sharp variation of the gray level intensity within these areas. Consequently, the fitted contour may deviate from the true edge. A refinement method based on local maximum average gradient search was introduced by Chakraborty et al. (2012). This method is based on the fact that the gradient is known to be higher at the edge of two regions. The average gradient AG is computed at each point of the image along the x-axis as follows:

$$AG(x_j) = \frac{2}{L-1} \sum_{1 \leq k \leq \frac{L-1}{2}} \frac{x_{j-k} - x_{j+k}}{2k} \tag{8}$$

Note that the formula above assumes that pixels are indexed from left to right, for each image line.

Thanks to its smoothing aspect, the average gradient is robust to spike structures observed in the gradient of a profile. It clearly depicts the location of edges with a prominent peak even in boundary area with low pixels intensities variation and thus make it easier to retrieve edge point.

The edge point of the pectoral muscle is the one whose average gradient value is maximal among pixels in the search path. It was observed in some cases that shortening the search path during boundary fitting as described in (Chakraborty et al., 2012; Kwok et al., 2004) does not produce accurate contour and more iterations were needed to obtain a good one. Therefore, instead of shortening the search path we used a fixed length of the search path and perfomed iterations until a convergence of points of the contour is obtained. The convergence is obtained when from one iteration to another the Haussdorf distance between the previous contour and the current one is less than a threshold (2 pixels or 0.8mm).

Pectoral muscle contour smoothing. After fitting the contour, a set of points detected as true pectoral edge pixels are found. The contour points are smoothed within a moving window to avoid the effect of noisy detected points. The smoothing is performed by a locally weighted least square fitting method. For each data point in the moving window, the regression weights are comptued with the following characteristics:

- the data points to be smoothed has the largest weight and the data outside of the window have zero weight.
- a linear least square regression is computed using a first degree polynomial.

The smoothed value is given by the weighted regression as a predictor value of interest. After the smoothing process, it may happen that a row in the area of pectoral muscle has more than one edge point or no edge point at all. If two contour points are detected in a ROI row, then the pixel with the maximum gradient is chosen as pectoral edge pixel.

For no edge pixel in a row, the pectoral pixel is estimated to lie at the same column as the one in the previous row. The pectoral region is finally enclosed by an extrapolated straight line if the smoothed contour does not reach the chest wall.

3. Results and performances evaluation

This section presents the results yielded by the proposed method on an open dataset of images as well as discussion and comparison of results with those of related works.

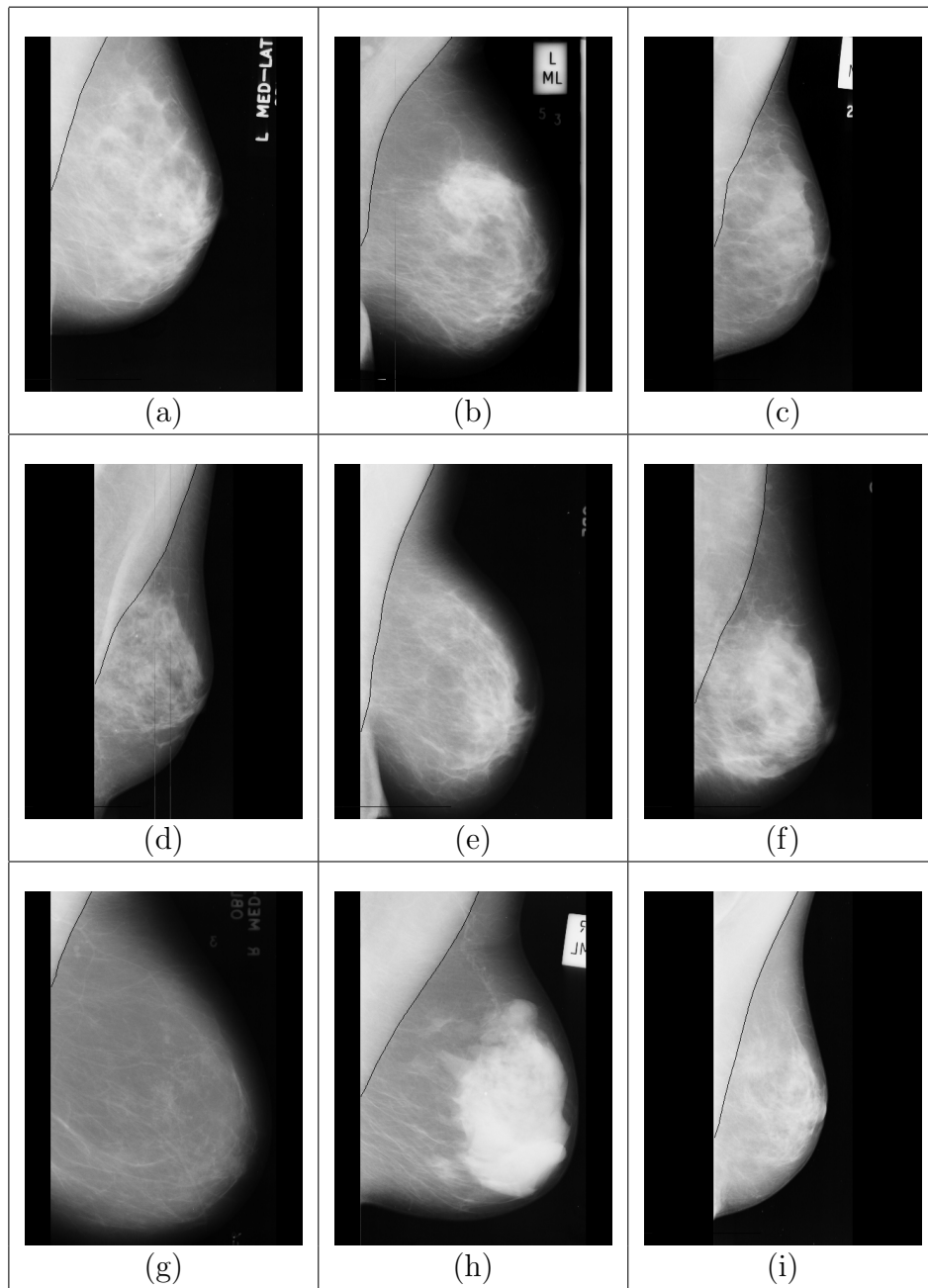### 3.1. Pectoral muscle extraction evaluation

Figure 8 shows some pectoral muscle extractions obtained with the proposed method. In comparison with radiologist's manually drawn pectoral muscle boundary, the following results can be highlighted:

- The POA mean and the standard deviation are 86.42±13.23%.
- The FN mean and the standard deviation are are $11.12 \pm 12.53\%$, while the FP mean and the standard deviation are are $3.35 \pm 8.72\%$.
- The Haussdorf distance mean and the standard deviation are 14.83±16.15 mm.
- 55.96% (155/277), 80.51% (223/277) and 89.89% (249/277) of the computed pectoral muscle edges had a greater POA than 90%, 80% and 70% respectively.
- 27.80% (77/277) and 53.07% (147/277) of the computed boundaries had Hausdorff distances within 5 mm and 10 mm from the reference boundaries respectively.

Identifying the pectoral muscle becomes difficult if dense tissue appears near the pectoral muscle. Consequently, as explained in section 2.2.3, a strategy to cope with contour deviation and refine the estimated contour is proposed to track down the true pectoral muscle boundary. Evolution of performances thanks to boundary refinement step based on overlap percentage is presented on figure 9. One can notice that the accuracy of the proposed algorithm has been significantly improved through boundary refinement. For instance, the rate of images with inaccurate ($POA < 80\%$) estimation of pectoral muscle drops from 41.16%(114/277) to 19.49%(54/277) while the one with high accurate estimation ($POA > 95\%$) raises from 0.36%(1/277) to 19.13%(53/277). In addition, the Hausdorff distance drops from 23.32 ± 18.30 mm before refinement process to 14.83±16.15 mm after refinement.

The inaccurate images obtained before boundary refinement were mainly those having strong overlapping between pectoral muscle and glandular tissue or those showing some artifacts or having non uniform gray level intensities in the pectoral muscle area. The above results prove that the pectoral muscle contour fitting step improves the accuracy

Figure 8. Some examples of pectoral muscle boundaries detected by the proposed method and superimposed on the mammograms. (a) mdb029, (b) mdb111, (c) mdb113, (d) mdb185, (e) mdb217, (f) mdb221, (g) mdb272, (h) mdb320 and (i) mdb191.

of the pectoral muscle extraction in those cases.

3.2. Observer variability for identifying pectoral muscle

Regarding the subset of 84 images used in (Ferrari et al., 2004), two hand-drawn pectoral boundaries were available. In the first case, the pectoral boundaries coordinates from a radiologist R2 were those used in Ferrari et al. (2004) and provided by Rangayyang
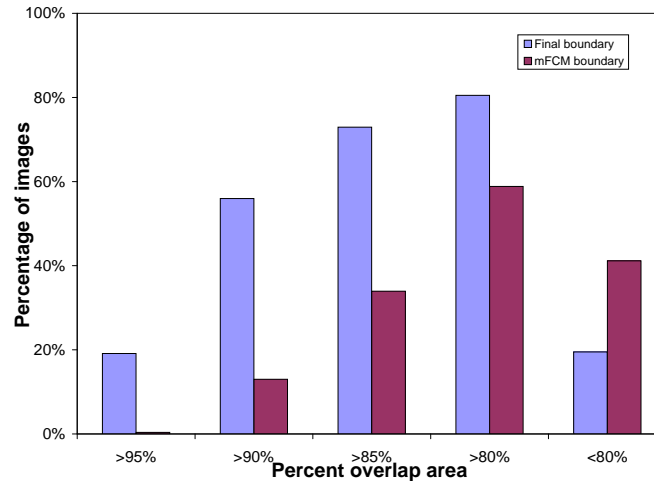
Figure 9. Histogram showing the proportion of images with percent overlap area between the computed and the reference standard pectoral muscle region before and after boundary refinement.

**Table 1.** Effect of the variation of reference standard on the performances of the algorithms.

		R1vsR2	R1vsmFCM	R1vsGabor	R2vsmFCM	R2vsGabor
POA	Mean	95.58	89.12	81.98	87.95	84.27
(%)	Std Dev	2.12	14.09	15.13	14.66	16.02
H_{dist}	Mean	3.89	9.30	10.51	8.50	12.47
(mm)	Std Dev	7.37	8.60	19.46	14.06	21.97

while in the second case, the pectoral boundaries correspond to that drawn by our radiologist R1. These two reference standards were used to study the effect of the reference standard variability on the performance evaluation. The same experiment was carried out on both our results and those of Ferrari et al. (2004) as the outputs of their algorithm were provided. Note that in one very particular image, the radiologist R2 identified a very small region as pectoral muscle while the radiologist R1 did not identified any. This is explained by the fact that the pectoral muscle region is composed of a very few pixels and consequently radiologist R1 considered it unworthy to delineate. This image was therefore discarded from the following study.

Table 1 shows the comparison between the hand drawn pectoral muscle boundaries of radiologist R1 and R2, the computed boundaries of our method and the ones obtained in (Ferrari et al., 2004). For the inter-observer variability evaluation, the results show that the average and the standard deviation of POA and Hdist are $95.58 \pm 2.12\%$ and 3.89 ± 7.37 mm respectively. Furthermore, the results of the proposed method are the closest to those obtained when two radiologists are compared. The Mc Nemar test was used to assess the statistical significance of the performances of these two approaches.

The images fulfilling the accuracy criterion (POA>90%) for both reference standard in one method and do not in the other are 45 (54.22%) with mFCM and only 4 (4.82%) with Gabor method. The Mc Nemar value is 34.31 which compared to χ^2 indicates that the mFCM improvement is statistically significant with a p-value less than 0.01. To evaluate the significance in observer variability, images having POA>90% when compared to one reference standard and less with the other are counted for each method. The Mc Nemar values found are 1.14 and 24 for mFCM method and Gabor method with p-values less than 0.25 and 0.01 respectively. This result shows that Gabor method is statistically sensitive to the reference standard whereas the mFCM which has is p-value greater than the threshold of 0.05 does not exhibit a significant sensitivity.

In figure 10, our approach and Ferrari et al. (2004) approach are further compared. The cumulative percentages of images with performance metrics greater than a given value for both methods compared to each radiologist reference standard are presented. For the POA criterion, the proposed method performs better than the Gabor method for both reference standard. However, for H_{dist} criterion, the performances of the two methods are quite close with minor advantage to Gabor method.
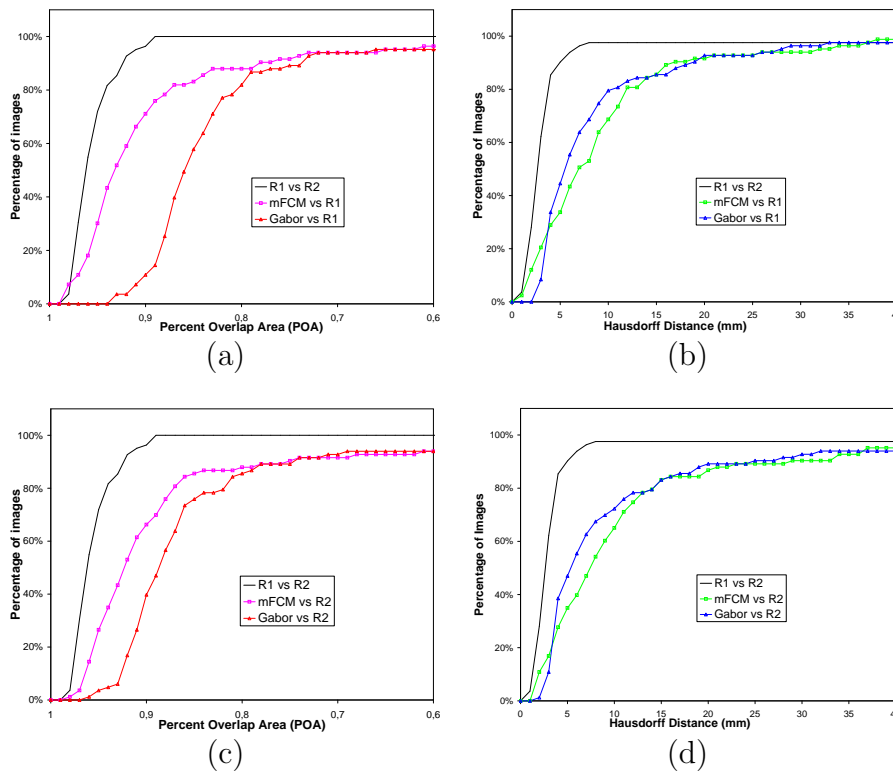
Figure 10. Cumulative percentages of images with performance metrics greater than a given value. Percent overlap area (a), (c) and Hausdorff distance (b), (d) performances of pectoral extraction using Gabor wavelets (Ferrari et al., 2004) and the mFCM compared to two expert reference standards R1 and R2 respectively.

Table 2. Comparison of reported studies using several performance metrics: mean and standard deviation and the percentage of images with the conditions set in the reported studies, where FP and FN rates are considered to be accurate ($< 5\%$), acceptable (5% to 10%) and unacceptable ($> 10\%$). Unprovided data in the publications are marked as NA.

Performance criterion	Hough	Gabor	AP	MST	mFCM
FP (Mean $\pm$ Std Dev)	1.98 ± 6.09	0.58 ± 4.11	3.71	2.55	2.58 ± 6.43
FN (Mean $\pm$ Std Dev)	25.19 ± 19.14	5.77 ± 4.83	5.95	11.68	8.78 ± 13.95
$FP < 5\%$ and $FN < 5\%$	10	45	50	40	38
$min(FP,FN) < 5\%$ and $5\% \leq max(FP,FN) \leq 10\%$	NA	NA	18	20	24
$min(FP,FN) < 5\%$ and $max(FP,FN) > 10\%$	NA	NA	11	18	18
$5\% < FP < 10\%$ and $5\% < FN < 10\%$	8	22	0	0	1
$5\% < min(FP,FN) < 10\%$ and $max(FP,FN) > 10\%$	NA	NA	0	1	1
$FP > 10\%$ and $FN > 10\%$	66	17	5	3	1

*Hough transform and Gabor filter were introduced by Ferrari et al. (2004), AP and MST were presented by Ma et al. (2007) and mFCM is the proposed method.

4. Discussion

4.1. Comparison study

For a coherent comparison, the performances of our algorithm were compared to those applied on same dataset of the same database. The performance criteria were computed on 83 images as we removed from the study all images identified by our radiologist without any pectoral muscle component. Table 2 shows the comparison between the performances of the proposed method and those of four reported methods performance metrics provided in their publications: The mean and standard deviation of the FP and FN rates in comparison to radiologist manually segmented pectoral muscle region. The four related works used for comparison are based on Hough Transform and Gabor filter (Ferrari et al., 2004), adaptive pyramids (AP) and minimum spanning trees (MST) (Ma et al., 2007).

From table 2, one can notice that Gabor wavelets exhibits the lowest rates of FP and FN. However, the proposed method perfoms as well as other methods presented in the table. A robust measure of pectoral muscle the segmentation can be expressed by the

counts of images for which both FP and FN are small (rows 4 and 5 of table 2). In these cases, AP and the proposed method exhibit the best performances. The rate of images for which both FP and FN are large (rows 8 and 9 of table 2) is smallest for the proposed method. When satisfying performaces regarding the two latter criteria are wanted, the proposed method is the best. This comparison tends to show that the proposed method globally achieves better performances than the previous methods.

4.2. Difficult cases and limitations

The main difficulties to overcome when dealing with pectoral muscle extraction are:
- presence of artifacts (skin folded),
- non uniform texture of the pectoral muscle region,
- overlapping tissues in the lower part of the pectoral muscle region.
In addition, the size and the shape of the pectoral muscle are different from one patient to another.
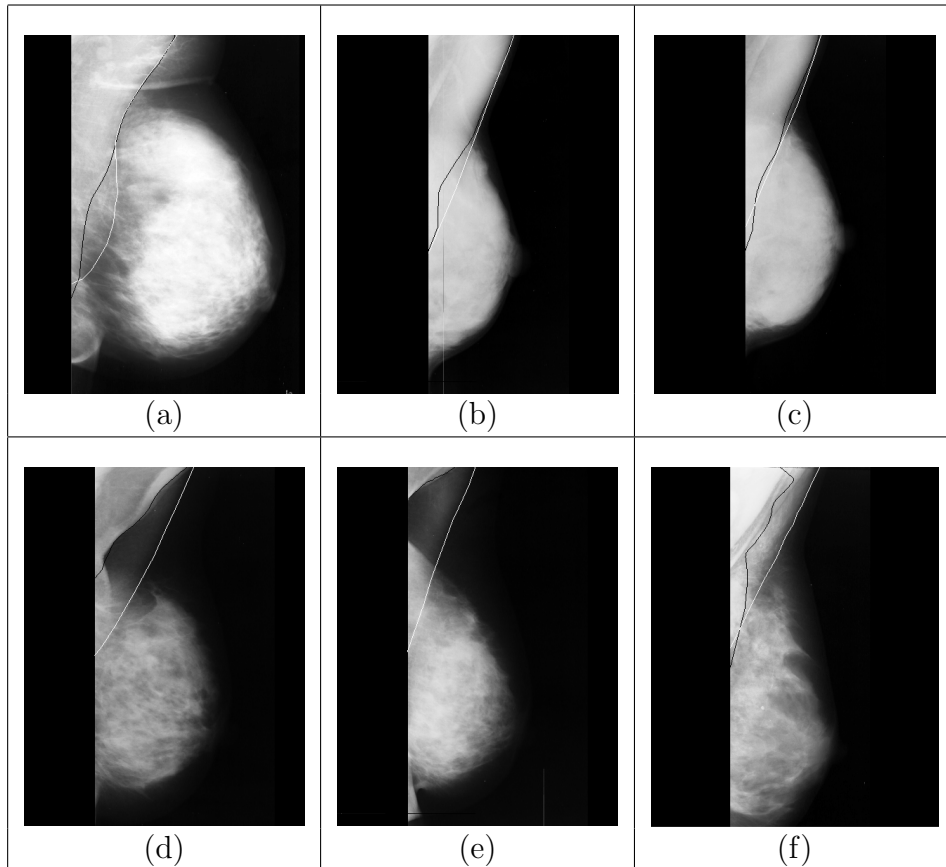
Figure 11. Some examples of pectoral muscle boundaries detected by the proposed method (black) and hand-drawn by the radiologist (white) superimposed on the mammograms. Cases with strong overlap of glandular tissue (a) mdb240, (b) mdb053, (c) mdb054 and cases with skin folded or non uniform texture (d) mdb066, (e) mdb065, (f) mdb034.

Figure 11 presents the outputs of our approach on very difficult cases. In some cases showing strong overlap with glandular tissue or non uniform contrast, the proposed method gives quite good pectoral mucle extraction results (see figure 11(a-c)). This is due to the fact that the proposed method takes advantage of contour validation to better localize the pectoral muscle region as it upper edge in such cases is mostly well estimated. The refinement of contour in these cases is effective only in the area where the overlap is not very strong or where the contrast is acceptable. However, in presence of severe artifacts like skin folded or poor contrast the algorithm fails to produce accurate results (see figure 11(d-f)). The main cause of this failure lies in the segmentation step. In these cases, the contour obtained after segmentation mainly follows the skin folded shape. Moreover, during the refinement step, the edge points found are those located on false edges since the latter show the nearest spike on the average gradient profile towards which the refinement process converges.

Although some significant deviations between the computed boundary and the actual one can be seen, the proposed solution is never completely incoherent. Note that these limitations are mostly observed in particular cases of images where mammograms were not performed under optimal conditions. Carelessness on patient's positioning during mammograms acquisition may result in blurred images with patterns like pectoral muscle partially obstructed or not well depicted (see figure 11(d-e)). In addition, the difficult cases represent less than 5% of processed images. The robustness of our approach is thus satifying. Better results on such difficult cases could be achieved by designing an additional post-processing or by imposing more stringent conditions into mammographic examination protocols. In such cases, a semi-supervised segmentation approach, in which some control points of pectoral muscle contour are defined by the operator, is likely to cope with this difficulty. This is an attempt to formalize the experts knowledge to segment these difficult cases thereby reducing the expert's workload considerably. This strategy will be investigated in further work.

5. Conclusion

Identification of pectoral muscle in MLO view of mammograms is a challenging task due to various patterns that make its automatic segmentation complicated. The main difficulties to overcome are the differences of shape, size and texture of the pectoral muscle region from one image to another, the variations in image quality and the overlapping between dense glandular tissue and pectoral muscle. In this work, we introduced an automated approach of pectoral muscle extraction using some basic a priori knowledge and gray level intensity information to delineate the pectoral boundary. The boundary of the pectoral muscle is segmentated using a modified clustering algorithm. Then a validation of the contour obtained from the segmentation is done to cope with overestimation of pectoral muscle region occuring in its lower part when its overlaps with dense glandular tissue. The final contour is obtained through iterative boundary refinement using average gradient. The performance of the proposed method

compared to a radiologist expert manual segmentation as reference shows that our method produces good results as compared to related works tested on the same database. The proposed method can be useful as a preprocessing step of applications dealing with mammogram analyses like registration, tissues characterization, and breast deformation modelling which are the key steps to build accurate CAD systems.

Acknowledgments

Appendix A. Distribution of images selected from the MIAS database

Table A1. Filenames of images rejected.

Amount : 45 images
mdb002; mdb010; mdb017; mdb024; mdb029; mdb061; mdb098; mdb137; mdb138; mdb146; mdb148; mdb151; mdb153; mdb154; mdb155; mdb158; mdb173; mdb179; mdb216; mdb224; mdb225; mdb236; mdb241; mdb252; mdb253; mdb254; mdb262; mdb269; mdb273; mdb283; mdb285; mdb287; mdb288; mdb289; mdb293; mdb295; mdb301; mdb304; mdb305; mdb313; mdb314; mdb315; mdb317; mdb321; mdb322.

References

Chakraborty J, Mukhopadhyay S, andNiranjan Khandelwal V S and Bhattacharyya P 2012 Automatic detection of pectoral muscle using average gradient and shape-based feature *J. Digit. Imaging* **25**, 387–99.

Domingues I, Cardoso J S, Amaral I, Moreira I, Passarinho P, Comba J S, Correia R and Cardoso M J 2010 Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs *in* 'Annual Int. Conf. on Engineering in Medicineand Biology Society' pp. 3158–61.

Eklund G W, Cardenosa G and Parsons W 1994 Assessing adequacy of mammographic image quality *Radiology* **190**, 297–307.

Ferrari R J, Rangayyan R M, Desautels J E L, Borges R A and Frère A F 2004 Automatic identification of the pectoral muscle in mammograms *IEEE Trans. Med. Imaging* **23**, 232–45.

Feudjio C, Tiedeu A, Gordan M, Domngang S and Vlaicu A 2012 Computerized detection and smoothing of contour in mammograms *in* '11th African Conference on Research in Computer Science and Applied Mathematics' pp. 299–306.

Table A2. Filenames of images used for inter-observer variability study.

Amount : 84 images
mdb003; mdb004; mdb007; mdb008; mdb033; mdb034; mdb035; mdb036; mdb037; mdb038; mdb039; mdb040; mdb041; mdb042; mdb043; mdb044; mdb045; mdb046; mdb047; mdb048; mdb049; mdb050; mdb051; mdb052; mdb055; mdb056; mdb067; mdb068; mdb071; mdb072; mdb073; mdb074; mdb075; mdb076; mdb077; mdb078; mdb081; mdb082; mdb083; mdb084; mdb085; mdb086; mdb087; mdb088; mdb089; mdb090; mdb091; mdb092; mdb095; mdb096; mdb097; mdb098; mdb099; mdb100; mdb101; mdb102; mdb103; mdb104; mdb105; mdb106; mdb107; mdb108; mdb109; mdb110; mdb111; mdb112; mdb113; mdb114; mdb115; mdb116; mdb117; mdb118; mdb119; mdb120; mdb121; mdb122; mdb123; mdb124; mdb125; mdb126; mdb127; mdb128; mdb129; mdb130.

Guo R C, Ye S S, Quan M and Shi H X 2009 Modified fast fuzzy C-means algorithm for image segmentation *in* '2nd Int. Symp. on Electronic Commerce and Security' Vol. 2 pp. 39–43.

Karnan M and Thangavel K 2007 Automatic detection of breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications *Comp. Method Prog. Biomed.* **87**, 12–20.

Karssemeijer N 1998 Automated classification of parenchymal patterns in mammograms *Phys. Med. and Biol.* **43**, 365–78.

Kwok S M, Chandrasekhar R, Attikiouzel Y and Rickard M T 2004 Automatic pectoral muscle segmentation on mediolateral oblique view mammograms *IEEE Trans. Med. Imaging* **23**, 1129–40.

Liu L, Wang J and Wang T 2011 Breast and pectoral muscle contours detection based on goodness of fit measure *in* '5th Int. Conf. on Bioinformatics and Biomedical Engineering' pp. 1–4.

Ma F, Bajger M, Slavotinek J P and Bottema M J 2007 Two graph theory based methods for identifying the pectoral muscle in mammograms *Pattern Recogn.* **40**, 2592–602.

Mirzaalian H, Ahmadzadeh M R and Sadri S 2007 Pectoral muscle segmentation on digital mammograms by nonlinear diffusion filtering *in* 'IEEE Pacific Rim Conference on Communications, Computers and Signal Processing' pp. 581–4.

Mustra M, Bozek J and Grgic M 2009 Nipple detection in craniocaudal digital mammograms *in* '51st Int. symp. ELMAR'09' pp. 15–8.

Raba D, Olivier A, Martí J, Peracaula M and Espunya J 2005 Breast segmentation with pectoral muscle suppression on digital mammograms *in* J Marques, N Pérez de la Blanca and P Pina, eds, 'Pattern Recognition and Image Analysis' Vol. 3523 of *Lecture Notes in Computer Science* Springer-Verlag, Berlin Heidelberg pp. 471–8.

Reddy M and Given-Wilson R 2006 Screening for breast cancer *Women's Health Medicine* **3**, 22–7.

Saha P K, Udupa J K, Conant E F, Chakraborty D P and Sullivan D 2001 Breast tissue density quantification via digitized mammograms *IEEE Trans. Med. Imaging* **20**, 792–803.

Suckling J, Dance D R, Moskovic E, Lewis D J and Blacker S G 1995 Segmentation of mammograms using multiple linked self-organizing neural networks *Med. Phys.* **22**, 145–52.

Suckling J, Parker J, Dance D R, Astley S, Hutt I, Boggis C R M, Ricketts I, Stamatakis E, Cerneaz N, Kok S L, Taylor P, Betal D and Savage J 1994 The mammographic images analysis society: Digital mammogram database *Experta Medica International Congress Series* **1069**, 375–8.

Wang P and Wang H 2008 A modified FCM algorithm for MRI brain image segmentation *in* 'Int. Seminar on Future BioMedical Information Engineering' pp. 26–29.

Table A3. Filenames of images selected for this study.

Amount: 277 images

mdb001; mdb003; mdb004; mdb005; mdb006; mdb007; mdb008; mdb009; mdb011; mdb012; mdb013; mdb014; mdb015; mdb016; mdb018; mdb019; mdb020; mdb021; mdb022; mdb023; mdb025; mdb026; mdb027; mdb028; mdb030; mdb031; mdb032; mdb033; mdb034; mdb035; mdb036; mdb037; mdb038; mdb039; mdb040; mdb041; mdb042; mdb043; mdb044; mdb045; mdb046; mdb047; mdb048; mdb049; mdb050; mdb051; mdb052; mdb053; mdb054; mdb055; mdb056; mdb057; mdb058; mdb059; mdb060; mdb062; mdb063; mdb064; mdb065; mdb066; mdb067; mdb068; mdb069; mdb070; mdb071; mdb072; mdb073; mdb074; mdb075; mdb076; mdb077; mdb078; mdb079; mdb080; mdb081; mdb082; mdb083; mdb084; mdb085; mdb086; mdb087; mdb088; mdb089; mdb090; mdb091; mdb092; mdb093; mdb094; mdb095; mdb096; mdb097; mdb099; mdb100; mdb101; mdb102; mdb103; mdb104; mdb105; mdb106; mdb107; mdb108; mdb109; mdb110; mdb111; mdb112; mdb113; mdb114; mdb115; mdb116; mdb117; mdb118; mdb119; mdb120; mdb121; mdb122; mdb123; mdb124; mdb125; mdb126; mdb127; mdb128; mdb129; mdb130; mdb131; mdb132; mdb133; mdb134; mdb135; mdb136; mdb139; mdb140; mdb141; mdb142; mdb143; mdb144; mdb145; mdb147; mdb149; mdb150; mdb152; mdb156; mdb157; mdb159; mdb160; mdb161; mdb162; mdb163; mdb164; mdb165; mdb166; mdb167; mdb168; mdb169; mdb170; mdb171; mdb172; mdb174; mdb175; mdb176; mdb177; mdb178; mdb180; mdb181; mdb182; mdb183; mdb184; mdb185; mdb186; mdb187; mdb188; mdb189; mdb190; mdb191; mdb192; mdb193; mdb194; mdb195; mdb196; mdb197; mdb198; mdb199; mdb200; mdb201; mdb202; mdb203; mdb204; mdb205; mdb206; mdb207; mdb208; mdb209; mdb210; mdb211; mdb212; mdb213; mdb214; mdb215; mdb217; mdb218; mdb219; mdb220; mdb221; mdb222; mdb223; mdb226; mdb227; mdb228; mdb229; mdb230; mdb231; mdb232; mdb233; mdb234; mdb235; mdb237; mdb238; mdb239; mdb240; mdb242; mdb243; mdb244; mdb245; mdb246; mdb247; mdb248; mdb249; mdb250; mdb251; mdb255; mdb256; mdb257; mdb258; mdb259; mdb260; mdb261; mdb263; mdb264; mdb265; mdb266; mdb267; mdb268; mdb270; mdb271; mdb272; mdb274; mdb275; mdb276; mdb277; mdb278; mdb279; mdb280; mdb281; mdb282; mdb284; mdb286; mdb290; mdb291; mdb292; mdb294; mdb296; mdb297; mdb298; mdb299; mdb300; mdb302; mdb303; mdb306; mdb307; mdb308; mdb309; mdb310; mdb311; mdb312; mdb316; mdb318; mdb319; mdb320.

Zhou C, Wei J, Chan H P, Paramagul C, Hadjiisaki L M, Sahiner B and Douglas J A 2010 Computerized image analysis: Texture-field orientation method for pectoral muscle identification on MLO-view mammograms *Med. Phys.* **37**, 2289–99.