Pectoral Muscle removal and Detection of masses in Digital Mammogram using CCL

T.S.Subashini
Annamalai University
Department of Computer Science & Engineering
Faculty of Engineering & Technology

V. Ramalingam
Annamalai University
Department of Computer Science & Engineering
Faculty of Engineering & Technology

S. Palanivel
Annamalai University
Department of Computer Science & Engineering
Faculty of Engineering & Technology

ABSTRACT

A mammogram is a radiograph of the breast tissue. It is an effective non-invasive means of examining the breast, commonly searching for breast cancer. Cancer is not preventable, but early detection leads to a much higher chance of recovery and lowers the mortality rate. Due to the high volume of images to be analyzed by radiologists, and since senior radiologists are rare, the accuracy rate tends to decrease. This is reflected in the high percentage of unnecessary biopsies that are performed and many deaths caused by late detection or poor diagnosis. This paper proposes a computer aided diagnosis system for detecting masses in mammograms using connected component labeling(CCL). This paper also addresses the problem of eliminating and pectoral muscles from the mammogram before the detection process so that further processing is confined to the breast region alone.

Categories and Subject Descriptors

I.4.3 [IMAGE PROCESSING AND COMPUTER VISION]: Enhancement- filtering, gray scale manipulation, smoothing.

I.4.6 [IMAGE PROCESSING AND COMPUTER VISION]: Segmentation – Edge and feature detection

General Terms

Algorithms

Keywords

Digital mammogram, morphological reconstruction, blur filter, connected component labeling, segmentation, mass detection.

INTRODUCTION

When viewing simple objects on a simple background of noise, humans perform nearly as well as 'ideal observers'. However, humans are limited in their ability to detect and diagnose disease during image interpretation due to their non-systematic search patterns and to the presence of structure noise. In addition, the vast amount of image data that is generated by some imaging devices makes the detection of potential disease a burdensome task and may cause oversight errors. Another

problem is that the similar characteristics of some abnormal and normal lesions may cause interpretational errors. Developments in computer vision and artificial intelligence in medical image interpretation have shown the potential for computers as providers of a 'second opinion' in image interpretation. Reading mammograms is a challenge for Radiologists. Diagnosis is truly subject to interpretation. Hence the concept of a 'second reader' is emerging in recent years. Although mammography is currently the best method for the detection of breast cancer, between 10% and 30% of women who have breast cancer and undergo mammography have negative mammograms.\cite In approximately two-thirds of these false-negative mammograms, the radiologist failed to detect the cancer that was evident retrospectively. Missed detections may be due to the subtle nature of the radiographic findings, poor image quality or eye fatigue or oversight by radiologists. In addition, it has been demonstrated that reading by two radiologists can increase sensitivity. Therefore, one aim of CAD is to increase the efficiency and effectiveness of screening procedures by using a computer system as a 'second reader', like a 'spell checker', to aid the radiologist by indicating locations of suspicious abnormalities in mammograms [18].

1. PREVIOUS WORK

A number of image processing methods are seen in the literature to isolate the abnormal regions in digital mammograms. [12],[17] employs a segmentation method based on the fuzzy sets theory to divide a mammogram into different regions and produces regions of mass candidates and then discrete texture features are calculated for the area of each mass candidate. A number of methods have already been used to detect circumscribed masses morphological operations and gray level thresholding [14],[11]. In [16] after enhancing the image the features are extracted and k-means clustering algorithm is applied for mass classification. Difference of Gaussians and derivative based feature saliency is employed in [13] and relative image intensity was used in [15] for the detection of masses. An iris filter is used to detect tumors in [8]. Some studies have justified the use of model based image processing techniques such as Markov random field and statistical model is used in [9] for abnormal area detection. Morphological band pass filters is used for detecting the regions of interest in [6]. Watershed Segmentation algorithm is applied in [5]) for detecting masses in digital mammograms.

In this paper image segmentation, filtering, contrast improvement and gray level thresholding techniques are applied for enhancing the whole image. Then connected component labeling and decision rules are applied to isolate the region of interest.

2. DATA SOURCE

The proposed work was done using Mini-Mias database. The Mammography Image Analysis Society (MIAS) [7], which is an organization of UK research groups interested in the understanding of mammograms, has produced a digital mammography database. The X-ray films in the database have been carefully selected from the United Kingdom National Breast Screening Program and digitized with a Joyce - Lobel scanning micro densitometer to a resolution of 50 μm × 50 μm, 8 bits represent each pixel. It has been reduced to 200 micron pixel edge and clipped/padded so that every image is 1024 × 1024 pixels.

3. PROPOSED WORK

The proposed method contains the following modules depicted in figure 1. Before carrying out feature extraction, mammogram must undergo preprocessing to remove the artifacts and to smoothen the image. The second module performs binarization to create a binary image. Binary image undergoes edge detection and segmentation to remove unwanted background region, which will reduce the processing time in subsequent image analysis. Then Gray level manipulation is performed to improve the contrast of suspected masses so that mass detection can be done effectively. This is followed by global thresholding and connected component labeling to detect masses. This section explains the various modules in detail.

```
┌─────────────────────────────┐
│ Preprocessing               │
│ Artifact Removal,Blur       │
│ Filter                      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Binerization and Edge       │
│ detection                   │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Detection and removal of    │
│ Pectoral Muscles            │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Gray Level                  │
│ Manipulation                │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Mass Detection Connected    │
│ component labeling          │
└─────────────────────────────┘
              │
              ▼
          **Tumor**
```
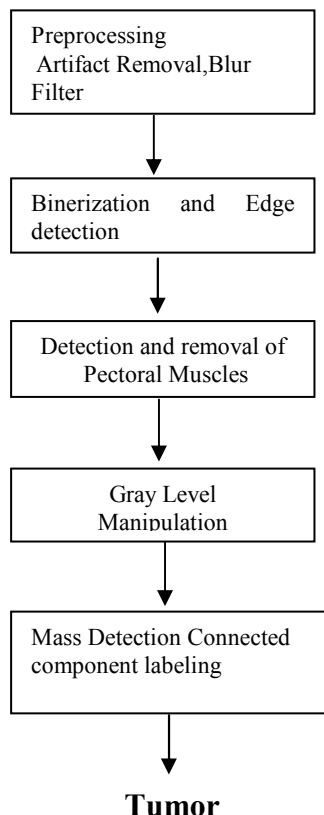
Figure 1. The block diagram of the proposed method

4. Preprocessing

Mammograms are difficult images to interpret, and a preprocessing phase is necessary to improve the quality of the images and make the feature extraction phase more reliable. In order to limit the search for abnormalities by computer aided diagnosis systems to the region of the breast without undue influence from the background of the mammogram, removal of artifacts and removal of pectoral muscle is necessary. Preprocessing stage consists of two parts. The first part involves the removal of unwanted parts from the image and the second part deals with reducing the high frequency components present in the image. Artifacts are removed by morphological open operation followed by reconstruction operation [4]. Blur filtering reduces high frequency components and the resulting image is smoothened [16].

4.2 Breast Contour Detection

Breast contour detection and pectoral muscle removal helps to confine further analysis to the breast region alone which otherwise could bias the detection procedures in consequent stages. According to mammogram images, the breast image is bright in the middle of the tissue and gradually becomes darker towards the skin air interface. So a binary image was created choosing a proper initial threshold level thus segmenting a large area of the breast region. Binerization is the process of dividing the image into two sub images or regions. Here the background region is made completely black (i.e. all the pixels with gray level between 0 and 15 are assigned 0) and the breast region is made completely white (i.e. all the pixels with gray level between 16 and 255 are assigned 255).

Edge can be detected by finding the intensity discontinuities. The pixels, which are lying between the black and the white regions, are extracted and linked. Raster scanning is done from the right corner to left corner on each row of the binary image and the pixels which are lying between the black and white regions are detected by intensity discontinuities and are assigned a grey value of 255. This edge image is checked for ascertaining that the breast region alone will be segmented to avoid unnecessary computational overheads. Almost 50% of the whole mammogram image comprises of a noisy background. This background region must be segmented to eliminate unwanted portion of the image. All the pixels, which are lying on the left side of the edge, are replaced with the original pixel value and rest of the image is made completely white.

4.3 Pectoral Muscle Detection and Removal

Pectoral muscles are the regions in mammograms that contain brightest pixels. These regions must be removed before detecting the tumor cells so that mass detection can be done efficiently. Pectoral muscles lie on the left or right top corner depending on the view of the image. We must detect the position of the pectoral muscles (left top corner or right top corner) before removing it. For this searching for nonzero pixels are simultaneously done from the left and right top corner. Width of the image in which the non zero pixel detected from both the corner were counted and compared. If the left width is smaller than the right width

then it is assumed that pectoral is on the left side of the image else it is on the right side.

From the detected corner pixel the intensity discontinuity is detected on each and every column of the same row. Coordinates of the pixel in which the intensity change is encountered is considered as width of the pectoral region. All the pixels, which lie inside pectoral width and half of the height of the whole image is segmented from the original image. This rectangle shaped image contains the entire pectoral muscles.

To extract the pectoral muscles from this image binary image should be obtained by simple thresholding. This binary image contains pectoral muscles and other tissues. To segment the pectoral muscles alone from the binary image raster scanning is done from the right or left side of the image to detect the intensity discontinuities.

The resulting image contains pectoral muscles alone and this region is completely removed from the original image.

4.4 Gray Level Manipulation

In this stage regions of interest are enhanced and the unwanted regions of the image are deemphasized. The enhancement procedure results in a better description of the objects of interest, thus improving the sensitivity of the detection system and leading to better classification of the abnormalities in the case of diagnosis. The enhancement of the contrast of the regions of interest and the suppression of noise is performed in this stage. Median filtering and gray level transformations are done in this work in order to enhance the regions of interest. Manipulation of all the gray level in image is done to enhance the suspected masses or the cancer cells present in the mammogram. The manipulation function [16] is given by equation (1)

$$P_1(i, j) = round (P(i, j)^3 / f * M^2) \qquad (1)$$

Where $P_1(i,j)$ is the gray level of manipulated pixel. $P(i,j)$ is the gray level of corresponding pixel in the original image. M is the maximum gray level in the image. Global thresholding is performed to find the exact region of interest form the manipulated image. f is the manipulation factor and it varies between 0.6 and 1 according to the maximum gray level of the image. If the maximum gray level is above 235, transformation factor is assigned 1 and with every 10 gray level decrease from the maximum gray level, transformation factor is reduced in a step of 0.1 till 0.6. These values are found out empirically.

4.5 Mass Detection

Region of interest in mammograms are the masses and these masses are extracted by global thresholding of the transformed image. Since masses are brightest region in the mammogram all the brightest pixels in the image are extracted till the number of pixels in the extracted image is greater than the 5% of nonzero pixels in the segmented image. For this all the pixels with maximum gray level (M) in the image (C) are counted. If it

is less than 5% of the total number of non zero pixels in the original image then the number of pixels with gray value M-1 are counted and checked for 5% pixels in original image and this process continues till C is greater than 5%. Resulting image contains tumour cells and unwanted noises. This image is processed using connected component modelling technique. Connected components labelling scans an image and groups its pixels into components based on pixel connectivity, *i.e.* all pixels in a connected component share similar pixel intensity values and are in some way connected with each other. Once all groups have been determined, each pixel is labelled with a gray level according to the component it was assigned to. The largest component in the image is extracted. The resulting image contains tumour image. Results of the proposed work are shown in figures 2, 3 and figure 4.

5. CONCLUSION

This work has been implemented and tested using Microsoft Visual C++ 6.0 and OpenCV Image Processing Library. The proposed work was done using Mini-Mias mammogram database. Here we have presented several aspects of image processing techniques that can be applied for detection of masses in digital mammography. In this paper, we have considered the problem of detecting cancer masses by the application of simple thresholding followed by connected component labeling and an algorithm to remove artifacts in digital mammograms using morphological open operation followed by reconstruction. Further the pectoral muscle was removed successfully using simple thresholding and raster scan methods.

6. ACKNOWLEDGMENTS

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

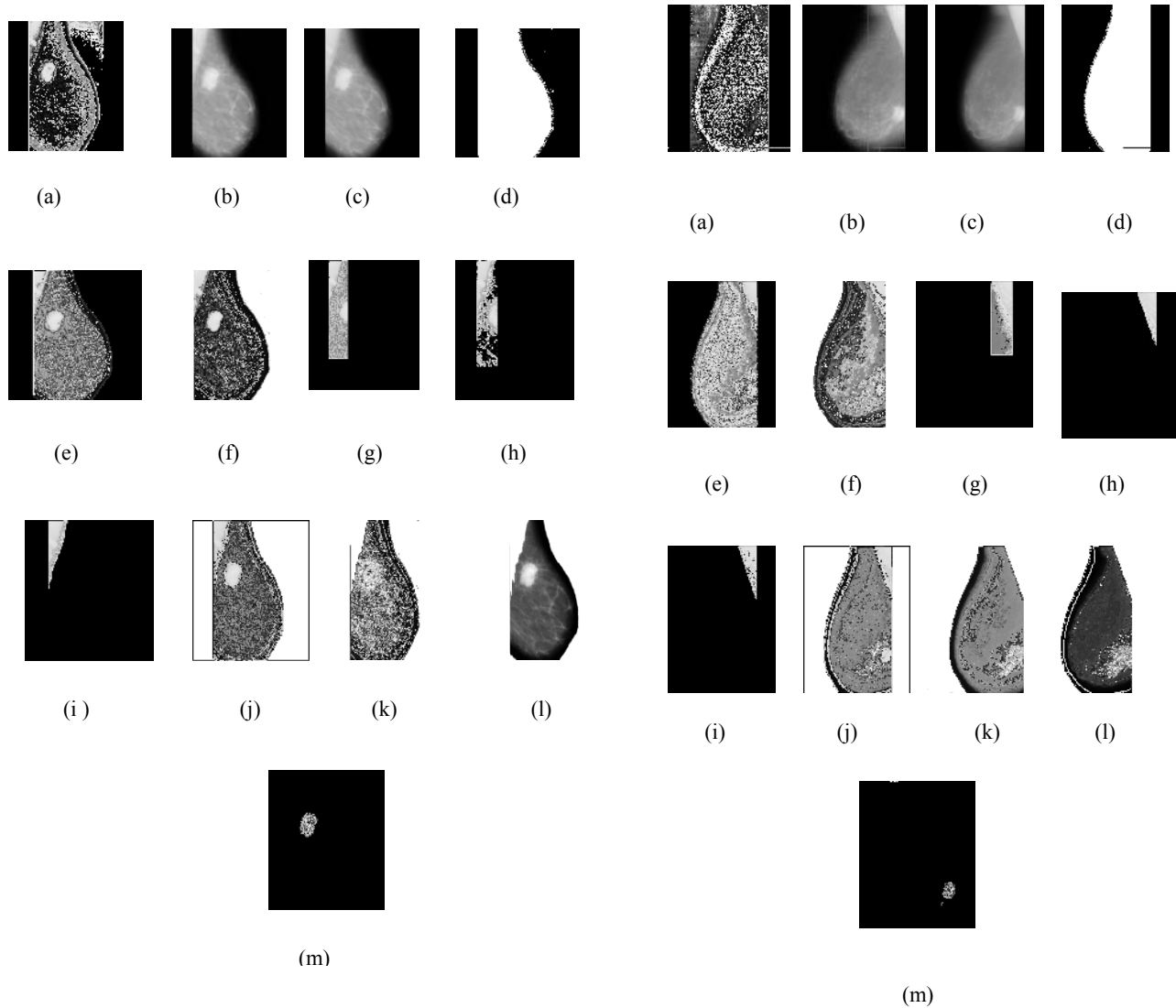(i)   (j)   (k)   (l)

(m)

Figure 1

(a)Input Image (b) Artifact Removed Image (c) Blurred Image
(d) Binary Image (e) Edge detected Image (f) Segmented Image
(g) pectoral suspected region (h)Pixels with 70% of maximum
gray value (i) Pectoral Image (j) Pectoral detected Image (k)
Pectoral removed Image (l) Transformed Image (m) Tumor Image

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

(i)   (j)   (k)   (l)

(m)

Figure2

(a) Input Image (b) Artifact Removed Image (c) Blurred Image
(d) Binary Image   (e) Edge detected Image (f) Segmented
Image   (g)suspected pectoral region (h) Pixels with 70% of
maximum gray value (i) Pectoral   Image (j) Pectoral detected
Image (k) Pectoral removed Image (l) Transformed Image (m)
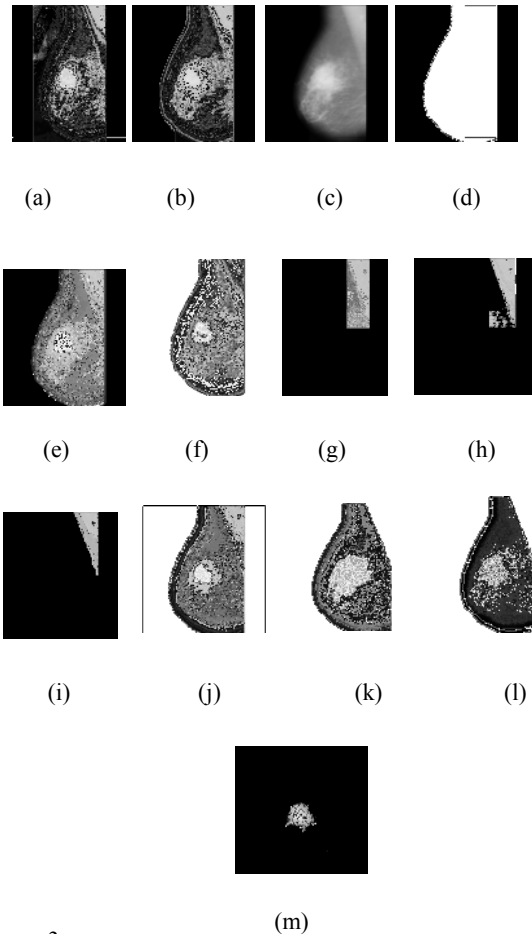Tumor Image

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

(i)   (j)   (k)   (l)

(m)

Figure 3

(a) Input Image (b) Artifact Removed Image (c) Blurred Image (d) Binary Image (e) Edge detected Image (f) Segmented Image (g)suspected pectoral region (h) Pixels with 70% of maximum gray value (i) Pectoral   Image (j) Pectoral detected Image (k) Pectoral removed Image (l) Transformed Image (m) Tumor Image

7. REFERENCES

[1]   Giger ML. 2000. Computer-aided diagnosis in mammography. Handbook of Medical Imaging, Medical Imaging Processing and Analysis. SPIE, 2(2000), 915-1004.

[2]   Maysam Shahedi B K,Rassoul Amirfattahi, Farah TorkamaniAzar, Saeed Sadri. Accurate Breast Region Detection In Digital Mammograms Using A Local   Adaptive Thresholding Method. Eighth International Workshop On Image Analysis For Multimedia Interactive Services(Wiamis'07), 2006.

[3]   Vincent L.   1993.   Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. IEEE Trans in Image Processing. 2(1993), 176-201.

[4]   Michael A. Wirth and Dennis Nikitenko. 2005. Suppression of Stripe Artifacts in Mammograms Using Weighted Median Filtering. ICIAR (2005), 966-973.

[5]   Thor Ole Gulsrud et al., 2005. Wareshed Segmentation of Detected masses in digital mammograms. IEEE, Engineering in medicine and 27th annual conference. 3304-3307,

[6]   Chang R. et. al., 2003. Detecting ROI using Morphological Bandpass Filter. IEEE Transaction on Information Technology in Biomedicine. 7-3(2003), 197-201

[7]   Suckling J. et al., 1994. The Mammographic Image Analysis Society digital mammogram database. 2nd International Workshop on Digital Mammography, York, England. 375-378.

[8]   Kobatake H. et al., 1999. Computerised detection of malignant tumors on digital mammograms. IEEE Transactions on Medical Imaging.. 18(1999), 378-396.

[9]   Zwiggelaar R. et al., 1999. Model based detection of spiculated lesions on digital mammograms. Medical Image Analysis. 3 -1(1999), 39-62.

[10] H.D. Li et al., 1995. Markov random field for tumor detection in digital mammography. IEEE Transactions on Medical Imaging. 14 (1995), 565-576.

[11] R. Mousa et al., 2005. Breast cancer diagnosis system based on wavelet analysis and fuzzy neural. Expert Systems with applications 28(2005), 713-723.

[12] N.R. Mudigonda et al., 2000. Gradient & Texture analysis for the classification of Mammographic Masses. IEEE Transactions on Medical Imaging. 10(200),1032-1042.

[13] W.E. Polakowski et al., 1997. Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative based feature saliency. IEEE Transactions on Medical Imaging. 16(1997), 811-819.

[14] R.P. Velthuizen. 2000. Computer Description of Mammographic Masses. Proceedings of the 5th International workshop on Digital Mammography.

[15] M.D. Heath and K.W. Bowyer. 2000. Computer Description of Mammographic Masses. Proceedings of 5th International Workshop on Digital Mammography.(2000)

[16] Maria Victoria et al., 2006.   Circumscribed Mass Detection in Digital Mammograms., Proceedings of the Electronics, Robotics and Automotive Mechanics Conference

[17] M. Sameti et al., 1997. Texture feature extraction for tumor detection in mammographic images. Pacific Rim Conference on Communications, Computers and Signal Processing. 3(1997), 831-834.

[18] Zhimin Huo, Maryellen L. Giger, Carl J. Vyborny, Charles E. Metz. 2002. Breast Cancer: Effectiveness of Computer-aided Diagnosis—Observer Study with Independent Database of Mammograms. *Radiology, 224(2002), 560-56.*