Dense Tissue Segmentation in Digitized Mammograms

Mario Mustra, Mislav Grgic
University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia
mario.mustra@fer.hr

Abstract - Determining the breast density in mammograms is important both in diagnostic and computer-aided detection applications. Knowing the right breast density and having knowledge of changes in breast density could give a hint of a process which started to happen within a patient. Breast density could be rather easily estimated by dividing mammogram into fibroglandular and fat tissue. Mammograms suffer from a problem of overlapping tissue which results in possibility of inaccurate detection of tissue types. Fibroglandular tissue has rather high attenuation of X-rays and is visible as brighter in the resulting image. Small blood vessels and microcalcifications are shown as brighter objects with similar intensities as dense tissue. In this paper we try to divide dense and fat tissue by suppressing scattered structures which do not represent glandular or dense tissue in order to divide mammograms more accurately in two major tissue types. For suppressing blood vessels we have used Gabor filters of different size and orientation to detect edges of blood vessels and subtract them from the original image. Microcalcifications have been suppressed by combination of morphological operations on filtered image with enhanced contrast. Dense tissue has been segmented using different thresholds to avoid false detection.

Keywords - Gabor Filter; Breast Density; CLAHE; Morphology

I. Introduction

Computer-aided diagnosis (CAD) systems are being developed more and more each day. Their aim is to help radiologists especially in screening examinations when a large number of patients are examined and radiologist often spend very short time for readings of non-critical mammograms. Mammograms are X-ray breast images of usually high resolution with moderate to high bit-depth which makes them suitable for capturing fine details. Because of the large number of captured details, computer-aided detection (CADe) systems have difficulties in detection of desired microcalcifications and lesions in the image. Since mammograms are projection images in grayscale, it is difficult to automatically differentiate types of breast tissue because different tissue types can have same or very similar intensity. The problem which occurs in different mammograms is that the same tissue type is shown with a different intensity and therefore it is almost impossible to perform histogram based thresholding. To overcome that problem different authors have came up with different solutions. Some used statistical feature extraction and classification of mammograms into different categories according to their density. Other approaches used filtering of images and then extracting features from filter response images. Among methods which use extraction of statistical features, Oliver et al. [1] obtained very good results by using combination of statistical features extracted not directly from the image, but from gray level co-occurrence matrices. Images were later classified into four different density categories, according to BI-RADS [2]. Authors who used image filtering techniques tried to divide, as precisely as possible, breast tissue into two main types: dense and fat tissue. With the accurate division of dense and fat tissue in breasts it would be possible to quantify the results of breast density classification and classification itself would become trivial. However, the task of defining the appropriate threshold for dividing breast tissue into two categories is far from simple. Each different mammogram captured using the same mammography device is being captured with slightly different parameters which will affect the final intensity of the image. These parameters are also influenced by the physical property of each different breast. In image acquisition process the main objective is to produce an image with very good contrast and no clipping in both low and high intensity region. Different mammograms will therefore suffer from different intensities for the corresponding tissue. Reasons for that, if we neglect usage of different imaging equipment, are difference in the actual breast size, difference in the compressing force applied to the breast during capturing process, different exposure time and anode current. Having this in mind authors tried to overcome this problem by applying different techniques which should minimize influence of capturing inconsistencies. Muhimmah and Zwiggelaar [3] presented an approach of multiscale histogram analysis having in mind that image resizing will affect the histogram shape because of detail removal when image is being downsized. In this way they were able to remove small bright objects from images and tried to get satisfactory results by determining which objects correspond to large tissue areas. Petroudi et al. [4] used Maximum Response 8 filters [5] to obtain a texton dictionary which was used in conjunction with the support-vector machine classifier to classify breast into four density categories. Different equipment for capturing mammograms produces resulting images which have very different properties. The most common division is in two main categories: SFM (Screen Film Mammography) and FFDM (Full-Field Digital Mammography). Tortajada et al. [6] have presented a work in which they try to compare accuracy of the same classification method on SFM and FFDM images. Results which they have obtained show that there is high correlation of automatic classification and expert readings and overall results are slightly better for FFDM images.

In this paper we present a method which should provide a possibility for division breast tissue between parenchymal tissue and fatty tissue without influence of blood vessels and fine textural objects which surround fibroglandular disc. Segmentation of dense or glandular tissue from the entire tissue will be made by setting different thresholds. Our goal is to remove tissue which interferes with dense tissue and makes the division less accurate because non-dense tissue is being treated as dense due to its high intensity when compared with the rest of the tissue. Gabor filters generally proved to be efficient in extracting features for breast cancer detection from mammograms because of their sensitivity to edges in different orientations [7]. Therefore, for the removal of blood vessels, we have used Gabor filter bank which is sensitive to brighter objects which are rather narrow or have high spatial frequency. Output of the entire filter bank is an image which is created of superimposed filter responses from different orientations. Subtraction of the image which represents vessels and different tissue boundaries from the original image produces a much cleaner image which can later be enhanced in order to equalize intensity levels of corresponding tissue types among different images. In that way we will be able to distinct dense tissue from fat more accurately. The proposed method has been tested on mammograms from the mini-MIAS database [8].

This paper is organized as follows. In Section II we present the idea behind image filtering using Gabor filter bank and explain which setup we will choose for filtering blood vessels and smaller objects out. In Section III we present results of filtering with the appropriate filter and discuss results of region growing after contrast enhancement and application of morphological operations. Section IV draws the conclusions.

II. Gabor Filters

Gabor filters are linear filters which are most commonly used for edge detection purposes as well as textural feature extraction. Each filter can be differently constructed and it can vary in frequency, orientation and scale. Because of that Gabor filters provide a good flexibility and orientation invariantism. Gabor filter in a complex notation can be expressed as:

$$G = \exp\left(-\frac{(x\cos\theta + y\sin\theta)^2 + \gamma^2(y\cos\theta - x\sin\theta)^2}{2\sigma^2}\right) \cdot \exp\left(i\left(\frac{2\pi(x\cos\theta + y\sin\theta)}{\lambda} + \psi\right)\right), \quad (1)$$

where θ is the orientation of the filter, γ is the spatial aspect ratio, λ is the wavelength of the sinusoidal factor, σ is the sigma or width of the Gaussian envelope and ψ is the phase offset. This gives a good possibility to create different filter shapes which are sensitive to different objects in images. To be able to cover all possible blood vessels and small linearly shaped objects it is necessary to use more than one orientation. In our experiment we have used 8 different orientations and therefore obtained angle resolution of 22.5°. Figure 1 (a)-(h) shows 8 different filter orientations created using (1) with the angle resolution of 22.5° between each filter, from 0 to 157.5° respectively.

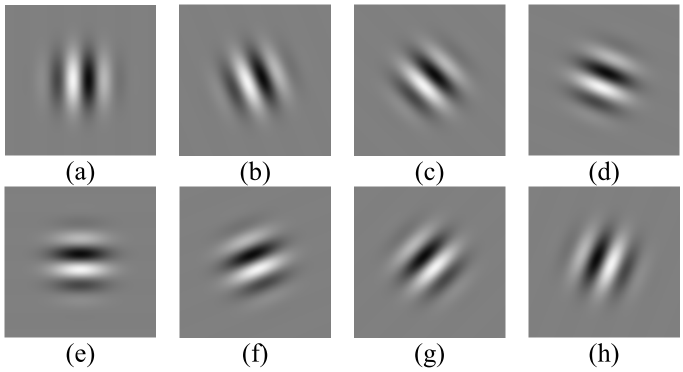


Figure 1. Gabor filters of the same scale and wavelength with different orientations: (a) 0°; (b) 22.5°; (c) 45°; (d) 67.5°; (c) 90°; (c) 112.5°; (c) 135°; (c) 157.5°.

Besides the orientation angle, one of the most commonly changed variables in (1) is the sinusoidal frequency. Usage of different sinusoidal frequencies will provide different sensitivity of the used filter for different spatial frequencies of objects in images. If the chosen filter contains more wavelengths, filtered image will correspond more to the original image because filters will be sensitive to objects of a high spatial frequency, e.g. details. In the case of smaller number of wavelengths, filtered image will contain highly visible edges. Figure 2 shows different wavelengths of Gabor filter with the same orientation.

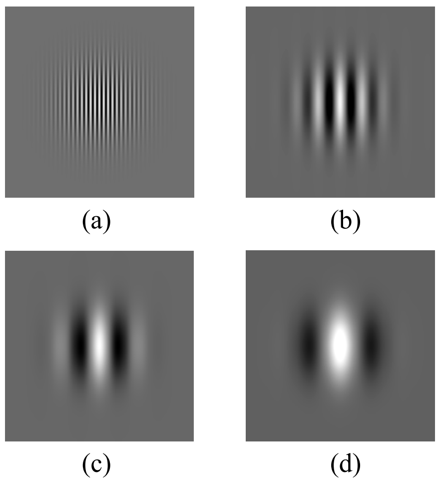


Figure 2. (a)-(d) Gabor filters which contains larger to smaller number of sinusoidal wavelengths respectively.

There is of course another aspect of the filter which needs to be observed and that is the actual dimension of the filter. Dimension of the filter should be chosen carefully according with the image size and the size of object which we want to filter out.

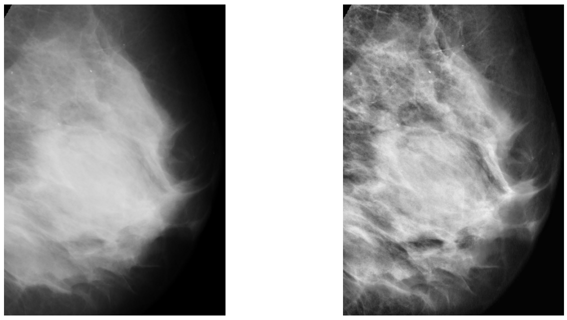
III. IMAGE FILTERING

Preprocessing of images is the first step which needs to be performed before filtering. Preprocessing steps include image registration, background suppression with the removal of artifacts and pectoral muscle removal. For this step we have used manually segmented masks drawn for all images in mini-MIAS database. These masks were hand-drawn by an experienced radiologist and, because of their accuracy, can be treated as ground truth. The entire automatic mask extraction process has been described in [9]. After the preprocessing we proceed with locating the fibroglandular disc position in each breast image. Fibroglandular disc is a region containing mainly two tissue types, dense or glandular and fat and according to their distribution it is possible to determine in which category according to density some breast belong. Dense tissue mainly has higher intensity in mammograms because it presents higher attenuation for X-rays than fat tissue. Intensity also changes with the relative position towards edge of the breast because of the change in thickness. Since the fibroglandular disc is our region of interest, we have extracted only that part of the image. Entire preprocessing step done for all images in the mini-MIAS database is described in [10]. Actual ROI boundaries are chosen to be V and H for vertical and horizontal coordinates respectively:

$$V = \left[\frac{\max(horizontal)}{2} : \frac{\max(horizontal)}{2} + \max(horizontal) \right]$$
$$H = \left[\frac{\max(vertical)}{3} : \max(vertical) \right] \tag{2}$$

where max(*horizontal*) is the vertical coordinate of the maximal horizontal dimension, and max(*vertical*) is the horizontal coordinate of the maximal vertical dimension. This approach gives a good isolation of the fibroglandular disc area with no need for the exact segmentation of it. It would be good if we could eliminate fibrous tissue and blood vessels and treat our ROI as it is completely uniform in the case of low density breasts. To be able to perform that task we can choose an appropriate Gabor filter sensitive to objects that we want to remove. A good Gabor filter for detection of objects with high spatial frequency contains less sinusoidal wavelengths, like the ones showed in Fig. 2 (c) and (d).

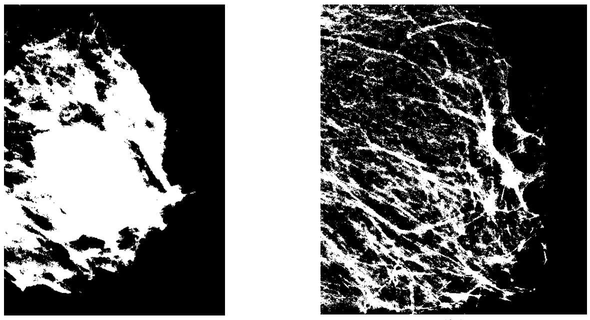
Contrast Limited Adaptive Histogram Equalization (CLAHE) [11] is a method for local contrast enhancement which is suitable for equalization of intensities in each ROI that we observe. Contrast enhancement obtained using CLAHE method will provide better intensity difference between dense and fat tissue. If we observe the same ROI before and after applying CLAHE enhancement it is clear that fat tissue can be filtered out easier after contrast enhancement. Figure 3 shows application of contrast enhancement using CLAHE on "mdb001".



(a) (b)

Figure 3. (a) Original ROI from "mdb001"; (b) Same ROI after contrast enhancement using CLAHE.

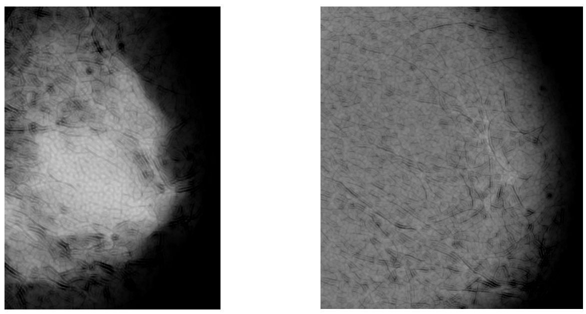
If we apply threshold on the enhanced ROI we will get the result for "mdb001" and "mdb006" as shown in Fig. 4 (a) and (b). These two images belong to opposite categories according to the amount of dense tissue. The applied threshold is set to 60% of the mean image intensity.



(a) (b)

Figure 4. (a) Enhanced ROI from "mdb001" with applied threshold; (b) Enhanced ROI from "mdb006" with applied threshold.

After applying threshold images have visibly different properties according to the tissue type. It is not possible to apply the same threshold because different tissue type has different intensity. Contrast enhancement makes the detection of fibrous tissue and vessels easier especially after Gabor filtering, Fig. 5 (a) and (b).



(a) (b)

Figure 5. (a) ROI from "mdb001" after applying Gabor filter; (b) ROI from "mdb006" after applying Gabor filter.

After contrast enhancement and filtering images using Gabor filter to remove fibrous tissue we need to make a decision in which category according to density each breast belongs. For that we will use binary logic with different threshold applied to images. We will apply two thresholds, at 60% and 80% of the maximal intensity and calculate the area contained in both situations. For that we will use logical AND operator, Fig. 6.

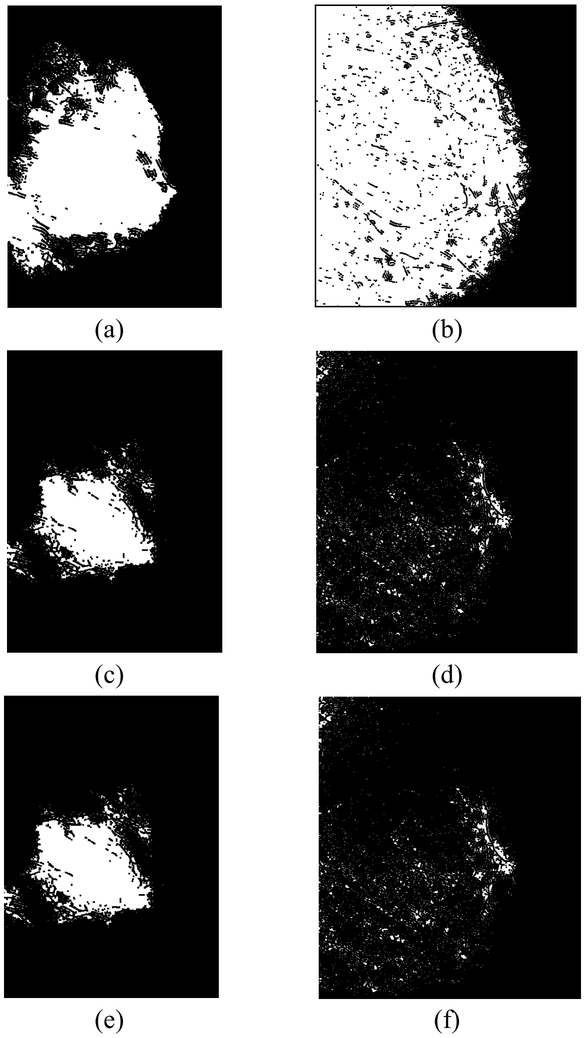


Figure 6. (a) "mdb001" after filtering out blood vessels and threshold at 60% of maximal intensity; (b) "mdb006" after filtering out blood vessels and threshold at 60% of maximal intensity; (c) "mdb001", threshold at 80%; (d) "mdb006", threshold at 80%; (e) "mdb001" threshold at 60% AND threshold at 80%; (f) "mdb006" threshold at 60% AND threshold at 80%.

IV. CONCLUSION

In this paper we have presented a method which combines usage of local contrast enhancement using CLAHE method and Gabor filters for removal of blood vessels and smaller portions of fibrous tissue. Combination of different thresholds in conjunction with logical AND operator provides a setup for determining whether we have segmented a fat or dense tissue. The advantage of Gabor filter over classical edge detectors is in easy orientation changing and possibility to cover all possible orientations by superpositioning filter responses. Usage of Gabor filter improves number of false positive results which come from blood vessels or small fibrous tissue segments and contrast enhancement provides comparability of the same tissue type in different mammograms. Our future work in this field will be development of automatic segmentation algorithms for dense tissue in order to achieve quantitative breast density classification by knowing the exact amount of dense tissue.

ACKNOWLEDGMENT

The work described in this paper was conducted under the research project "Intelligent Image Features Extraction in Knowledge Discovery Systems" (036-0982560-1643), supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

REFERENCES

[1] A. Oliver, J. Freixenet, R. Martí, J. Pont, E. Pérez, E.R.E. Denton, R. Zwiggelaar, "A Novel Breast Tissue Density Classification Methodology", IEEE Transactions on Information Technology in Biomedicine, Vol. 12, Issue 1, January 2008, pp. 55-65.

[2] American College of Radiology, "Illustrated Breast Imaging Reporting and Data System (BI-RADS) Atlas", American College of Radiology, Third Edition, 1998.

[3] I. Muhimmah, R. Zwiggelaar, "Mammographic Density Classification using Multiresolution Histogram Information", Proceedings of the International Special Topic Conference on Information Technology in Biomedicine, ITAB 2006, Ioannina - Epirus, Greece, 26-28 October 2006, 6 pages.

[4] S. Petroudi, T. Kadir, M. Brady, "Automatic Classification of Mammographic Parenchymal Patterns: a Statistical Approach", Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Cancun, Mexico, Vol. 1, 17-21 September 2003, pp. 798-801.

[5] M. Varma and A. Zisserman, "Classifying images of materials: Achieving viewpoint and illumination independence", Proceedings of the European Conference on Computer Vision, Copenhagen, Denmark, 2002, pp. 255–271.

[6] M. Tortajada, A. Oliver, R. Martí, M. Vilagran, S. Ganau, . Tortajada, M. Sentís, J. Freixenet, "Adapting Breast Density Classification from Digitized to Full-Field Digital Mammograms", Breast Imaging, Springer Berlin Heidelberg, 2012, pp. 561-568.

[7] Y. Zheng, "Breast Cancer Detection with Gabor Features from Digital Mammograms", Algorithms 3, No. 1, 2010, pp. 44-62.

[8] J. Suckling, J. Parker, D.R. Dance, S. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cernaez, S.L. Kok, P. Taylor, D. Betal, J. Savage, "The Mammographic Image Analysis Society Digital Mammogram Database", Proceedings of the 2nd International Workshop on Digital Mammography, York, England, pp. 375-378, 10-12 July 1994.

[9] M. Mustra, M. Grgic, R. Huzjan-Korunic, "Mask Extraction from Manually Segmented MIAS Mammograms", Proceedings ELMAR-2011, Zadar, Croatia, September 2011, pp. 47-50.

[10] M. Mustra, M. Grgic, J. Bozek, "Application of Gabor Filters for Detection of Dense Tissue in Mammograms", Proceedings ELMAR-2012, Zadar, Croatia, September 2012, pp. 67-70.

[11] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization", Graphic Gems IV. San Diego, Academic Press Professional, pp. 474-485, 1994.