

# A New Preprocessing Filter for Digital Mammograms

Peyman Rahmati<sup>1</sup>, Ghassan Hamarneh<sup>2</sup>, Doron Nussbaum<sup>3</sup>, and Andy Adler<sup>1</sup>,

<sup>1</sup> Dept. of System and computer Engineering, Carleton University, ON, Canada

<sup>2</sup> School of Computing Science, Simon Fraser University, BC, Canada

<sup>3</sup> Dept. of Computer Science, Carleton University, ON, Canada

{Prahmati, Adler, Nussbaum}@sce.carleton.ca , hamarneh@cs.sfu.ca

**Abstract.** This paper presents a computer-aided approach to enhancing suspicious lesions in digital mammograms. The developed algorithm improves on a well-known preprocessor filter named contrast-limited adaptive histogram equalization (CLAHE) to remove noise and intensity inhomogeneities. The proposed preprocessing filter, called fuzzy contrast-limited adaptive histogram equalization (FCLAHE), performs non-linear enhancement. The filter eliminates noise and intensity inhomogeneities in the background while retaining the natural gray level variations of mammographic images within suspicious lesions. We applied Catarious segmentation method (CSM) to compare the segmentation accuracy in two scenarios: when there is no preprocessing filter, and when the proposed preprocessing filter is applied to the original image. The proposed filter has been evaluated on 50 real mammographic images and the experimental results show an average increase of segmentation accuracy by 14.16% when the new filter is applied.

**Keywords:** Breast cancer, mammography, image denoising, and segmentation.

## 1 Introduction

Breast cancer is the leading type of cancer in women and the second most fatal type of cancer in women [1]. Mammography allows for the detection of intangible tumors and increases survival rate [2]. Digital mammography uses x-rays to project structures in the 3D female breast onto a 2D image [2]. The inherently noisy nature of digital mammograms, low contrast of suspicious areas, and ill-defined mass borders make mass segmentation a challenging problem. The design of an appropriate preprocessing filter is essential for segmentation algorithms to delineate masses with high accuracy.

Tumors appear as medium-gray to white areas on digital mammograms [3]. Tumor shapes are described by standardized keywords [4]: the shapes are grouped as oval, irregular, lobulated, or round and the margins are expressed as circumscribed, obscured, ill-defined or spiculated. The presence of irregularly-shaped masses and spicules correlate with increased likelihood of malignancy [5].

Given the low signal-to-noise ratio of mammography images, which greatly decreases the observable details, this paper develops a new strategy to enhance the original image. Previous works on mammographic-image preprocessing used methods such as Gamma correction [6], adaptive 2D Wiener filtering [7], contrast-limited adaptive histogram equalization (CLAHE) [8], and multi-scale wavelet-based

enhancement [9]. It is a difficult task to choose the single best technique for image enhancement, and the enhancement most often is evaluated based on the performance of the subsequent segmentation algorithm [10].

The role of a proper preprocessing filter is to provide an enhanced image to be fed into the subsequent blocks of a Computer Aided Diagnosis (CAD) system. To design an accurate CAD system to diagnose the tumors as either benign or malignant, the inherent nature of mammograms should not be affected by the preprocessing filter. Otherwise, important features which are probably useful for tumour classification will be lost. To retain the inherent characteristic intensity variations of mammograms, we propose a new preprocessor filter, named fuzzy contrast-limited adaptive histogram equalization (FCLAHE).

There are two goals for this work: enhancement of mammographic images to achieve better visibility to the human observer (radiologist), and smoothing inhomogeneities in the background of a main lesion under investigate to decrease the amount of probable false positive.

This paper is organized as follows: In the next section, we contrast some of the closely related works to our proposed method. In section 3, the suggested preprocessor filter (FCLAHE) is detailed. Finally, the experimental results are detailed in section 4. We draw the conclusion in section 5.

## 2 Related Works

In addition to the noise present in mammograms, some artifacts further complicate the diagnosis and introduce uncertainty into the image interpretation. These artifacts are related to the variability in tissue density and the inhomogeneous nature of tissue in some anatomical structures. These artifacts imply that designing algorithms for mammography preprocessing is a significantly more demanding task than for, say, medical images of homogeneous structures. The main goal is to decrease the inhomogeneities leading to increased accuracy in the subsequent mass segmentation algorithm. There are several works on mammographic-image preprocessing. Baeg *et.al* [6] used gamma correction for mammographic enhancement. Based on their texture-analysis method, classification of 150 biopsy-proven masses into benign and malignant classes resulted in an area under Receiver Operating Characteristic (ROC) curve of 0.91. Gamma correction is simple but the effects are localized and not global. The adaptive 2D Wiener filtering (A2DWF) noise-reduction algorithm [7] estimates the noise in the neighborhood around each pixel and then adjusts the surrounding region based on that noise estimate. Mayo *et.al* [11] compared A2DWF, a wavelet filter, a filter based on independent component analysis, and a diffusion-based filter. Although Mayo *et.al* did not extract features, the noise-removal aspects of the compared methods were shown similar, based on both visual observation and mean square error. To yield improved diagnostic performance, Mekle *et.al* [9] used an interactive multi-scale enhancement which incorporates dyadic spline wavelet functions and sigmoidal nonlinear enhancement functions.

In mammogram image, masses appear brighter and gradually darken as the image is traversed from the mass core toward the background. Since useful discriminatory features to classify tumors can be based on the natural intensity variations of

mammogram, it is important that the preprocessing filter retains such intensity characteristics.

To the best of our knowledge, designing a preprocessing filter to enhance the original mammographic image, while still preserving its natural intensity variations, has not been reported previously. In this work, we enhance the original mammograms while maximizing the preservation of their inherent characteristics. Pisano *et al.* [8] examined several digital mammograms using multiple methods of enhancement, including CLAHE. They concluded that image detail is good and that, in general, lesions appeared obvious compared to the background. They also found that graininess (inhomogeneity) was introduced due to the enhancement of image noise, which might mislead a radiologist to thinking that there are false microcalcifications. The advantage of CLAHE is that it is straightforward to implement and results in a high contrast image. As per the conclusions of Pisano *et al.* on the efficacy of CLAHE [8], the original mammographic images are first processed using CLAHE to eliminate noise and followed by the application of a nonlinear filter, which incorporates the advantages of a nonlinear fuzzy function.

### 3 Preprocessing Filter: FCLAHE

In this section, we develop a preprocessing filter to enhance digital mammograms, called fuzzy contrast-limited adaptive histogram equalization (FCLAHE). This work is inspired by CLAHE [8], with modifications to improve its performance. Pisano *et al.* [8] examined several image-enhancement methods, and expressed that CLAHE “might be helpful in allowing radiologists to see subtle edge information, such as spiculations”.

The main drawback of CLAHE is that the amount of the enhancement for the foreground and the background of the original image are similar (linearly filtered). The result is an image with high contrast in both the foreground and the background. In other words, it increases the visibility of the main mass at the cost of simultaneously creating small yet misleading intensity inhomogeneities in the background, i.e. leads the radiologist to the detection of more tumors than actually exists in the image (higher false positives). The presence of inhomogeneities in the background can also mislead the segmentation algorithm as to location of the real mass under investigation.

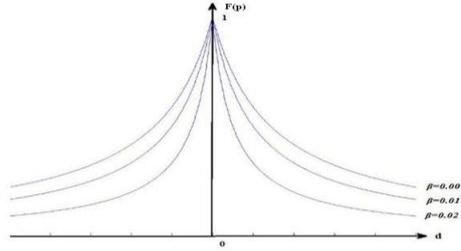
Since CLAHE eliminates noise in exchange for increasing the inhomogeneities in the background, there is a trade-off between the accuracy of the enhanced image and the inhomogeneities in the background. The proposed preprocessing filter tries to make an improved compromise between them. In the proposed filter, we retain the advantages of CLAHE, which are the ability to remove noise and obtaining a high contrast image, while improving on its deficiencies, which is the loss of inherent nature of enhanced mammography images. Therefore for simulating the natural gray levels variations of the original mammogram and providing a smoothed image with an acceptable brightness, a custom algorithm is used to supply non-linear enhancement.

We use fuzzy function proposed in [12] within a new form to provide a non-linear adjustment to the image based on the image gray level statistics. The fuzzy function in [12] attributes a membership value between 0 and 1 to the pixels based on the

difference of their gray levels than a seed point, located at the mass core. The fuzzy function defined as [12]:

$$F(p) = \frac{1}{1 + \beta d} , \quad (1)$$

where  $p$  is the intensity of the pixel being processed,  $d$  is the intensity difference between  $p$  and a seed point, and  $\beta$  controls the opening of the function. The larger the difference, the lower the membership function ( $F(p)$ ); and as  $\beta$  increases, the opening of  $F(p)$  decreases. Fig. 1 visualizes the behaviour of this function.



**Fig. 1.** Fuzzy function  $F(p)$  membership for various values of  $\beta$ .

The non-linear filter called fuzzy contrast-limited adaptive histogram equalization (FCLAHE) that we proposed is described as:

$$\hat{I}(x, y) = A_r(x, y) + (I(x, y) - A_r(x, y)) F(|I(x, y) - A_r(x, y)|) , \quad (2)$$

where  $A_r(x, y)$  denotes the average of the gray levels of the pixels placed radially in distance  $r = \sqrt{(x - x_{seed})^2 + (y - y_{seed})^2}$  away from a seed point, see the red circle in figure 2(b) with radial distance ( $r$ ) of 110 and  $A_r = 121$ . The seed point is selected by a radiologist at the center of the mass core, the brightest region in mammograms.  $F(d)$  is the fuzzy function defined in (1).  $I(x, y)$  is the original intensity of a pixel located at the coordinate  $(x, y)$ .  $\hat{I}(x, y)$  is the new intensity of the same pixel in the processed image.

In (2), it is assumed that the intensities of the pixels at a radial distance ( $r$ ) away from the seed point are similar unless there are inhomogeneities in the background. This assumption comes from the fact that masses in digital mammograms appear as brighter areas that gradually darken from the mass core outwards toward the background (see figure 2(a)), resembling an intensity variations profile similar to the presumed fuzzy function in figure 1 where  $d$  here indicates the radial distance from the seed point. Wherever there is a bright region in the background (due to background inhomogeneities),  $I(x, y)$  will be more different than  $A_r$ , i.e.  $I(x, y) - A_r(x, y) = \Delta I_r(x, y)$  will give a higher value. According to (1),  $F(\Delta I_r(x, y))$ , will therefore yield a lower value. This, in turn, will attenuate the second term in the right

hand side of (2), and  $\hat{I}(x, y)$  will be assigned  $A_r$ . If there are no inhomogeneities in the background, then  $A_r$  will be close to  $I(x, y)$ , i.e.  $\Delta I_r(x, y) \approx 0$ , and  $F(l)$  will yield unity, which in turn assigns  $\hat{I}(x, y)$  to  $I(x, y)$ , (i.e. the intensity of the pixel does not change). In fact, equation (2) removes inhomogeneities or the small regions with a high contrast in the background, which may be mistaken for a real lesion in the subsequent segmentation algorithm. In addition, the proposed equation keeps the brightness and the contrast of the mass core, see Fig. 2. As a result, the proposed FCLAHE algorithm provides a smoothed image so that has a close agreement with the nature of the original mammography image.

The major advantage of this filter compared to earlier approaches [6-9] is its ability to eliminate noise without affecting the intensity characteristics of mammographic image. One of the most common approaches to segment mammograms is by utilizing the statistical distribution of their intensities. Therefore, to achieve accurate lesion segmentation, it is important that preprocessing minimizes the effects of changing the statistical distribution or changes it properly to a predefined probability distribution function (PDF), see figure 5(c).

## 4 Experimental Results

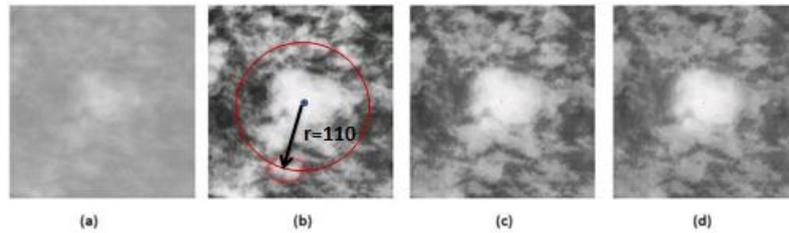
The proposed algorithm was run on a Pentium IV (PC), Intel 3.0 GHz, with Windows XP Professional, 3.0 GB RAM, in MATLAB 7.0. We executed our proposed method over 50 mammographic images which were selected from Digital Database of Screening Mammography (DDSM) [13]. Fig. 2 shows the enhanced images with FCLAHE for different  $\beta$  values. Note how the result from CLAHE introduce erroneous enhancement of background intensity patterns (red oval in Fig. 2b).

To investigate the effect of the FCLAHE over the accuracy of the subsequent segmentation algorithm, we performed the Catarious segmentation method (CSM) to delineate the masses [14]. We performed CSM segmentation method on both the original images and the enhanced images using either CLAHE or FCLAHE. Then, we evaluated the segmentation accuracy for them based on expert-validated ground truth delineation using overlap criterion defined as follows:

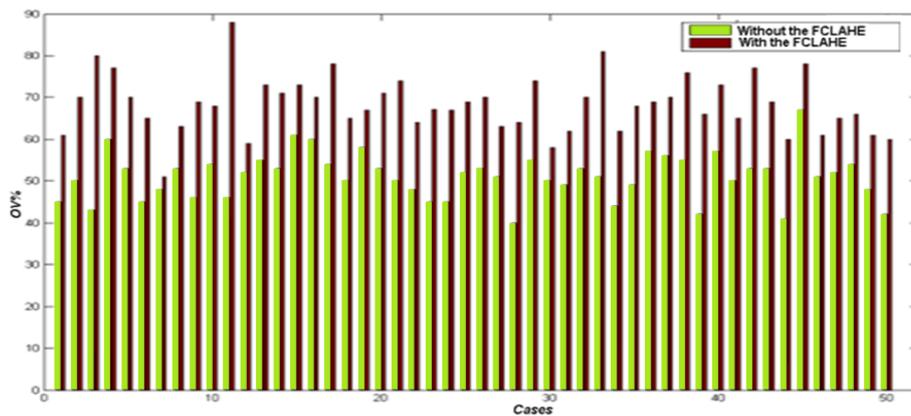
$$OV = \frac{A \cap B}{A}, \quad (3)$$

where  $A$  is the ground truth delineation,  $B$  is the competing delineated area, and  $OV = 1$  when  $A$  and  $B$  match perfectly. The bar chart in Fig. 3 shows the *overlap* values ( $OV\%$ ) (between the automatic segmentation and ground truth) for each study case with and without FCLAHE. The average overlap with expert increased from 52% to 68.36% when the proposed filter was employed (averaged over 50 images), i.e a 16.3% increase in accuracy. Moreover, the average overlap with expert was 54.2% when we applied CLAHE. Comparing with CLAHE average segmentation accuracy (54.2%), FCLAHE offers a 14.16% increase in accuracy. Fig. 4 depicts qualitative comparative segmentation results. The corresponding quantitative increase in accuracy for these cases were: from 41% to 67.1% for Fig. 4 (a), 49% to 80.1% for Fig. 4(b), 55% to 69% for Fig. 4(c), and 46% to 88.2% for Fig. 4 (d). It seems that

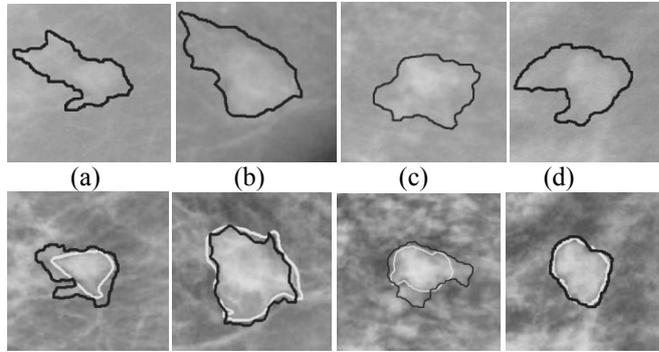
Fig. 4(a) and Fig. 4(c) have less accurate segmentation results after applying FCLAHE due to having several small bright regions (inhomogeneities) immediately attached to their main mass boundary. Fig. 5 depicts the effect of CLAHE and FCLAHE on the histogram specification of original mammogram for one case study. Fig. 5(a) shows the histogram of the original image which is totally spiky and has less symmetry while in Fig. 5(b), the histogram of CLAHE-based enhanced image has better symmetry but still suffers from spikes due to the presence of inhomogeneities in the background. FCLAHE-based enhanced image, shown in Fig. 5(c), offers the smoothest and the most symmetric histogram specification, which is better similar to an exponential distribution such as Poisson distribution which is already deemed to be effective for modeling the intensity variations of mammogram [15].



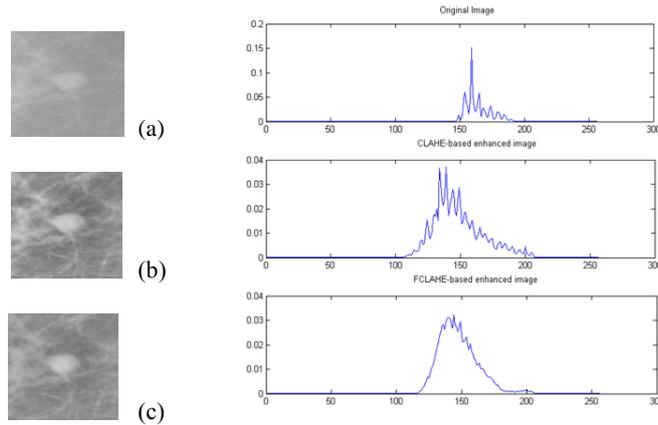
**Fig. 2.** An example of applying the FCLAHE algorithm to a mammographic image. (a) Original mammographic image. (b) The enhanced images of (a) with CLAHE, which has a good contrast but some small regions (which one of them is shown by a red oval) in the background may misleadingly be mistaken for a real mass. (c) Our proposed enhanced image (FCLAHE) with  $\beta = 0.009$  which has an acceptable contrast while there are no regions with a high contrast (high intensity inhomogeneity) in the background. (d) FCLAHE image with  $\beta = 0.02$ .



**Fig. 3.** The effect of the FCLAHE on the accuracy ( $OV\%$ ) of the CSM as a function of each case number.



**Fig. 4.** Examples of the CSM performance over mammographic images with and without the proposed pre-processing filter. First row: The delineated images achieved using the original mammograms without applying the FCLAHE. Second row: The final contours of CSM algorithm (black lines) performed on the enhanced images using the FCLAHE. The ground truth is shown in white.



**Fig. 5.** The effect of the proposed preprocessing filter on the histogram of an original mammogram image. (a) Original mammogram image along with its histogram specification. (b) CLAHE-based enhanced image with its histogram. (C) FCLAHE-based enhanced image with its histogram, resembling an exponential PDF.

## 5 Conclusion

We presented a novel preprocessing filter, referred to as FCLAHE, which was applied in the context of computer-aided delineation of lesions in mammographic images. The new model incorporates the advantages of the preceding predecessor of CLAHE and simultaneously addresses its deficiencies. The proposed preprocessing filter improved

the CLAHE algorithm by applying a nonlinear fuzzy function in a new form. It conserved the brightness of the mass core and improved the graininess and small inhomogeneous regions in the background in an appropriate manner. The enhanced images retained the inherent nature of mammograms. The overlap with expert delineations increased from 54.2% to 68.36% when the FCLAHE is applied, averaged over all 50 images in the database. A potential direction for future work is to apply a PDF-based segmentation method on the FCLAHE-enhanced mammography images, in anticipation of higher segmentation accuracy due to the FCLAHE advantage of preserving the intensity variations of mammogram.

## References

1. American Cancer Society, "American Cancer Society: Breast Cancer Facts & Figures 2005-2006," pp. 1-28, (2006)
2. Rangayyan, R.M., "Breast cancer and mammography," in *Biomedical Image Analysis*, M.R. Neuman, Ed. Boca Raton, FL: CRC Press, pp. 22-27, (2005)
3. Egan, R.L., *Breast Imaging: Diagnosis and Morphology of Breast Diseases*. Philadelphia, PA: W. B. Saunders Co., (1988)
4. The Mosby Medical Encyclopedia, Revised ed. New York, NY: The Penguin Group, (1992)
5. Jeske, J.M., Bernstein, J.R., and Stull, M.A., "Screening and Diagnostic Imaging," in *American Cancer Society Atlas of Clinical Oncology* London, UK: B.C. Decker, pp. 41-63, (2000)
6. Baeg, S. and Kehtarnavaz, N., "Texture based classification of mass abnormalities in mammograms," *Proc. of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS)*, Houston, TX, vol. 1, pp. 163-168, Jun. (2000)
7. Mayo, P., Rodenas, F., and Verdu, G., "Comparing methods to denoise mammographic images," *Proc. of the 26th Annual Intl. Conference of the Engineering in Medicine and Biology Society (EMBC)* vol. 1, pp. 247-250, (2004)
8. Pisano, E.D., Cole, E.B., Hemminger, B.M., Yaffe, M.J., Aylward, S.R., Maidment, A.D.A., Johnston, R.E., Williams, M.B., Niklason, L.T., Conant, E.F., Fajardo, L.L., Kopans, D.B., Brown, M.E., and Pizer, S.M., "Image Processing Algorithms for Digital Mammography: A Pictorial Essay," *RadioGraphics*, vol. 20, no. 5, pp. 1479-1491, (2000)
9. Mekle, R., Laine, A.F., Smith, S., Singer, C., Koenigsberg, T., and Brown, M., "Evaluation of a multiscale enhancement protocol for digital mammography," *Proc. of the Wavelet Applications in Signal and Image Processing VIII*, San Diego, CA, USA, vol. 4119, pp. 1038-1049, (2000)
10. Gonzalez, R.C., and Woods, R.E., *Digital Image Processing*, 2nd ed. Upper Saddle River, NY: Prentice Hall, (2002)
11. Mayo, P., Rodenas, F., and Verdu, G., "Comparing methods to denoise mammographic images," *Proc. of the 26th Annual Intl. Conference of the Engineering in Medicine and Biology Society (EMBC)* vol. 1, pp. 247-250, (2004)
12. Guliato D., Rangayyan R.M., Carnielli W.A., Zuffo J.A., and Desautels J.E.L., "Segmentation of breast tumors in mammograms using fuzzy sets," *Journal of Electronic Imaging*, 12(3):369-378, (2003)
13. Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, W. PH., Moore, R., Chang, K., and MunishKumaran, S., Current status of the Digital Database for Screening Mammography

<<http://marathon.csee.usf.edu/Mammography/Database.html>> 1998 (Accessed 15 Sep. 2009).

14. Catarious, D.M., Baydush, A.H., and Floyd, C.E., "Incorporation of an iterative, linear segmentation routine into a mammographic mass CAD system," *Medical Physics*, vol. 31, no. 6, pp. 1512-1520, Jun. (2004)
15. Rahmati, P., and Ayatollahi, A., "Maximum Likelihood Active Contours Specialized for Mammography Segmentation," *The 2nd IEEE International Conference on BioMedical Engineering and Informatics (BMEI'09)*, vol. 1, pp. 257-260, China, (2009)