Signal Processing 90 (2010) 952–958

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast communication

An automatic method to enhance microcalcifications using Normalized Tsallis entropy

Mohanalin *, Prem Kumar Kalra, Nirmal Kumar

Department of Electrical Engineering, IIT Kanpur, India

ARTICLE INFO

Article history:
Received 18 April 2009
Received in revised form
6 July 2009
Accepted 15 September 2009
Available online 23 September 2009

Keywords:
Mammogram
Normalized Tsallis entropy
Shannon entropy
Fuzzy
Non-extensive

ABSTRACT

This paper presents a new approach to enhance the contrast of microcalcifications in mammograms using a fuzzy algorithm based on Normalized Tsallis entropy. In phase I, image is fuzzified using Gaussian membership function. In Phase II, using the non-uniformity factor calculated from local information, the contrasts of microcalcifications are enhanced while suppressing the background heavily. Some of the previous works are related to Shannon entropy and Tsallis entropy. This is the first time in literature to propose an enhancement algorithm using Normalized Tsallis entropy. Normalized Tsallis entropy has an extra parameter q. Our proposed work is completely automatic and q values are selected empirically. The proposed approach improves the detection process vastly. Without Normalized Tsallis entropy enhancement, detection of MCs is meager 80.21% Tps with 8.1 Fps, whereas after introduction of Normalized Tsallis entropy, the results have surged to 96.25% Tps with 0.803 Fps.

1. Introduction

There is a rising incidence of breast cancer in India. According to The International Agency for Research on Cancer, which is a part of the World Health Organization, there were approximately 79,000 women affected by breast cancer in India in 2001 and over 80,000 women in 2002. The seriousness of the situation is apparent after going through recent data from Indian Council of Medical Research (ICMR). Mammography is preferred for early detection of breast cancer. But still, approximately 10–30% of women with breast cancer who had mammography tests show negative results. Roughly, there are two-thirds of false negatives (Fns) [1]. Mammograms with suspicious abnormalities which the radiologist unsuccessfully detected are called as false negatives [2]. Apart from this, some times, the harmless lesions could be mistaken for disease leading to false-positives (Fps), which would be recommended for biopsy which causes the patient mental and physical trauma that could often be evaded. Only 15–34% of the biopsies carried out are non-palpable, and suspicious lesions identified through mammogram actually prove to be malignant. Studies indicate that failure to detect breast cancer correctly could be due to technical reasons and human error. Technical reasons include inadequate compression or poor image quality due to improper positioning or faulty equipment. Human error is due to fatigue, oversight on the part of the radiologist, distraction by more prominent image features, and varying decision criteria as narrated by Giger [2]. Leading researchers including Papadopoulos et al. [3], Gavrielides et al. [4], Verma and Zakos [5], etc. have concluded that computer aided detection (CAD) algorithms play as second to the radiologists to locate suspicious spots on the mammogram. Microcalcifications (Mcs) are the early signs of breast carcinomas and their detection is one of the key issues for breast cancer control. Mcs, specks of calcium found in an area of rapidly dividing cells, are the first sign of breast cancer in more than half of all cases and mammogram is the best way to image it in earlier stage.

* Corresponding author.
E-mail address: mohanlee@iitk.ac.in (Mohanalin).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.09.012

Identifying Mcs is a challenging task for radiologist and the major reasons being the size of the Mcs which is being very small can lead to potential misidentification, Secondly there is no uniform size for Mcs. In some cases, the regions of interest may be of low contrast. The intensity difference between suspicious areas and their surrounding tissues is thin especially in denser breasts, which causes Mcs to be almost invisible.

Cheng et al. [6] have proposed a fuzzy based Mcs enhancement. The image is fuzzified using S membership function. Then using the local statistics computed, the ROI of image is enhanced. Jiang et al. [7] combines fuzzy logic and structure tensor for improved enhancement of Mcs in digital mammograms. This method suppresses non-Mcs regions while enhancing the Mcs regions. In [8], a multi-scale analysis method is proposed for the Mcs enhancement. This method enhances Mcs unseen or barely seen features of mammography without additional radiation. Sakellaropoulos et al. [9] uses local modification of multi-scale gradient magnitude values provided by the redundant dyadic wavelet transform to minimize image noise while optimizing contrast of image features. In [10], an algorithm for feature enhancement in mammograms is presented; using discrete wavelet enhances multi-scale structures in images. Portes de Albuquerque et al. [11] in their paper have proposed an image thresholding by nonextensive entropy, regarding the presence of nonadditive information content in some images. They have discussed some results to illustrate the influence of the parameter q in the thresholding. El-Fegh et al. [12] have proposed a new Three-level gray-scale images segmentation using non-extensive entropy. This method is based on maximizing non-extensive entropy. But it is a more or less similar work to Portes de Albuquerque. However a more appropriate method for Mcs detection based on Tsallis entropy (TE) is presented by Mohanalin et al. [13]. One major blockade of this algorithm is that it does not discuss the automatic 'q' computation. It deals with trial and error approach which makes the algorithm ineffective when applied to huge amount of data. Later they come up with an automatic algorithm which relates the density index calculated from the mammogram to the Tsallis parameter which gives an outstanding performance of results with 97% true positives (Tps) with 0.8 Fps. In fact, the automatic 'q' detection is a long standing issue. Our motivation for the proposed scheme arises from this work. Halkiotisa et al. [14] have used mathematical morphology and artificial neural networks to segment Mcs and achieved positive detection rate of 94.7% and 0.27 false positives per image.

Normalized Tsallis entropy (NTE) is often considered as nephew of TE. Both TE and NTE do not possess similar qualities and properties. Abe [15] proves TE, and the NTE all converge to the Boltzmann–Gibbs–Shannon entropy for the limit $q \to 1$. Also NTE is concave only for $q \in (0, 1)$, whereas the TE is always concave for any positive values of q. It naturally creates curiosity to check the performance of NTE for enhancement of Mcs. This paper shows the application of (NTE) as a new method of Mcs enhancement. The purpose of this paper is to automatically determine the fuzzy region and the bi-level threshold based on the maximum fuzzy entropy principle. Our approach involves a novel fuzzy partition on a 2D histogram accompanied with NTE for optimal results. Finally the performance of the proposed scheme is used to compare the results between traditional Shannon entropy based works and TE. A detailed free-response operating characteristic (FROC) analysis is shown where the detection results are shown with and without NTE based enhancement. Significant improvement can be observed from the graphs shown at the end of the Section 6.

Further to establish a relation between mammogram complexity and the Tsallis parameter, it is essential to understand the structure of breast. Breast tissue is a mixture of non-dense tissue (fat) and dense tissue (glands, ligaments and stromal tissue). The relative ratio of density to fat varies among women. Dense breast tissue appears as a solid white area on a mammogram film and fat emerging as a dark area. Mammogram X-rays do not penetrate or "see through" dense tissues as well as they do through the fat. Hence, in women with dense breasts, mammograms are more difficult to interpret. Tumors are dense tissues and appear as solid white areas on the film. In this work, it is assumed that complexity of the mammogram is linked with the amount of whiteness present in the mammogram. In other words, as the denseness (whiteness) of the mammograms increases the complexity of the mammogram increases. This is fairly an acceptable assumption, because as the denseness of mammogram increases the isolation of Mcs visually becomes tedious.

The paper is organized as follows. Section 2 explains how the thresholding can be done using NTE. Section 3 discusses the issues related to the selection of 'q' parameter. Section 4 narrates architect of proposed algorithm and Section 5 the fuzzy based enhancement technique. The achieved results and figures are displayed in Section 6. Section 7 contains the conclusion. We have imported mammograms for research from Lawrence Livermore National Laboratory. All films are digitized at 35 μ per pixel, and with 12 bits of grayscale per pixel. We have also procured the MIAS data from its official website.

2. Thresholding the mammograms using NTE

Mammogram can be seen as 3 individual objects fused together, background, tissues and MCs. We have redesigned NTE according to the need for mammograms. Global techniques tend to fail as they may contain background, pectoral muscle and unwanted tissues with Mcs. We neglect the region lesser than average of the image [10]. We have redesigned the limits of NTE in such a way to make the technique local.

The mean can be calculated by

$$k = (1/w) * \sum_{x,y \in G}^{X} \sum^{Y} g_{x,y} \tag{1}$$

where x and y are dimensions of the image, G be the intensities larger than 100, w be the number of pixels in this region, g_{xy} is the grey level at the coordinates x and y. Let p_1, p_{k+1},…, p_N be the probability distribution of grey levels in the mammogram. The tissue background of the

mammogram can be written as

$$p_k/P_t - P_{K-1}, p_{K+1}/P_t - P_{K-1}, \ldots, p_t/P_t - P_{K-1}$$

The region of interest i.e. MCs of the mammogram can be written in terms of equations as:

$$p_{t+1}/1 - P_t, p_{t+2}/1 - P_t, \ldots, p_N/1 - P_t.$$

where

$$P_{K-1} = \sum_{i=1}^{k-1} p_i, \quad P_t = \sum_{j=k}^{t} P_j.$$

Generally TE for Mcs problem can be written [13] as follows:

$$S^A{}_q[p] = (1/q-1) * \left(1 - \sum_{i=k}^{N} \left(\frac{p_i}{P_t - P_{K-1}}\right)^q\right) \tag{2}$$

If A and B denote the region above and below the threshold, then NTE can be written from TE as follows

$$S_A^{NT}{}_q[p] = S^A{}_q[p] / \sum_{i=k}^{t} \left(\frac{p_i}{P_t - P_{K-1}}\right)^q$$

$$= (1/1-q) * \left(1 - 1/\sum_{i=k}^{t} \left(\frac{p_i}{P_t - P_{K-1}}\right)^q\right)$$

$$S_B^{NT}{}_q[p] = S^B{}_q[p] / \sum_{i=t+1}^{N} \left(\frac{p_i}{1 - p_t}\right)^q$$

$$= (1/1-q) * \left(1 - 1/\sum_{i=t+1}^{N} \left(\frac{p_i}{1 - p_t}\right)^q\right)$$

where t–threshold values; q–Tsallis parameter € (0,1) and N–Max value of image.

NTE exhibits pseudo-additivity [15] and can be written as

$$S^{NT}[p]_q(A+B) = S^{NT}{}_q(A) + S^{NT}{}_q(B) + (1-q) \cdot S^{NT}{}_q(A) \cdot S^{NT}{}_q(B) \tag{3}$$

We have used the traditional energy maximization technique.

$$t^* = \text{Arg}\max\{(S_A^{NT}{}_q[p] + S_B^{NT}{}_q[p])\} \tag{4}$$

Eq. (4) is the optimal threshold which separates MCs from the tissues ideally.

3. Selection of optimal '*q*' parameters

An experiment is performed with 4 set of graded mammograms from MIAS and UCSF database. The purpose of this experiment is to pre-define optimal values of q, so that the algorithm will be automatic. This is done by relating the q parameter and the complexity of the mammogram through experiments. It is a known fact, that when the grade increases the amount of presence of whiteness of the mammogram also increases and vice versa. As the grade increases, the texture also gets complicated. This is not taken into account, since this process is to pick the optimal 'q' parameter and is not dedicated to grade the mammograms. Each graded set contains 20 hand picked mammograms with presence of suspicious Mcs. By fixing several values of q ranging from 0.1 to 0.9, the thresholding operation is performed and analysed visually with the help of a doctor. The results of the thresholding operation using the Eq. (4) are shown through Fig. 1. Only the results of significance have been shown. The images used for display are BARCC, ACRCC and AFRCC of UCSF database and Mdb219 of MIAS database. By studying the results obtained, we have concluded that q of 0.1 gives good thresholding results (99.07% Tps with 0.1 Fps) for grade 1 mammograms.

But when complexity of the mammograms increases (namely grade 4) the thresholding operation gives poorer results (84.23% Tps with 9 Fps) when compared to the other certain values of q. Similarly the value of q=0.9 has given good results (94.45% Tps with 0.5 Fps) for grade 4 images, but corroded important results (81.05% Tps with 0.3 Fps) of the less dense mammograms (grade 1) and so on.

Table 1 gives the list of the threshold values (T1, T2, T3, T4) obtained for the corresponding 'q' parameter used in NTE for various images. The quality of the results is obtained by completing the entire detection procedure followed by FROC analysis by varying the values of q. For the entire set of mammograms, the whole set of experiments are repeated iteratively to tune the best values for 4 different levels of complexity of mammograms. Finally, the best choices of q are defined by studying the amount of Fps and Tps the experiments have generated. i.e. for grade 1 mammograms q=0.1 is fixed and for grade 2 it is 0.5 and so on. Now we have established a relation between Tsallis parameter and grade of mammogram. The index calculator proposed by Mohanalin et al. [16] is built using the same principle explained above and it is used here. This index (ψ) calculator computes the amount of whiteness present in mammogram and relates it to appropriate grades and assigns suitable values of 'q' which we have computed empirically. Following are the values of 'q' assigned according to the index computed from mammogram.

$$\left.\begin{array}{l} \text{Index } 1 \rightarrow 0 < \Psi < 0.01 \rightarrow q = 0.1 \\ \text{Index } 2 \rightarrow 0.01 < \Psi < 0.1 \rightarrow q = 0.5 \\ \text{Index } 3 \rightarrow 0.1 < \Psi < 0.2 \rightarrow q = 0.7 \\ \text{Index } 4 \rightarrow 0.2 < \Psi < 1 \rightarrow q = 0.9 \end{array}\right\}$$

For the images under Index 1 category, the value of q is defined as 0.1 and so on.

4. Proposed algorithm

Fig. 2 shows the architect of proposed algorithm. In phase I, the computation of Tsallis parameter 'q' is done by calculating the index. For the computed 'q', TE is calculated for finding the optimal threshold between Mcs and Background. In phase II this information is utilized to enhance the Mcs by fuzzy technique. Fuzzy based technique is used to maximize the variation between tissue and MCs, because structure of the mammogram itself is fuzzy.

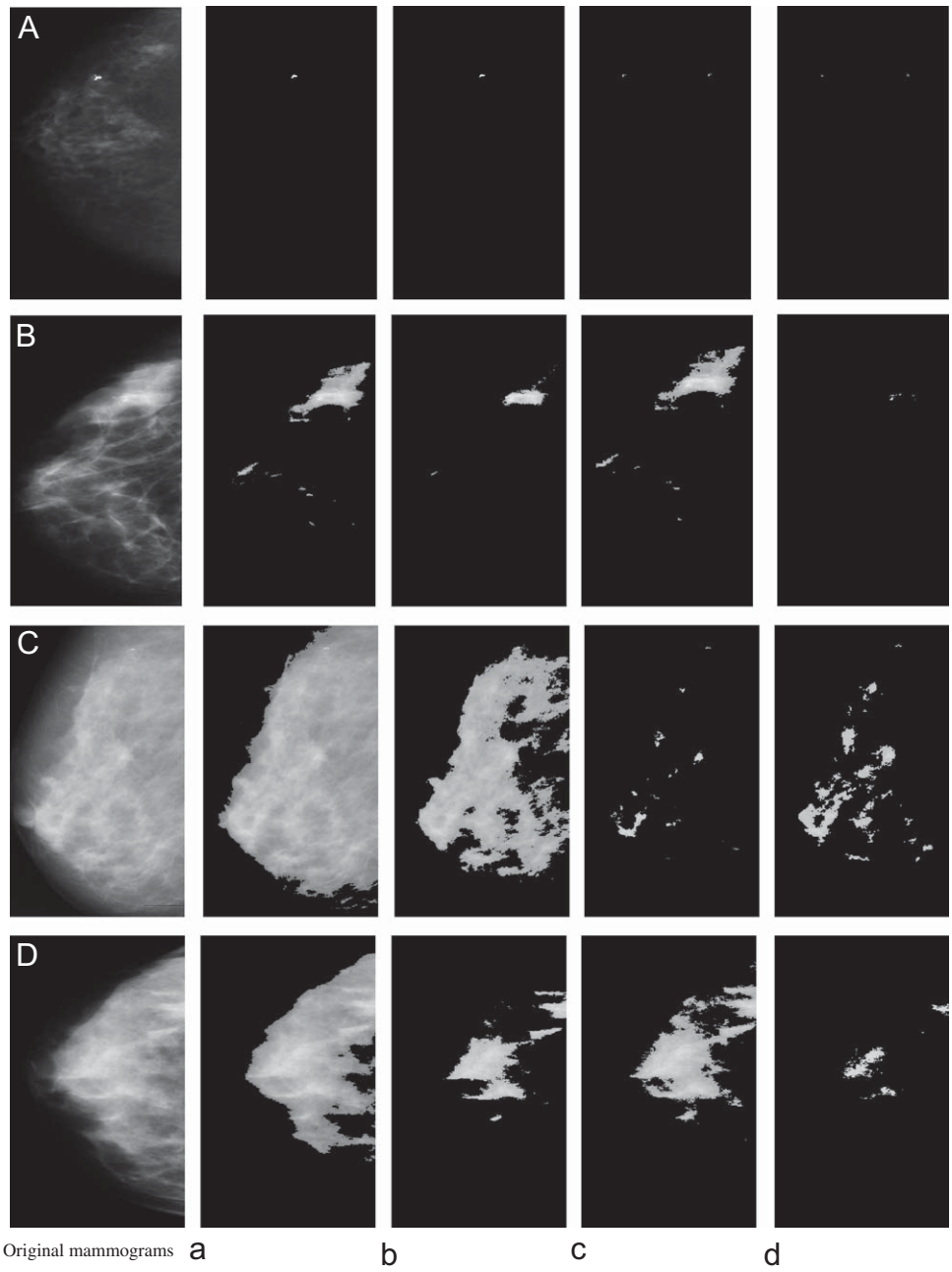

Fig. 1. (A–D) Displays original mammogram (BARCC, ACRCC and AFRCC and Mdb219) and thresholding results at (a) q=0.1, (b) q=0.5, (c) q=0.7, and (d) q=0.9, respectively.

Table 1
Value of 'q' and its corresponding thresholds obtained.

Image label	Thresholds obtained	At q=0.1	At q=0.5	At q=0.7	At q=0.9
BARCC	T1	173	191	228	235
ACRCC	T2	150	176	154	226
MDB219	T3	115	153	189	156
AFRCC	T4	107	171	156	199

5. Image fuzzification and enhancement

Mammograms are highly fuzzy in nature, having complicated structures like fatty acids and tissues running everywhere, making the life harder for radiologist. Due to this reason, fuzzy logic would be a better choice to deal the fuzziness of mammograms than traditional non-fuzzy methods. We have used Gaussian membership function because of its simplicity and robustness. A sigmoid

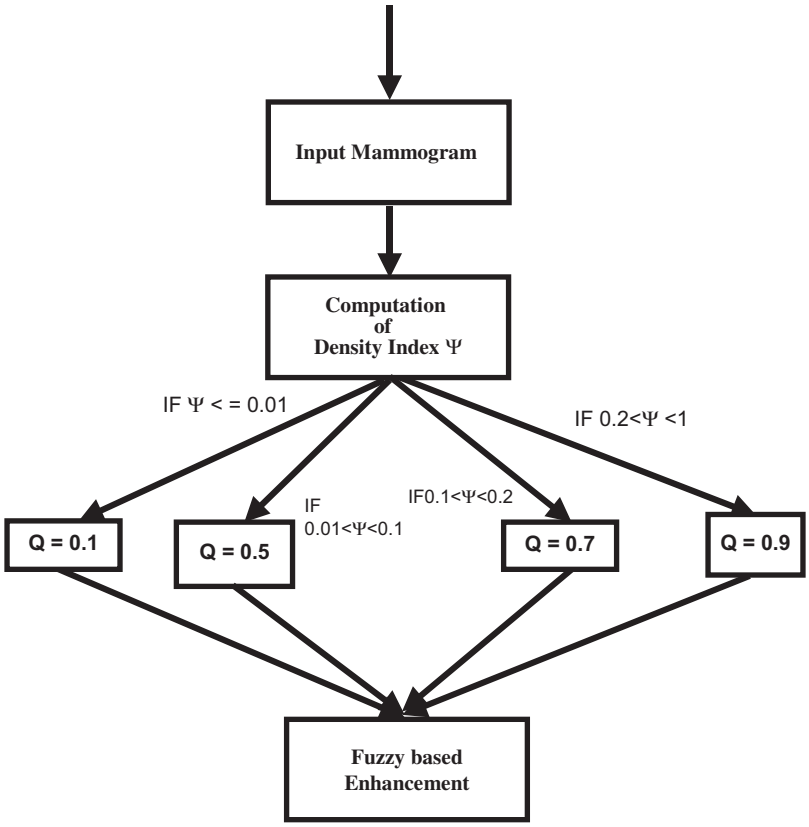

Fig. 2. Demonstrates our algorithm. t^*—threshold obtained by NTE for fuzzification.

function is not capable of modeling a range such as a class interval. A triangular function will not ensure that all inputs are fuzzified in some classes. These are the reasons why Gaussian function has been used. The intensities of an image are transformed to an interval [0, 1]. The proposed approach gives maximum membership value to the intensities of Mcs higher than the threshold (t^*) and suppress heavily the remaining. The Gaussian function can be written as

$$\pi_{m,n} = \exp\left(-\frac{|x - \mathrm{Max}\,N|^2}{f_h^2}\right) \tag{5}$$

X-Pixel value at (m,n) position, Max N—maximum intensity value and f_h^2—width of the Gaussian membership function. The bandwidth of the Gaussian function can be found out by the following equation.

$$f_h = \max\{(t^* - k),\ (\mathrm{Max}\,N - t^*)\} \tag{6}$$

k is the mean value calculated in (1) and t^*—threshold found using (5). This is a pixel by pixel process and not traditional windowing technique. By this way we suppress more the region which doesn't belong to the ROI and enhance the remaining. M*M moving window is used to find the local variances which measures the nonuniformity. It can be expressed as

$$\mu_{xy} = \frac{1}{\mathrm{M}\times\mathrm{M}}\sum_{j=1}^{M^2} g_j, \sigma_i = \frac{1}{\mathrm{M}\times\mathrm{M}}\sum_{j=1}^{M^2}[g_j - \mu_{xy}]^2 \tag{7}$$

where μ_{xy} is the local mean, σ_i is the local variance, x and y are the coordinates of the current pixel.

The non uniformity factor calculated will be normalized by using the optimum threshold:

$$V_i = \begin{Bmatrix} \sigma_i^2/t^* & \text{if } \sigma_i^2 \le t^* \\ 1 & \text{otherwise} \end{Bmatrix} \tag{8}$$

t^* is the optimal threshold found using NTE (3). The enhanced image can be obtained by the Eq. (9).

$$\text{Enhanced image} = \Pi^* V_i^* \,\mathrm{Max}\,N \tag{9}$$

Here Π denotes the Fuzzified image.

6. Experimental results

Now let us check the efficacy of the proposed algorithm by choosing 4 mammograms randomly (belonging to different grades) from the MIAS and UCSF. Since page is a constraint here, we show one example of enhancement result each for 4 different grades of mammogram. This paper concentrates on enhancement of mammograms and hence only the enhancement results alone are displayed. But FROC test which is traditionally used for analysis of detection algorithms are shown here to explain how the proposed NTE algorithm boosts up the performance for detecting Mcs. Analysis has been taken on a number of mammograms, all containing Mcs showing presence of tumors (either benign or malign).

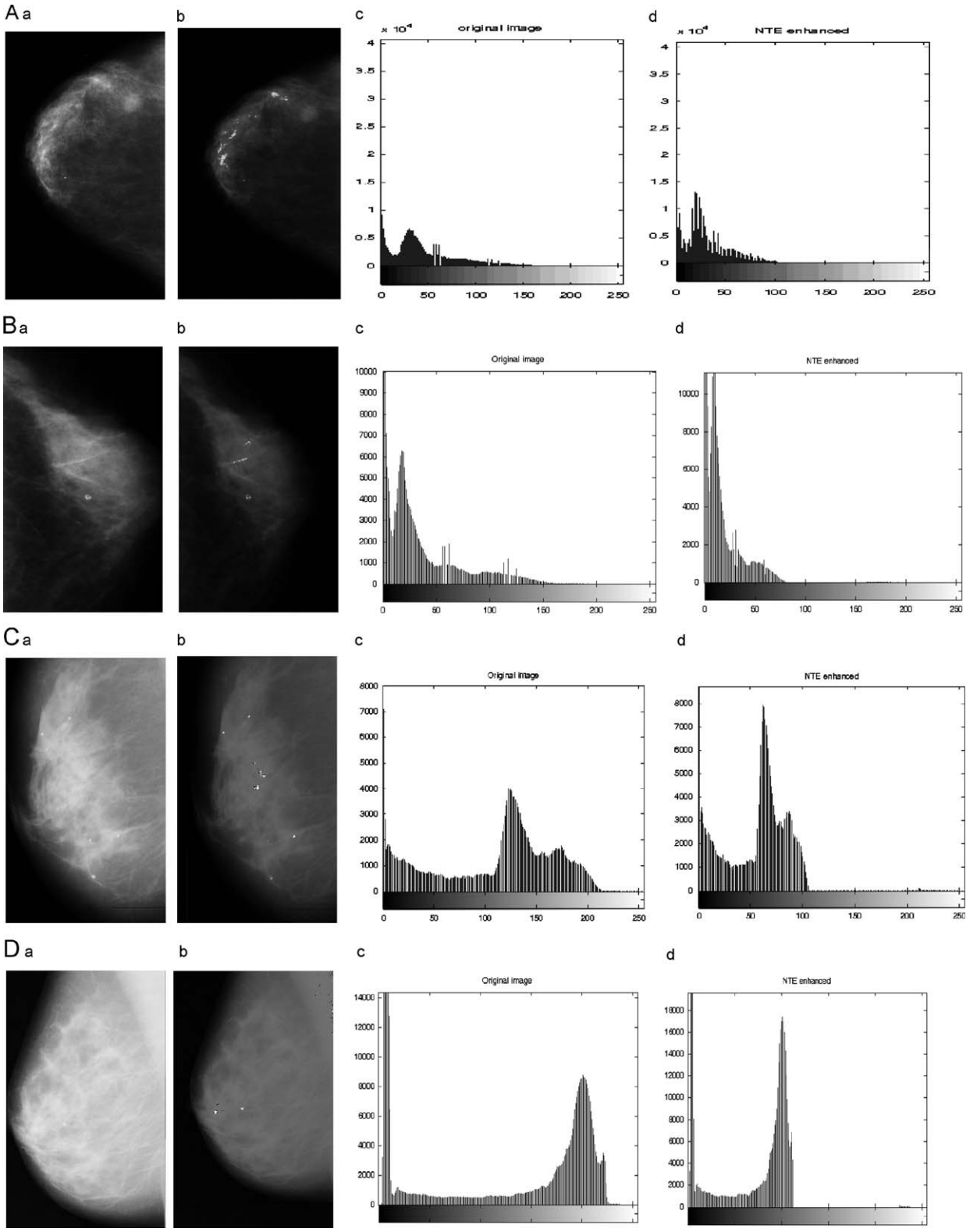

Fig. 3. (A–D) Displays (a) original (ARRC, AJLCC, mdb161 and mdb029) mammogram and (b) enhanced mammogram along with their histogram plots (c) before and (d) after NTE enhancement.

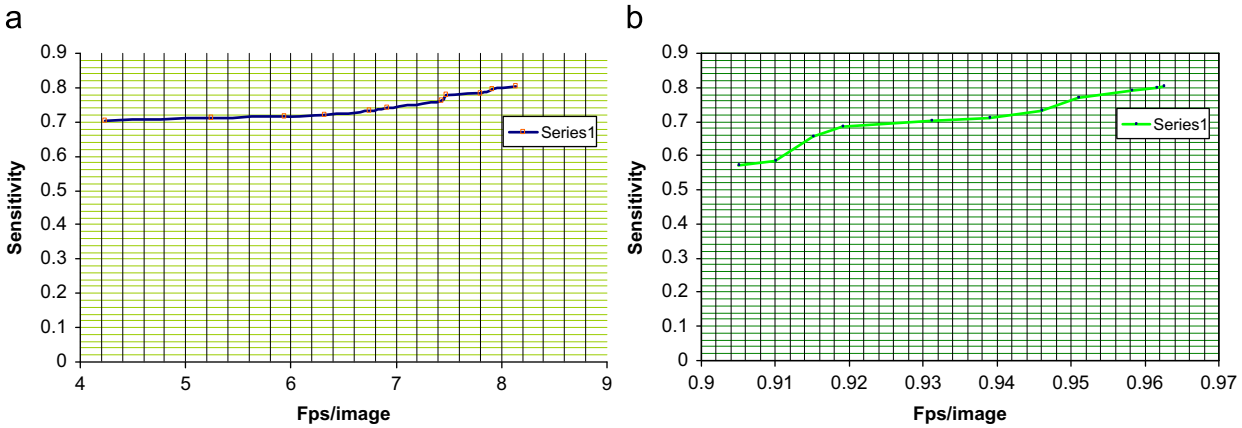

Fig. 4. Displaying the FROC curves without NTE in detection algorithm (a) and being included in the algorithm (b).

Fig. 3 displays both original and enhanced images along with their histograms. It is observed from the figures that our algorithm is precise for all the grades of mammogram. The displayed images are ARRC, AJLCC of UCSF database and mdb161, mdb029 of MIAS database.

7. Conclusion

In this paper, we use NTE and fuzzy technique to enhance Mcs in digitized mammograms. The proposed approach is very efficient for locating Mcs in the mammograms with various densities. The advantages of the proposed approach are: (1) the Mcs can be accurately enhanced even in mammograms of very dense breasts; (2) mammogram enhancement is more adaptive and robust and fits for various databases; (3) fuzzy based enhancement algorithm can enhance the main feature while suppress the noise; (4) without NTE enhancement the FROC test has yielded only 80.21% Tps with 8.1 Fps. But proposed approach has increased the detection rate to 96.25% Tps with 0.803 Fps. The proposed approach can be used by radiologist to detect Mcs more accurately and efficiently, and it will be useful for breast cancer control. Fig. 4 displays the FROC curve drawn for the experiments conducted on 50 MIAS and 197 UCSF mammograms.

References

[1] S.A. Feig, M.J. Yaffe, Digital mammography, computer-aided diagnosis and Telemammography. The Radiologic Clinics of North America, Breast Imaging 33 (6) (1995) 1205–1230.
[2] M.L. Giger, Computer-aided diagnosis in medical imaging—a new era in image interpretation, Medical Imaging Ultrasound (2000) 75–79.
[3] Papadopoulos, D.I. Fotiadis, A. Likas, Characterization of clustered microcalcifications in digitized mammograms using neural networks and support vector machine, Artificial Intelligence in Medicine 34 (2005) 141–150.
[4] M. Gavrielides, M. Kallergi, L.P. Clarke, Automatic shape analysis and classification of mammographic calcifications, SPIE 3034 (1997) 869–876.
[5] Verma, J. Zakos, A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques, IEEE Transaction on Information Technology in Biomedicine 5 (1) (2001) 46–54.
[6] H.-D. Cheng, Y.M. Lui, R.I. Freimanis, A novel fuzzy logic approach to mammogram contrast enhancement, Information Sciences 148 (1–4) (2002) 167–184.
[7] J. Jiang, B. Yao, A.M. Wason, Integration of fuzzy logic and structure tensor towards mammogram contrast enhancement, Computerized Medical Imaging and Graphics 29 (1) (2005) 83–90.
[8] A.F. Laine, S. Schuler, J. Fan, W. Huda, Mammographic feature enhancement by multiscale analysis, IEEE Transactions on Medical Imaging 13 (4) (1994) 725–740,.
[9] P. Sakellaropoulos, L. Costaridou, G. Panayiotakis, A wavelet-based spatially adaptive method for mammographic contrast enhancement, Physics in Medicine and Biology 48 (6) (2003) 787–803.
[10] P. Heinlein, J. Drexl, W. Schneider, Integrated wavelets for enhancement of microcalcifications in digital mammography, IEEE Transaction on Medical Imaging 22 (3) (2003) 402–413.
[11] M. Portes de Albuquerque, et al., Image thresholding using Tsallis entropy, Pattern Recognition Letters 25 (9) (2004) 1059–1065.
[12] I. - El.-Feghi, M. Galhoud, M.A. Sid-Aadhmed, M. Ahmadi, Three-level gray-scale images segmentation using non-extensive entropy, Computer Graphics, Imaging and Visualisation (CGIV 2007) (2007) 304–307.
[13] J. Mohanalin, P.K. Kalra, Nirmal Kumar, Tsallis entropy based microcalcification segmentation, ICGST-GVIP Journal 9 (I) (2009) 49–55.
[14] S. Halkiotisa, T. Botsis, M. Rangoussi, Automatic detection of clustered microcalcifications in digital mammograms using mathematical morphology and neural networks, Elsevier Signal Processing 87 (2007) 1559–1568.
[15] S. Abe, Tsallis entropy: How unique?, Continuum Mechanics and Thermodynamics (2003) [cond-mat/0305087], to appear.
[16] J. Mohanalin, P.K. Kalra, N. Kumar, Extraction of micro calcification using non extensive property of mammograms; IACC 2009, IEEE International Conference (2009) 636–641.