An Application of Topological Median Filters on Detection and Clustering of Microcalcifications in Digital Mammograms

İbrahim Kıvanç CİHAN

Havelsan Ehsim A.Ş. ODTÜ Teknokent Gümüş Bloklar B Blok Suit 2B 06531 Ankara, TURKEY
ikcehsim@havelsan.com.tr

Hakan Güray ŞENEL

Department of Electrical & Electronics Engineering, Anadolu University, 26470 Eskisehir, TURKEY
hgsenel@anadolu.edu.tr

ABSTRACT

Existence of microcalcification clusters on mammograms is one of the earliest signs of breast cancers. In this study, a method that is based on topological median filters is proposed for the automated detection of microcalcifications. The proposed algorithm consists of two steps. First, probable microcalcification pixels in the mammograms are segmented out by using topological top-hat transform. Then, individual microcalcifications are clustered by using a subtractive clustering algorithm. The method has been applied to Nijmegen database of 34 mammograms with a total of 72 microcalcification clusters. The results show that the proposed algorithm has a success rate of 93%.

Index Terms — Topological median filters, top-hat transform, microcalcification, subtractive clustering.

1.INTRODUCTION

Breast cancer is one of the leading factors in women mortality. A research shows that one woman in every seven who lives till 70 years old develops breast cancer [1]. Although, breast cancer can be fatal, patients have a chance to survive if the disease is diagnosed in early stages. Early diagnosis and treatment of breast cancer reduces the rate of fatality. A method to detect breast cancers in early stages is the mammography.

One of the most important signs of breast cancer is faint microcalcification clusters on mammograms. Therefore, automated detection of microcalcifications has attracted much scientific attention. To detect and classify the microcalcifications, many methods have been proposed. For references, readers may refer to [2][3][4][5].

In this research, an algorithm is developed to detect clusters of microcalcifications. The success rate of this algorithm is compared to a morphology based detection technique. Our technique is composed of a filter based on recently introduced topological median filters [6] and a fuzzy clustering method.

Topological median filters which are proposed by Senel [6] is a fuzzy topology based method for image smoothing with edge preserving properties. They are shown to be superior to conventional median filters.

In this paper, mammograms made available by University Hospital Nijmegen are used to evaluate the effectiveness of the proposed algorithm. The database contains 34 mammograms with microcalcifications clusters from 21 different patients.

2. TOPOLOGICAL MEDIAN FILTERS

Since its introduction by Tukey [7] in the 1970s, the median filter has been used extensively for image noise reduction and smoothing [6]. Conventional Median Filter (CMF) is especially good at removing impulsive noise from images with edge preserving properties. However, CMF is known to have side effects under noise. In a recently proposed method, Topological Median Filters (TMF) is shown to preserve the step edges in the presence of noise more accurately than CMFs [6]. Before giving the definition of TMF some topology based concepts must be reviewed.

2.1 Connectedness

Let Σ be an image and σ be a fuzzy subset of it. Let $\{P_0, P_1, ..., P_n\}$ where $P_0 \in Z^n$ be a set of n pixel locations and $\sigma(P)$ be a pixel value. Let ρ be any path between two pixels P and Q such that $\rho = \{P = P_0, P_1, ..., P_n = Q\}$. The strength of ρ is defined as $s_\sigma(\rho) = \min_{0 \le i \le n} \sigma(P_i)$. In other words "a path is as strong as its weakest link." In any digital image, there are a finite number of paths between any two pixels. In order to determine the degree of connectedness between two pixels, all possible paths that link the pixels must be considered [6]. The degree of connectedness (DOC) of two pixels P and Q is defined as $c_\sigma(P,Q) \equiv \max_{\rho \subset \sigma}\{s_\sigma(\rho)\}$. Thus, the degree of connectedness between two pixels is the strength of the strongest path between any two pixels. Two pixels P and Q are said to be connected in σ if and only if there exists a path $\rho^c = \{P = P_0, P_1, ..., P_n = Q\}$ where $P_i \in \sigma$ for $0 \le i \le n$, such that for every pixel $P_i \in \rho^c$ $s_\sigma(\rho) \ge \min(\sigma(P), \sigma(Q))$. In other words, two pixels are connected if the minimum value of the pixels along the strongest path that connects them is attained by one of the pixels itself.

2.2 Connectivity Map (Conn-Map)

If a pixel P is taken as the center pixel in an $(2m+1)\times(2n+1)$ rectangular window (*m>0*, *n>0*), then a map can be built indicating how the center pixel is connected to the others. We call such a map, a *connectivity map* or a *conn-map*. The conn-map of an image around a center pixel consists of pixels that are not related to the center [6]. Figure 1 shows two 7x7 windows and their conn-maps.

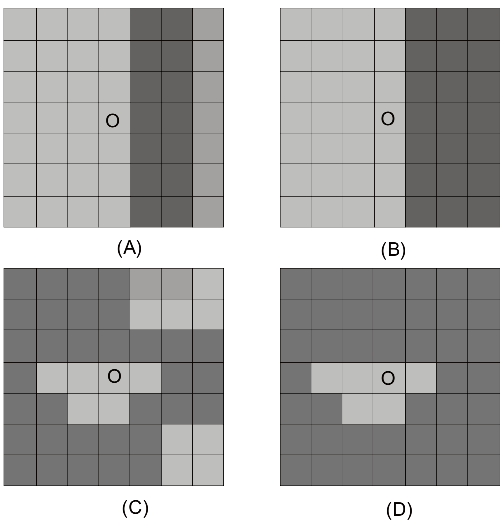

Figure 1. (a) and (c) are original image segments, (b) and (d) are the corresponding connectivity maps.

2.3 Topological Median Filters

If median operation is applied on the values of conn-maps for an image, a different type of filter, Topological Median Filter is obtained [6].

Let Σ be an image, σ be a subset of the image, *P* and *Q* be pixels and $c_\sigma(P,O)$ be the DOC between middle point O and others pixels. Topological Erosion (TE) and topological Dilation (TD) operations are defined by

$$TE(\Sigma;\sigma;O) = median_{P,O\in\Sigma}(c_\sigma(O,P)) \quad (1)$$

$$TD(\Sigma;\sigma;O) = 1 - median_{P,O\in\Sigma}(c_\sigma(O,P)) \quad (2)$$

TE operates directly on image, whereas the domain of TD is the negative image. By using TE and TD, two composite operations can be constructed, namely, topological opening (TED) and topological closing (TDE). TED is the method that is used in this research due to the fact that it operates on bright objects on a dark background, which resembles to a typical mammogram.

$$TED[\Sigma;\sigma] = TE[TD[\Sigma;\sigma];\Sigma] \quad (3)$$

$$TDE[\Sigma,\sigma] = TD[TE[\Sigma;\sigma];\Sigma] \quad (4)$$

2.4 Reconstruction

Edge and corner preserving smoothing is important during the segmentation of mammograms. Smoothing does not just remove microcalcifications alone but also some important details. Since calcifications are extracted by taking the difference between the original image and its smoothed version. If smoothing removes more image details besides calcifications, the difference image, also, contains deformed portions. In order to prevent such complications, a morphological reconstruction transformation can be used [8].

Original image (f) and smoothed image (g) are used in the reconstruction transformation. During this transformation, smoothed version of the image is dilated in an iterative way, until dilation is not possible. The reconstructed image is then subtracted from the original image to segment out the local maxima points that are eliminated by smoothing. Fig. 2 depicts the result of reconstruction. Then, a threshold is applied to convert the image into binary form.

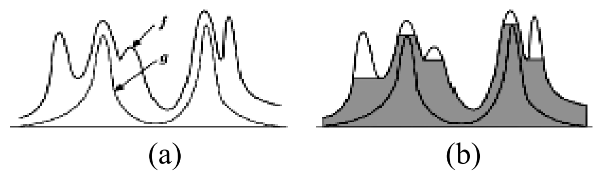

Figure 2. (a) Original image f and smoothed image g. (b) Reconstruction transformation.

2.5 Subtractive Clustering

Several methods have been developed to cluster points in black and white images [9]. Fuzzy C-Means (FCM) clustering method [11], ISODATA clustering method [10] and subtractive clustering algorithms [12] are widely used. For calcification clusters on mammograms, there is not any way to determine the number of clusters. However, for FCM and ISODATA methods, the number of iterations depends on the choice of the initial values of the cluster centers. Subtractive clustering is a fast one-pass algorithm for estimating the number of clusters and cluster centers in a set of data, if no prior knowledge of number of clusters is available.

The procedure for grouping *n* data point $\{X_1, X_2, X_3, ..., X_n\}$ clusters in the training set is described below.

1. Compute the initial potential value for each data point (X_i) by using the Eq. 5.

$$P_i = \sum_{j=1}^{n} e^{-\alpha\left\|X_i - X_j\right\|^2} \quad (5)$$

where α is a positive number and $\|.\|$ is the Euclidean distance operator. With this equation, the point with the highest number of neighbors has more potential value to be a center point than others. The point with the highest potential value is selected as the first cluster center.

2. Revise the potential values (P_k) of the remaining points (X_k) with respect to first cluster center according to Eq. (6).

$$P_k = P_k - P_{V_i} e^{-\beta \|x_k - v_i\|^2} \quad (6)$$

where β is positive number, v_i is the first cluster center and P_{v_i} is the potential value of v_i. With this operation, the potential value of the points near the cluster center is reduced. Thus, probability of being selected as center point of these points in the next iteration is minimum.

After revising the potential value measure for each point, the next cluster center is selected, and all the potential value measures are revised again. The process is repeated until the ratio of maximum potential value of the current iteration divided by the first cluster center point's potential value is equal or less then a threshold value. The threshold value has been determined to be 0.4 experimentally. α and β values are selected as 4.0.

2.6 Morphological Top-Hat Transform

In order to evaluate the effectiveness of the method proposed in this paper, a mathematical morphology based algorithm, top hat transform, is selected [13]. Top hat transform is the difference of the original image and its reconstructed opening result. The opening operation removes bright details from an image that are smaller than the size of the structuring element. Therefore, the residual image contains only those image features that have been removed by the opening operation. This operation is exactly what we did in this research. The only difference is the topological opening operation used instead of the morphological one.

3. METHOD

The block diagram of proposed algorithm is given in Figure 3. In the first stage of the proposed method, topological opening is applied to the original image to obtain a smoothed image. By using the original and smoothed images, reconstructed version of the smoothed image is formed (Fig. 4.b). A difference image between original and reconstructed image is then obtained. The difference image consists of small objects and microcalcifications. If the difference image is obtained without reconstruction transformation, there can be more features like edges and corners which do not exist in smoothed image in difference image. This situation can cause problems during the detection of microcalcifications. In the next step, small threshold value (usually between 6 and 10 depending on the contrast of the image) is applied to the difference. Last step is the clustering operation. Black and white image is considered as a set of data points and subtractive clustering algorithm applied on it. As a result, microcalcifications are clustered. Figure 5 shows the real microcalcification regions (Fig 5a-c) which are signed by doctors and microcalcification regions clustered by using our algorithm (Fig 5b-d).

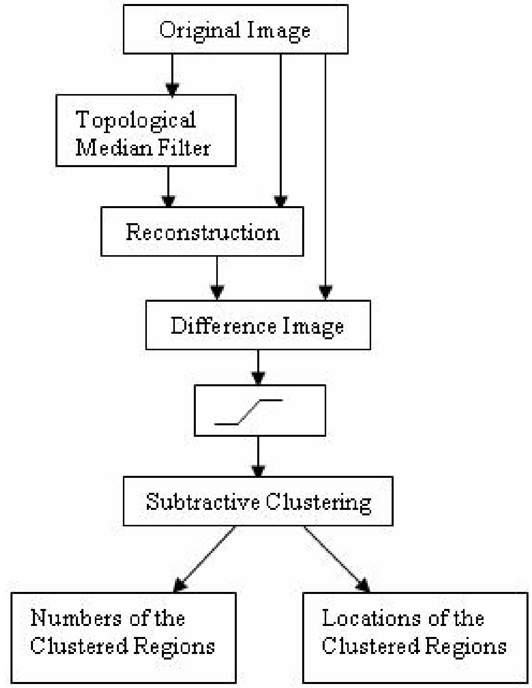

Figure 3. Block diagram of the proposed algorithm

4. RESULTS AND CONCLUSION

In this research, a method to detect and cluster the microcalcifications based on the topological median filters and subtractive clustering algorithm is proposed and comparison of the results with morphological top-hat transformation is given. During this study, a total of 34 mammograms are used to assess the performance of the proposed algorithm.

Number of false positives and the ratio of true positives to number of clusters are computed for each method and they are compared. For the topological method, the number of false positive clusters per image is computed to be 0.82. Sixty seven clusters are successfully detected in a total of 72 clusters, which corresponds to a success rate of 93%. On the other hand, for the morphology based method, the number of false positives per image is 1.18. Success rate, which is the ratio of true positives and number of positives, is 67%.

These results reveal that the topological method yields better results than the morphological top-hat transformation. This is mostly due to the fact that maximum and minimum operations used in morphology based methods are more susceptible to noise than other smoothing methods. This causes deformations that can not be restored by reconstruction. Therefore, the fitting concept of structuring element is not sufficient in order to detect clusters. With a 93% success rate, topological median filter can be effectively used to detect microcalcifications.

Some parameter values are experimentally determined. For future work, an adaptive scheme is going to be devised for those parameters.

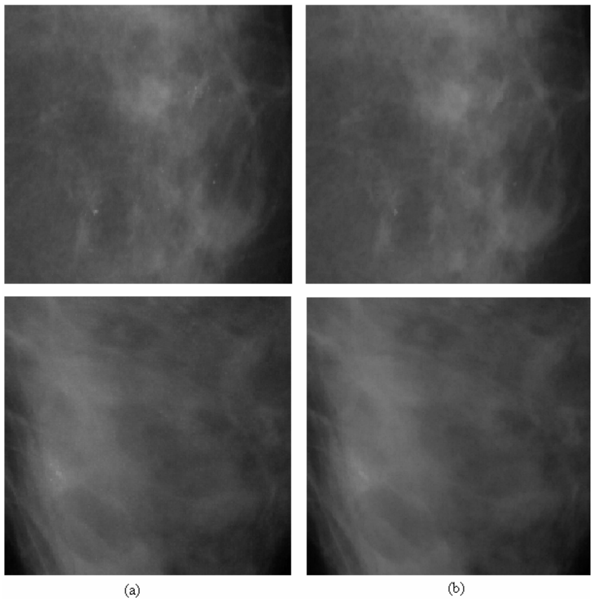

Figure 4. (a) Original (b) Reconstructed Images

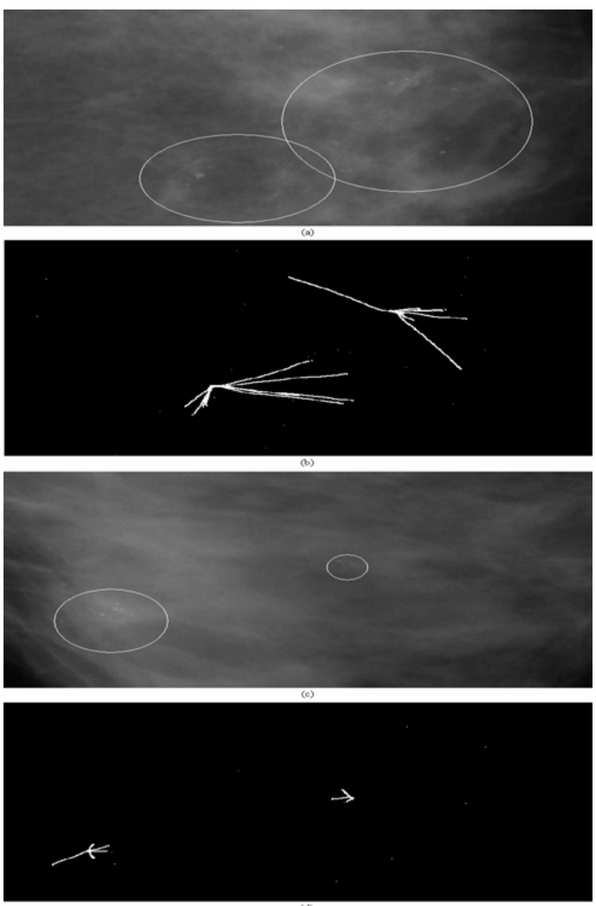

Figure 5. (a)-(c) Real Microcalcification Regions. (b)-(d) Results of Our Algorithm

5. REFERENCES

[1] Lawrence W. B., "The Radiologic Clinics of North America: Breast Imaging Current Status and Future Directions", W.B. Saunders, USA, 1992.

[2] McLeod G., Parkin G. et al "Automatic detection of clustered microcalcifications using wavelets", 3rd Int. Workshop on Digital Mammography, Chicago, 1996.

[3] Heinlein P., Drexl J., and Schneider W., "Integrated Wavelets for Enhancement of microcalcifications in digital mammography", IEEE Trans. Medical Img, 22 (3), 2003.

[4] Cheng HD, et al, "Microcalcification detection using fuzzy logic and scale space approaches", Pattern Recognition, vol. 37 (2004) pp. 363 – 375.

[5] Papadopoulosa A, et al., "An automatic microcalcification detection system based on a hybrid neural network classifier", Artifical Intelligence in Medicine, v25, 2002.

[6] Senel HG, Peters II RA., Dawant B., "*Topological Median Filters*", IEEE Trans. Image Processing, 11 (2), pp 89-104, 2002.

[7] Tukey JW, "Nonlinear (nonsuperposable) methods for smoothing data", Proc. Congr. Rec. EASCOM'74, p. 673, 1974

[8] Vincent L, "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms", IEEE Trans. Image.Proc., 2(2), 1993.

[9] Dunn JC, "A fuzzy relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", J. Cybernet. 3 (3), 1974.

[10] Bezdek JC, "A Convergence Theorem of the Fuzzy ISODATA Clustering Algorithms," IEEE Trans. PAMI., 2(1), 1980.

[11] Bezdek JC, "Fuzzy Mathematics in Pattern Classification", Ph.D. Diss., A.Math. Cornell Univ., Ithaca, NY, 1973.

[12] Chiu S, "Fuzzy Model Identification Based on Estimation," J. Intell. & Fuzzy Sys, 2(3), 1994

[13] Serra J, "Image Analysis and Mathematical Morphology", Academic Press, New York, pp. 436–437, 1982.