

Research Article

Automated Abnormal Mass Detection in the Mammogram Images Using Chebyshev Moments

¹Alireza Talebpour, ¹Dooman Arefan and ²Hamid Mohamadlou

¹Department of Radiation Medicine Engineering, Shahid Beheshti University, Tehran, Iran

²Department of Computer Science, Utah State University, Utah, US

Abstract: Breast cancer is the second leading cause of cancer mortality among women after lung cancer. Early diagnosis of this disease has a major role in its treatment. Thus the use of computer systems as a detection tool could be viewed as essential to helping with this disease. In this study a new system for automated mass detection in mammography images is presented as being more accurate and valid. After optimization of the image and extracting a better picture of the breast tissue from the image and applying log-polar transformation, Chebyshev moments can be calculated in all areas of breast tissue. Then after extracting effective features in the diagnosis of mammography images, abnormal masses, which are important for the physician and specialists, can be determined with applying the appropriate threshold. To check the system performance, images in the MIAS (Mammographic Image Analysis Society) mammogram database have been used and the results allowed us to draw a FROC (Free Response Receiver Operating Characteristic) curve. When compared the FROC curve with similar systems experts, the high ability of our system was confirmed. In this system, images of different thresholds, specifically 445, 450, 455 are processed and then put through a sensitivity analysis. The process garnered good results 100, 92 and 84%, respectively and a false positive rate per image 2.56, 0.86, 0.26, respectively have been calculated. Comparing other automatic mass detection systems, the proposed method has a few advantages over prior systems: Our process allows us to determine the amount of false positives and/or sensitivity parameters within the system. This can be determined by the importance of the detection work being done. The proposed system achieves 100% sensitivity and 2.56 false positive for every image.

Keywords: Automatic method, chebyshev moments, image processing, mammography, mass detection

INTRODUCTION

Breast cancer is the world's most common form of cancer among women after lung cancer and is the second leading cause of cancer mortality among women. The Cancer as with any disease can be best combated by prevention and avoiding growth in the early stages. It is more economical to catch diseases, such as breast cancer, in the early stages than in the more advanced stages. But it is clear that the number of groups and individuals being studied in disease prevention are much wider than those samples that are treated in more advanced stage of disease. Therefore, in detection systems and systems that are faced with a large number of samples, testing speed and accuracy are very important and influential. However, due to the tremendous amount of samples, adequate speed and accuracy by human is in most cases impossible. Therefore, the use of computer systems and artificial intelligence saves time in the calculations and the ability of specialists to respond to a large number of samples. In automatic tumor and suspicious mass area

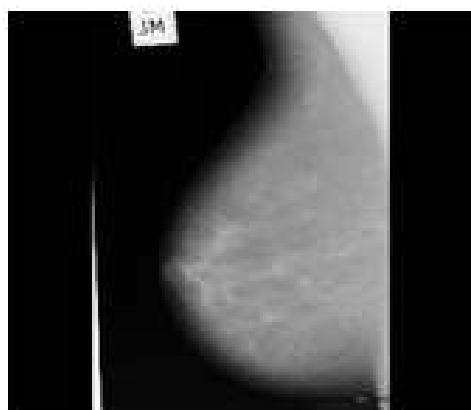
detection in mammogram images there are different methods and algorithms. The objective of this study is to improve of system parameters such as True Positive (TP), False Positive (FP) and other abilities of system than other works that has been done similar to our work. In this study, an automatic mass detection system using MATLAB software based on computed Chebyshev moments (Vibha and Priti, 2007) has been designed to extract effective features in the detection of abnormal mass areas. Then after testing several digital mammography images from the MIAS image database, system parameters such as sensitivity and the amount of False Positives (FP) have been calculated. In order to compare our work vs. other works, FROC curve has been drawn in Fig. 1.

LITERATURE REVIEW

Much works has been done to detect masses in the mammography images and they have used different methods for this purpose. Some of them are explained in this section in summery. Sampaio *et al.* (2011) has

Corresponding Author: Dooman Arefan, Department of Radiation Medicine Engineering, Shahid Beheshti University, Tehran, Iran

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).



(a)



(b)

Fig. 1: Thresholding with otsu method (a) Main image, (b) Image after thresholding with otsu method. It is an example of image thresholding after applying a median filter

applied geostatistical functions as texture signatures, Cellular Neural Networks (CNNs) and Support Vector Machine (SVM) for classification images. They have achieved TP = 80% and FP = 0.84 clusters/image. Bong-Ryul *et al.* (2010) has used breast segmentation by region growing and opening operations, decision of an initial seed with characteristics of masses and mass segmentation by a level set segmentation. A seed for mass segmentation is set based on mass scoring measure calculated by block-based variances and masked information in a sub sampled mammogram. They have achieved TP = 78% and FP = 4 clusters/image. Viet *et al.* (2009) has applied edge-based algorithm for mass detection, textural features based on Gray Level Co-occurrence Matrix (GLCM) as features for classification and artificial neural network is employed as the classifier. They have achieved TP = 85% and FP = 3.47. Alfonso and Asoke (2008) has applied multilevel thresholding segmentation and

is proposed local statistical measures of the mammograms for improving image contrast and after image enhancement, regions are segmented via thresholding at multiple levels and a set of features is computed from each of the segmented regions. They have achieved TP = 80% and Fp = 2.3 clusters/image. Qian *et al.* (2007) has used Ipsilateral Multiview CAD system, 2 wavelet based methods for image decomposition and enhancement, Tree Structured Wavelet Transform (TSWT) and Fuzzy C-Means (FCM) algorithm for suspicious region segmentation. They have achieved TP = 89.6% and FP = 1 clusters/image. Kom *et al.* (2007) they have used local adaptive thresholding filter for image classification and this algorithm has been tested on 61 images and they have achieved TP = 95.61% and Fp = 2 clusters/image. Sun *et al.* (2004) they have used adaptive Fuzzy C-Means (FCM) algorithm for segmentation, directional wavelet transform and tree structured wavelet transform. They have achieved TP = 90% and FP = 3 clusters/image. Cheng and Muiyi (2004) has used Fuzzy Neural Network (FNN) and co-occurrence matrix for feature extraction. They have achieved TP = 92% and FP = 1.33 clusters/image. Zheng and Andrew (2001) they have used Discrete Wavelet Transform (DWT), multi resolution markow random field, dogs and rabbits algorithm and other algorithm to segmentation. They have achieved TP = 97.3% and FP = 3.92 clusters/image.

METHODOLOGY

In this study, a cancerous-tissue detection system from digital mammography images has been designed and implemented using MATLAB software. In this system, digital mammography images are read by software as input and then after optimizing the image and extracting breast tissue from image, the Chebyshev moments are calculated in the entire breast tissue after applying log-polar transformation. Applying Chebyshev moments in whole image extracts some special properties of image like mass asymmetry and edge sharpness, which in turn is important for specialists and then after using an appropriate threshold, suspicious regions are determined.

Log-polar transformation: The Log-polar transformation has features which lead to simplification in data spatial exchange (Vibha and Priti, 2007):

- Maintains continuity in the images and angles can be measured
- Geometric features in images are measurable and invariant toward rotation and magnification. Linear magnification and rotation becomes linear shift along with axis

Given an image $f(x, y)$ in the Cartesian coordination (x, y) can be shown in log-polar coordination (λ, θ) as follows:

$$\exp(\lambda) = r = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}(x, y); \theta \in (-\pi, +\pi)$$

In the function $\tan x, y$ values are used to generate angle between $(-\pi, \pi)$. The drawback of applying these methods for invariant analysis is that the translation invariance is sacrificed. However, this is not appropriate; therefore, there should use another method which allows transformation independent from translation. An image will become independent from translation if the center of the mass (\bar{x}, \bar{y}) (can move to the center of image $f(x, y)$):

$$f(x + \bar{x}, y + \bar{y})$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

M_{00}, M_{10}, M_{01} are geometric moment of order 0, 1 respectively, Hu (1962) and Richard *et al.* (1993).

Chebyshev moments: Chebyshev moments, are discrete orthogonal moments (Hu, 1962). When they are applied to image with log-polar transformation, they are able to differentiate between normal and abnormal images (Dudani *et al.*, 1977). The Chebyshev moments is defined as follows Mukundan *et al.* (2001):

$$T_{pq} = \frac{1}{\rho(p, N)\rho(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} t_p(x)t_q(y)f(x, y), \quad p, q = 0, 1, \dots, N-1$$

$$\rho(p, N) = \frac{N \left(1 - \frac{1^2}{N^2}\right) \left(1 - \frac{2^2}{N^2}\right) \dots \left(1 - \frac{p^2}{N^2}\right)}{2p + 1}$$

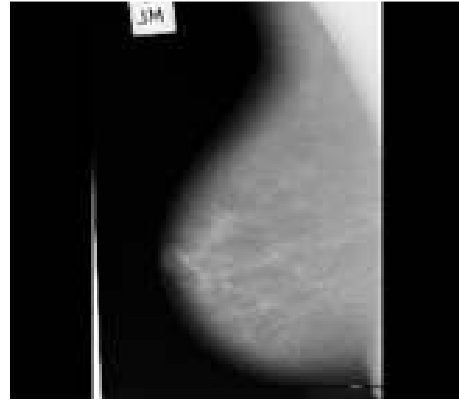
And Chebyshev moments polynomial is calculated using the below recursive equation:

$$t_0(x) = 1$$

$$t_1(x) = \frac{2x + 1 - N}{N}$$

$$t_p(x) = \frac{(2p - 1)t_1(x)t_{p-1}(x) - (p - 1) \left\{1 - \frac{(p - 1)^2}{N^2}\right\} t_{p-2}(x)}{p}, \quad p > 1$$

Suspicious areas detection in the image: To identify abnormal areas in mammography images, the image is



(a)



(b)

Fig. 2: (a) Main image (b) Extracting breast body. It shows the breast after labeling and applying maximum area logic

processed and analyzed according to flowchart indicated in Fig. 2.

Image Improvement: in the first stage, using image improvement techniques, which median filter is one of the most important of them Rafael *et al.* (2004), the images are optimized for the later stages.

Thresholding: In this study the Otsu method has been used for thresholding images and removing background (Rafael *et al.*, 2004). In the Otsu method, mathematical relationships maximize class variance, allowing the best threshold which separate background picture from image to be calculated and then after applying the threshold to the image, the image in later stages is converted into binary. Figure 3 is an example of image thresholding after applying a median filter.

Breast tissue detection: At this stage in order to detect breast tissue from other objects in the image, first the image is segmented and labeled entirely using region growing techniques (Region drawing). After labeling

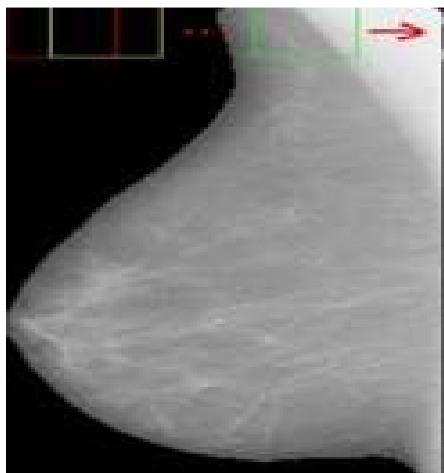


Fig. 3: Example of how to move the window to a 50×50 pixel image on a mammogram

different areas, using the maximum area logic, breast tissue is detected and its image is extracted from other areas. By combining this image with the original image,

we will have a mammography image which only includes the breast tissue. Figure 4 shows the breast after labeling and applying maximum area logic.

Restricting image: As image size increase, image analysis and calculations takes a longer time. To speed calculations, the image should be restricted as long as possible and the additional areas removed. In this study a heuristics approach based on finding borders, restricts the image as far as possible.

Applying chebyshev moments: the basis of abnormal areas detection in this study is the calculation of Chebyshev moments of order 2 (T11). To achieve this purpose T11 should be calculated and evaluated in all parts of breast tissue in the image within the center of all pixels. But to get independent images from rotation, translation and size changes which are problematic in the images analysis, before calculating Chebyshev moments in each area, the area becomes polarized (applying log-polar transformation). According to a survey conducted in the MIAS digital mammography

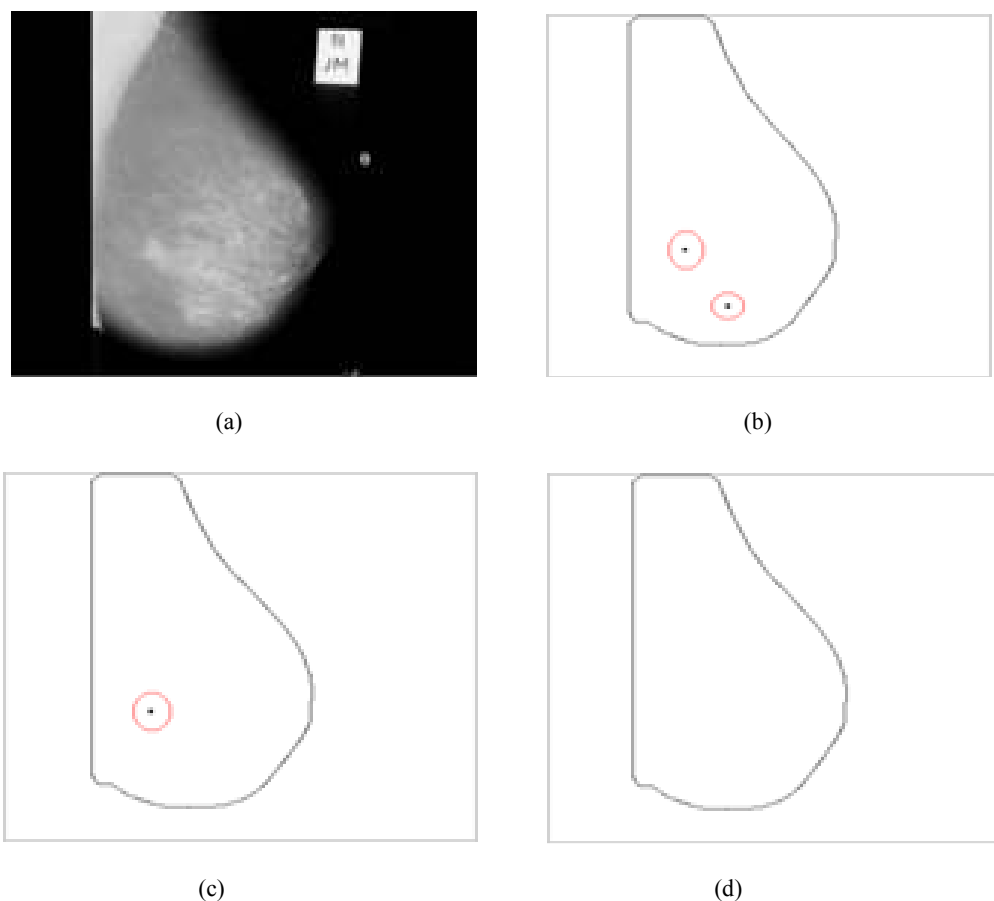


Fig. 4: Tumor detection in mammography image, (a) Original mammography image (image number 058 from MIA's database), (b) Output of system with threshold of 450, (c) Output of system with threshold of 455, (d) Output of system with threshold of 460. Depict a sample mammogram image from the MIAS database that is chebyshev moments specified for three different thresholds in abnormal areas

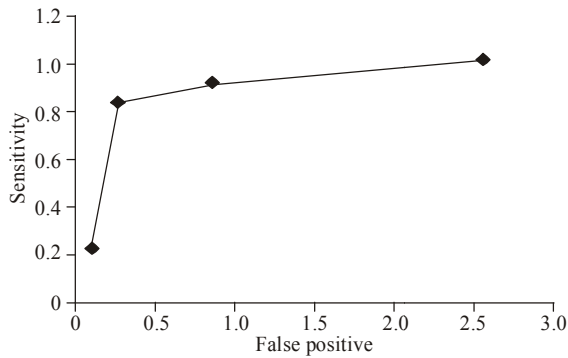


Fig. 5: Free Response Receiver Operating Characteristic (FROC) curve for 4 thresholds 445, 450, 455 and 460

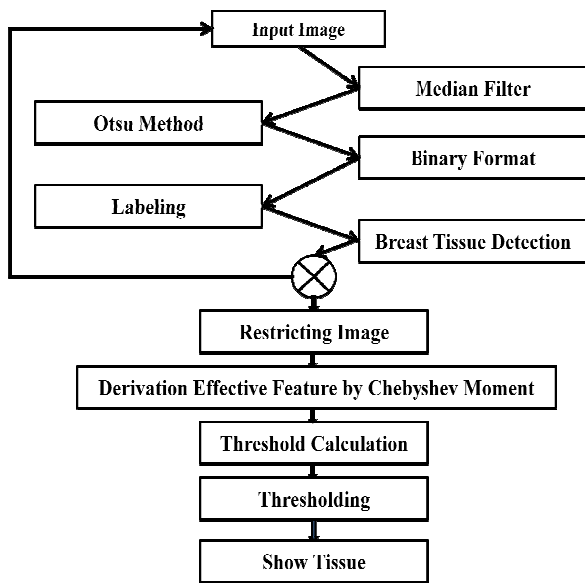


Fig. 6: Flowchart of designed system. Image analysis and abnormal mass extraction procedures are explained

images database the average size of an abnormal mass is 50 pixels in diameter. Based on these results, the value of the T11 is calculated in the window of 50×50 pixels (the best window size, 50×50 pixels, was determined by testing photos) to the center of all the pixels in the image. Figure 5 shows how this approach is applied.

Applying threshold on chebyshev moments and abnormal area detection: Using information in the images database (Livermore, CA, USA) and the results obtained from the Chebyshev moments, the best threshold for a T11 torque is obtained according to the rate of error considered to each threshold. After obtaining the best threshold for T11, abnormal areas in mammography image is automatically determined. Figure 6 depict a sample mammogram image from the MIAS database that is Chebyshev moments specified for 3 different thresholds in abnormal areas.

Table 1: Comparison of results obtained by the other works and our study

Type of system	Sensitivity (%)	FP/image
Sampaio <i>et al.</i> (2011)	80.00	0.84
Bong-ryul <i>et al.</i> (2010)	78.00	4.00
Viet <i>et al.</i> (2009)	85.00	3.47
Alfonso and Asoke (2008)	80.00	2.30
Qian <i>et al.</i> (2007)	89.60	1.00
Kom <i>et al.</i> (2007)	95.61	2.00
Sun <i>et al.</i> (2004)	90.00	3.00
Cheng and Muiyi (2004)	92.00	1.33
Zheng and Andrew (2001)	97.30	3.92
Chebyshev moments T = 455	84.00	0.20
Chebyshev moments T = 450	92.00	0.86
Chebyshev moments T = 445	100.0	2.56

RESULTS

In this study, we selected 30 random mammogram images (17 normal images and 13 abnormal images) from the MIAS images database (Heath *et al.*, 1998) and processed them using the designed system. After TP (True Positive), FP (False Positive), FN (False Negative) parameters are calculated for four thresholds of 460, 455, 450, 445, the system was evaluated.

For the threshold with a value of 460, TP equals to 3, FP equals to 3 and FN equals to 10. Considering these values, sensitivity parameter equals 0.23 and the average FP per image equals 0.1. For thresholds with the value of 455, TP equals 11, FP equals 8 and FN equals 2. Sensitivity parameters are calculated and it is equal 0.85 and the average FP per image equals 0.26. For threshold with value of 450, TP equals 12, FP equals 26 and FN equals 1. Sensitivity parameter is calculated and it is equal to 0.92 and the average FP per image equals 0.86. For threshold with value of 445, TP equals 13, FP equals 77 and FN equal 0. Sensitivity parameter is calculated and it is equal to 1.0 and the average FP per image equals 2.56. There exist different methods to identify tumor and abnormal masses in mammogram images and some work has been performed in this field. In Table 1 have been compared results of the other works and our study.

DISCUSSION AND CONCLUSION

As the results in the previous section show, analysis on images for four different amounts of threshold, 445, 450, 455, 460 with Chebyshev moments of order 2 (T11), allows abnormal areas to be calculated for each value separately. According to the results obtained by increasing the threshold from 445 to 460, the amount of sensitivity and FP is reduced per image, however a decreasing FP value is desirable but sensitivity value decreasing is undesirable. As seen in Table 1 for the threshold of 445, sensitivity and FP rate reported 1.0 and 2.56, respectively for each image.

REFERENCES

- Alfonso, R.D. and K.N. Asoke, 2008. Computerized medical imaging and graphics. *Official J. Comput. Med. Imag. Soc.*, 32(4): 304-315.
- Bong-Ryul, L., L. Jong-Doo and L. Myeong-Jin, 2010. Automated recommendation of initial mass positions for mass segmentation in digital mammograms. *International Conference on Electronics and Information Engineering (ICEIE)*, South Korea, 2: V2-202-V2-206.
- Cheng, H.D. and C. Muiyi, 2004. Mass lesion detection with a fuzzy neural network. *Pattern Recogn.*, 37: 1189-1200.
- Dudani, S.A., K.J. Breeding and R.B. McGhee, 1977. Aircraft identification by moment invariants. *IEEE Trans. Comput.*, 26: 139-145.
- Heath, M.D., K.W. Bowyer and D. Kopans, 1998. Current Status of the Digital Database for Screening Mammography. In *Digital Mammography*, Kluwer Academic Publishers, Dordrecht, pp: 457-460.
- Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theor. IT*, 3: 179-187.
- Kom, G., T. Alain and K. Martin, 2007. Automated detection of masses in mammograms by local adaptive thresholding. *Comput. Bio. Med.*, 37: 37-48.
- Mukundan, R., S.H. Ong and P.A. Lee, 2001. Discrete vs. continuous orthogonal moments for image analysis. *International Conference on Imaging Science, Systems and Technology (CISST)*, Las Vegas, Nevada Proceedings, pp: 23-29.
- Qian, W.D.S., L. Minshan, S. Ravi and E. Edward, 2007. Computer-aided mass detection based on ipsilateral multiview mammograms. *Acad. Radiol.*, 14: 530-538.
- Rafael, C.G., E.W. Richard and L.E. Steven, 2004. *Digital Image Processing Using MATLAB*. Prentice Hall, New Jersey.
- Richard, A.P., B. Maged and O. Tamara, 1993. On Computation of the Logpolar Transform. *Department of Electrical and Computer Engineering Vanderbilt University School of Engineering*, pp: 277-310.
- Sampaio, W.B., E.M. Diniz, A.C. Silva, A.C. De Paiva and M. Gattass, 2011. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. *Comput. Bio. Med.*, 41(8): 653-664.
- Sun, X.W.Q. and S. Dansheng, 2004. Ipsilateral-mammogram computer-aided detection of breast cancer. *Comput. Med. Imag. Graph.*, 28: 151-158.
- Vibha, S.V. and P.R. Priti, 2007. Malignancy texture classification in digital mammograms based on chebyshev moments and log polar transformation. *ICGST-BIME J.*, 7(1): 29-35.
- Viet, D.N., T.N. Duc, D.N. Tien and T.P. Van, 2009. An automated method to segment and classify masses in mammograms. *Int. J. Electr. Comput. Eng.*, 4: 499-505.
- Zheng, L. and K.C. Andrew, 2001. An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Trans. Med. Imag.*, 20(7): 559-567.