

Automatic detection of microcalcifications with multi-fractal spectrum

Yong Ding*, Hang Dai and Hang Zhang

Institute of VLSI Design, Zhejiang University, Hangzhou 310027, China

Abstract. For improving the detection of micro-calcifications (MCs), this paper proposes an automatic detection of MC system making use of multi-fractal spectrum in digitized mammograms. The approach of automatic detection system is based on the principle that normal tissues possess certain fractal properties which change along with the presence of MCs. In this system, multi-fractal spectrum is applied to reveal such fractal properties. By quantifying the deviations of multi-fractal spectrums between normal tissues and MCs, the system can identify MCs altering the fractal properties and finally locate the position of MCs. The performance of the proposed system is compared with the leading automatic detection systems in a mammographic image database. Experimental results demonstrate that the proposed system is statistically superior to most of the compared systems and delivers a superior performance.

Keywords: Automatic detection, mammographic calcifications, multi-fractal spectrum

1. Introduction

With the development of efficient and reliable radiological screening techniques for early detection of breast cancer, mammography plays a crucial role in the increasing breast cancer treatment options and the survival rate through mass screening. Since the detection of suspicious abnormalities is regarded as a fatiguing and time-consuming task, it is difficult for radiologists to make an accurate judgment from the enormous number of mammograms [1]. Thus, automatic detection of abnormalities systems has been proposed to help radiologists reveal breast cancer. Recent studies have demonstrated that these systems indeed achieve more accuracy and consume less time in breast cancer detection and serve as an auxiliary tool for radiologists to make final decisions [1–3].

An important symptom of breast cancer is the presence of tiny granule-like deposits of calcium named micro-calcifications (MCs) which appear on the mammogram as small bright spots and are characterized by their size, texture and high spatial frequencies [1]. Various approaches [1–3] have been applied to MC detection according to these characteristics, and this paper concentrates more on texture [4], one of the most important characteristics for MC detection.

However, the complex texture of MCs could not be quantified mathematically until Mandelbrot first introduced the concept of fractal [5]. The proposed system contains two major phases: i) Mammogram preprocessing. Some preprocessing methods of the mammogram are employed to enhance, segment

*Corresponding author: Yong Ding, Institute of VLSI Design, Zhejiang University, Hangzhou 310027, China. Tel.: 057187951071; Fax: 057187952486; E-mail: dingy@vlsi.zju.edu.cn.

and highlight the MCs according to their small size and high spatial frequencies. ii) Classification. The key to locating MCs is to create a criterion for classifying the suspicious abnormalities. Multi-fractal spectrum is a concept related to certain principal properties of fractals: self-similarity and scale-invariance. The efficient description of multi-fractal spectrum makes it possible to find out such a criterion. Evaluations in a database validate that the proposed system is statistically superior and applicable.

The rest of this paper is organized as follows. Section 2 illustrates the procedure of mammogram preprocessing. The classification method is presented in section 3. Results and discussion are rendered in section 4. At last, the conclusion is drawn in section 5.

2. Mammogram preprocessing with relevant image processing methods

The original mammogram has no salient presence of MCs resulting from their small size and other tissue masking. Direct detection is bound to consume much time, so a reasonable integration of image processing architecture illustrated in Figure 1 is applied to highlight MCs.

On small data regions rather than the entire mammogram, histogram equalization flats the peak of a histogram and redistributes the pixel value. This tuning mammogram intensity enhances the contrast of mammograms and clarifies the high spatial frequency texture for later filtering.

In order to reserve high spatial frequency information, Gaussian high-pass filter $F(x, y)$ is employed to dispose of the enhanced mammograms,

$$F(x, y) = 1 - \exp\left(-I^2(x, y)/2I_0^2\right) \quad (1)$$

where (x, y) is the pixel of mammogram, $I(x, y)$ is the threshold of high-pass filter altering with pixel, and I_0 is an average threshold. Gaussian high-pass filter facilitates the region of interest (ROI) segmentation because of the removal of unrelated information.

Sobel edge detection is adept at segmenting tiny region like MCs. The method is based on the principle that the absolute value of the image gradient is numerically determined for each pixel in the mammogram. If the absolute value of the gradient at a particular location is greater than a certain threshold, such point will be designated to a candidate edge point. Once these candidate edge points constitute a closed curve, Sobel edge detection results in the distinct boundary of MCs from other tissues.

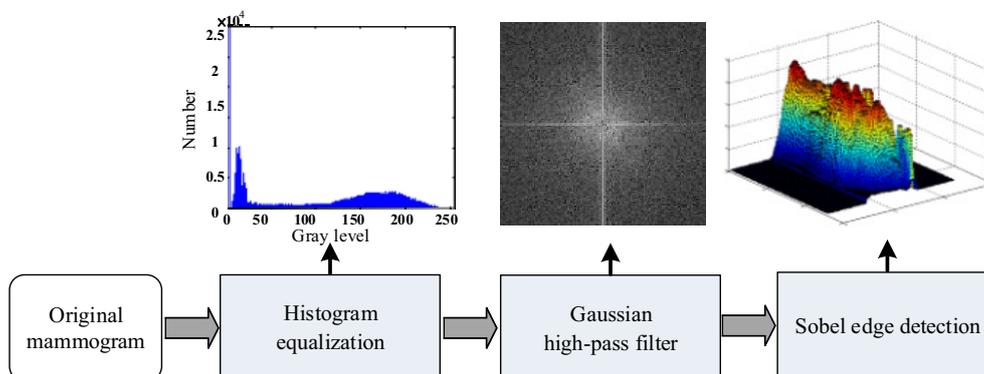


Fig. 1. Architecture of mammogram preprocessing.

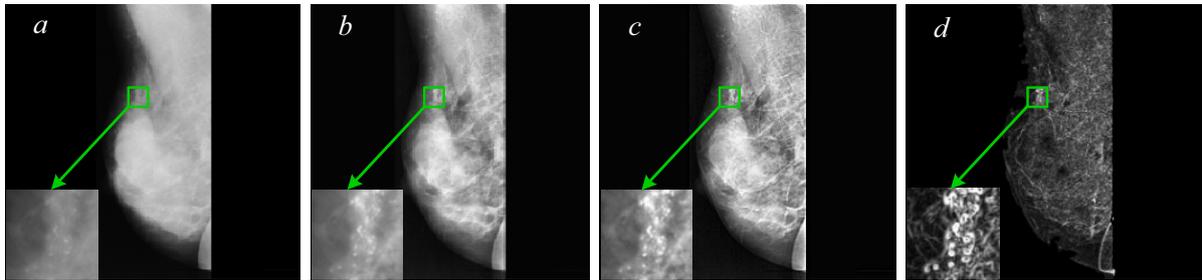


Fig. 2. Output mammograms from (a) Original mammogram, (b) Histogram equalization, (c) Gaussian high-pass filter and (d) Sobel edge detection.

The whole output mammograms are gradual in presence of MCs, the position of which are marked in Figure 2. Considering the high spatial frequencies and small size, the preprocessing methods mentioned above reveal MCs and furthermore establish a substantial foundation for locating MCs.

3. Classification based on multi-fractal spectrum comparison

The method of ROI classification plays a crucial role in finding and locating the position of MCs. In order to create a criterion that judges whether ROI contains MCs or not, the multi-fractal spectrum is employed for classification, the scheme of which is illustrated in Figure 3.

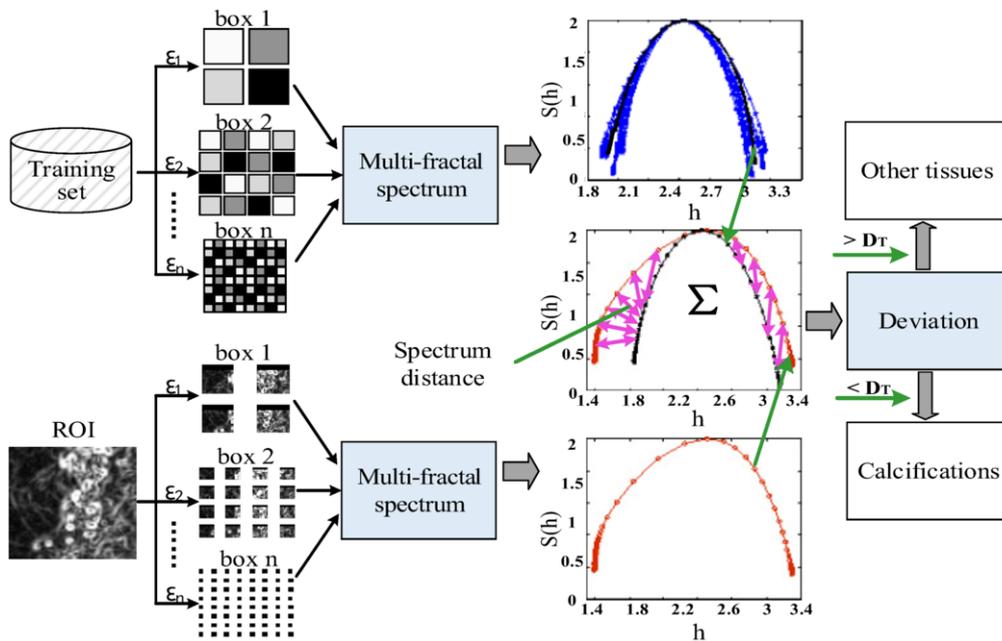


Fig. 3. Scheme of ROI classification ($D < D_T$).

MCs are highly structured and the complexity of texture is crucially important to distinguish the MCs from other textures. Texture irregularity and surface complexity is fundamental to MCs. From this perspective, fractal becomes a powerful and efficient tool to characterize MCs. In practice, complexity is measured by the box-counting method [6], which calculates multi-fractal spectrum with simplified computation.

First of all, the suspicious regions are divided recursively into some equal boxes and each step constitutes a decomposition scale. In each scale, the decomposed probability of the box is defined by,

$$P_{ij}(\varepsilon) = m_{ij} / \sum m_{ij} \quad (2)$$

where m_{ij} is the gray level of one pixel in the box and ε is the scale computed by the ratio of box size to mammogram size.

Furthermore, the partition function $\gamma_q(\varepsilon)$ is introduced as,

$$\gamma_q(\varepsilon) = \sum P_{ij}(\varepsilon)^t = \varepsilon^{\pi(t)} \quad (3)$$

where t is the moment order and $\pi(t)$ is the slope of curve $\ln\gamma_t(\varepsilon) - \ln\varepsilon$.

Finally, the multi-fractal spectrum $S(h)$, as shown in Figure 3, can be obtained by performing a Legendre transform,

$$h = d[\pi(t)] / dt, S(h) = ht - \pi(t) \quad (4)$$

where h is the singularity of probabilities.

The training set consists of MCs that are segmented from a mammographic image database. Multi-fractal spectrums from the training set seem to follow a common trend. In contrast, the deviations between the multi-fractal spectrums from ROI and the common trend can be calculated by,

$$D_i = \sum_{q=-N}^N \sqrt{[S_R(h) - S_C(h)]^2 + (h_R - h_C)^2} / (2N + 1) \quad (5)$$

where N is the number of moment order, R indicates the suspicious area and C is a common trend. From the pattern recognition theory, a threshold D_T of the deviation shown in Figure 3 must be chosen for judging whether ROI contains the calcifications or not. If the deviation is less than the D_T , the suspicious region will contain MCs, else it will contain other tissues instead of MCs. Such criterion helps locate MCs accurately by using multi-fractal spectrum and tuning the value of D_T .

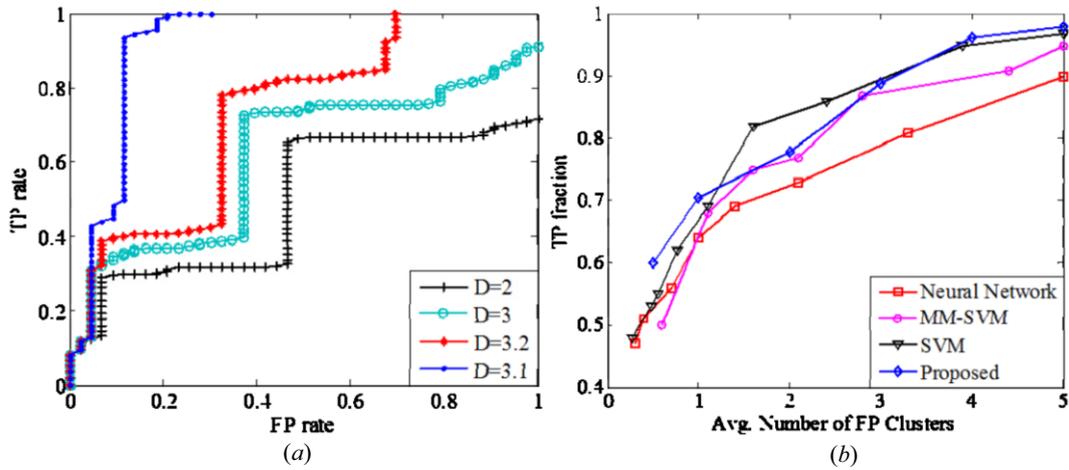


Fig. 4. (a) Receiver operating characteristic curve (b) Free-response receiver operating characteristic curve.

4. Results and discussion

The proposed system is tested in mammographic image analysis society (MIAS) database [7]. The MIAS Database has 23 mammographic images that contain MCs. In this paper, 10 images from the database are used to train the classifier and the remaining 13 images are used to evaluate the effectiveness of our proposed method. A fivefold cross-validation is performed to determine the threshold D_T of the classifier along with a receiver operating characteristic (ROC) curve [8]. And 28 positive samples with MCs and 22 negative samples are randomly into five subsets, four of which are used to train the classifier while the rest one is used for testing. The curve more approximate to true-positive (TP) rate axis shows better performance. Obviously in Figure 4(a), $D_T = 3.1$ has the relatively higher ROC curve than $D_T = 2, 3$ and 3.2 . Compared with many other candidates in ROC, $D_T = 3.1$ is selected for classification to ensure better performance.

For quantitative evaluation, free-response receiver operating characteristic (FROC) curves provide a comprehensive summary of the trade-off between detection sensitivity and specificity [8]. As can be seen in Figure 4(b), compared with Neural Network [9], SVM [10], and MM-SVM [11], FROC curve of the proposed system is obviously higher than most of other systems and it is statistically comparable to SVM, which is the most classical and practical method. The TP rate and the FP rate are shown in Table 1. Clearly, the proposed system has the highest TP rate and lowest FP rate. The evaluation validates that the proposed system is superior in detection sensitivity, specificity and accuracy.

From Table 2, one can see that the proposed system consumes less training time than relevance vector machine (RVM) and less testing time than SVM. The trade-off between training and testing time of proposed system outperforms RVM, which is the improved edition of SVM [11], resulting from proper mammogram preprocessing and simplified computation.

Table 1
Evaluation of the test results

Method	Proposed	[9]	[10]	[11]
TP	96.37%	81.7%	95.67%	94.85%
FP	1.91%	—	2.24%	2.5%

Table 2

Computation time of SVM, RVM and proposed system

	SVM	RVM	Proposed
Training time (s)	297.43	2063.20	806.32
Testing time (s)	249.33	30.04	41.05

5. Conclusion

In this paper, an innovative automatic detection system for MCs is presented and verified with MI-AS database. The proposed system integrates mammograms preprocessing for revealing the tiny MCs and multi-fractal spectrum for providing a reasonable classification criterion. As a result, the proposed system is highly accurate and timesaving in MC detection, which can serve as a helpful and reliable auxiliary tool for radiologists.

Acknowledgment

The authors would like to acknowledge the supports from national science & technology support program (2013BAH03B01) and Zhejiang provincial natural science foundation (LY14F020028).

References

- [1] Heng-Da Cheng et al., Computer-aided detection and classification of microcalcifications in mammograms: A survey, *Pattern Recognition* **36** (2003), 2967–2991
- [2] Heng-Da Cheng, Jingli Wang and Xiangjun Shi, Microcalcification detection using fuzzy logic and scale space approaches, *Pattern Recognition* **37** (2004), 363–375.
- [3] J.-Z. Cheng et al., Automated delineation of calcified vessels in mammography by tracking with uncertainty and graphical linking techniques, *IEEE Transactions on Medical Imaging* **31** (2012), 2143–2155.
- [4] Georgia D. Tourassi, David M. Delong and Carey E. Floyd Jr., A study on the computerized fractal analysis of architectural distortion in screening mammograms, *Physics in Medicine and Biology* **51** (2006), 1281–1299.
- [5] B. Mandelbr, How long is Coast of Britain, *Science* **156** (1967), 636–638.
- [6] D. A. Russell and J. D. Hanson, Dimension of strange attractors, *Physic Review Letters* **45** (1980), 1175–1178.
- [7] J. Suckling, J. Parker, D.R. Dance and S. Astley, The mammographic image analysis society digital mammogram database, *Elsevier Science* **1069** (1994), 375–378.
- [8] P.C. Bunch, J.F. Hamiton, G.K. Sanderson and A.H. Simmons, Free response approach to measurement and characterization of radiographic observer performance, *American Journal of Roentgenology* **130** (1978), 382–382.
- [9] Y. Songyang and G. Ling, A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films, *IEEE Transactions on Medical Imaging* **19** (2000), 115–126.
- [10] Liyang Wei, Yongyi Yang, R. M. Nishikawa, M. N. Wernick and A. Edwards, Relevance vector machine for automatic detection of clustered microcalcifications, *IEEE Transactions on Medical Imaging* **24** (2005), 1278–1285.
- [11] E. H. Zhang, F. Wang, Y. C. Li and X. N. Bai, Automatic detection of microcalcifications using mathematical morphology and a support vector machine, *Bio-Medical Materials and Engineering* **24** (2014), 53–59.