

Automatic Recognition of Spicules in Mammograms

Hao JIANG*, Wilson TIU**, Shinji YAMAMOTO** and Shun-ichi IISAKU*
*Communications Research Laboratory, MPT **Toyohashi University of Technology
Koganei, Tokyo, Japan Toyohashi, Aichi, Japan

Abstract: This paper presents a method of automatic recognition of spicules in mammograms. The method is consisted of two steps, enhancement and feature selection. First, spicule shadows are enhanced by using a newly developed operation. An opening operation is applied to remove noises and a direction map is made for feature selection. Second, a concentration expression is given with gray levels and two features are selected for recognition of tumors with spicules. In the method, the direction of spicules is not only considered, but the density is also utilized. The method was tested on 24 samples including seven tumors with spicules. The recognition rate for tumors with spicules was 100% without the false positives.

1 Introduction

Breast cancer is one of the leading cancers in the female population. Mammographic screening is the most effective method of early detection. Manual reading of mammograms is labor intensive and a computer system which assists experts in visual examination is required. Several approaches to analyzing mammograms by computer have been proposed in the past [1-20]. They can be loosely grouped into two categories, identifying clustered microcalcifications [1,6,15-17] and tumor detection [3,4,7-9,14,18-20].

Since the existence of spicules is one of most important visual signs for breast cancer diagnosis, there are several proposals on detection of tumors with spicules [7-9,19,20]. A spine-oriented method for the recognition of tumor with spicules has been presented in [7]. An approach to detect spicules based on analysis of local oriented edges has been given by [8,9]. A detection method of spicules based on skeleton analysis and Hough transformation has been proposed by [19]. A method to detection of stellate distortions that is based on statistical analysis of a map of pixel orientations has been also proposed in [20]. From those articles they suggested that the direction of spicules is useful information for the detection of tumors with spicules.

The objective of this paper is to propose a method to automatically recognize spicule shadows in mammograms. In this method, the direction of spicules is not only considered, but the density is also utilized for classification of tumors with spicules.

2 Method overview

Most breast carcinomas have the appearance of stellate lesions, consisting of a central tumor mass surrounded by spicules. Spicules have a star-shaped structure that emanate from the central mass with sharp, dense, and fine lines. The model of spicule can be illustrated as Fig.1 and considered as a ridge. The central tumor mass is a isolated shadow and appears as a bright blob on mammograms. The candidate shadow of central tumor mass can be detected by using the proposed method [21]. The input images for recognition of spicules are partitioned from original images based on the detected candidate location of central tumor mass. The recognition flowchart of spicules is shown in Fig.2.

The spicule recognition method we proposed consists of two steps, enhancement

of spicule shadows and feature selection. First, spicule shadows are enhanced by using a newly developed operation. An opening operation is applied to remove noises and a direction map is made for feature selection. Then, a concentration expression of spicules around central mass candidate region is given with gray levels and two features are selected based on the expression for classification.

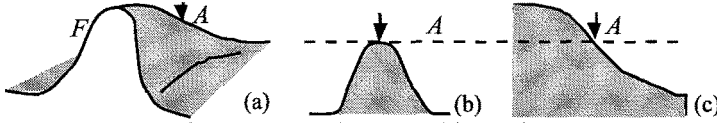


Fig.1 The model of spicule, (a) a bird view with gray levels, (b) gray levels intersecting spicule at A, (c) gray levels along spicule at A.

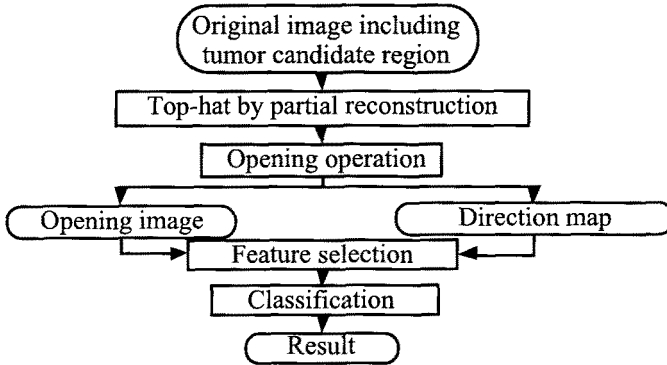


Fig.2 The recognition of spicules

3 Enhancement of spicules

A new operation, called top-hat by partial reconstruction (*TPR*), is proposed for enhancement of spicules. This operation is based on gray scale reconstruction of mathematical morphology [22]. In this section, we first review the gray scale reconstruction and provide a modified gray scale reconstruction, i.e., gray scale partial reconstruction and *TPR* operation. Then we apply an opening operation to remove noises and make a direction map of spicules for next step of classification.

3.1 Gray scale reconstruction

Two definitions of gray scale reconstruction are provided in [22]. Let F and S be two gray scale images defined on the same domain and $S(p) < F(p)$ for each pixel p . The gray scale reconstruction R can be given by

$$R(S) = \{ I_{n+1} = (I_n \oplus B) \wedge F, \text{ until } I_{n+1} = I_n \} \tag{1}$$

where $I_n \oplus B$ is the dilation of I_n by structuring element B , $I_0 = S$ and \wedge stands for pointwise minimum.

The gray scale reconstruction R of F from S is obtained by iterating gray scale geodesic dilations of S under F until stability is reached. The regional domes of image F can be extracted by subtracting the reconstructed image from original image, i.e., the components with high frequency can be effectively detected based on the gray scale reconstruction. This operation is called top-hat by reconstruction.

On the other hand, spicules are also the components with high frequency. However it is difficult to extract spicules by using the gray scale reconstruction because two dimensional flat structuring elements are used in the definitions. For example, Fig.3 illustrates gray scale reconstruction applied to spicule. At point A , it is not a peak appeared

in cross section along spicule (Fig.1(c)), although there is a peak in cross section intersecting spicule (Fig.1(b)). The top of the spicule (broken line of Fig.3(c)) can be extracted, but the shoulder with higher gradient can not be extracted and is reconstructed by iterating gray scale geodesic dilations.

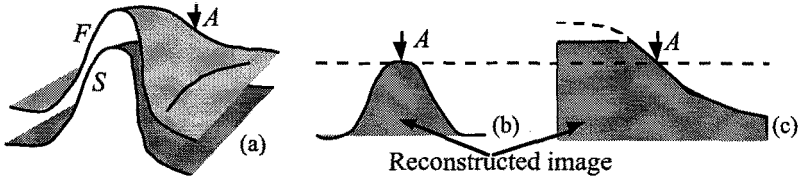


Fig.3 Gray scale reconstruction of spicule, (a) bird views of F and S with gray levels, (b) gray levels intersecting spicule at A , (c) gray levels along spicule at A .

3.2 Top-hat by partial reconstruction

Since spicules are long and slender shadows and have directions, an operation for enhancement of spicule shadows is proposed based on gray scale reconstruction that can preserve the direction, length and density of spicules. The operation is shown in Fig.4. The modified gray scale reconstruction, i.e., gray scale partial reconstruction PR , is given by

$$PR = \text{Min} \{ R_k (S, B_k) \} \quad (k = 1, 2, 3, 4) \tag{2}$$

where B_k is structuring element with direction as shown in Fig.5 and Min is minimum operation. The top-hat by partial reconstruction TPR is given by

$$TPR = F - PR. \tag{3}$$

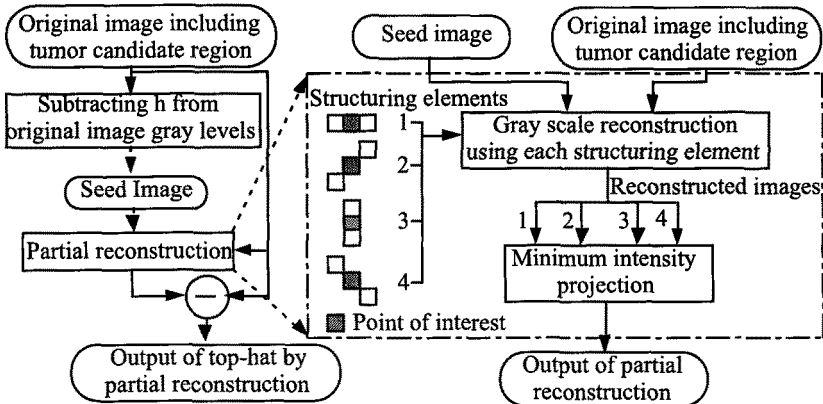


Fig.4 Top-hat by partial reconstruction

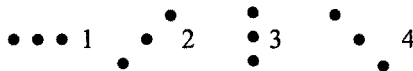


Fig.5 The structuring elements of four directions for gray scale partial reconstruction

Fig.6 illustrates the TPR operation. Let us first prepare 4 kinds of structuring element (Fig.5) and generate a seed image S by subtracting a constant h from original image F . Then carry out the gray scale reconstruction by each structuring element and operate pointwise minimum of four reconstructed images to obtain PR . Last, subtract the image of PR from original image F , the spicules are enhanced to be below the gray level h . For example, at point A of Fig.6, the output of PR operation by using the structuring element, which is parallel to the spicule, is higher than that of other three structuring elements. Original image is reconstructed (just like gray scale reconstruction

of Fig.3(c)). However, the output using the structuring element, which is vertical to the spread direction of spicule, is lower. Therefore, the spicule shadows are enhanced and the difference h is preserved by minimum and TPR operations (Fig.6).

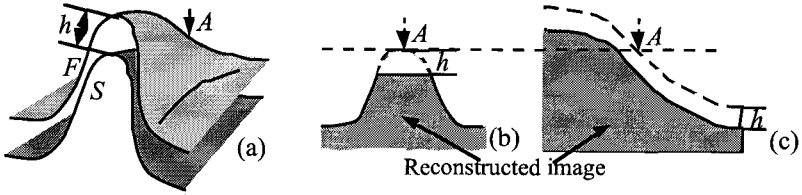


Fig.6 Gray scale partial reconstruction, (a) bird views of F and S with gray levels, (b) gray levels intersecting spicule at A , (c) gray levels along spicule at A .

3.3 Seed image selection

The seed image S is using the image subtracted from original image with constant h based on the following reasons. First, it can be automatically obtained and the details with respect to spicules are preserved. Then the seed image is included in original image and the gray levels of spicules enhanced by TPR operation are restricted to be below h . In our experiments, the h is selected based on the gray levels measured from spicules in original images.

3.4 Opening operation and direction map

After TPR operation, an opening operation is applied to suppress the noises and a direction map is made for classification of tumors with spicules. The flowchart of opening operation and direction map generation is shown in Fig.7.

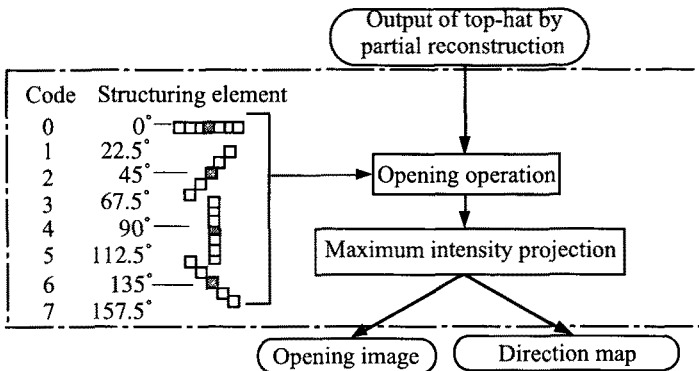


Fig.7 Opening operation and direction map

The structuring element of opening is eight bars with different direction and the length of bar is experimentally decided to remove noises and preserve the enhanced spicules. The opening operation is carried out by using each structuring element and obtained by pointwise maximum of intensities from eight images.

Different direction code is given according to each structuring element with different direction as shown in Fig.7 and the direction of spicule at each pixel is mapped based on opening operation. When maximum output of opening operation using a structuring element is obtained at a pixel among eight structuring elements, the direction code of the structuring element is considered as a direction unit vector of spicule at the pixel and recorded into the direction map (Fig.7).

4 Feature selection and classification

In this section, we give a concentration expression of spicules based on the enhanced image and direction map and select two features for classification.

4.1 Concentration expression of spicules

Let S_i express the concentration degree of spicules for a selected region i , the S_i is given by

$$S_i = \sum_{m,n} S_i(m,n) \tag{4}$$

$$S_i(m,n) = \frac{\sum_{x,y} \frac{\cos \alpha \cdot bo(x,y)}{r}}{\sum_{x,y} \frac{bo(x,y)}{r}} \tag{5}$$

where $bo(x,y)$ is the gray level of pixel (x,y) , α is the angle between line from pixel (m,n) to (x,y) and the direction unit vector at pixel (x,y) , and r is the distance between pixel (m,n) and (x,y) .

The S_i is to evaluate the concentration behavior of the radiating line pattern around a central region. The term $\cos\alpha$ is a measure of convergence from pixel (x,y) to pixel (m,n) . If it is close to 1.0, it means that direction vector points to the pixel (m,n) (Fig.8(a)). The direction map is applied to calculate the $\cos\alpha$. The term $bo(x,y)$ is the gray level of spicule at pixel (x,y) and the value is the output of opening operation. The term $1/r$ is to suppress the influence of pixel as the distance increases. The denominator is the term to normalize the degree bounded between 0 and 1.0.

4.2 Mask selection

Two masks are selected for classification of spicules. One is the concentration region of spicules and the other is the existence region. The concentration region is the domain of pixel (m,n) and the area of circle with radius r_1 from central point O of tumor candidate region, which is previously decided. The value of r_1 is experimentally decided. The discussion with respect to r_1 is also described in the following section. The existence region is the domain of point (x,y) and the bounded area of circles from radius r_1 to r_2 . The r_2 is also experimentally decided, but it is enough large to include central mass and spicules.

4.3 Classification

For suppressing the reflections of noises such as blood vessels and fibrous tissues, the existence region is not directly used to calculate the concentration degree. First, the existence region is partitioned to eight small regions R_i ($i=1, \dots, 8$) in the same size (Fig.8(b)) and the concentration degree of each region S_i ($i=1, \dots, 8$) is calculated by the equations (4) and (5) individually. Then two features are selected to classify spicules and others.

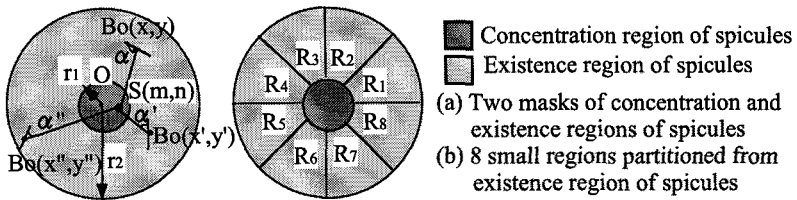


Fig.8 The masks and regions for feature selection

Let us denote the mean and the minimum of S_i by S_{mean} and S_{min} . The two features are given as

$$S_{mean} = (S_1 + S_2 + \dots + S_8) / 8 \tag{6}$$

$$S_{min} = \text{Min}(S_1, S_2, \dots, S_8) \tag{7}$$

where Min is minimum operation. It is clear that the values of S_{mean} and S_{min} are close and large for spicules and the values are small for others.

5 Experiments

5.1 Experimental data set

Mammograms used in the experiments are computed radiography (CR) images and a part of a data base published by Japan Society of Computer Aided Diagnosis of Medical Images. The spatial resolution is 0.1 mm and the size of each CR image is 2550×2100 pixels with 10-bit accuracy. 24 samples are used for evaluating the performance of the proposed method. The samples are partitioned from original image with size 512×512 based on the positions of the identified candidate regions of central tumor mass [21], i.e., 7 samples including tumors with spicules, 5 samples including tumors without spicules, 8 samples of mammary gland, and 4 samples of blood vessels.

5.2 Experimental conditions

The position of candidate region of central tumor mass is firstly detected by using the proposed method, called Quoit filtering [21]. The Quoit filtering is effective for detecting the isolated shadow like central tumor mass.

The h of generating seed image is 10 for enhancement of spicules and the value is the mean of gray levels of 10 places selected from original images with spicules. The length of bar filter used in opening operation is 20 pixels. The r_1 and r_2 of mask radius are 38 and 200 pixels, respectively. The central point O is the center of gravity of the tumor candidate region. The r_2 is decided based on the consideration that the maximum diameter of concentric circles formed by spicules is less than 40 mm. The value of r_1 is experimentally decided based on the following discussion.

The features of S_{mean} and S_{min} are calculated from samples with spicules by changing the value r_1 . The average of distance in class of tumors with spicules is calculated based on the two features. Fig.9 shows the result and there is a minimum at $r_1=38$, i.e., for the optimal classification of tumors with spicules the r_1 should be 38 pixels.

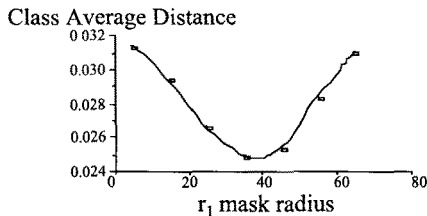


Fig.9 Relationship between the mask radius of concentration region and average distance in class.

5.3 Results

Fig.10 shows a serial images for classification based on the proposed method. Fig.10(a) is the input image and it is enhanced by TPR processing as shown in Fig.10(b). The gray levels of the image are bounded from 0 to 10. The values of spinal axes of spicules and local maxima are close to 10. The result of opening operation is Fig.10(c) and the noises such local maxima are mostly suppressed. The Fig.10(d) shows the direction code at each pixel in direction map with respect to right lower frame of Fig.10(c). The code of Fig.10(d) is corresponding on the direction code of Fig.8. The code at a pixel is the mean, if the number of times appeared with the same maximum in opening operations of eight directions is from 2 to 4. The sign - means that there is not spicule pixel and the + is given if the number of times appeared with the same maximum is more

than 5 among eight directions of opening operation. Fig.10(e) and (f) are the two masks, concentration region and existence region.

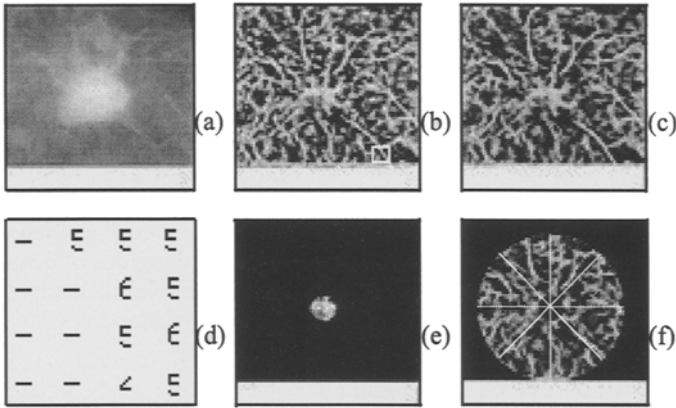


Fig.10 Serial images for recognition of spicules, (a) original image, (b) enhancement, (c) opening operation, (d) direction map, (e) concentration region of spicules, (f) partitioned existence region.

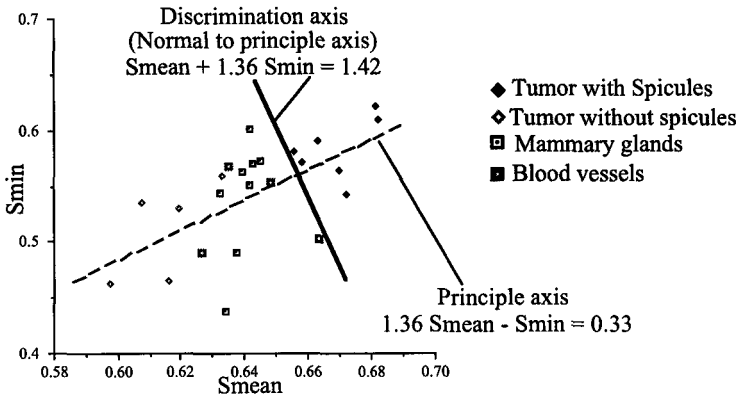


Fig.11 Classification graph

Table 1 Classification results

Input images	Classification	Tumors with spicules	The others
Tumors with spicules		7/7	0/7
The others		0/17	17/17

Experimental results applied to 24 samples are shown in Fig.11 and the axes are S_{mean} and S_{min} . Tumors with spicules and others are classified based on the following condition. The sample is classified to the set of tumors with spicules if the features S_{mean} and S_{min} are larger than 0.652 and 0.521 respectively. The Table 1 shows the classification results based on the condition above.

The discrimination function is the normal to principle axis obtained based on features and is given by

$$S_{mean} + 1.36 S_{min} = 1.42. \tag{8}$$

6 Conclusion

In this paper, a recognition method of spicules has been proposed. The spicule shadows in mammograms are first enhanced by using a newly developed operation called top-hat partial reconstruction. The operation can preserve the length of spicules. An opening operation is applied to remove noises and a direction map is made for feature selection. Second a spicule concentration expression is given and two features are selected for classification that can characterize the radiating structure of spicules. The direction of spicules is not only considered but the density is also utilized for classification of tumors with spicules. Experiments to test the performance of the proposed method have been presented. 24 samples including seven tumors with spicules have been classified perfectly without false positives. In the next step, the samples will be increased for verifying the robustness of the proposed method.

References:

- [1] W.Spiesberger: Mammogram inspection by computer, *IEEE Trans. Biomed. Engin.*, Vol.BME-26, No.4, pp.213-219, 1979
- [2] A.P.Dhawan, G.Buelloni and R.Gordon: Enhancement of mammographic features by optimal adaptive neighborhood image processing, *IEEE Trans. Med. Imaging*, Vol.MI-5, No.1, pp.8-15, 1986
- [3] S.-M.Lai, X.Li and W.F.Bischof: On techniques for detecting circumscribed masses in mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-8, No.4, pp.377-386, 1989
- [4] D.Brzakovic, X.M.Luo and P.Brzakovic: An approach to automated detection of tumors in mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-9, No.3, pp.233-241, 1990
- [5] W.M.Morrow, R.B.Paranjape, R.M.Rangayyan and J.E.L.Desautels: Region-based contrast enhancement of mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-11, No.3, pp.392-406, 1992
- [6] J.Dengler, S.Bebine and J.F.Desaga: Segmentation of microcalcifications in mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-12, No.4, pp.634-642, 1993
- [7] S.L.Ng and W.B.Bischof: Automated detection and classification of breast tumors. *Computers and Biomedical Research*, Vol.25, pp.218-237, 1992
- [8] W.P.Kegelmeyer, Jr.: Computer detection of stellate lesions in mammograms, *Proc. 92 SPIE Conf. on Biomedical Image Processing and 3-D Microscopy*, Vol.1660, pp.446-454, 1992
- [9] W.P.Kegelmeyer, Jr.: Evaluation of stellate lesion detection in a standard mammogram data set, in K.W.Bowyer and S.Astley eds: *State of the art in digital mammographic image analysis*, World Scientific, pp.262-279, 1993
- [10] W.Qian, L.P.Clark, M.Kallergi and R.A.Clark: Tree-structured nonlinear filters in digital mammography, *IEEE Trans. Med. Imaging*, Vol.MI-13, No.1, pp.25-36, 1994
- [11] L.Shen, R.M.Rangayyan and J.E.L.Desautels: Application of shape analysis to mammographic calcifications, *IEEE Trans. Med. Imaging*, Vol.MI-13, No.2, pp.263-274, 1994
- [12] R.P.Highnam, J.M.Brady and B.J.Shepstone: Computing the scatter component of mammographic images, *IEEE Trans. Med. Imaging*, Vol.MI-13, No.2, pp.301-313, 1994
- [13] A.F.Laine, S.Schuler, J.Fan and W.Huda: Mammographic feature enhancement by multiscale analysis, *IEEE Trans. Med. Imaging*, Vol.MI-13, No.4, pp.725-740, 1994
- [14] N.Petrick, H.-P.Chan, B.Sahiner and D.Wei: An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.1, pp.59-67, 1996
- [15] R.N.Strickland and H.I.Hahn: Wavelet transforms for detecting microcalcifications in mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.2, pp.218-229, 1996
- [16] A.P.Dhawan, Y.Chitre, C.Kaiser-Bonasso and M.Moskowitz: Analysis of mammographic microcalcifications using gray-level image structure features, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.3, pp.246-259, 1996
- [17] B.Zheng, W.Qian and L.P.Clark: Digital mammography: Mixed feature neural network with spectral entropy decision for detection of microcalcifications, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.5, pp.589-597, 1996
- [18] B.Sahiner, H.-P.Chan, N.Petrick, D.Wei, M.A.Helvie, D.D.Adler and M.M.Goodsitt: Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.5, pp.598-610, 1996
- [19] H.Kobatake and Y.Yoshinaga: Detection of spicules on mammograms based on skeleton analysis, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.3, pp.235-245, 1996
- [20] N.Karssemeijer G.M.te Brake: Detection of stellate distortion in mammograms, *IEEE Trans. Med. Imaging*, Vol.MI-15, No.5, pp.611-619, 1996
- [21] S.Yamamoto, M.Matsumoto, Y.Tateno, T.Iinuma and T.Matsumoto: Quoit filter -- a new filter based on mathematical morphology to extract the isolated shadow, and its application to automatic detection of lung cancer in X-ray CT, *Proc. of 13th ICPR*, Vol.II, pp.3-7, 1996
- [22] L. Vincent: Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms, *IEEE Trans. Image Processing*, Vol.IP-2, No.2, pp.176-201, 1993