



ELSEVIER

International Journal of Medical Informatics 60 (2000) 29–57

International Journal of
**Medical
Informatics**

www.elsevier.com/locate/ijmedinf

A computer-aided design mammography screening system for detection and classification of microcalcifications

San-Kan Lee ^{a,b,c}, Chien-Shun Lo ^d, Chuin-Mu Wang ^d, Pau-Choo Chung ^d,
Chein-I Chang ^{e,*}, Ching-Wen Yang ^f, Pi-Chang Hsu ^e

^a Department of Radiology, Taichung Veterans General Hospital, 40705 Taichung, Taiwan ROC

^b Department of Diagnostic Radiology, National Defense Medical Center, 100 Taipei, Taiwan ROC

^c Department of Radiology, Chung Shan Medical and Dental College, 402 Taichung, Taiwan ROC

^d Department of Electrical Engineering, National Cheng Kung University, 70101 Tainan, Taiwan ROC

^e Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering,
University of Maryland Baltimore County, Baltimore, MD 21250, USA

^f Computer Center, Taichung Veterans General Hospital, Taichung, Taiwan ROC

Received 25 October 1999; received in revised form 15 May 2000; accepted 15 May 2000

Abstract

This paper presents a prototype of a computer-aided design (CAD) diagnostic system for mammography screening to automatically detect and classify microcalcifications (MCCs) in mammograms. It comprises four modules. The first module, called the Mammogram Preprocessing Module, inputs and digitizes mammograms into 8-bit images of size 2048×2048 , extracts the breast region from the background, enhances the extracted breast and stores the processed mammograms in a data base. Since only clustered MCCs are of interest in providing a sign of breast cancer, the second module, called the MCCs Finder Module, finds and locates suspicious areas of clustered MCCs, called regions of interest (ROIs). The third module, called the MCCs Detection Module, is a real time computer automated MCCs detection system that takes as inputs the ROIs provided by the MCCs Finder Module. It uses two different window sizes to automatically extract the microcalcifications from the ROIs. It begins with a large window of size 64×64 to quickly screen mammograms to find large calcified areas, this is followed by a smaller window of size 8×8 to extract tiny, isolated microcalcifications. Finally, the fourth module, called the MCCs Classification Module, classifies the detected clustered microcalcifications into five categories according to BI-RADS (Breast Imaging Reporting and Data System) format recommended by the American College of Radiology. One advantage of the designed system is that each module is a separate component that can be individually upgraded to improve the whole system. Despite that it is still a prototype system a preliminary clinical evaluation at TaiChung Veterans General Hospital (TCVGH) has shown that the system is very flexible and can be integrated with the existing Picture Archiving and Communications System (PACS) currently implemented in the Department of Radiology at TCVGH. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

* Corresponding author. Tel.: +1-410-4553502; fax: +1-410-4553969.

E-mail address: cchang@umbc.edu (C.-I Chang).

Keywords: Classification; Computer-aided design (CAD) diagnostic system; Detection; Mammography; Microcalcifications (MCCs); Mammogram Preprocessing Module; MCCs Classification Module; MCCs Detection Module; MCCs Finder Module; Nijmegen database; Shape cognitron (S-Cognitron); TaiChung Veterans General Hospital (TCVGH)

1. Introduction

In its January 14, 1991 issue, Time Magazine published a cover story: *One American Woman in Ten Will Get Breast Cancer, Why and What Can Be Done* [1]? Since then, the rate has risen. According to recent statistics [2], one woman in every eight in the United States will develop breast cancer at some point during their lifetime. A 1987 study showed that for women whose tumors were discovered early by mammograms, the 5-year survival rate was about 82% as opposed 60% for a control group [3]. However, while eight out of ten masses detected by mammography turn out to be false alarms, this technique may miss as many as 9% of actual tumors. The cause of this may be largely attributed to the poor quality of mammograms. Even high quality film can only capture gross variations. For young women whose breasts tend to be dense, firm and cluttered with natural fibrous masses, spotting tumors using standard mammography is generally difficult. Recently, the American Cancer Society has recommended that all women between the ages of 40 and 49 should get their mammograms every 2 years and every year thereafter. If this recommendation is followed, radiologists will be required to examine 170 million new mammograms each year and the workload will be overwhelmed. In this case, radiologists must reach their verdicts in a few minutes. For example, as pointed out in [2] by Laurie Fajardo with the University of Arizona's Tucson Breast Cancer Center, each radiologist in the center must scrutinize about 75 mammograms per day with only a few revealing abnormalities. In a recent survey [3], it

was pointed out that one of the major problems expected with any mammography screening program is that it would be the requirement of interpreting a large volume of mammograms produced. Due to a shortage of trained radiologists it is difficult to maintain interest and concentration when only a small number of occasional abnormalities are encountered. With advances of computer technology, radiologists have an opportunity to improve their diagnosis with the aid of computer capabilities that can enhance image quality, select regions of interest, detect suspicious spots, help to develop interpretive tools, etc. [4,5]. Mammography is a potential area where advanced image processing can improve the odds of mammograms in detecting breast cancer early.

Microcalcifications generally present an early sign of breast cancer. Past screening studies showed that 90% of impalpable in situ ductal carcinomas and 70% of impalpable minimal carcinomas were visible as microcalcifications alone [6]. Accordingly, detecting impalpable malignant calcifications within the breast can improve survival rate of breast cancer patients. Mammography screening remains an effective technique for early breast cancer detection and cannot be replaced by other diagnostic modalities such as sonography, thermography and MRI [7]. Calcifications are tiny clustered particles and probably the smallest structures that can be identified on mammograms and are best visualized using high-resolution imaging techniques or direct radiological magnification. In order to assist radiologists in detecting such microcalcifications, developing reliable computer aided diagnostic (CAD) systems for microcalcifications detection is highly desirable. Al-

though early reports from 1960s suggested that clustered microcalcifications associated with benignancy and malignancy usually have distinct characteristics, however, more recent studies in the 1980s involving a large number of cases indicated that these characteristics do considerably overlap. As a result, most radiologists encourage biopsies, even when only 20–30% of cases proved to be cancer [8]. So, it becomes a challenge for radiologists to not only recognize the presence of these tiny particles, but also to assess the likelihood of malignancy in order to avoid unnecessary biopsies [9]. Therefore, an important issue is how to reduce the false-positive biopsy rate for mammographically detected abnormalities such that the number of mammography-generated biopsy for benign lesions can be substantially decreased.

Despite the fact that the great majority of breast calcifications found in mammograms are clearly benign, difficulties in interpretation arise primarily when both benign and malignant lesions have similar mammographic appearances. Some experienced radiologists take advantage of fine-resolution images and their interpretative expertise to replace biopsy for those calcifications unlikely to be malignant. However, for less trained radiologists, experience comes with extensive practice. One way to alleviate this dilemma is to use CAD systems as a second reader to improve the diagnostic accuracy [10]. Over the past years, many CAD algorithms and methods have been proposed for detection and segmentation of microcalcifications [11–24], such as global and local thresholding techniques, artificial neural network and wavelet approaches. However, each method has its own strengths and weaknesses. This is mainly due to the nature of mammographic characteristics and appearances. In order for a CAD systems to be diagnostically useful in detection and segmen-

tation of microcalcifications, it is very important to design and develop effective preprocessing techniques to smooth inhomogeneous background and remove structured noise that is caused by parenchyma tissues and texture variations.

In this paper, a prototype of a computer-aided design mammography screening system is proposed for detection and classification of microcalcifications (MCCs). The system is made up of four modules, each of which is designed for a particular task. The first module is called the Mammogram Preprocessing Module that takes a mammogram and digitizes it into an 8-bit image of size 2048×2048 . It then extracts the breast region from the mammogram, enhances the extracted breast image and stores the processed mammogram along with the original unprocessed mammogram in a database of the system for future reference. Since only clustered MCCs provide useful diagnostic information about malignancy, the second module is designed to find and locate suspicious clusters of MCCs and then segment these clustered MCCs from the background as regions of interest (ROIs) that will be used for further MCCs detection. So, it is called the MCCs Finder Module. As soon as ROIs are identified, they will be fed to the third module whose task is to detect MCCs in ROIs. Therefore, it is called the MCCs Detection Module and is a real-time processing system that uses two different window sizes to extract MCCs. It begins with a large window of size 64×64 to quickly screen mammograms to find large calcified areas. This is followed by a smaller window of size 8×8 to extract small and tiny MCCs. Finally, the fourth module is the MCCs Classification Module that classifies each of the segmented clustered MCCs into five categories, 'negative' (no further operation), 'benign finding' (MCCs found to be negative), 'probably benign finding' (short interval fol-

low-up suggested), ‘suspicious abnormality’ (biopsy should be considered) and ‘highly suggestive of malignancy’ (appropriate action should be taken) to represent different stages of MCCs as suggested in [25].

Despite that individual modules have been reported previously in various conferences [26–30], this paper is the first to present a prototype of a mammography screening system that integrates all the four modules as a single entity for clinical applications. One of strengths of the proposed system is its open architecture where each module is an object-oriented, plug-in component and can be upgraded individually to improve the whole CAD system. Another is that the prototype system can be included in the Picture Archiving and Communication System (PACS) currently implemented in TaiChung Veterans General Hospital (TCVGH), Taichung, Taiwan, Republic of China [31,32]. This inclusion allows radiologists to have easy access to the system and assist them in making their diagnosis. In order to evaluate the designed system, a preliminary study was conducted using the public Nijmegen database provided by the Department of Radiology at the Nijmegen University Hospital, Netherlands. These experimental results were further used to build up the TCVGH database and improve the system.

In recent reports [33,34], many commercial companies such as R2 Technology, Inc., GE, Hewlett Packard Co., MedDetect/Lockheed Martin, Siemens, Sterling Diagnostic Imaging, have invested their money and resources in developing and designing a mammography screening system for clinical applications. In particular, R2 Technology, Inc. produces a system for MCCs and mass detection, named ImageChecker M1000 that recently received approval from the US Food and Drug Administration (FDA) in June 1998 [34]. Previously, it received CE Mark certification from

the European Union. The system was developed in collaboration with the Department of Radiology at the University of Chicago. It cannot be used for diagnosis, but can only be used for MCCs and mass detection. Thus, it provides a second opinion for radiologists. This system corresponds to the third module, MCCs Detection Module, in our system. Although there is no proprietary information available about the R2 system except for their recently awarded patent [35], a series of publications by this group [17–20] suggests a lead that the system might have been developed based on the model referred to as shift-invariant artificial neural networks [35,36] that was also based on Neocognitron [37]. Coincidentally, the Shape Cognitron [38] proposed in the clustered MCCs Classification Module in our system was first envisioned in [39] and also designed based on Neocognitron and Tricognitron in [40,41]. To the authors’ best knowledge, most of such commercial efforts are devoted to the detection of MCCs. The proposed system extends its ability beyond only the MCCs detection. By integrating four modules into a fully computer-automated system. The technique developed for each module is unique in the sense that it is developed to perform a specific data processing. The Mammogram Pre-processing Module implements a block region method in conjunction with the K-means cluster-based thresholding to segment the breast tissue from a mammogram. It is then followed by the MCCs Finder Module that makes use of the blanket method to find suspicious areas for clustered MCCs. As soon as the clustered MCCs are located, the MCCs Detection Module applies a sequence of enhancement techniques and entropy-based thresholding methods to detect MCCs. Finally, the detected MCCs are classified by the MCCs Classification Module where the use of Shape Cognitron in this classification

module is a novel approach. It is particularly designed to capture geometric orientations and shape curvatures of MCCs, which are crucial features in classification. The potential and usefulness of the Shape Cognitron is further demonstrated in detection of venous beading in retinal images [42].

2. System and architecture

Due to the rapid growth in number of women suffering from breast cancer in Taiwan, the Department of Radiology in TCVGH began an initiative in 1995 to undertake a task to design a mammography screening system for the purpose of clinical trials and training. The system presented in this section is a result of a multi-year effort devoted to this development and its block diagram is depicted in Fig. 1. It is comprised of four modules, Mammogram Preprocessing

Module, MCCs Finder Module, MCCs Detection Module, and MCCs Classification Module. Each module of the system was originally developed individually in a different phase. When one module was developed, it was also evaluated by radiologists at TCVGH and other hospitals for future improvement. For example, the MCCs Detection Module was first developed [26,27], then followed by the MCCs Classification Module [28,29] and MCCs Finder Module [29,30]. Finally, the Mammogram Preprocessing Module was incorporated for data storage and future data retrieval. The structure of the system is particularly designed for flexibility. It enables users to include new modules if there is a need. In addition, each module can also be upgraded and improved individually and separately. In what follows, each module of the system is described with details in [26–30,38,39].

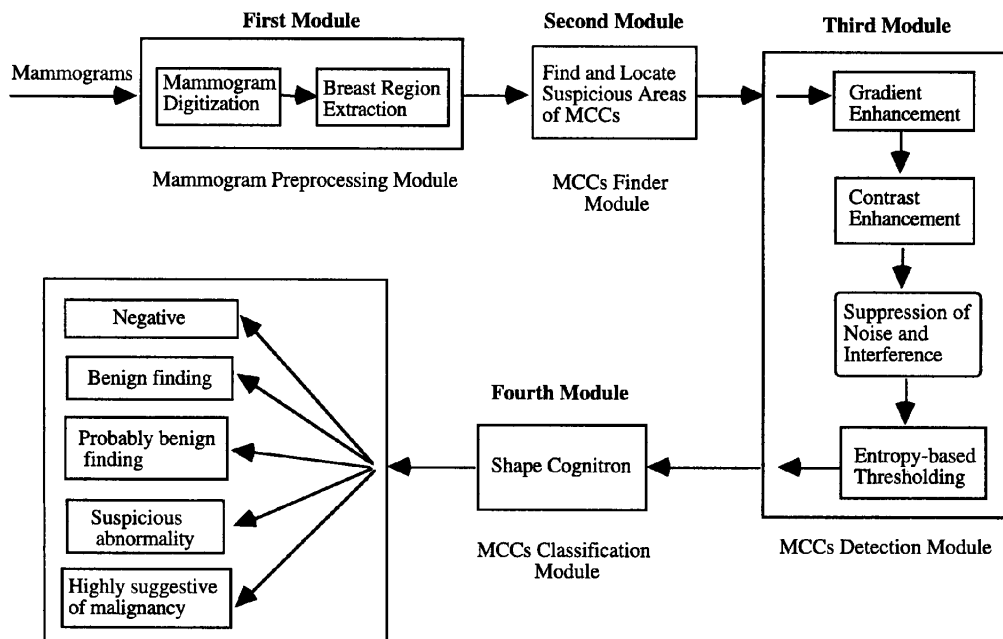


Fig. 1. A block diagram of the mammography screening system.

2.1. First module Mammogram Preprocessing Module

The purpose of the first module is to digitize mammograms and preprocess the digitized mammograms for enhancement and breast region extraction.

2.1.1. Mammogram digitization

First of all, all mammograms are digitized into an 8-bit image of size 2048×2048 by a Truvel film digitizer made by Vidar System Corporation with 260 DPI (approximately 0.1 mm/pixel).

2.1.2. Breast region extraction

The need of breast region extraction arises from several main reasons. One is that approximately or more than one-third of a mammogram is dark breast background, which provides very little information for diagnosis. Additionally, for the purpose of storage and fast retrieval this background should not be included for diagnosis. Another is computational efficiency. Since the size of a mammogram is generally 16 times as large as CT and MR images, it will be highly desirable if the dark breast background can be removed while retaining only the breast region for future data processing. More importantly, in order to make the system more efficient, extracting regions of interest (ROIs) is the first step of computer automation. In mammography screening, the breast is the region in which radiologists are interested. So, the first module of the system is designed for this purpose where a region growing method is developed to extract the breast region and is described as follows.

2.1.2.1. Breast region extraction algorithm.

1. Divide a mammogram into pq blocks of size $M \times N$, $\{B_{11}, B_{12}, \dots, B_{pq}\}$.

2. Use block region growing method to eliminate background.
 - 2.1. For each block B_{ij} , compute its mean μ_{ij} , variance σ_{ij}^2 and energy function defined by $E_{ij} = (\mu_{ij})^2 + \sigma_{ij}^2$.
 - 2.2. Begin with the block with smallest energy, then use the four-neighbor connectivity rule to grow a region with a prescribed a tolerance T_e .
 - 2.3. The region obtained by step (2.2) will be considered to be the breast background and will be eliminated.
3. Apply K-means clustering-based thresholding method to extract breast region.

After step 2, a silhouette of the breast was segmented. But it only provides an outline and a rough estimate of the breast region. In this step, a K-means clustering-based thresholding technique developed in [43] is applied to further refine and smooth the estimated breast region obtained by step 2, denoted by BREAST. Let $\text{Min}(T_e)$ and $\text{Max}(T_e)$ be the smallest and largest means of all image blocks in BREAST. We also let $\text{Mid}(T_e) = (\text{Min}(T_e) + \text{Max}(T_e))/2$ be the mid-value between $\text{Min}(T_e)$ and $\text{Max}(T_e)$. Using these three values $\text{Min}(T_e)$, $\text{Max}(T_e)$ and $\text{Mid}(T_e)$ as seeds for clustering, three classes, denoted by $\{C_0, C_1, C_2\}$ are generated with C_0, C_1, C_2 corresponding to $\text{Min}(T_e)$, $\text{Mid}(T_e)$ and $\text{Max}(T_e)$ respectively. In this case, we assume that C_2, C_1 and C_0 represent the breast region, breast boundary and breast background within BREAST respectively. Suppose that μ_i is the mean of class C_i . The desired threshold value μ to extract the breast region can be selected by $\mu = (\mu_0 + \mu_1)/2$.

In step 2 of the breast extraction algorithm, a block region growing method is used to eliminate the breast background. It

requires a seed block to start with. Two selections can be made for this purpose. One is to select a block that has the lowest average intensity. This selection is based on the assumption that the darkest image block must be part of the breast background. Unfortunately, this is not always the case with lepto-breasts in which the darkest block may occur within the breast region. Under this circumstance, using the block with averaged darkest intensity as a starting block will not work. A second selection is to pick one of four corner blocks of a mammogram as a seed block. Although a mammogram may have a transparent fringe when it is scanned, this fringe generally produces high gray level values and can be easily detected. In this case the corner on the opposite side can be selected as a seek block. In our system, the latter was used to select a seed block and the threshold T_e for region growing was set to $T_e = 15$.

2.2. Second module: MCCs Finder Module

Detection of MCCs is crucial to success in detecting early breast cancer and has been investigated extensively [11–24,26–30]. However, in most cases, the regions of interest (ROIs) for possible MCCs are preselected manually by radiologists. Furthermore, from a diagnostic point of view, only clustered MCCs are of interest. This is because single MCC blobs or sparse MCCs are generally caused by breast tissues and noises, and do not provide much useful information for diagnosis. In order to make the system fully computer automated, a second module is introduced to the system with the goal finding and locating areas consisting of possible clustered MCCs. In the following, we describe a process, referred to as blanket method, that can be used to automatically find and locate ROIs for possible MCCs.

The blanket method was originally developed by Mandelbrot [44] to estimate the length of an irregular curve and was further extended by Peleg [45] to estimation of a surface area. Let $[(x,y),I(x,y)]$ be the surface area of an object at (x,y) with the gray level $I(x,y)$. The surface area can be estimated by measuring the volume between an upper blanket, $U_r(x,y)$ defined by

$$U_r(x,y) = \max \left\{ U_{r-1}(x,y) + 1, \max_{|(s,t)-(x,y)| \leq 1} U_{r-1}(s,t) \right\} \quad (1)$$

and a lower blanket, $L_r(x,y)$ defined by

$$L_r(x,y) = \max \left\{ L_{r-1}(x,y) - 1, \min_{|(s,t)-(x,y)| \leq 1} L_{r-1}(s,t) \right\} \quad (2)$$

where $U_0(x,y) = L_0(x,y) = I(x,y)$ and r is a distance above or below the surface and is a scaling factor of the fractal dimension. Peleg defined the surface area $V(r)$ as the half of the volume increment by

$$V(r) = \frac{1}{2} \sum_{(x,y)} \{ [U_r(x,y) - U_{r-1}(x,y)] + [L_{r-1}(x,y) - L_r(x,y)] \} \quad (3)$$

Since $V(r)$ is proportional to r , it can be represented as

$$V(r) = kr^{2-D}. \quad (4)$$

Using Eq. (4) we can calculate the fractal dimension D based on $\log V(r)$ versus $\log r$ as follows.

$$D = 2 - \frac{\log V(r) - \log k}{\log r} \quad (5)$$

The volume $V(r)$ can be viewed as an image surface with the variance specified by the scaling factor r while D can be used as a measure of image texture characterization.

It should be noted from Eq. (5) that the D is not directly related to r because the texture characteristics are weakly associated with r . The magnitude of r only determines the extent of the area covered by the image texture. For example, clustered MCCs usually have high gray gradients and variances in texture. In this case, D provides an important indication of the existence of clustered MCCs. More importantly, a cluster of MCCs matches the fractal property, which is directly determined by the scaling factor r of the fractal model. Because of that, two window sizes, 64×64 and 8×8 are used in this module to detect and find suspicious areas of clustered MCCs, referred to as regions of interest (ROIs). Using the 64×64 window enables users to quickly screen mammograms to find large calcified areas, while the 8×8 window is used to detect small, isolated microcalcifications. For example, the 64×64 window using the fractal dimension D with $r = 16$ can detect the size of the area of clustered MCCs up to 128×128 pixels. However, a small lesion within a large window may sometimes fail to satisfy the fractal property. In this case, they may be passed and go undetected by a large window such as 64×64 . So, the use of a small window ensures that such small lesions will meet the fractal property and can be extracted by the fractal dimension D . Recently, the concept of the fractal dimension D was also found to be very useful for CT liver image classification where it was used to detect three different types of liver regions, normal liver, hepatoma and liver boundary [46]. For details of the application and implementation of fractal dimension we refer to reference [46].

2.3. Third module: MCCs Detection Module

The goal of developing the fractal dimension D in MCCs Finder Module is to find and locate suspicious clustered MCCs and to provide radiologists with ROIs that require their attention. But these ROIs are not necessarily all calcified. In particular, some detected pixels may be noise or breast tissues. So, the next step is to extract possible MCCs from these ROIs for diagnosis. One challenge presented to this module is that MCCs are embedded in or obscured by the inhomogeneous background within the breast. So, in order to cope with this difficulty, the MCCs Detection Module is developed for this purpose.

The techniques described below are developed based on the assumption that the gray-level intensity of calcified pixels is generally brighter than that of uncalcified pixels. However, two situations generally occur in detection of MCCs. One is that calcified pixels themselves already have low intensities. The other is that even when calcified pixels have higher intensities, their neighboring pixels may also have high intensities so that the relative contrast of these calcified pixels is significantly reduced. Therefore, as a first step the module enhances the low intensity of calcified pixels. This is followed by a step that improves the low contrast of enhanced calcified pixels. Finally, the third step is to remove suppressed undesired high intensity uncalcified pixels, particularly noise pixels using a Gaussian filter. Since the resulting Gaussian filtered images are generally gray scaled, MCCs can only be detected by visual inspection. Therefore, thresholding is necessary to produce binary images that show the locations of MCCs. The implementations of MCCs Detection Algorithm is given as follows.

2.4.1. Gradient enhancement

Let $I(x,y)$ be the gray level of the pixel at the spatial location (x,y) of a mammogram. We first find its gradient $g(x,y)$ given by

$$g(x,y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 |I(x+i,y+j) - I(x,y)|, \quad (6)$$

then add $g(x,y)$ to $I(x,y)$

$$I_1(x,y) = I(x,y) + g(x,y). \quad (7)$$

so that $I(x,y)$ is enhanced by $I_1(x,y)$.

2.4.2. Contrast enhancement

While Eq. (7) enhancing intensities of gray levels it does not necessarily increase gray level contrast, in this case a contrast enhancement technique is employed in the second stage. The idea is to use a 3×3 window to average eight-neighbor connectivity pixels of every pixel to reduce the intensities of uncalcified pixels so that the contrast can be increased. This averaging processing can be repeatedly applied as many times as described in Eq. (8) until a desired outcome is achieved.

$$\mu^m(x,y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 \mu^{m-1}(x+i,y+j) \quad (8)$$

The superscript m in Eq. (8) indicates the number of times of the averaging processing was performed. The contrast enhancement process scales the mean calculated in Eq. (8) by $L-1$ where L is the highest gray level used in the mammogram, then it is multiplied to the left with the image produced by the first stage preprocessing, $I_1(x,y)$. The resultant image $I_2(x,y)$ is given by

$$I_2(x,y) = \frac{\mu^m(x,y)}{L-1} I_1(x,y). \quad (9)$$

2.4.3. Suppression of interference and noise

In this stage, a Gaussian filter is applied to eliminate noisy and interfering pixels that can be caused by breast tissues. It is done by subtracting a Gaussian-filtered image, $G_\sigma \otimes I_2(x,y)$ from $I_2(x,y)$, i.e.

$$I_3(x,y) = I_2(x,y) - G_\sigma \otimes I_2(x,y) \quad (10)$$

where G_σ is a Gaussian filter with S.D. σ and \otimes is the convolution operator.

2.4.4. Segmentation of MCCs from the background by entropy-based thresholding

After step 3 is completed, the image quality of the mammograms is enhanced. However, one problem remains. The processed mammograms are gray scale and still require visual inspection of radiologists. In order for a system to be fully computer automated and not require human intervention, an automatic segmentation method is needed to segment detected MCCs from the background. This step is crucial in detection of MCCs. Many thresholding techniques have been proposed in the literature [47]. In this module, the thresholding methods used in this step are entropy-based thresholding methods [48–50], which were shown to perform better in terms of capturing the characteristics of MCCs than other popular thresholding methods such as Otsu's method [51].

Entropic thresholding is a technique that adopts entropy as a criterion to threshold an image. The concept of entropy has been widely used in data compression to measure information content of an information source. Suppose that an L -symbol source X is governed by a probability distribution $\mathbf{p} = (p_1, \dots, p_L)$. The information generated by the source X can be described by its entropy $H(X)$ defined as follows.

$$H(X) = - \sum_{j=1}^L p_j \log p_j. \quad (11)$$

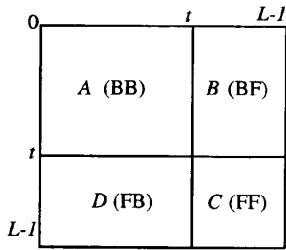


Fig. 2. Four quadrants of a co-occurrence matrix divided by a threshold t .

Since an image can be viewed as an information source with the probability distribution given by its gray-level image histogram, the information contained in the image can be characterized by the entropy of the histogram. Pun and Kapur et al. used this concept to derive entropic thresholding methods.

Given a digital image of size $M \times N$ with L gray levels denoted by $G = \{0, 1, \dots, L-1\}$, let $f(x,y)$ be the gray level of the pixel at the spatial location (x,y) . Then the image can be represented by a matrix $F = [f(x,y)]_{M \times N}$. A co-occurrence matrix of an image is an $L \times L$ square matrix, denoted by $W = [t_{ij}]_{L \times L}$ whose elements are specified by the numbers of transitions between all pairs of gray levels in $G = \{0, 1, \dots, L-1\}$ in a particular way. A widely used co-occurrence matrix is an asymmetric matrix, which only considers the gray level transitions between two adjacent pixels. More specifically, let t_{ij} be the (i, j) th element of the co-occurrence matrix W . Following the definition given in [48,49],

$$t_{ij} = \sum_{m=1}^M \sum_{n=1}^N \delta_{mn} \quad (12)$$

where

$$\delta_{mn} = 1 \text{ if } \begin{cases} f(m,n) = i \text{ and } f(m+1,n) = j \\ \text{and/or} \\ f(m,n) = i \text{ and } f(m,n+1) = j \\ = 0; \text{ otherwise} \end{cases}$$

It should be noted that the co-occurrence matrix defined above considers only the pixels on the right and bottom transitions since it was found that including the pixels on the left and top transitions does not provide significant information and improvement. By normalizing the total number of transitions in the co-occurrence matrix, a desired transition probability from gray level i to j is obtained by

$$p_{ij} = \frac{t_{ij}}{\sum_{k=0}^{L-1} \sum_{l=0}^{L-1} t_{kl}} \quad (14)$$

Assume that t is a threshold used to threshold an image. Then t partitions the co-occurrence matrix defined by Eq. (13) into four quadrants, namely, A , B , C , and D , shown in Fig. 2. These four quadrants can be further grouped into two classes. If we assume that pixels with gray levels above the threshold are assigned to the foreground (objects), and those equal to or below the threshold are assigned to the background quadrants A and C correspond to local transitions within background and foreground respectively, whereas quadrants B and D represent transitions across boundaries between background and foreground. The probabilities associated with each quadrant are then given by

$$\begin{aligned} P_A^t &= \sum_{i=0}^t \sum_{j=0}^t p_{ij}, & P_B^t &= \sum_{i=0}^t \sum_{j=t+1}^{L-1} p_{ij}, \\ P_C^t &= \sum_{i=t+1}^{L-1} \sum_{j=0}^t p_{ij}, & P_D^t &= \sum_{i=t+1}^{L-1} \sum_{j=t+1}^{L-1} p_{ij} \end{aligned} \quad (15)$$

The probabilities in each quadrant can be further obtained by so called 'cell probabilities',

$$p'_{ij|A} = \frac{p_{ij}}{P'_A}, p'_{ij|B} = \frac{p_{ij}}{P'_B}, p'_{ij|C} = \frac{p_{ij}}{P'_C}, p'_{ij|D} = \frac{p_{ij}}{P'_D} \quad (16)$$

which are probabilities conditioning on a specific quadrant.

Three definitions can be derived based on the cell probabilities given by Eqs. (4) and (5), each of which yields a different method.

2.4.5. Local entropy (LE)

Since quadrant *A* and quadrant *C* contain the local transitions from background to background (BB) and objects to objects (FF) respectively, the local entropy of BB, denoted by $H_{BB}(t)$ and local entropy of FF, denoted by $H_{FF}(t)$ can be defined respectively as follows.

$$H_{BB}(t) = - \sum_{i=0}^t \sum_{j=0}^t p'_{ij|A} \log p'_{ij|A} \quad (17)$$

$$H_{FF}(t) = - \sum_{i=t+1}^{L-1} \sum_{j=t+1}^{L-1} p'_{ij|C} \log p'_{ij|C} \quad (18)$$

By summing up the local within-class transition entropies of the foreground and the background, a second-order local entropy, denoted by $H_{LE}(t)$ was derived by N.R. Pal and S.K. Pal [48]

$$H_{LE}(t) = H_{BB}(t) + H_{FF}(t). \quad (19)$$

The LE method proposed by Pal and Pal is the one that selects a threshold, t_{LE} maximizing $H_{LE}(t)$ defined by Eq. (19) over t , namely,

$$t_{LE} = \arg \left\{ \max_{t \in G = \{0,1,\dots,L-1\}} H_{LE}(t) \right\} \quad (20)$$

2.4.6. Joint entropy (JE)

Alternatively, quadrant *B* and quadrant *D* provides edge information about transitions from background to foreground (BF) and foreground to background (FB). In analogy with LE, another second-order entropy method, called joint entropy (JE) was also derived by N.R. Pal and S.K. Pal by finding

$H_{JE}(t)$, the sum of the entropy $H_{BF}(t)$ resulting from quadrant *B*, and the entropy $H_{FB}(t)$ from quadrant *D*, which are defined as follows.

$$H_{BF}(t) = - \sum_{i=t+1}^{L-1} \sum_{j=0}^t p'_{ij|B} \log p'_{ij|B} \quad (21)$$

$$H_{FB}(t) = - \sum_{i=0}^t \sum_{j=t+1}^{L-1} p'_{ij|D} \log p'_{ij|D} \quad (22)$$

$$H_{JE}(t) = H_{BF}(t) + H_{FB}(t). \quad (23)$$

Similarly, a method finding t_{JE}

$$t_{JE} = \arg \left\{ \max_{t \in G = \{0,1,\dots,L-1\}} H_{JE}(t) \right\} \quad (24)$$

that the maximizes $H_{JE}(t)$ defined by Eq. (24) over t is called JE method, which is the second algorithm developed by N.R. Pal and S.K. Pal [48].

2.4.7. Global entropy (GE)

The global entropy $H_{GE}(t)$ is simply the sum of the local entropy $H_{LE}(t)$ and the joint entropy $H_{JE}(t)$, i.e.

$$\begin{aligned} H_{GE}(t) &= H_{LE}(t) + H_{JE}(t) \\ &= H_{BB}(t) + H_{FF}(t) + H_{BF}(t) + H_{FB}(t). \end{aligned} \quad (25)$$

A value, t_{GE} that maximizes $H_{GE}(t)$ defined by Eq. (25) over t ,

$$t_{GE} = \arg \left\{ \max_{t \in G = \{0,1,\dots,L-1\}} H_{GE}(t) \right\} \quad (26)$$

is called the global entropy threshold.

The three entropic thresholding methods described above do not take into account the thresholded images. Relative entropy-based thresholding described below takes care of this issue. Let two sources with L symbols be described by probability distributions $\mathbf{p} = (p_1, \dots, p_L)$ and $\mathbf{h} = (h_1, \dots, h_L)$ respectively. The relative entropy between \mathbf{p} and \mathbf{h} (or the entropy of \mathbf{h} relative to \mathbf{p}) denoted by $J(\mathbf{p};\mathbf{h})$ is defined by [52]

$$J(\mathbf{p};\mathbf{h}) = \sum_{j=0}^{L-1} p_j \log \frac{p_j}{h_j} \quad (27)$$

So, the smaller the relative entropy $J(\mathbf{p};\mathbf{h})$, the more closer the two sources in terms of their probability distributions \mathbf{p} and \mathbf{h} .

Now if we assume that $\{p_{ij}\}$ and $\{h_{ij}^t\}$ represent the co-occurrence matrices of an original image and an image thresholded by t , respectively. Then the relative entropy between $\{p_{ij}\}$ and h_{ij}^t is defined by

$$J(\{p_{ij}\};\{h_{ij}^t\}) = \sum_{i,j} p_{ij} \log \frac{p_{ij}}{h_{ij}^t} \quad (28)$$

By minimizing Eq. (28) over the threshold value $t \in \{0, 1, \dots, L-1\}$ three relative entropic thresholding methods similar to the above entropic thresholding methods, called local relative entropy (LRE), joint relative entropy (JRE) and global relative entropy (GRE) can be also derived as follows.

$$J_{LRE}(\{p_{ij}\};\{h_{ij}^t\}) = \sum_{(i,j) \in BB \cup FF} p_{ij} \log \frac{p_{ij}}{h_{ij}^t} \quad (29)$$

$$J_{JRE}(\{p_{ij}\};\{h_{ij}^t\}) = \sum_{(i,j) \in BF \cup FB} p_{ij} \log \frac{p_{ij}}{h_{ij}^t} \quad (30)$$

$$J_{GRE}(\{p_{ij}\};\{h_{ij}^t\}) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} \log \frac{p_{ij}}{h_{ij}^t} \quad (31)$$

The LRE, JRE and GRE are thresholding methods that find thresholds minimizing Eqs. (29)–(31) respectively. Namely,

$$t_{LRE} = \arg \left\{ \min_{t \in G = \{0,1,\dots,L-1\}} J_{LRE}(\{p_{ij}\};\{h_{ij}^t\}) \right\} \quad (32)$$

$$t_{JRE} = \arg \left\{ \min_{t \in G = \{0,1,\dots,L-1\}} J_{JRE}(\{p_{ij}\};\{h_{ij}^t\}) \right\} \quad (33)$$

$$t_{GRE} = \arg \left\{ \min_{t \in G = \{0,1,\dots,L-1\}} J_{GRE}(\{p_{ij}\};\{h_{ij}^t\}) \right\}. \quad (34)$$

These three relative entropies can be viewed as the counterparts of LE, JE and GE

in entropic thresholding. Nevertheless, both entropic thresholding and relative entropic thresholding perform differently and have different advantages. So, in order to allow radiologists to make various comparison among different thresholding methods, this module includes three entropic thresholding methods (LE,JE,GE), three relative entropic thresholding methods (LRE,JRE,GRE), one popular threshold method (Otsu's method [51]) plus a manual threshold adjustment (that allows radiologists to be able to manually adjust the threshold value themselves). For detailed treatments of these methods we refer to references [48–50].

2.5. Fourth module: MCCs Classification Module

After clustered MCCs have been identified by the MCCs Detection Module, the follow-up task is classification of the segmented clustered MCCs. In this section, we introduce a classification module, which is based on a neural network-like shape recognition system, called Shape Cognitron (SC). The SC was previously developed for classification of MCCs in [27,39] and later used for venous beading detection in retinal images [42]. The structure of SC is given by Fig. 3. It has an input unit U_0 , two simple layer-complex layers combined units, referred to as (S_1, C_1) (called shape orientation extraction unit) and (S_2, C_2) (called shape feature extraction-classification unit), and one 3-D figure layer, called shape orientation display layer lying between (S_1, C_1) and (S_2, C_2) units. Each of the units is briefly described as follows. We refer details to refs. [38,42].

2.5.1. Shape Cognitron (SC)

Shape Cognitron (SC) is derived from Tricognitron and Fukushima's Neocognitron. It was particularly designed to classify clus-

tered microcalcifications into malignancy and benignancy using a set of shape features it generates. It is known that malignant clustered microcalcifications generally have irregular shapes as opposed to round shape or egg-shaped benign clustered microcalcifications. SC captures the shape curvatures of clustered microcalcifications and provides a crucial indication of malignancy.

The SC is a neural network-like system and consists of two major components, each of which has two layers, called simple layer and complex layer and a mid-layer between them, called 3-D figure layer. The first component is similar to that used in Neocognitron, but it uses 20 orientation spatial patterns to specify eight degree spatial patterns, $0 = 360^\circ$: East (E), 45° : Northeast (NE), 90° : North (N), 135° : Northwest (NW), 180° : (West), 225° : Southwest (SW), 270° : South (S), 335° : Southeast (SE). It implements a two-layer operation, a simple layer denoted by layer S_1 and a complex layer denoted by layer C_1 . Layer S_1 contains 20 cell planes of size $N \times N$ resulting from 20 orientation spatial patterns

operating on the input pattern. Layer C_1 contains eight cell planes of size $N \times N$ obtained by merging 20 cell planes in layer S_1 that represent 20 orientation spatial patterns. They result from different weight assignments generated by two masking processes using the 20 orientation spatial patterns in layer S_1 and a particularly designed merging procedure in layer C_1 . The layer following the first component is 3-D figure layer. It is a feature extraction-display layer that extracts and stores the information of the shape orientations of an input pattern in the third dimension. It displays the input pattern as a 3-D figure using the numeric values generated in layer C_1 as the elevation of the pattern to represent eight different degrees in the third dimension. The second component can be viewed as a joint feature selection and classification system that is made up of a feature selection layer, layer S_2 , which generates a desired set of shape features from the 3-D figure layer and a classification layer, layer C_2 , which employs a probabilistic neural network (PNN) [53] as a classifier with the shape features produced

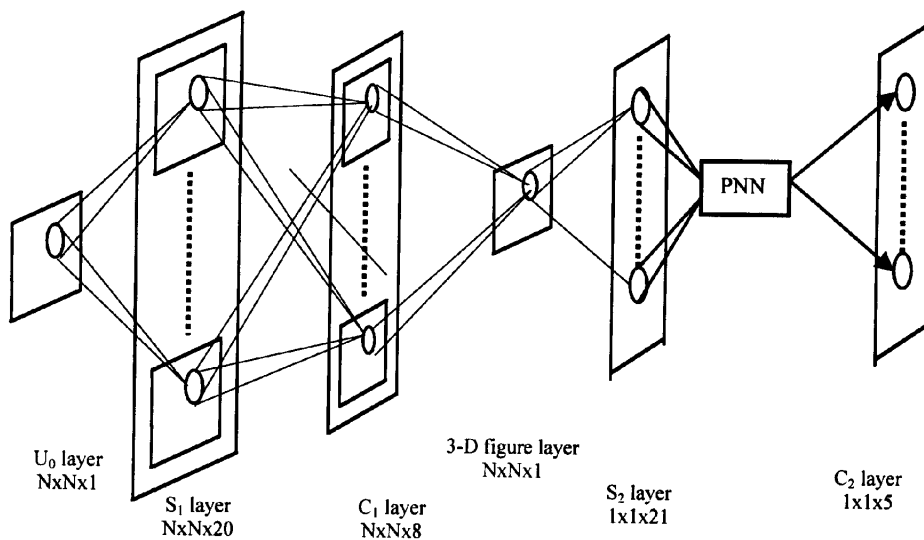


Fig. 3. Structure of shape cognition.





























by layer S_2 as inputs. The details can be summarized as follows.

The input unit U_0 takes an input pattern of size $N \times N$, which may vary with applications. In our case, the input patterns are clustered MCCs of size 256×256 produced by the MCCs Detection Module. The (S_1, C_1) unit contains a simple layer S_1 followed by a complex layer C_1 . It is a shape information extraction unit that extracts geometric shape orientations of an input pattern, then converts them into numeric values for computer processing. The idea of using (S_1, C_1) is similar to that used in Neocognitron proposed in [37] and Tricognitron in [40,41]. Layer S_1 uses a set of 20 orientation spatial patterns (eight 2×2 spatial patterns numbered from 1 to 8 and 12 3×3 spatial patterns numbered from 9 to 20) shown in the first column of Fig. 4(a) operating on the input pattern to capture eight degree spatial patterns as shown in the top row of Fig. 4(a) that represent eight different degrees: $< 45^\circ$ (i.e. degrees less than $< 45^\circ$), 45° , 90° , 135° , 180° , 225° , 270° and $0 = 360^\circ$ in the second top row. The pixel labeled by 'x' is the seed pixel currently being examined during a masking process. The first 12 orientation patterns in layer S_1 of Fig. 4(a) are designed to extract eight different degree spatial patterns that correspond to multiples of 45° , 45° , 90° , 135° , 180° , 225° , 270° , 335° and 360° . If there is a match between an orientation spatial pattern and a degree spatial pattern, a weight '1' is assigned; a weight '0', otherwise. For example, pattern 1 specifies the East orientation. It matches all degree spatial patterns except those representing degrees $< 45^\circ$ and 45° . As a result, in the row of pattern 1, there are two '0's and six '1's. The next eight orientation spatial patterns in layer S_1 from pattern 13 to pattern 20 are designed to extract degrees 167.5° , 22.5° , 202.5° , 247.5° , 67.5° , 112.5° , 292.5° and 337.5° respectively. Since these

eight patterns describe smaller degrees that are multiples of 22.5° but not multiples of 45° that are already specified by the first 12 orientation patterns, they are crucial to measure subtle differences among geometric shape features. So, if there is a match, a weight '2' will be assigned; a weight '0', otherwise. For example, pattern 13 matches degree spatial patterns, 225° , 270° , 335° . So, there are three '2's appearing in the row of pattern 13 under the columns of degrees, 225° , 270° , 335° . As a result, a total of 20 256×256 cell planes are produced in layer S_1 , each of which represents a specific orientation spatial pattern with assigned weights given in Fig. 4(a). These 20 cell planes will be then input to the next complex layer C_1 .

The task of layer C_1 is to fuse the shape orientation information produced by Fig. 4(a) by merging the 20 cell planes generated in layer S_1 into a set of eight cell planes so that each cell plane in layer C_1 represents the information generated by one specific orientation spatial pattern that corresponds to one of eight degrees, $< 45^\circ$, 45° , 90° , 135° , 180° , 225° , 270° and 360° . Fig. 4(b) shows how the eight orientation spatial patterns in layer C_1 are generated by merging the 20 orientation patterns in layer S_1 where a merge is described by a \oplus . It should be noted that the weight of a merged pattern is not obtained by summing all the weights of merging patterns. Instead, we adopt a rule suggested in [37,41] that when two patterns are merged, the higher weight of the pattern will be assigned to the weight of the resulting pattern obtained by merging these two patterns. Fig. 4(c) shows the weights obtained for these eight orientation spatial patterns in layer C_1 resulting from the merging process shown in Fig. 4(b). For example, pattern 1 in layer C_1 in Fig. 4(b) is a result of merging six patterns 1–4, 14–15 in layer S_1 shown in the first row in Fig. 4(b). Therefore, the weight of the

8 degree spatial patterns

								
20 Orientation spatial patterns in layer S_1	< 45°	45°	90°	135°	180°	225°	270°	360°
1 	0	0	1	1	1	1	1	1
2 	0	0	0	0	1	1	1	1
3 	0	0	0	0	0	0	0	1
4 	0	0	0	1	1	1	1	1
5 	0	0	0	0	0	0	1	1
6 	1	1	1	1	1	1	1	1
7 	0	0	0	0	0	1	1	1
8 	0	1	1	1	1	1	1	1
9 	0	0	0	0	1	1	1	1
10 	0	0	0	0	0	0	1	1
11 	0	0	0	0	0	1	1	1
12 	0	0	0	0	0	0	0	1
13 	0	0	0	0	0	2	2	2
14 	0	0	0	0	0	0	0	2
15 	0	0	0	2	2	2	2	2
16 	0	0	0	0	0	0	2	2
17 	0	0	0	0	0	0	0	2
18 	0	0	0	0	0	2	2	2
19 	0	0	0	0	0	0	2	2
20 	0	0	0	0	2	2	2	2

(a)

Fig. 4. (a) Weight assignments of 20 orientation spatial patterns in layer S_1 corresponding to eight-degree spatial patterns. (b) Eight orientation spatial patterns in layer C_1 obtained by merging the 20 orientation spatial patterns in layer S_1 . (c) Weight assignments of eight orientation spatial patterns in layer C_1 obtained by summing each column to represent eight-degree spatial patterns corresponding to < 45°, 45, 90, 135, 180, 225, 270 and 360°.

degree pattern corresponding to 135° of pattern 1 in layer C_1 was obtained by the highest weight produced by pattern 15 in layer S_1 ,

which was ‘2’ according to Fig. 4(a). As another example, the weight of the degree pattern corresponding to 90° of pattern 1 in

layer C_1 was '1' because from Fig. 4(a) the highest weight among the six merging patterns, 1–4 and 14–15 in layer S_1 is '1' produced by the weight of pattern 1 in layer S_1 . Similarly, the weights of the remaining degree patterns of pattern 1 corresponding to $< 45^\circ$, 180° , 225° , 270° and 360° were obtained by 0, 0, 2, 2, 2, 2 respectively as shown in Fig. 4(c).

The layer following the first (S_1, C_1) unit is a 3-D figure layer, which is a shape information display layer. It uses a 3-D figure to represent the shape orientations of an input

pattern in the third dimension, called elevation. The magnitude of the elevation of each degree spatial pattern is expressed by the sum of the weights produced by the eight orientation patterns in layer C_1 that correspond to this particular degree pattern. For example, the degree spatial pattern represented by $< 45^\circ$ only matches patterns 2 and 3 in layer C_1 . So, the elevation of $< 45^\circ$ degree spatial pattern is the sum of the column under $< 45^\circ$ degree spatial pattern, which is $1 + 1 = 2$. Similarly, 3, 4, 5, 7, 9, B(11) and C(12) are

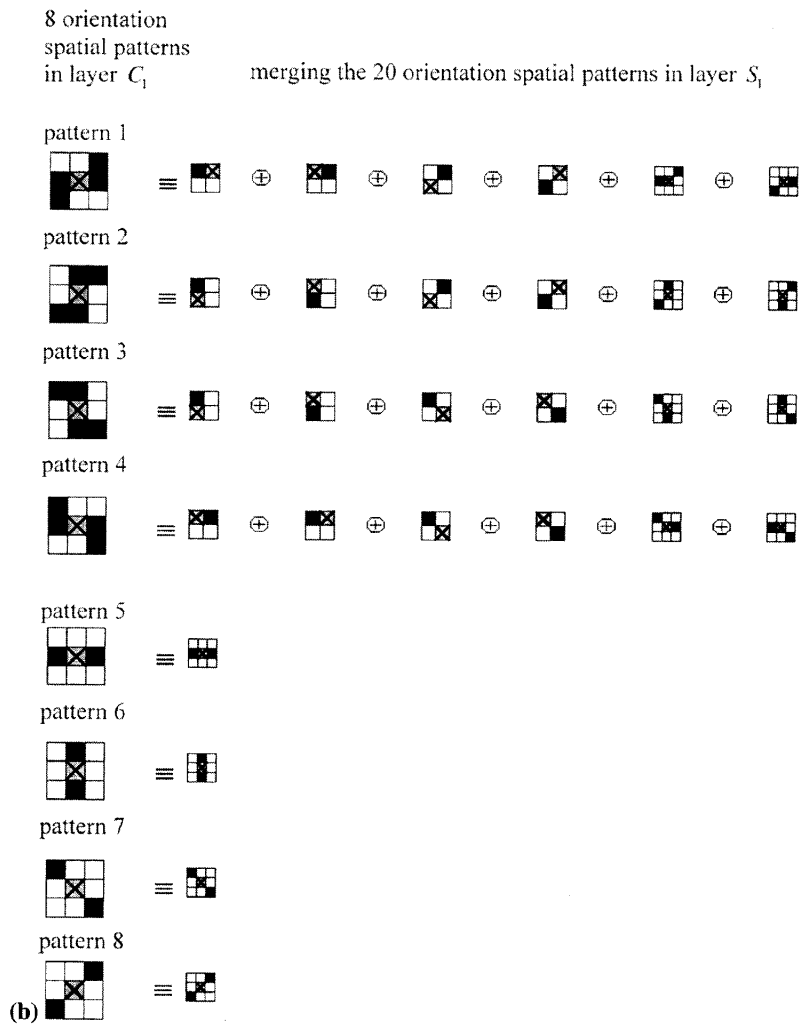


















Fig. 4. (Continued)

8 degree spatial patterns

								
8 orientation spatial patterns in layer C_1	< 45°	45°	90°	135°	180°	225°	270°	360°
1 	0	0	1	2	2	2	2	2
2 	1	1	1	1	1	1	2	2
3 	1	1	1	1	1	2	2	2
4 	0	1	1	1	2	2	2	2
5 	0	0	0	0	1	1	1	1
6 	0	0	0	0	0	0	1	1
7 	0	0	0	0	0	1	1	1
8 	0	0	0	0	0	0	0	1
numeric values = sum of columns	2	3	4	5	7	9	B(11)	C(12)

(c)

Fig. 4. (Continued)

obtained for degree spatial patterns corresponding to 45°, 90°, 135°, 180°, 225°, 270° and 360°.

The second simple-complex layers combined unit, (S_2, C_2) is a shape feature extraction-classification unit where layer S_2 extracts and selects an appropriate set of shape features for classification and layer C_2 performs classification based on features generated by layer S_2 . Layer S_2 produces a set of shape features on the basis of the shape orientation information provided by the 3-D figure in 3-D figure layer. Each of these features represents a specific shape characteristic. It should

be noted that layer S_2 is very flexible and varies with feature selections. C_2 layer is a classification layer that employs a probabilistic neural network (PNN) [53] as a classifier. It takes as input the shape features produced by layer S_2 and its outputs are used for classification. According to the MCCs assessment categories suggested in [25], five outputs are used in layer C_2 , which are 'negative' (no further operation), 'benign finding' (MCCs found to be negative), 'probably benign finding' (short interval follow-up suggested), 'suspicious abnormality' (biopsy should be considered) and 'highly suggestive of malig-

nancy'. Each of these outputs reflects different stages of MCCs. It is worth noting that the number of cell planes in layer C_2 is determined by the number of patterns needed to be classified or recognized. Because layer C_2 can be designed separately, it offers S-Cognitron great flexibility to adapt different applications. In this paper, a PNN is implemented to perform classification task. However, a backpropagation neural network (BNN) can also be used as was done in [42] to detect the venous beading in retinal images. In addition, layer C_2 is an application-dependent layer and can be designed by specific criteria for classification. For details of S-Cognitron, we refer to references [38,42].

In summary, S-Cognitron works as follows. (1) It first takes the clustered MCCs produced by the MCCs Detection Module as input patterns in layer U_0 . (2) It employs the shape orientation extraction unit (S_1, C_1) to extract shape orientations and convert them to numeric representations. (3) It then displays the numeric values of shape orientations in the 3-D figure layer. (4) The shape feature extraction-classification unit (S_2, C_2) is used to extract and select shape features in S_2 layer and then classifies clustered MCCs in layer C_2 . The outputs of layer C_2 produce five-category classification results for diagnosis of the input clustered MCCs. The details of each layer are described in [42].

3. Experimental results using Nijmegen database

In order to evaluate the proposed system, the Nijmegen database was used for experiments. The choice of this database is based on the availability of the Nijmegen database in the public domain and biopsy results are also provided for each case in the database. There are 40 mammograms from 21 patients

in the Nijmegen database collected by the Department of Radiology, Nijmegen University Hospital, Netherlands. Each of the mammograms in the database was digitized into the size of 2048×2048 by an Eikonix 1412 12-bit CCD camera with a fixed calibration. The maximum output level (4095) corresponds to the optical density 0.18. A sample aperture of 0.5 mm in diameter and 0.1 mm sampling distance were used for digitization. All the mammograms were corrected for inhomogeneity of the light source (Gordon planar 1417) and recorded by a Kodak MINR/SO1777 screen/film combination. Each mammogram shows one of more clustered MCCs. The 40 mammograms contain a total of 102 clustered MCCs and the detailed locations and radii of these clustered MCCs were also provided by radiologists.

Since the first module is a Mammogram Preprocessing Module designed to segment the breast region from the entire mammogram, it has little impact on the system performance. So, the experiments were specifically designed to evaluate the performance for the following three modules, the second module-MCCs Finder Module, the third module-MCCs Detection Module and the fourth module-MCCs Classification Module. In this case, 104 positive ROIs and 41 negative ROIs were selected from the Nijmegen data base where a positive ROI means that it contains clusters of microcalcifications and a negative ROI implies that no cluster of microcalcifications is found in the region.

As noted in the MCCs Finder Module, the areas of containing suspicious MCCs were located by the fractal dimension using a 64×64 window to screen the breast region extracted by the Mammogram Preprocessing Module as shown in Fig. 5(a) where the breast region was divided into blocks of 64×64 small regions. Then the MCC Finder Module determined whether or not each

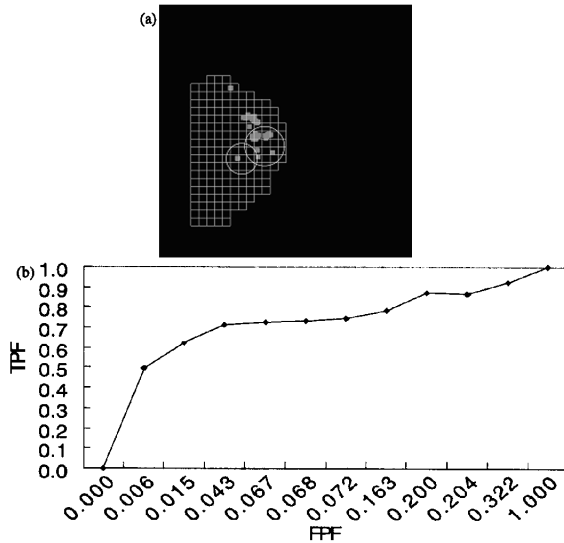


Fig. 5. (a) Breast regions divided into small blocks with size 64×64 . (b) ROC curve of TPF versus FPF.

64×64 small region contains MCCs. In order to demonstrate its detection performance, the receiver operating characteristic (ROC) analysis was used for performance evaluation [54,55]. Assume that A_{MCCs} is the area of blocks that contain MCCs and A_{normal} is the area of blocks that contain normal regions. In addition, we also assume that A_D is the area of blocks detected for MCCs that actually contain MCCs and A_F is the area of blocks detected for MCCs, but did not have MCCs. Two criteria were used for evaluation, referred to as true positive fraction (TPF) and false positive fraction (FPF), which are defined by $TPF = A_D/A_{MCCs}$ and $FPF = A_F/A_{normal}$ respectively [25]. The ROC curve of TPF versus FPF was plotted in Fig.

5(b) where the detection rate can reach as high as 90% at the false alarm rate 1%.

To evaluate the performance of the MCCs Detection Module, here, we adopt the criteria suggested in [25] to define sensitivity (SS), specificity (SP) and positive predictive value (PPV) for performance evaluation. Let N_p be the total number of positive ROIs and N_n be the total number of negative ROIs. Then we define TPN to be the number of ROIs that contains clusters of microcalcifications and are actually detected, and FPN to be the number of ROIs that contains no clusters of microcalcifications but were falsely detected. Similarly, the true negative number (TNN) and false negative number (FNN) can be defined by $TNN = N_n - FPN$ and $FNN = N_p - TPN$ respectively. According to [25], we can further define sensitivity (SS), specificity (SP) and positive predictive value (PPV) by $SS = TPN/N_p$, $SP = TNN/N_n$, $PPV = TPN/(TPN + FPN)$ where $N_p = TPN + FPN = 104$ and $N_n = TNN + FNN = 41$. Since the JRE has been shown more effective than other thresholding methods in our experiments and also in retinal images in [42], only the JRE was used for evaluation and compared to Otsu's method. The detection results produced by the MCCs Detection Module are tabulated in Table 1 that shows that JRE is indeed a better method than Otsu's method.

Finally, these detected clustered MCCs were then fed to the MCCs Classification Module for benign-malignancy classification. It should be noted that all the 41 negative

Table 1
Detection results of the MCCs Detection Module using the Nijmegen database

	N_p	N_n	TPF	FPF	TNF	FNF	SS	SP	PPV
JRE	104	41	96	16	25	8	0.92	0.61	0.86
Otsu's method	104	41	30	14	27	74	0.29	0.66	0.34

Table 2
Confusion matrix of the MCCs Detection Module using the Nijmegen database^a

	Training cases		Negative	Benign finding	Probably benign finding	Suspicious abnormality	Highly suggestive of malignancy
	Benign	Malignant					
1	7	19	29	13	14	20	28
2	8	25	16	8	19	27	34
3	10	31	17	9	8	35	35

^a Total cases: 104, benign cases: 29, malignant cases: 75.

ROIs were not included here because they were filtered out by the second module, MCCs Finder Module, which located potential positive ROIs for MCCs while eliminating negative ROIs that contain no MCCs. As a result, only 104 ROIs needed to be classified. Among these 104 positive ROIs 29 were benign and 75 malignant. Three sets of training data were selected to evaluate the classification performance. The first training data set consists of seven benign and 19 malignant cases, whereas the second and third training data sets were made up of eight benign + 25 malignant cases and ten benign + 31 malignant cases, respectively. These training cases were selected from the 104 ROIs. The resulting confusion matrix is given in Table 2 where the 104 positive ROIs were classified according to the following five categories: 'negative', 'benign finding', 'probably benign finding', 'suspicious abnormality' and 'highly suggestive of malignancy'. Since the information provided by the Nijmegen data base about each cluster of MCCs is based on its biopsy report, all the clusters of MCCs in the Nijmegen data base must be classified into either benign or malignant and cannot be classified in accordance with five categories suggested by the MCCs Classification Module. In this case, we declared a cluster of MCCs to

be malignant only if it fell in the categories of 'suspicious abnormality' and 'highly suggestive of malignancy'; benign, otherwise. That is, if we let n_{neg} , n_{bf} , n_{pbf} , n_{abn} and n_{hsm} denote the number of ROIs classified to category categories, 'negative', 'benign finding', 'probably benign finding', 'suspicious abnormality' and 'highly suggestive of malignancy' respectively, the number of cases to be classified to be benign, denoted by n_B is given by $n_B = n_{neg} + n_{bf} + n_{pbf}$, while the number of cases to be classified to be malignant, denoted by n_M is $n_M = n_{abn} + n_{hsm}$. Using TPN, FPN, TNN and FNN defined above, we can derive three rates for system performance evaluation, which are detection rate (DR) defined by $DR = TPN / (\text{number of malignant cases})$, false alarm rate (FAR) by $FAR = FPN / (\text{number of benign cases})$ and correct classification rate (CR) by $CR = (TPN + TNN) / (\text{total number of cases})$. These results are also tabulated in Table 3. For example, when the MCCs Classification Module was trained by the third training data set, $n_{neg} = 17$, $n_{bf} = 9$, $n_{pbf} = 8$, $n_{abn} = 35$ and $n_{hsm} = 35$ with $n_B = 34$ and $n_M = 70$, which result in $TPN = 70$, $FPN = 0$, $TNN = 29$ and $FNN = 5$. So, $DR = 70/75 \approx 93\%$, $FAR = 0/29 = 0\%$ and $CR = (70 + 29)/104 \approx 95\%$. This experiment showed that if 41 train-

ing cases were used to test 104 cases, we could achieve as high as 95% classification rate with 93% detection rate and 0% false alarm rate. As we can see from Tables 2 and 3, the more training data that were used, the better the classification. It is worth noting that the classification performance is not linearly proportional to the number of training cases used.

The processing time of each module in the system using PC Pentium 200 MHz is given in Table 4. From the table, the total amount of the time required for the entire system to process the mammogram in Fig. 6(b) is averaged about 72 s.

4. Demonstration of system

In this section, a demonstration of the system is presented. The system can be run on a PC platform. Fig. 6(a) shows a system

frame where the larger window, called display window, is used to display the original mammogram and a smaller window, called ROI window, is used to show a particular ROI that can be selected by the system or radiologists. There is a tool bar across the top of the frame. It consists of various functions, 'File', 'Screening', 'Sensitivity', 'Refresh', 'Biopsy', 'Image processing'. When 'File' is opened, the Mammogram Preprocessing Module (i.e. the first module of the system) allows users to access the mammogram database stored in the system. As soon as a mammogram is selected, it will be shown in the display window in Fig. 6(b) where the ROI window also shows a ROI selected by a small movable square window, called processing window, that is designed to allow users to select any region of interest in the ROI window without appealing for the automatic 'Screening'. However, when 'Screening' is used, the MCCs Finder Module

Table 3
Classification results of Table 2^a

	TPN	FPN	TNN	FNN	DR (%)	FAR (%)	CR (%)
1	43	5	24	32	57	17	64
2	58	3	26	17	77	10	81
3	70	0	29	5	93	0	95

^a Total cases: 104, benign cases: 29, malignant cases: 75.

Table 4
Processing time for each module of the system

	Mammogram Preprocessing Module	MCCs Finder Module	MCCs Detection Module (256 × 256 pixel size per cluster)	MCCs Classification Module (256 × 256 pixel size per cluster)	Total processing time
Average time (s)	37	30	4	1	72
Variance (s)	2.5	8.4	1.3	0.1	12.3
Rate of variance (%)	7	28	33	10	17

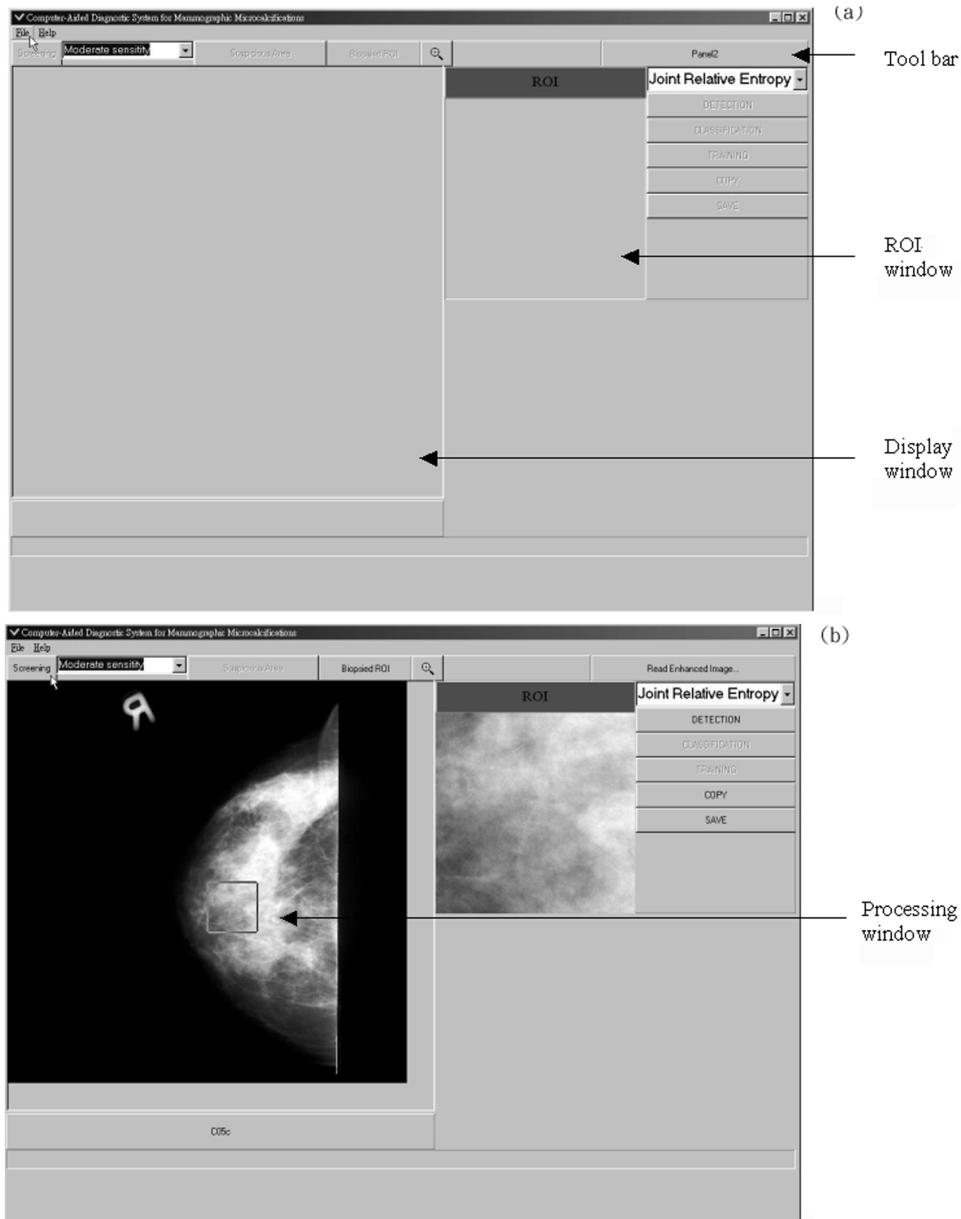
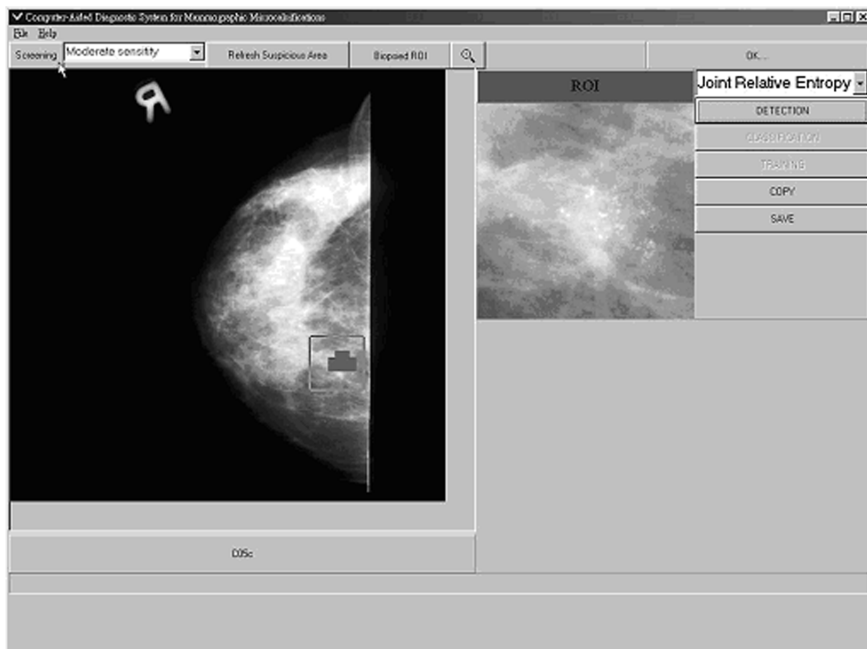


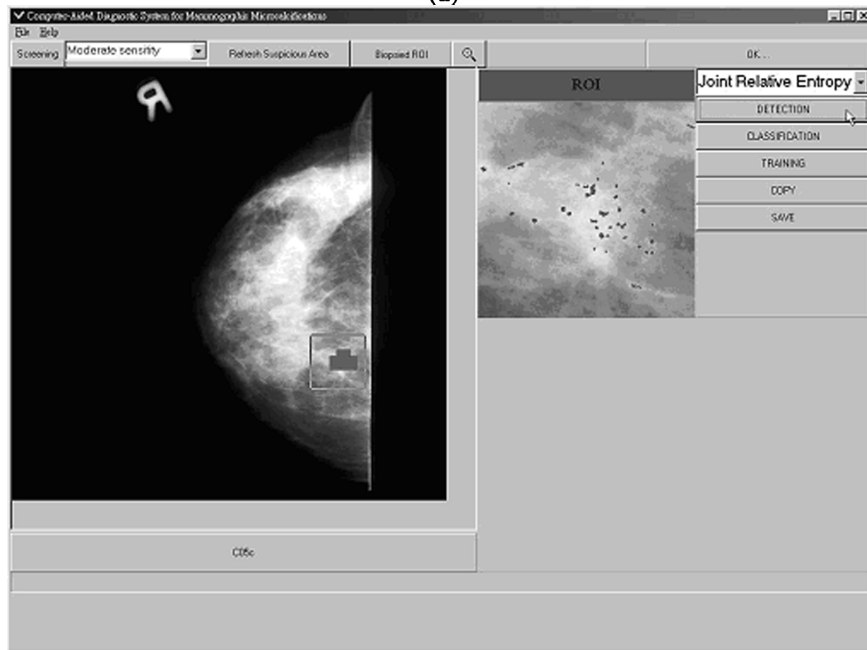
Fig. 6. (a) A system frame of the demonstration. (b) A selected mammogram and an ROI shown in the windows

(i.e. the second module of the system) processes the selected mammogram in real time. It automatically locates all suspicious areas of MCCs as shown in Fig. 7(a). ‘Sensitivity’ can be implemented in three fashions, ‘Low sensitivity’, ‘Moderate sensitivity’ and ‘High sensitivity’, which are determined by the fractal

dimension D through the scaling factor r in Eq. (5). They can be used to control the sensitivity of locating true positive MCCs. When ‘Low sensitivity’ is selected, it finds all areas of containing suspicious MCCs that can be easily located by visual inspection. So, it produces low false positive rates but may



(a)



(b)

Fig. 7. (a) Suspicious MCCs area located by the MCCs Finder Module. (b) Detection result of Fig. 8(a) resulting from applying the JRE to the image in Fig. 7(a).

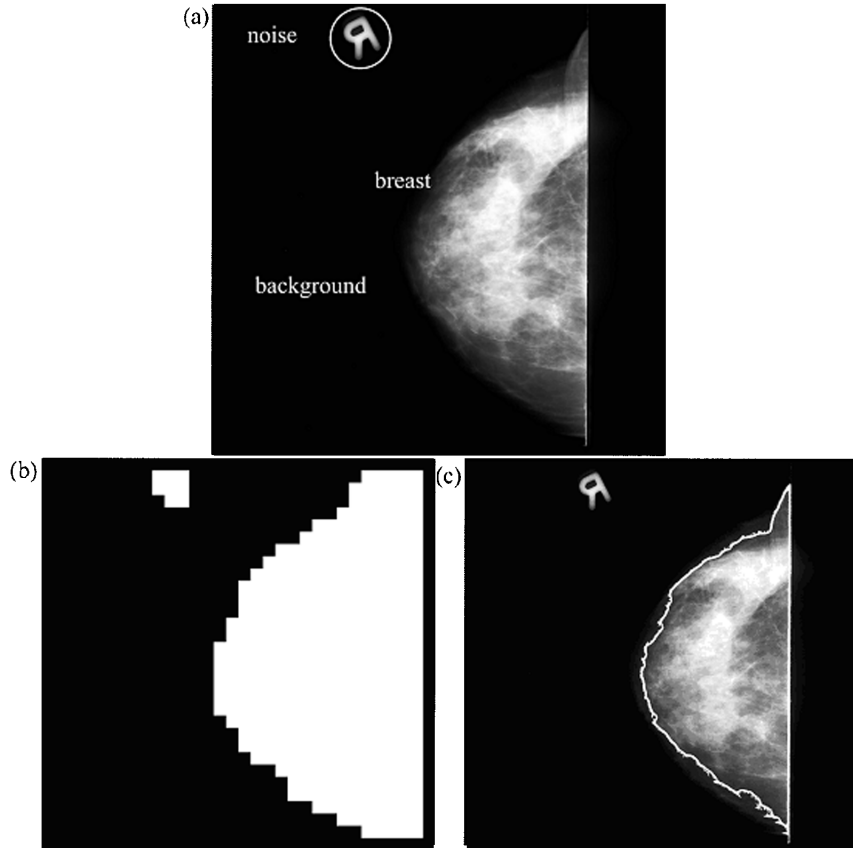


Fig. 8. (a) A representative mammogram from the Nijmegen data base. (b) Background elimination by the block region growing method. (c) Breast region extracted by the K-means clustering-based thresholding method.

miss many areas of true positive MCCs. By contrast, when ‘High sensitivity’ is selected, MCCs that are invisible to human eyes can be now extracted so as to achieve high true positive rates. However, this also results in many unnecessary reviews required for radiologists. The ‘Moderate sensitivity’ is a compromise between ‘low sensitivity’ and ‘high sensitivity’. ‘Refresh’ is simply to clear previous identified suspicious areas. ‘Biopsy’ gives the true diagnosis of the MCCs in the ROI window. ‘Image processing’ provides several functions, ‘Detection’, ‘Classification’, ‘Training’, ‘Save’ and ‘Copy’. When ‘Detection’ is used, the MCCs Detection Module (i.e. the third module of the system) is activated. In this

case, the users need to select a particular thresholding method to segment clustered MCCs from the background as shown in the upper right ROI window of Fig. 7(b). ‘Classification’ classifies a cluster of detected MCCs into five categories: ‘Negative’, ‘Benign finding’, ‘Probably benign finding’, ‘Suspicious abnormality’, ‘Highly suggestive of malignancy’ (see the right bottom of Fig. 11). The three functions ‘Training’, ‘Save’ and ‘Copy’ are yet to be developed but will be included for future reporting and data system.

In order to demonstrate how the entire system works module by module, a representative mammogram shown in Fig. 8(a) was selected from the Nijmegen database to eval-

uate each of the modules implemented in the system. As we can see, approximately half of the mammogram is the breast background. It provides very little diagnostic information except for the letter ‘R’ that indicates the right breast. Fig. 8(b)–(c) is the result of the first module, Mammogram Preprocessing Module where Fig. 8(b)–(c) were obtained by steps 1–2 and step 3 in the Breast Extraction Algorithm respectively. Step 1 divided the mammogram into a set of image blocks of size 64×64 and step 2 applied a block region growing method with the threshold set to $T_e = 15$ to eliminate the breast background. Step 3 used K-means clustering method to refine the breast estimate in Fig. 8(b). The result is shown in Fig. 8(c) where three regions were classified, C_0 corresponded to the breast background, C_2 extracted the full breast and C_1 was the breast boundary delineated by the bright contour.

After the breast region was extracted, it was fed to the second module, MCCs Finder Module, to locate and extract suspicious ar-



Fig. 9. Image resulting from applying blanket method to the image in Fig. 8(c).

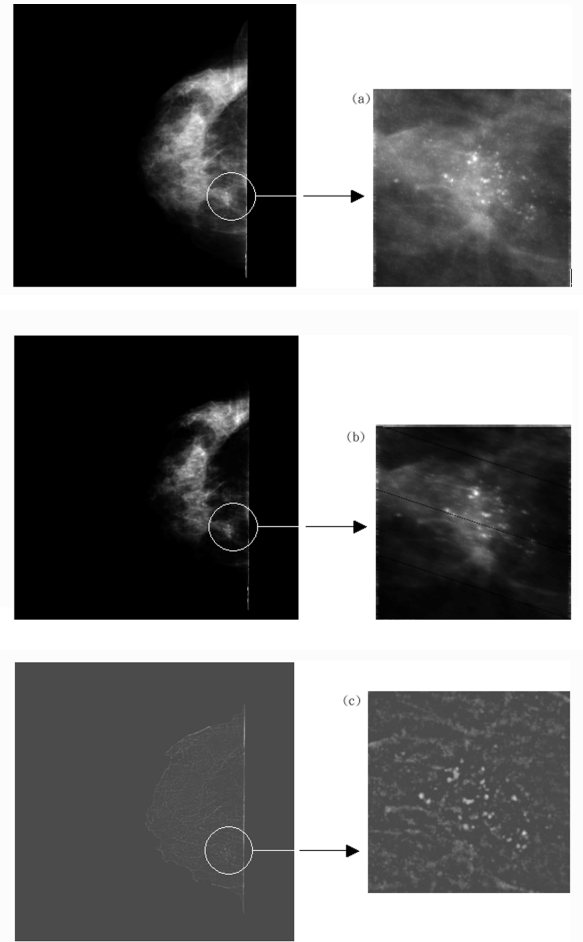


Fig. 10. (a) Image resulting from applying the gradient enhancement to Fig. 9. (b) Image resulting from applying the contrast enhancement to Fig. 10(a). (c) Image resulting from applying the Gaussian filtering to Fig. 10(b).

eas of MCCs. Fig. 9 is the result of applying the blanket method to the image in Fig. 8(c) where a suspicious cluster of MCCs was located and marked by a circle. This clustered MCCs was further fed to the third module, MCCs Detection Module.

Fig. 10(a)–(c) were the results produced by applying steps 1–4 respectively in the MCCs Detection Algorithm to Fig. 9 where the left column shows the full image while the right

column shows only ROI for better illustration. Fig. 10(a) enhanced the pixels with high intensities and Fig. 10(b) improved their relative contrast with $m = 2$. Since some pixels that are not calcified might also be enhanced by the gradient enhancement, the contrast improvement allowed us to eliminate these pixels by Gaussian filtering as shown in Fig. 10(c). This is because the pixels surrounding uncalcified pixels will bring down the contrast while the pixels surrounding calcified pixels will bring up the contrast. The upper right ROI window of Fig. 11 shows the detection result of Fig. 10(c) segmented by the JRE.

Finally, the clustered MCCs was classified by the fourth classification module, MCCs Classification Module into the category ‘Highly suggestive of malignancy’ as shown in the right bottom of Fig. 11 where the circled area in the left display window of Fig.

9 is the biopsy report provided by the Nijmegen database and is indeed malignant.

5. Conclusion

In this paper, we present a prototype of a computer-aided design mammography screening system. The initial development of this system began in 1995 under a pilot project funded by TCVGH over the past years. It is the result of a multi-year effort led by the Department of Radiology in TCVGH. The third MCCs Detection Module was first developed to assist radiologists to improve their diagnosis in detection of MCCs. As the progress evolved, the classification of clustered MCCs immediately became indispensable. This need resulted in the fourth MCCs Classification Module to classify clustered MCCs. In order to make the system fully

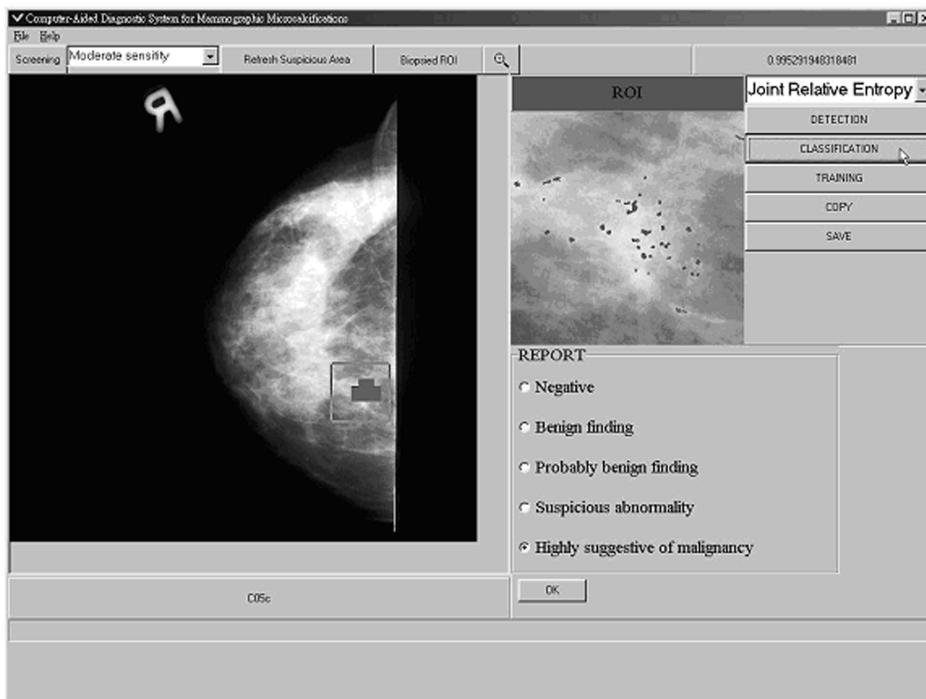


Fig. 11. Detection and classification results of Fig. 10(c).

computer automated, the second MCCs Finder Module was then developed to automatically locate suspicious areas for possible MCCs, a task usually done by radiologists' visual inspection. Finally, to integrate these three modules in the existing PACS in TCVGH, it required a computer graphics interface and a build-up of mammogram database for storage and fast retrieval. To meet this challenge, the first Mammogram Preprocessing Module was further developed for this purpose. The system has been evaluated by radiologists in TCVGH and other hospitals in Taiwan as each module was developed. A demonstration version of the system was also prepared for radiologists in Taiwan who would like to participate in our project. The feedback received from radiologists was used to further improve the system. Although the present system is still a prototype and requires many more clinical trials, its open architecture offers great flexibility. It enables the system to upgrade its individual module as well as to add new modules to make the system more reliable and adapt rapid changing computer environment. It should be noted that the system presented in this paper is only a trial clinical system. The techniques used in each module can be further improved or replaced if there exist more effective methods. For example, a module for mass detection currently being investigated in TCVGH can be added to the system as a companion module of the MCCs Detection Module since the techniques used to detect MCCs are generally different and cannot be used for mass detection. Another important module yet to develop for our proposed system is a reporting and data system module. It will be based on the BI-RADS (Breast Imaging Reporting and Data System) format recommended by the American College of Radiology [25]. It will allow doctors to prepare their reports and can be also used for

purpose of training and future pathology. We anticipate that this module will require a large database build-up as well as an extensive study that includes various statistical tests and tedious book-keeping data analysis.

Acknowledgements

The authors would like to thank Taichung Veterans General Hospital for their support under a research grant (TCVGH 865501C) and would also like to thank the National Expert and Training Centre for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, Netherlands for providing the Nijmegen mammogram database. They would also like to thank Althouse and Daniel C. Heinz for proof reading this manuscript.

References

- [1] Time Magazine, January 14 (1991) 48–55.
- [2] K.F. Schmidt, A better breast test, *Sci. News* 143 (1993) 392–393.
- [3] X. Zhou, R. Gordon, Detection of early breast cancer: an overview and future prospects, *Crit. Rev. Biomed. Eng.* 17 (1989) 203–255.
- [4] Newsweek, May 10, (1993) 64.
- [5] D.J. Vyborny, Can computers help radiologists read mammograms, *Radiology* 191 (1994) 315–317.
- [6] S.A. Feig, B.M. Galkin, Breast microcalcifications: early warning for cancer?, *Diagn. Imag.* November (1990) 132–138.
- [7] L.W. Bassett, Mammographic analysis of calcifications, *Radiol. Clin. North Am.* 30 (1) (1992) 93–105.
- [8] E.A. Sickles, Breast calcifications: mammographic evaluation, *Radiology* 160 (1986) 289–293.
- [9] M. Lanyi, Microcalcifications in the breast — a blessing or a curse?, *Diagn. Imag. Clin. Med.* 54 (1985) 126–145.
- [10] M.P. Kegelmeyer, J.M. Pruneda, P.D. Bourland, A. Hillis, M.W. Riggs, M.L. Nipper, Computer-aided mammographic screening for spiculated lesions, *Radiology* 191 (1994) 331–337.

- [11] B.W. Fam, S.L. Olson, P.F. Winter, F.J. Scholz, Algorithm for the detection of fine clustered calcifications on film mammograms, *Radiology* 169 (1988) 333–337.
- [12] D.H. Davies, D.E. Dance, Automatic computer detection of clusters of calcifications in digital mammograms, *Phys. Med. Biol.* 35 (1990) 1111–1118.
- [13] H.P. Chan, K. Doi, S. Galhotra, C.J. Vyborny, H. MacMahon, P.M. Jokich, Image feature analysis and computer-aided diagnosis in digital radiology. I. Automated detection of microcalcifications in mammography, *Med. Phys.* 14 (1987) 538–548.
- [14] H.P. Chan, K. Doi, C.J. Vyborny, R.A. Schmidt, C.E. Metz, K.L. Lam, T. Ogura, Y. Wu, H. MacMahon, Improvement in radiologists' detection of clustered microcalcifications on mammograms, *Invest. Radiol.* 20 (1990) 1102–1110.
- [15] J. Dengler, J. Behrens, J.F. Desaga, Segmentation of microcalcifications in mammograms, *IEEE Trans. Med. Imag.* 12 (1993) 634–642.
- [16] L. Shen, R.M. Rangayyan, J.E.L. Desautels, Application of shape analysis to mammographic calcifications, *IEEE Trans. Med. Imag.* 13 (1994) 263–274.
- [17] Y. Wu, K. Doi, M.L. Giger, R.M. Nishikawa, Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks, *Med. Phys.* 19 (1992) 555–560.
- [18] Y. Wu, M.L. Giger, K. Doi, R.A. Schmidt, C.E. Metz, Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, *Radiology* 187 (1993) 81–87.
- [19] W. Zhang, K. Doi, M.L. Giger, Y. Wu, R.M. Nishikawa, R.A. Schmidt, Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural networks, *Med. Phys.* 21 (1994) 517–524.
- [20] W. Zhang, K. Doi, M.L. Giger, Y. Wu, R.M. Nishikawa, R.A. Schmidt, An improved shift-invariant artificial neural networks for computerized detection of clustered microcalcifications in digital mammograms, *Med. Phys.* 23 (1996) 595–601.
- [21] S.C.B. Lo, H.P. Chan, J.-S. Lin, H. Li, M.T. Freeman, S.K. Mun, Artificial convolution neural network for medical image pattern, *Neural Networks* 7/8 (1995) 1201–1214.
- [22] H.P. Chan, S.-C. Lo, B. Sahiner, K.L. Lam, M.A. Helvie, Computer-aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network, *Med. Phys.* 22 (1995) 1555–1567.
- [23] A.F. Laine, S. Schuler, J. Fan, W. Huda, Mammographic feature enhancement by multiscale analysis, *IEEE Trans. Med. Imag.* 13 (1994) 725–740.
- [24] *IEEE Mag. Eng. Med. Biol.* 14 (1995) 536–577.
- [25] BI-RADS (Breast Imaging Reporting and Data System), third ed., American College of Radiology, 1998.
- [26] P.-S. Liao, B.C. Hsu, C.-S. Luo, P.-C. Chung, T.-S. Chen, S.-K. Lee, L. Cheng, C.-I Chang, Automatic detection of microcalcifications in digital mammograms, in: Proceedings of the 18th Annual International Conference IEEE Engineering, Medical and Biological Society, 1996, pp. 88–89.
- [27] B.-C. Hsu, P.-C. Chung, C.-I Chang, Automated system for detection and classification of microcalcifications in digital mammograms, in: Proceedings of CVGIP '96, Taiwan, 1996, pp. 127–134.
- [28] S.-C. Lo, P.-C. Chung, B.-C. Hsu, C.-I Chang, S.K. Lee, B.-S. Liao, An algorithm for detection and segmentation of clustered microcalcifications on mammograms, in: Proceedings of the 2nd Medical Engineering Week of the World, 3rd Asian–Pacific Conference on Medical and Biological Engineering, Taiwan, 1996, p. 102.
- [29] C.-S. Lo, P.-C. Chung, C.-I Chang, S.K. Lee, A computerized system for detection and segmentation of clustered microcalcifications, in: Proceedings of the Joint Conference International Computer Symposium, Taiwan, 1996, pp. 247–253.
- [30] C.-S. Lo, S.-K. Lee, P.C. Chung, C.-I Chang, An automatic computerized system for detection and segmentation of clustered microcalcifications on mammograms, in: Proceedings of the 46th Annual Meeting of Radiological Society of Republic of China, 1997, p. F157.
- [31] C.-W. Yang, P.-C. Chung, C.-I Chang, An image capture and communication system for emergency computed tomography, *Comput. Methods Prog. Biomed.* 52 (1997) 139–145.
- [32] C.-W. Yang, P.-C. Chung, C.-I Chang, S.-K. Lee, L.-Y. Kung, A hierarchical model for PACS, *Comput. Methods Med. Imag. Graph.* 21 (1) (1997) 29–37.
- [33] *Vision Systems Design*, May (1999) 25–31.
- [34] *Optical Engineering Reports*, June (1999) 7.
- [35] W. Zhang, H.M. Romsdahl, J.R. Roehrig, Method and system for using local attention in the detection of abnormalities in digitized medical images, U.S. Patent and Trademark Office, patent number 6 014 452 January, 2000.

- [36] W. Zhang, K. Doi, Method and system for the detection of microcalcifications in digital mammograms, U.S. Patent and Trademark Office, patent number 5 492 627 May 13, 1993.
- [37] K. Fukushima, N. Wake, Handwritten alphanumeric character recognition by the neocognitron, *IEEE Trans. Neural Networks* 2 (3) (1991) 355–365.
- [38] C.-W. Yang, P.-C. Chung, C.-I Chang, A new neural model for shape recognition: shape cognitron, in: *Proceedings of the 1996 International Symposium on Multi-Technology Information Processing*, Taiwan, 1996, pp. 195–200.
- [39] P.-C. Hsu, An Automatic System for Detection and Classification of Microcalcifications in Digital Mammograms, Department of Electrical Engineering, National Cheng Kung University, 1996 M.S. Thesis.
- [40] Y. Xu, C.-I Chang, Implementation of a 3-D model for neocognitron, *Proc. Int. Conf. Neural Network* (1996) 794–799.
- [41] Y. Xu, C.-I Chang, Method and Apparatus for Automatic Pattern Recognition, U.S. Patent and Trademark Office, patent number 5 784 485 July 21, 1998.
- [42] C.-W. Yang, D.-J. Ma, S.-C. Chao, C.-M. Wang, C.H. Wen, S.C. Lo, P.-C. Chung, C.-I Chang, A computer-aided diagnostic detection system of venous beading in retinal images, *Optical Eng.* 39 (5) (2000) 1293–1303.
- [43] J. MacQueen, Some methods for classification and analysis of multivariate observations, *Proc. Fifth Berkeley Symp. Math. Stat. Prob.* 1 (1967) 281–296.
- [44] B.B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, New York, 1977.
- [45] S. Pleg, J. Naor, R. Hartley, D. Avnit, Multiple resolution texture analysis and classification, *IEEE Trans. Pattern Anal. Machine Intelligence* 6 (1984) 518–523.
- [46] E.-L. Chen, P.C. Chung, C.-L. Chen, H.M. Tsai, C.-I Chang, An automatic diagnostic system for CT liver image classification, *IEEE Trans. Biomed. Eng.* 45 (6) (1998) 783–794.
- [47] P.K. Sahoo, S. Soltani, A.K.C. Wong, A survey of thresholding techniques, *Comput. Vis. Graph. Image Process.* 41 (1998) 233–260.
- [48] N.R. Pal, S.K. Pal, Entropic thresholding, *Signal Process.* 16 (1989) 97–108.
- [49] C.-I. Chang, K. Chen, J. Wang, M.L.G. Althouse, A relative entropy approach to image thresholding, *Pattern Recognit.* 27 (9) (1994) 1275–1289.
- [50] C.-W. Yang, P.-C. Chung, C.-I. Chang, J. Wang, M.L.G. Althouse, Entropic and relative entropic thresholding, in: *Proceedings of the Joint Conference 1996 International Computer Symposium*, 1996, pp. 82–89.
- [51] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man Cybernetics* 9 (1) (1979) 62–66.
- [52] T. Cover, J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
- [53] D.F. Specht, Probabilistic neural networks, *Neural Networks* 3 (1990) 109–118.
- [54] J.A. Swets, R.M. Pickett, *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*, Academic Press, New York, 1982.
- [55] C.E. Metz, ROC methodology in radiological imaging, *Radiology* 21 (1986) 720–733.