ELSEVIER

doi:10.1016/S0301-5629(02)00788-3

● *Original Contribution*

IMPROVEMENT IN BREAST TUMOR DISCRIMINATION BY SUPPORT VECTOR MACHINES AND SPECKLE-EMPHASIS TEXTURE ANALYSIS

RUEY-FENG CHANG,* WEN-JIE WU,* WOO KYUNG MOON[†] and DAR-REN CHEN[‡]

*Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan;
[†]Department of Diagnostic Radiology, Seoul National University Hospital, South Korea; and [‡]Department of
General Surgery, China Medical College & Hospital, Taichung, Taiwan

Abstract—Recent statistics show that breast cancer is a major cause of death among women in developed countries. Hence, finding an accurate and effective diagnostic method is very important. In this paper, we propose a high precision computer-aided diagnosis (CAD) system for sonography. We utilize a support vector machine (SVM) to classify breast tumors according to their texture information surrounding speckle pixels. We test our system with 250 pathologically-proven breast tumors including 140 benign and 110 malignant ones. Also we compare the diagnostic performances of three texture features, i.e., speckle-emphasis texture feature, nonspeckle-emphasis texture feature and conventional all pixels texture feature, applied to breast sonography using SVM. In our experiment, the accuracy of SVM with speckle information for classifying malignancies is 93.2% (233/250), the sensitivity is 95.45% (105/110), the specificity is 91.43% (128/140), the positive predictive value is 89.74% (105/117) and the negative predictive value is 96.24% (128/133). Based on the experimental results, speckle phenomenon is a useful tool to be used in computer-aided diagnosis; its performance is better than those of the other two features. Speckle phenomenon, which is considered as noise in sonography, can intrude into judgments of a physician using naked eyes but it is another story for application in a computer-aided diagnosis algorithm. (E-mail: dlchen88@ms13.hinet.net)   © 2003 World Federation for Ultrasound in Medicine & Biology.

Key Words: Speckle, Support vector machine, Computer-aided Diagnosis, Breast ultrasound.

INTRODUCTION

According to a recent report (Bothorel et al. 1997), breast cancer has affected one of every 10 women in Europe and one of every eight women in the USA. In other words, breast cancer has become a major cause of death among women in developed countries. The most useful way to reduce breast cancer deaths is to treat the disease as early as possible. However, earlier treatment requires the ability to detect breast cancer in early stages. Early diagnosis requires an accurate and reliable diagnosis procedure that allows physicians to distinguish benign breast tumors from malignant ones. Thus, how to find an accurate and effective diagnosis method is very important. Biopsy is the best way accurately to determine whether the tumor is benign or malignant. However, it is invasive and expensive and positive findings at biopsy for cancer are low, between 10% and 31% (Gisvold and

Martin 1984; Bassett et al. 1991; Rubin et al. 1997). Medical imaging plays an important role in the clinical study and was dominated by X-radiography in early stages. Owing to remarkable advances in physics and engineering, the use of ultrasound (US) is getting popular. According to the statistics, more than one out of every four medical image studies is performed with ultrasound and the proportion is increasing now (Wells 2000). Comparing with other imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI), ultrasound is lower cost, higher efficacy, real-time, convenient and safe; therefore, it gradually becomes the most important diagnostic tool. However, ultrasound imaging has its limitations, such as the low contrast and interference with noise and speckle. Speckle is a common phenomenon in coherent imaging systems. It occurs when a coherent source and a coherent or noncoherent detector are used to interrogate a medium, which is rough on the scale of the wavelength. Speckle is an artifact degrading target visibility. It will appear as a granular structure superimposed on the image and is one of the primary limitations on detecting low contrast le-

Corresponding Author: Dar-Ren Chen Department of General Surgery, China Medical College & Hospital 2 Yer-Der Rd, Taichung, Taiwan. Tel No: (886)-4-22052121 Ext. 1629 Fax: (886)-4-7267167 E-mail: dlchen88@ms13.hinet.net

sions in ultrasonic imaging. Hence, it is thought to be one kind of noise. Many authors have proposed methods to reduce it in ultrasonic images (Trahey et al. 1986; Dutt and Greenleaf 1996). Nevertheless, speckle may include information behind and it can be utilized to track internal motion due to applied deformation (Skovoroda et al. 1995; Steinberg et al. 1998) and to position the 3D ultrasound frames (Tuthill et al. 1998).

Support vector machines (SVM) (Vapnik 1995; Pontil and Verri 1998; Chapelle et al. 1999) have been recently proposed as a very effective method for pattern recognition, machine learning and data mining. It is considered to be a good candidate because of its high generalization performance. Intuitively, given a set of points which belong to either one of two classes, a SVM can find a hyperplane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyperplane. According to Vapnik (1995), this hyperplane, called optimal separating hyperplane (OSH), can minimize the risk of misclassifying examples of the test set.

There are many successful applications of the CAD strategy reported for medical imaging. These CAD systems include mammography CAD (Petrick et al. 1996; Huo et al. 2002) and sonographic CAD (Garra et al. 1993; Valckx and Thijssen 1997; Chen et al. 2000, 2002). Although many people were still skeptical of these CAD systems in the past, it is still undergoing great development and utilization within the field of medical imaging by showing its potential in many clinical areas.

In this paper, we evaluated breast masses in pathologically-proven tumors using SVM based on analysis of speckles for classifying breast tumors. Regions-of-interest (ROIs) of ultrasonic images selected by the physician and autocovariance texture features were used in our experiments. We first classify these pixels into speckle and nonspeckle pixels by a speckle detector. Then we compute the speckle-emphasis texture feature, namely, autocovariance texture feature of speckle pixels. To be contrasted, we also compute autocovariance texture features of nonspeckle pixels and all pixels, respectively. The reason of computing autocovariance is that we suppose spatial relationships of pixels surrounding speckles can help to improve breast tumor discrimination. Finally, we compare the diagnostic performances using these three features to prove that the speckle-emphasis texture feature is superior to the others.

MATERIALS AND METHODS

Image data acquisition

Generally, an experienced physician can easily identify a tumor on an ultrasonic image by the shape of tumor and contrast of internal echoes. However, it is hard auto-
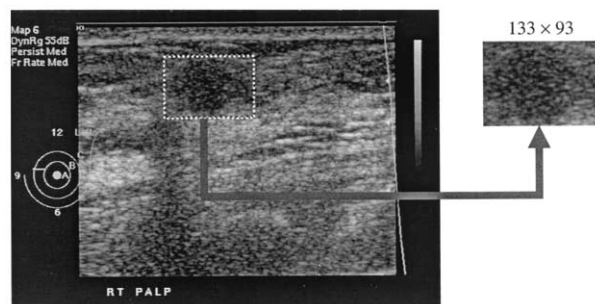

Fig. 1. A 640 $\times$ 480 digital image is captured from the ultrasound scanner. In a 1 cm $\times$ 1 cm rectangle, there are 100 $\times$ 100 = 10000 pixels. The ROI rectangle is 1.33 cm $\times$ 0.93 cm and 133 $\times$ 93 pixels.

matically to segment a tumor from the ultrasonic image by a computer system, owing to the noise and contrast. In this study, a physician first extracts the subimage of an ROI of an ultrasonic image and then, using the texture information of speckles for further differential diagnosis, uses the CAD system to analyze the subimage.

There are 250 ultrasonic images of pathologically-proven (either by fine needle cytology, core-needle biopsy or open biopsy) benign breast tumors from 140 patients and carcinomas from 110 patients in the ultrasonic image database. The database contains only one image from each patient. The ultrasonic images are captured at the largest diameter of the tumor. The images were collected from August 1, 1999 to May 31, 2000; the patients' ages ranged from 18 to 64 y. Sonography is performed using an ATL HDI 3000 system with a L10-5 small parts transducer. The L10-5 transducer is a linear-array transducer with a frequency of 5-10 MHz and a scan width of 38 mm. When acquiring images, the patient is in the supine position with arms extended overhead. No acoustic stand-off pad is used. The acquired images are quantized into 8 bits, i.e., 256 gray levels and the features were stored via the magnetooptical (MO) disks. Then these images can be read and analyzed in a personal computer and serve as our experimental data. W.K.M. (one of authors) supplied the entire database. All solid nodules at US belong over C3 according to ACR BI-RADS category and all images were fed into the algorithm without selection bias. The subimages of ROI were manually selected by a breast surgeon (one of authors) retrospectively from the database. Tissue diagnosis was unknown to the physician before ROI selections. The selection ROIs were outlined, leaving a 1-2 mm given border around the lesion. The ROI has to include the entire extent of the spiculations if this characteristic of a lesion present. The ROI subimage was then saved as a file for later analysis. Figure 1 illustrates an example of real-time digitized monochrome ultrasonic

image of a benign tumor. ROI images are used in our breast images database further to investigate the texture characteristics of benign and malignant tumors.

Feature extraction

In this study, autocovariance coefficients of speckles are used in the texture analysis. First, we will give an introduction to speckle then we describe the definition of autocorrelation and autocovariance coefficients. Finally, we introduce how to combine the speckle information and autocovariance coefficients.

Speckle

Speckle is a common phenomenon in ultrasonic pulse-echo measurements. It is caused by the coherent nature of the imaging source and also occurs in laser and radar imaging. When an acoustic wave propagates through a scattering medium, all the wavelets generated by the reflection of the incident wave on randomly distributed scatters interfere on the surface of the receiving transducer in a constructive or destructive way, depending on their relative phase. In other words, speckles are produced by the superposition of numerous waves scattered from the various surface elements of the object.

In order to utilize the speckle information, we must detect speckle patterns first. For a fully-developed speckle, the intensity image should have an exponential distribution and a constant ratio of mean to standard deviation (SD) of 1.0 (Tuthill et al. 1998). Due to this property, a detector based on first-order statistics is used to detect speckles. When the ratio of mean intensity to SD for a region with a moving 5×5 mask is between 0.8 and 1.2, the region is classified as a speckle region and the central pixel is marked as a speckle. However, there is one inverse-log step to do before detecting speckles. That is, each pixel value needs to be converted to a linear scale (Smith and Fenster 2000) by

$$I(x, y) = 10^{P(x,y)/P_o}, \qquad (1)$$

where $I(x,y)$ is the acoustic intensity, $P(x,y)$ is the raw B-mode pixel value in dB and the linearization factor P_0 is used to convert from a 0-255 log scale to a linear scale. P_0 relates to the frequency of the transducer. The reason for this conversion is that the image we acquired is processed with a logarithmic compression to reduce the dynamic range. To detect speckles more precisely, we must decompress it first. In Fig. 2, we show an original tumor image and the result after the speckle detection, respectively. Meanwhile, we also show the gray scales of pixels which are classified as speckles in Fig. 2.
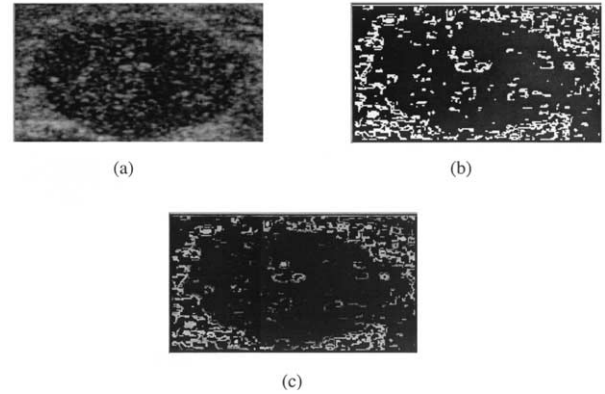

Fig. 2. (a) Original tumor image. (b) The result after the speckle detector of (a). (c) The gray scales of pixels which are classified as speckles.

Autocorrelation and autocovariance coefficients

The autocorrelation coefficients are a kind of statistical method. They can be used to reflect the inner-pixel correlation within an image. The normalized autocorrelation coefficient (Gonzalez and Woods 1992) between pixel (i, j) and pixel $(i + \Delta m, j + \Delta n)$ in an image with size $M \times N$ can be defined as

$$\gamma(\Delta m, \Delta n) = \frac{A(\Delta m, \Delta n)}{A(0,0)}, \qquad (2)$$

where

$$A(\Delta m, \Delta n) = \frac{1}{(M - \Delta m)(N - \Delta n)}$$

$$\times \sum_{x=0}^{M-1-\Delta m} \sum_{y=0}^{N-1-\Delta n} f(x, y) f(x + \Delta m, y + \Delta n). \qquad (3)$$

However, there is a disadvantage in the autocorrelation method. It is usually affected by brightness. The autocorrelation coefficients of two images which have similar textures but different brightnesses may be different. Thus, we replace autocorrelation coefficients with autocovariance ones, as our image features. The autocovariance coefficient is defined as follows:

$$A(\Delta m, \Delta n)$$

$$= \frac{1}{(M - \Delta m)(N - \Delta n)} \sum_{x=0}^{M-1-\Delta m} \sum_{y=0}^{N-1-\Delta n} (f(x, y) - \bar{f})$$

$$\times (f(x + \Delta m, y + \Delta n) - \bar{f}) \qquad (4)$$

where $\bar{f}$ is the mean value of $f(x, y)$.

Combination of speckle information and autocovariance coefficients

To utilize the speckle information, we first use a speckle detector which is described above to locate where the speckle is. Then we calculate the autocovariance coefficients of those pixels which are speckle. To calculate the speckle autocovariance coefficients, eqn (4) can be rewritten as follows:

$$A(\Delta m, \Delta n) = \frac{1}{NS} \sum_{x=0}^{M-1-\Delta m} \sum_{y=0}^{N-1-\Delta n} S_1(f(x, y) - \bar{f})S_2(f(x + \Delta m, y + \Delta n) - \bar{f})$$

$$S_1 = \begin{cases} 0 & \text{if } f(x, y) \text{ is not a speckle} \\ 1 & \text{if } f(x, y) \text{ is a speckle} \end{cases} \quad (5)$$

$$S_2 = \begin{cases} 0 & \text{if } f(x + \Delta m, y + \Delta n) \text{ is not a speckle} \\ 1 & \text{if } f(x + \Delta m, y + \Delta n) \text{ is a speckle,} \end{cases}$$

where $\bar{f}$ is the mean value of $f(x, y)$ and NS is the total number of speckles.

In this work, each ultrasonic image will produce a 5 × 5 speckle autocovariance matrix (i.e., 25 speckle autocovariance coefficients; both Δm and Δn are 5) as the texture features. However, the value of $\gamma(0,0)$ is always 1 for the normalized speckle autocovariance matrix. Hence, we discard $\gamma(0,0)$ and the other 24 speckle autocovariance coefficients are formed as a 24-D image feature vector.

Table 1 lists the mean and standard deviation of the speckle autocovariance coefficients for benign cases and malignant ones and the mean difference between the two groups. The significant mean differences of the two groups can prove that speckle autocovariance coefficients are good features for distinguishing benign cases from malignant ones. Moreover, Fig. 3 shows that, the larger Δm and Δn are, the larger the mean difference is.

Classifying with Support Vector Machine Model

SVM is a very good classification tool. It has been proved as a very effective method for many applications such as pattern recognition, machine learning and data mining. In this section, we will give a brief introduction to SVM.

Optimal separating hyperplane. Let $\{(\mathbf{x}_i, y_i), i = 1, \ldots, N\}$ be a training example set S; each example $x_i \in R^n$ belongs to a class labeled by $y_i \in \{-1,1\}$. The goal is to define a hyperplane which divides S, such that all the points with the same label are on the same side of the hyperplane while maximizing the distance between the two classes and the hyperplane. This means to find a pair

Table 1. The mean, SD and mean difference of the speckle autocovariance coefficients for benign cases and malignant ones

	Benign case		Malignant case		Mean diff.
	Mean	SD	Mean	SD	
A(1,0)	0.961	0.012	0.968	0.007	0.007
A(2,0)	0.879	0.033	0.898	0.019	0.019
A(3,0)	0.791	0.056	0.823	0.030	0.032
A(4,0)	0.723	0.074	0.764	0.040	0.041
A(0,1)	0.889	0.028	0.903	0.016	0.014
A(1,1)	0.862	0.033	0.881	0.019	0.019
A(2,1)	0.803	0.047	0.832	0.026	0.029
A(3,1)	0.738	0.062	0.777	0.035	0.039
A(4,1)	0.687	0.075	0.734	0.043	0.047
A(0,2)	0.710	0.063	0.753	0.038	0.043
A(1,2)	0.699	0.064	0.745	0.038	0.046
A(2,2)	0.673	0.069	0.725	0.041	0.052
A(3,2)	0.643	0.075	0.701	0.044	0.058
A(4,2)	0.618	0.082	0.682	0.049	0.063
A(0,3)	0.582	0.080	0.669	0.049	0.087
A(1,3)	0.577	0.080	0.667	0.049	0.090
A(2,3)	0.567	0.081	0.661	0.049	0.094
A(3,3)	0.553	0.084	0.651	0.049	0.098
A(4,3)	0.540	0.089	0.641	0.051	0.101
A(0,4)	0.502	0.085	0.633	0.055	0.131
A(1,4)	0.499	0.085	0.632	0.055	0.133
A(2,4)	0.493	0.087	0.628	0.055	0.135
A(3,4)	0.483	0.090	0.620	0.055	0.137
A(4,4)	0.472	0.094	0.610	0.056	0.138

$(\mathbf{w}, b)$ such that

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1, \quad i = 1,\ldots, N \quad (6)$$

where $\mathbf{w} \in R^n$ and $b \in R$. After finding the pair $(\mathbf{w}, b)$, we can achieve a separating hyperplane (see Fig. 4) and its equation is defined as follows:

$$\mathbf{w} \cdot \mathbf{x} + b = 0. \quad (7)$$

Maybe there are many separating hyperplanes; the OSH is a hyperplane for which the distance to the closest point is maximal. Hence, in order to find the OSH, we must minimize $\|\mathbf{w}\|^2$ under constraint eqn (6).
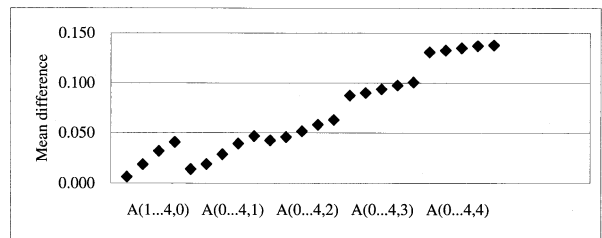

Fig. 3. The diagram of the mean difference between the speckle autocovariance coefficients for benign cases and malignant ones.
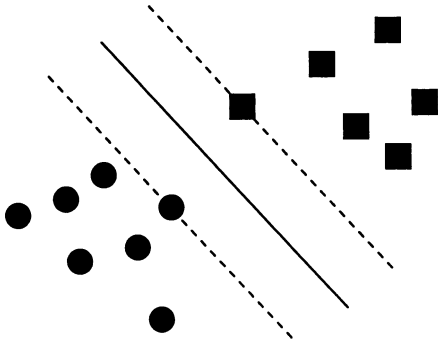
Fig. 4. Separating hyperplane. The dashed lines identify the margin.

If we denote with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)$ the N nonnegative Lagrange multipliers associated with constraint eqn (6), the problem of finding OSH is equivalent to the maximization of the function

$$W(\boldsymbol{\alpha}) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j, \qquad (8)$$

where $\alpha_i \geq 0$ and under constraint

$$\sum_{i=1}^{N} y_i \alpha_i = 0. \qquad (9)$$

Once we achieve the vector $\bar{\boldsymbol{\alpha}} = (\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_N)$, the solution of eqn (7), the OSH ($\bar{\mathbf{w}}$, b), has the following expansion:

$$\bar{\mathbf{w}} = \sum_{i=1}^{N} \bar{\alpha}_i y_i x_i, \qquad (10)$$

while $\bar{b}$ can be determined from $\bar{\boldsymbol{\alpha}}$ and from the Kühn-Tucker conditions (Kuhn and Tucker 1961)

$$\bar{\alpha}_i(y_i(\bar{\mathbf{w}} \cdot \mathbf{x}_i + \bar{b}) - 1) = 0, \quad i = 1,2, \ldots, N. \quad (11)$$

At last, the decision function of classifying a new data point $\mathbf{x}$ can be written as:

$$f(\mathbf{x}) = \mathrm{sgn}(\sum_{i=1}^{N} \bar{\alpha}_i y_i \mathbf{x}_i \cdot \mathbf{x} + \bar{b}). \qquad (12)$$

Linearly nonseparable case. If the set S is not linearly separable, we must introduce N nonnegative slack variables $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_N)$, such that

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad i = 1,2, \ldots, N. \quad (13)$$

Hence, we can regard the generalized OSH as the solution of the following minimizing problem:

$$\frac{1}{2}\mathbf{w} \cdot \mathbf{w} + C\sum_{i=1}^{N}\xi_i, \qquad (14)$$

where C is a regularization parameter. When the parameter C is small, the OSH tends to maximize the distance $1/\|\mathbf{w}\|$; on the contrary, a larger C will lead the OSH to minimize the number of misclassified points.

Nonlinear Support Vector Machines

The training example set that we want to classify is usually linearly nonseparable. To achieve better generalization performance, the input data can first be mapped into a high-dimensional feature space. Then the OSH is constructed in the feature space.

If $\Phi(\mathbf{x})$ denotes a mapping function that maps $\mathbf{x}$ into a high-dimensional feature space, eqn (7) is then rewritten as follows:

$$W(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j). \quad (15)$$

Now, let $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$, we can rewrite eqn (13) as

$$W(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j), \qquad (16)$$

where K is called a kernel function and must satisfy Mercer's theorem (Vapnik 1995). Finally, the decision function becomes

$$f(\mathbf{x}) = \mathrm{sgn}(\sum_{i=1}^{N}\alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b). \qquad (17)$$

In this study, we use nonlinear SVM with Gaussian radial basis kernel as our classifier. The speckle autocovariance texture information is used as the inputs to find an OSH for distinguishing the benign tumors from malignant ones.

SIMULATIONS AND RESULTS

In our experiments, there are 250 pathology-proven cases (including 140 benign breast tumors and 110 malignant ones) in the database. These ultrasonic images are randomly divided into five groups. We first set the first

Table 2. Classification of breast tumors by SVM with speckle-emphasis, conventional all-pixels, and nonspeckle-emphasis texture features

Sonographic classification	Speckle-emphasis		Conventional all pixels		Nonspeckle-emphasis	
	Benign*	Malignant*	Benign*	Malignant*	Benign*	Malignant*
Benign	TN 128	FN 5	TN 116	FN 12	TN 112	FN 15
Malignant	FP 12	TP 105	FP 24	TP 98	FP 28	TP 95
Total	140	110	140	110	140	110

* Histological finding %TP = true positive; TN = true negative; FP = false positive; FN = false negative.

group as a testing group and use the remaining four groups to train the SVM. After training, the SVM is then tested on the first group. Then we set the second group as a testing group and the remaining four groups are trained and then the SVM is tested on the second. This process is repeated until all five groups have been set in turn as a testing group.

In this work, we use a nonlinear SVM with Gaussian radial basis kernel as our classifier where C and γ are 975 and 0.061, respectively. In order to prove that our method is effective, we also calculate autocovariance coefficients of all pixels and pixels excluding speckle, respectively, and then utilize these coefficients as input of SVM.

These simulations are made on a single CPU Intel Pentium-VI 2 GHz personal computer with Microsoft Windows XP operating system. We list the experimental result in Table 2.

To estimate the performance of the experimental result, five objective indices are used. These indices are accuracy, sensitivity, specificity, positive predictive value and negative predictive value. In our experiment, accuracy of SVM with speckle information for classifying malignancies is 93.2% (233/250), the sensitivity is 95.45% (105/110), the specificity is 91.43% (128/140), the positive predictive value is 89.74% (105/117) and the negative predictive value is 96.24% (128/133). Besides, to prove that the speckle information is helpful in classifying tumors, we also use the chi-squared test. Sum-

mary of performances including the accuracy, sensitivity, specificity, positive predictive value, the negative predictive value and p-value among these different features are listed in Table 3. In Table 3, P1 is p-value of speckle-emphasis texture feature and conventional all pixels' texture feature; P2 is the p-value of speckle-emphasis texture feature and texture feature of nonspeckle pixels. We also show the ROC analysis of three features in Fig. 5. According to the Table 3 and Fig. 5, the classification ability for breast tumors using autocovariance coefficients of only speckles is the better choice among these three features.

DISCUSSION

Texture features are helpful to classify benign and malignant tumors on sonography. The potential of sonographic texture analysis to improve breast tumor diagnosis has already been demonstrated (Goldberg et al. 1992; Garra et al. 1993). Texture analysis can be briefly classified in three main groups: these are the models, the mathematical morphology and the statistical method. Autocovariance texture is a kind of the statistical method. It is easy to implement and not affected by brightness. In the proposed diagnostic system, autocovariance texture features and a SVM model are used to classify tumors. We combine speckle information with autocovariance as features of sonography to classify breast tumors. Surprisingly, from the result of this study

Table 3. Summary of performance among different texture features

	Speckle-emphasis (%)	Conventional all-pixels (%)	Nonspeckle-emphasis (%)	P1	P2
Accuracy	93.20	85.60	82.80	<0.01	<0.001
Sensitivity	95.45	89.09	86.36	NS	<0.025
Specificity	91.43	82.86	80.00	<0.05	<0.025
PPV	89.74	80.33	77.24	<0.05	<0.025
NPV	96.24	90.63	88.19	NS	<0.025

Accuracy = (TP + TN)/(TP + TN + FP + FN); sensitivity = TP/(TP + FN); specificity = TN/(TN + FP); PPV = TP/(TP + FP); NPV = TN/(TN + FN); P1: p value of speckle-emphasis and conventional all-pixels texture features; P2: p value of speckle-emphasis and nonspeckle-emphasis texture features.
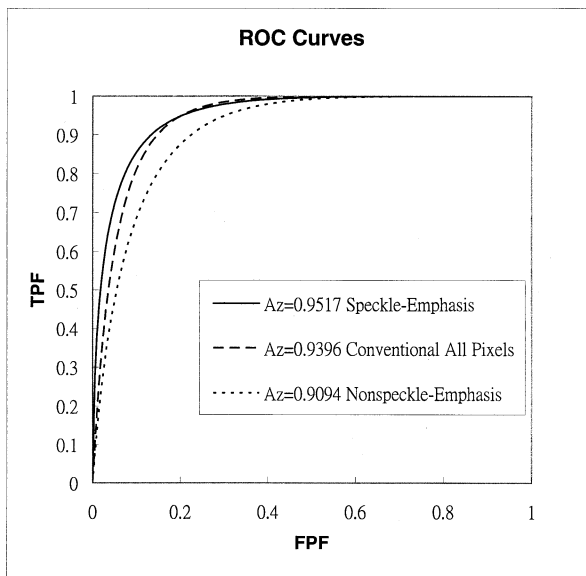
Fig. 5. The ROC analysis of three features.

(see Table 3), we found that the diagnostic performance using speckle information is better than those of other two features. Speckle is a special characteristic of sonography and it can be found when imaging breast tissue at typical frequency. Speckle occurs when the size of scatterers is small compared with a wavelength and there are many scatterers within one sample volume. The distribution pattern of backscatters may have related with surrounding tissue and therefore the texture information around the speckle may be used as a classification characteristic. For this purpose, although speckle is considered to be noise, there exists information behind the scene. Some authors had utilized it to track internal motion due to applied deformation (Skovoroda et al. 1995; Steinberg et al. 1998) and to position the 3D ultrasound frames (Tuthill et al. 1998). From the result of this study, we confirmed that texture information surrounding speckle pixels provides important clues for the applications of texture analysis in the future development of CAD systems. The conventional texture analysis methods using all pixels may include much unrelated information, such as nontexture areas for an ultrasonic image and this unrelated information may interfere with the diagnostic accuracy. Nonspeckle pixels exclude the texture information surrounding speckle pixels which may have diagnostic aids according to this study, so as to make the relatively suboptimal performance among three methods. Speckle phenomenon in sonography is considered as noise that can intrude on judgments of a physician using naked eyes but it is another story for application in a computer-aided diagnosis algorithm.

The main difference among our study and those of (Goldberg et al. 1992 and Garra et al. 1993) is the range of ROI. For the sake of rectangular ROI, we analyze not only the internal texture but beyond the boundary of a breast tumor. Besides, a direct comparison between the results of our study with the results of their earlier studies is inaccessible, because of different dataset, classifier and ROI selections. Incongruities between our study and previous ones could be due to either the specific population of cases examined or the difference methods employed. Successful applications of the CAD strategy have been reported for other types of texture analysis, but some difference in the evaluating results among this study and other ones is that we utilize the speckle information to improve diagnostic performance.

SUMMARY

In this paper, we propose a novel diagnosis system based on a SVM model combining with speckle information to differentiate benign and malignant tumors. In our experiment, the accuracy of our study was 93.2% and the sensitivity was up to 95.45%. In addition, owing to the higher PPV and NPV of this study, the number of biopsies of benign lesions by using these methods can be lower. This would increase the diagnostic confidence, offering a second reading to help reduce misdiagnosis.

REFERENCES

Bassett LW, Liu TH, Giuliano AE, Gold RH. The prevalence of carcinoma in palpable vs impalpable, mammographically detected lesions. AJR 1991;157:21–24.
Bothorel S, Meunier BB, Muller SA. Fuzzy logic based approach for semilogical analysis of microcalcification in mammographic images. International Journal of Intelligent System 1997;12(11–12):819–848.
Chapelle O, Haffner P, Vapnik VN. Support vector machines for histogram-based image classification. Neural Networks, IEEE Transactions on 1999; 10(5):1055–1064.
Chen DR, Chang RF, Huang YL. Breast cancer diagnosis using self-organizing map for sonography. Ultrasound Med Biol 2000;26(3):405–411.
Chen DR, Kuo WJ, Chang RF, Moon WK, Lee CC. Use of the bootstrap technique with small training sets for computer-aided diagnosis in breast ultrasound. Ultrasound Med Biol 2002;28(7):897–902.
Dutt V, Greenleaf JF. Adaptive speckle reduction filter for log-compressed B-scan images. Medical Imaging, IEEE Transactions on 1996; 15(6):802–813.
Garra BS, Krasner BH, Horii SC, et al. Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. Ultrason Imaging 1993;15(4):267–285.
Gisvold JJ, Martin JK Jr. Prebiopsy localization of nonpalpable breast lesions. AJR Am J Roentgenol 1984;143(3):477–481.
Goldberg V, Manduca A, Ewert DL, Gisvold JJ, Greenleaf JF. Improvement in specificity of ultrasonography for diagnosis of breast tumors by means of artificial intelligence. Med Phys 1992;19(6):1475–1481.

Gonzalez RC, Woods RE. Image compression. Digital Image Processing. Reading, Massachusetts: Addison-Wesley, 1992:312–315.

Huo Z, Giger ML, Olopade OI, et al. Computerized Analysis of Digitized Mammograms of BRCA1 and BRCA2 Gene Mutation Carriers. Radiology 2002;225(2):519–526.

Kuhn HK, Tucker AW. Nonlinear programming. Berkeley, CA: Berkeley Symp. Math. Stat. Probab., 1961:481–492.

Petrick N, Chan HP, Wei D, et al. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996;23(10):1685–1696.

Pontil M, Verri A. Support vector machines for 3D object recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on 1998; 20(6):637–646.

Rubin M, Horiuchi K, Joy N, et al. Use of fine needle aspiration for solid breast lesions is accurate and cost-effective. Am J Surg 1997;174(6):694–696.

Skovoroda AR, Emelianov SY, O'Donnell M. Tissue elasticity reconstruction based on ultrasonic displacement and strain images. Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 1995; 42(4):747–765.

Smith WL, Fenster A. Optimum scan spacing for three-dimensional ultrasound by speckle statistics. Ultrasound Med Biol 2000;26(4): 551–562.

Steinberg BD, Sullivan DC, Carlson DL. Disparity mapping applied to sonography of the breast: technical note. Radiology 1998;207(2): 545–550.

Trahey GE, Allison JW, Smith SW, von Ramm OT. A quantitative approach to speckle reduction via frequency compounding. Ultrason Imaging 1986;8:151–164.

Tuthill TA, Krucker JF, Fowlkes JB, Carson PL. Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. Radiology 1998;209(2):575–582.

Valckx FM, Thijssen JM. Characterization of echographic image texture by cooccurrence matrix parameters. Ultrasound Med Biol 1997;23(4):559–571.

Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer-Verlag, 1995.

Wells PNT. Current status and future technical advances of ultrasonic imaging. IEEE Engineering in Medicine and Biology Magazine 2000;19(5):14–20.