ORIGINAL PAPER

# Computer Aided Diagnosis System for Breast Cancer Based on Color Doppler Flow Imaging

Yan Liu  $\cdot$  H. D. Cheng  $\cdot$  J. H. Huang  $\cdot$  Y. T. Zhang  $\cdot$  X. L. Tang  $\cdot$  J. W. Tian  $\cdot$  Y. Wang

Received: 18 May 2012 / Accepted: 26 June 2012 / Published online: 13 July 2012  $©$  Springer Science+Business Media, LLC 2012

Abstract Color Doppler flow imaging takes a great value in diagnosing and classifying benign and malignant breast lesions. However, scanning of color Doppler sonography is operator-dependent and ineffective. In this paper, a novel breast classification system based on B-Mode ultrasound and color Doppler flow imaging is proposed. First, different feature extraction methods were used to obtain the texture and geometric features from B-Mode ultrasound images. In color Doppler feature extraction stage, several spectrum features are extracted by applying blood flow velocity analysis to Doppler signals. Moreover, a velocity coherent vector method is proposed based on color coherence vector, which is helpful for designing to the optimize detection of flow indices from different blood flow velocity fields automatically. Finally, a support vector machine classifier with selected feature vectors is used to classify breast tumors into benign and malignant. The experimental results demonstrate that the proposed computer-aided diagnosis system is useful for reducing the unnecessary biopsy and death rate.

Y. Liu · H. D. Cheng · J. H. Huang · Y. T. Zhang · X. L. Tang School of Computer Science and Technology, Harbin Institute of Technology, Harbin, No. 92, Xidazhi Street,

Harbin 150001, People's Republic of China

Y. Liu  $(\boxtimes)$ College of Science, Harbin Institute of Technology, Harbin, No. 92, Xidazhi Street, Harbin 150001, People's Republic of China e-mail: snacly@163.com

H. D. Cheng Department of Computer Science, Utah State University, Logan, UT 84322, USA

J. W. Tian : Y. Wang Second Affiliated Hospital of Harbin Medical University, Harbin, China

Keywords Color Doppler flow imaging . Feature extraction . Color coherence vector . Computer-aided diagnosis system

### Introduction

Breast cancer is a leading cause of death among women, and its incidence is continuing rising [[1\]](#page-6-0). Early detection is an effective way to control the disease [\[2](#page-6-0)]. Due to its noninvasive, practically harmless, and inexpensive characteristics, ultrasound (US) has become one of the most effective and popular approaches for the early detection of breast cancer [\[3](#page-6-0)]. Computer-aided diagnosis (CAD) system can automatically process images and help radiologists in interpreting sonography [[4\]](#page-6-0). Many CAD systems for distinguishing between benign and malignant breast lesions based on B-Mode ultrasound (BUS) images have been published [\[5](#page-6-0)–[7](#page-7-0)]. BUS has a high sensitivity for the detection of breast abnormalities, but it has limited specificity in differentiating benign and malignant lesions [[8\]](#page-7-0).

Doppler ultrasound takes a great value in diagnosing and classifying benign and malignant breast lesions. It has been proved that Doppler ultrasonic imaging can improve the diagnostic accuracy of gray-scale imaging in the prenatal detection of abnormalities of the blood flow and tumor vascularity [\[9](#page-7-0)]. Previous CAD systems focused on quantitative the Doppler image by calculating the density of color pixels [\[10](#page-7-0)–[12](#page-7-0)]. However, since both benign and malignant tumors may be vascularized, the detection of vessel amount alone may be insufficient to accurately differentiate benign and malignant tumors. The flow indices derived from Doppler sonograms, resistance index (RI) and pulsatility index (PI), are of high clinical value for the evaluation of breast tumors [[13\]](#page-7-0). In traditional clinic examinations, the sample volume was chosen to optimize detection expected from all parts of tumor, include margins. However, scanning of color Doppler sonography is highly dependent on operator's experience and is also not effective. [[14](#page-7-0)] demonstrated that the power spectrum of the Doppler signal is completely determined by the blood flow velocity field. Therefore, quantitative of blood flow velocity field information of color Doppler sonography is important for designing the optimize detection of power spectrum for CAD system.

In this paper, a HSV-space based, computer-aided classification system combining B-mode ultrasound imaging with color Doppler flow imaging (CDFI) is proposed. Different from the most existing systems, the proposed system not only extract the static features, but also extract the dynamic features from color Doppler videos to study blood flow characteristic. First, morphologic and textural features of breast lesion are extracted from BUS images by using different kinds of feature extraction methods. For extracting color Doppler features, a dynamic velocity coherence vector algorithm is proposed for quantitative the blood velocity field information. Then, a novel method of mapping color Doppler imaging color Doppler indices instead of ultrasonic frequency shifts is proposed and several hemodynamic features are extracted. Finally, these features are selected and employed to discriminate benign masses from malignant masses by using support vector machine classifier. The experiment results show that by combining the B-Mode features with the CDFI features, the proposed method is effective and useful for classifying breast lesions. The block diagram is illustrated in Fig. 1.



Fig. 1 The diagram of the proposed system

#### The proposed method

## Materials

A data base including 105 cases (50 benign and 55 malignant cases) is utilized for evaluating the performance of the proposed system. Each case includes one color Doppler video and one BUS image. These cases are obtained from 2nd Affiliated Hospital of Harbin Medical University collected by using HITACHI Vision 900 system (Hitachi Medical System, Tokyo, Japan) equipped with a liner probe having central frequency of 6~13 MHz, and captured from the video signals directly. From each case, 3 morphology features and 4 texture features were extracted from the B-Mode image, 2 color Doppler flow features were extracted from the color Doppler video. The cases were obtained from Feb. 2008 to Mar. 2011, and all the classes of the cases were confirmed by surgery, or pathological examinations, or biopsy. The length of the color Doppler video has 2–7 s with the frame rate, 10 frames per second. Special attention has been paid to protect the privacy of the patients.

Feature extraction from B-mode images

Breast ultrasound (BUS) image segmentation is a crucial and essential step in CAD systems. Because of its low signal/noise ratio, low contrast and blurry boundaries, ultrasound (US) image segmentation is a difficult task. To overcome the problems, in our system, breast lesions are segmented by using an effective segmentation method based on cellular automata (CA) principle [[15](#page-7-0)]. In approach [\[15](#page-7-0)], a 2-D ultrasound image is considered as a cell space, the state of each cell belongs to a finite set and the dynamic changes of the cell states are calculated by the local energy transition function operating on a given neighborhood. In its energy transition function, both global and local image information, and the relationship between neighboring pixels can be integrated conveniently by using different energy transfer strategies, which is helpful for handing blurry boundaries. At each discrete time step, CA updates the entire lattice with the evaluation rule, and the states of cells at the current time step are updated according to the states of the cells at the previous time step. The result of local competitions is that the strongest bacteria occupy the neighboring sites and gradually spread over the image, and then, the class of each pixel as foreground or background can be specified accordingly. The calculation continues until none of cell's state is changed, the automaton converges to a stable configuration and the segmentation is finished. Figure  $2(a)$  is a B-Mode US image with a complicated structure, and the boundary of the tumor has multiple connected regions with the surrounding tissues. Comparing to the manual delineation produced by an experienced radiologist (Fig. [2\(c\)\)](#page-2-0), it is seen that by using approach [[15\]](#page-7-0), the blurry boundary can be treated well (Fig. [2\(b\)](#page-2-0)).

<span id="page-2-0"></span>

Fig. 2 The segmentation result of an B-Mode ultrasound image with blurry boundaries. a. Original B-Mode ultrasound image. b. Segmentation result of (a). c. Lesion regions delineated by the radiologist

From tumor contour, it is possible to establish the first malignant diagnostic hypothesis. Generally, benign tumors usually have smooth shape and malignant tumor tend to have irregular border. In [[16](#page-7-0)], several morphometric parameters including standard deviation (SD), area ratio (AR) and roughness (RH) are proved to be effective in classifying breast tumors and they are used to account for the sonographic observations of tumor. Each feature is defined as:

$$
SD = \sqrt{\frac{1}{N_{cou} - 1} \sum_{c_i=1}^{N_{cou}} \left( d(c_i) - \overline{d(c_i)} \right)^2}
$$
 (1)

$$
AR = \frac{1}{\overline{d(c_i)} \cdot N_{cou}} \sum_{c_i=1}^{N_{cou}} \left( d(c_i) - \overline{d(c_i)} \right)
$$
 (2)

$$
RH = \frac{1}{N_{cou}} \sum_{c_i=1}^{N_{cou}} |d(c_i) - d(c_{i+1})|
$$
\n(3)

where  $d(c_i)$  is the normalized radial length calculated as:

$$
d(c_i) = \frac{\sqrt{(x_{c_i} - X_0)^2 + (y_{c_i} - Y_0)^2}}{\max(d(c_i))}, 1 \le c_i \le N_{cou}
$$
 (4)

where  $X_0$  and  $Y_0$  are the coordinates of the centroid, and  $(x_{ci},y_{ci})$  are the coordinates of boundary pixel at the c<sub>i</sub>th location, respectively.  $N_{cou}$  is the number of contour pixels and max  $(d(c_i))$  is the maximum value of the radial length of the tumor.

Textural information extracted from sonography has been found to be useful to classify breast tumors. In this study, B-Mode texture information is derived from the co-occurrence matrix [[4](#page-6-0)]. Each element of the cooccurrence matrix  $C(i, j, d, \theta)$  is defined as the joint probability of the gray levels separated by distance d and along direction  $\theta$  of the image.

$$
C_{(d,\theta)}(i,j,d,\theta)
$$
  
=  $||{((x_1,y_1),(x_2,y_2)|x_2 - x_1 = d\cos\theta,y_2 - y_1 = d\sin\theta,I(x_1,y_1) = i,I(x_2,y_2) = j}||$ 

where  $(x_1, y_1)$  and  $(x_2, y_2)$  are the pixels in the lesion region.  $I(\cdot)$  is the intensity and  $\|\cdot\|$  is the number of pixel pairs matching to the conditions. In propose method, 4 texture features: energy  $(f_{ENG})$ , entropy  $(f_{ENT})$ , contrast  $(f_{CON})$  and homogeneity  $(f_{HOM})$  are extracted from cooccurrence matrix with 4 distance  $d(d=1, 2, 3, 4)$  and along 4 directions  $\theta$  ( $\theta$ =0°,45°,90°,135°).

$$
f_{ENG} = \sum_{i=0}^{ng} \sum_{j=0}^{ng} P(i,j)^2
$$
 (6)

$$
f_{ENT} = -\sum_{i=0}^{ng} \sum_{j=0}^{ng} P(i,j) \log P(i,j)
$$
 (7)

$$
f_{CON} = \sum_{i=0}^{ng} \sum_{j=0}^{ng} (i-j)^2 P(i,j)
$$
 (8)

$$
f_{HOM} = \sum_{i=0}^{ng} \sum_{j=0}^{ng} \frac{P(i,j)}{1 + |i - j|}
$$
(9)

where  $P(i, j)$  is an element of the normalized co-occurrence matrix. The above texture descriptors of the same distance will be averaged to reduce the dimension of the feature vectors. The number of gray levels ng for establishing the co-occurrence matrix is:

$$
ng = \max(I(x, y)/4)
$$
 (10)

 $(5)$ 

Feature extraction from color Doppler flow images

#### Vascular detection

A 2-D color Doppler ultrasound image is essentially an array of pixel set represented in red, greed and blue (RGB) color system. The magnitude of the blood velocity is represented by different color shades. Conventionally, the red color is assigned to indicate the blood flow toward the transducer, and blue is represented the flow away from the transducer. To extract the vascular points, we converted the original color Doppler image into binary by applying a predefined threshold value  $T_c$  to examine the primary color channel for each pixel. The value of  $T_c$  was selected to be 40 by the experiment. For each pixel pin a color Doppler image, if the intensity differences between R, G and B are equal or greater than  $T_c$ , pis a vascular pixel  $(p_c)$ . Otherwise, pis a tissue (the lesion tissue  $\&$  the normal tissue) pixel.

#### Vascularity feature extraction

Color Doppler sonography and their flow spectrum indices are measured from the spectral Doppler tracing [[13\]](#page-7-0). The power spectrum of the Doppler signal is completely determined by the blood flow velocity field. For a color Doppler, in the center of large vessels with plug flow, where most flow is at a uniform velocity, the mean frequency is close to the peak. Nonuniform velocity within a sample volumesuch as in areas of turbulence, at curves, bifurcations, or confluences, near vessel walls, or where the diameter of a vessel changes is depicted in a Doppler spectral display as spectral broadening. This causes the mean frequency to be considerably less than the peak frequency [\[17](#page-7-0)]. According to this clinic factor, we develop a velocity coherence vector method based on color coherence vector to measure whether local velocities in different regions are uniform. Then, to assess the optimize detection of blood flow indices of the lesion.

Color coherence vector is defined as a degree to which pixels of that color bin are members of large similarlycolored regions [[18\]](#page-7-0). It measures whether some pixels with a discretized color are coherent or not. The classification of pixels as either coherent or incoherent depends on the size in pixels of its connected component. Pixels are coherent if the number of pixels in connected region exceed a threshold value  $\tau$ ; Otherwise, they are incoherent. For establishing velocity coherence vector, it is necessary to extract the velocity information from continuous wave Doppler signal firstly. Usually, low-velocity region in color Doppler is much duller than high-velocity region. According to this observation, in this study, progressively increasing velocities are encoded in varying value (V) component of brightness by transforming the color Doppler image from RGB color space to HSV color space [[19\]](#page-7-0). Then a local discrete Doppler wave signal (DDW) at the vascular point  $p_c$  is constructed from the a series of spitted color Doppler video frames:  $DDW_{p_c} = [V_{p_c}(1), V_{p_c}(2), V_{p_c}(3), ..., V_{p_c}(F)]$ , where  $F$  is the number of frames (Fig. [3\)](#page-4-0). The problem of selecting a pair of vascular points  $(p_c, q_c)$  with the uniform velocity is handled by using a similarity-based criterion as the following [[20\]](#page-7-0).

$$
\frac{1}{F} \cdot \left| \sum_{t=1}^{F} V_{p_c}(t) - \sum_{t=1}^{F} V_{q_c}(t) \right| \le T \tag{11}
$$

where  $T$  is the threshold value of similarity criterion set to:  $T = \sqrt{2\sigma}$ , where  $\sigma$  is the local standard deviation of the neighborhood. The coherent region with different veloc neighborhood. The coherent region with different velocity is determined by computing the size of connected components. A connected component  $C$  is the maximal set of points such that for any two pixels  $p, pnewapos;$  $\in C$  with the uniform velocity, there is a path in C between  $p$  and  $p'$ . A path in C is a sequence of pixels  $p = p_1, p_2, \dots, p_{nnewapos} = p'$  such that each pixel  $p_i'$  is in C and any two sequential pixels  $p_{i'}$ ,  $p_{i'+1}$  are adjacent to each other.

For a series of color Doppler video frames, let  $N_k$ be the set of vascular points with the velocity  $V_{\kappa} = \frac{1}{F} \sum_{t=1}^{F} V_{p_c}(t)$ ,  $\alpha_{\kappa}$ be the number of coherent elements in $N_{\kappa}$  and  $\beta_{\kappa}$ be the number of incoherent elements. Thus, the number of total number of vascular points with mean velocity  $V_{\kappa}$ is  $V_k = \alpha_k + \beta_k$ , and so a 'velocity' histogram can be summarized an image as:  $V_1, V_2, ..., V_{n_c}$ . Instead, for each velocity we compute the pair( $\alpha_{\kappa}$ ,  $\beta_{k}$ ), which we call the coherence pair for the velocity  $V_{\kappa}$ . The velocity coherence vector for the color Doppler image consists of  $\langle (\alpha_1, \beta_1), ..., (\alpha_{nc}, \beta_{nc}) \rangle$ , where ncis the number of discretized velocities. For reducing the computational complexity and preserving the details, the number of value levels *nc* for establishing the velocity coherence vector is:

$$
nc = \max(V(x, y)/4)
$$
 (12)

The threshold value  $\tau$  for determining the coherent region of the image is selected according to the size of sample volume on the color Doppler image. For a  $600 \times 400$  image, the value of  $\tau$  is set to 15×15 pixels by the experiment.

The amplitude of color Doppler signal is stored in colorencoded information that serves as index. The indices derived from the sonograms defined as receptivity (RI) index, pulsatility index (PI) and they are widely used indices for evaluating Doppler sonograms [[21\]](#page-7-0). Both the RI and PI are reflections of the resistance to blood flow, downstream from the point of insonation. So in this study, we evaluate the <span id="page-4-0"></span>Fig. 3 DDW derived from a series of color Doppler video frames



tumor vascularity by extracting the features of RI and PI. Two spectrum features are defined as:

$$
RI = \frac{V_S - V_D}{V_S} \tag{13}
$$

$$
PI = \frac{V_S - V_D}{V_M} \tag{14}
$$

where  $V_S$   $(V_S = \max(DDW(t)))$  is systolic peak velocity,  $V_D$   $V_D$  = min $(DDW(t))$  is diastolic terminal velocity and  $V_M$   $(V_M = mean(DDW(t)))$  in a corresponding cardiac cycle. Then, from different velocity coherence regions, the highest values of  $V_{\rm S}$ , RI and PI are found and used for analysis [\[21](#page-7-0)].

#### Support vector classification

tics (syst1[Ref. [6\]](#page-6-0), syst2

[Ref. [12\]](#page-7-0))

In this paper, support vector machine (SVM) is used for classifying the breast lesions into two classes (benign and malignant) because of its high generalization performance without need to add a priori knowledge [\[22](#page-7-0)]. The aim of SVM is to devise of learning separating hyperplanes in a high-dimensional feature space. The unknown sample  $x'$  is determined:

$$
f(x) = sign\left(\sum_{i'}^{Nsv} \alpha_{i'} y_{i'} k(s_{i'}, x') + b\right)
$$
\n(15)

where  $\alpha_i$  is a positive LaGrange multiplier, Nsvis the number of support vectors, and  $k(s, x')$  is the convolution number of support vectors, and  $k(s_i, x')$  is the convolution<br>of the kernel of the decision function. BBE kernel nonof the kernel of the decision function. RBF kernel nonlinearly maps samples into a higher dimensional space, unlike the linear kernel; it can handle the case when the relation between class labels and attributes is nonlinear. Furthermore, the linear kernel is a special case of RBF. In addition, the sigmoid kernel behaves like RBF for certain parameters. The polynomial kernel has more hyper parameters than the RBF kernel and so high complexity than RBF. So, RBF kernel defined in Eq. (16) is selected as the kernel function of proposed system.

$$
k(x, x_s) = \exp\left(-\gamma |x' - x'_s|^2\right) \tag{16}
$$

where  $x'_{s}$  is a support vector obtained from training and  $\gamma$  is the kernel parameters.

# Classifier combination

In this paper, we combine the prediction variable of each individual feature (one per classifier) to obtain the benefit of mutual compensation to improve the classification performance. The ensemble decision of each unknown sample  $l(x)$  is designed by employing an optimal combination rule of weighted majority voting algorithm [[23\]](#page-7-0):



<span id="page-5-0"></span>Table 2 Classification performances of the features with different characteristics [Ref. (4, 6, 12 and 16)]

| Features              | ACC(%) |
|-----------------------|--------|
| Color Doppler feature | 87.0   |
| Morphologic feature   | 81.3   |
| Texture feature       | 77.5   |

$$
l(x) = \frac{1}{C} \sum_{k=1}^{tk} l'_k(x) D_k
$$
 (17)

where C is the constraint parameter bounding  $l(x)$  to  $\{1, -1\}$ ,  $l'(x)$  is the prediction variable of kth classifier. To weight each classifier, the distance  $D_k$  of each feature between benign tumor and malignant tumor is calculated to evaluate the effectiveness of each input features [\[12\]](#page-7-0).

$$
D_k = \frac{|f_{b_k} - f_{m_k}|}{\sqrt{\sigma_{b_k}^2 + \sigma_{m_k}^2}}
$$
(18)

where  $f_{b_k}$  is the average and  $\sigma_{b_k}$  is the variance of the kth feature of the benign, while  $f_{m_k}$  and  $\sigma_{m_k}$  are the average and variance of the kth feature of the malignant, respectively. The larger the distance  $D_k$  is, the more important the *kth* classifier is. To compare the features with different characteristics on a common scale, each feature is normalized as:

$$
f_k(x) = \frac{f_k(x) - \min(f_k(ns))}{\max(f_k(ns)) - \min(f_k(ns))}
$$
\n(19)

where *ns* is the number of training vectors.

Fig. 4 The ROC curve of the classification ratio

## Results

To evaluate the classification performance, five error descriptive statistics: accuracy (ACC), sensitivity (SEN), specificity (SPE), positive value (PPV) and negative predictive value (PPV). Define the number of correctly and incorrectly classified malignant tumors as true positive TP and false negative FN, the number of correctly and incorrectly classified benign tumors as true negative TN and false positive FP, respectively. Five descriptive statistics are calculated as:  $ACC = \frac{TP+TN}{TP+TN+FP+FN}$ ;  $SEN = \frac{TP}{TP+FN}$ ;  $SPE = \frac{TN}{TN+FP}$ ;  $PPV = \frac{TP}{TN+FP}$  $\frac{TP}{TP+FP}$ ;  $NPV = \frac{TN}{TN+FN}$ 

In the experiments, leave-one-out method was utilized for SVM training and testing: one case is for testing, and the others are used for training, the process is iterated until all the cases have been used to testing and training. Given a set of training vectors (ns in total) belonging to separate classes:  $(x_1', y_1), (x_2', y_2), \ldots, (x_s', y_s)$ , where  $x_s'$  denotes the *sthinput* vector and  $y_s \in \{1, -1\}$  is the corresponding desired output.<br>When the output of SVM is 1, the system will classify the When the output of SVM is 1, the system will classify the tumor in the image as benign. Conversely, the tumor will be diagnosed as malignant. The error rates or predictive accuracy on the different iterations are averaged to yield an overall error rate, or predictive accuracy. The kernel parameter of SVM classifier is selected according to the classification performance with different  $\gamma$  values. By the experiment, we found that with the  $\gamma$  ranging from 0.01 to 0.023, the SVM classifier obtained a stable and highest accuracy.

In this paper, two other CAD systems: syst 1 [[6\]](#page-6-0), syst 2 [\[12](#page-7-0)] are also utilized to the same cases for comparison. Syst



<span id="page-6-0"></span>1 extracted the autocovariance texture features and solidity morphologic features from B-Mode ultrasound images, then, a SVM classifier was utilized to classify the tumor as benign or malignant. In syst 2, geometric features, textural features and blood flow features were extracted from the B-Mode and the color Doppler images and employed to differentiate the benign and malignant lesions. The statistical overall performances of three systems are shown in Table [1](#page-4-0).

From Table [1](#page-4-0) it can be seen that the proposed CAD system achieves a highest overall accuracy  $(ACC=$ 91.42 %). Its sensitivity is 92.73 %, and it means the more malignant tumors can be detected by the proposed approach than two other approaches. This is because, different to other systems, the benefit of mutual compensation of proposed system can be obtained from both global features (morphologic and texture features) and local features (local blood flow features). The specificity value (90.00 %) demonstrates that the number of benign tumors incorrectly classified as malignant is lower than other approaches, i.e., by using proposed CAD system, the fewer benign tumors need further examination. Furthermore, the lower NPV indicate that fewer malignant tumors were wrongly classified, which is helpful to reduce the death rate.

To offer a more comprehensive comparison, we compare the classification performances of different classifiers trained by integrating the morphology features, texture features and color Doppler flow features extracted by works [4, 6, [12,](#page-7-0) [16\]](#page-7-0) and the proposed system to validate the efficiency of the proposed feature extraction strategy. The classification accuracies with different feature combinations are listed in Table [2.](#page-5-0)

It is seen from Table [2](#page-5-0) that the corresponding classification accuracy of color Doppler features extracted by proposed system is much larger than those of the features in other related works. This demonstrates that the proposed feature extraction strategy is reliable and useful in discriminating benign and malignant tumors.

The second criterion for evaluating the performance of the classifier system is the receiver operator characteristic curve (ROC) [[24\]](#page-7-0). The curves were constructed by computing the sensitivity and specificity of increasing numbers of clinical findings in predicting step. The most commonly used for global index of diagnostic accuracy is the area under the ROC curve (AUC). Values of AUC close to 1.0 indicate that the marker has a high diagnostic accuracy. From Fig. [4](#page-5-0) it can be seen that the AUC value of the proposed method is the much higher (0.9455) than that of other two approaches (0.8842, 0.9040), this demonstrates that the proposed method is more accurate.

#### **Discussion**

In this paper, an effective classification system of breast ultrasound image was proposed. According to the principle of color Doppler imaging, the blood flow velocity information is encoded in varying value component of brightness by transforming the color Doppler flow image from RGB color space to HSV color space. In addition, a velocity coherence vector algorithm was proposed based on color coherence vector algorithm, which is helpful for estimating the blood velocity field information quantitatively. By using proposed system, the optimize detection of blood flow indices can be designed automatically without visual inspection. Ensemble of classifiers with different feature characteristics to obtain the benefit of mutual compensation was also dealt with. The diagnosis using different inputs (feature vectors) in classification of benign and malignant tumors was performed by support vector machine classifier, and the assessment of feature extraction and classification was executed by evaluating the classification performances of classifiers with different feature characteristics and their combinations. The accuracy for classifying benign and malignant tumors by integrating the B-Mode features with color Doppler features was 91.42 % with 92.73 % sensitivity and 90.00 % specificity. The AUC value of the proposed method is 0.9455. The experimental results demonstrate that the proposed CAD system is very effective and useful for classifying the breast tumors.

Acknowledgments Financial support from the National Nature Science Foundation of China (NSFC) greatly appreciated; Grant numbers: 81071216, 61100097, and 81101103.

#### References

- 1. Li, J. B., Mammographic image based breast tissue classification with kernel self-optimized fisher discriminant for beast cancer diagnosis. Journal of Medical Systems. doi[:10.1007/s10916-011-](http://dx.doi.org/10.1007/s10916-011-9691-4) [9691-4](http://dx.doi.org/10.1007/s10916-011-9691-4).
- 2. Manglem Singh, Kh, Fuzzy rule based median filter for gray-scale images. Journal of Information Hiding and Multimedia Signal Processing 2(2):108–122, 2011.
- 3. Li, J. B., Yu, Y., Yang, Z. M., Tang, L. L., Breast tissue image classification based on semi-supervised locality discriminant projection with kernels. Journal of Medical Systems. doi[:10.1007/](http://dx.doi.org/10.1007/s10916-011-9754-6) [s10916-011-9754-6](http://dx.doi.org/10.1007/s10916-011-9754-6)
- 4. Cheng, H. D., Shan, J., Ju, W., Guo, Y. H., and Zhang, L., Automated breast cancer detection and classification using ultrasound images: A survey. Pattern Recognition 43:299–317, 2010.
- 5. Drukker, K., Giger, Maryellen L., Vyborny, Carl J., and Mendelson, Ellen B., Computerized detection and classification of cancer on breast ultrasound. Acad. Radiol. 11:526–535, 2004.
- 6. Chang, R. F., Wu, W. J., Moon, W. K., et al., Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. Breast Cancer Res Treat 89(2):179–185, 2005.
- <span id="page-7-0"></span>7. Liu, B., Cheng, H. D., Huang, J. H., et al., Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis. Pattern Recognition 43(1), 2010.
- 8. Madjar, H., Prompeler, H. J., Del Favero, C., Hackeloer, B. J., and Llull, J. B., A new Doppler signal enhancing agent for flow assessment in breast lesions. Clin. Sci. 1:123–130, 2000.
- 9. Madjar, H., Contrast ultrasound in breast tumor characterization: present situation and future tracks. Eur. Radiol. 11(3):41–46, 2001.
- 10. Wu, C. H., Hsu, M. M., Chang, Y. L., et al., Vascular pathology of malignant cervical lymphadenopathy: Qualitative and quantitative assessment with power doppler ultrasound. Cancer 183(6):1189– 1196, 1998.
- 11. Hsiao, Y. H., Huang, Y. L., Kuo, S. J., et al., Characterization of benign and malignant solid breast masses in harmonic 3D power Doppler imaging. Eur. J. Radiol. 71:89–98, 2009.
- 12. Diao, X. F., Zhang, X. Y., Wang, T. F., et al., Highly sensitive computer aided diagnosis system for breast tumor based on color Doppler flow images. J. Med. Syst 35:801–809, 2011.
- 13. Choi, H. Y., Kim, H. Y., Baek, S. Y., et al., Significance of resistive index in color Doppler ultrasonogram: Differentiation between benign and malignant breast masses. Clinical imaging 23:284–288, 2000.
- 14. Bastos, C. C., Fish, P. J., and Vaz, F., Spectrum of Doppler ultrasound signals from nonstationary blood flow. Ultrasonics, ferroelectrics and frequency control 46(5):1201–1217, 1999.
- 15. Liu, Y., Cheng, H. D., Huang, J. H., et al., An effective approach of lesion segmentation within the breast ultrasound image based on the cellular automata principle. Journal of digital imaging, 2012. doi[:10.1007/s10278-011-9450-6](http://dx.doi.org/10.1007/s10278-011-9450-6).
- 16. Ikeda, O., Nishimura, R., Miyayama, H., et al., Evaluation of tumor angiogenesis using dynamic enhanced magnetic resonance

imaging: comparison of plasma vascular endothelial growth factor, hemodynamic, and pharmacokinetic parameters. Acta Radiologica 45(4):446–452, 2004.

- 17. Mitchell, D. G., Color Doppler imaging: principles, limitations, and artifacts. Radiology 177(1):1–10, 1990.
- 18. Kang, K. H., Yoon, Y. I., Choi, J. S., et al., Additive texture information extraction using color coherence vector. Proceedings of the 7th WSEAS International Conference on Multimedia Systems & Signal Processing 15–17, 2007
- 19. Jessee, E., Wiebe, E., Visual perception and the HSV color system: Exploring Color in the Communications Technology Classroom," International Technology and Engineering Educators Association 7–11, 2008
- 20. Gordon, R., and Rangayyan, R. M., Feature enhancement of film mammograms using fixed and adaptive neighborhoods 23(4):560– 564, 1984.
- 21. Chao, T. C., Lo, Y. F., Chen, S. C., et al., Color Doppler ultrasound in benign and malignant breast tumors. Breast Cancer Research and Treatment 57:193–199, 1999.
- 22. Huang, Y. L., and Chen, D. R., Support Vector machines in sonography application to decision making in the diagnosis of breast cancer. Journal of Clinical Imaging 29:179–784, 2005.
- 23. Sergey, T., Stefan, J., Venu, G., et al., Review of classifier combination methods. Studies in Computational Intelligence: Machine Learning in Document Analysis and Recognition 90:361–686, 2008.
- 24. Zhu, W., Zeng, N., Wang, N., Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations NESUG, 2010.