Intensity-Invariant Texture Analysis for Classification of BI-RADS Category 3 Breast Masses

Chung-Ming Lo[1], Woo Kyung Moon[2*], Chiun-Sheng Huang[3], Jeon-Hor Chen[4,5], Min-Chun Yang[1], and Ruey-Feng Chang[1,6]*

[1]Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan
[2]Department of Radiology, Seoul National University Hospital, Seoul, Korea
[3]Department of Surgery, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan
[4]Center for Functional Onco-Imaging and Department of Radiological Science
University of California Irvine, California, USA
[5]Department of Radiology, E-Da Hospital and I-Shou University, Kaohsiung, Taiwan
[6]Graduate Institute of Biomedical Electronics and Bioinformatics
National Taiwan University, Taipei, Taiwan

* Corresponding Authors:

Professor Woo Kyung Moon
Department of Radiology
Seoul National University Hospital
101 Daehangno, Jongno-gu
Seoul 110-744, KOREA
Telephone: 82-2-2072-3928
Fax: 82-2-743-6385
E-mail: moonwk@snu.ac.kr

Professor Ruey-Feng Chang
Department of Computer Science and Information Engineering
National Taiwan University
Taipei, Taiwan 10617, R.O.C.
Telephone: 886-2-33661503
Fax: 886-2-23628167
E-mail: rfchang@csie.ntu.edu.tw

Abstract

Radiologists likely incorrectly classify benign masses into Breast Imaging Reporting and Data System (BI-RADS) category 3. A computer-aided diagnosis (CAD) system was developed in this study as a second viewer to avoid misclassifying carcinomas. 69 biopsy-proven BI-RADS category 3 cases including 21 malignant and 48 benign masses were used to evaluate the CAD system. To improve the texture features, the gray-scale variations between images were reduced by transforming pixels into intensity-invariant ranklet coefficients. The textures of the tumor and speckle pixels were extracted from the transformed ranklet images to provide more robust features than conventional CAD systems. In the result, tumor texture and speckle texture with ranklet transformation achieved significantly better areas under the receiver operating characteristic curve (Az) compared with those without it (Az=0.83 vs. 0.58 and Az=0.80 vs. 0.56, *p*-value<0.05). The improved CAD system can be a second reader to confirm the classification of BI-RADS category 3 masses.

Keywords: Breast cancer; ultrasound; BI-RADS; computer-aided diagnosis; ranklet

Introduction

Breast ultrasound (US) has been used in cancer detection to distinguish between benign and malignant lesions (Stavros et al. 1995; Kelly et al. 2010; Weigel et al. 2013). US examination has been shown to detect additional breast cancers compared with conventional mammography (Weigel et al. 2013). The sonographic appearance of lesions is interpreted by radiologists for clinical diagnosis. The American College of Radiology (ACR) developed the Breast Imaging Reporting and Data System (BI-RADS) lexicon (Mendelson et al. 2013) to standardize the sonographic descriptors. The descriptors were then quantified in various computer-aided diagnosis (CAD) systems to automatically evaluate the likelihood of malignant tumors (Drukker et al. 2005; Kim et al. 2005; Moon et al. 2012a; Moon et al. 2012b). The quantitative features extracted from the US B-mode images included the tumor shape, texture, and speckle texture features (Moon et al. 2012a; Moon et al. 2012b). The tumor texture was extracted from the delineated tumor area, whereas the speckle texture correlated with detected speckle pixels inside and surrounding the tumor area. With the assistance of CAD systems, clinicians have demonstrated improved performance in diagnosing different masses (Kashikura et al. 2013). In the study by Kashikura et al., the performance indices of the area under the receiver operating characteristic (ROC) curve significantly improved for seven observers (Kashikura et al. 2013).

The ACR also suggested BI-RADS assessment categories for tumor classification and management strategies when US examination is used. Lesions labeled as BI-RADS category 2 (benign) will not be biopsied, and those labeled as BI-RADS category 4 (suspicious abnormality) will definitely be biopsied. Between 2 and 4, BI-RADS category 3 (probably benign) lesions are in a clinical gray zone. Follow up is recommended instead of core needle biopsy due to the low likelihood of malignancy (less than 2%) (Sickles 1994; Berg 2004; Leung and Sickles 2007). This management strategy was supported in the results of previous studies (Graf et al. 2007; Berg et al. 2008; Kim et al. 2012; Gruber et al. 2013). Tumors with circumscribed margins and parallel orientations are considered to likely be benign. However, radiologists should evaluate tumors with more objective suggestions to avoid misclassifying carcinomas into BI-RADS category 3 (Lazarus et al. 2006; Lee et al. 2008). Acting as a second reader, the CAD system (Moon et al. 2013a) distinguished more malignant lesions from benign lesions among category 3 cases. However, the performance of the texture features was not as good as that of the shape features in the CAD system. The texture analysis of ultrasound patterns was system-dependent (Chang et al. 2005). CAD systems based on texture analysis may only perform well in one specific ultrasound system. Different US system settings and different US scanners may result in different performances in texture analysis (Tsui et al. 2010).

With the development of image processing, previous studies (Masotti et al. 2009; Yang et al. 2013) have extracted invariant texture features from the ranklet transformed region of interest (ROI) for tumor detection and diagnosis. The ranklet transform based on the multi-resolution and orientation-selective analysis was invariant to linear/nonlinear grayscale variations (Masotti et al. 2009). Masotti et al. used the gray-scale invariant texture features to detect breast tumors in mammography for false positive reduction (Masotti et al. 2009). Yang et al. extracted the texture features from US images to classify breast tumors (Yang et al. 2013). However, the texture features were extracted from ROIs which may not accurately contribute tumor information in the methods of Masotti et al. and Yang et al.. The current work proposes extracting intensity-invariant texture features from automatically delineated tumor area to obtain more specific tumor characteristics. Compared to Yang et al., the US database used in the current work is BI-RADS category 3 breast masses. Classifying BI-RADS category 3 cases having less than 2% malignancy assessed by the radiologists is more challenging to CAD systems. The performance result would be closer to the real clinical examination using CAD systems as second readers.

Materials and Methods

Fig. 1 shows the flowchart of the proposed CAD system. First, tumor contour segmentation is performed by the CAD system to separate the specific tumor area from the background tissues in the US image. According to the delineated tumor contour, the area enclosed by the tumor contour was defined as the tumor area and the extended area within a distance of 5 pixels to the tumor contour was the area for speckle detection. The ranklet-transformed images were then submitted to the procedure after tumor area segmentation to provide the transformed pixel values of the defined tumor area and speckle area for intensity-invariant texture extraction. Based on the biopsy-proven pathology, the diagnostic performances of the two ranklet-transformed texture sets were calculated using binary logistic regression.

Patients and data acquisition

Approval was obtained from the institutional review board of Seoul National University Hospital, and informed consent was waived for this retrospective study. The breast US data were collected using an ATL HDI 3000 scanner (Philips, Bothell, WA) or a Medison Voluson 530 scanner (Kretz Technik, Zipf, Austria) during a 2-year period. One hundred consecutive tumors were acquired before needle biopsy. Initially, the cases were assessed as 32 BI-RADS category 3, 56 category 4, and 12 category 5 masses. Five radiologists performed blinded retrospective interpretation to assess the BI-RADS category for each tumor. Sixty-nine masses including 21 malignant and 48 benign masses were assigned to BI-RADS category 3 by at least one of the five radiologists in the experiment. Twenty-one malignant lesions were classified as histological grade 2 (6 cases) or grade 3 (15 cases) invasive ductal carcinomas (IDC). The size ranged from 1.2-4.7 cm (mean=2.7 cm). The benign lesions were composed of 34 fibroadenomas (FA), 13 fibrocystic changes (FCC), and 1 papilloma. The size ranged 1.4-4.3 cm (mean=2.6 cm). The patients' ages ranged from 20-84 years (mean=43 years). Patients with malignant lesions had age range of 28-84 years (mean=47.2 years). For benign cases, the patients' ages ranged from 20-53 years (mean=39.9 years). The age difference between the benign and malignant groups was statistically significant (p-value=0.001). Fifteen lesions were palpable and 54 lesions were nonpalpable. The illustrations in Fig. 2 (a) show the cropped tumors with surrounding fat and the posterior area from the acquired US images. The figures shown in Fig. 2 (a) from left to right are a FA, an IDC acquired by ATL, a FCC, and an IDC obtained by Medison. Whether the image cases were acquired from either ATL or Medison, the brightness and contrast of the acquired US images were various.

Tumor segmentation

First, the tumor area was delineated for quantitative feature extraction. To reduce operator dependence, the level-set segmentation method (Moon et al. 2013a) was used to automatically delineate the tumor contour in the original US images. Using the grayscale gradient as the criterion in the differential equation, the level-set function developed the user-defined seeds into a complex shape with changing topology to obtain the tumor contour. To address the tumors with weak edges, a sigmoid filter (Suri 2008) was used to enhance the image contrast and to make the tumor boundary more distinct. A gradient magnitude filter (Deriche 1990) was then applied to generate a gradient image that showed the magnitude in the horizontal and vertical directions. Using the level-set function on the gradient image accomplished the segmentation procedure.

The level-set function $\psi(x, t)$, which is a high-dimension function, uses the initial contour $\gamma(t)$ as the zero level set $\Gamma(x, t) = \{\psi(x, t) = 0\}$ where x is a point in $\Re^N$. Based on the partial differential equation, the level-set function $\psi(x, t=0)$ evolved from the initial contour $\gamma(t=0)$ is defined as

$$\psi(x, t = 0) = \pm dis \tag{1}$$

The distance between the point x and $\gamma(t=0)$ is defined as dis with a sign to indicate the position. A positive or negative sign means that the point is outside or inside the initial contour, respectively. Then, the partial differential equation is defined by the given values of $\psi(x, t=0)$ and F, which provides the propagation speed from the initial contour to the outer region. The equation is

$$\psi + F\left|\nabla \psi\right| = 0 \qquad (2)$$

Fig. 2 (b) shows the tumor contours of Fig. 2(a) automatically delineated by the level-set segmentation.

Speckle detection

After acquiring the tumor area via tumor segmentation, the speckle pixels around the tumor area were detected to extract speckle features. The speckle patterns in the B-mode images were analyzed regarding their tissue characteristics for breast tumor classification (Moon et al. 2012a; Moon et al. 2012b). The inherent speckle pattern in US images is generated by microstructures that cause scattering, which are contained in tissues such as the breast parenchyma including ducts and glands. Scattered US pulses result in granular appearances in the interference pattern with small grayscale differences in the B-mode images. For fully developed speckle, the intensity image has an exponential distribution with a mean-to-standard deviation (SD) ratio of 1.0. To extract the speckle pixels, the 0-255 pixel values in the US images were log decompressed to the raw intensity value defined as

$$\mathrm{I}(\mathrm{x},\ \mathrm{y}) = 10^{B(x,y)/B_0} \qquad (3)$$

where $B(x,y)$ is the acquired B-mode grayscale value and B_0 is a linearization factor related to the transducer frequency (Berg et al. 2008) that converts $B(x,y)$ to a linear scale. A moving 5×5 window was then used to detect the speckle pixels with a mean-to-SD distribution of 0.8-1.2 as the tolerance range in the raw intensity image. The criterion is defined as following:

$$W_{mSD}(\mathrm{x},\mathrm{y}) = \frac{W_{mean}(x,y)}{\sqrt{\sum_{i=-a}^{a}\sum_{j=-a}^{a}\left(I(x+i,y+j)-W_{mean}(x,y)\right)^2}} \quad (4)$$

$$W_{mean}(x,y) = \sum_{i=-a}^{a}\sum_{j=-a}^{a} I(x+i,y+j) \Big/ (2a+1)^2 \quad (5)$$

where a=2 was used for the window size. Speckle pixels satisfying W_{mSD}=0.8-1.2 within a distance of 5 to the tumor boundary were gathered as the speckle map for texture analysis in the experiment. Fig. 2 (c) shows the speckle pixels detected around the tumors in Fig. 2 (a) for speckle texture calculation.

Texture features

Texture analysis has been widely used for pattern recognition in digital images (Gonzalez 2008). In US images, the texture information is based on the echo pattern presented by the grayscale echogenicity. Previous studies (Moon et al. 2012b; Lo et al. 2014) have suggested extracting texture features inside or surrounding the tumor area for tissue characterization. The quantitative texture features used in CAD systems can be classified into tumor textures and speckle textures. Both texture feature sets provide useful information in distinguishing between benign and malignant breast lesions. However, the gray-scale variations between US images from different system settings and scanners may affect the performances of texture features. In this study, we propose to extract texture features from the ranklet-transformed US images to obtain intensity-invariant tumor textures and speckle textures. The target areas for texture extraction are the previously defined tumor area and speckle area.

Ranklet transform

The ranklet transform was used in grayscale medical images for intensity-invariant texture extraction (Masotti and Campanini 2008; Yang et al. 2013). In US B-mode images, analyzing the grayscale tissue echogenicity in the tumor region and speckle pixels (Moon et al. 2012a; Moon et al. 2012b) has been demonstrated to be useful in breast tumor classification. In clinical use, US images were not always acquired using the same scanner settings and scanner models. Fig. 3 shows the effects of different grayscale distributions on the shape and texture features. Regardless of the case that was selected from ATL (Fig. 3(a)) or Medison (Fig. 3(b)), the segmentation results on the original, contrast enhanced, gamma corrected, and histogram equalized images (Gonzalez 2008) were similar. The SDs of four NRL entropy values were 0.03

and 0.02 for the ATL and Medison cases, respectively. However, the texture value of the cluster shade highly depended on the grayscale distribution results in the SD (36.33 and 45.46 for the ATL and Medison cases, respectively).

In this study, to reduce the effect of intensity variation, the ranklet transform (Masotti and Campanini 2008) was applied to the US images to achieve intensity-invariant texture features. Calculating the relative rank of the pixel values rather than original grayscale pixel values is the key methodology of the ranklet transform. The use of multi-resolution and orientation-selective transformations achieves the analysis of different scales and angles. Using a resolution value of R, a number of overlapping crops with $R \times R$ are extracted by shifting an $R \times R$ crop window. The resolutions of 2, 4, and 8 for an 8×8 image generate 49, 25, and 1 crops, respectively. For each resolution, the image is separated into two subsets, X and Y, according to the selected orientations. The divisions are based on the Haar functions used in the wavelet transform (Mallat 1989) and shown in Fig. 4 for the vertical, horizontal, and diagonal orientations. For each resolution and orientation, the number of pixel pairs (P_H, P_L) in each crop is calculated when the relative rank of pixels of P_H in the subset X is higher than that of P_L in the other subset Y. If there are C pixels in a crop, $C/2 \times C/2 = C^2/4$ comparisons are needed. The resulting number, which originally ranges from 0 to $C^2/4$, can be normalized to be between -1 and 1. The derived ranklet transform coefficient R_O is formulated as following:

$$R_O = \frac{\sum_{p \in Y_O} \pi(p) - C/4\,(C/2 + 1)}{C^2/8} - 1, O = V, H, D \quad (6)$$

In the subset Y_O, the pixel ranks $\pi(p)$ are summed. While more pixels in Y_O have higher grayscale values than those in X_O, the R_O is close to 1. Otherwise, it is close to -1. The ranklet coefficient for a crop is close to 0 if there is no global value variation. After performing the ranklet transform, the pixel values are replaced by the ranklet coefficient to express the regularity correlation in the area. Specifically, the ranklet images are texture patterns extracted from different scales and angles. Fig. 5 shows the ranklet images of the cases in Fig. 3. Using a resolution value of 4 as an example, the original US images with different grayscale distributions have very similar ranklet images regardless of the orientation (vertical, horizontal, or diagonal). The SDs of the cluster shades in $R4_D$ were 2.78 and 4.59 for the ATL (Fig. 5(a)) and Medison (Fig. 5(b)) cases, respectively. In the experiment, five image resolutions (2, 4, 8, 16, and 32) and three orientations (vertical, horizontal, and diagonal) were used. The minimum resolution was 2 to extract least local information. The maximum resolution was 32 because not all tumors had sizes more than 32 pixels. Each resolution scale was the double of the prior scale to generate new contrast information. For each tumor case, the texture calculation was performed 15 times in the ranklet images for the different combinations of resolution and orientation and normalized using the grayscale values of 0-255.

GLCM texture

The defined tumor and speckle areas are the clusters of similar biological structures. The texture information inside the tumor and speckle areas can be extracted by analyzing the correlations between pixel values. The gray-level co-occurrence matrix (GLCM) (Haralick et al. 1973) is the second-order statistic describing the joint frequencies of pair-wise combinations. Co-occurrence matrices ($P=[p(i,j|d,\theta)]$) are constructed by scanning each pixel and the adjacent pixels. The matrix element $P=[p(i,j|d,\theta)]$ describes the frequencies of two adjacent pixels at a distance of d and a direction θ, one with gray-scale i and the other with gray-scale j. In the experiment, the distance for the occurrence of two pixels at a distance d=1 and four offset directions, θ=0°, 45°, 90°, 135°, was used. For rotation invariance, the results from different directions were summed and averaged to be an element in the matrix (Haralick et al. 1973). The direction average and SD of eight GLCM metrics defined below were quantified as the texture features.

$$Energy = \sum_i \sum_j p(i,j|d,\theta)^2 \tag{7}$$

$$Entropy = -\sum_i \sum_j p(i,j|d,\theta)\log\left(p(i,j|d,\theta)\right) \quad (8)$$

$$Correlation = \frac{\sum_i \sum_j (i-\mu_x)(j-\mu_y)p(i,j|d,\theta)}{\sigma_x \sigma_y} \quad (9)$$

$$Local\ Homogeneity = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j|d,\theta) \quad (10)$$

$$Inertia = \sum_i \sum_j (i-j)^2 p(i,j|d,\theta) \quad (11)$$

$$Cluster\ Shade = \sum_i \sum_j \left(i+j-\mu_x-\mu_y\right)^3 p(i,j|d,\theta) \quad (12)$$

$$Cluster\ Prominence = \sum_i \sum_j \left(i+j-\mu_x-\mu_y\right)^4 p(i,j|d,\theta) \quad (13)$$

$$Haralick's\ Correlation = \frac{\sum_i \sum_j (i \cdot j)p(i,j|d,\theta) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (14)$$

where μ_x, μ_y, σ_x and σ_y are mean and SD of the marginal distributions of $p(i,j|d,\theta)$.

$$\mu_x = \sum_i i \sum_j p(i,j|d,\theta), \mu_y = \sum_j j \sum_i p(i,j|d,\theta) \quad (15)$$

$$\sigma_x^2 = \sum_i (i-\mu_x)^2 \sum_j p(i,j|d,\theta), \sigma_y^2 = \sum_j \left(j-\mu_y\right)^2 \sum_i p(i,j|d,\theta) \quad (16)$$

Fig. 2 (b) and (c) show the tumor and speckle areas from which the GLCM texture were extracted, respectively. The tumor texture and speckle texture were GLCM features extracted from five scales and three orientations of the tumor and speckle areas after ranklet transformation, respectively. Both feature sets had 240 texture features. To evaluate the effectiveness of ranklet transform, the performances of the two feature sets were compared to those from the original US images in the experiment.

Statistical analysis

The quantitative features were evaluated if they were significant in distinguishing between benign and malignant lesions. Kolmogorov-Smirnov test (Field 2009) was used to determine whether the value distribution of a feature is normal distribution or not. Features with normal distribution were evaluated using Student's t-test (Field 2009). Otherwise, the Mann-Whitney U test (Field 2009) was used. The resulting *p*-values less than 0.05 were interpreted as statistically significant. Significant features of tumor texture and speckle texture were combined to be feature sets, respectively. In the classifier of binary logistic regression (Hosmer 2000), stepwise backward elimination evaluated the features in a feature set. While the least error rate was obtained, the corresponding features were selected to be the most relevant. After feature selection, leave-one-out cross-validation was used to evaluate the generalization ability. Each time, one individual case was separated from the total cases and was used to test the result trained by the remaining cases.

After classification, five performance indices including accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated according to the biopsy-proven pathology. The performance differences between two feature sets were evaluated using chi-square test. The trade-offs between sensitivity and specificity were analyzed using ROC curve. Az, area under the ROC curve, was compared using the z-test in ROCKIT software (C. Metz, University of Chicago, Chicago, IL, USA). Other test methods were performed using SPSS software (version 16 for Windows; SPSS, Chicago, IL, USA).

Results

Using the Kolmogorov-Smirnov test, 186 tumor textures and 181 speckle textures were normally distributed. The number of significant features for the tumor texture and speckle texture were 46 and 35, respectively. Ten features of tumor texture with ranklet transformation were selected in the classifier. They were Cluster Prominence ave. ($R16_D$), Energy std. ($R32_H$), Entropy ave. ($R32_H$), Correlation std. ($R32_H$), Inverse Difference Moment ave. ($R32_H$), Inverse Difference Moment std. ($R32_H$), Cluster Prominence ave. ($R32_H$), Cluster Prominence std. ($R32_H$), Haralick Correlation ave. ($R32_H$), and Haralick Correlation std. ($R32_H$). The 11 selected speckle features with ranklet transform were Cluster Prominence std. ($R16_H$), Haralick Correlation std. ($R16_H$), Energy ave. ($R32_H$), Entropy ave. ($R32_H$), Correlation ave. ($R32_H$), Correlation std. ($R32_H$), Inverse Difference Moment ave. ($R32_H$), Inertia std. ($R32_H$), Cluster Shade std. ($R32_H$), Cluster Prominence std. ($R32_H$), and Haralick Correlation ave. ($R32_H$). Most selected features were from $R32_H$, the resolution of 32 with the horizontal orientation. The common features were Entropy ave. ($R32_H$), Correlation std. ($R32_H$), Cluster Prominence std. ($R32_H$), and Haralick Correlation ave. ($R32_H$).

The performances of the selected texture features with ranklet transformation were compared to those of the original texture features in Table 1. Regardless of the location the texture features were extracted from (tumor area or speckle pixels), the original intensity-based textures performed poorly in diagnosis. Only sixty percent of the cases were correctly classified by the texture features. In contrast, the texture features extracted from images after ranklet transformation performed better than those from the original B-mode images. For the tumor texture, the features with ranklet transformation achieved an accuracy of 80% (55/69) and a sensitivity of 76% (16/21), which are significantly better than an accuracy of 58% (40/69) and a sensitivity of 38% (8/21) for the original texture features. **The specificities were not significantly different**. The accuracy and sensitivity of the speckle texture with ranklet transformation are also significantly better than the original speckle texture (83% (57/69) vs. 62% (43/69) and 71% (15/21) vs. 33% (7/21), respectively). Similarly, the differences **in** specificity and NPV were not significant. With respect to the trade-offs between sensitivity and specificity, the Az differences between the texture sets with and without ranklet transformation were significant (p-value=0.0009 and 0.02, respectively). The corresponding ROC curves are shown in Fig. 6.

Assessment of combined performance showed that the use of two feature sets together resulted in essentially equal performance.

Discussion

Various CAD systems have been developed to distinguish between malignant and benign tumors using US images (Drukker et al. 2005; Kim et al. 2005; Moon et al. 2013b). Acting as a second viewer, a CAD system that can confirm the classification of malignant tumors initially classified as BI-RADS category 3 by radiologists is important. To provide more objective suggestions, a CAD system reviewed the quantitative characteristics of BI-RADS category 3 masses to avoid misclassifying carcinomas in a previous study (Moon et al. 2013a). The elliptical-shaped features performed well in the CAD system to confirm malignant tumors. However, the performance of the texture features was not as good as that of the shape features. One possible reason for this result is that the texture features are easily affected by the grayscale distribution.

In this study, we proposed using ranklet transformation (Masotti and Campanini 2008) to extract robust texture features independent of the intensity variation. Using ranklet transformations with multiple resolutions and orientations, the original B-mode US images were converted to ranklet images that represented the regularity correlation in the area. As shown in Fig. 3 and 5, the SDs of the cluster shades in the four different US images were reduced from 36.33 and 45.46 to 2.78 and 4.59 for the ATL and Medison cases, respectively. This result indicates that extracting texture features from ranklet images is intensity-invariant. Based on the effectiveness of this technique, the performances of tumor textures and speckle textures extracted from ranklet images were analyzed to confirm the classification of malignant tumors initially classified as BI-RADS category 3. The diagnostic result showed that texture features with ranklet transformation are significantly better than those of features without it. The Az values improved from 0.56-0.58 to 0.80-0.83. The texture features extracted from the ranklet images were demonstrated to be intensity-invariant and to provide diagnostic information in classifying tumors examined using different US equipment (ATL and Medison). Compared with other existing CAD systems (Kim et al. 2005; Moon et al. 2012b), our database composed of different scanner settings and models is more reliable. Additionally, the effectiveness of extracting intensity-invariant speckle texture after ranklet transform indicates that the method can be applied to discrete pixels. Detecting and analyzing speckle features for tumor diagnosis can be performed in an ROI (Moon et al. 2012a; Moon et al. 2012b) which is simpler than tumor segmentation. More time can be saved by using speckle texture than tumor texture.

The success is in agreement with a previous study using ranklet transformation in breast tumor classification (Yang et al. 2013). The Az of the tumor texture was 0.90-0.94, which is higher than the value of 0.83 obtained in this study. Nevertheless, the BI-RADS category 3 breast masses used in this study had less than 2% malignancy as assessed by radiologists and were more challenging to assess using CAD systems. Additionally, the Az improvement from 0.58 to 0.83 in this study is higher than that from 0.81 to 0.90 in the literature. The superior Az improvement achieved here demonstrates that the ranklet transformation is a promising tool in developing superior CAD systems.

According to the selected features in the classifier, most selected features were from $R32_H$, the resolution of 32 with the horizontal orientation. The result indicates that these features had more relevant diagnostic information than others. A possible reason is that the contrast difference in large-scale resolution is clearer than small-scale in the ranklet transform and provides more contrast information for texture analysis. **This may also be explained by the typical shape of breast tumors, which tend to be ellipsoid with their longest axis horizontal.**

A limitation of this study is that the number of specimens in the experiment is only 69. More BI-RADS category 3 breast masses collected from different scanner models using different settings should be included in future studies to evaluate the ability to generalize the proposed CAD system. **Another limitation is the significant difference in the ages between the benign and malignant tumors. Age can be a feature to estimate the likelihood of tumors being carcinomas. To investigate the usefulness of the proposed texture features, the comparisons between benign and malignant cases should be performed for similar age cases to reduce the effect of age. Nevertheless, the proposed texture features also describe the composition of echo patterns in tumors such as homogeneity and heterogeneity which may not completely correlate with age. Collecting the cases with similar ages to explore the correlation between textures and ages would be an interesting future study. However, acquiring hundreds of benign and malignant cases with same ages would need many years.**

To improve the CAD performance, a novel method of feature selection may be helpful. The ranklet transform used multiple resolutions and orientations to generate hundreds of texture features. Determining the most relevant subset of these features is an interesting topic. Using the least features to achieve the best performance would improve the efficiency. Combining relevant features with complementary diagnostic information is the key aspect of this methodology.

Acknowledgments

The authors thank the Ministry of Science and Technology (MOST 103-2221-E-002-170-MY3), Ministry of Economic Affairs (102-EC-17-A-19-S1-164), and Ministry of Education (AE-00-00-06) of the Republic of China for the financial support. This study was also supported by the grant no. 04-2011-0380 from the SNUH Research Fund.

References

Berg WA. Supplemental screening sonography in dense breasts. Radiol Clin North Am 2004;42:845-+.

Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, Pisano ED, Jong RA, Evans WP, Morton MJ, Mahoney MC, Larsen LH, Barr RG, Farria DM, Marques HS, Boparai K, Investigators A. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. Jama-J Am Med Assoc 2008;299:2151-63.

Chang RF, Wu WJ, Moon WK, Chen DR. Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. Breast Cancer Res Treat 2005;89:179-85.

Deriche R. Fast Algorithms for Low-Level Vision. IEEE Trans Pattern Anal Mach Intell 1990;12:78-87.

Drukker K, Giger ML, Metz CE. Robustness of computerized lesion detection and classification scheme across different breast US platforms. Radiology 2005;237:834-40.

Field AP. Discovering statistics using SPSS, 3rd ed. Los Angeles: SAGE Publications, 2009.

Gonzalez RC. Digital image processing, third ed. Upper Saddle River, New Jersey:

Pearson Prentice Hall, 2008.

Graf O, Helbich TH, Hopf G, Graf C, Sickles EA. Probably benign breast masses at US: Is follow-up an acceptable alternative to biopsy? Radiology 2007;244:87-93.

Gruber R, Jaromi S, Rudas M, Pfarl G, Riedl C, Flöry D, Graf O, Sickles E, Helbich T. Histologic work-up of non-palpable breast lesions classified as probably benign at initial mammography and/or ultrasound (BI-RADS category 3). Eur J Radiol 2013;82:398-403.

Haralick RM, Shanmuga.K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern 1973;Smc3:610-21.

Hosmer DW. Applied logistic regression. 2nd edition. New York: Wiley, 2000.

Kashikura Y, Nakayama R, Hizukuri A, Noro A, Nohara Y, Nakamura T, Ito M, Kimura H, Yamashita M, Hanamura N, Ogawa T. Improved Differential Diagnosis of Breast Masses on Ultrasonographic Images with a Computer-Aided Diagnosis Scheme for Determining Histological Classifications. Acad Radiol 2013;20:471-7.

Kelly KM, Dean J, Comulada WS, Lee S-J. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol 2010;20:734-42.

Kim KG, Cho SW, Min SJ, Kim JH, Min BG, Bae KT. Computerized scheme for assessing ultrasonographic features of breast masses. Acad Radiol 2005;12:58-66.

Kim SJ, Chang JM, Cho N, Chung S-Y, Han W, Moon WK. Outcome of breast lesions detected at screening ultrasonography. Eur J Radiol 2012;81:3229-33.

Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: Interobserver variability and positive predictive value. Radiology 2006;239:385-91.

Lee HJ, Kim EK, Kim MJ, Youk JH, Lee JY, Kang DR, Oh KK. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. Eur J Radiol 2008;65:293-8.

Leung JWT, Sickles EA. The probably benign assessment. Radiol Clin North Am 2007;45:773-+.

Lo C-M, Chen R-T, Chang Y-C, Yang Y-W, Hung M-J, Huang C-S, Chang R-F. Multi-dimensional Tumor Detection in Automated Whole Breast Ultrasound using Topographic Watershed. IEEE Trans Med Imaging 2014;33:1503-11.

Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 1989;11:674-93.

Masotti M, Campanini R. Texture classification using invariant ranklet features. Pattern Recognit Lett 2008;29:1980-6.

Masotti M, Lanconelli N, Campanini R. Computer-aided mass detection in mammography: False positive reduction via gray-scale invariant ranklet texture features. Med Phys 2009;36:311-6.

Mendelson E, Bohm-Velez M, Berg W. ACR BI-RADS Ultrasound. ACR BI-RADS atlas, Breast Imaging Reporting and Data System Reston, VA: American College of Radiology 2013;

Moon WK, Lo C-M, Chang JM, Huang C-S, Chen J-H, Chang R-F. Quantitative Ultrasound Analysis for Classification of BI-RADS Category 3 Breast Masses. J Digit Imaging 2013a;1-8.

Moon WK, Lo C-M, Cho N, Chang JM, Huang C-S, Chen J-H, Chang R-F. Computer-aided diagnosis of breast masses using quantified BI-RADS findings. Comput Methods Programs Biomed 2013b;111:84-92.

Moon WK, Lo CM, Chang JM, Huang CS, Chen JH, Chang RF. Computer-aided classification of breast masses using speckle features of automated breast ultrasound images. Med Phys 2012a;39:6465-73.

Moon WK, Lo CM, Huang CS, Chen JH, Chang RF. Computer-Aided Diagnosis Based on Speckle Patterns in Ultrasound Images. Ultrasound Med Biol

2012b;38:1251-61.

Sickles EA. Nonpalpable, Circumscribed, Noncalcified Solid Breast Masses - Likelihood of Malignancy Based on Lesion Size and Age of Patient. Radiology 1994;192:439-42.

Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions. Radiology 1995;196:123-34.

Suri JS. Advances in diagnostic and therapeutic ultrasound imaging. Boston, London: Artech House, 2008.

Tsui PH, Yeh CK, Liao YY, Chang CC, Kuo WH, Chang KJ, Chen CN. Ultrasonic Nakagami Imaging: A Strategy to Visualize the Scatterer Properties of Benign and Malignant Breast Tumors. Ultrasound Med Biol 2010;36:209-17.

Weigel S, Biesheuvel C, Berkemeyer S, Kugel H, Heindel W. Digital mammography screening: how many breast cancers are additionally detected by bilateral ultrasound examination during assessment? Eur Radiol 2013;23:684-91.

Yang MC, Moon WK, Wang YCF, Bae MS, Huang CS, Chen JH, Chang RF. Robust Texture Analysis Using Multi-Resolution Gray-Scale Invariant Features for Breast Sonographic Tumor Diagnosis. IEEE Trans Med Imaging 2013;32:2262-73.

Figure Captions

Fig. 1. The flowchart of the proposed CAD system

Fig. 2. Examples of acquired US images and the corresponding contours and speckle pixels. (a) A fibroadenoma, an invasive ductal carcinoma from the ATL scanner, a fibrocystic change, and an invasive ductal carcinoma from the Medison scanner (b) The corresponding segmentation results (c) The corresponding detected speckle pixels.

Fig. 3. Quantification results of the US images obtained from the ATL and Medison scanners. (a) The segmentation results and feature values of different grayscale distributions of an US image from ATL. (b) The segmentation results and feature values of different grayscale distributions of a US image from Medison.

Fig. 4. An image is separated into two subsets according to the vertical, horizontal, and diagonal orientations for ranklet transformation.

Fig. 5. Ranklet images and feature values of the US images obtained from the ATL and Medison scanners. (a) The ranklet images (resolution=4) with three orientations and feature values in different grayscale distributions of a US image from ATL. (b) The ranklet images (resolution=4) with three orientations and feature values in different grayscale distributions of a US image from

Medison.

Fig. 6. The receiver operating characteristic (ROC) curves of the original tumor texture, tumor texture with ranklet transformation, original speckle texture, and speckle texture with ranklet transformation.

Table 1. The performance differences of tumor texture and speckle texture with and without ranklet transformation using the chi-square test.

	Tumor Texture	Tumor Texture (Ranklet)	Tumor Texture vs. Tumor Texture (Ranklet)	Speckle Texture	Speckle Texture (Ranklet)	Speckle Texture vs. Speckle Texture (Ranklet)
Accuracy	58% (40/69)	80% (55/69)	0.0058*	62% (43/69)	83% (57/69)	0.0076*
Sensitivity	38% (8/21)	76% (16/21)	0.0126*	33% (7/21)	71% (15/21)	0.0134*
Specificity	67% (32/48)	81% (39/48)	0.1035	75% (36/48)	88% (42/48)	0.1167
PPV	33% (8/24)	64% (16/25)	0.0318*	37% (7/19)	71% (15/21)	0.0281*
NPV	71% (32/45)	89% (39/44)	0.0396*	72% (36/50)	88% (42/48)	0.0570
Az	0.58	0.83	0.0009*	0.56	0.80	0.02*

* p-value<0.05 indicates a statistically significant difference.