Classification of Benign and Malignant Thyroid Nodules Using Wavelet Texture Analysis of Sonograms

Ali Abbasian Ardakani, MSc, Akbar Gharbali, PhD, Afshin Mohammadi, MD

Objectives—The purpose of this study was to evaluate a computer-aided diagnostic system using texture analysis to improve radiologic accuracy for identification of thyroid nodules as malignant or benign.

Methods—The database comprised 26 benign and 34 malignant thyroid nodules. Wavelet transform was applied to extract texture feature parameters as descriptors for each selected region of interest in 3 normalization schemes (default, $\mu \pm 3\sigma$, and 1%–9%). Linear discriminant analysis and nonlinear discriminant analysis were used for texture analysis of the thyroid nodules. The first–nearest neighbor classifier was applied to features resulting from linear discriminant analysis. Nonlinear discriminant analysis features were classified by using an artificial neural network. Receiver operating characteristic curve analysis was used to examine the performance of the texture analysis methods.

Results—Wavelet features under default normalization schemes from nonlinear discriminant analysis indicated the best performance for classification of benign and malignant thyroid nodules and showed 100% sensitivity, specificity, and accuracy; the area under the receiver operating characteristic curve was 1.

Conclusions—Wavelet features have a high potential for effective differentiation of benign from malignant thyroid nodules on sonography.

Key Words—computer-aided diagnosis; head and neck ultrasound; sonography; texture analysis; thyroid nodules; wavelet

Received September 24, 2014, from the Student Research Committee, Urmia University of Medical Sciences, Urmia, Iran (A.A.A.); and Departments of Medical Physics (A.G.) and Radiology (A.M.), Faculty of Medicine, Imam Khomeini Hospital, Urmia University of Medical Sciences, Urmia, Iran. Revision requested December 12, 2014. Revised manuscript accepted for publication February 10, 2015.

Address correspondence to Akbar Gharbali, PhD, Department of Medical Physics, Faculty of Medicine, Urmia University of Medical Sciences, Urmia 1138, Iran.

E-mail: gharbali@yahoo.com

Abbreviations

A_z, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV positive predictive value; ROI, region of interest

doi:10.7863/ultra.14.09057

A thyroid nodule is an abnormal cell growth in the thyroid gland and appears as a palpable or nonpalpable mass. Thyroid nodules are more common in women than in men. The National Cancer Institute estimated that 62,980 new cases of thyroid cancer and 1890 thyroid cancer–related deaths would occur in 2014.[1] Studies indicate a 3% increase in the annual incidence of thyroid cancer.[2,3] Although fine-needle aspiration biopsy is the standard method for evaluation of thyroid nodules, it is painful, incurs health care costs, and contains the risk of infection and bruising. Approximately 10% to 20% of fine-needle aspiration biopsies are nondiagnostic, in which the thyroid nodule shows a high probability of malignancy, and aspiration should be repeated.[4]

The structure of the thyroid tissue changes when nodules develop, and the attributes of the nodule differ from those of surrounding thyroid tissue. Sonography can show this change by using echogenicity features. At present, sonography is an accepted imaging

modality for diagnosing and classifying benign and malignant nodules because it is inexpensive, easy to use, noninvasive, and allows real-time scanning.[5,6]

In the last decade, computer-aided diagnosis has been used to classify thyroid nodules using sonographic features. Sonograms contain textural, morphologic, and elastographic features. Ultrasound elastography is a noninvasive technique for obtaining information about tissue stiffness. Elastography represents tissue deformation in response to compression; this technique requires 2 sonograms.[7] Many studies have demonstrated that elastography has great potential for classification of benign and malignant thyroid nodules, especially for nodules with indeterminate cytologic characteristics.[8–10] Hong et al[9] used elastography to distinguish benign and malignant thyroid nodules. They achieved a best performance with sensitivity and specificity of 88% and 90%, respectively. Zhang et al[11] found that virtual touch tissue elasticity imaging using acoustic radiation force impulse elastography was useful for differentiating between benign and malignant thyroid nodules. They were able to classify thyroid nodules with sensitivity of 87.0%, specificity of 95.8%, and accuracy of 92.7%. The morphologic features describe symmetry, margin irregularity, and the shape of the nodule surface.[12] In this regard, Tsantis et al[13] designed support vector machines to discriminate between low- and high-risk thyroid nodules. They achieved the best results using combined "smoothness–symmetry–standard deviation of local maxima" features with 93% sensitivity, 98% specificity, and an area under the receiver operating characteristic curve (A_z) of 0.96.

Generally, texture shows different gray-level values, brightness, coarseness, and color, among other features, in an image. Tissues from benign and malignant thyroid nodules have different texture features. At times, patterns within an image may be different but are perceived by the human eye as having the same texture. Since radiologists usually assess texture qualitatively, computerized texture analysis can increase the accuracy of assessment. This process is a mathematical technique that provides increased quantification and information on spatial gray-level intensities in the pixels within an image.[14–16] In this regard, Ding et al[17] reported that sensitivity, specificity, and accuracy were 94.6%, 92.3%, and 93.6%, respectively, when the energy features of a co-occurrence matrix were used. Acharya et al[18] combined wavelet texture analysis and a gray-level co-occurrence matrix from 3-dimensional high-resolution sonograms and achieved excellent performance, with accuracy, sensitivity, and specificity of 100% and an A_z of 1.

This study proposed a new computer-aided diagnostic system based on wavelet transform to differentiate benign from malignant thyroid nodules. No previous published works were found that used texture analysis with wavelet features for sonography.

Materials and Methods

The data used in this study consisted of 60 fine-needle aspiration biopsy–proven lesions (26 benign and 34 malignant). Sonograms were obtained with a high-resolution Accuvix V20 sonography system (Medison, Seoul, Korea) equipped with an L5-13IS (5–13-MHz) linear array transducer. Sonograms were inputted into MaZda version 4.6 software (Institute of Electronics, Technical University of Lodz, Lodz, Poland) for texture analysis. MaZda software was developed in 1998 for texture analysis of magnetic resonance imaging,[15] and it was used here for sonography. The texture analysis method used in this study was based on wavelet transform, which is often called wavelet texture analysis. In transform texture analysis methods, texture features are analyzed in a new space on an image, with coordinates associated with characteristics such as frequency and size.[19] Fourier, Gabor, and wavelet transformation are 3 popular methods of texture analysis. Fourier transform is unsuitable for most medical applications because of its inability to analyze nonstationary signals, and it represents the global frequency content of an image with no time localization. Gabor transform provides better time localization, but its practice is limited because natural textures have multiple resolutions.[20,21] Wavelet transform is a popular method of texture analysis because it can analyze natural nonstationary signals and be localized in both spatial and frequency domains. Wavelets decompose the image signals into frequency components using independent spatially oriented frequency filters cascaded in a pyramidal structure. Wavelet texture analysis comprises sets of high-pass and low-pass filters or the Haar function to increase frequency resolution and produce a set of wavelet coefficients corresponding to the different scales and directions.[14,20,22]

In this study, discrete wavelet transform was used. Discrete wavelet transform is a linear transformation using a cascade of filters, followed by a 2-factor subsampling lattice. It was initially performed for all rows and then for all columns of an image (Figure 1).[23]

Six levels of discrete wavelet transform features were measured, and the wavelet coefficients were computed as

$$a^{j+1}[p] = \sum_{n=-\infty}^{+\infty} l[n-2p]a^{j}[n];$$

$$d^{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n-2p]a^{j}[n],$$

where d_j is the wavelet coefficient; a_j is the next level of transform; and h and l denote coefficients for the high-pass and low-pass filters, respectively. The Harr wavelet was calculated for the whole image to compute the wavelet features. The image was decomposed into 4 sub-band images at each level. Figure 2 shows the discrete wavelet transform tree structure for the 6 decomposition levels with the 2-band signal in the vertical (columns) and horizontal (rows) directions.

The wavelet coefficients were calculated at each sub-band level for all image rows and columns. The energy of the Haar wavelet sub-bands was calculated. The energy of the wavelet coefficient (a_{LL}, d_{HH}, d_{HL}, and d_{LH}) was computed at each level and sub-band in each region of interest (ROI) as

$$E_{\text{sub-band, level}} = \frac{\sum_{x,y \in \text{ROI}} (d_{\text{sub-band}; x,y})^2}{n},$$

where n is the number of pixels in the ROI at a given level and sub-band.

There are 6 levels of wavelet decomposition, and each level has 4 sub-band images; 24 energies of wavelet coefficients were derived from the 24 wavelet coefficients for wavelet texture analysis (Figure 3). These 24 wavelet texture features were calculated for each ROI in 3 normalizations for wavelet texture analysis: N_1 default, using default images with the same appearance at an intensity range of 1 to 2^k, where k is the number of bits per pixel; N_2: 3σ, in which the image intensities were located inside a normalization range (μ – 3σ and μ + 3σ), where μ and σ, respectively, are the mean value and standard deviation of the gray-level intensity inside the ROI; and N_3: 1%–99%, in which the ROI gray-level ranged between the darkness level (accumulated histogram = 1% of the total) and the brightness level (accumulated histogram = 99% of the total). The intensity levels outside the normalization range were not considered for ROI analysis.

B11 version 3.4, which is a companion software program for MaZda (Institute of Electronics, Technical University of Lodz), was used for evaluation of texture features and classification. The B11 program implements linear discriminant analysis and nonlinear discriminant analysis to transform raw data to lower-dimensional spaces and increase discriminative power.[24,25] The first–nearest neighbor classifier was used for features resulting from linear discriminant analysis, and an artificial neural network classifier was used for features resulting from nonlinear discriminant analysis.[26] To compare the accuracy of discrimination of the linear discriminant analysis and nonlinear discriminant

Figure 1. Filter bank for 2-dimensional discrete wavelet transform at the (*j* + *1*)th decomposition level, consisting of 2-band signals using high- and low-pass wavelet filters in 2 row and column directions.

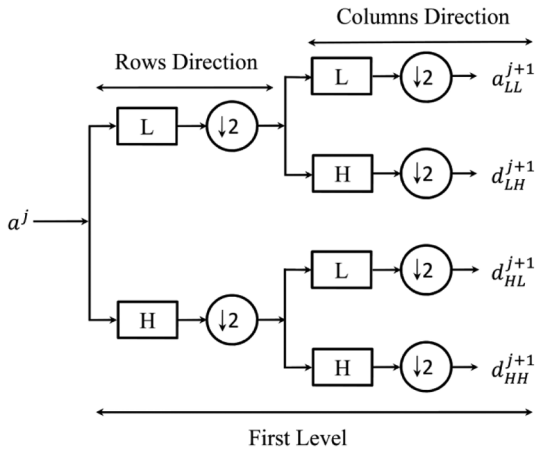
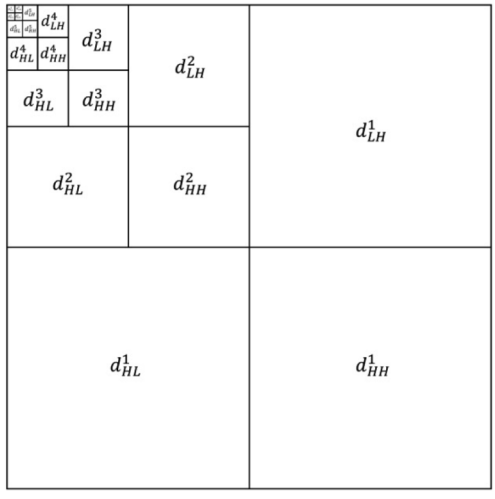
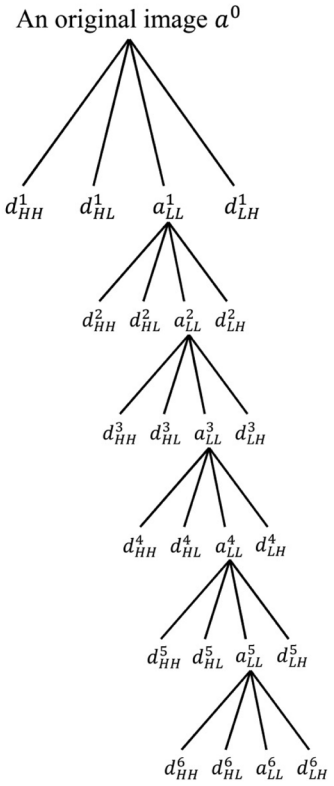


Figure 2. Discrete wavelet transform tree structure for the sixth decomposition level. For each level, the Haar wavelet was performed, and the image was decomposed to 4 sub-bands.



analysis methods, receiver operating characteristic analyses were applied.[27] Receiver operating characteristic analysis is a statistical technique that assesses discrimination performance for any 2-group classification task in terms of sensitivity, specificity, overall accuracy, positive predictive value (PPV), negative predictive value (NPV), and A_z.

Figure 3. Total decomposition tree of images using discrete wavelet transform for wavelet texture analysis. Twenty-four wavelet coefficients comprise a^n_{LL}, d^n_{HH}, d^n_{LH}, and d^n_{HL} for n = 1, 2, 3, 4, 5, and 6.



Results

Six options for texture analysis were used: 3 normalization schemes for 2 texture analysis methods. A total of 60 fine-needle aspiration biopsy–proven samples with 26 benign and 34 malignant thyroid nodules were selected.

The diagnostic performance of the texture analysis methods and normalization schemes are shown in Table 1. Considering the normalization role, default wavelet features represented higher performance, with sensitivity, specificity, accuracy, PPV, and NPV of 100% by nonlinear discriminant analysis in comparison with linear discriminant analysis. The classification task performed by linear discriminant analysis yielded lower discrimination performance, with sensitivity, specificity, accuracy, PPV, and NPV of 100%, 96.15%, 98.33%, 97.14%, and 100%, respectively.

In 3σ normalization, the nonlinear discriminant analysis showed better discrimination power, with sensitivity, specificity, accuracy, PPV, and NPV of 94.12%, 100%, 96.67%, 100%, and 92.86%, respectively. The features analyzed by linear discriminant analysis showed lower discrimination power than nonlinear discriminant analysis, with sensitivity, specificity, accuracy, PPV, and NPV of 94.11%, 88.46%, 91.66%, 91.43%, and 92.00%, respectively.

In 1%–99% normalization, features analyzed by nonlinear discriminant analysis had higher discrimination power than linear discriminant analysis to distinguish between benign and malignant thyroid nodules, with sensitivity, specificity, accuracy, PPV, and NPV of 91.18%, 100%, 95.00%, 100%, and 89.65%, respectively. Linear discriminant analysis had worse performance than nonlinear discriminant analysis, with sensitivity, specificity, accuracy, PPV, and NPV of 88.23%, 84.61%, 86.67%, 88.23%, and 84.61%, respectively.

The best results for linear discriminant analysis were for default normalization, with sensitivity, specificity, accuracy, PPV, and NPV of 100%, 96.15%, 98.33%, 97.14%, and 100%, respectively. Table 1 shows that the default dis-

Table 1. Summary of Performance for Wavelet Texture Analysis by Different Methods and Normalizations

Normalization	Texture Analysis Method	SEN, %	SPC, %	ACC, %	PPV, %	NPV, %	A_z
Default	LDA	100	96.15	98.33	97.14	100	0.98
	NDA	100	100	100	100	100	1
3σ	LDA	94.11	88.46	91.66	91.43	92.00	0.9128
	NDA	94.12	100	96.67	100	92.86	0.9706
1%–99%	LDA	88.23	84.61	86.67	88.23	84.61	0.8642
	NDA	91.18	100	95.00	100	89.65	0.9559

ACC indicates accuracy; LDA, linear discriminant analysis; NDA, nonlinear discriminant analysis' SPC; specificity; and SEN, sensitivity.

criminative power of nonlinear discriminant analysis was higher than the 2 other normalization schemes, with sensitivity, specificity, accuracy, PPV, and NPV of 100%.

The discrimination distributions for linear discriminant analysis and nonlinear discriminant analysis are illustrated in Figure 4 and show that nonlinear discriminant analysis had the greatest discriminative power to distinguish between benign and malignant thyroid nodules. Figure 5 shows receiver operating characteristic curves of the proposed computer-aided diagnostic system plotted on the same graph for each normalization to compare the discriminative power for classification of benign and malignant thyroid nodules. In general, the nonlinear discriminant analysis method had an advantage over linear discriminant analysis in each normalization with respect to a greater A_z value. As shown, linear discriminant analysis had the best performance in default normalization, with an A_z of 0.98. Nonlinear discriminant analysis also showed the best accuracy in default normalization, with an A_z of 1 (Figure 5a).

Discussion

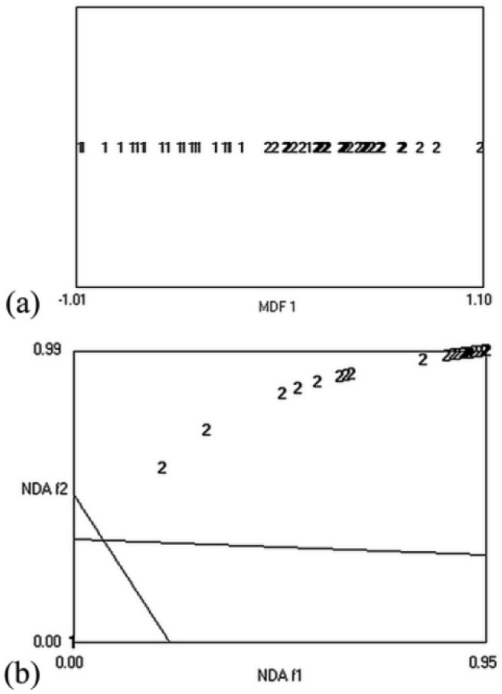
The results of this study have shown that wavelet texture analysis can differentiate benign from malignant thyroid nodules with high accuracy. Figure 5 reveals that nonlinear discriminant analysis had the best performance in all normalization schemes according to A_z. It also shows that in 3σ, nonlinear discriminant analysis was more accurate than linear discriminant analysis.

Table 1 indicates that normalization had a disruptive effect on performance of the classifier. The best performance was obtained for default normalization with nonlinear discriminant analysis.

Many studies have been conducted to distinguish between benign and malignant thyroid nodules using texture analysis. In this regard, Ding et al[17] computed first-order statistical features to describe the histographic elasticity properties and computed a co-occurrence matrix to describe the spatial distribution of elasticity. The features based on the co-occurrence matrix were extracted along 4 directions (0°, 45°, 90°, and 135°) and with 4 distances (1, 2, 3, and 4). The best result, with sensitivity of 94.6%, specificity of 92.8%, and accuracy 93.6%, was reached when 2 energy features of the co-occurrence matrix with a distance of 3 were used. Gopinath and Shanthi[28] used texture features derived from the Gabor filter bank at various wavelengths and angles for classification of benign and malignant thyroid nodules. They used a support vector machine classifier and achieved sensitivity of 95%, specificity of 100%, and accuracy of 96.7%. Gopinath and Shanthi et al[29] reported that sensitivity, specificity, and accuracy were 95%, 100%, and 96.66%, respectively, when 2-level wavelet decomposition based on the texture characteristics of thyroid cells was used to differentiate between benign and malignant thyroid nodules. Acharya et al[18] extracted texture features from 3-dimensional high-resolution sonograms. They combined a co-occurrence matrix (contrast, entropy, and homogeneity) and wavelet texture features and achieved excellent performance, with accuracy, sensitivity, and specificity of 100% and an A_z of 1. The proposed method showed more effective performance than the results of previous studies,[9,11,13,17,18,28,29] with sensitivity ranging from 87% to 95%, specificity ranging from 90% to 100%, and accuracy ranging from 92% to 96.7%.

The results of the proposed system and those from Acharya et al[18] were similar in all but 2 aspects. This study used only wavelet texture features to classify benign and malignant thyroid nodules, whereas Acharya et al combined wavelet and gray-level co-occurrence matrix–based texture features. Also, Acharya et al used 3-dimensional high-resolution sonograms for evaluation, and this study used 2-

Figure 4. Sample distributions after 2 texture analysis methods: linear discriminant analysis (**a**) and nonlinear discriminant analysis (NDA; **b**). Areas on the nonlinear discriminant analysis (f1 and f2) plane that form a common part of sets belong to 2 different categories (benign and malignant nodules). MDF indicates most discriminating features.

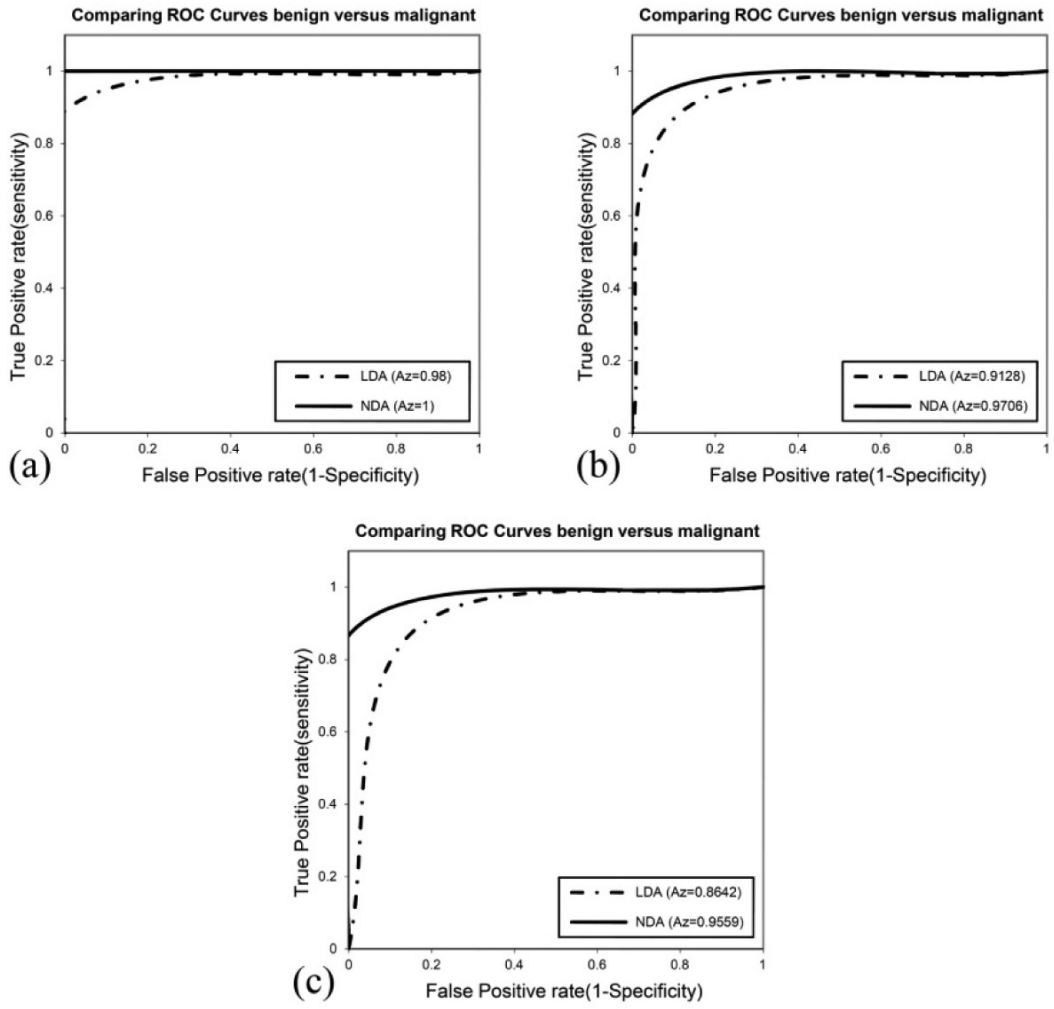


dimensional sonograms, which are conventionally used in clinical centers. Our research indicates that wavelet features had higher accuracy than the previous study, which used a combination of texture feature groups. Our results also indicate that wavelet texture analysis has substantially more discriminative ability than other methods using morphology and elastography.[9,11,13]

This method will not be an alternative to biopsy, but it can help radiologists identify individual nodules that are suspicious for malignancy and should undergo fine-needle aspiration biopsy in patients with multinodular goiters. One of the clinical challenges is determining which nodules in patients with multinodular goiters should undergo fine-needle aspiration biopsy, and our method can aid radiologists in targeting nodules among multiple nodules in these patients.

There were several limitations in our study. First, the data group was small; further investigation using a larger data set is needed. Second, our sonographic classification was compared with fine-needle aspiration biopsy results. Although fine-needle aspiration biopsy is highly sensitive for diagnosis of neoplastic thyroid nodules, definitive results require surgical pathologic analysis. Third, some data in this study were excluded because of indeterminate fine-needle aspiration biopsy diagnoses. Fourth, although MaZda software is useful for texture analysis, the ROI is selected manually. Through further software development, MaZda can be equipped with automated or semiautomated ROI selection.

Figure 5. Receiver operating characteristic curves for wavelet texture analysis in 3 normalization schemes: default (**a**), 3σ normalization (**b**), and 1%–99% normalization (**c**). LDA indicates linear discriminant analysis; and NDA, nonlinear discriminant analysis.



In conclusion, a new method based on wavelet transform is proposed to differentiate between benign and malignant thyroid nodules using 2-dimensional sonography. The results indicate that wavelet texture analysis is a useful technique for discrimination of thyroid nodules by sonography and can be an auxiliary tool to improve radiologists' diagnostic accuracy.

References

1. National Cancer Institute. Thyroid cancer—for patients. National Cancer Institute website. http://www.cancer.gov/cancertopics/types/thyroid. Accessed September 14, 2015.
2. US Cancer Statistics Working Group. *United States Cancer Statistics: 2000 Incidence*. Atlanta, Georgia: Centers for Disease Control and Prevention; 2003.
3. National Cancer Institute. SEER stat fact sheets: thyroid cancer. National Cancer Institute website. http://seer.cancer.gov/statfacts/html/thyro.html. Accessed September 14, 2015.
4. Chow LS, Gharib H, Goellner JR, van Heerden JA. Nondiagnostic thyroid fine-needle aspiration cytology: management dilemmas. *Thyroid* 2001; 11:1147–1151.
5. Koike E, Noguchi S, Yamashita H, et al. Ultrasonographic characteristics of thyroid nodules: prediction of malignancy. *Arch Surg* 2001; 136:334–337.
6. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *J Clin Endocrinol Metab* 2002; 87:1941–1946.
7. Ophir J, Cespedes I, Ponnekanti H, Yazdi Y, Li X. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrason Imaging* 1991; 13:111–134.
8. Dighe M, Bae U, Richardson ML, Dubinsky TJ, Minoshima S, Kim Y. Differential diagnosis of thyroid nodules with US elastography using carotid artery pulsation. *Radiology* 2008; 248:662–669.
9. Hong Y, Liu X, Li Z, Zhang X, Chen M, Luo Z. Real-time ultrasound elastography in the differential diagnosis of benign and malignant thyroid nodules. *J Ultrasound Med* 2009; 28:861–867.
10. Rago T, Santini F, Scutari M, Pinchera A, Vitti P. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. *J Clin Endocrinol Metab* 2007; 92:2917–2922.
11. Zhang YF, Xu HX, He Y, et al. Virtual touch tissue quantification of acoustic radiation force impulse: a new ultrasound elastic imaging in the diagnosis of thyroid nodules. *PLoS One* 2012; 7:e49094.
12. Thiran JP, Macq B. Morphological feature extraction for the classification of digital images of cancerous tissues. *IEEE Trans Biomed Eng* 1996; 43:1011–1020.
13. Tsantis S, Dimitropoulos N, Cavouras D, Nikiforidis G. Morphological and wavelet features towards sonographic thyroid nodules evaluation. *Comput Med Imaging Graph* 2009; 33:91–99.
14. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. *Clin Radiol* 2004; 59:1061–1069.
15. Materka A, Strzelecki M. *Texture Analysis Methods: A Review*. COST B11 Report. Lodz, Poland: Institute of Electronics, Technical University of Lodz; 1998.
16. Szczypiński PM, Strzelecki M, Materka A, Klepaczko A. MaZda: a software package for image texture analysis. *Comput Methods Programs Biomed* 2009; 94:66–76.
17. Ding J, Cheng H, Ning C, Huang J, Zhang Y. Quantitative measurement for thyroid cancer characterization based on elastography. *J Ultrasound Med* 2011; 30:1259–1266.
18. Acharya UR, Faust O, Sree SV, Molinari F, Suri JS. ThyroScreen system: high resolution ultrasound thyroid image characterization into benign and malignant classes using novel combination of texture and discrete wavelet transform. *Comput Methods Programs Biomed* 2012; 107:233–241.
19. Cohen L. Time-frequency distributions: a review. *IEEE Ultrason Proc* 1989; 77:941–981.
20. Livens S, Scheunders P, Van de Wouwer G, Van Dyck D. Wavelets for texture analysis: an overview. Paper presented at: Sixth International Conference on Image Processing and Its Applications; July 14–17, 1997; Dublin, Ireland.
21. Unser M. Local linear transforms for texture measurements. *Signal Processing* 1986; 11:61–79.
22. Unser M. Texture classification and segmentation using wavelet frames. *IEEE Trans Image Processing* 1995; 4:1549–1560.
23. Kociołek M, Materka A, Strzelecki M, Szczypiński P. Discrete wavelet transform-derived features for digital image texture analysis. Paper presented at: International Conference on Signals and Electronic Systems; September 18–21, 2001; Lodz, Poland.
24. Webb AR. *Statistical Pattern Recognition*. Hoboken, NJ: John Wiley & Sons; 2003.
25. Fukunaga K. *Introduction to Statistical Pattern Recognition*. Amsterdam, the Netherlands: Elsevier; 1990.
26. Anderson JA, Rosenfeld E. *Neurocomputing*. Vol 2. Cambridge, MA: MIT Press; 1993.
27. Van Erkel AR, Pattynama PMT. Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. *Eur J Radiol* 1998; 27:88–94.
28. Gopinath B, Shanthi N. Support vector machine based diagnostic system for thyroid cancer using statistical texture features. *Asian Pac J Cancer Prev* 2013; 14:97–102.
29. Gopinath B, Shanthi N. Development of an automated medical diagnosis system for classifying thyroid tumor cells using multiple classifier fusion [published online ahead of print November 26, 2014]. *Technol Cancer Res Treat*. doi:10.7785/tcrt.2012.500430.