Ultrasound in Med. & Biol., Vol. 30, No. 10, pp. 1397-1407, 2004

0301-5629/04/$–see front matter

ELSEVIER

doi:10.1016/j.ultra.medbio.2004.08.023

● *Original Contribution*

BREAST TISSUE CHARACTERIZATION USING FARMA MODELING OF ULTRASONIC RF ECHO

Burak Alacam,* Birsen Yazici,* Nihat Bilgutay,† Flemming Forsberg,‡
Catherine Piccoli‡

*Rensselaer Polytechnic Institute Electrical, Computer, and Systems Engineering, Troy, NY, USA;
†Drexel University Electrical and Computer Engineering, Philadelphia, PA, USA; and
‡Thomas Jefferson University Department of Radiology, Philadelphia, PA, USA

(*Received* 15 *January* 2004; *revised* 16 *August* 2004; *in final form* 17 *August* 2004)

Abstract—A number of empirical and analytical studies demonstrated that the ultrasound RF echo reflected from tissue exhibits 1/*f* characteristics. In this paper, we propose to model 1/*f* characteristics of the ultrasonic RF echo by a novel parsimonious model, namely the fractional differencing auto regressive moving average (FARMA) process, and evaluated diagnostic value of model parameters for breast cancer malignancy differentiation. FARMA model captures the fractal and long term correlated nature of the backscattered speckle texture and facilitates robust efficient estimation of fractal parameters. In our study, in addition to the computer generated FARMA model parameters, we included patient age and radiologist's prebiopsy level of suspicion (LOS) as potential indicators of malignant and benign masses. We evaluated the performance of the proposed set of features using various classifiers and training methods using 120 *in vivo* breast images. Our study shows that the area under the receiver operating characteristics (ROC) curve of FARMA model parameters alone is superior to the area under the ROC curve of the radiologist's prebiopsy LOS. The area under the ROC curve of the three sets of features yields a value of 0.87, with a confidence interval of [0.85, 0.89], at a significance level of 0.05. Our results suggest that the proposed method of ultrasound RF echo model leads to parameters that can differentiate breast tumors with a relatively high precision. This set of RF echo features can be incorporated into a comprehensive computer-aided diagnostic system to aid physicians in breast cancer diagnosis. (E-mail: yazici@ecse.rpi.edu) **© 2004 World Federation for Ultrasound in Medicine & Biology.**

Key Words: **Ultrasound RF echo modeling, Fractional differencing, Tissue characterization, Breast cancer, Computer aided diagnosis.**

INTRODUCTION

Breast cancer is currently the second cause of mortality among women and affects more women in the world than any other type of cancer (IARC, 2001). In the past 50 y, the lifetime risk of breast cancer has nearly tripled in the USA. In the 1940s, a woman's lifetime risk of breast cancer was 1 in 22 (NCI, 2002). In the year 2003, the risk is 1 in 8 (ACS, 2003). Currently, X-ray mammography is the primary imaging modality for breast cancer detection, but in many cases the findings are not sufficiently specific and subsequent diagnostic testing is required. It has been reported that the likelihood of a lesion found by X-ray mammography being malignant is less than 30% (NCI, 2002).

Ultrasound is currently the most useful adjunctive imaging modality to X-ray mammography. Its real-time, high resolution imaging capability has made ultrasound an increasingly integral part of evaluation and diagnosis of breast tumors. Typically, malignancy diagnosis is based on the radiologist's qualitative evaluation of the morphological features in B-scan images. This approach, however, is prone to variability, due to lack of quantitative measures and varying degree of experience among radiologists. In an attempt to quantify ultrasound tumor evaluation, some cancer centers have adapted a scoring system for research purposes, whereby radiologists assign a value between a minimum and a maximum number, minimum being benign and maximum being malignant tumor (TJU Radiology Department, UPENN Radiology Department). This scoring system, known as level

Corresponding Author: Dr. Birsen Yazici, Rensselaer Polytechnic Institute, Electrical, Computer, and Systems Engineering, 110 8th Street, Jonsson Engineering Building JEC 7008, Troy, NY 12180, USA, PHONE: (518) 276-2905, FAX: (518) 276-6261. E-mail: yazici@ecse.rpi.edu

of suspicion (LOS), reflects the radiologist's qualitative evaluation of B-scan images of breast tumor, patient's overall case, as well as radiologist's level of experience.

Many studies point out that computer-aided diagnosis (CAD) can improve the radiologist's ability to differentiate between benign and malignant tumors, thereby potentially reducing unnecessary biopsies. CAD systems attempt to extract quantitative attributes from image data and automate lesion classification using various pattern recognition techniques. These attributes or features are typically based on the morphological features of tumors, such as shape, sharpness of borders, texture, as well as the RF echo signal reflected from tissue. Neither morphological nor RF echo features alone may be sufficient to characterize a tumor. Therefore, both humans and computers must consider multiple features to make a reliable diagnosis.

In this work, we developed a set of features extracted from ultrasound RF echo and evaluated its performance in breast cancer diagnosis. The proposed set of features capture the fractal and scaling characteristics of backscattered RF echo in a robust and parsimonious fashion. The motivation for our study stems from a number of empirical and analytical studies reported in the literature, demonstrating fractal or 1/*f* characteristics of the backscattered ultrasound RF echo signals. Karaoguz et al. (2000) has shown empirically that RF echo signal exhibits a 1/*f*-like scaling characteristics. Kutay et al. (2000, 2001) proposed a heavy-tailed shot noise model to capture the 1/*f*-like scaling properties of RF echo and reported relatively high accuracy for breast tissue characterization. Fractional differencing autoregressive moving average (FARMA) is a time series model developed to capture the long-term correlated fractal nature of 1/*f* processes observed for a broad range of scientific and engineering data. These include electronic noise in solid state devices, texture in natural and remote sensing imagery, internet traffic and financial data, to mention a few (Bonanno et al., 2001; Ilow, 2000; Kashyap and Lapsa, 1984; Kashyap and Eom, 1989; Xiong et al., 2002; Wen and Acharya, 1996). FARMA model has several advantages compared with other 1/*f* models (Kutay et al., 2001; MandelBrot and Van Ness, 1968; Yazici and Kashyap, 1996). FARMA is a discrete stationary model that can capture both long- and short-term correlations. Furthermore, stationary nature of FARMA model leads to robust, efficient and consistent estimation of model parameters that govern the fractal and scaling properties of 1/*f* processes. These properties are particularly pertinent for 1/*f* modeling of ultrasound RF echo. This is because ultrasound RF echo has a short-term correlated component due to transducer filter response and a long-term correlated component due to tissue response. Thus, FARMA modeling of RF echo allows direct parametrization of the tissue response from the RF echo signal. Furthermore, estimation of model parameters can be performed in a statistically efficient and robust manner with minimal computational requirement, due to the stationary nature of the underlying model (Hosking, 1981).

Apart from the studies involving 1/*f*-type modeling, there are a number of alternative RF echo models reported in the literature. Wagner et al. (1983) and Tuthill et al. (1988) proposed a Rician distribution to model envelope statistics of ultrasound RF echo. Cohen and Georgiou (1995, 1997) proposed WOLD decomposition of the echo to estimate mean scatterer spacing in tissue. Shankar? (1995, 1996) proposed *K*-distribution for ultrasound RF echo by observing the deviations from Rayleigh statistics. Narayanan et al. (1994) and Dutt and Greanleaf (1994) proposed generalized *K*-distribution and homodyned *K*-distribution models. Abeyratne et al. (1997a, 1997b) used higher order statistics for feature extraction from RF echoes. Recently, Shankar et al. (2001a, 2001b) proposed the Nakagami distribution and Donohue and Huang (2001) proposed the generalized spectrum to investigate discrimination of model parameters for benign and malignant lesions. These studies have reported a varying degree of complexity in RF echo modeling and feature estimation, as well as diagnostic accuracy.

Our study demonstrates that the FARMA modeling of ultrasound RF echo leads to relatively high diagnostic accuracy, with a statistically robust and efficient feature extraction procedure. We modeled transducer filter response as an autoregressive moving average (ARMA) process and tissue response as a fractional differencing (FD) process with 1/*f* parameter *d,* leading to the FARMA modeling of ultrasound RF echo. We estimated ARMA model parameters from phantom data, based on the final prediction error (FDE) and residual time series methods. Next, we estimated the mean and the variance of the FD model parameter *d* from the deconvolved RF echo using a log periodogram technique. The diagnostic value of various combinations of features are evaluated using 120 *in vivo* ultrasound breast images containing both benign and malignant tumors (75 benign and 45 malignant). The set of feature vectors include FD model parameters, radiologist's prebiopsy LOS and patient's age. In the training and classification stage, three different training techniques and three different classifiers were used. These include leave-one-out, hold-out and resubstitution methods and linear, quadratic and nonlinear classifiers. Finally, the receiver operator characteristics (ROC) for the proposed set of feature vectors, with respect to various classification methods and training technique, were derived and compared.

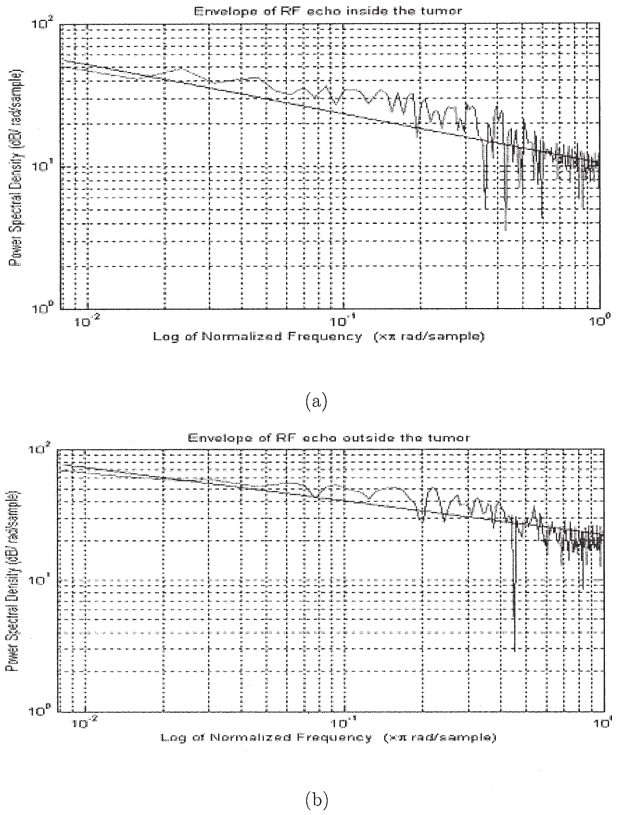


Fig. 1. The double logarithmic plot of the spectrum of the envelope of RF echo taken from (a) Inside and (b) Outside of the tumor region.

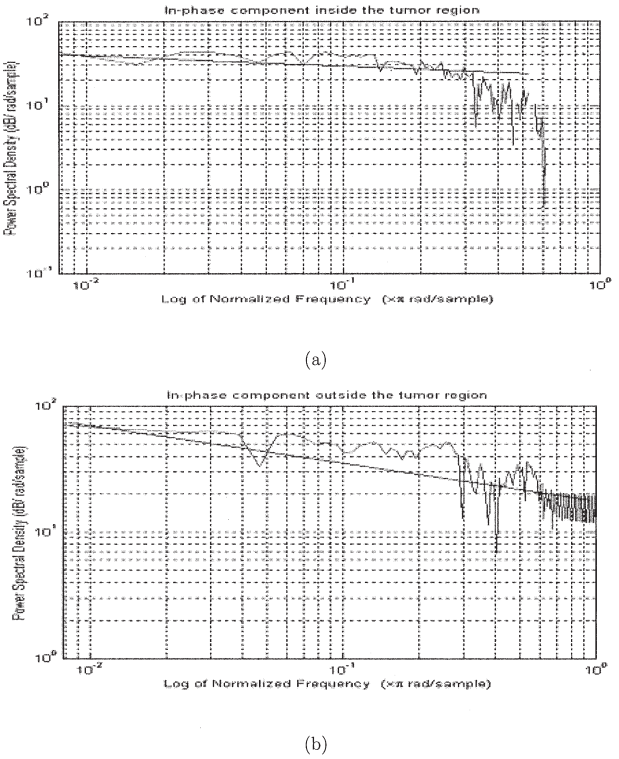


Fig. 2. The double logarithmic plot of the spectrum of the in-phase component of RF echo taken from (a) Inside and (b) Outside of the tumor region.

MATERIALS AND METHODS

FARMA model for RF echo

In ultrasonic applications, the RF echo scattered from tissue is typically modeled as a convolution integral of the ultrasonic pulse and the scattering structure as follows:

$$y(t) = \int_{-\infty}^{+t} x(\tau)h(t-\tau)d(\tau) \tag{1}$$

where $h(t)$ is the impulse response of the transducer and $x(t)$ is the ultrasonic tissue response.

It was empirically observed by Karaoguz et. al (2000) and Kutay et. al (2001) that the tissue response exhibits $1/f$ characteristics, due to the complex structure of tissue scatterers. $1/f$ processes are typically characterized by an empirical spectrum following a scaling behavior of the form $S(f) = 1/f^{\beta}$, where $0 < \beta < 2$. Figures 1 and 2 illustrate the log-log periodogram of the envelope and the in-phase components of the RF echo taken from inside and outside of the tumor region for a randomly chosen B-scan image. This linear behavior of the log-log spectrum is a manifestation of $1/f$ nature of ultrasound RF signals reflected from tissue.

From the point of view of texture modeling, similar observations were also made for natural terrain and texture data in remote sensing imagery (Cetin and Karl, 1998). In the latter case, self-similar and long-term correlated time series models were successfully used to model and classify image texture (Langer, 2000; Rainville and Kingdom, 1999). Recently, the FD model was used to capture the self-similar nature of the network traffic (Ilow 2000). FD model is a discrete stationary process with self-similar and long-term correlated structure (Anderson, 1976; Box and Jenkins, 1976; Hosking, 1981). It can be compactly represented as follows:

$$x(n) = (1 - z^{-1})^{-d}\sqrt{\rho}w(n) \tag{2}$$

where z^{-1} is the unit delay operator, $w(n)$ is discrete white Gaussian noise sequence with zero mean and unit variance, $\rho > 0$ is a constant and $0 < d < 0.5$ is the FD model parameter. The spectral density function of FD process is given by:

$$S_x(f) = (2\sin\frac{1}{2}f)^{-2d}, \quad for \;\; 0 < f \leq \pi. \tag{3}$$

For $f \rightarrow 0$, it can be approximated as $S_x(f) \sim f^{-2d}$. As a result, the FD model parameter d captures the scaling and fractal nature of the process.

We model tissue response as a fractional differencing (FD) model and transducer response as an ARMA (p,q) model, which leads to the FARMA modeling of the ultrasonic RF echo. FARMA is a discrete model that can capture both long- and short-term correlations and leads to robust, efficient and consistent estimation of the model parameters that govern the fractal and scaling properties of $1/f$ processes. Since the ultrasound RF echo has a short-term correlated component due to transducer filter response and a long-term correlated component due to tissue response, these properties are particularly pertinent for $1/f$ modeling of the ultrasound RF echo. As a result, FARMA modeling of the RF echo allows direct parametrization of the tissue response from the RF echo signal.

FARMA model can be represented as follows:

$$A(z^{-1})x(n) = B(z^{-1})(1 - z^{-1})^{-d}w(n) \quad (4)$$

where z^{-1} is a unit delay operator, $A(z^{-1})$ and $B(z^{-1})$ are the autoregressive and moving-average polynomials of orders p and q, respectively.

The transfer function of the RF echo is given by:

$$H(z^{-1}) = \frac{1}{(1 - z^{-1})^d} H_{ARMA}(z^{-1}) \quad (5)$$

where

$$H_{ARMA} = \frac{B(z^{-1})}{A(z^{-1})}$$

is the auto-regressive, moving-average part representing the transducer response and

$$\frac{1}{(1 - z^{-1})^d}$$

is the FD model representing the tissue response.

Estimation of FARMA model parameters

We estimate the FARMA model parameters in two steps. First, ARMA parameters of the transducer response are estimated. Next, estimated ARMA parameters are utilized to estimate the FD parameters of the tissue response. The B-scan images used throughout the project were obtained at the Radiology Department of Thomas Jefferson University Hospital in Philadelphia, PA, USA. Images were obtained using a flat linear broadband array transducer with a central frequency of 7.5 MHz. The ultrasound imaging system is Ultramark-9, HDI, manufactured by Advanced Technology Laboratories, Bothell, WA, USA. The data were sampled at 20 MHz using 12 bits for quantization after applying analog time-gain control. The digital images consist of 1024 axial points by 192 lateral points.

Estimation of ARMA parameters

ARMA parameter estimation is based on the transducer impulse response data, obtained using pulse-echo measurements from a flat surface reflector in water. Figure 3 presents (a) B-scan, (b) A-scan and (c) greater detailed A-scan around first reflection region measurements of the transducer response for a randomly chosen image.

Final prediction error (FPE) criterion is used to estimate the order of the ARMA type transducer response. Akaike's final prediction error (Akaike, 1970) is a well-known statistical measure for the goodness-of-fit of ARMA (p, q) model. FPE is a function of residuals given by:

$$FPE = \frac{1 + n/N}{1 - n/N} V \quad (6)$$

where V is the variance of model residuals, N is the length of the time series and $n = p + q$ is the number of estimated parameters in the ARMA model.

Two windows, composed of 30 consecutive scan lines with data length of 1×256, were used to estimate the ARMA model order and parameters. A window selection criterion for the estimation procedure is given in Figure 4. The windows were selected from the first reflection region of the ultrasonic RF pulse. The FPE is computed for various candidate model orders and the model order with the lowest FPE was selected as the best-fit model. Based on our numerical studies and experimental data, we concluded that the best estimated model order is an ARMA (3,1) for the transducer response obtained from a flat surface reflector in water.

After the estimation of ARMA model orders, ARMA model parameters p and q were estimated using the well-known residual time series model estimation technique (Brockwell and Davis, 1991; Chatfield, 1975; Roger, 1999). In this estimation procedure, the ARMA parameters p and q were estimated for 30 scanlines taken for the two windows. Sixty parameters obtained from two selected windows for each coefficient are averaged to estimate the ARMA model coefficients. The estimation resulted in the following AR and MA polynomials:

$$A(z^{-1}) = 1 - 1.192z + 0.06404z^2 - 0.4332z^3 \quad (7)$$

$$B(z^{-1}) = 1 + 0.7347z \quad (8)$$

where z^{-1} is a unit delay operator.

Estimation of fractional differencing model parameter

After the estimation of ARMA model parameters, the transducer response is deconvolved from the RF echo

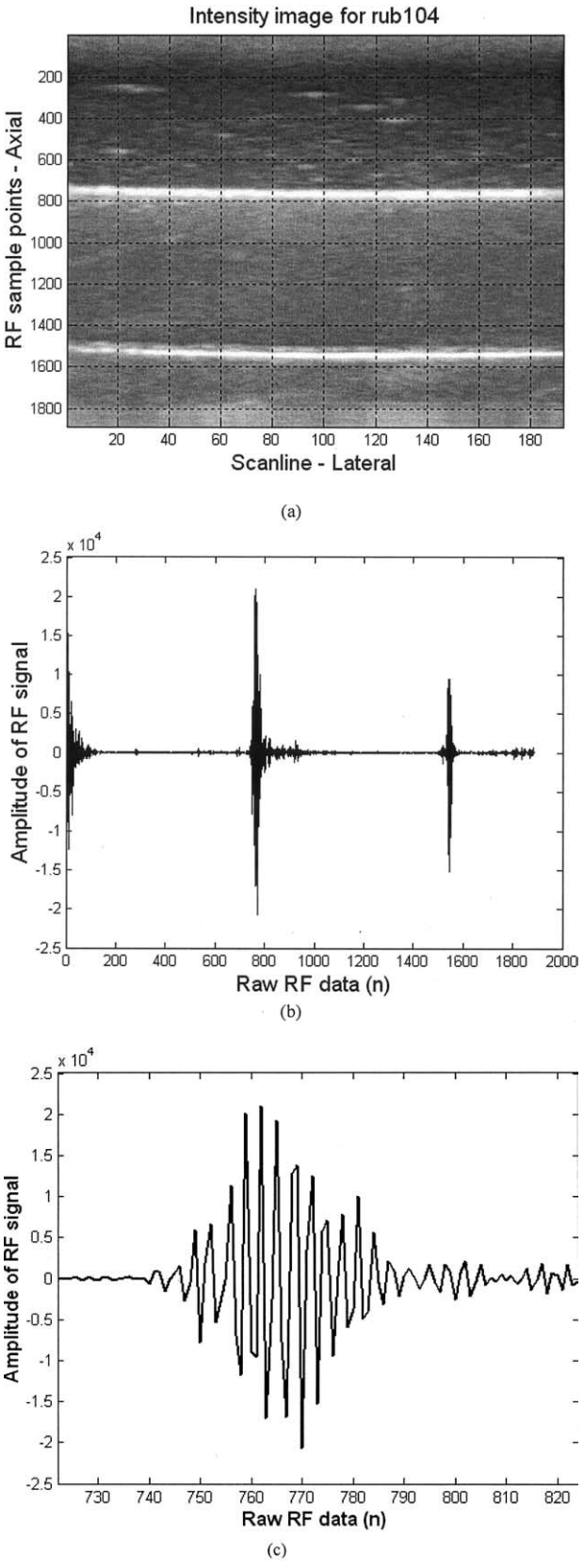


Fig. 3. (a) B-scan image of RF echo obtained from water/flat reflector surface. (b) A-scan data sequence of RF echo obtained from a flat-surface reflector in water. (c) A-scan data sequence zoomed around first reflection region.

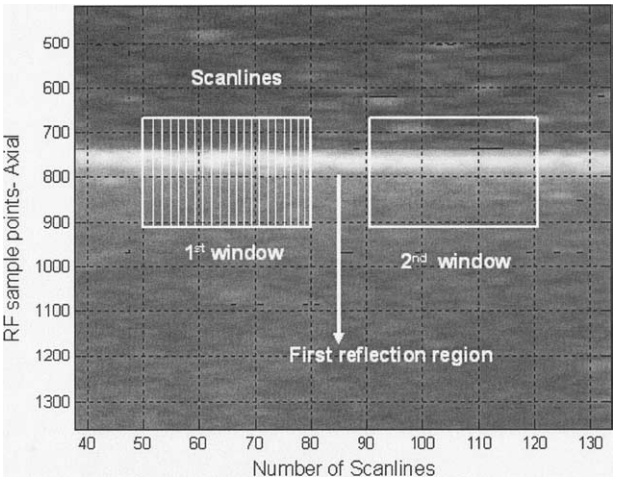


Fig. 4. Window selection for the estimation of ARMA parameters. The windows were chosen from the first reflection region from water surface.

and the FD model parameter d is estimated from the resulting data. The estimation procedure is based on the log-periodogram method that uses linear least-squares procedure (Kashyap and Eom, 1987). Recall that the FD model is given by:

$$x(n) = (1 - z^{-1})^{-d}\sqrt{\rho}w(n) \tag{9}$$

where $w(n)$ is a zero mean, unit variance discrete white noise process. The periodogram of x(n) is given by:

$$f_x(k/N) = |X(k)|^2/N \tag{10}$$

where X(k) is the N point discrete Fourier transform (DFT) of x(n). Using the FD model, it is straightforward to show that:

$$f_x(k/N) = |2sin\ (\pi k/N)|^{-2d}\rho f_w(k/N) \tag{11}$$

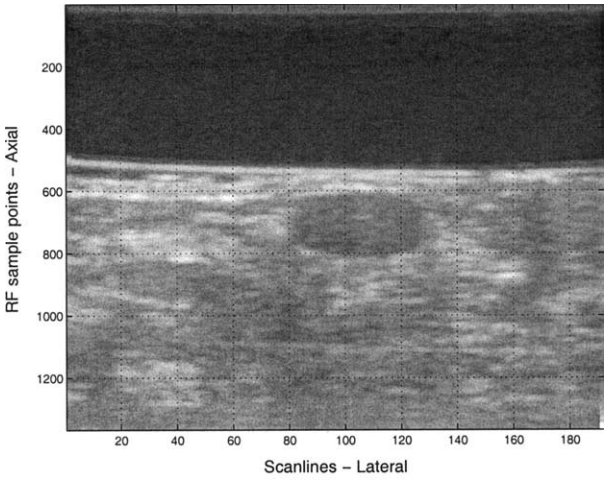
where f_w is the N point DFT of the white noise process $w(n)$. Thus, combining eqns (11) and (12) and taking the logarithms of both sides, we obtain:

$$\log |X(k)|^2/N = \log f_x(k/N)$$
$$= -2d\log |2\sin (\pi k/N)| + \log \rho + \log |f_w(k/N)|. \tag{12}$$

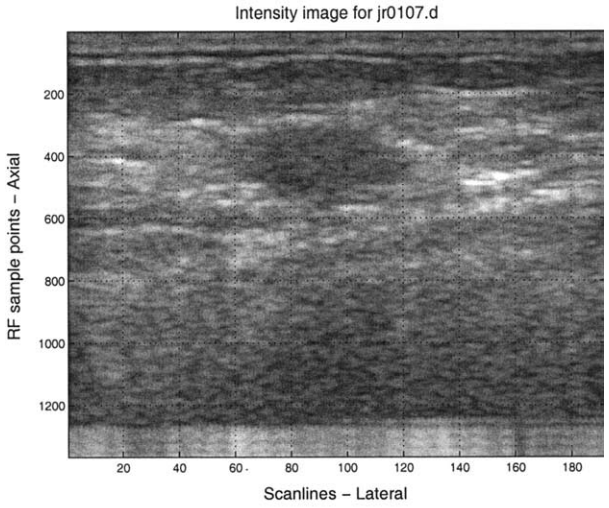
Note that the FD model parameter d and $log\rho$ are linear in $logf_x(k/N)$. Thus, d and $log\rho$ can be estimated using linear least-squares technique. The estimates are given by the following standard formula:

$$[d, log\ \rho] = \left[\sum_{k=1}^{N2} Z(k)Z^T(k)\right]^{-1}\left[\sum_{k=1}^{N2} Z(k)log\ f_x(k/N)\right] \tag{13}$$

where $Z(k) = [-2log|2sin(\pi k/N)| -1]^T$.



(a)



(b)

Fig. 5. Ultrasonic B-scan images of a benign and a malignant tumor. (a) Benign tumor. (b) Malignant tumor.

The FD parameter, *d,* is estimated from 120 different B-scan images acquired from 90 patients. Thirty-five of these patients have malignant breast tumors and 55 of them have benign breast tumors. Figure 5a and b shows two sample B-scan images with benign and malignant tumor regions, respectively. The smallest tumor is 6×5 mm^2 and the largest is 15×8 mm^2. We have 30 scan lines and 128 RF points per scan line, providing 3840 points per estimation. This corresponds to a window size of 5.925×4.8 mm^2. For each B-scan image, 30 scan lines from inside and outside the tumor region were taken with data lengths of 1×128 to estimate the fractional differencing parameter *d.* Hence, we have 30 estimates of *d* from inside and 30 estimates of *d* from

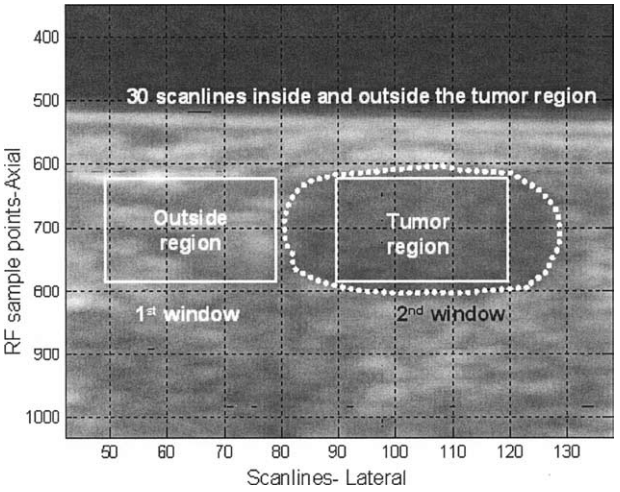


Fig. 6. Window selection for the estimation of FD parameter *d.* The windows were chosen from inside and outside the tumor region.

outside the tumor region. Figure 6 illustrates the windows taken from inside and outside the tumor region for a random B-scan image. The selection of windows can be automated in a CAD system, once the radiologist identifies a region-of-interest. For each B-scan image, mean and variance of the fractional differencing parameter *d* were estimated for classification purposes. In Figures 7 and 8, the mean values of the parameter *d* obtained from inside and outside the suspected tumor region are presented for 31 patients with benign and 29 patients malignant tumors, respectively.

Feature analysis and classification for tissue characterization

In this subsection, we present the set of tumor classification features and the malignancy differentiation criteria. F1 and F2 denote the mean and variance of FD parameter *d,* F3 denotes radiologist's prebiopsy level of suspicion (LOS) and F4 denotes the patient age information. Number of lesions were evaluated by experienced

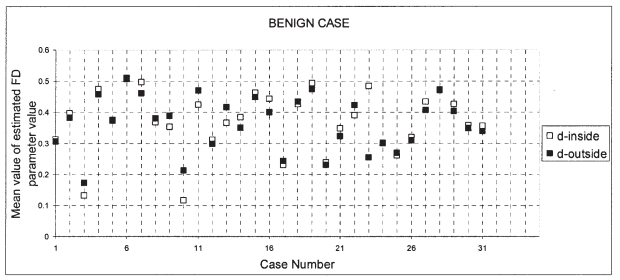


Fig. 7. Mean value of fractional differencing parameter *d* for inside and outside the tumor region for 31 patients with benign tumor in breast.

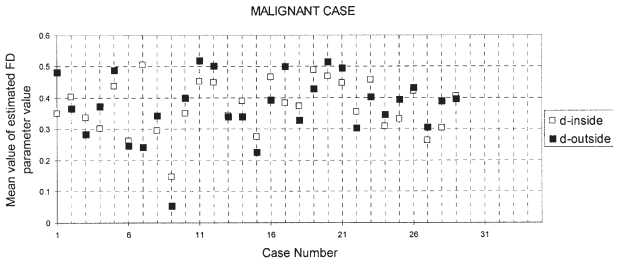


Fig. 8. Mean value of fractional differencing parameter *d* for inside and outside the tumor region for 29 patients with malignant tumor in breast.

radiologists. Lesions were categorized based on shape, orientation, margin sharpness and posterior acoustic features. The level of suspicion was classified as (1) normal, (2) benign, (3) probably benign, (4) moderate suspicious and (5) highly suspicious for malignancy. The ages of the patients ranged from 18 to 86 y, with the average being 47 y. We evaluate the malignancy differentiation capability of the individual features and various combinations of the these features.

As a malignancy differentiation criterion, we used the Neyman-Pearson hypothesis testing method (Fukunaga, 1990). Neyman-Pearson decision rule maximizes true positive rate subject to a given false positive rate. True positive rate is defined as the ratio between the number of correct malignant decisions and total number of malignant cases. False positive rate is the ratio between the number of incorrect malignant decisions and total number of benign cases. Let F be a set of features chosen from F1 to F4. The Neyman-Pearson statistics is given by the following likelihood ratio:

$$\frac{P(F|H_1)}{P(F|H_1)} = l(F) \tag{14}$$

where $P(F|H_0)$ is the conditional probability of F, given that the tumor is benign and $P(F|H_1)$ is the conditional probability of F, given that the tumor is malignant. In Neyman-Pearson likelihood ratio test, a threshold value τ_α is chosen so that the false positive probability is equal to α, i.e.,

$$\alpha = P(l(F) > \tau_\alpha | H_0). \tag{15}$$

This results in the following malignancy differentiation criteria:

$$H_0: \quad l(F) < \tau_\alpha \quad \Rightarrow \quad \text{Benign tumor}$$

$$H_1: \quad l(F) > \tau_\alpha \quad \Rightarrow \quad \text{Malignant tumor}$$

Neyman-Pearson test is implemented using a linear classifier (minimum least-squares linear classifier), a quadratic classifier (Gaussian normal density based quadratic classifier) and a nonlinear classifier (Parzen classifier) (Fukunaga, 1990). The 95% confidence interval for the area under the ROC curves is obtained by using three different training techniques, namely resubstitution, leave-one-out and hold-out methods (Duda, 2000; Fukunaga, 1990; Hastie, 2001). Figures 9, 10, and 11 show the distribution of various subsets of features extracted from benign and malignant tissues and the thresholds computed using linear and quadratic classifiers. Figure 12 shows the data scatter information for the complete set of features.

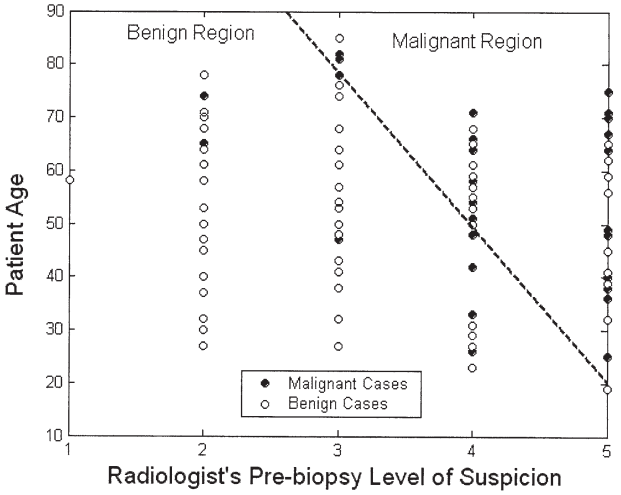


Fig. 9. Linear classifier and LOS vs patient age information 2-D data clustering.

RESULTS AND DISCUSSION

ROC analysis of individual and combined features

We evaluated the malignancy differentiation capability of the following individual and combined features:

F1-F2: Computer generated FD parameters
F3: Radiologists' prebiopsy LOS
F4: Patient age information
F1-F2-F3: FD parameters and LOS
F1-F2-F4: FD parameters and patient age
F3-F4: LOS and patientage
F1-F2-F3-F4: All features.

The evaluation is based on receiver operating characteristics (ROC) methodology. The ROC curve is obtained by plotting the probability of false positive versus the probability of detection. Figure 13 shows the ROC curves of individual features using the linear classifier and the leave-one-out method of training. The observed area under the ROC curve ranges from 0.62 to 0.79 for four different features. The best-performing feature set is the combination of the computer generated features, that

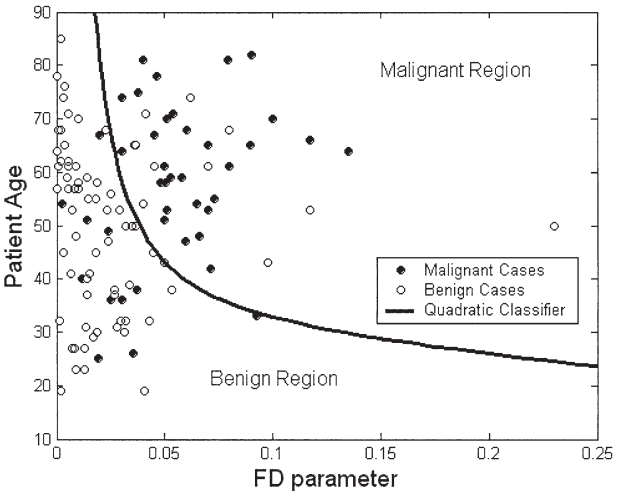


Fig. 10. Quadratic classifier and FD parameter vs patient age information 2-D data clustering.

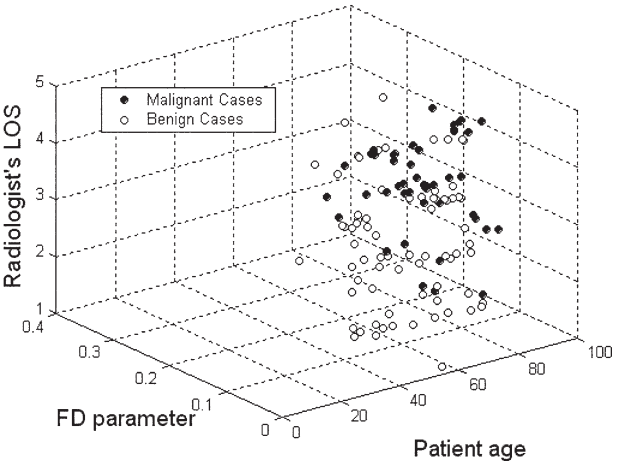


Fig. 12. LOS vs FD parameter vs patient age 3-D data clustering.

is, the mean and variance of the FD parameter *d,* with an A_z (area under ROC) of 0.7973. The A_z values of radiologist's prebiopsy level of suspicion (LOS) and patient age are 0.7411 and 0.6243, respectively.

The area under the ROC curve indicates that the computer-generated FD features provides approximately 6% improvement over the radiologist's prebiopsy LOS and 18% improvement over the patient age information. However, since A_z, the area under the ROC curve, is obtained from patient data, we proposed a hypothesis-testing method to compare the ROC performance of the computer generated features, F1 and F2, with that of F3, i.e., radiologist's prebiopsy level of suspicion. The method is based on the difference of the area of ROC curve statistics, $A_z(F1, F2)$–$A_z(F3)$. Here, $A_z(F1, F2)$ denotes the area under the ROC curve of F1 and F2 and $A_z(F3)$ denotes the area under the ROC curve of F3. The null hypothesis, H_0, is that there is no difference in performance between FD features, F1 and F2, and radiologist's LOS, F3, against the alternative hypothesis, H_1, that the performance of F1 and F2 is better than the performance of F3.

H_0: $A_z(F1, F2) = A_z(F3)$ $\Rightarrow$

No difference in performance

H_1: $A_z(F1, F2) > A_z(F3)$ $\Rightarrow$

Statistical difference in performance.

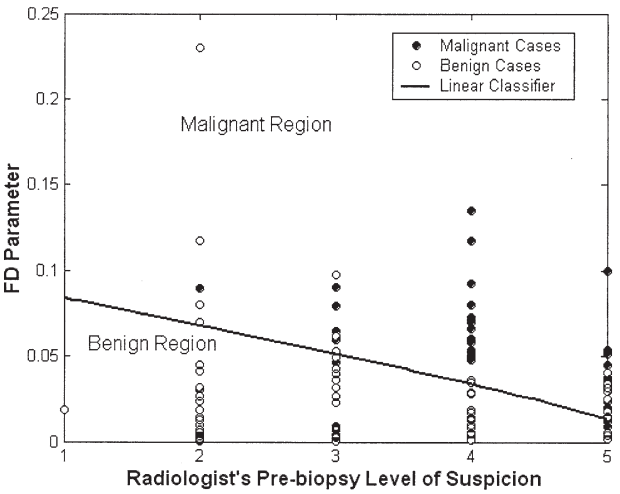


Fig. 11. Linear classifier and LOS vs FD parameter 2-D data clustering.

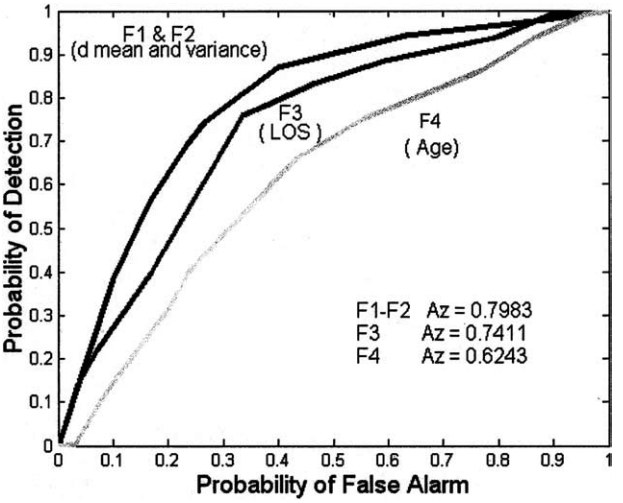


Fig. 13. ROC curves for individual features using a linear classifier with leave-one-out technique.

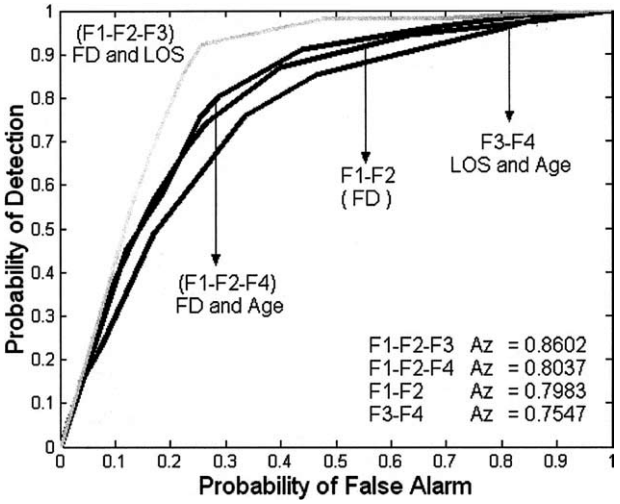


Fig. 14. Four ROC curves for the mean, variance of fractional differencing parameter d (F1 and F2) and radiologist's prebiopsy level of suspicion, patient age (F3 F4), F1-F2-F3 and F1-F2-F4 using a linear classifier with leave-one-out technique.

We use the P-value in the hypothesis-testing. A P-value is a measure of how much evidence there is against the null hypotheses. We reject the null hypothesis H_0 if the P-value is smaller than α (a predefined significance level), otherwise we accept the null hypothesis. In our study, we used a significance level of 5%. As a result, the hypothesis test described above is reexpressed as follows:

H_0: $P > 0.05 \Rightarrow$ No difference in performance

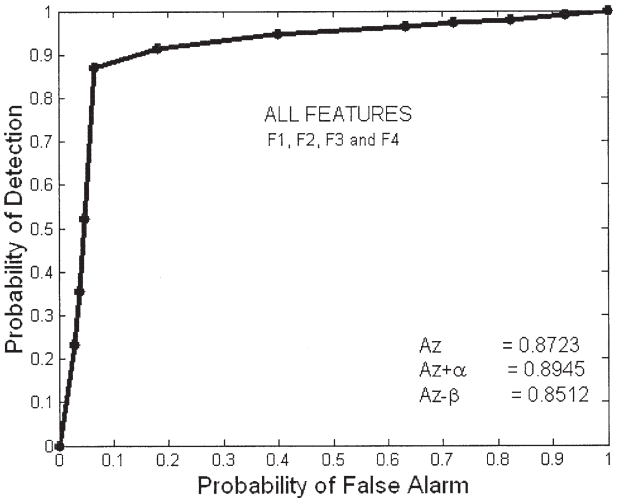


Fig. 15. ROC curve for the complete feature set, which are mean and variance of fractional differencing parameter (F1 and F2), radiologist's prebiopsy level of suspicion (F3) and patient age (F4) using a linear classifier with leave-one-out technique.

Table 1. Performance of combined features fused with linear classifier at a significance level of 0.05. A_z is the area under the ROC curve, α and β are positive and negative error bounds, respectively.

	Linear resubstitution	Linear leave-one-out	Linear-hold-out
A_z	0.88	0.87	0.80
$A_z + \alpha$	0.90	0.89	0.83
$A_z - \beta$	0.86	0.85	0.79

H_1: $P < 0.05 \Rightarrow$ Statistical difference in performance.

The A_z difference between F1-F2 and F3, using a linear classifier and leave-one-out technique, is 0.0572 with a P-value of 0.0246. Similarly, the ROC area difference between F1-F2 and the patient age, F4, is 0.1740 with a P-value of 0.0012. Therefore, we concluded that the ROC performance of computer-generated FD parameters is statistically better than that of LOS and patient age information.

Figure 14 presents a comparison between various subsets of features, including F1-F2, F1-F2-F3 and F1-F2-F4 and F3-F4. The performance of fractional differencing parameter d is better than the radiologist's decision and patient age information by 5%, which is a crucial percentage when we think of unnecessary biopsies. When the radiologists' prebiopsy LOS is combined with the patient age information, the area under the ROC curve is better than ROC area of the LOS parameter alone. This result indicates that the radiologist's prebiopsy LOS does not contain the patient's age information. When LOS or age is combined with FD parameter, the performance of classification increases. The addition of different features to the computer-generated features improves the results that made us think increasing the number of features. This study is currently underway and will be reported in the near feature.

Figure 15 shows the ROC curve for the feature set F1-F2-F3-F4. The best classification results are obtained when all four features are fused. Combining the features F3 and F4 with the features extracted from RF echo modeling increased the area under the ROC curve from

Table 2. Performance of combined features fused with quadratic classifier at a significance level of 0.05. A_z is the area under the ROC curve, α and β are positive and negative error bounds, respectively.

	Quadratic resubstitution	Quadratic leave-one-out	Quadratic hold-out
A_z	0.85	0.82	0.77
$A_z + \alpha$	0.86	0.84	0.80
$A_z - \beta$	0.82	0.79	0.75

Table 3. Performance of combined features fused with nonlinear classifier at a significance level of 0.05. A_z is the area under the ROC curve, α and β are positive and negative error bounds, respectively.

	Nonlinear resubstitution	Nonlinear leave-one-out	Nonlinear hold-out
A_z	0.84	0.81	0.76
$A_z + \alpha$	0.83	0.86	0.78
$A_z - \beta$	0.83	0.80	0.75

0.79 to 0.87. Alternatively, one can conclude that the addition of RF echo model parameters to radiologist's LOS increased the value from 0.74 to 0.87, an approximate 13% improvement over the LOS feature alone.

Tables 1, 2 and 3 tabulate the performance of the combined features, i.e., F1-F2-F3-F4, with respect to linear, quadratic and nonlinear classifiers for three different training techniques. Our study shows that the linear classifier yields better results in comparison with the other two classifiers. As expected, the resubstitution method provides the best results in comparison with hold-out and leave-one-out method. Leave-one-out-method provides better results than hold-out method which divides the total number of cases into two as the training and test sets.

In conclusion, our study indicates that the proposed set of features form a good candidate to extract diagnostic information from the RF echo. We believe that this set of features will lead to a computer-aided diagnosis system that can reduce the number of unnecessary biopsies performed on benign masses.

SUMMARY

In this paper, we showed that features obtained by statistical modeling of RF echo can be used as a decision criterion for tissue characterization. In particular, we propose FARMA model to capture *1/f* characteristics of the RF echo backscattered from tissue and evaluate their potential malignancy differentiation. Additionally, we used radiologist's prebiopsy LOS and patient age information as potential malignancy differentiators. We used a database of *in-vivo* B-scan images to generate the features. Features are evaluated using different classifiers and training techniques, based on ROC methodology. The observed area under the ROC of the computer-generated features yields an approximate 6% improvement over that of radiologist's prebiopsy LOS criteria. The best performance is obtained when all features are fused using a linear classifier, yielding an area of 0.87 with 95% confidence interval under the area of the ROC curve. This is an improvement of 13% over that of radiologist's prebiopsy LOS alone. Our study indicates that the proposed RF echo model and associated model parameters provide relatively high malignancy differentiation information. We believe that, when this set of features are combined with morphological image analysis features, it will lead to a comprehensive computer-aided diagnosis system that can aid radiologists in breast cancer diagnosis.

Acknowledgments This work was supported in part by the National Institutes of Health, National Cancer Institute, grant number CA 52823.

REFERENCES

Abeyratne UR, Petropolu AP, Reid JM. Higher-order statistics for tissue characterization from ultrasound images. Proceedings of the IEEE Signal Processing Workshop 1997;72–76.

Abeyratne UR, Petropolu AP, Reid JM, Golas T, Conant E, Forsberg F. Higher order versus second order statistics in ultrasound image deconvolution. IEEE Trans Ultrason Ferroelect Freq Contr 1997; 44–6:1409–1416.

Akaike H. Statistical Predictor Identification. Annual Ins Statist Math 1970;22:203–217.

American Cancer Society. Cancer Facts and Figures 2003.

Anderson O. Time Series Analysis and Forecasting: The Box-Jenkins Approach. London:Butterworths, 1976.

Bonanno G, Lillo F, Mantegna RN. $1/f$ and $1/f^2$ noise in financial time series. Proceedings of 16th International Conference on Noise in Physical Systems and $1/f$ Fluctuations (World Scientific). October 22-25, 2001, Gainesville, FL. 2001; pp 791–796.

Box GEP, Jenkins GM. Time Series Analysis: Forecasting and Control. San Francisco: Holdan Day, 1976.

Brockwell PJ, Davis RA. Time Series: Theory and Methods. New York: Springer-Verlag, 1991.

Cetin M, Karl WC. A statistical method for discrimination of natural terrain types for SAR data. Proceedings of IEEE Conference on Image Processing. October 4-7, 1998, Chicago, IL. 1998;1:587–591.

Chatfield C. The Analysis of Time Series: Theory and Practice. London: Chapman and Hall, 1975;2:488–491.

Cohen FS, Georgiou G. Detecting and estimating structure regularity of soft tissue organs from ultrasound images. Proceedings of the IEEE International Conference on Image Processing. October 23-26, 1995, Washington, DC. 1995;2:488–491.

Cohen FS, Georgiou G. WOLD decomposition of the backscatter echo in ultrasound images of soft tissue organs. IEEE Trans Ultrason Ferroelec Freq Contr 1997;44–2:460–472.

Donohue KD, Huang L. Ultrasonic scatterer structure classification with the generalized spectrum. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. May 7-11, 2001, Salt Lake City, UT. 2001;6:3401–3404.

Duda RO, Hart PE, Stork DG. Pattern Classification. New York: Wiley-Interscience, 2000.

Dutt V, Greenleaf JF. Ultrasound echo envelope analysis using a homodyned K distrubution signal model. Ultrason Imaging 1994; 16:265–287.

Fukunaga K. Introduction to Statistical Pattern Recognition. New York: Academic Press, 1990.

Hastie T. The Elements of Statistical Learning. New York: Springer Verlag, 2001.

Hosking JRM. Fractional differencing. Biometrika 1981;V68:165–176.

Ilow J. Parameter estimation in FARIMA processes with applications to network traffic modeling. IEEE Trans Com 2000;505–509.

International Agency for Research on Cancer. Cancer Incidence, Mortality and Prevalence Worldwide. IARC Cancer Base No. 5. Lyon, IARC Press, 2001.

Karaoguz M, Bilgutay N, Onaral B. Modeling of scattering dominated ultrasonic attenuation using power-law function. IEEE Ultrasonics Symposium 2000;1:793–796.

Kashyap RL, Lapsa PM. Synthesis and estimation of random fields using long correlation models. IEEE Transactions on Pattern Analysis and Machine Intelligence. PAMI 1984;V6:800–809.

Kashyap RL, Eom KB. Estimation in Long-memory Time Series Model. Journal of Time Series Analysis 1987;9–1:35–41.

Kashyap RL, Eom KB. Texture boundary detection based on the long correlation model. IEEE Transactions on Pattern Analysis and Machine Intelligence. PAMI 1989;V11:58–67.

Kutay MA, Petropulu AP, Reid JM, Piccoli K. Malignant versus benign tumor classification based on ultrasonic B-scan images of the breast. IEEE ultrasonics symposium 2000;1383–1386.

Kutay MA, Petropulu AP, Piccoli W. On modeling a biomedical ultrasound RF echoes using a power-law shot-noise model. IEEE Trans Ultrason Ferroelect, Freq Cont 2001;48–4:953–968.

Langer MS. Large scale failures of f^{α} scaling in natural image spectra. J Opt Soc Am 2000;17–1:28–33.

Mandelbrot BB, Van Ness IW. Fractional Brownian motions, fractional noises, and applications. Society Ind. Appl Math (SIAM) 1968;10:422–437.

Narayanan VM, Shankar PM, Reid JM. Non-Rayleigh statistics of ultrasonic backscattered signals. IEEE Trans Ultrason Ferroelect Freq Contr 1994;41–6;845–852.

National Cancer Institute. Cancer Facts: Lifetime probability of breast cancer in American women. 2002.

Rainville SJ, Kingdom FA. Spatial-scale contribution to the detection of mirror symmetry in fractal noise. J Opt Soc Am 1999;16–9:2112–2123.

Roger EK. Statistics: An Introduction, New York: Baylor University, 1999.

Shankar PM, Molthen RC, Narayanan VM. Further studies on the use of non-Rayleigh statistics for ultrasonic tissue characterization. Ultrasound Med Biol 1996;22–7:873–882.

Shankar PM. A model for ultrasonic scattering from tissues based on K-distribution. Phys Med Biol 1995;40:1633–1649.

Shankar PM, Dumane VA, Reid JM. Classification of ultrasonic B-mode images of breast masses using Nakagami distribution. IEEE Trans Ultrason Ferroelect Freq Contr 2001;48–2:569–580.

Shankar PM, Dumane VA. Use of frequency diversity and Nakagami statistics in ultrasonic tissue characterization. IEEE Trans Ultrason Ferroelect Freq Contr 2001;48–4:1139–1145.

Thomas Jefferson University, Radiology Department, Philadelphia, PA, USA.

Tuthill TA, Sperry RH, Parker KJ. Deviations from Rayleigh statistics in ultrasonic speckle. Ultrason Imaging 1988;10:81–89.

University of Pennsylvania, Radiology Department, Philadelphia, PA, USA.

Wagner RF, Smith SW, Sandrick JM, Lopez H. Statistics of speckle in ultrasound B-scans. IEEE Trans Son Ultrason 1983;30:156–163.

Wen CY, Acharya R. Fractal Analysis of Self-Similar Textures Using a Fourier-Domain Maximum Likelihood Estimation Method. Image Processing International Conference. Lausanne, Switzerland, September 1996. 1998;A:165–168.

Xiong HD, Fleetwood DM, Choi DK, Sternberg AL. Temperature dependence and irradiation response of 1/f-noise in MOSFETs. IEEE Trans Nuc Sci 2002;49–6:2718–2723.

Yazici B, Kashyap RL. Signal modeling and parameter estimation for 1/*f* processes using scale stationary models. Proceedings of ICASSP May 7-10, 1996, Atlanta, GA 1996;2841–2844.