BREAST CANCER DETECTION BASED ON ULTRASOUND B-SCAN TEXTURE ANALYSIS AND PATIENT AGE INFORMATION

Burak Alacam, Birsen Yazici, and Nihat Bilgutay

Department of Electrical and Computer Engineering, Drexel University, Philadelphia, PA 19104

Abstract— In this paper we apply the Fractional Differencing Autoregressive Moving Average (FARMA) model for ultrasonic breast tissue characterization using RF echo signals. We present estimation techniques to extract the model parameters, namely features, for classification purposes and tissue characterization. Along with the model parameters, we use patient age information as an additional feature to improve ROC results. We evaluate the performance of the proposed method using *in vivo* ultrasound breast images with benign and malignant tumors.

I. INTRODUCTION

The information contained in ultrasound RF echo can be used to distinguish between tumors thorough tissue characterization features [1]. In this paper we model ultrasound RF echo as a Fractional Differencing Auto-regressive Moving-average (FARMA) process to capture the speckle texture. It has been observed empirically by a number of researchers that RF echo exhibits $1/f$ characteristics [2], [3]. This is consistent with the $1/f$ behavior of image textures in computer vision [4]. Transducer response is modeled as an ARMA process, and tissue response is modeled as a fractional differencing process (FDM), leading to a FARMA process for RF echo.

Transducer response, i.e., ARMA parameters, are estimated using phantom data based on the final prediction error (FDE) and residual time series methods. Next, the transducer response is deconvolved from the RF echo and the FD parameter d is estimated from the resulting signal, based on a log periodogram technique. The sample mean and the variance of the parameter d from the inside and the outside of the suspected tumor are estimated next using multiple adjacent scanlines. A feature vector based on parameter d and patient age is formed. Classification is performed using a quadrature classifier under Gaussian assumption. Finally, the receiver operator characteristics (ROC) for the proposed method is derived using 60 *in vivo* ultrasound breast images containing both benign and malignant tumors. The resulting ROC has an area of 0.83, shows that the proposed method can be used effectively to differentiate between benign and malignant breast tumors.

II. FARMA MODEL FOR RF ECHO

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

$$y(n) = h(n) * x(n) \qquad (1)$$

where $h(n)$ is the impulse response of the transducer and $x(n)$ is the response from tissue.

It was empirically observed that the tissue response exhibits $1/f$ characteristics due to the complex structure of tissue scatterers [3]. In this paper, we propose to model ultrasonic tissue response as a fractional differencing process [5]. FD model has a number of advantages over other $1/f$ models [6], [7]. FD is a discrete, stationary process with self-similar and long-term correlated structure. It is governed by a long-term correlation parameter d, $0 < d < 0.5$. We model the transducer response as an ARMA(p,q) model, which leads to the FARMA modeling of the ultrasonic RF echo. FARMA process can be represented as in (2):

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

where $w(n)$ is a white Gaussian noise sequence with zero mean and unit variance, z^{-1} is unit delay operator, $A(z^{-1})$, and $B(z^{-1})$ are the autoregressive and moving-average polynomials of orders p and q, respectively.

III. ESTIMATION OF FARMA MODEL PARAMETERS FOR TISSUE CHARACTERIZATION

The ARMA parameter estimation is based on the transducer impulse response data, obtained using pulse-echo measurements from a flat surface reflector in water. The Final Prediction Error (FPE) criterion is used to estimate the order of the ARMA transducer response. The best model order for our transducer is an ARMA (3,1). The ARMA (3,1) parameters are estimated using the residual time series model.

Fractional differencing model parameter. d of the tissue response was estimated from the FARMA process. The estimation procedure is based on the log-periodogram method [8] that uses a linear least square procedure. In order to estimate the fractional differencing parameter d, 60 different data sets from 60 different B-scan images were used. 29 of these patients have malignant tumor and 31 of them have a benign breast tumor. The B-scan (grey scale) breast images were obtained at the Radiology Department of Thomas Jefferson University Hospital Philadelphia, PA.

For each B-scan image, 30 scanlines from the inside and outside the tumor region were taken with data lengths of 1x128, and each of these scanlines were used to estimate the fractional differencing parameter d. Hence, we have 30 values of d from inside and 30 values of d from outside the
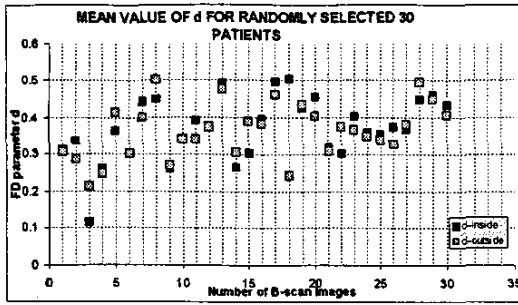
Fig. 1. Mean values of fractional differencing parameter d for inside and outside the tumor region for randomly selected 30 patient
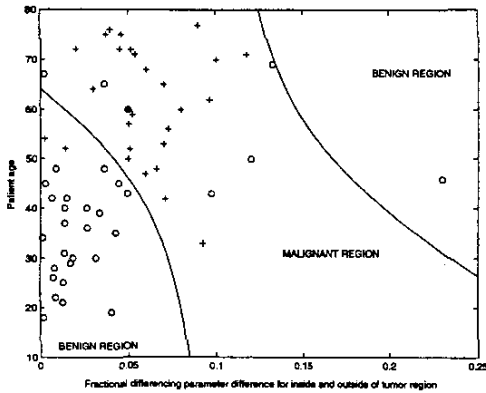

Fig. 2. Quadratic classifier, and "d" vs "age" DATA clustering

Fig. 3. ROC curve of fractional differencing parameter d and patient age.

of benign cases. The area under ROC curve was calculated as 0.8334 with a standard deviation of 0.0921.

V. CONCLUSION

In this paper, we modeled the RF echo as a FARMA process. We showed that the FARMA model parameters can be used to discriminate between benign and malignant tumors. The ROC analysis based on 60 *in vivo* B-scan images shows that the proposed method yields results that are comparable to most methods reported in the literature. In our future work, we will include morphological features such as tumor size and smoothness to improve our ROC results.

ACKNOWLEDGMENT

REFERENCES

[1] S. Gefen, O. J. Tretiak, C. W. Piccoli, K. D. Donohue, A. P. Petropulu, P. M. Shankar, V. A. Dumane, L. Huang, M. A. Kutay, V. Genis, F. Forsberg, J. M. Reid ,"ROC Analysis of Ultrasound Tissue Characterization Classifiers for Breast Cancer Diagnosis ," *IEEE Trans. Medical Imaging*, V22, January 2003.

[2] M. A. Kutay, A. P. Petropulu, and W. Piccoli, "On modeling a biomedical ultrasound RF echoes using a power-law shot-noise model," *IEEE Trans. on Ultrason., Ferroelect., Freq. Contr.*, Vol. 48, No.4, pp.953-968, July 2001.

[3] M. Karaoguz, N. Bilgutay, and B. Onaral, "Modeling of scattering dominated ultrasonic attenuation using power-law function," *Ultrasonics Symposium*, 2000 IEEE , Volume: 1, pp. 793 -796, 2000

[4] R.L. Kashyap, K. B. Eom ,"Texture boundary detection based on the long correlation model," *IEEE Trans. Pattern. Anal. Machine Intell.*, PAMI V11, January 1989, pp 58-67.

[5] J.R.M. Hosking, "Fractional differencing," *Biometrika*, V68, 1981, pp165-176.

[6] B. Yazici, and R. L. Kashyap, "Signal modeling and parameter estimation for 1/f processes using scale stationary models," *Proc. of ICASSP'96*, pp.2841-2844, 1996.

[7] B. B. Mandelbrot and I. W. Van Ness, "Fractional Brownian motions, fractional noises, and applications," Soc. for Ind. and Appl. Math. (SIAM) Review, vol. 10, pp. 422-437, 1968

[8] R.L. Kashyap and K. B. Eom, *Estimation in Long-memory Time Series Model*, Journal of Time Series Analysis, Vol. 9, No 1, June 1987, pp 35-41.

[9] B. Alacam, B. Yazici, N. Bilgutay, " Breast tissue characterization based on fractional differencing model of ultrasonic RF echo ," *Proc. of Medical Imaging 2003: Ultrasonic Imaging and Signal Processing*, February, 2003

tumor sample. For each B-scan image, mean and variance of the fractional differencing parameter d were calculated for classification purposes. In Figure 1, we present the mean values of the parameter d obtained from inside and outside the suspected tumor for randomly selected 30 patients.

IV. CLASSIFICATION OF B-SCAN IMAGES FOR TISSUE CHARACTERIZATION

In [9] area under the ROC was calculated as 0.79 with classification using only the fractional differencing parameter d. Classification quality can be improved by including patient age information as a second feature. Assuming that patient age and d parameter are independent and Gaussian distributed, a quadratic classifier can be used to classify benign and malignant tissues. 15 objects both from malignant and benign classes were chosen randomly as the training set and classification discriminant was computed for the quadratic classifier. In Figure 2, results for the classification are displayed and plotted. The classes were plotted in a scatter diagram together with the discriminant and the curve for the quadratic classifier. The ROC curve, given in Figure 3, was obtained by plotting the probability of false alarm versus the probability of detection. Probability of detection is the ratio between the number of correct malignant decisions and total number of malignant cases. The probability of false alarm is the ratio betwe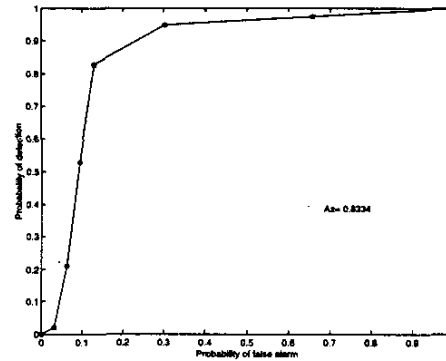en the number of incorrect benign decisions and total number