Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Detection of Prostate Cancer from RF Ultrasound Echo Signals Using Fractal Analysis

Mehdi Moradi, Purang Abolmaesumi, Phillip A. Isotalo,
David R. Siemens, Eric E. Sauerbrei, Parvin Mousavi*, *IEEE Member*

Abstract— **In this paper we propose a new feature, Average Higuchi Dimension of RF Time series (AHDRFT), for detection of prostate cancer using ultrasound data. The proposed feature is extracted from RF echo signals acquired from prostate tissue in an in vitro setting and is used in combination with texture features extracted from the corresponding B-scan images. In a novel approach towards RF data collection, we continuously recorded backscattered echoes from the prostate tissue to acquire time series of the RF signals. We also collected B-scan images and performed a detailed histopathologic analysis on the tissue. To compute AHDRFT, the Higuchi fractal dimensions of the RF time series were averaged over a region of interest. AHDRFT and texture features extracted from corresponding B-scan images were used to classify regions of interest, as small as $0.028cm^2$, of the prostate tissue in cancerous and normal classes. We validated the results based on our histopathologic maps. A combination of image statistical moments and features extracted from co-occurrence matrices of the B-scan images resulted in classification accuracy of around 87%. When AHDRFT was added to the feature vectors, the classification accuracy was consistently over 95% with best results of over 99% accuracy. Our results show that the RF time series backscattered from prostate tissues contain information that can be used for detection of prostate cancer.**

I. INTRODUCTION

As the most common malignancy among men, prostate cancer (PCa) is expected to affect 254,960 men in North America and kill about 31,650 in 2006 [2], [6]. Due to the inconsistent appearance of prostate tumors in medical images, the process of screening and diagnosis of the disease is controversial [5]. For example, on Transrectal Ultrasound (TRUS) which is the standard imaging modality to study PCa, cancer lesions can be hypoechoic, hyperechoic or even isoechoic. The presence of Benign Prostatic Hyperplasia (BPH) further complicates the inspection of ultrasound images, as BPH is typically associated with hyperplastic nodules that may mimic the appearance of malignant areas. The gold standard for detection of PCa is pathological analysis of tissue samples acquired through TRUS guided biopsy. However, the multi-focal nature of the disease and the limited number of biopsies cause high rates of false negative diagnoses.

Tissue characterization based on the acoustic parameters extracted from the raw RF ultrasound echo signals (before being transformed to B-scan images) has been studied since the early 70's (see [9] for a review). Frequency-dependent nature of ultrasound scattering and attenuation phenomena can characterize different tissue types and is studied through frequency spectrum of RF signals. Along with texture and co-occurrence based features extracted from B-scan images, RF spectrum parameters have been used to form hybrid feature vectors for detection of PCa. Such features are utilized as the input to neural networks and neuro-fuzzy inference systems [11], self organizing Kohonen maps [13], and quadratic Bayes classifiers [12] for characterization of prostate tissue. Nevertheless, despite the long history of studies in this field the results of RF-based tissue classification methods are not promising enough for clinical applications.

This fact encouraged us to investigate a new approach to feature extraction from RF signals for detection of PCa. It is known that backscattered ultrasound signal is affected by the geometry and spatial distribution of scatterers. In the case of prostate cancer, the progression of the malignancy is associated with geometrical deregulation of the architectural structure of the cellular network (this is in fact the basis for pathologic indices used for detecting and grading of the disease). This structural change can potentially affect the ultrasound signals backscattered from the tissue. Due to the self-replicating nature of the cellular structure of the tissue [10], we hypothesized that the interactions of tissue and ultrasound follow a chaotic dynamic pattern and therefore fractal analysis of the backscattered signals can reveal the effects of cancer on tissue.

Our proposed approach for fractal analysis of the interaction of the ultrasound and tissue is to continuously apply ultrasound beams and analyze the time series of RF signals backscattered from the tissue. If the interactions follow a chaotic dynamic pattern, the time series formed by each sample of backscattered signal should have a meaningful fractal dimension. This fractal dimension can be used to distinguish between cancerous and normal tissue. To perform the fractal analysis on RF signals, we acquired continuous RF data frames from the prostate tissue of patients undergone radical prostatectomy, and extracted the Higuchi fractal dimensions of the time series of RF signals. We used the average of this parameter on ROIs of size $0.028cm^2$ as a feature for tissue classification. We used neural networks to classify the

* Corresponding author.
M. Moradi is with the School of Computing, Queen's University, Kingston, ON, K7L 3N6, Canada (moradi@cs.queensu.ca).
P. Abolmaesumi is with the School of Computing and the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, K7L 3N6, Canada (purang@cs.queensu.ca).
P. A. Isotalo is with the Department of Pathology and Molecular Medicine, Queen's University, Kingston, ON, K7L 3N6, Canada.
D. R. Siemens is with the Department of Urology, Queen's University, Kingston, ON, K7L 3N6, Canada.
E. E. Sauerbrei is with the Department of Diagnostic Radiology, Queen's University, Kingston, ON, K7L 3N6, Canada.
P. Mousavi is with the School of Computing, Queen's University, Kingston, ON, K7L 3N6, Canada (pmousavi@cs.queensu.ca).

1-4244-0033-3/06/$20.00 ©2006 IEEE.

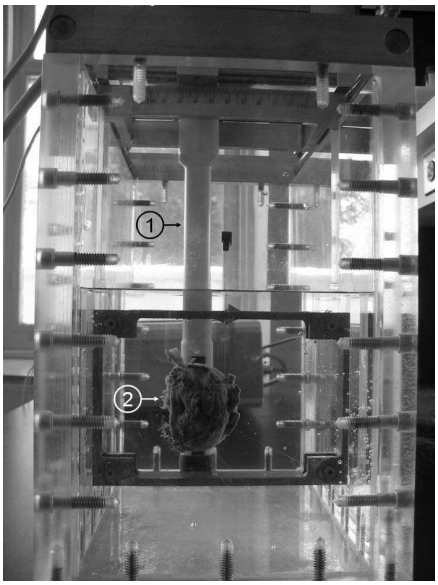

Fig. 1. RF signal and B-scan acquisition setup, the probe (marked with 1 in the image) and prostate tissue (marked with 2 in the image) are fixed in position for continuous acquisition of the RF time series.

ROIs and observed that when the Higuchi fractal dimensions of the RF time series is added to a combination of B-scan based texture and co-occurrence features, the accuracy of neural network based classification of prostate tissue increases considerably. We conclude that the RF time series backscattered from prostate tissues contains information that can be used for detection of prostate cancer.

This paper is organized as follows: in Section II we introduce our data collection and feature extraction techniques, in Section III we present our neural-network-based classification procedure and the results. Section IV concludes the paper and includes interpretation of the results.

II. Methods

A. Data

For ultrasound data collection, we used a Sonix RP (Ultrasonix Inc., Vancouver, Canada) ultrasound machine which has the capability of collecting and recording raw RF signals. Figure 1 demonstrates our data collection setup; the probe[1] was mounted on a rail which could be moved along the prostate while the tissue was fixed in a frame and immersed in water. We collected B-scan images and time series of RF signals from parallel cross sections of tissue each one $1mm$ apart. We ensured that the orientation of these cross sections was as close as possible to the orientation of the slices to be made for specimen histopathologic analysis. To mark the position of the first cross section along the prostate we placed an ultrasound visible needle inside the tissue (Figure 2). This needle was used to determine the starting point for acquiring tissue slices for histopathologic analysis.

To form the RF time series, we continually acquired 63 frames of RF data, at the rate of 8 frames per second,

[1]Probe model: BPSL9-5/55/10, frequency range: 5-9MHz, set to 6.6MHz for our experiments. We used the linear transducer on this probe.

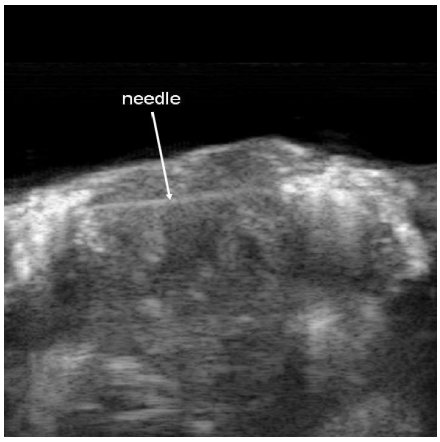

Fig. 2. The first imaging position is marked with a needle.

from each cross section of tissue. We also recorded the B-scan image of the cross section. In our settings, each RF frame consisted of 256 lines of RF signal each with 2064 samples. We defined regions of interest (ROIs) as small as $0.028cm^2$ in the tissue which corresponds to a window of size 16×16 pixels in the B-scan ultrasound and a window of size 8×88 samples in the RF data. After acquisition of ultrasound data, a detailed histopathologic analysis was performed on tissue slices each 5mm apart. Multifocal prostatic carcinoma was confirmed histologically in each prostate examined. Malignancy maps were produced for each slice and were used as the gold standard for validation in this study. We have acquired the ultrasound data and histopathologic analysis from the prostate of four patients from which two were utilized in the current study. For the other two patients the areas of malignancy were too small and the process of registering the histopathologic results to ultrasound data could not be performed reliably.

B. Features

The first-order statistical moments [11], [8] such as mean, standard deviation, skewness and kurtosis and features extracted from the gray level co-occurrence matrices of B-scan images [11], [4] have been reportedly used for classification of Regions of Interest (ROIs) in prostate ultrasound images. In our study, each ROI is described with 13 features:

- Four statistical moments of the pixel intensities in the B-scan image (mean, standard deviation, skewness and kurtosis),
- Eight features extracted from the gray level co-occurrence matrices of each ROI.
- AHDRFT: Average Higuchi Dimension of RF Time series in the ROI.

Higuchi fractal dimension (AHDRFT): The self-replicating properties of chaotic systems result in a fractal pattern in their output when recorded as a time series. Higuchi has proposed a stable methodology to compute the fractal dimension of the irregular output time series of natural phenomena [7]. Higuchi dimension has reportedly been applied to mental task classification and sleep stage scoring based on electroencephalography (EEG) signals [1],

[3] and to oceanographic data [15]. As a novel approach towards tissue classification, we used this methodology to extract the fractal dimension of ultrasound RF time series acquired from prostate tissue.

Consider N frames of RF data acquired at a regular rate while the probe and the tissue are fixed in position. Each sample of the RF data forms a time series $X(1),X(2),...,X(N)$. From this time series, we first construct k new time series X_k^m as follows:
$X_k^m : X(m),X(m+k),X(m+2k)...,X(m+[\frac{N-m}{k}].k)$, where $k < N$, $m = 1,2,...,k-1$ and both are integers. The length of each time series is defined as:

$$L_m(k) = 1/k \times (\frac{N-1}{[\frac{N-m}{k}].k}) \times \sum_{i=1}^{[\frac{N-m}{k}]} |X(m+ik) - X(m+(i-1).k)| \quad (1)$$

The average value of $L_m(k)$ over k sets (denoted with $< L(k) >$) is the so-called length of the time series. If the condition $< L(k) > \propto k^{-d}$ holds, the time series is fractal with the dimension d. In other words, to compute d, a line is fitted to values of $\ln(L_m(k))$ versus $\ln(1/k)$ and the slope of this line is considered as the Higuchi fractal dimension of the time series. In our implementation, $k = 16$ and $N = 64$ (the 63 point time series was augmented with one sample equal to the last sample to increase the computational efficiency). We defined AHDRFT as the average of Higuchi dimensions computed for all the RF samples in the corresponding RF window of an ROI.

Features extracted from the gray level co-occurrence matrices: Co-occurrence matrices of an image are histogram representations, each describing the co-occurrence of a pair of intensities at a certain distance (l) along a certain direction (θ). For feature extraction from gray-scale images, usually only the co-occurrence matrices corresponding to directions $\theta = 0$ and $\theta = 90$ are used [4], [16]. Features extracted from these matrices statistically describe the relationships between pixels in the image (where as the statistical moments only depend on individual intensities).

In this work, we used eight bins of gray levels, co-occurrence distance of $l = 1$ and directions $\theta = 0$ and $\theta = 90$ (vertical and horizontal adjacency) to extract two 8×8 co-occurrence matrices from each ROI. More precisely, the values of these two matrices were described by:

$$\theta = 0^\circ \quad : \quad p(I(m,n) = i, I(m \pm 1, n) = j) = \frac{no.\ of\ H.\ adjacent\ pixel\ pairs\ with\ values\ (i,j)}{total\ number\ of\ possible\ pairs} \quad (2)$$

and,

$$\theta = 90^\circ \quad : \quad p(I(m,n) = i, I(m, n \pm 1) = j) = \frac{no.\ of\ V.\ adjacent\ pixel\ pairs\ with\ values\ (i,j)}{total\ number\ of\ possible\ pairs} \quad (3)$$

(H and V stand for Horizontally and Vertically). The following four features were extracted from each co-occurrence matrix, (see [14] for more details):

Contrast (CON0, CON90): A measure of the intensity contrast between a pixel and its neighbors in the image:

$$CON = \sum_i \sum_j p(i,j) \times |i-j|^2 \quad (4)$$

Correlation (COR0, COR90): A measure of correlation between the intensity of a pixel and its neighbor:

$$COR = \frac{(\sum_i \sum_j (ij) p(i,j)) - \mu_i \mu_j}{\sigma_i \sigma_j} \quad (5)$$

where μ_i and σ_i are the mean and standard deviation of row i in the co-occurrence matrix.

Energy (ENE0, ENE90): A measure of the smoothness of the image, also known as the angular second moment:

$$ENE = \sum_i \sum_j p(i,j)^2 \quad (6)$$

Entropy (ENT0, ENT90): A measure of randomness in the image which takes low values for smooth images:

$$ENT = -\sum_i \sum_j p(i,j) \log_2 p(i,j) \quad (7)$$

C. Classification

Different combinations of the previously described features were extracted from ROIs and used with Artificial Neural Networks (ANNs) to classify ROIs into cancerous and normal classes. We employed multi-layer perceptron neural networks trained with back propagation algorithm. We found a multi-layer feedforward network with two hidden layers and ten neurons in each hidden layer to be the optimal network for our problem. Networks trained on a set of ROIs were tested on separate unseen ROIs and the results were validated based on the histopathologic analysis of the tissue. We designed two sets of experiments: In one set of experiments, data acquired from one patient was used for training and testing the networks. In the second set of experiments, networks were trained on ROIs acquired from one patient and tested on another one.

III. Results

Training and testing on one patient: Based on the histopathology results, we selected 213 normal ROIs and 185 cancerous ROIs in 20 different frames of ultrasound data acquired from a 57 year old patient. The cancerous samples were from two different lesions (one hypoechoic and one isoechoic) in two different regions of the prostate. Two thirds of the data samples were randomly selected for training and the rest were used for testing the trained network. Table 1 summarizes the results. Accuracy, sensitivity and specificity values reported in each row are the average of six rounds of training and testing experiments, each time with a new randomly selected training and testing set. As shown in table 1, after including AHDRFT in the feature vector, a significant

TABLE I

THE PERFORMANCE OF DIFFERENT GROUPS OF FEATURES. ADDITION OF AHDRFT HAS CONSIDERABLY INCREASED THE ACCURACY.

Features	Accuracy	Sensitivity	Specificity
mean, std, kurtosis, skewness	87	86.1	84.5
mean, std, kurtosis, skewness, **AHDRFT**	93.9	94.5	93.4
mean, std, ku, sk, CON0, CON90, COR0, COR90, ENE0, ENE90, ENT0, ENT90	87.1	88.8	85.7
mean, std, ku, sk, CON0, CON90, COR0, COR90, ENE0, ENE90, ENT0, ENT90, **AHDRFT**	95.4	95.5	96.1

increase in the classification accuracy is witnessed (compare row 1 with row 2 and row 3 with row 4). It was observed that the co-occurrence features did not change the classification accuracy (compare row 1 with row 3). The overall best results were acquired by using the feature vector that contained all thirteen features (Our best recorded result shows around 99% accuracy). In an effort for further statistical validation of the results, we performed 100 rounds of training and testing each time with a random split of training and testing sets with the 13 dimensional feature vector and witnessed an average accuracy of 93.1%.

Two patient cross validation experiments: For validation purposes, we repeated the experiments this time with training and testing samples each acquired from a different patient. The ANN was trained with the data used in the previous experiment. For testing, we used 50 cancerous ROIs extracted from two different cancerous regions and 54 ROIs extracted from the non-cancerous areas in the prostate of a 54 year old patient. The sensitivity of the tests was constantly equal to or higher than 96%. However, the specificity dropped to almost 70%. In other words, although almost all areas of cancer were identified, a high rate of false positives was witnessed. We observed that relatively large lesions of BPH were present in the areas labeled as normal in data acquired from the second patient. Although BPH tissue is not cancerous, it shows a different cellular structure. The high rate of false positive detections is likely due to the fact that networks were not trained on such benign ROIs.

IV. DISCUSSION AND CONCLUSIONS

We proposed a novel feature, Average Higuchi Dimension of RF Time series (AHDRFT), for prostate cancer detection from ultrasound RF signals and compared its performance with statistical moments and co-occurrence based image texture features. A combination of image statistical moments and features extracted from co-occurrence matrices of the B-scan images resulted in ANN-based classification accuracy of around 87%. When AHDRFT was added to the feature vectors, the accuracy was consistently over 95% with best results of over 99%. Furthermore, the neural networks trained with our proposed set of features on data acquired from one patient were almost perfectly capable of detecting all cases of cancer in the data acquired from another patients. The detection of normal tissues on the second patient resulted in a lower accuracy. Our investigation showed that this phenomena is due to the presence of areas with BPH in the prostate tissue which was not considered as a separate class in our training data.

Results of this study show that AHDRFT is a helpful feature for detection of prostate cancer (although as an independent feature AHDRFT fails to effectively discriminate prostate tumors from normal tissue, it increases the classification performance of texture features considerably). Due to the computationally expensive nature of the process of extracting AHDRFT, in future implementations of this system, parallel processing should be taken into consideration. Further validation of the proposed methodology will require larger datasets to challenge the separability capability of the features on different varieties of prostate tissues. In our future work we will extract other fractal and non-fractal based features from the ultrasound RF time series and use them with different classification methodologies to maximize the detection rate of prostate cancer.

REFERENCES

[1] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet, "Use of the fractal dimension for the analysis of electroencephalographic time series," *Biological Cybernetics*, vol. 77, no. 5, pp. 339–350, 1997.

[2] American Cancer Society. [Online]. Available: http://www.cancer.org.

[3] A. Anier, T. Lipping, S. Melto, and S. Hovilehto, "Higuchi fractal dimension and spectral entropy as measures of depth of sedation in intensive care unit," in *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, Sept. 2004, pp. 526–529.

[4] O. Basset, Z. Sun, J. L. Mestas, and G. Gimenez, "Texture analysis of ultrasonic images of the prostate by means of cooccurrence matrices," *Ultrasonic Imaging*, vol. 15, no. 3, pp. 218–237, 1993.

[5] P. Boyle, "Prostate specific antigen (PSA) testing as screening for prostate cancer: The current controversy," *Annals of Oncology*, vol. 9, no. 12, pp. 1263–1264, 1998.

[6] Canadian Cancer Society. [Online]. Available: http://www.cancer.ca.

[7] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277–283, 1988.

[8] A. G. Houston, S. B. Premkumar, D. E. Pitts, and R. J. Babaian, "Prostate ultrasound image analysis: localization of cancer lesions to assist biopsy," in *Proceedings of the Eighth IEEE Symposium on Computer-Based Medical Systems*, June 1995, pp. 94–101.

[9] F. L. Lizzi, M. Greenebaum, E. J. Feleppa, M. Elbaum, and D. J. Coleman, "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," *Journal of the Acoustic Society of America*, vol. 73, no. 3, pp. 1366–1373, 1983.

[10] F. Pansera, "Fractals and cancer," *Med. Hypo.*, vol. 42, p. 400, 1994.

[11] U. Scheipers, H. Ermert, H. J. S. M. Garcia-Schurmann, T. Senge, and S. Philippou, "Ultrasonic multifeature tissue characterization for prostate diagnosis," *Ultrasound in Medicine and Biology*, vol. 20, no. 8, pp. 1137–1149, 2003.

[12] G. Schmitz and H. Ermert, "Tissue-characterization of the prostate using radio frequency ultrasonic signals," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 46, no. 1, pp. 126–138, 1999.

[13] G. Schmitz, H. Ermert, and T. Senge, "Tissue characterization of the prostate using Kohonen-maps," in *Ultrasonics Symposium*, 1994, pp. 1487–1490.

[14] S. Theodoridis and L. koutroumbas, *Pattern Recognition*. San Diego, CA: Academic Press, 1999.

[15] A. Tsuda, "Fractal distribution of an oceanic copepod neocalanus cristatus in the subarctic pacific," *Journal of Oceanography*, vol. 51, pp. 261–266, 1995.

[16] M. J. F. Valckx and M. J. Thijssen, "Characterization of echocardiographic image texture by coocurence matrix parameters," *Ultrasound in Medicine and Biology*, vol. 23, no. 4, pp. 559–571, 1997.