Tissue classification using depth-dependent ultrasound time series analysis: In vitro animal study

Farhad Imani[a], Mohammad Daoud[c], Mehdi Moradi[c], Purang Abolmaesumi[c], Parvin Mousavi[a,b]

[a]Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada;
[b]School of Computing, Queen's University, Kingston, ON, Canada;
[c]Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada

ABSTRACT

Time series analysis of ultrasound radio-frequency (RF) signals has been shown to be an effective tissue classification method. Previous studies of this method for tissue differentiation at high and clinical-frequencies have been reported. In this paper, analysis of RF time series is extended to improve tissue classification at the clinical frequencies by including novel features extracted from the time series spectrum. The primary feature examined is the Mean Central Frequency (MCF) computed for regions of interest (ROIs) in the tissue extending along the axial axis of the transducer. In addition, the intercept and slope of a line fitted to the MCF-values of the RF time series as a function of depth have been included. To evaluate the accuracy of the new features, an *in vitro* animal study is performed using three tissue types: bovine muscle, bovine liver, and chicken breast, where perfect two-way classification is achieved. The results show statistically significant improvements over the classification accuracies with previously reported features.

Keywords: RF time series, ultrasound RF data, tissue classification, mean central frequency

1. INTRODUCTION

Tissue classification using ultrasound imaging has been an active field of study in the last three decades[1,2]. A recent approach proposed for distinguishing various tissue types using both high-frequency (20-60 MHz) and clinical-frequency (2-10 MHz) ultrasound scanners, is by acquiring radio-frequency (RF) time series data[3,4]. To acquire RF time series, the transducer is used to acquire a few hundred frames from a fixed region of interest inside the tissue. Previous studies have shown that features extracted from spectral analysis of the time series of the backscattered RF signals are indicative of anatomical variations between different tissue types[3,4,7].

Moradi et al. used RF echo time series at both high-frequency and clinical-frequency to distinguish between different animal tissue types[3,4]. Classification accuracies of 68% to 96% are obtained between various types of animal tissue by using spectral features of the RF time series at clinical frequencies[7]. It has been shown that the spectral features extracted from the RF time series are sensitive to the anatomical variations between pig liver, bovine liver, chicken breast, and bovine muscle. Features of RF time series that are used in the analysis include the integral of the power spectrum, the intercept and the slope of the regression line fitted to the spectrum, and the Higuchi fractal dimension of the time series. In a previous *ex vivo* human study[4], the time series features are also applied at clinical frequencies to differentiate between healthy and cancerous tissue in 35 human prostate specimens. The study shows that the time series features are significantly more sensitive and accurate compared to the texture-based and spectral features proposed in the literature[10,11] for detecting cancerous tissue. In order to improve the classification accuracy between healthy and cancerous specimens of the prostate, Aboofazeli et al. used wavelet transform approximation and detail sequences of RF time series of prostate tissue as a feature set[9]. The results show that wavelet features probably cannot be recognized as an alternative feature set for RF time series features.

In this paper, the classification accuracy of the time series method at clinical frequencies has been significantly improved by incorporating new features extracted from the RF time series. The features include the mean central frequency of the RF time series. Moreover, the variations of the new features as a function of depth along the axial direction of the

transducer are studied, and applied to improve tissue classification accuracy. The tissue classification capabilities of the proposed features are investigated to classify three animal tissue types. The results reported in this paper show that the proposed features lead to substantial and statistically significant improvement in classification accuracy.

2. METHODS

2.1 Data Collection

A Sonix MDP ultrasound scanner (Ultrasonix Inc., Richmond, BC, Canada) is used to collect RF time series signals from three types of tissue: bovine muscle, bovine liver, and chicken breast. The transducer used in imaging is an Ultrasonix L14-5/38 probe with a central frequency of 7.2 MHz. The scanner is configured to acquire 1024 frames composed of 5 scan lines from each tissue type extending 30 mm along the axial dimension of the transducer with a focal length of 20 mm. Each scan line is the average of 20 adjacent acquired scan lines. Imaging of each location in the tissue is repeated over 5 trials. The rate of ultrasound acquisition is 500 frames per second. A water-like coupling medium of thickness 10 mm is placed between the surface of the transducer and the tissue specimen (Fig. 1). A set of 41 equally-spaced ROIs are selected in the tissue that extend axially between 2 mm before and after the focal point. The size of each ROI in depth is 5 samples, and a total of 205 ROIs for each tissue type is selected. The RF signals received from the tissue are sampled using an analog to digital converter with a sampling frequency of 40 MHz.
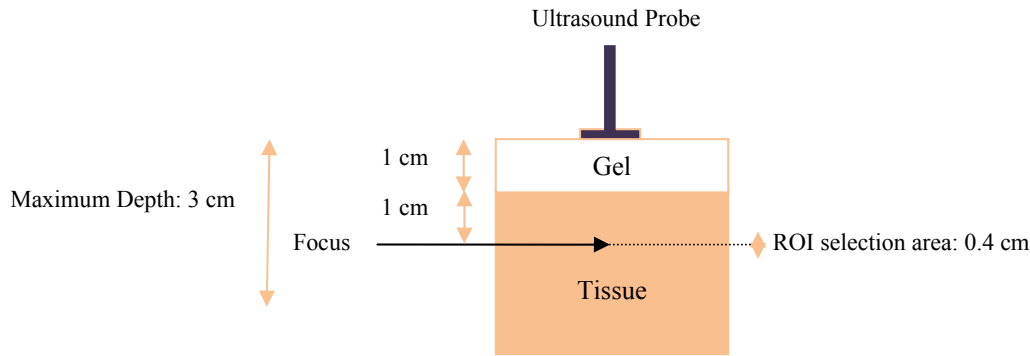

Fig. 1. Data collection setup.

2.2 Feature Extraction

MCF of the RF time series is calculated and used to extract new tissue typing features. MCF is defined as the mean of the bandwidth of the power spectrum of an RF signal, and has been used in the literature to characterize tissue types based on their attenuation coefficients[5]. In this paper, the MCF is calculated for the time series of the RF signals received from a specific ROI in the tissue. The value of the MCF for an ROI is taken as the mean central frequency of the power spectrum computed for the time series of the signal acquired from that ROI, as expresses below:

$$MCF = \frac{\sum_{BW} f.PSD(f)}{\sum_{BW} PSD(f)} \qquad (1)$$

where the PSD is the average value of the power spectrum densities computed for the RF time series located in an ROI as a function of the frequency of the time series, f is the frequency of the RF time series, and BW is the bandwidth of the RF time series spectrum. MCF values computed for the individual ROIs are analyzed using linear regression to fit a line to the experimental MCF values as a function of depth. Four features, named F1, F2, F3, and F4, are extracted from the data and fitted to the values. These features are summarized in Table 1.

Table 1. The RF time series features proposed to improve tissue classification accuracy

Feature	Description
F1	The MCF of an individual ROI in the tissue
F2	The slope of regression line fitted to experimental MCF values as a function of depth along the axial direction. This feature represents the change rate of MCF in depth.
F3	The intercept of the regression line fitted to the MCF values as a function of depth.
F4	The mean of the MCF values computed for the ROIs.

In addition to the features presented in this study, the time series spectral features proposed previously[3] are used in the analysis. These spectral features are: the sum of the power spectrum computed over the four quarters of the frequency spectrum of the time series (F5, F6, F7, F8), the intercept (F10) and slope (F11) of the regression line fitted to the power spectrum of the time series, and the fractal dimension of the time series (F12) calculated using Higuchi's algorithm. Moreover, the sum of the power spectrum computed over the entire spectrum of the time series (F9) is calculated as another time series feature.

In order to calculate the sum of the power spectrum over different frequency ranges (F5, F6, F7, F8, and F9), zero-mean Discrete Fourier Transform (DFT) of the time series of each point in an ROI is computed *(i.e. X(k) where $k = 1, .., N/2$)*. To reduce computational complexity, only the positive frequency components are used in the analysis. Such a simplification should not affect the results due to symmetry of the spectrum. Each feature is normalized within the selected ROI as below:

$$F5 = \sum_{k=1}^{N/8} |\hat{X}_{ROI}[k]|, \ F6 = \sum_{k=N/8}^{N/4} |\hat{X}_{ROI}[k]|, \ F7 = \sum_{k=N/4}^{3N/8} |\hat{X}_{ROI}[k]|, \ F8 = \sum_{k=3N/8}^{N/2} |\hat{X}_{ROI}[k]| \quad (2)$$

$$F9 = \sum_{k=1}^{N/2} |\hat{X}_{ROI}[k]| \quad (3)$$

Fractal dimension (F12) is computed using Higuchi's algorithm. In order to do this, first k new time series for N frames are constructed as below:

$$X_k^m = (X(m), X(m+k), X(m+2k), .., X\left(m + \left[\frac{N-m}{k}\right].k\right)\}; \ m = 1,2, ..., k-1 \quad (4)$$

Then, the length of each time series (i.e. $L_m(k)$) is as follows:

$$L_m(k) = \frac{1}{k} \times \left(\frac{N-1}{\left[\frac{N-m}{k}\right].k}\right) \times \sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X(m+ik) - X(m+(i-1)k)| \quad (5)$$

where $L(k)$ is the length of time series and can be calculated as the average of $L_m(k)$ over k. The procedure is repeated from 1 to k_{max}. Fractal dimension is the slope of fitted line to $ln(L(k))$ as a function of $ln(1/k)$.

2.3 Classification

A support vector machines (SVM) classifier[6] is used to classify various animal tissue types. Two- and three-way classifications are carried out between the three tissue types. The SVM classifier solves the following optimizing problem where the parameters of the optimal hyper plane with maximum margin (weight vector W, and margin b) are found to separate the data into different classes.

$$\min_{W,b,\zeta} \left(\frac{W^T W}{2} + C \sum_{i=1}^{l} \zeta_i \right) \qquad (6)$$

which is subject to:

$$y_i(W^T \phi(x_i) + b) > 1 - \zeta_i \ ; \ \zeta_i \geq 0 \qquad (7)$$

where ζ_i is the slack variable and indicates the error that is made for each data point, and $C > 0$ is the coefficient of the error term. x_i are training vectors, and function ϕ maps x_i to a higher dimension space to achieve maximum separation. For two- and three-way classifications, class labels, y_i, are selected from $\{-1,1\}$, and $\{1,2,3\}$, respectively. $K(x_i, x_j) = \phi^T(x_i) \phi(x_j)$ is the kernel function; we tried various kernel functions to map the training vectors to higher dimensional space. Radial basis function outperforms other functions in terms of accuracy and initialization.

The parameters of the classifier are tuned through an exhaustive search to maximize the classification accuracy. The classifier is trained using the features of 90% of the ROIs, while the rest of the data is used for testing the classification performance. Data division process into training and testing portions is repeated for 500 trials, to ensure there is no bias towards the division of the data. The mean and standard deviation of the accuracy of the test data in 500 trials is reported. We report classification accuracies using single features, and using the four proposed depth-dependent features in this paper. We also exhaustively search for the best classification accuracies using a combination of two to four of these features.

3. RESULTS

Two- and three-way classification accuracies obtained using the individual depth-dependent time series features proposed in this paper are shown in Table 2. The highest classification accuracy with individual features between bovine muscle and bovine liver is achieved using the intercept of the fitted line to the MCF values, F3. When the classifier is used to distinguish between the bovine muscle and chicken breast, the highest classification accuracy (100%) is obtained using the intercept of the fitted line to the MCF values. The mean of the MCF, F4, yields the highest classification accuracy between bovine liver and chicken breast. It is worth mentioning that the classification accuracy between bovine muscle and bovine liver obtained using a single feature is lower than that of the other pairs of tissue. When the classifier is used to perform three-way classification between the three tissue types, the F3 feature yielded the highest accuracy as shown in Table 2.

Table 2. Mean and standard deviation of classification accuracies obtained by configuring the classifier to use one of the time series features proposed in this study

Tissue types	F1	F2	F3	F4
Bovine Muscle and Bovine Liver	53.47±1.03%	81.80±0.27%	82.16±1.10%	77.19±1.52%
Bovine Muscle and Chicken Breast	61.74±0.84%	84.59±1.51%	100±0%	79.47±0.27%
Bovine Liver and Chicken Breast	67.53±0.51%	54.64±0.41%	81.60±0.26%	90.64±0.22%
Bovine Muscle and Bovine Liver and Chicken Breast	48.72±0.55%	72.16±1.48%	73.69±1.56%	68.54±0.81%

Table 3. Mean and standard deviation of two-way classification accuracies obtained using both the feature combinations of new features and feature combinations that incorporate previously-proposed features.

Tissue types	Combination of previously proposed features (classification accuracy)	Combination of new features (classification accuracy)
Bovine Muscle and Bovine Liver	F5, F6, F10, and F12 (95.59±0.42%)	F1, F2, F3, and F4 (99.77±0.04%)
Bovine Muscle and Chicken Breast	F8, F10, F11, and F12 (95.83±0.37%)	F3 (100±0%)
Bovine Liver and Chicken Breast	F5, F6, F7, F11, and F12 (95.50±0.41%)	F3 and F4 (100±0%)

The best combinations of the newly proposed features for tissue classification are presented in Table 3. For comparison, the two-way classification accuracies are computed for the same groups of tissue types using the combination of the previously-proposed features[7] as optimal feature combination for tissue classification. The classification accuracies achieved by the proposed features substantially outperform the accuracies obtained by previously-proposed features. The results also show statistically significant improvement in classification accuracy ($p \cong 0$; two-way t-test analysis using 500 classification accuracies for each test). Moreover, the three-way classification accuracy between the three tissue types obtained using a combination of new features (F2, F3, and F4) is 99.57±0.07%.

Fig. 2 shows the histogram of two-way classification accuracies between bovine muscle and chicken breast, bovine liver and chicken breast, and bovine muscle and bovine liver. Distributions of classification accuracies obtained by combination of new proposed features and combination of previously proposed features are compared for each pair of different tissue types. The figure demonstrates improvement of tissue classification accuracy when applying the proposed features.

4. CONCLUSION

Acquisition and analysis of RF time series has been proposed to improve the sensitivity and accuracy of ultrasound-based tissue typing, compared to conventional methods at both high-frequency and clinical-frequency ultrasound. Improvement of the tissue classification accuracies at clinical frequencies will extend the use of the RF time series method for clinical tissue characterization.

In this paper, a new set of features extracted from the power spectrum of the RF time series is introduced to improve tissue classification. The MCF for the time series of the RF signals contains information that can be used to improve tissue classification accuracy at the clinical frequency band. The proposed features include the slope and intercept of a regression line fitted to experimental MCF values as a function of axial depth, the mean of the MCF values, and the value of the MCF computed for individual ROIs in the tissue. To evaluate the accuracy of the new features, an *in vitro* animal study is performed using three animal tissue types. Combinations of the new features presented in this paper obtain classification accuracies between chicken breast, bovine liver, and bovine muscle that substantially outperform the accuracies obtained by earlier feature combinations suggested in the literature.

The analysis of new proposed features in this paper to distinguish between cancerous and healthy tissue regions, and to grade cancer is a topic of our future studies. In order to achieve this goal, we need to extend the proposed approach to understand its effectiveness in differentiating cancerous and healthy tissue types for clinical purposes. Moreover, since the features presented in this paper are depth dependent, all of the features extracted from a scan line are presumed to have similar information indicating a specific tissue type. In real clinical applications, tissue type in depth could be

variable; therefore, we need to overcome the limitation of selecting appropriate ROIs from each scan line to produce the features that can differentiate between cancerous and healthy tissue regions.
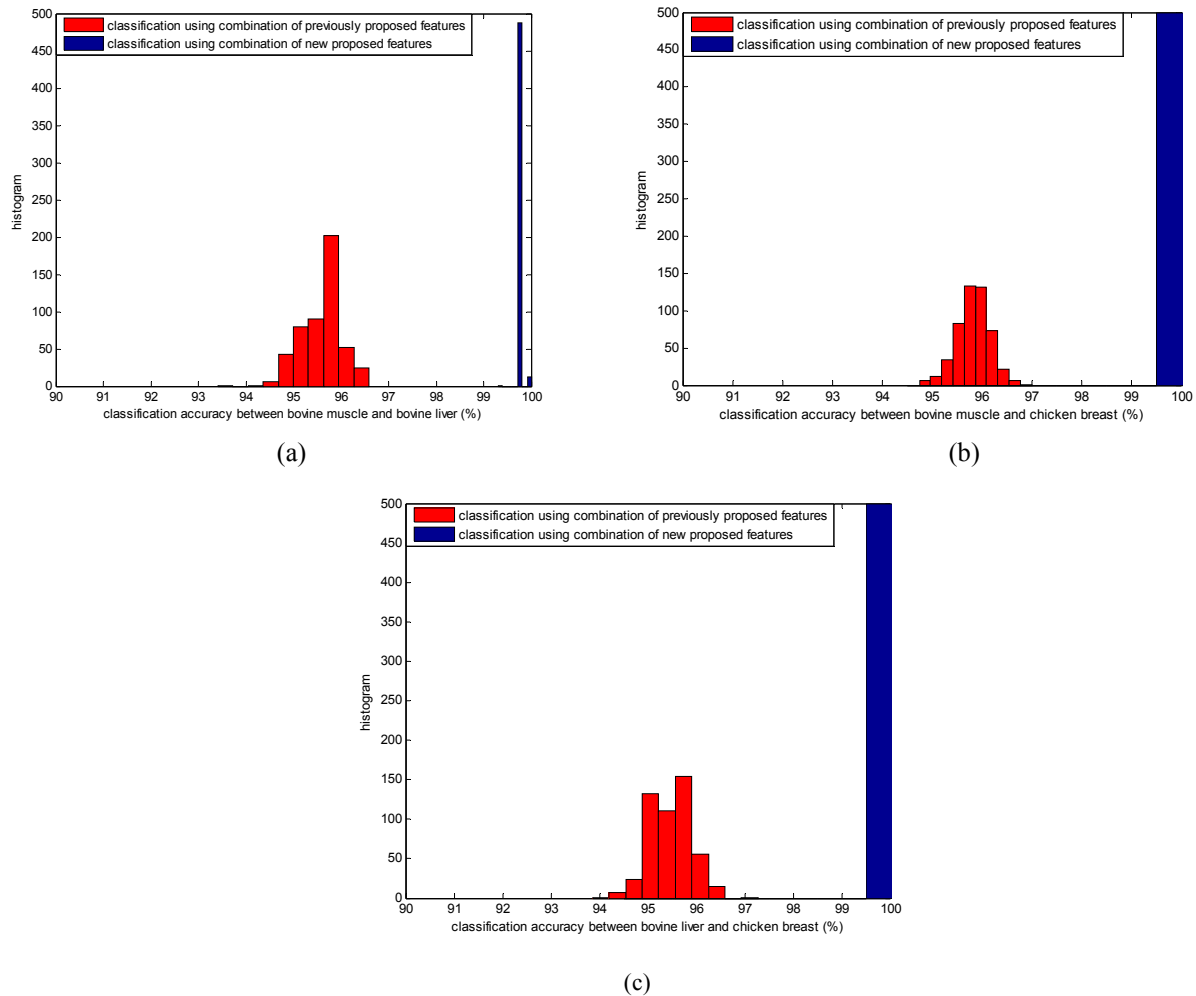

(a)

(b)

(c)

Fig. 2. Histogram of two-way classification accuracies between bovine muscle and chicken breast, bovine liver and chicken breast, and bovine muscle and bovine liver. Figure shows improvement in classification accuracies obtained by combination of new proposed features compared to combination of previously proposed features.

REFERENCES

[1]   Lizzi, F. L., Greenebaum, M., Feleppa, E. J., Elbaum, M. and Coleman, D. J.," Theoretical framework for spectrum analysis in ultrasonic tissue characterization," *Journal of the Acoustical Society of America* 73, 1366-1373 (1983).

[2]   Mamou, J., Coron, A., Hata, M., Machi, J., Yanagihara, E., Laugier, P. and Feleppa, E. J.," Three-dimensional high-frequency characterization of cancerous lymph nodes," *Ultrasound in Medicine and Biology* 36, 361-375 (2010).

[3]   Moradi, M., Abolmaesumi, P. and Mousavi, P.," Tissue typing using ultrasound RF time series: Experiments with animal tissue samples," *Journal of Medical Physics* 37(8), 4401-4413 (2010).

[4] Moradi, M., Abolmaesumi, P., Siemens, D.R., Sauerbrei, E.E., Boag, A. and Mousavi, P.," Augmenting detection of prostate cancer in transrectal ultrasound images using SVM and RF Time Series," *IEEE Transactions on Biomedical Engineering* 56(9), 2214-2223 (2009).

[5] Noble, J.A.," Ultrasound image segmentation and tissue characterization," *Journal of Engineering in Medicine* 224, 307-316 (2010).

[6] http://www.csie.ntu.edu.tw/~cjlin/libsvm

[7] Moradi, M.," A New paradigm for Ultrasound-Based Tissue Typing in Prostate Cancer," PhD Thesis, School of Computing, Queen's University (2008).

[8] Moradi, M., Mousavi, P., and Abolmaesumi, P., " Tissue characterization using fractal dimension of high frequency ultrasound RF time series," *Medical Image Computing and Computer Assisted Intervention* 4792, 900-908 (2007).

[9] Aboofazeli, M., Abolmaesumi, P., Moradi, M., Sauerbrei, E., Siemens, R., Boag, A., and Mousavi, P. ," Automated detection of prostate cancer using wavelet transform features of ultrasound RF time series," *Proc. SPIE* 7260, 72603J-1-8 (2009).

[10] Basset, O., Sun, Z., Mestas, J. L., and Gimenez, G., " Texture analysis of ultrasonic images of the prostate by means of cooccurrence matrices, " *Ultrasonic Imaging* 15(3) , 218–237 (1993).

[11] Feleppa, E. J., Kalisz, A., Sokil-Melgar, J. B., Lizzi, F. L., Liu, T., Rosado, A. L., Shao, M. C., Fair, W. R., Wang, Y., Cookson, M. S., Reuter, V. E., and Heston, W. D. W.," Typing of prostate tissue by ultrasonic spectrum analysis," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 43(4), 609–619 (1996).