Regular Article

Prostate Tissue Characterization/Classification In 144 Patient Population Using Wavelet and Higher Order Spectra features from Transrectal Ultrasound images

Gyan Pareek¹, U Rajendra Acharya², G Swapna³, S Vinitha Sree⁴*, Ratna Yantri⁵, Roshan Joy Martis⁵, Luca Saba⁶, Ganapathy Krishnamurthi⁷, Giorgio Mallarini⁶, Ayman El-Baz⁸, Shadi Al Ekish¹, Michael Beland⁹, Jasjit S. Suri¹⁰ Sr. Member IEEE, Fellow AIMBE

¹Section of Minimally Invasive Urologic Surgery, The Warren Alpert Medical School of Brown University, Providence, RI 02905
²Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore; Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia
³Department of Applied Electronics and Instrumentation, Government Engineering College, Kozhikode, Kerala 673005, India
⁴Visiting Scientist, Global Biomedical Technologies Inc., Roseville, CA, USA
⁵Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore;
⁶Department of Radiology, Azienda Ospedaliero Universitaria di Cagliari, Cagliari, Italy
⁷Mayo Clinic, Rochester, MN, USA
⁸Bioengineering Department, Speed School of Engineering, University of Louisville, Louisville, KY 40292
⁹The Alpert Medical School of Brown University, Director of Ultrasound, Rhode Island Hospital
¹⁰Fellow AIMBE, CTO, Department of Biomedical Engineering, Global Biomedical Technologies, Roseville, CA, USA; Biomedical Engineering Department, Idaho State University (Affl.), ID, USA

*Corresponding Author
Postal Address:
Global Biomedical Technologies, Inc., 208 Otter Glen CT, Roseville, CA, USA, 95661
Telephone: (916) 749-5628;
Email Address: vinitha.sree@gmail.com

Abstract

In this work, we have proposed an on-line computer-aided diagnostic system called “UroImage” that classifies a Transrectal Ultrasound (TRUS) image into cancerous or non-cancerous with the help of non-linear Higher Order Spectra (HOS) features and Discrete Wavelet Transform (DWT) coefficients. The UroImage system consists of an on-line system where five significant features (one DWT-based feature and four HOS-based features) are extracted from the test image. These on-line features are transformed by the classifier parameters obtained using the training dataset to determine the class. We trained and tested six classifiers. The dataset used for evaluation had 144 TRUS images which were split into training and testing sets. Three-fold and ten-fold cross-validation protocols were adopted for training and estimating the accuracy of the classifiers. The ground truth used for training was obtained using the biopsy results. Among the six classifiers, using 10-fold cross-validation technique, Support Vector Machine and Fuzzy Sugeno classifiers presented the best classification accuracy of 97.9% with equally high values for sensitivity, specificity and positive predictive value. Our proposed automated system, which achieved more than 95% values for all the performance measures, can be an adjunct tool to provide an initial diagnosis for the identification of patients with prostate cancer. The technique, however, is limited by the limitations of 2D ultrasound guided biopsy, and we intend to improve our technique by using 3D TRUS images in the future.
Running Title: A data mining technique for prostate tissue characterization/classification

Keywords: Prostate Cancer, Transrectal Ultrasound, Classification, Discrete Wavelet Transform, Higher Order Spectra, Features, Accuracy.

Abbreviations: CAD, Computer Aided Diagnosis; CAP, Prostate Cancer; DRE, Digital Rectal Examination; DT, Decision Tree; DWT, Discrete Wavelet Transform; EM, Expectation Maximization; FN, False Negative; FP, False Positive; FS, Fuzzy Sugeno; GMM, Gaussian Mixture Model; HOS, Higher Order Spectra; KNN, K-Nearest Neighbor; MRI, Magnetic Resonance Imaging; PPV, Positive Predictive Value; PSA, Prostate-Specific Antigen; RBPNN, Radial Basis Probabilistic Neural Network; RF, Radio Frequency; ROI, Region of Interest; SVM, Support Vector Machine; TN, True Negative; TP, True Positive; TRUS, Transrectal Ultrasound; WCOB, Weighted Centre of Bispectrum
Introduction

Prostate cancer (CaP) is due to the abnormal and uncontrolled cell mutation and replication in the prostate gland. CaP is the most prevalent malignancy in men and it has been found that one in six men have a lifetime risk of a CaP diagnosis and a 3.4% chance of death due to CaP (1). In the USA, it was estimated that there would be 241,740 new cases in 2012, with around 28,170 deaths (2). As with any cancer, early detection and treatment is necessary for good survival rates. Digital Rectal Examination (DRE) and Prostate-Specific Antigen (PSA) testing have been commonly used for CaP screening (3). However, both DRE (4) and PSA (5) testing lack specificity, and hence, patients have to undergo unnecessary biopsies. The specificity of the PSA blood test is low due to false-positive elevation of PSA levels under some benign conditions such as inflammation or benign prostatic hyperplasia (6). Therefore, other biomarkers are also being studied for CaP diagnosis (7).

At the microscopic level, the features that differentiate CaP from benign prostate tissue are the loss of normal glandular architecture, increased cellular density, and altered microvasculature (8). Due to the loss of normal glandular architecture in some high-grade CaP, there would be fewer reflective interfaces and reduced echo-texture on conventional ultrasound (9). Increased cellular density results in the reduction in elasticity, which may be analyzed using real-time elastography (10). Changes in cellular density can be studied using Magnetic Resonance Imaging (MRI) (11). Finally, increased local perfusion can be visualized with contrast-enhanced computed tomography (12) or MRI (13) or contrast-enhanced ultrasound (9). Thus, it is evident that there is no single modality that can adequately capture all changes between benign and malignant prostate tissue.

Prostate Transrectal Ultrasound (TRUS) images can be easily acquired real-time at lower cost, and hence is widely used for CaP diagnosis. However, the prostate regions in TRUS images are characterized by a weak texture, speckle, short gray scale ranges, and shadow regions (14). Moreover, there is also a potential for benign processes to have texture characteristics quite similar to that of CaP (15). Hence, there is a need for better image analysis frameworks to effectively quantify the subtle textural changes in cancerous and non-cancerous TRUS prostate images in order to accurately detect CaP. Visual analysis and study of any medical image is time-consuming and prone to subjective evaluations. Hence, Computer Aided Diagnostic (CAD) tools are being employed to obtain a more objective and reproducible diagnosis. In this work, we have used several feature extraction algorithms and classifiers to build a CAD framework called "UroImage" for CaP identification. The objective of this work is two-fold: first, to determine representative features that can quantify even the subtle differences between the two classes: non-cancerous and cancerous; second, to study several classifiers using the selected features to determine the best classifier-feature set combination that will result in good classification accuracy in CaP detection.

Early studies using TRUS in CAD frameworks studied the use of first and second order statistics textural parameters (16-18). Subsequently, texture and spectra features extracted from Radio Frequency (RF) data were used (19, 20). PSA value and patient’s age as features along with spectral parameters were used (21). The limitation of these studies is that the methodology is expensive, time consuming and requires several steps such as segmentation of the prostate Region of Interest (ROI), registration of the slices to build the three dimensional model, probability map generation and validation. Moreover, all these steps have to be adequately evaluated to establish the validity of the protocol. Morphologic features and multiresolution textural features were also used for malignancy detection (22), and a multi-feature classifier based on generalized discriminant analysis with Gaussian kernels was also developed (23). The results of recent studies show that combining features extracted from RF analysis of ultrasound signals and texture would result in effective classification (24).
Medical images may contain non-linear variations which may be different in cancerous and non-cancerous classes, and hence, detecting and quantifying these changes will help in CaP detection. In such images, the phase contains important information about the grayscale intensity variations. Higher Order Spectra (HOS) features precisely use this information, and we extracted HOS and Discrete Wavelet Transform (DWT) based features from manually segmented ROI from the TRUS images and used them for classification. The ground truth of whether an image was non-cancerous or cancerous was found with the help of biopsy, the gold-standard.

The proposed CAD system "UroImage" consists of an on-line system (shown on the right side of Figure 1) which processes a test image. This system predicts the class label based on the transformation of the on-line grayscale feature vector by the training parameters calculated by an off-line learning system (shown on the left side of Figure 1). The off-line classification system consists of a classification phase which produces the classifier training parameters with the aid of the grayscale off-line training features and the respective off-line ground truth training class labels (0/1 for non-cancerous/cancerous). In both systems, the grayscale features are non-linear HOS and DWT based features extracted from the prostate ROI that is manually segmented using ImgTracer™ (25, 26). Significant features are selected using the t-test. We evaluated several supervised learning based classifiers such as Decision Tree (DT), Fuzzy Sugeno (FS), Gaussian Mixture Model (GMM), K-Nearest Neighbor (KNN), Radial Basis Probabilistic Neural Network (RBPNN), and Support Vector Machine (SVM) as off-line learning classifiers. To test the performance of the system, we split the data set into a training set and a test set. The training set images were used to develop the classifiers. The built classifiers were evaluated using the test set. For evaluation, we used a k-fold cross validation protocol. The predicted class labels of the test images and the corresponding ground truth labels (0/1) were compared to determine the performance measures such as sensitivity, specificity, accuracy, and Positive Predictive Value (PPV).

Insert Figure 1 here

Materials and Methods

Data Acquisition

We retrospectively evaluated the ultrasound and biopsy results of 144 asymptomatic men who had either an abnormal PSA or DRE. Approval was obtained from the ethical committee of the Rhode Island Hospital where the study was conducted from August to October 2011. A signed consent was obtained from the subjects prior to the study. All patients had average serum PSA (tPSA) levels drawn before DRE and TRUS-guided biopsy of the prostate. The tPSA level (ng/mL) was measured using a monoclonal radioimmunoassay (Tandem R, Hybritech, San Diego, California). Ultrasound-estimated prostate volume was calculated using the volume of an ellipsoid \((\pi/6 \times \text{length} \times \text{width} \times \text{height})\). Biopsies were performed using a spring-loaded 18-gauge needle in men with suspicious DRE, tPSA greater than 2.5 ng/mL, increasing PSA velocity and/or atypia or high-grade prostatic intraepithelial neoplasia found in a previous biopsy specimen. Among the patients analyzed, the average age at biopsy was 62.6 years (median: 62, min: 43 years, max: 86). tPSA was 7.5 ng/ml (median: 4.8). The average TRUS-determined prostate volume at the biopsy session was 60.7 cc, (median: 50.3 cc). The total number of core biopsies obtained was 12 taken as sextant needle biopsies and lesion-directed biopsies. Gleason 6 was the most common pattern.

Among the 144 patients, 83 were found to have clinically localized adenocarcinoma of the prostate, and 61 had no cancer. The ultrasound machine utilized to perform the biopsies was the B&K Flex Focus with a frequency of 5-10 MHz. The probe model was 8667. The patient’s position was always left lateral decubitus. During image acquisition, we captured the transverse and longitudinal planes at the exact location where the radiologist/urologist predicted the presence of a cancerous lesion. A lesion directed biopsy was also carried
out at this location. This technique made sure that the imaging planes contained the probable cancerous lesion. In the case of non-cancerous cases, the two planes were acquired at a random location once the urologist confirmed via ultrasound that there is no cancer in the imaged area. If the lesion-directed biopsy results were positive, we included that case in the “cancer” labeled cases of our database. This is how the database was formed.

Data Pre-processing and ROI Determination

Prostate gland capsule was chosen as the ROI and was manually segmented using an off-line interactive user-friendly software system called ImgTracer™ (Global Biomedical Technologies, Inc., CA, USA) (25, 26). In this work, Luca Saba, MD, and Giorgio Mallarini, MD, from the Department of Radiology AOU Cagliari, did the tracings on the TRUS images. Both have rich experience of about 15 years in radiology and ultrasound, and thus, we ensured very high accuracy in the segmentation process. Figure 2 presents studied ROIs of cancerous and non-cancerous cases.

Feature Extraction

Discrete Wavelet Transform (DWT)-based features: A wavelet is a small wave of finite duration and finite energy. Wavelet coefficients are obtained by correlating a mother wavelet with the signal (27). This wavelet is shifted continually along the time scale to obtain a set of coefficients at all instants of time. Then, the wavelet is dilated and shifted along the time scale and corresponding set of coefficients is calculated. The wavelet transform using a given wavelet $\psi_{a,b}(t)$ is given by

$$W(i, j) = \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt$$  \hspace{1cm} [1]$$

where, $a$ is the scale factor (related to dilation or compression of wavelet) and $b$ is the translation factor (related to shifting of the wavelet). In this work, we used DWT to decompose each image into several scales by successive low pass and high pass filtering. The coefficients obtained after low pass filtering reflect the slow varying details of the image while high pass filtering results in coefficients that give details about the abrupt changes. Low and high pass filtering are done recursively initially on the input image and then on the approximation coefficients obtained. In this work, we performed two iterative decompositions to obtain the coefficients at two levels. Figure 3 explains the DWT computation of images into various approximation and detail coefficients. The typical DWT images of non-cancerous and cancerous prostate images are shown in Figures 4 (a) and (b) respectively. It is very difficult to decipher the subtle changes in these images using the naked eye. Hence, we have evaluated the average energy of approximate coefficients at level 2 ($a_2$), detail coefficients i.e. horizontal ($h_1$ and $h_2$), vertical ($v_1$ and $v_2$) and diagonal coefficients ($d_1$ and $d_2$) at level 1 and level 2.

Higher Order Spectra (HOS)-based features: HOS is an effective method to get useful information from non-linear signals (28, 29). HOS are moments and cumulants of third and higher order statistics. Radon transform is applied on the TRUS images prior to HOS feature extraction (30, 31). This transform rotates the image around its centre through a chosen angle $\theta$ and computes line integrals along many parallel paths in
the image. The intensities along these lines are transformed into points of the resultant signal. Thus, an image is transformed to a one-dimensional signal from which the HOS features are extracted. Bispectrum \( B(f_1, f_2) \), the third order statistics of the signal, is given by

\[
B(f_1, f_2) = E\left[ X(f_1)X(f_2)X(f_1 + f_2) \right]
\]  

[2]

Here, \( X(f) \) is the Fourier transform of the signal \( x(nT) \), \( n \) is an integer index, \( T \) is the sampling interval and \( E[.] \) is expectation operator. The region \( \Omega \) of computation of bispectrum and bispectral features of a real signal is uniquely given by a triangle \( 0 \leq f_2 \leq f_1 \leq f_1 + f_2 \leq 1 \) as shown in Figure 5.

**Insert Figure 5 here**

In this work, we determined the following HOS features.

**Normalized Bispectral entropy**: \( \text{ent}1 = -\sum_n p_n \log p_n \) \[3\]

where \( p_n = \frac{|B(f_1, f_2)|}{\sum_{\Omega} |B(f_1, f_2)|} \)

**Normalized Bispectral cube entropy**: \( \text{ent}3 = -\sum_n q_n \log q_n \) \[4\]

where \( q_n = \frac{|B(f_1, f_2)|^3}{\sum_{\Omega} |B(f_1, f_2)|^3} \)

The weighted centre of bispectrum (WCOB):

\[
wcobx = \frac{\sum_{\Omega} iB(i, j)}{\sum_{\Omega} B(i, j)}, \quad wcoby = \frac{\sum_{\Omega} jB(i, j)}{\sum_{\Omega} B(i, j)}
\]  

[5]

where \( i \) and \( j \) are the frequency bin indices in the non-redundant region \( \Omega \). \( awcobx \) and \( awcoby \) were computed by substituting the absolute values for each component in the above equations. All these features were extracted for every one degree of Radon Transform from 1 to 180 degrees. Thus, there were a total of 720 extracted HOS features. The entropies characterize the regularity or irregularity of the variations in the bispectrum.

**Feature Selection and Classification**

The significance of the extracted features was tested by finding the \( p \)-value using the Student’s \( t \)-test (32). The lower the \( p \)-value is, the more statistically significant the feature is. Out of the extracted features, we observed that four HOS features (\( \text{ent}1, \text{ent}3, \text{awcobx} \) and \( \text{awcoby} \)) for angle 180º were statistically significant with a \( p \)-value of less than 0.05. Using the DWT method, we have evaluated the average energy of seven features (\( a_2, h_1, h_2, v_1, v_2, d_1 \) and \( d_2 \)) at level 1 and level 2 (see Fig. 3). Among them, we found that the average of the horizontal coefficients at level 2 decomposition, denoted as \( h_2 \) (\( h_2 \)) was statistically significant.
(p<0.05). Hence, we have chosen only $h_2$, which gives useful information about the sudden changes in the prostate image.

The significant features were given as input to the six classifiers mentioned earlier. The DT classifier partitions the training set in a recursive way until each partition consists of dominant samples from one class. A series of rules are extracted from the tree and these rules are used to recognize the class of the test data (33). RBPN is a form of radial basis network which consists of four layers namely the input layer, pattern layer, summation layer and output layer. The pattern layer generates the product of the weight vector with the input data set. The output of the pattern layer is fed into the summation layer which receives the outputs associated with a given class. The final output layer gives the classification decisions (34).

Fuzzy Sugeno (FS) classifier is a rule based classifier where fuzzy logic is used to take imprecise observations for inputs and arrive at precise values for outputs. In this work, subtractive clustering technique was used to generate a Fuzzy Inference System (FIS) (35-37). An FIS structure containing a set of fuzzy rules that cover the feature space is generated after the training. These rules are used to perform fuzzy inference calculations of the test data. KNN classifier classifies an unknown sample by relating the unknown to a known sample according to some distance or similarity criteria (38). A sample is assigned a class which is the most common among its K nearest neighbors.

SVM maps samples as points in a space. The mapping is done in such a way that samples belonging to different classes are separated by a very clear gap that is as wide as possible (39-41). New samples are mapped to the same space and the decision of which class the sample belongs to is done based on which side of the gap the sample is mapped to. GMM is a generalized basis function network where the basis functions are Gaussian functions. Expectation Maximization (EM) is one of the commonly used maximum likelihood algorithms that are used for fitting the mixture model to the training data (42). The trained GMMs are used to classify the test data.

The performance of these classifiers was evaluated by calculating sensitivity, specificity, PPV and accuracy. Let TN (True Negative) be the number of non-cancerous cases identified as non-cancerous, FN (False Negative) be the number of cancerous cases incorrectly identified as non-cancerous, TP (True Positive) be the number of cancerous cases correctly identified as they are, and FP (False Positive) be the number of non-cancerous cases incorrectly identified as cancerous. Sensitivity is defined as TP / (TP+FN), specificity as TN / (TN+FP), accuracy as (TP + TN) / (TP + FN + TN + FP), and PPV as TP / (TP+FP).

In a $k$-fold cross validation technique, the dataset is randomly divided into $k$ parts, with each part containing the same proportion of data belonging to the two classes. During the training phase, $(k-1)$ parts with their corresponding labels are used to train the classifiers and obtain the classifier parameters. During the test phase, the trained classifiers are used to predict the labels of the remaining one part and for the calculation of the performance measures. This procedure is then repeated for another $(k-1)$ times with a different test set each time. The overall performance measures are taken as the average of the measures obtained in each of the $k$-folds. In this study, we performed cross validation trials with two different fold sizes to demonstrate the impact of the training size set on the results. For 3-fold cross validation, 101 images (58 non-cancerous and 43 cancerous) were used for training each time, and 43 samples (25 non-cancerous and 18 cancerous) were used for testing. In the case of 10-fold cross validation, 130 images (75 non-cancerous and 55 cancerous) were used for training each time, and 14 samples (8 non-cancerous and 6 cancerous) were used for testing.

**Results**

**Significant Features**
Among all the extracted features, t-test indicated that only five were clinically significant. Table I lists these features. Medical images with non-uniform pixels cannot be fully described by second-order measures. High-order statistic information would be able to reveal some information about non-linearity and deviation of Gaussianity which would most likely be present in abnormal cases. Hence, HOS based features are more discriminative than the second-order measures from power spectrum. Moreover, the HOS performs better even in noisy conditions (43-45). It can be seen from Table I that all the four significant HOS features, measured at the Radon Transform angle of 180°, are higher for cancerous ROIs indicating that there is a general increase of entropy in the cancerous cases. DWT helps to quantify sudden changes in the pixels, and it is evident that the value of the DWT feature h2 is lower for cancerous ROIs indicating that the high frequency components are low in the cancerous cases. From Figure 2, it can be seen that there is a change in the pixel intensity values in the horizontal direction. So the DWT coefficients in the horizontal direction (h2) have provided highest discrimination among other competing sub bands. Also the HOS features measured at 180°, which is the horizontal direction, proved to be significant. Figure 6 provides mean and standard deviation of two groups (non-cancerous and cancerous) for different extracted features.

Insert Table I here
Insert Figure 6 here

Classification Results

In DT, during the training phase, the tree parameters were learnt. Rules are generated as a result of training and these rules are used to classify the test dataset. In RBPNN, the weights of the neural network were determined using the gradient descent algorithm. In Fuzzy Sugeno classifier, the center and width of the Gaussian membership function were derived during training of the classifier. In the case of SVM, the support vectors (which are the parameters of the SVM) were computed by maximizing the margin between the two classes. The SVM with a RBF kernel function has two training parameters: cost (C) which controls over fitting of the model, and sigma (σ) to control the degree of nonlinearity of the model. The values of these training parameters were determined using a ‘grid search’ approach (46). The final SVM parameters obtained were: C: 75 and sigma: 0.001. In the case of GMM, the Gaussian probabilities were derived during training.

On evaluating the classifiers with the three-fold classification technique, we observed that the SVM classifier resulted in the highest accuracy of 97.7%, sensitivity of 96.3%, specificity of 98.7% and PPV of 98.1%. Table II lists the classifier parameters obtained using 10-fold cross validation. When the cross-validation parameter k was increased from 3 to 10, higher accuracy was recorded by the Fuzzy and SVM classifiers. SVM gave the best accuracy of 97.9% and sensitivity of 98.3% and the second best values of 96.9% and 97.5% for PPV and specificity, respectively. Thus, it is evident that in both cross validation protocols, SVM performed consistently well, resulting in an accuracy over 97.5%.

Insert Table II here

Discussion

CaP is the second leading cause of cancer death in men (47). Early detection of CaP is the key to increased survival rate. PSA has a high false-positive rate of about 76% (48). Visual analysis of TRUS images is limited by low reliability due to the prominent overlap of benign and malignant lesion characteristics and by inter-observer variabilities. The standard test for CaP detection is TRUS guided symmetrical needle biopsy. This test is painful and invasive. Due to speckle and shadow regions in the TRUS images, sometimes tissue is extracted from non-cancerous prostate regions of a CaP affected person giving a false negative biopsy result. This leads to the need for an alternate adjunct non-invasive prostate detection method which
has high classification accuracy. Once CaP is detected through this reliable, painless method, the subject can undergo needle biopsy for a second opinion. For this purpose, many studies were conducted to develop techniques for processing TRUS images to detect CaP. Table III provides an overview of different studies including the methods/features used and the performance achieved.

Ours is the first work to use a combination of DWT and HOS non-linear features for analyzing TRUS images and we have achieved a high accuracy of 97.9% with SVM and Fuzzy classifiers along with high specificity and sensitivity. We also found the system’s reliability to detect CaP is not affected much by the choice of $k$ in $k$-fold cross-validation. SVM recorded the best performance since it was able to determine a good separating hyperplane with the maximum margin that could clearly discriminate both classes. The performance of Fuzzy classifier was also equally good. We used clustering algorithm to reduce the number of fuzzy rules, improve the system interpretability and accuracy, and to reduce problem complexity. Fuzzy classifiers employ parallel processing, and hence, require less computational time. The key features of this work can be summarized as follows:

- The proposed technique has presented good accuracy, sensitivity, specificity and PPV (more than 96%) which are higher than those reported in previous studies.
- If needle biopsies are to be taken randomly without any knowledge of tissue morphology, then the specificity of the test will be affected. Crawford et al. (58) observed that more than 50% of cancers remain undetected during initial needle biopsies. To overcome this, an atlas was developed to locate the statistically probable cancerous regions in prostate (59). Atlas is a probabilistic map of prostate region built by taking tissue samples from thousands of CaP affected people. This atlas is used as a guide to decide where to drive the needle into prostate regions. Though this atlas is precise, it is very expensive, time consuming and tedious to design and build it for a long term purpose. Our work is an effective substitute to atlas guided needle biopsy. Our work can tell with good accuracy whether the prostate is affected by cancer or not.
- Our technique is economical, non-invasive and can be considered a precursor to atlas guided needle biopsy. If the test is positive, the person can then go for needle biopsy for confirmation. Thus, in the clinical sense, the proposed technique can work as an adjunct tool for CaP detection.
- This is the first work for CaP detection which uses the methods of DWT and HOS. We combined DWT and HOS together and selected significant parameters from both of these methods. DWT and HOS can capture sudden changes in the images. They are standard, generalized methods used to capture maximum information from non-linear signals or data. HOS is more robust to noise and can capture the phase information in the image. We tried several combinations of non-linear features and found that that selected combination (Table I) delivered the best performance. In our recently published study (60), we have demonstrated the effectiveness of the HOS features for medical image analysis by obtaining over 95% accuracy for oral cancer detection using microscopic images. We have also demonstrated that DWT, in combination with texture, can be very effectively used to detect thyroid malignancy by analyzing 3D HRUS images (61) and CEUS images (62) with classification accuracies of 100% and 98.9% respectively.
- We demonstrated that we could obtain very high accuracy with a small number of significant features that were fed to the classifier (just five features – one DWT feature and four HOS features). This makes the design and training of the classifier simpler, and computationally fast.
- In both 10-fold and 3-fold cross validation protocols, an almost equal accuracy of over 97.5% was achieved. This clearly demonstrates that the system is able to sustain the accuracy level even with lower training data sets. This underlines the reliability of the system whose performance is almost untouched by the variation in the number of training population.

There are certain limitations of the study, which are listed below.
• The current study does not take multiple tracings into account where multiple readers can segment the prostate region. Even though this is a shortcoming, it does not cause a detrimental effect on the system’s performance. We, however, are working towards establishing a protocol for multiple tracings and performance differences in the inter-observer analysis.

• We have not benchmarked our algorithm with other prostate CAD algorithms because we do not have access to their datasets.

• Our technique is not completely automated as urologists with decades of experience performed manual segmentation of the prostate capsule using ImgTracerTM. In future, we intend to make this segmentation process automated and still obtain the same accuracy level as obtained now.

• We would also like to associate the histology of the patients and not just biopsy. This would provide a strong correlation to the validation of our classification system. Currently, we are using biopsy as the gold standard.

• Finally, we do believe that the data size can be increased over time. However, the system, built and evaluated using 144 patients, has shown encouraging results in both three-fold and ten-fold cross validation protocols because of the use of the novel combination of five DWT and HOS features which have adequately captured the subtle differences between the cancerous and non-cancerous prostate images.

• A key limitation in this technique is that the plane to be examined is selected based on the urologist’s knowledge. Thus, there is a possibility of mismatch between the examined location and the actual nature of the prostate. This issue can be alleviated in two ways: (1) by training the system with a larger database to make the learning process robust, and then evaluate many planes (instead of only one chosen by the urologist) from a new patient’s prostate in order to predict the presence of cancer; (2) by going for 3-D TRUS imaging. Since this is only a preliminary study, we intend to conduct future research in these directions to improve our system.

Conclusions

Prostate cancer diagnosis is a difficult task. The standard test of needle biopsy is a painful and invasive test which has low specificity. If we have to improve the specificity of needle biopsy, we have to resort to building the expensive and time consuming atlas. In this paper, we have proposed an affordable, painless and very reliable on-line automated system called “UroImage” to classify prostate tissue into cancerous and non-cancerous. We extracted DWT and HOS features from TRUS images and fed them to classifiers. With both 3-fold and 10-fold cross-validation data resampling protocols, our system achieved more than 95% for accuracy, PPV, sensitivity and specificity measures using SVM and fuzzy classifiers. Our work can be used as an initial test to reliably detect whether the prostate of a person is cancer-affected or not.

Conflict of interest statement

None of the authors have any financial or personal conflict of interest that could inappropriately influence the writing or publication of this manuscript.

Ethical approval

Approval was obtained from the ethical committee of the Rhode Island Hospital, Providence, RI, USA, where the study was conducted over a three-month period (Aug-Oct 2011). IRB Registration #: 0000396, 00004624

References


**Figure Captions**

**Figure 1:** Block diagram of the proposed CAD technique “UroImage” for prostate tissue characterization and classification; The blocks outside the dotted shaded rectangular box represent the flow in the off-line system, and the blocks within the dotted box indicate the on-line system

**Figure 2:** Typical prostate image ROIs: (a) non-cancerous cases; (b) cancerous cases. The areas within the blue contours (indicated by white arrows) are the cancerous regions in the cancerous ROIs

**Figure 3:** Illustration of DWT decomposition of prostate image

**Figure 4:** DWT decomposition of prostate image: (a) Non-cancerous image and (b) Cancerous image

**Figure 5:** Principal domain or non-redundant region Ω of computation of the bispectrum. Frequencies are shown normalized by the Nyquist frequency

**Figure 6:** Mean and standard deviation of features depicted as ranges

**Table Captions**

**Table I:** Table of feature values and t-test results for statistical significance

**Table II:** Classification Results obtained using 10-fold cross validation

**Table III:** Summary of previous studies on CAD based CaP detection
Prostate Ultrasound Training Database

ROI Segmentation using ImgTracer™

Higher Order Spectra
Discrete Wavelet Transform

Ground Truth
Feature Selection
Classification
Training Parameters

Calculation of significant features as determined in the training phase

Classification of test images

Non-Cancerous
Cancerous

Off-line System

On-line Real-Time System

Prostate Ultrasound Test Database

ROI Segmentation using ImgTracer™

Figure 1
Figure 2
LL coefficients are obtained by applying low pass filtering to both rows and columns. The output image is similar to the original image. These coefficients are called approximation coefficients.

HL coefficients are obtained by applying low pass filtering to the rows and high pass filtering to columns. These coefficients show diagonal details of the image. They represent the finest-scale wavelet coefficients.

LH coefficients are obtained by applying high pass filtering to the rows and low pass filtering to columns. These coefficients show horizontal details of the image.

HH coefficients are obtained by applying high pass filtering to the rows and columns. These coefficients show vertical details of the image.

Figure 3
Figure 4
Figure 6
Table I

<table>
<thead>
<tr>
<th>Feature</th>
<th>Non-Cancerous</th>
<th>Cancerous</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DWT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_2$</td>
<td>15.4 ±6.49</td>
<td>13.0 ±3.50</td>
<td>0.011</td>
</tr>
<tr>
<td><strong>HOS 180°</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ent1</td>
<td>0.376±6.895E-02</td>
<td>0.407±7.416E-02</td>
<td>0.0085</td>
</tr>
<tr>
<td>ent3</td>
<td>3.060E-02±3.504E-02</td>
<td>0.04665±4.987E-02</td>
<td>0.025</td>
</tr>
<tr>
<td>awcobx</td>
<td>8.02±1.89</td>
<td>8.9±2.03</td>
<td>0.0062</td>
</tr>
<tr>
<td>awcoby</td>
<td>2.96±0.402</td>
<td>3.08±0.354</td>
<td>0.048</td>
</tr>
</tbody>
</table>
Table II

<table>
<thead>
<tr>
<th>Classifier</th>
<th>TN</th>
<th>FN</th>
<th>TP</th>
<th>FP</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>85.7</td>
<td>83</td>
<td>85</td>
<td>86.3</td>
</tr>
<tr>
<td>Fuzzy</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td><strong>97.9</strong></td>
<td><strong>98.3</strong></td>
<td><strong>96.7</strong></td>
<td><strong>98.8</strong></td>
</tr>
<tr>
<td>KNN</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>90.7</td>
<td>92.5</td>
<td>88.3</td>
<td>92.5</td>
</tr>
<tr>
<td>PNN</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>90.7</td>
<td>92.5</td>
<td>88.3</td>
<td>92.5</td>
</tr>
<tr>
<td>SVM</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td><strong>97.9</strong></td>
<td><strong>96.9</strong></td>
<td><strong>98.3</strong></td>
<td><strong>97.5</strong></td>
</tr>
<tr>
<td>GMM</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>96.4</td>
<td>95.7</td>
<td>96.7</td>
<td>96.3</td>
</tr>
<tr>
<td>Authors</td>
<td>Modality/ Features</td>
<td>Performance (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------------------------------------------------</td>
<td>-------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basett et al. (16)</td>
<td>Ultrasound/ Co-occurrence matrices</td>
<td>Accuracy: 78%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynen et al. (17)</td>
<td>Ultrasound/ Tissue characterization</td>
<td>Sensitivity: 80.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Houston et al. (18)</td>
<td>TRUS/ Statistical distribution of digital grayscale values</td>
<td>Accuracy: 78%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmitz et al. (19)</td>
<td>Ultrasound/ Tissue characterization parameters</td>
<td>Sensitivity: 82%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheipers et al. (20)</td>
<td>TRUS/ Spectral, texture, first order and morphologic parameters</td>
<td>Specificity: 88%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Han et al. (22)</td>
<td>Ultrasound/ Multiresolution autocorrelation texture features and clinical features</td>
<td>Sensitivity: 92-96%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petricoin et al. (49)</td>
<td>Serum proteomic mass spectra/ Serum proteomic pattern</td>
<td>Specificity: 90-95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabesh et al. (50)</td>
<td>Hematoxylin-and-Eosin (H&amp;E)-stained tissue/ Several texture and wavelet features</td>
<td>Accuracy: 94.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spurgeon et al. (51)</td>
<td>Predictors of aggressive cancer analyzed using CART</td>
<td>Sensitivity: 91.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keles et al. (52)</td>
<td>Medical data using NEFCLASS tool</td>
<td>Specificity: 100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim et al. (53)</td>
<td>Near infrared (NIR) spectroscopic data/ Spectral data based features</td>
<td>Specificity: 98.33%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al. (54)</td>
<td>Selection of principal genes</td>
<td>Sensitivity: 94%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bouatmane et al. (55)</td>
<td>Multispectral images of microscopic prostatic tissue textures from needle biopsy/ Spectral band based features</td>
<td>Specificity: 64%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aboofazeli et al. (56)</td>
<td>TRUS/ wavelet transform based features of ultrasound radiofrequency (RF) time series</td>
<td>Accuracy: 99.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. (57)</td>
<td>TRUS/ texture features using wavelet decomposition filter</td>
<td>Accuracy: 91.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>This work</strong></td>
<td><strong>TRUS/ DWT and HOS</strong></td>
<td><strong>Accuracy: 97.9% (SVM and Fuzzy)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Sensitivity: 98.3% (SVM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Specificity: 98.8% (Fuzzy)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>PPV: 98.3% (Fuzzy)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>