ULTRASOUND-GUIDED INTERVENTION FOR PROSTATE CANCER DETECTION

by

AMIR KHOJASTE

A thesis submitted to the
School of Computing
in conformity with the requirements for
the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada
November 2014

Copyright © Amir Khojaste, 2014
Abstract

Prostate cancer is the second most common cancer diagnosed in North American men. It can be managed with a survival rate of over 90% if diagnosed early. Although diagnosis of prostate cancer is important, however not all the diagnosed cancers are life-threatening. Hence, accurate prognosis of prostate cancer is crucial to avoid overtreatment of patients with indolent disease. Recently, a novel ultrasound tissue typing technique was proposed that utilizes ultrasound RF data acquired from a stationary location of the tissue over time. The goal in this thesis is to present the feasibility of characterizing aggressive prostate cancer using ultrasound RF time series. We pursue this goal by analyzing data from two ex vivo and one in vivo studies involving prostatectomy patients.

In almost all of the ultrasound-based interventions, calibration of the ultrasound probes is of crucial importance to determine the location of image plane in a global coordinate system. Calibration is routinely performed by imaging a geometrically known object. The calibration problem is then solved by relating the feature locations in this object in the ultrasound images and their true locations. Recently a calibration technique was proposed that eliminated the need for fabricating complex phantoms. This technique is based on imaging a flat plate that is immersed into a water tank. In this thesis, we also evaluate the robustness of this calibration by performing extensive
experimental evaluations.

In the ex vivo studies, in a cross-validation framework, areas under the accumulated receiver operating characteristic curve (AUC) of 0.8 and 0.85 were obtained from 15 and 6 patients, respectively. The results are confirmed by performing a gold-standard pathology to ultrasound registration. We also provide likelihood maps showing the probability and extent of higher grade cancer in the entire cancerous area of each patient. In the in vivo study on prostatectomy cases, an AUC of 0.88 is achieved on characterizing higher grade prostate cancer. The experimental evaluations of single wall approach yielded reproducibility accuracy of $3.14 \pm 1.5 \text{ mm}$. Having a well-defined hand motion pattern for ultrasound probe movements and digitizing the wall plane first can improve the robustness of this calibration framework.
Statement of Co-Authorship

The work presented in this thesis was accomplished under the supervision of Dr. Parvin Mousavi.

Acknowledgments

I would first like to thank my supervisor, Dr. Parvin Mousavi for her guidance throughout my thesis project. Thank you for providing me with the opportunity for growth. Also I would like to thank Dr. Gabor Fitchinger and Dr. Purang Abolmaesumi for their helpful advice and guidance. Thank you Dr. Berman and Dr. Boag for your clinical insights.

Thanks to members of Med-i Lab: Dr. Farhad Imani, Lili, Nathan, Layan and Sinthu. Special thanks to Dr. Imani for his support, helpful advice and interest in this thesis.

Finally, I would like to thank my friends and family. I have had two phenomenal years in Kingston. Thanks for the laughter and thanks for the tears.
Contents

Abstract i

Statement of Co-Authorship iii

Acknowledgments iv

Contents v

List of Tables ix

List of Figures x

Glossary 1

Chapter 1: Introduction 3

1.1 Motivation 4

1.1.1 Thesis Objectives 9

1.1.2 Thesis Contribution 10

1.2 Thesis Outline 11

Chapter 2: Background 13

2.1 Prostate cancer 13
2.1.1 Prostate anatomy ........................................... 13
2.1.2 Grading and staging of prostate cancer .......................... 14
2.1.3 Active Surveillance ........................................... 15
2.1.4 Diagnosis of prostate cancer .................................... 16
2.2 Prognosis of prostate cancer ....................................... 18
2.3 Ultrasound-based characterization of prostate cancer and its aggressiveness .................................................. 21
  2.3.1 Ultrasound physics and image formation ....................... 21
  2.3.2 Texture-based methods for prostate cancer diagnosis .......... 23
  2.3.3 Spectral-based methods for prostate cancer diagnosis .......... 24
  2.3.4 Elastography-based methods for diagnosis and prognosis of prostate cancer .............................................. 25
  2.3.5 Color Doppler methods for prostate cancer diagnosis and prognosis ............................................................... 27
  2.3.6 Time series-based methods for prostate cancer diagnosis and prognosis ............................................................... 30
2.4 Calibration ...................................................... 32
2.5 Chapter summary and conclusion .................................... 34

Chapter 3: Characterization of aggressive prostate cancer, \textit{ex vivo}
clinical study ...................................................... 37
  3.0.1 Data collection .............................................. 37
  3.0.2 ROI selection .............................................. 39
  3.0.3 Feature extraction ............................................ 41
  3.0.4 Classification .............................................. 45
List of Tables

3.1 Total number of segmented ROIs from designated areas with different Gleason scores .......................... 43
3.2 Average accuracy, sensitivity, specificity and AUC in the leave-one-patient-out cross-validation and its comparison with previously reported features in the literature. .......................... 51

4.1 \textit{in vivo} study: total number of segmented ROIs from designated areas with different Gleason scores .......................... 62
4.2 \textit{ex vivo} study: total number of segmented ROIs from designated areas with different Gleason scores .......................... 63
4.3 \textit{in vivo} experiment: average accuracy, sensitivity, specificity and AUC in the 10-fold cross-validation .......................... 71
4.4 \textit{ex vivo} experiment: average accuracy, sensitivity, specificity and AUC in the leave-one-patient-out cross-validation .......................... 71
## List of Figures

2.1 Prostate Anatomy .................................................. 14
2.2 RF time series ......................................................... 29

3.1 *Ex vivo* prostate specimen data collection setup ............... 39
3.2 Registration of pathology and ultrasound data ..................... 41
3.3 Registration of pathology and ultrasound data ..................... 42
3.4 Spectrum of a sample ROI of a patient ......................... 46
3.5 ROC curve for leave-one-patient-out cross validation .......... 52
3.6 colormaps .......................................................... 53

4.1 Cutting box used for data collection ......................... 57
4.2 Registration flow-chart ............................................. 59
4.3 ROI selection for *ex vivo* study ................................ 61
4.4 Results of the *in vivo* experiments ............................ 66
4.5 Results of the *in vivo* experiments ............................ 67
4.6 Results of the *in vivo* experiments ............................ 68
4.7 *ex vivo* experiment: feature distribution .................. 69
4.8 *ex vivo* experiment: K-means ................................ 69
4.9 *ex vivo* experiment: ROC curve .............................. 70
5.1 General SW Calibration Framework ........................................ 74
5.2 Intersection of image plane and the wall plane .......................... 76
5.3 RF cross-correlation automatic line segmentation approach ............ 81
5.4 First round of data collection .................................................. 82
A.1 Colormaps of all slices of patients in chapter 3: part I .................. 112
A.2 Colormaps of all slices of patients in chapter 3: part II ................ 113
A.3 Colormaps of all slices of patients in chapter 3: part III ............... 114
A.4 Colormaps of all slices of patients in chapter 3: part IV ............... 115
Glossary

2D  2 Dimensions/Dimensional.

3D  3 Dimensions/Dimensional.

ADC  Apparent Diffusion Coefficient.

AJCC  American Joint Committee on Cancer.

AS  Active Surveillance.

AUC  Area Under receiver Characteristic.

CR  Calibration Reproducibility.

CT  Computed Tomography.

DCE  Dynamic Contrast Enhanced.

DCED  Dynamic Contrast Enhanced Doppler.

DFT  Discrete Fourier Transform.

DRE  Digital Rectal Examination.

DWI  Diffusion Weighted Imaging.
Glossary

**GSU**  Gray Scale Ultrasound.

**IGD**  Image Guided Diagnosis.

**KGH**  Kingston General Hospital.

**MR**  Magnetic Resonance.

**MRSI**  Magnetic Resonance Spectroscopic Imaging.

**MWW**  Mann-Whitney-Wilcoxon.

**PLUS**  Public software Library for UltraSound imaging research.

**PSA**  Prostate Specific Antigen.

**RF**  Radio Frequency.

**ROC**  Receiver Operating Curve.

**ROI**  Region Of Interest.

**RTE**  Real Time Electrography.

**SNR**  Signal to Noise Ratio.

**SW**  Single Wall.

**T1WI**  T1-weighted imaging.

**T2WI**  T2-weighted imaging.

**TRUS**  TransRectal UltraSound.
Chapter 1

Introduction

The type of research described in this thesis is a combination of two independent projects in the area of Image-Guided Diagnosis (IGD). Using computer technology, the goal in IGD is to guide the intervention providing a second diagnostic opinion for the physicians.

The first part of this thesis is about prognosis of prostate cancer utilizing ultrasound radio frequency (RF) time series method. Here, we have investigated the feasibility of RF time series approach to characterize multiple malignancy levels of prostate cancer. Results of two ex vivo and one in vivo experiments show promise in using RF time series to characterize clinically significant prostate cancer.

Probe calibration is an inevitable prerequisite in almost all of the ultrasound guided interventions such as prostate biopsy. Hence, the second part of this thesis evaluates a simple calibration method (Single Wall - (SW)) to be used as a reliable and robust calibration approach across multiple centers, as a user-independent and center-independent method for calibration of 2D ultrasound probes.
1.1 Motivation

Prostate cancer is the most common cancer diagnosed in North American men, excluding skin cancers [31], [47]. It is estimated that in 2014, approximately 233,000 new cases of prostate cancer will occur in North America. On average 1 out of 7 North American men will develop prostate cancer during their lifetime and 1 out of 36 will die of it [87]. This disease accounts for 10% of male cancer-related deaths in North America [47]. Prostate cancer can be managed with a survival rate of over 90% if diagnosed at the early stage of cancer development [39]. Current treatment options for prostate cancer are, radiotherapy brachytherapy, cryotherapy, thermal ablation and surgery [23]. One of the major issues in the treatment of prostate cancer is the over-diagnosis and over-treatment of this disease. Over-diagnosis occurs when an indolent cancer is recognized as an aggressive or clinically significant cancer. One of the biggest challenges after diagnosis of prostate cancer is to determine the aggressiveness of the disease [103]. Patients with aggressive prostate cancer would benefit from the treatment options such as radiotherapy compared to the patients with benign cancer. One of the important prognostic factors in determining the aggressiveness of the cancer is the Gleason score of the tumor. In this system a grade in the range of 1-5 is assigned to a cancerous area based on its similarity to normal cells, with 1 being the most similar to normal cells. Since usually more than one pattern of cancer is present in the prostate, a primary and a secondary grade are assigned to the most prevalent and the second most prevalent pattern of cancer in the tissue. The summation of these two grades is referred to as Gleason score [90].

The current approach for early diagnosis of prostate cancer depends on a test to measure the prostate specific antigen (PSA) level in the blood. Usually a PSA
test is accompanied by a digital rectal examination (DRE) test which evaluates the stiffness and the size of prostate cancer [52]. The ultimate approach for diagnosis of prostate cancer is the core needle biopsy which is performed under transrectal ultrasound (TRUS) guidance. TRUS-guided biopsy is then followed by histopathological analysis of the extracted cores. Due to the limitation of TRUS-guided biopsy in detecting prostate cancer with high sensitivity and specificity, biopsy process is usually performed at predefined anatomical locations [34]. Repeated biopsies are typical consequences of performing blind biopsy. Patients will undergo repeated biopsy while cancer is advancing. Hence, a need for additional guidance during the biopsy process seems inevitable.

Although diagnosis of prostate cancer is important, however, not all the diagnosed cancers are lethal. Hence, characterizing the aggressiveness of the tumors is crucial to avoid over-treatment of patients with indolent cancers. One of the biggest challenges in the clinical management of patients with newly diagnosed prostate cancer is to shift from detecting the cancer alone to characterizing its aggressiveness [103]. Patients with clinically significant prostate cancer would benefit more from radical treatment options compared to patients with indolent cancers. The treatment option for patients with indolent prostate cancers is to undergo active surveillance (AS) or watchful waiting [50]. AS include monitoring the PSA level of the patients and performing DRE and re-biopsy on a regular basis. AS enhances the quality of life of the patients by postponing the invasive treatment and its side effects and in some cases preventing the treatment all together. AS provides definitive clinical management of prostate cancer by avoiding invasive treatment for the the majority of patients with low-risk cancer [81].
Several techniques have been proposed to augment diagnosis of prostate cancer using different imaging modalities. However, in the literature there exist limited studies for prognosis of prostate cancer that report the aggressiveness of the detected tumors. Yet, an extensive study to differentiate between multiple aggressiveness levels of prostate cancer is lacking in the literature. The ultimate goal of methods developed for cancer diagnosis is to provide cancer distribution maps which highlight the most suspected cancerous area in the prostate. This can help augment biopsy by targeting the process versus performing it blindly. Ultrasound-based methods for prostate cancer diagnosis and prognosis have been a main approach for this purpose. This is due to the economic advantage of ultrasound, the fact that it is a non-ionizing modality and more importantly that it is already part of the biopsy process.

The current ultrasound-based methods utilize the information of B-mode or RF ultrasound images for tissue characterization. Texture features of B-mode images have been previously used only for diagnosis of prostate cancer [73]. The major limitation of these methods is the dependency on imaging parameters; in practice texture-based features apparently are usually combined with other tissue characterizing methods for prostate cancer detection. Spectral features of RF signals have also been utilized to classify prostate cancer [29]. These methods require a standard approach to calibrate ultrasound RF signals. Spectral-based and texture-based approaches have not been previously applied for prognosis of prostate cancer. An alternative ultrasound-based approach for prostate cancer detection that has been an active field of study in the past decades is elastography. Here, the mechanical properties of the tissue are estimated by tissue compression obtained either by an external mechanical force or by an acoustic radiation [91]. Compared to conventional ultrasound, conflicting detection
rates for aggressive tumors using elastography are reported [101], [115]. The clinical application of elastography-based approaches is limited due to the need for modifying the hardware in the ultrasound machines. Ultrasound Doppler imaging has emerged as an alternative approach for augmenting the conventional ultrasound imaging for diagnosis and prognosis of prostate cancer. This approach operates based on the fact that the received ultrasound echoes from a moving object will undergo a frequency shift which is proportional to the object’s speed [13]. It was shown that ultrasound Doppler outperforms TRUS on prognosis of prostate cancer [61].

Recently, it was shown that the time series of ultrasound echoes taken from a fixed location of the tissue can be used for tissue characterization [65]. This method was successfully applied to classify benign and malignant prostate tissues in 35 ex vivo specimens [65] and in two independent in vivo studies (with 14 [45] and 6 [65] patients, respectively.) This approach outperforms other ultrasound-based methods for detecting prostate cancer. RF time series has not been previously used to grade prostate cancer. Here, for the first time, we demonstrate the performance of RF time series for differentiating between lower and higher grade of prostate cancer.

Calibration of ultrasound probes is an inevitable prerequisite in almost all ultrasound data acquisition and ultrasound-guided interventions. Probe calibration refers to the process of determining the spatial transformation between the image pixels and a global coordinate system. Several calibration approaches were proposed in the past two decades [19], [104]. The most accurate calibration methods are performed by imaging a geometrically known object refereed to as a phantom. By combining the prior knowledge of image features in the phantom and their true locations, calibration parameters are found. Points and lines are features of the phantoms that are
1.1. MOTIVATION

used for calibration purposes. Single-point phantom [95], single cross-wire phantom [100] and multi N-wire phantoms [20] are examples of point-based phantoms. When using point-based calibrations, points in the ultrasound images should be segmented manually; it is difficult to focus the ultrasound beam at a single point. One way to increase the calibration accuracy is to use automatic instead of manual feature selection. In plane-based phantoms a plane is imaged; the intersection of this plane and the imaging plane appearers as a line in ultrasound images. These approaches take advantage of the fact that lines in ultrasound images can be segmented automatically using edge detection approaches. SW [84] and Cambridge [85] phantoms are examples of plane-based phantoms. In particular, SW is composed of a single flat plate that makes it attractive as a phantom. In addition to automatic feature selection, measuring differences of features such as measuring the slope of the line fitted to a set of points instead of absolute positioning of the points, will increase the accuracy of calibration [69].

Recently, a calibration technique was proposed by Najafi et al. [68] that is based on the SW phantom. This approach takes advantage of both of automatic feature selection and differential feature positioning ideas. Compared to other methods, it is attractive to evaluate the SW calibration approach proposed by Najafi et al. to seek the possibility of having an operator-independent and center-independent robust approach of calibration. In this thesis an extensive evaluation of the enhanced SW calibration technique is performed and the precision of this calibration routine is reported to evaluate the possibility of considering this method as a reliable calibration approach.
1.1. MOTIVATION

1.1.1 Thesis Objectives

As it has been discussed before, determining the clinical significance of prostate cancer can help avoid over-treatment of the disease. The main scope of this thesis is to evaluate the performance of ultrasound RF time series for characterizing the aggressiveness of prostate cancer. Given the proven tissue typing capabilities of ultrasound RF time series, for the first time in this thesis, RF time series is applied to characterize higher grade and lower grade prostate cancer. To be specific, our hypothesis is that RF time series can differentiate between higher (Gleason score $\geq 4+3$) and lower (Gleason score $\leq 3+4$) grade of prostate cancer.

Calibration of an ultrasound probe is an inevitable prerequisite in almost all of the ultrasound-based interventions. SW technique was shown to be a good alternative to current ultrasound calibration approaches [68]. In this thesis, an extensive experimental evaluation of the robustness of this approach was performed as a possible alternative for other calibration routines such as multi N-wire [20]. This method has been integrated within the Public software Library for UltraSound imaging research (PLUS). PLUS is an open-source software product for ultrasound image calibration, acquisition and processing [56].

In this thesis, the objectives are:

- *ex vivo* characterization of aggressive prostate cancer for patients who underwent radical prostatectomy, as the first step towards assessing the performance of RF time series in differentiating higher and lower grade prostate cancer.

- *in vivo* characterization of aggressive prostate cancer. This validates the tissue typing abilities of RF time series and represents an application of RF time series that is closer to the clinical translation.
1.1. MOTIVATION

- evaluation of the robustness of SW calibration as an alternative to the current calibration routines and integration of the SW calibration technique within the PLUS library.

1.1.2 Thesis Contribution

The major achievements of this thesis are as follows:

- contribution to a novel method of data acquisition from the prostate tissue in both \textit{ex vivo} and \textit{in vivo} environments. This data collection was designed to efficiently tackle the registration problem between gold standard pathology and ultrasound images.

- development of novel frequency domain features of RF time series. These features are the ratios of the summations of the spectrum of the RF time series in ten frequency bands.

- characterization of \textit{ex vivo} aggressive prostate cancer using ultrasound RF time series data. In a leave-one-patient-out cross-validation strategy, the thesis demonstrates that the proposed features of RF time series from a group of patients can be utilized to accurately characterize higher grade tumors in never-before-seen patients.

- \textit{in vivo} characterization of aggressive prostate cancer using ultrasound RF time series data. The thesis demonstrated that the combination of the proposed RF time series features with the previously reported features can efficiently characterize aggressive prostate cancer.
1.2. THESIS OUTLINE

- evaluation of the robustness of SW calibration routine as a possible alternative for the current calibration techniques.

- integration of enhanced SW calibration technique within PLUS library.

1.2 Thesis Outline

This thesis is organized as follows:

Chapter 2, Background: provides a review on the current prostate cancer diagnosis and prognosis approaches using ultrasound imaging. The advantages and disadvantages of each of these methods are highlighted. It also presents a review of the calibration of 2D ultrasound probes.

Chapter 3, Characterization of aggressive prostate cancer, \textit{ex vivo} clinical study: describes the application of RF time series to characterize aggressive prostate cancer for the first time in \textit{ex vivo} prostate specimens. It also details the methodology of RF time series, the proposed novel features, the cross-validation experiments and the generation of aggressive cancer likelihood maps.

Chapter 4, Characterization of aggressive prostate cancer, \textit{ex vivo} and \textit{in vivo} clinical study: demonstrates a novel data collection approach from prostate specimens in both \textit{ex vivo} and \textit{in vivo} environments. Here, a validation study is performed to evaluate the tissue typing capabilities of RF time series in the \textit{ex vivo} prostate specimens. Also for the first time, RF time series has been used to differentiate higher grade and lower grade prostate cancer in \textit{in vivo} prostate specimens prior to the prostatectomy.

Chapter 5, Single Wall Calibration: Methods and Experimental Results: represents the implementation and experimental evaluation of the enhanced
SW calibration technique. It also discusses the robustness of this method and evaluates its reproducibility.

Chapter 6, Conclusion and Future Work: presents the major outcomes of this thesis and describes possible future work.
Chapter 2

Background

This chapter presents an introduction to diagnosis of prostate cancer and reviews the most recent prognosis approaches for detecting prostate cancer using ultrasound imaging modality. In section 2.2 the available ultrasound-based methods for detecting and grading of prostate cancer are reviewed and the shortcomings and advantages of each of these techniques are highlighted. Finally, an introduction to calibration of ultrasound probes as an inevitable prerequisite in almost all of the ultrasound-based intervention techniques is presented in section 2.3.

2.1 Prostate cancer

2.1.1 Prostate anatomy

The prostate is a walnut-sized gland in the male urinary and reproductive system located between the bladder and the penis, just in front of the rectum (Figure 2.1). Prostate cancer is the most common cancer diagnosed in north American men, excluding skin cancers [31], [47]. It is estimated that in 2014, approximately 233,000 new cases of prostate cancer will occur in North America [47]. On average 1 out of 7
2.1. PROSTATE CANCER

Figure 2.1: Prostate is a gland of male reproductive system. It is located just below the bladder and in front of the rectum. Reference: http://www.faithandhealthconnection.org/category/prostate-cancer/

North American men will develop prostate cancer during their lifetime and 1 out of 36 will die of it [87]. This disease accounts for 10% of male cancer-related deaths in North America [47].

2.1.2 Grading and staging of prostate cancer

One approach to assess the aggressiveness of a prostate cancer is to use the Gleason scoring system. This system assigns a grade between of 1 and 5 to a cancerous area of the prostate based on its similarity to normal cells, with 1 being the most similar to normal cells. Since usually more than one pattern of cancer is present in the prostate, a primary and a secondary grade are assigned to the most prevalent and the second
most prevalent pattern of cancer in the tissue. The summation of these two grades is referred to as a Gleason score spanning the range from 2 to 10 [90].

The stage of prostate cancer is defined based on the biopsy results (which includes the Gleason scores) and the prostate specific antigen (PSA) level [6]. The cancer stage shows the advancement of the cancer. The most ubiquitous standard for staging of prostate cancer is the American Joint Committee on Cancer (AJCC) TNM system which demonstrates how far the cancer has spread. This system assesses the extent of the primary tumor (T category), the presence of the advancement of the cancer to nearby lymph nodes (N category) and the presence or absence of distant metastasis (M category) [6].

2.1.3 Active Surveillance

An important factor in selecting the best treatment option for prostate cancer is the aggressiveness and the stage of the tumor. Cancers with Gleason scores ≤ 6 with no primary or secondary Gleason grades of 4 or 5, PSA levels ≤ 10 ng/ml and cancer stage of T1c (cancer found in needle biopsy) to T2a (cancer is spread in only less than one half of the prostate) are considered as low-risk or indolent prostate cancers [50]. The treatment option for patients with indolent prostate cancers is to undergo surveillance [50]. AS include monitoring the PSA level of the patients and performing DRE and re-biopsy on a regular basis. Invasive treatment and its side effects are postponed. This is the definitive clinical management of prostate cancer for majority of patients with low-risk cancer [50].

Hence, AS increases the quality of life in patients with non-aggressive prostate cancer. A 10 years survival rate has been reported for 99% of patients with indolent
2.1. PROSTATE CANCER

cancers [81], [67]. Factors such as doubling of PSA levels of the patients in less than 3 years, Gleason scores $\geq 4+3$ in the biopsy results and the cancer stage of T3 and above will abort the AS and initiate the interventions [50]. Accurate grading and staging of prostate cancer therefore helps avoid over-treatment and under-treatment of the disease.

2.1.4 Diagnosis of prostate cancer

As mentioned before, early diagnosis of prostate cancer and accurate staging and determination of the Gleason score will assist with the clinical decision making, including choosing the best treatment option such as radiation therapy, brachytherapy, cryotherapy, thermal ablation and surgery [23]. Over-diagnosis occurs when low grade, low stage cancer is diagnosed as an aggressive cancer which can result in harmful or even unnecessary treatment. Under-diagnosis of prostate cancer is also an important issue when the existed aggressive cancer is not diagnosed which will have the inevitable consequence of the advancement of the disease.

The current approach for diagnosis of prostate cancer is based on the PSA, digital rectal examination (DRE) and ultimately the biopsy of the prostate. PSA is a protein produced by the cells of the prostate gland. It is present in small quantities in the serum of normal men, and is often elevated in the presence of prostate cancer and in other prostate disorders. Usually in patients with prostate cancer, the PSA level is elevated to over 4 ng/ml. Detecting the existence of prostate cancer by measuring PSA alone suffers from both false negatives and false positives [111]. A PSA test is usually accompanied by a DRE test which evaluates the stiffness of the cancerous tissue but is limited to superficial large palpable tumors [52].
2.1. PROSTATE CANCER

The gold standard for detecting prostate cancer is the core needle biopsy which is performed under transrectal ultrasound (TRUS) guidance and is followed by histopathological analysis of the extracted cores. Prostate cancer is multifocal and does not develop tumors like other cancers; as a result, there is a high chance of missing the cancerous area during the biopsy. The sensitivity of a TRUS-guided needle biopsy is reported to be between 40% and 60% in multiple different studies [75]. Since TRUS is not capable of distinguishing between normal and cancerous tissues with high sensitivity and specificity, a prostate biopsy is performed at predefined standard anatomical locations. This makes the biopsy procedure blind to alternative cancer locations in the tissue. Overestimation and underestimation of prostate cancer are the consequences of blind biopsy. Several studies reported that the Gleason score archived by radical prostatectomy is underestimated in 19% to 57% of cases and overestimated 5% to 50% in of cases [32], [34], [35].

The high rate of false negatives during the TRUS-guided biopsy, exposes the patients to a repeated biopsy process and of course allows the advancement of the cancer during this time. Therefore, there is a need for additional guidance that can help physicians during biopsies. The final goal of computer aided biopsy techniques are to provide the cancer distribution maps using information from imaging modalities. These maps indicate the areas with high probability of cancer, so the physicians can take samples from those areas in addition to standard protocols. As a result, computer-based methods can help to further improve the biopsies by replacing the blind biopsy with guided biopsy.
2.2 Prognosis of prostate cancer

Diagnosis of prostate cancer is important; however, not all of the diagnosed cancers are life-threatening. Hence, assessment of the aggressiveness of the detected tumors is crucial to avoid over-treatment of low-risk or indolent cancers. Using the current techniques for prostate cancer treatment, for every man saved from prostate-cancer related death, 1400 men are screened and 48 men undergo radical treatment [80]. One of the biggest challenges in the clinical management of patients with newly diagnosed prostate cancer is to shift from detecting the cancer alone to characterizing its aggressiveness [103]. Patients with clinically significant prostate cancer would benefit more from radical treatment options compared to patients with indolent cancers.

Clinically significant prostate cancer is refereed to a state where immediate treatment is necessary. The Gleason score of the tumor, the staging of the tumor and the microvessel density of the cancerous lesions are good prognostic factors in determining the aggressiveness of the cancer [12], [22]. It has been reported that the mortality rate and advancement of prostate cancers with Gleason score 6 is much lower than that of Gleason scores $\geq 7$ [61], [94]. Gleason scores $\geq 6$ tumors have a continuously increasing fatality rate for the period of up to 15 years after conservative management of the disease [4].

Several techniques have been proposed to augment diagnosis of prostate cancer by performing the targeted biopsy procedure using different imaging modalities such as magnetic resonance (MR) imaging [9] and ultrasound imaging [65]. However, in the literature there exist limited studies to determine the prognosis of prostate cancer. The Gleason scores of detected cancerous lesions utilizing imaging modalities such as ultrasound and MR are reported and conclusions are drawn between the correlation
of the accuracy of cancer detection and its aggressiveness. Yet, an extensive study to differentiate between multiple aggressiveness levels of prostate cancer is lacking in the literature.

The following subchapter reviews the state of the art advancements in diagnosis and prognosis of prostate cancer using MR and ultrasound imaging. The main focus of this thesis is on ultrasound-based prognosis of prostate cancer. Section 2.1.4 reviews the ultrasound physics and the image formation techniques used in ultrasound scanners. Thereafter, the state of the art ultrasound-based methods for diagnosis and prognosis of prostate cancer are reviewed in section 2.3.

MR imaging has also been utilized for characterization of aggressiveness of prostate cancer. Morphological MR imaging such as T1-weighted imaging (T1WI) and T2-weighted imaging (T2WI) suffer from lack of high sensitivity and specificity to detect prostate cancer [109]. Hence, they are accompanied by functional MR imaging to detect the physiological or molecular parameters associated with cancer-related abnormalities. Diffusion-weighted imaging (DWI), dynamic contrast-enhanced (DCE) MR imaging and MR spectroscopic imaging (MRSI) are three functional imaging parameters that are utilized in combination with morphological MR imaging to characterize aggressiveness of prostate cancer.

DWI examines the diffusivity of water molecules in the tissues [26]. The net displacements of water molecules is represented by apparent diffusion coefficient (ADC) parameter. Since the cancerous lesions have higher cellular density they have lower ADC values compared to normal tissues [112]. Tamada et al. reported a significant inverse correlation between the ADC values and the Gleason scores of the cancerous foci [98]. Similarly the mean ADC values of tumors with Gleason scores ≥ 7 have
been shown to be significantly lower than those with Gleason scores $\leq 6$ [27]. In DCE-MR imaging a contrast agent is administered into the vessels. The washout time of this contrast agent from the capillaries changes the T1WI relaxation time [40]. Compared with normal prostate tissues, cancerous tissues are denser and have more permeability. Hence, the washout time of contrast agents in cancerous tissues is lower than in normal tissues. Inconsistent results for the correlation of DCE-MR imaging and the Gleason scores of the prostate cancer are reported [46], [55], [78], [79]. MRSI reveals biochemical information of the cancerous tissues by assessing the levels of metabolites such as choline, creatine and citrate [26]. It has been shown that choline and citrate signals increase and decrease in cancerous prostate tissues, respectively [26]. Hence, the ratios of choline to citrate or (choline + creatine) to citrate are representative of prostate cancer. Zakian et al. reported that MRSI has a higher sensitivity of 89.5% to detect cancers with Gleason scores $\geq 8$ compared to a sensitivity of 44.4% for Gleason scores $\leq 6$ [110]. Moreover a trend has been observed between the metabolic ratios and the tumor grade.

Although MR imaging can provide a high resolution and clear depiction of the borders of organs, it suffers from the lack of ability to detect the cancer tumors that have a diagonal length of less than 1 cm or are lower grade [71]. Also studies have shown that the diagnostic value of MR is not superior to ultrasound [114]. Moreover, MRI-guided biopsy is expensive, difficult to perform and not widespread [82]. Fusion of pre-operative MR images of the prostate with intra-operative ultrasound data is used to overcome these challenges [59]. This procedure needs registration of the biopsy core locations determined in the pre-operative MR images, to the 3D TransRectal UltraSound (TRUS). The registration is performed either by software
only [107] or by using a sophisticated mechanical system [8]. Compared to other imaging modalities, the advantages of an ultrasound-based prostate cancer detection system are several folds: TRUS is already accepted as the standard prostate biopsy guidance tool; different ultrasound data and image acquisition modalities are simultaneously available on ultrasound machines, and ultrasound imaging is among the most accessible and least harmful medical imaging approaches.

2.3 Ultrasound-based characterization of prostate cancer and its aggressiveness

2.3.1 Ultrasound physics and image formation

An ultrasound scanner sends and receives ultrasound waves to the underlying tissue through a set of embedded piezoelectric crystals. The task of the transducer is to transform the electrical signal to ultrasound beam and also to convert the returning ultrasound echoes to electrical signals. Commercial medical ultrasound transducers for abdominal and prostate imagery generally operate at frequency ranges of 2 - 14 MHz. The ultrasound beams generated by these machines undergo different interactions with the structural elements of the tissue. The main interactions are the scattering and the absorption of the ultrasound signal. The net result of each interaction mechanism with the signal is a function of the ultrasound frequency, temperature, and the characteristics of the tissue. Therefore, the ultrasound echo signals, which, due to the similarity of their frequency range to radio-frequency waves are called RF echo signals, contain valuable information about the physical properties of the tissue. To generate the RF signals, the received analogue signal at the transducer are filtered and then passed through an analog-to-digital converter. Afterwards, they
go thorough a process of applying time delays and weighting amplitude refereed to as beam-forming [106]. Beam-forming is used to enhance the quality of ultrasonic images. To form the B-mode images, the RF signals are passed through an envelop detection followed by log-compression and further post-processing. Since the envelop detection process decreases the frequency bandwidth of the signal, some of the tissue characterizing parameters can only be extracted from the RF echo signals before they go through the envelop-detection process [16].

Due to the multiple advantages of ultrasound imaging modalities, ultrasound-based methods for augmenting prostate cancer detection and prognosis have received much attention during the past three decades. As mentioned previously, the net result of ultrasound beam interaction with the tissue is a function of multiple parameters including the tissue properties. Consequently the ultrasound data (including the RF echo signals and the B-mode images) contain valuable tissue type information that can be utilized for tissue discriminations and tissue characterization. Ultrasound-based methods for detecting the prostate cancer fall into five main categories: i) texture-based; ii) spectral based; iii) elastography; iv) Doppler ultrasound and v) RF time series analysis. Texture-based methods utilizes B-mode images, however, spectral-base, elastography and time series are based on using the RF ultrasound signals. Doppler ultrasound visualizes the frequency shifts of the ultrasound echoes. To the best of our knowledge, texture-based and spectral-based ultrasound approaches were utilized for prostate cancer diagnosis only and characterizing aggressiveness of prostate cancer using these approaches is lacking in the literature.
2.3. ULTRASOUND-BASED CHARACTERIZATION OF PROSTATE CANCER AND ITS AGGRESSIVENESS

2.3.2 Texture-based methods for prostate cancer diagnosis

Among the ultrasound-based prostate cancer detection techniques in the literature, there are methods using B-mode images to detect the cancer. Prognosis of prostate cancer using texture-based approaches is lacking in the literature. Prostate tumor usually appears in the form of hyper-echoic, iso-echoic, hypo-echoic or mixed-echogenecity area in the ultrasound images of the prostate [63], [96]. It has been shown that the speckle pattern in an ultrasonic image can carry tissue dependent information. Features representing the first order statistical moments, such as: mean, skewness, standard deviation and kurtosis were utilized to differentiate between multiple tissue types [41]. Houston et al. [41] conducted a study to investigate the performance of first order statistical moments of Bscan images to localize prostate cancer lesions in a group of 25 patients. In their study, the object image is divided into region of interests (ROIs) of sizes between 0.1 cm and 1.45 cm. It was reported that the noncancerous ROIs have a lower mean intensity feature compared to cancerous regions.

However, using the first order statistical moments, suffers from three drawbacks: i) It has been shown that as the scatterer density increases in an ultrasound image, the signal to noise (SNR) ratio reaches a saturation level. This indicates the shortcomings of first-order statistical moments in characterizing the high density tissues. Albeit, in many biological tissues the scatterer density will not exceed the saturation level. ii) These features are bound to change in the case of using different imaging parameters in ultrasound scanners such as different gain adjustments [7]. iii) It might be possible that tissues with the same pathology demonstrate different acoustical properties. For instance, although most prostate cancerous lesions appear as hypo-echoic areas in the ultrasound images, however, in some cases the tumors are appeared as hyper-echoic
or even iso-echoic.

Due to the limitations of the first order statistical moments to differentiate between different tissues, researchers were prompted to use the second order statistical moments as a tissue characterizing tool. Second-order features that code the spatial properties of the image were proposed such as co-occurrence matrices. Co-occurrence matrices are basically the histogram representation of the image that demonstrate the co-occurrence of a pair of intensities [7], [63], [76]. Although the texture features are valuable tissue characterizing parameters, due to the limitation of texture-based tissue typing methods such as dependency to imaging parameters, texture-based features usually are not solely used for tissue typing and are combined with other features extracted from different tissue typing approaches [73], [77].

2.3.3 Spectral-based methods for prostate cancer diagnosis

Using RF signals as a tissue characterizing tool was first introduced in the late 1960s. It was shown that frequency dependence of backscattering when depositing ultrasound beams can be utilized to derive tissue typing parameters from the RF signals [58]. Spectral-based methods were utilized for diagnosis of prostate cancer and prognosis of prostate cancer using spectral-based methods is lacking in the literature. The most popular tissue typing features extracted from the RF signals are those representing the scattering and attenuation of ultrasound beam in the tissue. The features such as axis-intercept, slope, mid-band values and integrated power spectrum which are extracted from a linear regression line fitted to the frequency spectrum of the RF signal, represent the most popular attenuation and backscattering parameters. Several groups have investigated the application of spectral analysis on characterizing
prostate cancer. Based on a study by Fleppa et al., features extracted from the normalized power spectrum of a RF echo signal are powerful tissue typing features [29]. In another study of spectral analysis conducted by Vlad et al. [92], prostate tumor responses to radiotherapy were assessed utilizing the ultrasound signals at 10-30 MHz central frequencies. It was shown that the ultrasound-integrated backscatter is related to the acoustic impedance of the tissue as well as the concentration and scatterer size. In several studies, the RF and Bscan features have been combined to obtain the optimal ultrasound-based feature vector for prostate cancer detection.

One of the limitations of spectral-based approaches for prostate cancer detection, is the need for a standard method for calibration of ultrasound RF signals. To partly address this issue, a high reflector such as glass, steel or a wire phantom are scanned at different depths. The high reflector causes high amplitude echo signals which ruins the ultrasound input image. Wire phantoms also make the calibration process dependent on the geometry of the wires [92].

2.3.4 Elastography-based methods for diagnosis and prognosis of prostate cancer

Ultrasound based methods using elastography for diagnosis [91] and prognosis [15] of prostate cancer, have been an active field of study in the past decades. This method estimates the mechanical properties of the tissue by tissue compression obtained either by an external mechanical force or by an acoustic radiation [113]. Real-time elastography (RTE) is a technique that has substantially improved recently, and enables characterizing the areas of higher cell density [53]. Since cancerous tissues have higher cell density hence are stiffer than the normal tissues their displacement as a
result of an external mechanical force is less compared to normal tissues [28], [53]. Tissue displacement is measured by using the RF echo signals before and after excitation [93].

Several researchers have used elastography for prostate cancer detection. Krouskp et al. [54] evaluated the elastic properties of the prostate specimens by applying a manual excitation force. These properties were further utilized by Konig et al. to perform targeted prostate biopsy [51]. The limitation of these studies were that the excitation process was operator dependant. Salcudean et al. designed an operator independent vibro-elastography (a dynamic ultrasound elastography) approach [91] for detecting the stiffness of the tissue using an ultrasound transrectal probe. It has been reported that ultrasound elastography has a higher specificity rate of cancer detection compared to DRE [62]. Sonoelastography is also another approach that has been used in the literature for prostate cancer detection [18]. This approach uses low frequency shear waves for tissue displacements estimation [99]. Sonoelastography has been shown to be more accurate and sensitive than conventional ultrasound approach for prostate cancer detection in 10 prostatectomy specimens [89]. In a recent study by Castaneda et al., three dimensional sonoelastography was used for prostate cancer detection. An accuracy of 80% was achieved for detecting the tumor volumes larger than 4 mm in 11 ex vivo and in vivo specimens [17].

Marko Brock et al. compared RTE versus texture-based methods on prostate cancer detection. Their results show that cancerous lesions with Gleason score of \( \leq 6 \), 7, > 7 are detected with accuracies of 59.9%, 55.3% and 67.7% for RTE and 45.3%, 43.5% and 42.5% for GSU. RTE significantly outperforms texture-based methods in detecting the cancerous foci; tumors with higher Gleason score were more likely to be
identified using RTE [15]. Ahn et al. reported a reverse correlation between the tumor elasticity and Gleason score [3]. Similar achievements were reported by Zhu et al [115]. They have observed that using RTE, the detection rate for Gleason scores > 7 was 80.4% whereas for Gleason scores < 7, an accuracy of 42.6% was obtained. While most of the studies reported that the detection rate of RTE increases for more aggressive prostate cancers with higher Gleason scores, Tsutsumi et al. observed higher detection rate for low-stage tumors with Gleason scores \( \leq 3+3 \) [101]. Recently, Xu et al. also reported no significant difference in the detection rate of Gleason scores < 7 and that for Gleason scores \( \geq 7 \) using RTE [36].

Ultrasound elastography has been widely used for tissue typing purposes. Yet, the clinical application of this approach is limited due to the need for modifying the ultrasound machine hardware. Modification is performed to obtain a relatively strong acoustic radiation force. In the case of external excitation, additional external hardware is also needed. Studies have reported higher detection rate for more aggressive cancers using RTE but the results are not consistent. In addition, to the best of our knowledge, there has not been any work directly on prediction of high grade cancer.

2.3.5 Color Doppler methods for prostate cancer diagnosis and prognosis

Cancer development is associated with an increase in the uptake of nutrients in the blood vessels of the cancerous lesions. This leads to a dense microvascular network formation referred to as "angiogenesis". It has been reported that angiogenesis correlates with the aggressiveness of the prostate cancer [105] and is one of the key factors for cancer growth greater than 1 \( mm^3 \) [14]. Hence, monitoring the blood flow at the capillary level has been a subject of attention. For this purpose, ultrasound Doppler
imaging is utilized. Doppler imaging operates based on the fact that the received ultrasound echoes from a moving object will undergo a frequency shift which is proportional to the object’s speed [13]. A color flow diagram is overlaid on the B-mode image to visualize the frequency shift. Alternatively, the total power of the doppler signal from a volume can be measured and displayed which is referred to as a power doppler image [63].

A wide range of tissue typing features is extracted from the power doppler images and color flow doppler images. Among these are features such as power doppler pixel density and speed weighted pixel density (SWD) [66]. Potdevin et al. conducted a study on detecting the prostate cancer in the peripheral and periurethral regions. Their results show that speed and SWD can be used for tissue characterizing of benign and malignant lesions in ROIs as small as 1 mm². An area under the receiver operating curve (ROC) of 0.8 was reported [83]. Recently, Boukadoum et al. evaluated the correlation between Gleason score of detected prostate tumors and the hypervascularity using power doppler ultrasound. Their study included a cohort of 105 patients. They have reported that the power doppler ultrasound might help guiding the biopsy process toward more aggressive cancer foci [13]. However, conflicting results are reported regarding the correlation of ultrasound Doppler flow and the aggressiveness of prostate cancer. Kahraman et al. reported no significant relationship between Gleason score and the flow grading of power Doppler [49]. The neovascularization associated with the formation of prostate tumors usually occurs at the microvessel levels which may limit the application of doppler ultrasound in the prostate cancer detection [66]. To address this issue, ultrasound contrast enhanced materials were utilized to increase the intensity of the scattered signals from blood
Figure 2.2: RF time series is collected from a fixed location of the tissue over time cells. Frauscher \textit{et al.} showed that utilizing microbubble ultrasound contrast agents will improve prostate cancer detection [33]. Mitterberger \textit{et al.} [61] compared the performance of contrast enhanced color Doppler (CECD) with the conventional systematic biopsy on detecting prostate cancer lesions. Out of 690 men recruited in their study, conventional systematic biopsy and CECD could detect cancer in 24% and 26% of the patients, respectively. The mean Gleason scores of cancers using conventional systematic biopsy and CECD were reported to be 5.4 and 6.8, respectively. Therefore, CECD detected more aggressive prostate cancers compared to conventional systematic biopsy. Nelson \textit{et al.}, utilized color Doppler imaging for prostate cancer detection [72]. It has been reported that the abnormal color flow strongly correlates with Gleason score of 8 or 9 and is not associated with lower grade tumors.
2.3.6 Time series-based methods for prostate cancer diagnosis and prognosis

The concept of using RF time series as a tissue characterization tool was proposed by Moradi et al. [65]. It was shown that a series of ultrasound radio frequency (RF) signals collected in time from a fixed tissue location (referred to RF time series), carry tissue typing information. RF time series has been shown to be an effective tissue characterizing tool in both high frequency (20-60 MHz) and low frequency (2-10 MHz) [64]. RF time series tissue typing approaches were used to successfully differentiate chicken breast, pig liver, bovine liver and bovine muscle [64].

Moradi et al. showed that the frequency domain features of RF time series are promising tissue discriminators. These features were the integral of the power spectrum in the four quarters of the frequency band of the signal, the slope and intercept of the regression line fitted to the power spectrum and the fractal dimension (FD) of the time series obtained using Higuichi’s approach [65].

RF time series were applied to successfully differentiate between cancerous and normal tissues in 35 ex vivo prostate specimens. Using the combination of these features, an accuracy of 87% and an area under the ROC of 0.82 in the leave-one-patient-out cross validation were reported [65]. The results demonstrate that the RF time series features differentiate between cancerous and normal tissues with significantly higher sensitivity and specificity compared to texture-based and spectral-based features. Abofazeli et al. proposed novel features of a RF time series signal. These features include the wavelet transform approximation and the sequences of RF time series [1]. High classification accuracies were reported using the combination of these features with the features proposed by Moradi et al., however, the results show that
the wavelet features are not an alternatives for Moradi et al. features [65] to detect prostate cancer.

The first in vivo clinical study to validate the tissue typing capabilities of RF time series was also conducted by Moradi et al. for a cohort of six patients. In the leave-one-patient-out cross-validation, an average area under ROC curve (AUC) of 0.76 using RF time series features was reported compared to AUC of 0.66 for using texture-based features. Imani et al. proposed novel features of RF time series [44]. These features were the mean central frequency and wavelet features of the signal. Using the two proposed features in a cohort of seven in vivo prostate specimens a classification accuracy and AUC of 78% and 0.83, respectively were reported in the leave-one-patient-out cross validation [44]. Recently, N. Uniyal also conducted a study for detection of prostate cancer in 18 MRI selected biopsy targets using ultrasound RF time series [102]. In a leave-one-patient out cross-validation, an area under ROC curve of 0.91 was reported for identification of prostate cancer in ROIs as small as 1 mm × 1 mm.

Based on the recent studies in the literature, RF time series is a powerful tissue typing approach that significantly outperforms other conventional ultrasound-based methods for diagnosis of prostate cancer.

In this thesis, the performance of RF time series on characterizing aggressive from non-aggressive prostate cancer is investigated by differentiating between higher and lower grade cancer. Two ex vivo and one in vivo studies are performed.
2.4 Calibration

Compared to other volumetric imaging modalities such as MR and computed tomography (CT), conventional ultrasound imaging is a 2D imaging modality. The advantages of a 3D ultrasound over 2D ultrasound are several folds: it provides direct visualization of the 3D organs; allows for generating 2D image slices at any arbitrary orientation and provides more precise volume measurements [42]. 3D ultrasound has been used for prostate volume measurement [24] and prostate cancer radiation therapy [43]. Moreover, it has been reported that 3D ultrasound enhances ultrasound-guided prostate biopsy by providing real time visualization of the organs [97], [37].

3D ultrasound volumes can be constructed either using a 3D probe or by collecting a series of tracked 2D ultrasound images and then constructing the 3D volume [60]. The second approach is refereed to as freehand ultrasound data acquisition. Compared to the 3D probe, in free hand 3D ultrasound, a fairly large volume can be scanned. More importantly, the subject being scanned can be located in an externally fixed coordinate system. This is particularly useful in applications such as radiotherapy [60]. The freehand systems generally consist of a position sensor that is attached to the ultrasound 2D probe to track the position of the probe at any time point. However, having the position of the ultrasound probes does not necessarily determine the position of the imaging plane. The position of each pixel of the scanning plane is related to the position sensor mounted on the probe through a process refereed to as probe calibration. Probe calibration is important for accurate 3D ultrasound reconstruction. Also, it is necessary for intra-operative navigation of the internal anatomy and/or surgical tools [97]. In many of the ultrasound-guided interventions such as prostate biopsy [97], breast biopsy [30], gynaecology [5], fetal cardiology and
surgery [108], probe calibration is essential. Accurate probe calibration is a key factor in the accuracy of ultrasound-based interventions. The clinician relies on the precise orientation of the ultrasound volume and surgical tools provided by probe calibration.

Several calibration techniques have been introduced in the past two decades [19], [104]. One typically precise approach for ultrasound probe calibration is to image a geometrically known object referred to as a Phantom. The calibration problem is solved by combining the prior knowledge of the physical locations of phantom features in the reference coordinate system and their corresponding locations in the ultrasound images. Calibration phantoms generally consist of strings [10], [11] and planes [86], [88]. The intersection of the imaging plane with strings and planes would appear as points and lines in ultrasound images, respectively. Point-based phantoms such as single point [95] and single cross wire phantoms [100] were early phantoms utilized for calibration of ultrasound images. In single point phantoms, a physical point such as a spherical object is imaged and in single cross-wire phantoms, intersection of two wires is imaged at multiple different angles. Trobaugh et al proposed a multiple cross-wire phantom that consist of multiple string crossings [100]. Recently, Chen et al. also proposed a multi N-wire approach for calibration of ultrasound images [20].

The drawbacks of point-based calibration are that it is difficult to focus the ultrasound beam at a single point; also, the points in the ultrasound images should be segmented manually which makes the approach operator dependent [43]. Other types of phantoms include plane-based phantoms where a plane is imaged instead of points. Here, the phantom features to be segmented are lines which are the intersection of the phantom plane with the imaging plane. The Single Wall (SW) phantom [84] a flat plane immersed into a water tank, and the Cambridge phantom [85] are examples
of plane-based phantoms. The plane-phantoms take advantage of the fact that lines in an ultrasound image can be automatically segmented [43]. In particular, SW is composed of a single flat plate that makes it attractive as a phantom. In addition to automatic feature selection (automatic line detection), measuring differences of features such as the slope of the line fitted to a set of points instead of absolute positioning of points is less prone to errors and hence leads to higher calibration accuracy [69].

Recently, a novel calibration method was proposed by Najafi et al. [69], [68] that takes advantage of using both automatic feature selection along with differential measurements. Here, a closed-form differential approach was proposed to calibrate ultrasound 2D probes based on the conventional SW phantom. Unlike the absolute positioning of the points in the conventional SW technique, here the slope of the line is utilized as a differential measurement to enhance the accuracy of line detection, hence calibration.

In this thesis an extensive experimental evaluation of the proposed method by Najafi et al. is performed. This method was integrated within PLUS library. Afterwards, a series of extensive data collections was conducted. The goal here was to evaluate this approach as a center-independent and user-independent calibration routine that can be used across multi-centers.

### 2.5 Chapter summary and conclusion

Being the second most common cancer in north American men, prostate cancer accounts for 10% of male cancer-related deaths in North America. Early diagnosis of
prostate cancer depends on PSA test followed by DRE. The gold standard for detecting prostate cancer is core needle biopsy under ultrasound guidance and is followed by histopathological analysis of the extracted cores. Diagnosis of prostate cancer is important; however, over-diagnosis of indolent cancers is a critical problem since it yields unfavorable and unnecessary treatment of low-risk cancers at no benefit. Assessing the aggressiveness of the cancer helps avoid over-treatment of indolent or low-risk cancers. Due to the advantages of ultrasound imaging such as being relatively cheap, non-ionizing and real-time, ultrasound-based techniques for prostate cancer diagnosis and prognosis have had attention during the past three decades.

Texture-based approaches utilize the speckle pattern in B-mode images for cancer detection. These methods are sensitive to the imaging parameters and are usually not used for tissue characterization on their own. Spectral-based methods show promising results for characterizing malignant from benign prostate tissue. They require a standard method for calibration of ultrasound RF signals. Texture-based and spectral-based methods were only utilized for diagnosis of prostate cancer and no studies to evaluate their performance for prognosis of prostate cancer have been done. Elastagrophy-based methods are based on the fact that cancerous lesions are stiffer than the normal tissues. The mechanical properties of the tissues are calculated and used for characterization. This method requires an external excitation source and modification to the ultrasound hardware. While some studies here reported that the detection rate of cancer using RTE increases for more aggressive prostate cancers with higher Gleason scores, there have been conflicting results showing higher detection rates for low-stage tumors. Doppler imaging is another ultrasound-based technique used for prostate cancer detection. Cancer development is associated with an increase
in the uptake of nutrients in the blood vessels of the cancerous lesions which leads to angiogenesis. Using the fact that ultrasound echoes from a moving object will undergo a frequency shift which is proportional to the objects speed, ultrasound Doppler visualizes the angiogenesis in the cancerous foci. The Neovascularization associated with the formation of prostate tumors usually occurs at the microvessels. Hence, ultrasound contrast enhanced materials were utilized to increase the intensity of the scattered signals from blood cells. Conflicting results on the sensitivity of ultrasound Doppler to aggressive prostate cancer are reported.

One of the other ultrasound-based methods for prostate cancer diagnosis and prognosis is the RF time series approach. It was shown that the time series of ultrasound echoes from a fixed location of the tissue contain tissue typing information [65]. RF time series was shown to outperform other conventional ultrasound-based methods for prostate cancer diagnosis. The performance of RF time series on prognosis of prostate cancer was not previously sought. Here we use RF time series to differentiate between lower and higher grade prostate cancer.

Probe calibration is an inevitable prerequisite in almost all of the ultrasound-guided interventions such as prostate biopsy. Probe calibration registers the ultrasound imaging plane to the surgical tools and external coordinate systems. In this thesis the performance of SW calibration techniques proposed by Najafi et al. [69] was also evaluated. Compared to other calibration methods such as multi N-wire, SW takes advantage of automatic feature selection and the fact that an SW phantom is easy to build.
Chapter 3

Characterization of aggressive prostate cancer, \textit{ex vivo} clinical study

This chapter presents the methodology for characterization of aggressive prostate cancer using ultrasound RF time series. We utilize the information that exists in the RF time series to characterize higher grade prostate cancer in a cohort of 15 \textit{ex vivo} prostate specimens. In addition, novel frequency domain features of the RF time series are proposed. We demonstrate that a non-linear classifier that is built using the RF time series information of a group of patients can successfully predict higher grade prostate cancer in the never-before-seen patient.

3.0.1 Data collection

Previously, ultrasound and histopathologic images were collected from a cohort of 35 \textit{ex vivo} specimens of human prostate tissues from patients that chose prostatectomy as their treatment option at Kingston General Hospital [65]. Patients’ consents were acquired to perform ultrasound scanning on extracted prostate specimens prior to their participation in the study. Following surgery, prostate specimens were fixed and
suspended in a water bath and scanned (Figure 3.1 a, b). Ultrasound scanning was performed along transverse planes at 4 mm apart. The location of the first imaging plane corresponded to the first plane of the histopathology dissection, marked by two needles that were visible as two lines in the ultrasound image. These needles were used to ensure the quality of manual registration between ultrasound and pathology images. Consequent images were collected along cross sections with 4 mm intervals. The needles were only used to ensure the location of first cross section and in fact due to the strong echoes from the needles, the data from first frame was practically not used. Depending on the size of the prostate, four to six consequent cross sections of data were collected from each specimen.

A Sonix RP (Ultrasonix Inc., Richmond, BC, Canada) ultrasound scanner with a transrectal probe (BPSL9-5/55/10) at central frequency of 6.6 MHz and maximum imaging depth of 4.5 cm was used to collect data. To form a time series of data, 112 frames at a frame rate of 22 fps at each cross section of the tissue were collected. Also, the sampling frequency of the RF data was 20 MHz. The water temperature was measured prior to data collection to ensure all specimens were scanned at the temperature range of 22-24 °c. The ultrasound probe was placed into the water with only one focal point at 1.5 cm depth.

Following ultrasound scanning, the tissue specimens were dissected along the scanning planes and the histopathological analysis of the whole mount slides was performed. Tumor regions in these images were contoured by a senior pathologist. A subgroup of 15 patients out of the original 35 that had primary cancerous areas larger than 5 mm × 5 mm in at least one of their dissected planes, were gridded into digitized ROIs of size 2.5 mm × 2.5 mm. Grids within the encircled cancer areas were
Figure 3.1: a) and b) setup used for data collection; c) a sample of a gridded pathology.

Further graded by the pathologist and the a Gleason score and estimated of its percentage was assigned to each grid. To avoid overlaps between multiple Gleason scores, a threshold level of 60% was considered for ROI selection. This means that ROIs are only segmented from digitized grids with more than 60% of a certain Gleason score. Based on this criterion, 3 patients out of the original 15 patients were not considered for ROI selection and the rest of the analysis in this chapter. These patients had an estimated Gleason score of less than 60% in their digitized grids of size $2.5 \text{ mm} \times 2.5 \text{ mm}$.

In this study, we aim to separate between cancerous tissue with Gleason score $\leq 3+4$ and that of Gleason score $\geq 3+4$. Figure 3.1, c. shows a sample of a gridded histopathology image.

3.0.2 ROI selection

The histology images were manually registered to RF frames using the 3D Slicer software (www.slicer.org). The registration process involved: matching the first histopathology cross section with the first imaging plane, and following thereafter
with consequent slices. Landmarks such as the urethra or the prostate boundary irregularities were used to ensure the correspondence of two images. There is bound to mis-registration between the histopathology images and tumor boundaries due to inevitable errors of tissue dissection and ultrasound beam width. To partly address this we only chose images with cancerous foci size of greater than $5 \text{ mm} \times 5 \text{ mm}$ in pathology images. Moreover, we only chose ROIs close to the center of the cancerous areas in the ultrasound images and avoided boundaries. Figure 3.2, a shows a sample of the registration result of pathology to ultrasound image as a snapshot of 3D Slicer software (www.slicer.org). Figure 3.2 b demonstrates the overlay of pathology Gleason scores on the corresponding ultrasound slice after performing the registration between pathology and ultrasound. Also Figure 3.2, c shows the gridded ultrasound image.

In total, 1440 ROIs of size $1 \text{ mm} \times 1 \text{ mm}$, were segmented from designated areas with different Gleason scores. Out of these, 604 ROIs were from cancerous areas with Gleason score $\leq 3+4$ and 836 ROIs were from areas with Gleason score $\geq 4+3$. Each ROI comprises a three dimensional signal with 26 axial, 5 lateral and 112 time samples (figure 3.3). Table 3.1 shows the total number of ROIs segmented from designated areas with different Gleason scores.

To extract the features of each ROI, first the mean of each RF time series was removed and it has been zero-padded to the closest power of 2. Following that its spectrum was calculated and the features of the spectrum as described below were extracted. Finally these features were averaged for any given ROI.
Figure 3.2: The registration of pathology and ultrasound data that was performed in the 3D Slicer software. a) pathology image registered to ultrasound image of the same cross-section; b) overlay of available Gleason scores on the ultrasound image and c) gridded ultrasound image for ROI selection

3.0.3 Feature extraction

We used 60 spectral features of the RF time series. Previously, Moradi et al. proposed novel frequency domain features of RF time series that included the summation of the spectrum in four quarters of the frequency bands and a subset of features that demonstrate the general appearance of the spectrum [65]. Here, to extract more information from the frequency domain of the RF time series, we divided the frequency band into 10 sections instead of four parts. The extracted features include: i) the summation of the spectrum in ten frequency bands (F1 to F10) following Discrete
Figure 3.3: Each ROI comprises a three dimensional signal that consist of 26 axial, 5 lateral and 112 time samples.

Fourier Transform (DFT) of the RF time series signal; ii) the intercept and slope of the regression line fitted to the spectrum (F11 and F12) [65] iii) the fractal dimension of time series computed using the Higuchi approach (F13) [65]; iv) the mean central frequency of the spectrum (F14) [44]; v) the first approximation coefficient following wavelet transformation (F15), and vi) the ratios of the summations of the spectrum in 10 frequency bands (F16 to F60), e.g., F16 is the ratio of the summation of the spectrum in the second frequency band to the summation of the spectrum in the first frequency band. The description of each feature is as follows:
Table 3.1: Total number of segmented ROIs from designated areas with different Gleason scores

<table>
<thead>
<tr>
<th>Gleason score</th>
<th>number of ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 3+3 6</td>
<td>168</td>
</tr>
<tr>
<td>GS 3+4 7</td>
<td>436</td>
</tr>
<tr>
<td>GS 4+3 7</td>
<td>232</td>
</tr>
<tr>
<td>GS 4+4 8</td>
<td>36</td>
</tr>
<tr>
<td>GS 4+5 9</td>
<td>380</td>
</tr>
<tr>
<td>GS 5+5 10</td>
<td>188</td>
</tr>
</tbody>
</table>

i) The mean of the signal was removed since we are only interested in the variation of the signal. The DFT of the zero-mean signal is taken. This transform is computed as:

\[
X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x_t[n] e^{-j \left( \frac{2\pi}{N} \right) kn}
\]  

(3.1)

Here \( X \) is the DFT of the time series \( x_t \) at frequencies \( k \). The DFT was zero padded to the next power of 2. DFT was computed using the Fast Fourier Transform algorithm implemented in MATLAB (Mathworks Inc.). To calculate the first 10 features for a given ROI, the frequency band of each RF time series was divided into 10 equal intervals. The summation of the amplitude of the Fourier transform at each interval was calculated, averaged and max-normalized over the 26 \( \times \) 5 RF time series in that ROI. Due to the symmetry, only positive frequency components were used which refer to the frequency range of 0-11 Hz. Equation 3.2 shows the calculation of these first 10 features:
\[ Feature(k) = \sum_{i=1+(k-1)\frac{N}{20}}^{\frac{N}{20}} X[k], k = 1, \ldots, 10 \] (3.2)

ii) We have also computed the intercept and slope of a regression line fitted to the spectrum values versus normalized frequency [65].

iii) Another feature extracted from the time series data was the fractal dimension of the time series computed using the Higuichi approach. This feature has previously been used for classification between cancerous and normal ROIs in 35 ex vivo prostate specimens [65]. Fractal dimension is a measure of the complexity of a signal and it is commonly used for biomedical signal analysis such as analysis of EEG signals [2], [38]. The implementation of Higuichi’s method to compute the fractal dimension is detailed elsewhere [2], [64], [65].

iv) Mean central frequency (MCF) of the RF time series is also calculated. MCF is in fact the averaged mean of the frequency of the RF time series [44]. Equation 3.3 shows the mathematical formulation of MCF.

\[ MCF = \frac{\sum_{BW} f \times SD(f)}{\sum_{BW} SD(f)} \] (3.3)

where BW is the bandwidth of the signal and SD is the average value of the spectrum density for each RF time series of an ROI. The MCF values of RF time series of an ROI are then averaged to generate the MCF value of each ROI.

v) To combine both time and frequency domain information, the wavelet transform of the RF signal is calculated using Daubechies-4 filter bank. Here, at each
decomposition level, high pass and low pass information of the signal were separated from each other while the low pass signal was the input to the next decomposition level. The frequency band of the output of each decomposition level is half of the band width of its input. The first approximation coefficient at the maximum level of decomposition was considered as feature F15. Wavelet and MCF features have been previously used as tissue typing features to classify cancerous and normal in vivo prostate specimens [44].

vi) We also proposed novel frequency domain features of RF time series for tissue typing. These features are the ratios of the summation of the spectrum in 10 frequency bands as below:

\[ Feature(k) = \frac{SS(n)}{SS(m)}, n = 1, \ldots, 10; m = 1, \ldots, n - 1, k = 16, \ldots, 60 \]  \hspace{1cm} (3.4)

where \( SS(n) \) is the summation of the spectrum in the \( n \)th frequency band (3.4).

### 3.0.4 Classification

For classification, a support vector machine (SVM) classifier is used to separate between lower grade and higher grade of prostate cancer. Here, lower and higher grade prostate cancers refer to ROIs with Gleason scores \( \leq 3+4 \) and \( \geq 4+3 \), respectively. The SVM classifier is a nonlinear classifier that uses a kernel function to map data into a higher dimension where they are linearly separable. This classifier finds the best hyperplane resulting in separating the two classes with the maximum margin. The optimization problems that SVM solves, given a set of training vectors: \( X_i; i = 1, \ldots, n \) with the true labels; \( Y_i \in \{-1, 1\} \), are:
Figure 3.4: A demonstration of the normalized spectrum of a sample ROI corresponding to Gleason score of 4+5 from a sample patient. The novel proposed features of RF time series are the ratios of the summation of the spectrum in 10 frequency bands. The probe frame rate was 22 fps which determines the principal frequency of RF time series.

\[ \min \left( \frac{1}{2} W^T W + C \sum_{i=1}^{n} \xi_i \right) \]  

\[ \text{subject to } Y_i(W^T \Phi(X_i) + b) \geq 1 - \xi_i \]  

In the equation 3.6, \( C > 0 \) is the penalty factor and controls the overfitting of the classifier to the training vector, and \( \xi_i \) is the slack variable which measures the misclassification degree of the training vector. The function \( \Phi(X_i) \) is the hyperplane that separates the data into two classes and \( K(X_i, X_k) = \Phi(X_i)^T \Phi(X_k) \) is called the kernel function.
The kernel function that has been used here is the radial basis function. The selection of radial basis was due to its ease of initialization and better classification accuracy as reported in the literature [44]. This function only has one variable to initialize, $\gamma$, as below:

$$K(X_i, X_k) = e^{-\gamma\|X_i - X_k\|^2}, \gamma > 0$$ (3.7)

We tune $C$ and $\gamma$ through an exhaustive search to maximize the classification accuracy in the test data set. We have followed the leave-one-patient-out cross-validation strategy. In this approach, the classifier is constructed using the features from $N - 1$ patients with known labels where $N$ is the total number of patients in the study, and is tested on the never-before-seen patient. To find the best subset of features resulting in the highest classification accuracy in the test data set, an exhaustive search is performed.

Moreover, in addition to the binary labels of the classifier, we compute the posterior class probability of the classifier. This probability displays the likelihood of each ROI of the test image belongs to the higher grade cancer. This estimated likelihood for the $i$th ROI is computed as follow:

$$p_{ROI_i} = p(y = i|X), i = -1, 1$$ (3.8)

where $p_{ROI_i}$ is the estimated likelihood for the $i$th ROI for any input $X$. Using a sigmoid function, the estimates likelihood is approximated as follows:

$$p(y = i|X) \approx \frac{1}{1 + \exp(\alpha \hat{f} + \beta)}$$ (3.9)
In the above equation, \( p(y = 1|X) \) denotes the higher grade estimated likelihood and \( \hat{f} \) is the binary output of the classifier. Using the training dataset and their corresponding true labels, parameters \( \alpha \) and \( \beta \) are chosen to minimize the negative likelihood functions. This approximation is detailed here [57].

### 3.0.5 Colormaps

In addition to classifying ROIs in leave-one-patient-out cross-validations, the colormaps showing the likelihood of higher grade cancer in the entire cancerous area of each patient are also generated. For creating these colormaps, a classifier is trained to distinguish between classes of Gleason score \( \leq 3+4 \) and Gleason score \( \geq 4+4 \); the classifier is then applied to predict the likelihood of ROIs belonging to a higher grade class in the entire delineated cancerous area for a never-before-seen patient.

### 3.1 Results

#### 3.1.1 Leave-one-patient-out evaluation

For the cross validation, we have followed the leave-one-patient-out strategy using both individual and combinations of RF time series features. Classification parameters were tuned through an exhaustive search to achieve the highest classification accuracy. These parameters were \( C = 6.5 \) and \( \gamma = 0.5 \). The low value of \( \gamma \) makes the classifier very close to a linear classifier which at the same time ensures the prevention of overfitting of the classifier to the training dataset. The accumulated ROC curve was generated based on the results of the classification. Following testing the classifier on the left-out patient, the testing ROIs along with the posterior class probabilities of the classifier are accumulated and the final ROC curve is generated with a likelihood
3.1. RESULTS

threshold of 0.5. The total sensitivity and specificity of the method along with the results of the classification are reported. Sensitivity and specificity demonstrate the frequency of the true diagnosis in the patients [74] and are defined as:

\[
\text{sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}} \quad (3.10)
\]

\[
\text{specificity} = \frac{\text{number of true negative}}{\text{number of true negatives} + \text{number of false positives}} \quad (3.11)
\]

The subset of features F16, F17 and F21 resulted in the highest accuracy and area under accumulated ROC curve in the leave-one-patient-out cross-validation. These features are \( \frac{SS(2)}{SS(1)} \), \( \frac{SS(3)}{SS(1)} \) and \( \frac{SS(7)}{SS(1)} \), respectively where \( SS(i) \) represents the summation of the spectrum in the \( i \)th frequency band. Table 3.2 shows the classification accuracy, the area under accumulated ROC (AUC), sensitivity and specificity for the best individual features and their combination in separating Gleason scores \( \leq 3+4 \) and Gleason scores \( \geq 4+3 \). The highest classification accuracy and AUC were achieved for the combination of the best individual features. For ROIs with Gleason score of 4+3, we have noticed that the distribution of selected features do not completely follow those of the other higher grade ROIs. If the classification is only performed between Gleason scores \( \leq 3+4 \) and Gleason scores \( \geq 8 \), an accuracy of up to 83%, AUC of 0.87, and sensitivity and specificity of 80% and 97% are achieved, respectively. The high rate of specificity here indicates that most higher grade cancerous ROIs are correctly classified.
The performance of previously introduced features is also investigated. The subset of features introduced by Moradi et al. [65] were used. They yielded accuracy, AUC, sensitivity and specificity of 72%, 0.77, 80% and 70%, respectively (table 3.2). Features introduced by Imani et al. resulted in the accuracy, AUC, sensitivity and specificity of 72%, 0.67, 70% and 77%, respectively. We have also performed Mann-Whitney-Wilcoxon (MWW) statistical test to see if the differences between the 12 classification accuracies obtained by previously reported features and the proposed features in this thesis are significant. MWW is a non-parametric test that does not assume a particular distribution for the data. We compared two sets of medians; it is possible to correct for multiple comparisons using e.g., a Bonferroni correction but since the number of of comparisons is only two, the correction will not affect the significance of findings. The p-values for these test are less than 0.001 which confirms statistically significant difference between accuracies obtained by Imani et al. [44] features (p-value = 0.0005), Moradi et al. [65] features (p-value = 0.0008) and the features proposed in this thesis. Figure 3.5 show the ROC curve for leave-one-patient-out cross validation using Moradi et al. [65] features, Imani et al. [44] features and the proposed features in this thesis.

The colormaps demonstrating the likelihood of a cancerous area to be higher grade prostate cancer for four sample patients are also depicted. Figure 3.6, a. shows the likelihood of an ROI being a higher grade cancer with the entire cancerous areas encircled in white. The colormaps only show ROIs that belong to higher grade class with likelihood of ≥ 0.5. Figure 3.6, b. shows the true grades overlaid on the ultrasound images as reported by the pathologist. Figure 3.6, c. displays the pathology image with cancerous regions encircled in black. As it is evident from Figure 3.6, a.
3.2. CHAPTER SUMMARY AND CONCLUSION

Table 3.2: Average accuracy, sensitivity, specificity and AUC in the leave-one-patient-out cross-validation and its comparison with previously reported features in the literature.

<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>F16</td>
<td>70%</td>
<td>77%</td>
<td>0.53</td>
<td>67% ± 31%</td>
</tr>
<tr>
<td>F17</td>
<td>76%</td>
<td>80%</td>
<td>0.72</td>
<td>72% ± 26%</td>
</tr>
<tr>
<td>F21</td>
<td>75%</td>
<td>79%</td>
<td>0.73</td>
<td>78% ± 20%</td>
</tr>
<tr>
<td>F16,F17,F21</td>
<td>81%</td>
<td>80%</td>
<td>0.8</td>
<td>79%* ± 19%</td>
</tr>
<tr>
<td>Moradi <em>et al.</em> features [65]</td>
<td>72%</td>
<td>69%</td>
<td>0.7</td>
<td>69%* ± 25%</td>
</tr>
<tr>
<td>Imani <em>et al.</em> features [44]</td>
<td>69%</td>
<td>72%</td>
<td>0.67</td>
<td>65%* ± 27%</td>
</tr>
</tbody>
</table>

the classifier is more specific to higher grade cancer yielding low rates of misclassified higher grade ROIs. The colormaps for all the slices of all patients used in this study are presented in Appendix A.

Figure 3.6 also shows an example of a patient image with mostly lower (patient 4). The colormap for this patient shows that there is no ROI with likelihood of $\geq 0.5$ to be classified as higher grade.

3.2 Chapter summary and conclusion

In this chapter, we proposed a feature extraction and classification approach from RF time series data for characterizing aggressive prostate cancer. Novel frequency domain features of RF time series were introduced and the performance of these features compared to the performance of the previously reported RF time series features were shown for separating Gleason scores $\leq 3+4$ 7 versus $\geq 4+3$ 7. We showed that using the novel features, we could highlight the higher grade cancers in areas of size $1 \text{ mm} \times 1 \text{ mm}$ with an accuracy of up to 79%. Moreover, colormaps demonstrating the
likelihood of the entire cancerous area of the never-before-seen patient were depicted and overlaid on the ultrasound images. The results of the current study shows promise in using RF time series to characterize clinically significant prostate cancer in the future.
Figure 3.6: Colormaps showing the likelihood of belonging to higher grade cancer
a) the result of classification b) true overlaid cancer c) pathology image encircling the cancerous areas.
Chapter 4

Characterization of aggressive prostate cancer, *in vivo* and *ex vivo* feasibility study

The ultimate goal of ultrasound-based methods for prostate cancer detection is to augment the TRUS-guided biopsy process by performing targeted instead of blind biopsy. In this chapter, we study the performance of RF time series and the features proposed in chapter 3 for differentiating between lower and higher grades of prostate cancer in an *in vivo* environment. For this purpose, *in vivo* B-mode and RF time series data were collected prior to the surgery. We also had the opportunity to collect *ex vivo* RF ultrasound data from the same patients following the surgery. Using this data we demonstrate RF time series can effectively characterize aggressive prostate cancer in ROIs as small as $1 \text{ mm} \times 1 \text{ mm}$. In the *in vivo* study classification accuracy and AUC of 81% and 0.88 were archived, respectively using 10-fold cross-validation. Also in the *ex vivo* study a classification accuracy of 78% and 0.83 were obtained in the leave-one-patient-out cross-validation.

Hence, following the *in vivo* data collection, *ex vivo* RF time series were also collected. Moreover, Computed Tomography (CT) images of the *ex vivo* prostate
specimens were collected for registering the gold standard pathology to the *in vivo* ultrasound data. Following that the histopathological analysis of the whole mount slides was performed by a senior pathologist. The details of each step of data collection are provided, ROI selection and feature extraction are explained and the results are demonstrated.

4.1 Data Collection

4.1.1 *in vivo* data collection

To further evaluate the tissue typing capabilities of RF time series for characterizing higher and lower grade prostate cancer, following institutional medical research ethics board approval, we collected *in vivo* data from patients who chose prostatectomy as their treatment option at Kingston General Hospital (KGH) in 2013 and 2014. Patients’ consents were acquired to perform ultrasound scanning on extracted prostate specimens prior to their participation in the study. Prior to the surgery, a series of 2D B-mode and RF time series data were collected from a cohort of seven patients. Informed consents were provided by the patients. A SonixTouch ultrasound scanner (Ultrasonix Inc., Richmond, BC, Canada) with a BPL9-5/55 transrectal side firing transducer was utilized. The transducer was mounted on a stepper which was held by a clinician. This stepper was responsible to track the position of each image plane. This tracking information was further utilized for 3D volume reconstruction of prostate specimens in the registration of gold standard pathology to the *in vivo* ultrasound. The details of volume reconstruction are provided in section 4.2.1. During scanning patients were in the supine position having their knees raised. A series of 2D B-mode and RF time series data with a 5° rotational interval were collected. The
B-mode ultrasound images had the lateral and axial pixel sizes of 0.43 mm and 0.019 mm, respectively. The RF time series data consisted of 128 frames of RF ultrasound data with 0.43 mm and 0.019 mm lateral and axial pixel sizes. The ultrasound data was collected with a frequency of 6.67 MHz, a depth of 6 cm and a constant time gain compensation.

We noticed that in two cases, due to the incorrect moving of the probe, the tracking information was not recorded properly. As it was explained previously, ultrasound probe calibration is an important prerequisite for free hand 3D volume reconstruction. The motion artifact of patients led to improper probe calibration results which further yielded in incorrect 3D volume reconstruction. 3D volume reconstruction is one of the important steps toward registering the gold standard pathology results into in vivo ultrasound. Hence, we have excluded the information of these two patients for the in vivo study. Therefore, the in vivo study was performed on a cohort of 5 prostate specimens. However, we have utilized data from these two patients for the ex vivo study. The details of ex vivo data collection protocol will be explained next.

4.1.2 ex vivo data collection

Following in vivo data collection, extracted prostate specimens were fixed. Each fixed specimen was then placed in a cutting box containing a solution of water, 8% glycerol and 3.5% agar and were refrigerated for a day. This cutting box contains slits at 5 mm apart (figure 4.1). The selection of a cutting box was to mitigate the challenges in the registration process between gold standard pathology and ultrasound data. Afterwards, B-mode and RF time series data were collected along the transverse planes at 5 mm apart at the locations of the slits. The slit numbers of the scanning cross
4.1. DATA COLLECTION

Figure 4.1: Cutting box used for data collection. The prostate was fixed and refrigerated for a day in a gel. The probe was set to be vertical to the bottom of the box and parallel to the slits. The tissue was dissected in the same cross-section (slit-numbers) as the ultrasound data were collected.

section were used for manual registration between pathology slicing and ultrasound ex vivo images.

Depending on the size of the prostate, three to five cross sections of data were collected from each specimen. A SonixTouch ultrasound scanner (Ultrasonix Inc., Richmond, BC, Canada) and a L14-5/38 linear transducer were used for data collection. Data were collected with a transducer central frequency of 6.6 MHz, a depth of 5 cm and with a constant time gain compensation. To form the RF time series of data, 159 frames with a frame rate of 43 fps were collected at each cross section of the tissue. The ultrasound probe was touching the posterior part of the prostate and only one focal point was set at 1.5 cm depth (figure 4.1). Moreover, ex vivo CT images of the prostate were collected using a 16-slice G.E. Lightspeed CT scanner.

For one of the ex vivo specimens, due to improper adjustment of frame rate of the ultrasound scanner, the length of the RF time series was different from the rest of the
patients in this study. This yielded in a different frequency bandwidth of the signal. Since the features that are extracted from the signal are mainly frequency domain features, we have excluded this patient from the \textit{ex vivo} experiment. Therefore, the \textit{ex vivo} study was performed on a cohort of six prostate specimens.

Following ultrasound and CT scanning the specimens were dissected in the same ultrasound scanning planes (same slit numbers) and the histopathological analysis of the whole mount slides was performed. The histology images were gridded into digitized ROIs of size $2.5 \, mm \times 2.5 \, mm$. The gridded ROIs were graded by a senior pathologist and Gleason scores were assigned to each digitized ROI. In this study we aimed to differentiate between cancerous lesions with Gleason scores of $\leq 3+4\,7$ and that of Gleason score $\geq 4+3\,7$.

### 4.2 Registration

Figure 4.2 is a flow chart of the registers process between the gold standard pathology data and the \textit{in vivo} and \textit{ex vivo} ultrasound images. As it was mentioned previously, the tissue specimen was dissected at the same cross-section where the ultrasound data was collected. Therefore, pathology is closely aligned with slices of CT and \textit{ex vivo} ultrasound images from the same slits. This helps enhance the registration process and minimize the registration error. In figure 4.2, "Registration 1" and "Registration 2" only consist of simple rigid manual alignments between the corresponding pathology and ultrasound slices since they are from the same cross-section (same slit numbers).
4.2. REGISTRATION

Figure 4.2: The flow-chart demonstrating the steps of registering the gold-standard pathology data into the in vivo ultrasound data. Since the tissue is dissected at the same slit number as the ultrasound data were collected, Registration 1 and 2 consist of simple alignments between the gold-standard pathology and the ultrasound and CT data.

4.2.1 Pathology to in vivo ultrasound registartion

To register the pathology data with in vivo ultrasound data, we have decomposed the registration into two more tractable problems: i) first, the histology images were registered to ex vivo CT data and then ex vivo CT images were registered to in vivo ultrasound data. These registrations were performed using the 3D Slicer software (www.slicer.org). CT data is used as a middle step to help with matching the boundaries of ex vivo and in vivo ultrasound.
### 4.2. REGISTRATION

**Pathology to ex vivo CT registration**

Post-prostatectomy pathology images were used as our gold standard. These pathology images were gridded into $2.5\, mm \times 2.5\, mm$ digitized ROIs and the Gleason scores for each ROI were provided by a senior pathologist. The registration between pathology data and CT images was performed manually. First using the slit numbers of the cutting planes the pathology and CT images were matched. Afterwards these images are overlaid in the 3D Slicer software (www.slicer.org) and using landmarks such as the urethra or the prostate boundary irregularities the registration process is performed. This registration is a rigid registration that only consist of in plane rotations and translations.

**ex vivo CT to in vivo ultrasound registration**

For this purpose, first the 3D B-mode ultrasound volume was reconstructed, using the PLUS library [56]. PLUS is an open-source library which is used for ultrasound image acquisition, calibration and processing. This software provides various functionalities like temporal and spatial calibration of ultrasound images, ultrasound data acquisition, volume reconstruction, etc. Here we have used the volume reconstruction module of PLUS to reconstruct the 3D volume of the tracked ultrasound B-mode images. We also used a hole filling algorithm within the volume reconstruction module to minimize the interpolation between various slices during the reconstruction process. By using the weighted average of surrounding pixels the holes are filled [25].

Using the 3D Slicer software (www.slicer.org) the prostate boundaries were segmented in both modalities by two trained undergraduate and graduate students. Using the orientation of the prostate on both ex vivo CT (determined from the prostate
Figure 4.3: *ex vivo* study: The pathology image is registered to the ultrasound image using the 3D Slicer software. a) an ultrasound image b) pathology image is overlaid on the ultrasound image from the same cross-section c) results of overlaying the available Gleason scores.

layout orientation in the cutting box) and *in vivo* ultrasound data (which can be discerned easily), we rigidly aligned the generated contours. Afterwards an intensity-based affine registration was performed using a multi-resolution global optimization on the generated labelmaps of the prostate in two modalities. The affine registration is detailed in [48].

### 4.2.2 Pathology to *ex vivo* ultrasound registration

The registration process between the histology and ultrasound images was performed manually using the 3D Slicer software (www.slicer.org). First the histology and ultrasound data from the same cross section of the tissue were matched and afterwards using landmarks such as urethra or prostate boundary irregularities, registration of
ex vivo ultrasound to pathology was performed. We have tried to minimize misregistration problems by utilizing a cutting box which enhances the dissection precision of the specimen and ultimately ensures the mechanical registration between ultrasound and the histology images. Moreover, we only chose ROIs close the center of the cancerous areas in the ultrasound images and avoided boundaries. ROIs of size $1 \text{ mm} \times 1 \text{ mm}$ were chosen for tissue characterization purposes. Figure 4.3, a and b show a sample of an ultrasound image and the pathology image registered to that.

### 4.3 ROI selection

Table 4.1: *in vivo* study: total number of segmented ROIs from designated areas with different Gleason scores

<table>
<thead>
<tr>
<th>Gleason score</th>
<th>number of ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 3+3 6</td>
<td>16</td>
</tr>
<tr>
<td>GS 3+4 7</td>
<td>28</td>
</tr>
<tr>
<td>GS 4+3 7</td>
<td>60</td>
</tr>
<tr>
<td>GS 4+4 8</td>
<td>368</td>
</tr>
<tr>
<td>GS 4+5 9</td>
<td>428</td>
</tr>
<tr>
<td>GS 5+5 10</td>
<td>0</td>
</tr>
</tbody>
</table>

For the *in vivo* study, identifying higher and lower grades of prostate cancer was performed in the ROIs of size $1 \text{ mm} \times 1 \text{ mm}$. Each ROI comprises a three dimensional signal of $52 \times 2 \times 280$ samples with a distance between two samples of $0.0192 \text{ mm}$, $0.2148 \text{ mm}$ and $27 \text{ ms}$ in the axial, lateral and temporal directions, respectively. Total number of ROIs of 44 for lower grade cancer (Gleason score $\leq 3+4$) and 856 for higher grade cancer (Gleason score $\geq 4+3$) were segmented (table 4.1). It is
Table 4.2: *ex vivo* study: total number of segmented ROIs from designated areas with different Gleason scores

<table>
<thead>
<tr>
<th>Gleason score</th>
<th>number of ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS 3+3 6</td>
<td>0</td>
</tr>
<tr>
<td>GS 3+4 7</td>
<td>160</td>
</tr>
<tr>
<td>GS 4+3 7</td>
<td>264</td>
</tr>
<tr>
<td>GS 4+4 8</td>
<td>612</td>
</tr>
<tr>
<td>GS 4+5 9</td>
<td>288</td>
</tr>
<tr>
<td>GS 5+5 10</td>
<td>0</td>
</tr>
</tbody>
</table>

expected to have such a high number of higher grade cancer ROIs compared to lower grade as the data is from patients that underwent prostatectomy hence had a likely high volume or grade of cancer.

In the *ex vivo* study, regions of interest (ROIs) of size 1 mm × 1 mm were segmented from each designated area of cancer. In total, 160 ROIs from the cancerous areas with Gleason scores of ≤ 3+4 7 and 1164 ROIs from the cancerous areas with Gleason scores of ≥ 4+3 7 were segmented (table 4.2). Each ROI comprises a three dimensional signal of 104 × 7 × 128 samples with a distance between two samples of 0.0192 mm, 0.2969 mm and 23 ms in the axial, lateral and temporal directions, respectively.

### 4.4 Feature Extraction

In both of the *in vivo* and *ex vivo* experiments, sixty features were extracted from the RF time series as explained in Chapter 3. These features are the summation of the spectrum in 10 frequency bands, fractal dimension of the signal, mean central
frequency and wavelet decomposition features along with the features representing the ratios of the summations of the spectrum. To extract these features, first the mean of the RF time series was removed and it was zero-padded to the closest power of 2. Afterwards, using the DFT technique, the spectrum of RF time series were calculated and the features of each time series were extracted. These features were averaged for a given ROI.

4.5 K-means clustering

As it was explained, the number of available ROIs for the in vivo study is very unbalanced. Although in the 10-fold classification the classifier is trained with a balanced number of ROIs from each class. To visualize the performance of the best individual features to separate higher and lower grade ROIs in the in vivo experiment and to get a feeling for the separation of the data using these best individual features, we have also performed a simple K-means clustering. In K-means clustering, $n$ observations are partitioned into $K \leq n$ clusters by minimizing the the distance between the each observations and the corresponding centroid of each clusters. The objective of the K-means is to find:

$$\arg\min_S \sum_{i=1}^{k} \sum_{x \in S_i} \| x - \mu_i \|^2$$

(4.1)

where $\mu_i$ is the mean of the points in $S_i$. A simple Euclidean distance metric was used here.
4.6 Classification

For both of the *in vivo* and *ex vivo* experiments, a support vector machine (SVM) classifier with a radial basis function was used to classify between lower and higher grades of prostate cancer. The parameters of the classifier (C and $\gamma$) were tuned through an exhaustive search to achieve the best classification accuracy in the test data set. The details of classification parameters are explained in Chapter 3.

For the *in vivo* experiment, 10-fold cross-validation was performed by pooling all ROIs of patients. The reason to perform this cross-validation instead of leave-one-patient-out cross-validation is due to the limited number of available ROIs for lower grade prostate cancer. In the 10-fold cross-validation, the two data sets from the two higher and lower grade classes are randomly permuted first. At each permutation a balanced number of ROIs from each of the classes are passed through the classifier. The fact that the classifier is trained with a balanced number of ROIs from two classes ensures prevention of over-fitting the classifier to one class. The classifier is trained using 90% of the data and is tested on the remaining 10%. The permutation process is repeated for 1000 times and the averaged accuracy and area under the ROC curve are reported in the test portion of the data. An exhaustive search was performed to find the best subset of features. Here the classification parameters were set to $C = 5.5$ and $\gamma = 0.3$. As it was explained previously, a low value of $\gamma$ resembles a linear classifier, which at the same time ensures avoiding the overtraining.

In the *ex vivo* experiment, leave-one-patient-out cross-validation was performed. The classification parameters were set to $C = 0.9$ and $\gamma = 0.1$. Here, the classifier is trained with ROIs of 5 patients and is tested on the ROIs of the never-before-seen patient. An exhaustive search is performed to find the best subset of features.
4.7 Results

For both ex vivo and in vivo, the subset of features F13 and F17 yielded the best accuracies through leave-one-patient-out and 10-fold cross validations. Feature F13 represents the fractal dimension of the time series and feature F17 is the ratio of the summation of the spectrum in the third frequency range to the summation of the spectrum in the first frequency range. For the in vivo experiment, the distribution of the normalized RF time series features confirms that they are good discriminators of higher grade prostate cancer. Figure 4.4 shows the distribution of the best individual features. Higher and lower grade ROIs have mostly higher and lower values for fractal dimension and \( \frac{SS(3)}{SS(1)} \) features, respectively. For the combination of the best individual features, the highest classification accuracy, AUC, sensitivity and specificity of 86%, 0.9, 85% and 93% were achieved, respectively (table 4.3).

In the ex vivo experiment, for leave-one-patient-out cross validation between ROIs
Figure 4.5: *in vivo* experiment: K-means clustering to visualize the performance of best individual features in separating the lower grade and higher grade ROIs.
with Gleason scores $\leq 3+4$ and those of Gleason scores $\geq 4+3$, an accuracy, area under accumulated ROC curve, sensitivity and specificity of 79%, 0.85, 90% and 76% were achieved, respectively (table 4.4). Figure 4.7, shows the distribution of the best individual features for the lower and higher grade ROIs. K-means clustering is performed to visually demonstrate the separation of higher and lower grade ROIs using the best individual features (4.5 and 4.8). Here, higher and lower grade ROIs are shown in red and blue stars, respectively. The results of K-means clustering is shown in red squares and blue circles for higher and lower grade ROIs. Figures 4.6 and 4.9 show the ROC curves for 10-fold cross-validation and leave-one-patient-out cross-validations in the \textit{in vivo} and \textit{ex vivo} studies.
4.7. RESULTS

Figure 4.7: a) ex vivo experiment: distribution of fractal dimension feature for the lower grade (blue) and the higher grade (red) ROIs, b) distribution of \( \frac{SS(3)}{SS(1)} \) for the lower grade (blue) and the higher grade (red) ROIs.

Figure 4.8: ex vivo experiment: K-means clustering to visualize the performance of best individual features in separating the lower grade and higher grade ROIs.
As it has been mentioned earlier, an exhaustive search was performed to find the best subset of features. This yielded different best individual features compared to chapter 3. Here, we have also investigated the performance of the best subset of features reported in chapter 3. These features are $\frac{SS(2)}{SS(1)}$, $\frac{SS(3)}{SS(1)}$, and $\frac{SS(7)}{SS(1)}$, respectively where $SS(i)$ represents the summation of the spectrum in the $i$th frequency band. Using these features, in the $in vivo$ experiment, a classification accuracy, AUC, sensitivity and specificity of of 81%, 0.88, 83% and 87% were obtained, respectively (table 4.3).

Also in the $ex vivo$ experiment, the best features reported in Chapter 3 yielded
4.7. RESULTS

Table 4.3: *in vivo* experiment: average accuracy, sensitivity, specificity and AUC in the 10-fold cross-validation

<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>F13</td>
<td>79%</td>
<td>78%</td>
<td>0.81</td>
<td>77% ± 2%</td>
</tr>
<tr>
<td>F17</td>
<td>78%</td>
<td>72%</td>
<td>0.8</td>
<td>76% ± 1%</td>
</tr>
<tr>
<td>F13, F17</td>
<td>85%</td>
<td>93%</td>
<td>0.9</td>
<td>86%* ± 1%</td>
</tr>
<tr>
<td>F16, F17, F21</td>
<td>83%</td>
<td>87%</td>
<td>0.88</td>
<td>81%* ± 1%</td>
</tr>
</tbody>
</table>

Table 4.4: *ex vivo* experiment: average accuracy, sensitivity, specificity and AUC in the leave-one-patient-out cross-validation

<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>F13</td>
<td>81%</td>
<td>75%</td>
<td>0.71</td>
<td>70% ± 35%</td>
</tr>
<tr>
<td>F17</td>
<td>89%</td>
<td>68%</td>
<td>0.81</td>
<td>75% ± 26%</td>
</tr>
<tr>
<td>F13, F17</td>
<td>90%</td>
<td>76%</td>
<td>0.85</td>
<td>79%* ± 25%</td>
</tr>
<tr>
<td>F16, F17, F21</td>
<td>88%</td>
<td>78%</td>
<td>0.83</td>
<td>78%* ± 21%</td>
</tr>
</tbody>
</table>

classification accuracy, AUC, sensitivity and specificity of 78%, 0.83, 88% and 78%, respectively (table 4.4). These results show that however the best individual features reported in Chapter 3 have not been nominated as the best features in this study, but they have demonstrated almost equal performance in classifying higher and lower grade prostate cancer with no significant statistical difference (*\(p\)-value = 0.31 for the *in vivo* (for 5 patients) and *\(p\)-value = 0.35 for the *ex vivo* (for 6 patients)). In another words, the same features from Chapter 3 could have been used here as well with statistically similar results. An important factor for augmenting clinical applicability of RF time series is to have certain tissue typing features that work well in every clinical set up. Although for this purpose larger scale studies need to be
4.8. Chapter summary and conclusion

This chapter shows the results of a feasibility study on utilizing RF time series for an application closer to clinical translations. In this chapter, we have demonstrated the performance of RF time series for characterizing higher (Gleason score ≥ 4+3 7) and lower grade (Gleason score ≤ 3+4 7) prostate cancer for the first time in in vivo prostate specimens. Moreover, we had the opportunity to collect ex vivo RF time series data. We showed that using the combination of previously reported features [65] and the novel frequency domain features introduced in chapter 3, we could effectively characterize aggressive prostate cancer in ROIs of size 1 mm × 1 mm with an accuracy of 86% and 79% in the in vivo and ex vivo experiments, respectively. The feature histograms confirm the tissue typing capabilities of features extracted from RF time series. In the in vivo experiment, due to the lack of available lower grade ROIs, we have performed 10-fold instead of leave-one-patient-out cross-validation. Also using a simple K-means clustering, we have visualized the performance of best features on separating ROIs with Gleason scores ≤ 3+4 7 and those of Gleason scores ≥ 4+3 7. We have also shown that the performance of features introduced in Chapter 3, on characterizing aggressive prostate cancer is not statistically different with the best individual features here. The results of the current study show promise in using RF time series to characterize clinically significant prostate cancer in the future.
Chapter 5

Single Wall Calibration: Methods and Experiment Results

In almost all ultrasound-guided interventions, calibration is of crucial importance. In this chapter, the framework for evaluating the single wall (SW) calibration in order to test it as a center-independent, operator independent method of calibration is explained. First the mathematical methodology of this approach is detailed. This technique is integrated within the PLUS library. Moreover, an extensive experimental evaluation is performed to test the robustness and reproducibility of this approach.

5.1 Mathematical Framework for the Single Wall Calibration

The mathematical framework behind SW calibration is detailed here. All the derivations (equation 5.1 to 5.17 and 5.20 to 5.21) are proposed by Najafi et al. [68] and are mentioned here to clarify the calibration problem and the solution Najafi et al. [68] presented.

In calibration, the goal is to find the spatial transformation between the image
5.1. MATHEMATICAL FRAMEWORK FOR THE SINGLE WALL CALIBRATION

Figure 5.1: General SW calibration framework which shows the reference (R), transducer (T) and image (I) coordinate systems. Intersection of image and wall plane is a line [68].

coordinate system and the probe coordinate system (figure 5.1, $T_T^r$). This transformation consists of six degrees of freedom (three rotations around x, y and z axes and three translations about these axes). Once this transformation is known, the location of each pixel of the image is known in the global coordinate system referred to as reference coordinate system. Since the transducer is tracked using a tracking sensor mounted on it, the transformation between the probe and reference coordinate system is known (figure 5.1, $R_T^r$). Figure 5.1 demonstrates the schematic view of transducer, image plane, reference coordinate system and the wall plane.

A closed-form solution for the SW technique is proposed by Najafi et al. [68] to calculate the transformation between image and probe coordinate systems (figure 5.1,
5.1. MATHEMATICAL FRAMEWORK FOR THE SINGLE WALL CALIBRATION

Compared to other calibration routines such as multi N-wire, the SW technique takes advantage of the ease of phantom construction. The SW phantom basically is a flat plane (referred to as a wall) immersed into a water tank. In this approach a set of tracked ultrasound images of the wall is collected (usually around 100 images is enough to converge to the solution) in various positions of the probe. The intersection of the image plane and the wall plane is a line that is segmented automatically. The detail of the line segmentation algorithm will be explained later in this chapter.

The image plane is spanned using two free vectors of unit length, $\vec{U}_i$ and $\vec{V}_i$ which are in the lateral and axial directions, respectively. If we consider the lateral and axial directions to be the $\vec{x}$ and $\vec{y}$ directions, the two aforementioned vectors can be re-written in the reference coordinate system as follows:

$$\vec{U}_i = R_T R_I^{T} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{V}_i = R_T R_I^{T} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$ (5.1)

If we name the $R_T = R_i$ to be the rotation from transducer to the reference at $ith$ position of the transducer ($R_0$ being the first rotation), we can re-write the equation 5.1 as:

$$\vec{U}_i = R_i(R_0)^{-1} \vec{U}_0 = R_i^{d} \vec{U}_0, \quad \vec{V}_i = R_i(R_0)^{-1} \vec{V}_0 = R_i^{d} \vec{V}_0$$ (5.2)

Based on the assumptions above, the origin of the image can be written as:
Figure 5.2: The wall plane would be represented as a line in the image. Any pixel along this line lies in the wall plane too and should satisfy the wall plane equation [68].

Here $\vec{t}$ is the translation from image to transducer which is unknown to us and also $\vec{t}_i$ is the translation from transducer to the reference coordinate system which is known since the transducer is tracked via a tracking sensor. So each pixel of the image $(x, y)$ in the reference coordinate system at each position of the transducer could be written as follow:

$$
\begin{bmatrix}
0 \\
0 \\
0 \\
1
\end{bmatrix}
= ^R_R T \vec{t}_I + ^R_i \vec{t}_T = R \vec{t}_I + \vec{t}_i
$$

(5.3)
\[ P = P_0 + S_x \vec{U}_i + S_y \vec{V}_i \] (5.4)

The variables \( S_x \) and \( S_y \) describes the pixel to \( mm \) ratios in the \( \vec{x} \) and \( \vec{y} \) directions of the image plane. The intersection of the image plane (which is a two dimensional plane) with the wall plane would be a line. Every pixel of the image that lies along this line should satisfy the plane equation of the wall. Figure 5.2 shows the imaging plane, wall plane and the intersected line. Suppose we have two points \((x_1, y_1)\) and \((x_2, y_2)\) of the image plane that lie along the intersected line, if we write the wall plane equations for these two points, we will have:

\[
[P_0 + S_x x_1 \vec{U}_i + S_y y_1 \vec{V}_i - Q] \vec{n} = 0 \] (5.5)

\[
[P_0 + S_x x_2 \vec{U}_i + S_y y_2 \vec{V}_i - Q] \vec{n} = 0 \] (5.6)

Here \( Q \) is any arbitrary point in the wall plane and \( \vec{n} \) is the normal vector of the wall in the reference coordinate system. Subtracting the two equations 5.5 and 5.6 will result in the following equation:

\[
\Delta x \vec{U}_i \cdot \vec{n} + k \Delta y \vec{V}_i \cdot \vec{n} = 0 \Rightarrow \vec{U}_i \cdot \vec{n} + km \vec{V}_i \cdot \vec{n} = 0 \] (5.7)

where \( k = \frac{S_y}{S_x}, \Delta x = x_2 - x_1, \Delta y = y_2 - y_1, \) and \( m = \frac{\Delta y}{\Delta x}. \)

If we substitute equation 5.2 into equation 5.6, we will have:

\[
R_i^d \vec{U}_0 \cdot \vec{n} + km R_i^d \vec{V}_0 \cdot \vec{n} = 0 \] (5.8)
This is the closed-form solution for the SW calibration technique which was proposed by Najafi et al. [69]. In the above equation, the vectors $\vec{u}_0$ and $\vec{v}_0$ and $\vec{n}$ are three unknown vectors that are the unit vectors in the lateral and axial directions of the image plane and also the normal vector of the wall plane in the reference coordinate system, respectively. Using properties of Kronecker product, we can re-write the above equation as:

$$vec(AXB) = (B^t \otimes A)vec(X)$$  \hspace{1cm} (5.9)

Applying the above equation to equation 5.7 will result in:

$$(\vec{u}_0^t \otimes \vec{n}_i^t)vec(R_i^d) + km_i(\vec{V}_i^t \otimes \vec{n}_i)vec(R_i^d) = 0$$  \hspace{1cm} (5.10)

The matrix format of the equation 5.8 in all different positions of the probe is:

$$\begin{bmatrix}
R_{11}^d & \cdots & R_{133}^d & m_1 R_{11}^d & \cdots & m_1 R_{133}^d \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
R_{N1}^d & \cdots & R_{N33}^d & m_N R_{N11}^d & \cdots & m_N R_{N33}^d \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
\vdots \\
x_9 \\
ky_1 \\
\vdots \\
kky_9 \\
\end{bmatrix}
= \begin{bmatrix}
0 \\
\end{bmatrix}_{18 \times 1}$$  \hspace{1cm} (5.11)

where we have: $\vec{x}_{9 \times 1} = (\vec{u}_0^t \otimes \vec{n}_i^t)^t$ and $\vec{y}_{9 \times 1} = (\vec{V}_i^t \otimes \vec{n}_i)^t$

The right side matrix of the equation 5.11 which is our unknowns to be found, will be
the solution to the null space of the left side matrix of the equation 5.11. It should also be noted that since the \( \vec{V}_0 \) and \( \vec{U}_0 \) are the unit vectors of image plane, the answer to the equation 5.11 should be scaled such that:

\[
\| X(1:9) \| = \| \vec{x}_{9 \times 1} \| = 1
\] (5.12)

and so the parameter "\( k \)" can be found from:

\[
\| X(10:18) \| = k \| \vec{y}_{9 \times 1} \| = k
\] (5.13)

So far the lateral and axial unit vectors of image plane are known in the reference coordinate system. So the rotation part of transformation \( T^T I \) is solved. Recalling from equations 5.3 and 5.5, if we substitute equation 5.3 into equation 5.5, we’ll have:

\[
(R_i \vec{l} + \vec{t}_i).\vec{n} + S_x x_i \vec{U}_i.\vec{n} + S_y y_i \vec{V}_i.\vec{n} = d,
\] (5.14)

where \( d \) is the parameter of the plane to be computed. Rewriting equation 5.14 as:

\[
c_i = \vec{t}_i.\vec{n} + S_x x_i \vec{U}_i.\vec{n} + S_y y_i \vec{V}_i.\vec{n}, \quad \vec{n}_i = \vec{n}^i R_i, \quad G = \begin{bmatrix} [n_1^i, -1] \\
\vdots \\
[n_N^i, -1] \end{bmatrix}
\] (5.15)

in which we can rewrite equation 5.15 as:
5.1. MATHEMATICAL FRAMEWORK FOR THE SINGLE WALL CALIBRATION

\[
G \begin{bmatrix} \vec{t} \\ d \end{bmatrix} = - \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}
\]  \hspace{1cm} (5.16)

Equation 5.16 can be easily solved as follow:

\[
\begin{bmatrix} \vec{t} \\ d \end{bmatrix} = -(G^tG)^{-1}G^t \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}
\]  \hspace{1cm} (5.17)

So far the mathematic framework of the SW enhanced calibration routine proposed by Najafi et al. [68] was explained. In the following section, the line segmentation approach that was proposed by Najafi et al. [70] and is utilized here is detailed.

**Line segmentation using RF cross-correlation**

Najafi et al. [70] proposed a new automatic line detection algorithm based on the RF signal [70]. Here, first the cross correlation of two columns of RF signal is calculated. Afterwards, by simply calculating the peak of the RF signal along each column of the image and finding the cross correlation between the columns, the slope of the line can be found. Since in the RF image, the axial resolution is high, a very accurate measurement of the slope of the line is possible [68].

Figure 5.3 shows the schematic view of how to calculate the slope of a line using the RF cross correlation. Here, the location of the peak of the signal will determine the \( \Delta x \) and \( \Delta y \). The slope of the line can be easily computed accordingly.
5.2. EXPERIMENTAL RESULTS

Figure 5.3: Slope measurements using RF cross-correlation proposed by Najafi et al. [70]

5.2 Experimental results

To move towards making SW a center-independent and user-independent calibration routine, extensive evaluation is required. For this purpose extensive calibration experiments have been conducted to seek the robustness and limitation of the SW approach.

Data collection was performed in two rounds: i) for the first round, 17 data collection experiments were performed. For each experiment, twenty trials of independent data collections were performed in the same session using the same setup. A SonixTablet ultrasound scanner (Ultrasonix Medical Corporation, Richmond, BC, Canada) was used at a central frequency of 10 MHz. A linear 2D L14-5/38 transducer with 4.5 cm depth was used along with an Optotrak Polaris optical tracker (Northern Digital Inc, Waterloo, Ontario, Canada). For the wall plane an aluminium plate was immersed into a water bath (Figure 5.4 , a). At each trial the image and tracking information at 100 different positions of the probe was recorded. The hand motion
5.2. EXPERIMENTAL RESULTS

Figure 5.4: Experimental setup for the first round of data collection.

consisted of rotations and translations of the probe as suggested in [86]. An alternative calibration technique based on the multi-N-wire method [21] was used as the gold standard. The multi-N-wire calibration routine is integrated into the PLUS library.

To quantify the precision of the calibration routine, a metric called Calibration Reproducibility (CR) is used [42]. Calibration matrices acquired from several calibration experiments (in one-session) are used to back-project some points in the image plane to the probe coordinate system. Usually the corners of the image along with the center of the image are used. If we suppose point $I P$ be on the image and then apply the transformation $T_T I$ (calculated from the calibration) to transform the points from the image coordinate system to the transducer coordinate system, CR is calculated as follow:
5.2. EXPERIMENTAL RESULTS

\[ CR = \frac{1}{N} \sum_{i=1}^{N} \| T \bar{P} - T T_i I P_i \| \]  
\[ (5.18) \]

where \( T \bar{P} \) is as follows:

\[ T \bar{P} = \frac{1}{N} \sum_{i=1}^{N} T T_i I P_i \]  
\[ (5.19) \]

Robust and reproducible calibration routines usually have CR values of less than 1 \( mm \) [20].

Throughout the first round of experiments, SW was very sensitive to the level of noise in the image. In another words, the line in the US image should be a thin sharp line. Furthermore, this approach was very inconsistent and rather incorrect. One way to enhance the robustness of this calibration approach is to have a well-defined hand motion pattern for moving calibration probes while scanning the wall. The calibration reproducibility of the first round of the experiment was far away from the acceptable range. Another approach to improve the robustness of this calibration approach is to digitize the wall plane first using a tracked pointer or \textit{Stylus}. Prior to calibration, the plane equation of the wall can be determined in the reference coordinate system. By transforming plane parameters into the transducer coordinate system, the normal of the plane is no longer a fixed vector but it changes with the transducer’s pose. Recalling from equation 5.8:

\[ n_i = T R_r n_r \]  
\[ (5.20) \]

Considering equations 5.8 and 5.20 for different poses of the transducer, \( i \), will result
in a matrix system of equations:

\[
\begin{bmatrix}
n_i^T & m_i n_i^T \\
\vdots & \vdots \\
kV^T & kV^T
\end{bmatrix}
\begin{bmatrix}
U \\
kV
\end{bmatrix} = 0
\]  

(5.21)

The null space of equation 5.21 will yield the image plane vectors \( \vec{U} \) and \( \vec{V} \).

ii) The second round of data collection was performed to investigate the robustness of the pre-tracked-wall enhanced SW calibration routine. In this round five data collection experiments were performed. This round of data collection was performed using an Optotrak Certus optical tracker (Northern Digital Inc, Waterloo, Ontario, Canada) for more precise tracking, a SonixTouch ultrasound scanner (Ultrasound Medical Corporation, Richmond, BC, Canada) at central frequency of 10 MHz, a linear 2D L14-5/38 transducer with 4.5 \( cm \) depth, an optical Stylus (Northern Digital Inc, Waterloo, Ontario, Canada) and an aluminium plate immersed into a water tank. Twenty trials of independent data collections were performed in the same session, and with the same setup. At each trial using the PLUS library, the Stylus is calibrated. This results in a spatial transformation that relates the tip of the stylus to the optical marker mounted on it. Then the wall plane is digitized using the calibrated Stylus. The image and tracking information at 100 different positions of the probe were recorded. The hand motion consisted of rotations and translations of the probe.

Pre-tracking the wall plane significantly improved the accuracy of the SW. However, the best CR that was achieved through the experiments performed in this thesis was 3.14 \( mm \) ± 1.5 \( mm \). Pre-tracked-wall enhanced SW utilizes an easily constructible
phantom and also uses an automatic method of feature segmentation. In our experiments, this method was sensitive to the results of image segmentation and wall digitization. A well-defined hand motion pattern that makes this method easier to learn for a novice user and digitizing the wall plane first using a tracked pointer or Stylus can improve the robustness of this calibration approach.

5.3 Summary and conclusion

Probe calibration is an inevitable prerequisite in almost all of the ultrasound-guided interventions such as prostate biopsy. Probe calibration registers the ultrasound imaging plane to tools and an external coordinate system. In this thesis, the practical implementation of enhanced SW calibration techniques proposed by Najafi et al. [69] was evaluated. The scope of this project was to investigate if this approach can be used as a center-independent and user-independent method of calibrating ultrasound 2D probes. Extensive experiments in two rounds were performed to evaluate the precision of this method. CR of the method was used as a criterion for the precision of the calibration routine. The first round of data collections was performed without pre-digitizing the wall plane. During this round, the SW approach was sensitive to the segmentation errors and CR values were not in the acceptable range (less than 1 mm). Prior to calibration the wall plane should be digitized using a tracked pointer or Stylus. A second round of data collection was performed to evaluate the robustness of a pre-tracked enhanced SW technique. During the second round of data collections, the best CR values improved significantly. The SW technique proposed by Najafi et al. has advantages such as ease of phantom construction and having a closed-form solution to converge to the calibration answer. In practice having a well-defined hand
motion pattern and digitizing the wall plane first using a tracked pointer can improve
the robustness of this calibration approach.
Chapter 6

Summary and Future Work

Accurate diagnosis of prostate cancer and its aggressiveness helps decision making by allowing for the best treatment options including, radiotherapy, brachytherapy, surgery and active surveillance. The biggest challenge after diagnosing cancer is to determine if it is clinically significant. Determining the aggressiveness of prostate cancer avoids over treatment in patients with indolent cancers. Currently, the gold standard for diagnosis of prostate cancer is core needle biopsy which is performed under ultrasound guidance. Since ultrasound-guided biopsy is performed from pre-defined anatomical locations, it suffers from a lack of sensitivity and specificity for detecting prostate cancer. Several techniques are proposed to enable targeting the biopsy using different imaging modalities. An ultrasound-based prostate cancer detection system is advantageous as ultrasound is already part of the prostate biopsy procedure, it is widely available and relatively safe. The available ultrasound-based techniques for prostate cancer detection suffer from either ubiquitous clinical acceptance and/or lack of accuracy. Recently, a novel ultrasound tissue typing technique was proposed that was shown significantly outperform other conventional ultrasound-based techniques for prostate cancer detection. It was shown that the time series of
ultrasound echoes from a stationary location of the tissue (referred to as RF time series) carry tissue dependent information.

Chapter 3 of this thesis presented a feasibility study to investigate the performance of RF time series, for the first time, to characterize higher and lower grade prostate cancer in 15 ex vivo prostate specimens using novel RF time series features. We showed that utilizing novel frequency features of RF time series, we could effectively highlight higher grade prostate cancer in ROIs of size 1 mm × 1 mm with an accuracy, AUC, sensitivity and specificity of 79%, 0.8, 81% and 80% in the leave-one-patient-out cross-validation, respectively.

In chapter 4, we demonstrated the performance of RF time series for characterizing higher grade prostate cancer for the first time in in vivo prostate specimens. We also performed ex vivo RF time series data collection from the same patients. We have used CT data as a middle step to help with matching the boundaries of ex vivo and in vivo ultrasound. We showed that the combination of the proposed novel RF time series features introduced in chapter 3 and the Moradi et al.’s feature [65] efficiently characterize higher grade prostate cancer in ROIs of size 1 mm × 1 mm. In the in vivo study, an accuracy, AUC, sensitivity and specificity of 86%, 0.9, 85% and 93% were achieved in 10-fold cross-validation, respectively. The selection of 10-fold cross-validation was due to the lack of available lower grade ROIs. Hence the cross-validation was performed by pooling all the available higher and lower grade ROIs. We also showed the novel features of RF time series introduced in Chapter 3, have similar performance to characterizing aggressive prostate cancer with an accuracy, AUC, sensitivity and specificity of 81%, 0.88, 83% and 87%, respectively. Moreover, in the ex vivo study we showed that the combination of the proposed novel
6.1. LIMITATIONS

RF time series features introduced in chapter 3 and Moradi et al.’s feature [65] can effectively characterize aggressive prostate cancer with an accuracy, AUC, sensitivity and specificity of 79%, 0.85, 76% and 90% in leave-one-patient-out cross-validation, respectively. Also the novel features introduced in chapter 3 showed similar tissue typing performance with an accuracy, AUC, sensitivity and specificity of 78%, 0.83, 88% and 78% in leave-one-patient-out cross-validation, respectively.

In almost all of the ultrasound-based interventions, calibration of the ultrasound probes is of crucial importance. Usually calibration is performed by imaging a phantom (a geometrically known object) and by relating the phantom feature locations to their true locations in a reference coordinate system. Recently a calibration technique was proposed that eliminated the need for fabricating complex phantoms. This technique is based on imaging a flat plate (so called single wall) that is immersed into a water tank. In chapter 5 of this thesis the robustness of this method was evaluated. The scope of this work was to investigate an alternative for the current calibration routines such as the multi N-wire technique. CR was used as a criterion to show the precision of the calibration routine in the experiments. Prior to calibration, the wall plane should be digitized using a tracked pointer or Stylus. Our results show that a solid pattern for hand motion is required and this approach is sensitive to the results of segmentation of the lines.

6.1 Limitations

As discussed in chapters 3 and 4, we have registered the pathology data to ex vivo and in vivo ultrasound images, respectively. This task has been performed to define a gold standard for the ROIs that are segmented from multiple malignancy levels
of prostate cancer. One of the limitations of this work was to ensure the perfect match between the pathology and ultrasound images. This was challenging due to inevitable errors of tissue dissection in the pathology lab and ultrasound beam width. We have tried to mitigate this registration error by using a cutting box in chapter 4 which enhances the dissection precision. However, a more precise data collection protocol can help to further mitigate this issue. In chapter 4, we also have registered the pathology data with \textit{in vivo} ultrasound images. For this purpose, first we have manually registered the pathology data into \textit{ex vivo} CT images using the slit numbers of the dissected planes in the cutting box. Thereafter, we registered the \textit{ex vivo} CT images to \textit{in vivo} ultrasound images by using an intensity-based affine registration on the prostate contours in two modalities. One challenge here was to measure this registration error since in practice it was not easy to define the anatomical landmarks in the \textit{in vivo} ultrasound. Moreover, \textit{in vivo} tissue deformation in the posterior part of the prostate is inevitable due to the ultrasound probe. During data collection, more accurate deformable registration of \textit{ex vivo} CT images to \textit{in vivo} ultrasound can help with accurate classification as well.

6.2 Future direction

Future work in this area should focus on including the normal lesions in the classification process. This will increase the clinical contribution of this work. Moreover, larger scale studies need to be performed to evaluate the best set of RF time series features. RF time series is sensitive to tissue and hand motion. In all the data collection setups, significant care was taken to avoid gross motion of the hand or tissue. To tackle this problem, data collection should be performed by significantly increasing
6.2. FUTURE DIRECTION

the acquisition frame rate. This task can be performed by imaging only a small ROI in the tissue. Finally, registration between \textit{ex vivo} CT images and \textit{in vivo} ultrasound data was challenging. Although substantial care was taken to avoid prostate boundaries while taking ROIs from multiple malignancy levels of the cancer, future work on this area should focus on developing a robust deformable registration framework between ultrasound and CT images.
Bibliography


[81] T. Pickles and Prostate Cohort Outcomes Initiative. Low risk prostate cancer carries a minimal risk of prostate mortality and intensification of treatment should be questioned. *In ASCO-ASTRO Prostate Symposium*, 2006.


[87] Prostate cancer Canada: http://www.prostatecancer.ca/.


Appendix A

Colormaps

In this appendix, the results of generated colormaps for all the slices of all patients included in the study in Chapter 3 are provided. For all the slices, the generated colormaps along with the overlaid true Gleason scores on the ultrasound images as well as the pathology images are demonstrated. As it has been mentioned in chapter 3, a threshold level of 60% was applied to digitized pathology grids for ROI selection. This means that ROIs are only segmented from digitized grids with more than 60% of a certain Gleason score. This is to avoid overlaps between multiple Gleason scores while selecting ROIs. Based on this criteria, 3 patients were not considered for ROI selection and the rest of the analysis in Chapter 3. Figures A.1, A.2, A.3 and A.4 demonstrate the results of colormaps generation for all the slices of 12 patients included in the study in chapter 3.
Figure A.1: Colormaps showing the likelihood of belonging to higher grade cancer a) the result of classification b) true overlaid cancer c) pathology image encircling the cancerous areas.
Figure A.2: Colormaps showing the likelihood of belonging to higher grade cancer
a) the result of classification b) true overlaid cancer c) pathology image encircling the cancerous areas.
Figure A.3: Colormaps showing the likelihood of belonging to higher grade cancer
a) the result of classification b) true overlaid cancer c) pathology image encircling the cancerous areas.
Figure A.4: Colormaps showing the likelihood of belonging to higher grade cancer a) the result of classification b) true overlaid cancer c) pathology image encircling the cancerous areas.