Accuracy Assessment of Time Delay Estimation in Ultrasound Elastography

Mohamad Ghasemi Amidabadi

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of

Master of Applied Science (Electrical and Computer Engineering) at

Concordia University

Montréal, Québec, Canada

February 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Mohamad Ghasemi Amidabadi**

Entitled: **Accuracy Assessment of Time Delay Estimation in Ultrasound Elastography**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Electrical and Computer Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

_____ Chair
Dr. Dr. Rabin Raut

_____ External Examiner
Dr. Dr. Terry Fancott

_____ Examiner
Dr. Dr. Wei-Ping Zhu

_____ Co-supervisor
Dr. Hassan Rivaz

_____ Co-supervisor
Dr. M.Omair Ahmad

Approved by _____
Dr. William E. Lynch, Chair
Department of Electrical and Computer Engineering

_____ 2018

Dr. Amir Asif, Dean
Faculty of Engineering and Computer Science

Abstract

Accuracy Assessment of Time Delay Estimation in Ultrasound Elastography

Mohamad Ghasemi Amidabadi

The accuracy of time-delay estimation (TDE) in ultrasound elastography is usually measured by calculating the value of the normalized cross correlation (NCC) at the estimated displacement. NCC value is usually high if the TDE is correct. However, it could be very high at a displacement estimate with large error, a well-known problem in TDE referred to as peak-hopping. Furthermore, NCC value could suffer from jitter error, which is due to electric noise and signal decorrelation. In this thesis, we propose a novel method to assess the accuracy of TDE by investigating the *NCC profile* around the estimated time-delay in a supervised approach. First, we extract seven features from the NCC profile, and utilize a linear support vector machine (SVM) to classify the peak-hopping and jitter error. The results on simulation, phantom and *in-vivo* data show the significant improvement in the classification accuracy realizing from the proposed algorithm compared to the obtained form the state of the art techniques. Second, we build on our model by utilizing the continuity features in the axial and lateral directions as a prior knowledge. We show that these features also improve the sensitivity and specificity of the classifier. After extracting the continuity features in addition to the seven features, we show the performance improvement of the proposed model on the available data sets. Furthermore, we show that our proposed model could be trained by other elastography methods in future, since we use a new elastography algorithm to train the model. Third, we compare the performance of the method developed using well-known classifiers in the literature and then study the importance of the proposed features using the mean decrease impurity method of the random forest classifier.

Acknowledgments

This work would not have been possible without the technical and financial supports of my professors, Dr. Hassan Rivaz and Dr. M.Omair Ahmad. Therefore, I would like to express my sincere gratitude to them who gave me the golden opportunity to do this wonderful project. I really thank them for their continuous support of research, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing the papers. I could not have imagined having a better advisors for my MASc study.

I should also thank my nice and smart colleagues in the IMPACT group at Concordia University, productive discussion with them was sometimes the key for me to answer tough questions in my research. Also, the group has a very nice, genius and hardworking director, Dr. Rivaz who spends a lot of time for his students. I still remember the discussions that we had in his office, and time was really flying in those hours until his wife was calling! All of these made my two years studies at Concordia a great and memorable experience in my life.

I finish with Iran, where the main source of my life energy resides: my family. I have an amazing family, unique in many ways. They have cherished me in every moment of life and loved me unconditionally throughout my studies. I would like to express my deepest appreciation to them for their support even from 6000 miles away!

Contents

List of Figures ix

List of Tables xii

List of Abbreviations xiv

1 Introduction 1

1.1 Ultrasound Elastography 1

1.1.1 Quasi-static Elastography 3

1.1.2 Shear-wave Elastography 3

1.2 Accuracy Assessment in Ultrasound Elastography 5

1.2.1 Fundamental Limit on Time Delay Estimation 7

1.2.2 SNR and CNR 8

1.2.3 Normalized Cross-Correlation Based Methods 9

1.3 Investigation of the NCC Profile 10

1.4 Motivation and Objective of the Thesis 12

1.5 Organization of the Thesis 13

2 Supervised Classification of the Accuracy of the Time Delay Estimation in Ultrasound Elastography 14

2.1 Introduction 14

2.2 Method and Material 15

2.2.1 Feature Selection . . . 16
2.2.2 Supervised Learning . . . 17
2.2.3 Training Sets . . . 18
2.2.4 Classification . . . 18
2.3 Results . . . 18
2.3.1 Running time . . . 19
2.3.2 Simulation Data . . . 19
2.3.3 Phantom Data . . . 20
2.3.4 *In-Vivo* Data . . . 21
2.3.5 Accuracy Map in Region of Interest (ROI) . . . 22
2.4 Summary . . . 26

3 Automatic Accuracy Assessment of Ultrasound Elastography Using Correlation Profile and Prior Information of Displacement Continuity 27
3.1 introduction . . . 27
3.2 Method and Material . . . 28
3.2.1 Supervised Learning . . . 28
3.2.2 Training Sets . . . 29
3.2.3 Classification . . . 29
3.3 Results . . . 30
3.3.1 Data Sets . . . 31
3.3.2 Accuracy Map in Region of Interest (ROI) . . . 32
3.4 Summary . . . 37

4 Performance Evaluation of Well-known Classifiers and Importance Measurement of the Proposed Features 38
4.1 Introduction . . . 38
4.2 Accuracy of Classifiers . . . 39
4.3 Feature Importance Measurement . . . 42
4.4 Summary . . . 42

5 Discussion and Conclusion **46**

5.1 Discussion 46

5.2 Conclusion 48

5.3 Scope for Further Investigation 48

List of Figures

Figure 1.1 The Alpinion ultrasound machine at the PERFORM Centre. 2
Figure 1.2 (a) Pre-compressed image on the left, (b) Post-compressed image on the right. 4
Figure 1.3 (a) , (b) Axial displacement field and axial strain image, respectively. 4
Figure 1.4 Propagation of shear waves in the tissue. 5
Figure 1.5 Jitter error and its impact in shifting the position of maximum value of the cross-correlation function. 6
Figure 1.6 Correct estimated sample in post-compressed image is shown in black and peak-hopping sample which is far from the actual estimate is shown in red. 7
Figure 1.7 Strain filters are shown for center frequencies of 3.5MHz, 5MHz, 7.5 MHz, respectively. Bandwith is 60 % and T= 1.3 μs. The SNR is improved by increasing the pulse center frequency. This figure is borrowed from [1]. 9
Figure 1.8 Three different strain images which show incorrect estimated areas in red. This figure is borrowed from [2]. 10
Figure 1.9 Joint probability density function of the skewness and variance for 1000 correct TDEs. 11
Figure 1.10 Joint probability density function of the skewness and variance for 200 peak-hopping TDEs. 11
Figure 1.11 NCC curve obtained by shifting the post-compressed window by ± 1 samples in both axial and lateral directions. NCC windows of pre- and post-compressed images are shown in black and red respectively. 12

Figure 2.1 Displacement between pre-and post compressed images. I_1 and I_2 are pre- and post-compressed images, respectively. Z, X, Y are axial, lateral and out-of-plane directions, respectively. The coordinate system is attached to the ultrasound probe. The sample P (i,j) in I_1 has been moved by ($a_{i,j}$, $l_{i,j}$) in I_2. 15

Figure 2.2 (a), (b) I_1 and I_2 are pre- and post-compressed images, respectively. Windows $\mathbf{X}$ and $\mathbf{Y_i}$, $(i = 1, 2, ..., 9)$ correspond to the sample (i,j) in I_1. (c) Nine NCC values are illustrated, NCC_1 was the only value which was used to asses accuracy. 16

Figure 2.3 Flow chart of the proposed method. I_1 and I_2 are pre- and post-compressed images, respectively. Displacement field is either available from the simulation data, or calculated from a displacement estimation algorithm. 17

Figure 2.4 ROC curve for the simulation data. 20

Figure 2.5 ROC curve for the CIRS phantom data. 21

Figure 2.6 ROC curves for three patients. 23

Figure 2.7 Accuracy map in tumor region for patient 3 is shown in red boxes. 24

Figure 2.8 Accuracy map in tumor region for patient 3 is shown in red boxes for true and false positive cases. 25

Figure 3.1 Displacement between pre-and post compressed images. I_1 and I_2 are pre- and post-compressed images, respectively. Z, X, Y are axial, lateral and out-of-plane directions, respectively. The coordinate system is attached to the ultrasound probe. The sample P (i,j) in I_1 has been moved by ($a_{i,j}$,$l_{i,j}$) in I_2. 29

Figure 3.2 Pre-compressed image on the left and the corresponding TDE between I_1 and I_2 on the right. For sample P (i,j) in I_1, eight neighboring TDE values are shown as black points on the right image. 30

Figure 3.3 Flowchart of the proposed method. I_1 and I_2 are pre- and post-compressed images, respectively. Displacement field is available from a displacement estimation algorithm. In this flowchart, feature selection section involves extracting 9 features instead of 7 features proposed in Fig 2.3 and Chapter II. 31

Figure 3.4 ROC curve for phantom data. 33

Figure 3.5 ROC curve for patient 1. 33

Figure 3.6 ROC curve for patient 2. 34
Figure 3.7 ROC curve for patient 3. 34
Figure 3.8 Accuracy map in tumor region for patient 3 is shown in red boxes. 35
Figure 3.9 Accuracy map in tumor region for patient 3 is shown in red boxes for true and false positive cases. 36
Figure 4.1 Feature importance measurement for simulation data. 43
Figure 4.2 Feature importance measurement for phantom data. 43
Figure 4.3 Feature importance measurement for patient 1. 44
Figure 4.4 Feature importance measurement for patient 2. 44
Figure 4.5 Feature importance measurement for patient 3. 45

List of Tables

Table 2.1 Classification accuracy using one NCC value and the proposed method for the simulation data. 20

Table 2.2 Classification accuracy using one NCC value and the proposed method for the CIRS phantom data. 21

Table 2.3 Classification accuracy using one NCC value and proposed seven features for three patient data. 22

Table 3.1 Classification accuracy using one NCC value, 7 features, and 9 features for phantom and the three patient data using SVM classifier. 32

Table 4.1 Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the simulation data. The variance of the classifier is shown in the parentheses beside the mean value. 39

Table 4.2 Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the phantom data. The variance of the classifier is shown in the parentheses beside the mean value. 40

Table 4.3 Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 1. The variance of the classifier is shown in the parentheses beside the mean value. 40

Table 4.4 Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 2. The variance of the classifier is shown in the parentheses beside the mean value. 41

Table 4.5 Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 3. The variance of the classifier is shown in the parentheses beside the mean value. 41
Table 5.1 The Mean Squared Error (MSE) of displacement values obtained using DPAM for all data sets. 47

List of Abbreviations

TDE:	Time-Delay Estimate
RF:	Radio-Frequency
CRLB:	Cramer-Rao Lower Bound
SNR :	Signal-to-Noise Ratio
CNR :	Contrast-to-Noise Ratio
NCC:	Normalized Cross-Correlation
SSD :	Sum of Squared Difference
SVM:	Support Vector Machine
PDF:	Probability Density Function
DPAM:	Dynamic Programming and Analytic Minimization
FEM:	Finite Element Method
ROC:	Receiver Operating Characteristic
ROI:	Region of Interest
TP:	True Positive

FP:	False Positive
GLUE:	Global Time-Delay Estimation in Ultrasound Elastography
ANN:	Artificial Neural Network
KNN:	K-Nearest Neighbor
LDA:	Linear Discriminant Analysis
MSE:	Mean Squared Error
CNN:	Convolutional Neural Network

Chapter 1

Introduction

1.1 Ultrasound Elastography

Ultrasound elastography is an emerging medical imaging modality that involves measuring tissue deformation field caused by an external or internal forces [3, 4]. Several ultrasound-based techniques have been established in past years in the literature to find the deformation field [5, 6, 7, 8] also referred to as Time-Delay Estimates (TDE). TDE is calculated between two frames of ultrasound Radio-Frequency (RF) data, and is used to infer tissue mechanical properties, in particular Young's modulus. On one hand, elastography has been applied in several clinical trials in breast, liver and prostate cancer, and is rapidly finding new clinical applications [9, 10, 11, 12, 13]. On the other hand, it has evolved into several different techniques such as shear-wave elastography [14, 15] and quasi-static elastography [16, 17]. In quasi-static elastography, either the spatial gradient of the displacement field (i.e. a strain image) is used, or an inverse problem is solved to calculate the Young's modulus [18, 19, 20]. Ultrasound machines are typically portable, which makes them a convenient imaging modality for both diagnostics and surgical guidance. In Fig. 1.1, a typical ultrasound imaging machine is shown.

Figure 1.1: The Alpinion ultrasound machine at the PERFORM Centre.

1.1.1 Quasi-static Elastography

This work focuses on quasi-static elastography, wherein tissue deformation is slow and is created by pushing the tissue by a hand-held ultrasound probe [17]. In quasi-static elastography, two images play an important role in measuring the mechanical properties of the tissue. First, the image that is taken without any pressure on the tissue and is called pre-compressed image. Second, the image that is taken while the tissue undergoes deformation due to an applied stress and is called post-compressed image. By measuring the deformation between the two images, we could get useful information about mechanical properties of the tissue for medical diagnosis and treatment. In Fig. 1.2, pre- and post-compressed images are shown.

There have been proposed several methods in the literature to find the deformation field between the two images. In fact, the displacement field is estimated by tracking the speckles's movement in the pre- and post compressed images. Many different types of deformation estimators are developed which can be mainly categorized into window-based and regularized cost function methods. Window-based estimators, either maximize the similarity functions (usually cross-correlation) or find the zero-crossing of the cross-correlation function [21, 22, 23, 24]. Minimization of a regularized cost function is the second popular approach compared to correlation-based methods due to their robustness to signal decorrelation [16, 25, 26, 27, 28, 29]. Prior knowledge that the tissue deformation field is smooth, is the key assumption in the proposed methods.

After finding the deformation field, axial gradient estimation is usually applied to generate axial strain image (Fig. 1.3). Since taking derivative of consecutive TDEs, amplifies the noise in the strain image, filtering is also needed to end up with smoother strains [3, 5, 27].

1.1.2 Shear-wave Elastography

Shear-wave elastography provides quantitative values of tissue elasticity or shear moduli, wherein tissue deformation is generated by ultrasonic focused beams [14, 30, 31]. By generating radiation sources remotely in the tissue, low frequency shear-waves are propagated through the medium

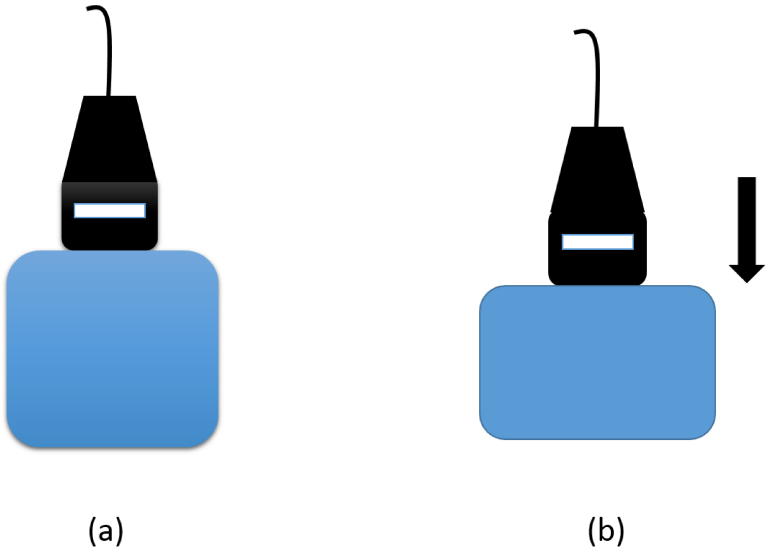
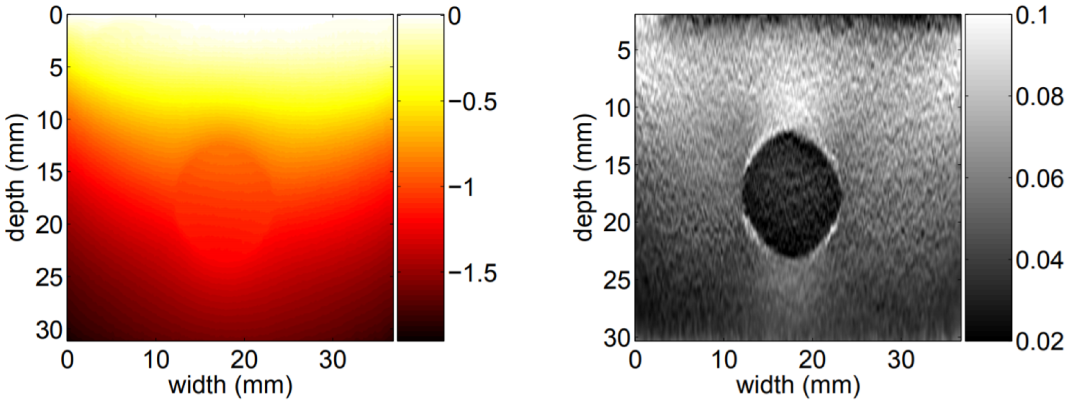


Figure 1.2: (a) Pre-compressed image on the left, (b) Post-compressed image on the right.



(a) Axial displacement (mm) (b) Axial strain

Figure 1.3: (a) , (b) Axial displacement field and axial strain image, respectively.

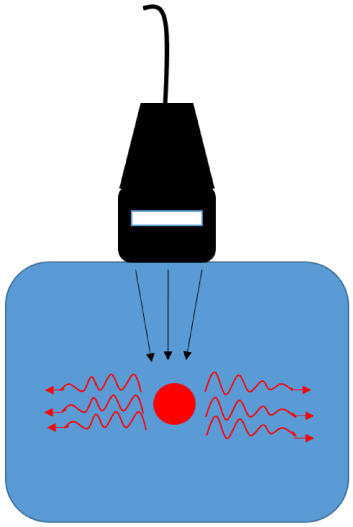


Figure 1.4: Propagation of shear waves in the tissue.

which is progressively distorted due to tissue heterogeneities. Insensitivity to patient motion and boundary condition artifacts are main advantages of shear-wave elastography. However, their application is limited due to having heavy and bulky vibrators with a unique shear spatial directivity pattern [14], thereby an alternative approach is proposed in [32, 33, 34]. This method consists of focusing an ultrasonic beam deep in tissues during a long time (around 100 μs) and uses ultrasound correlation-based techniques to measure the resulting displacement field at the focused medium . In Fig 1.4, propagation of shear waves in the vibrated tissue is schematically shown.

1.2 Accuracy Assessment in Ultrasound Elastography

TDE is challenging due to signal decorrelation between the two ultrasound frames. Failure in TDE creates artifacts in elastography, which can adversely affect diagnosis or surgical operations. The importance of assessment of the accuracy of TDE is four-fold. First, it can be used to mask out erroneous areas of the elasticity image. Second, ultrasound frame rate is very high, and therefore, several TDEs can be calculated and utilized every second. Recent work has, in fact, focused on

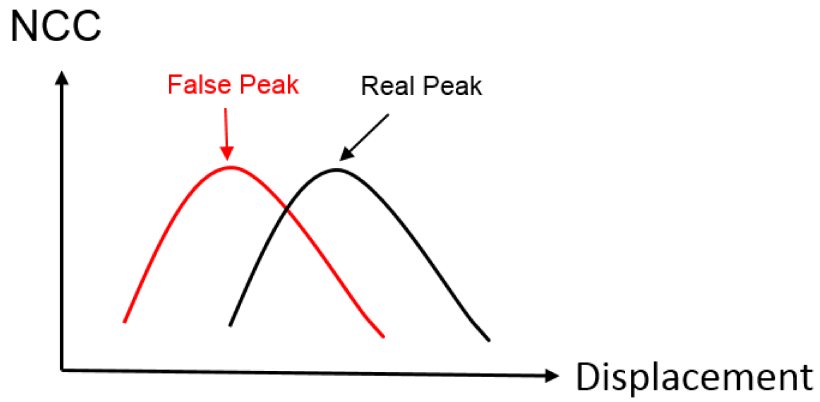


Figure 1.5: Jitter error and its impact in shifting the position of maximum value of the cross-correlation function.

exploiting multiple images to improve the quality of elastography [17],[35, 36]. An accuracy map can be exploited to perform weighted averaging of these elasticity images, assigning smaller weights to uncertain TDEs. Third, quasi-static elastography methods are user-dependent, and rely highly on the skill of the sonographer. An accurate assessment method can be used to generate accuracy maps alongside strain images, which can help train sonographers to obtain elastography images of higher quality. And fourth, elastography algorithms often use the displacement estimation of previous samples to reduce the search range and computational complexity, which can lead to propagation of errors. The proposed method can be used to prevent such propagation of errors.

TDE methods are always subject to small and large error values. Small error is widely referred to as jitter and can be quantified by studying the fundamental limits on the performance evaluation of TDE [35]. In Fig. 1.5, jitter error results in a shift in position of maximum value of the cross-correlation function in axial or lateral directions. Large errors in the displacement field, also called peak-hopping, create outlier data in the displacement field. Peak-hoping happens when maximum value of the cross-correlation function is found somewhere far from the actual displacement estimate within a predefined search region (Fig. 1.6). Several algorithms have been proposed to reduce the occurrence of large errors and limit their effect in both quasi-static [21, 36, 37] and shear-wave elastography [38, 39]. Although accuracy assessment of TDE is essential in both quasi-static and shear-wave elastography, this work focuses on the former.

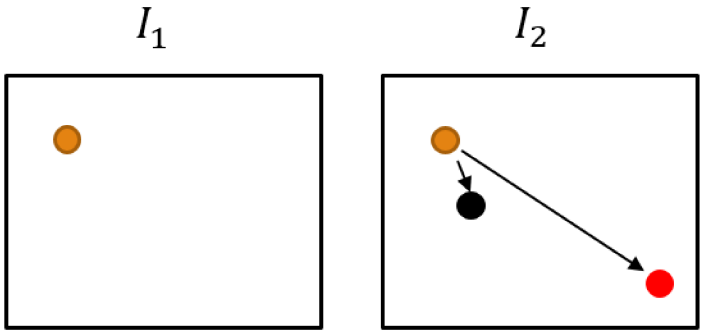


Figure 1.6: Correct estimated sample in post-compressed image is shown in black and peak-hopping sample which is far from the actual estimate is shown in red.

1.2.1 Fundamental Limit on Time Delay Estimation

Widespread adoption of ultrasound elastography relies on reliable quality assessment of TDE. Early work is focused on prediction of a lower bound for the variance of the displacement estimate errors using Cramer-Rao Lower Bound (CRLB) [40, 41, 42, 43]. These contributions derive a closed-form expression for the TDE error variance as following:

$$\sigma(\Delta t - \Delta\hat{t}) \geq \sqrt{\frac{1}{T\int_{-\infty}^{\infty}(2\pi f)^2 \frac{C_{r_1,r_2}(f)}{1-C_{r_1,r_2}(f)}df}}$$

where Δt is the correct displacement, $\Delta\hat{t}$ is the estimated displacement, T is the window length and $C_{r_1,r_2}(f)$ is the magnitude-squared coherence function defined as following:

$$C_{r_1,r_2}(f) = |\frac{C_{r_1,r_2}(f)}{\sqrt{C_{r_1,r_1}(f)C_{r_2,r_2}(f)}}|^2$$

It is parameterized in terms of the TDE algorithm and ultrasound system configuration, and obtains the minimum achievable error by any unbiased TDE algorithm [40, 41]. Although mathematically elegant, this approach does not provide the accuracy of the TDE and instead produces a lower bound value for the error variance.

1.2.2 SNR and CNR

To overcome the shortcomings of the lower bound error analysis, general performance measurements such as signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) of the strain image are widely used [1, 44, 45, 46, 47]. These measures are often called strain filter since they behave similar to a bandpass filters when they are measured as a function of axial strain. For example in [1], SNR is defined as the ratio of the mean m and standard deviation σ of the estimated strain over a small window as following:

$$SNR = \frac{m}{\sigma}$$

The strain filter proposes a framework that allows a limited range of strains to be included in the strain image. The reasons why the filter deviates from an all-pass characteristics in the strain domain are ultrasound system parameters, the finite value of the SNR, and the effect of signal decorrelation for high strain values. In fact, the strain filter approach provides valuable insight to design strain-imaging systems to generate high quality results. In Fig. 1.7, three strain filters are shown at different center frequencies of 3.5MHz, 5MHz, 7.5 MHz, respectively, with a bandwith of 60 % and duration of data window of T= 1.3 μs. The SNR is improved by increasing the pulse center frequency. However, there are three disadvantages in using general performance measurements. First, they should be estimated in small windows that must be regions of constant strain [1]. Such regions are not necessarily available in real tissue that is largely heterogeneous. Second, these methods do not directly estimate the accuracy of TDE but rely solely on strain, the spatial derivative of TDE. As such, the derivative operator has a strong impact on these measures, such that higher SNR and CNR can be obtained by smoother derivative operators [48]. Third, strain filters only provide an upper bound for the quality of the estimated strain value, and do not produce the quality of the specific TDE. For example, these filters usually predict a high value of SNR and CNR at around 1% strain for most displacement estimation algorithms. However, an incorrect displacement may be estimated at 1% strain due to a large signal decorrelation.

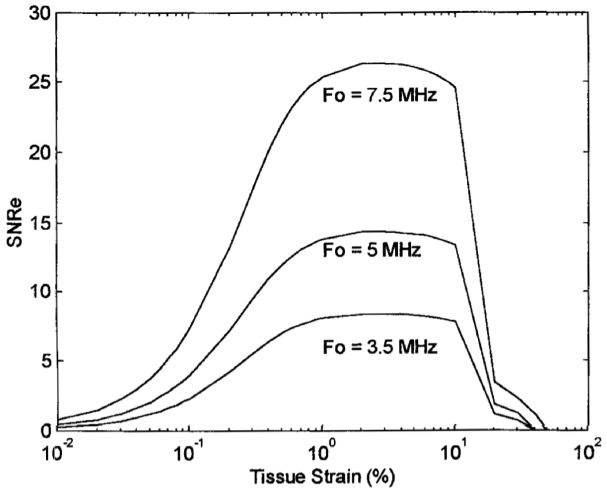
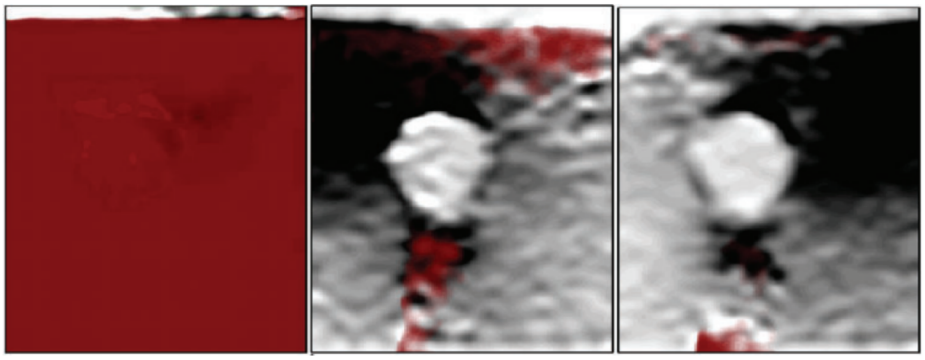


Figure 1.7: Strain filters are shown for center frequencies of 3.5MHz, 5MHz, 7.5 MHz, respectively. Bandwith is 60 % and T= 1.3 μs. The SNR is improved by increasing the pulse center frequency. This figure is borrowed from [1].

1.2.3 Normalized Cross-Correlation Based Methods

The application of normalized cross-correlation (NCC) similarity metric in performance evaluation of TDE is an active field of research. NCC is generally calculated between two corresponding windows in the two RF frames to quantify the accuracy of the displacement field. In [2], a method has been proposed to mask out the incorrect areas of the strain image where NCC falls below a user-defined threshold. Fig. 1.8 shows three different strain images, wherein incorrect estimated samples are washed out in red. In [49, 50, 51], the value of NCC is linked to the precision of the displacement field, which is used later to blur out the areas of the strain image that are not accurate. In [46], standard deviation of the jitter error is evaluated by finding NCC and the sum of squared difference (SSD) values of the corresponding windows. In [52], a technique has been introduced that incorporates the consistency information of consecutive frames as a measure to evaluate the quality of strain images. Finally, a frame quality indicator has been presented in [53] that selects a few representative frames from a large pool of axial strain images based on the value of NCC and performs weighted averaging of the strain images based on the NCC value. While these methods improve the quality of strain images, they utilize only NCC at the TDE. We will show that NCC profile around the TDE contains information that is otherwise not available from the peak NCC



(a) (b) (c)

Figure 1.8: Three different strain images which show incorrect estimated areas in red. This figure is borrowed from [2].

value.

1.3 Investigation of the NCC Profile

In Fig. 1.9, joint probability density functions of the skewness and variance for 1000 correct estimated samples and 200 peak-hopping samples are shown. For each sample, nine NCC values are calculated by shifting the center of the post-compressed window by ± 1 samples in the axial and lateral directions (i.e. nine values in a window of size 3×3). As it is obvious, the surface in Fig. 1.9 is concentrated in a specific area whereas it is spread randomly in Fig. 1.10. This difference in the behavior of the probability density functions enables us to use the skewness and variance of the neighboring samples around the estimated one as two invaluable features to train the classifier. The skewness and variance of the NCC profile is calculated as following:

$$Variance = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

$$Skewness = \frac{\sum_{i=1}^{n}(x_i - \mu)^3}{n}$$

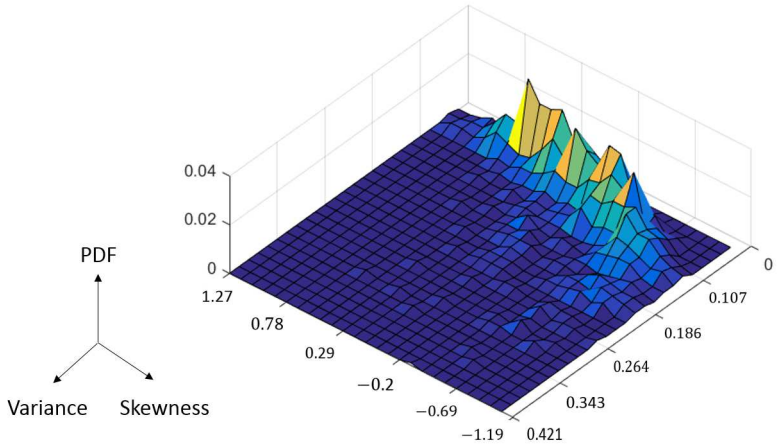


Figure 1.9: Joint probability density function of the skewness and variance for 1000 correct TDEs.

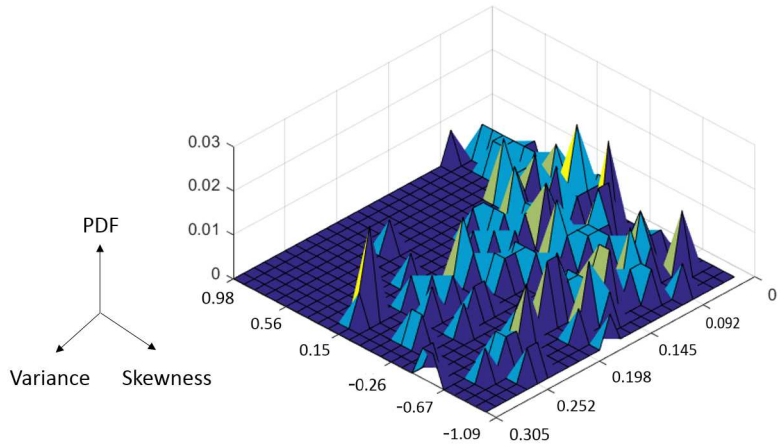


Figure 1.10: Joint probability density function of the skewness and variance for 200 peak-hopping TDEs.

where n and μ represent the number and average of the samples, respectively. Moreover, NCC curve by shifting ± 1 samples in the axial direction in the post-compressed image is shown in Fig. 1.11. It is expected to have a local maximum at the correct estimated displacement. In 2D, this curve becomes a 2D surface, and its shape can help in the assessment of the accuracy of the TDE.

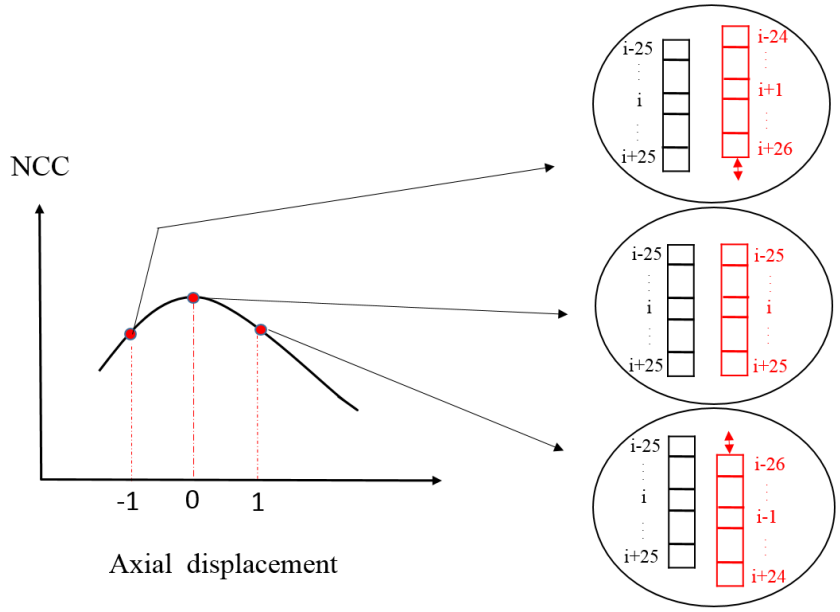


Figure 1.11: NCC curve obtained by shifting the post-compressed window by ± 1 samples in both axial and lateral directions. NCC windows of pre- and post-compressed images are shown in black and red respectively.

1.4 Motivation and Objective of the Thesis

Accuracy assessment of time-delay estimates (TDE) in ultrasound elastography is a critical step in ensuring reliable strain images, which is of critical importance for diagnosis and treatment. In quasi-static elastography, the accuracy of TDE is measured mainly by calculating the value of the normalized cross correlation (NCC) at the estimated TDE. Since this approach is extremely vulnerable to jitter and peak-hopping noise available in TDEs, in this thesis, we investigate the possibility of using the statistics around TDEs and develop a reliable method to address the issue. To this end, we aim to propose several features that would help us in assessing the accuracy of TDE in an most efficient approach. We will investigate the feasibility of real-time automatic accuracy assessment of TDEs with a view to help faster training of sonographers and decision making of doctors. This method could be later used to develop real-time applications reliably used in clinical settings for diagnosis and treatment.

1.5 Organization of the Thesis

In this thesis, we present a novel technique for assessing the accuracy of TDE that relies on multiple NCC measurements. Instead of relying on the value of NCC at the estimated displacement, we look at the NCC profile around the estimated displacement. Our techniques identify locations that contain errors larger than an accepted value, so that these regions can be marked out of the strain image. In Chapter II [54], we look at the 4 NCC values obtained by varying the displacement by ± 1 sample in either axial or lateral directions in addition to the NCC value at the estimated displacement. Five aforementioned NCC values and skewness and variance of nine samples in the 3×3 neighborhood around the estimated TDE are concatenated into a feature-vector of size seven. We show that these seven features are significantly superior in discriminating the correct versus incorrect displacement estimates using simulation, phantom and *in-vivo* experiments. In Chapter III [55], we show that continuity of the displacement field due to homogeneity properties of the tissue also could improve the distinguishability of the classifier. Therefore, each feature vector involves 9 features containing two continuity features in axial and lateral directions in addition to the seven features. Moreover, we use a new method as silver standard to obtain accurate displacement values for training our model which proves that our proposed could be updated by any accurate elastography method for use in future. In Chapter IV, we evaluate the performance of our proposed model by using state-of-the-art classifiers in the literature to find the best model for the TDE accuracy assessment. After finding the most useful classifier, we study the importance of the proposed features based on the Mean Decrease Impurity method of the random forest algorithm [56].

Chapter 2

Supervised Classification of the Accuracy of the Time Delay Estimation in Ultrasound Elastography

2.1 Introduction

In this chapter, we will introduce seven features using the *NCC profile* around the estimated time delay. Then, we use them to train a supervised classification model to do accuracy assessment of time delay estimates which are of high importance for diagnosis and treatment. Herein, SVM is used as a binary supervised classifier to assess the accuracy of TDEs in ultrasound elastography. We show that using the seven proposed features significantly increase the distinguishability of the classifier compared to using only one NCC value.

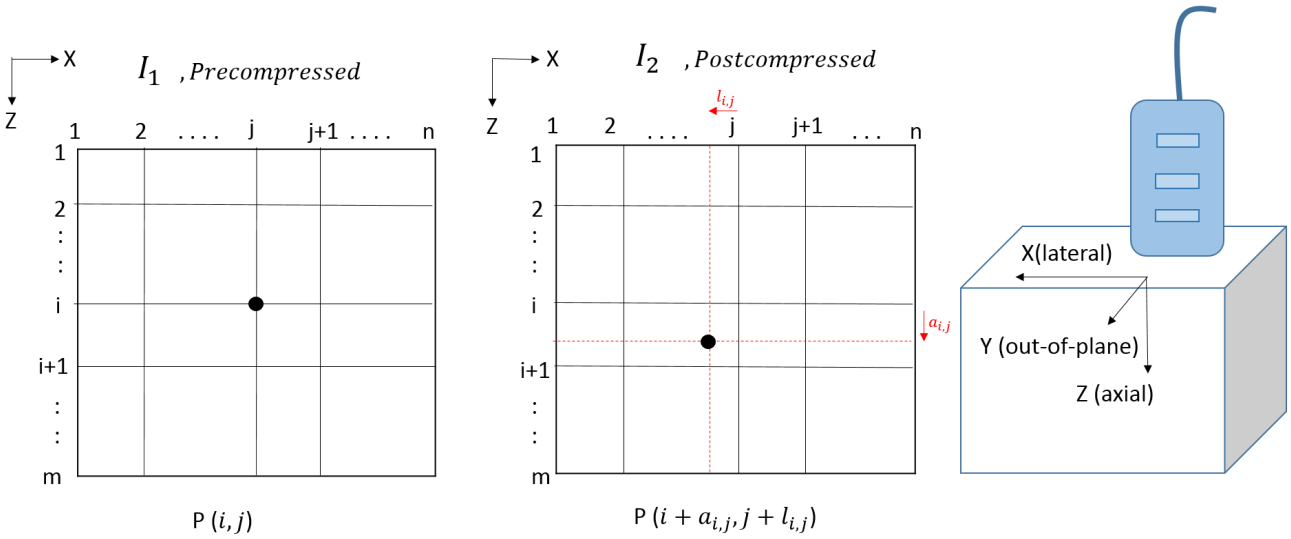


Figure 2.1: Displacement between pre-and post compressed images. I_1 and I_2 are pre- and post-compressed images, respectively. Z, X, Y are axial, lateral and out-of-plane directions, respectively. The coordinate system is attached to the ultrasound probe. The sample P (i,j) in I_1 has been moved by ($a_{i,j}$, $l_{i,j}$) in I_2.

2.2 Method and Material

In elastography, pre- and post-compressed images correspond to each other via a displacement field. Some popular methods for estimation of the displacement field are discussed in detail in [16, 57]. Assume that for each sample (i,j) in the pre-compressed image, axial and lateral displacement values are a and l as illustrated in Fig. 2.1. Our goal is to quantify the accuracy of the displacement estimates.

NCC is a widely used similarity metric in TDE and it evaluates the degree of similarity between pre- and post-compressed images. The main advantage of NCC is its sensitivity to linear changes in the intensity of the two images compared to the ordinary cross correlation metric [58]. In addition, NCC is confined to be in the range between -1 and 1, which avoids dealing with very small or large values in training a classifier. We, therefore, use the NCC profile as our feature set.

Let I_1 and I_2 be two matrices of dimension (m×n) representing the pre- and post-compressed ultrasound images, respectively. For each sample (i,j) of the pre-compressed image, a window

centered at (i,j) is considered for the calculation of the NCC. Suppose that $\mathbf{X}=(x,y)$ is a vector containing sample coordinates such that $(x, y) \in \{i-25, ..., i, ..., i+25\} \times \{j\}$ (Fig. 2.2a). Moreover, $\mathbf{Y_1}=(x,y)$ is a shifted and linearly interpolated window in the post-compressed image containing sample coordinates such that $(x,\ y) \in \{i + a_{i,j} -25, ..., i + a_{i,j}\ , ..., i + a_{i,j} +25\}\ \times \{j - l_{i,j}\ \}$. These windows and their corresponding NCC value that is called NCC_1 are illustrated in Fig. 2.2c. The goal of this scheme is to investigate the behavior of the similarity metric in the neighborhood of the estimated sample. After finding eight neighboring windows of $\mathbf{Y_1}$ which are called $\mathbf{Y_2}$, $\mathbf{Y_3}$, $\mathbf{Y_4}$, $\mathbf{Y_5}$,$\mathbf{Y_6}$, $\mathbf{Y_7}$, $\mathbf{Y_8}$, $\mathbf{Y_9}$ according to Fig. 2.2b, we calculate the corresponding NCC between each window and $\mathbf{X}$, separately. These nine NCC values of each sample (i,j) in image I_1 are called $NCC_{i,i=1,2,...,9}$, which are shown in Fig. 2.2c.

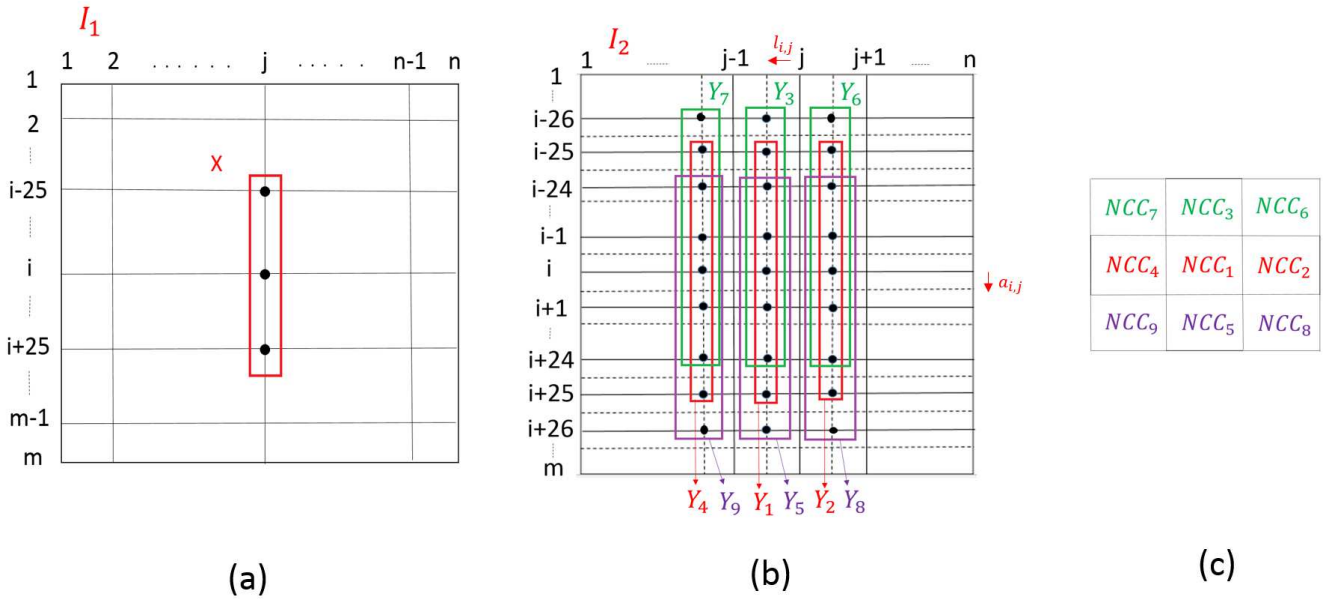


Figure 2.2: (a), (b) I_1 and I_2 are pre- and post-compressed images, respectively. Windows $\mathbf{X}$ and $\mathbf{Y_i}$, $(i = 1, 2, ..., 9)$ correspond to the sample (i,j) in I_1. (c) Nine NCC values are illustrated, NCC_1 was the only value which was used to asses accuracy.

2.2.1 Feature Selection

Fig. 1.10 indicates that the probability density function (PDF) of the peak-hopping samples is relatively randomly distributed, whereas the PDF of the correct estimated samples is more compact. Thus, the variance and skewness of the nine neighboring samples around the peak-hopping and

correct estimated samples are calculated to be considered as two features for recognizing the peak-hopping error.

Assuming that a correct TDE is available for the sample X at (i, j), NCC_1 must be larger than the other four neighboring NCC values, which are NCC_2, NCC_3, NCC_4, and NCC_5 (NCC_1 is the similarity calculated at the correct TDE). In other words, the NCC profile has a local maximum at the correct TDE as it is shown in Fig. 1.11. The steepness of the maximum and convexity of the NCC profile will be used for identification of the jitter error.

2.2.2 Supervised Learning

In this work, Support Vector Machine (SVM) [59] is used as a binary classifier to find the accuracy map of the elastography algorithm. SVM is a supervised classification method, and as such, requires training data. SVM performs non-linear classification and, therefore, often outperforms linear classification techniques. In addition, while training a SVM is computationally intensive, the testing stage is very fast. This is ideal in ultrasound elastography, wherein SVM can be trained offline and test the TDE results in real-time. The flow chart of our approach for obtaining the training data is shown in Fig. 2.3. Preparation of the training data, shown as true and false classes in this figure, is described in the following section.

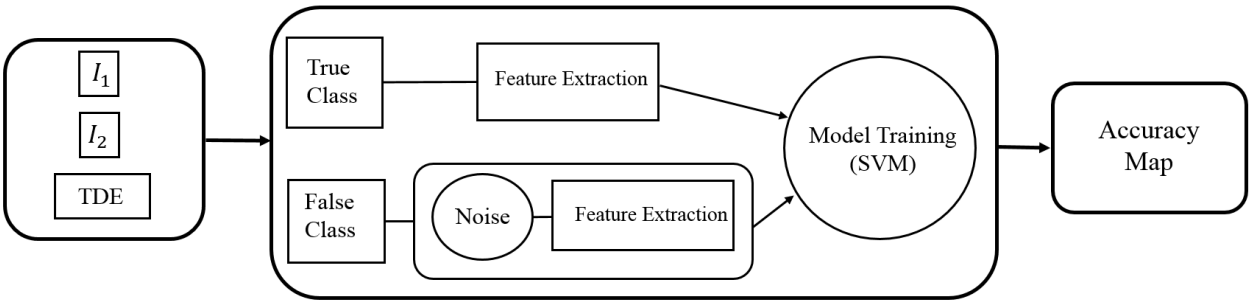


Figure 2.3: Flow chart of the proposed method. I_1 and I_2 are pre- and post-compressed images, respectively. Displacement field is either available from the simulation data, or calculated from a displacement estimation algorithm.

2.2.3 Training Sets

For simulation data, the ground truth TDE is available from our finite element simulation. For phantom experiments and *in-vivo* data, the ground truth TDE is not available. Therefore, we use a real-time elastography algorithm [60] to find TDE between pre- and post-compression RF data to obtain a silver standard. This method is based on dynamic programming and analytic minimization (DPAM) of a regularized cost function. We visually inspect the results to make sure that the algorithm has successfully calculated the displacement field. Therefore, for all the samples in image I_1, seven features are extracted by the displacement field that is available from either FEM or DPAM (Fig. 2.3).

Next, we generate samples for both training and testing that represent either the peak-hopping or jitter errors. For peak-hopping, we find samples in I_1 which have larger NCC value somewhere far from the correct estimated displacement. For jitter, uniform noise is added to the samples of the correct displacement field. These samples represent the incorrect displacement field data, and for all of these samples, we calculate seven features similar to the true class (Fig. 2.3).

2.2.4 Classification

The main idea in the proposed scheme is to employ information of the neighboring pixels to evaluate the accuracy of TDE. Therefore, instead of using one NCC value that is called NCC_1 as the only feature, four neighboring NCC values in addition to skewness and variance of nine NCC values are used according to Fig. 2.2(c). Therefore, for each sample (i,j) in I_1, seven features are calculated for training and validation procedure.

2.3 Results

The proposed scheme is implemented in MATLAB and is evaluated employing simulated, phantom and clinical data. In this study, all the samples of the pre-compressed image are considered as

the training and testing data set. For each sample (i,j) in I_1, windows of size 51 by 1 are taken in calculating the NCC values. As discussed earlier, the true class is constructed by the available ground truth or by utilizing the DPAM method. False class samples are formed by finding the peak-hopping samples or adding a uniform random noise to the rest of samples in the axial and lateral displacements. The uniform noise is in the range of [0.4 0.6] sample in the axial and [-0.1 0.1] sample in the lateral direction. Therefore, there is a minimum of 0.4 sample error in the axial direction in the false class.

2.3.1 Running time

A critical and computationally expensive step in SVM classification is finding the support vectors, i.e. the training samples that are close to the decision boundary. As a result, training the proposed method on 100,000 samples takes 381.02 sec on a single core of an i7 3.4 GHz Intel CPU. This training can be performed offline. The testing step of SVM is usually very fast, since the decision boundary is determined in the training step. In our implementation, evaluating the accuracy of an image of size 100×100 takes 0.78 sec on the same CPU, which means that the method can show an accuracy map in real-time by optimizing our implementation and using parallel computing resources.

2.3.2 Simulation Data

Ultrasound data has been created by Field II [61], and the digital phantoms are deformed using the Finite Element Method (FEM) by ABAQUS Software (Providence, RI). The parameters of the ultrasound probe are as follows: The sampling rate is 40 MHz and probe frequency is 7.27 MHz, and the fractional bandwidth is 60%. A Hanning window is used for apodization, the single transmit focus is at 22.5 mm, equi-distance receive foci are from 5 mm to 45 mm at each 5 mm, the transmit is sequential, and the number of active elements is 64. This phantom is uniform and is assumed to be isotropic and homogeneous.

Table 2.1: Classification accuracy using one NCC value and the proposed method for the simulation data.

Data set	1-NCC	7-Features
Simulation data	93.9	98.4

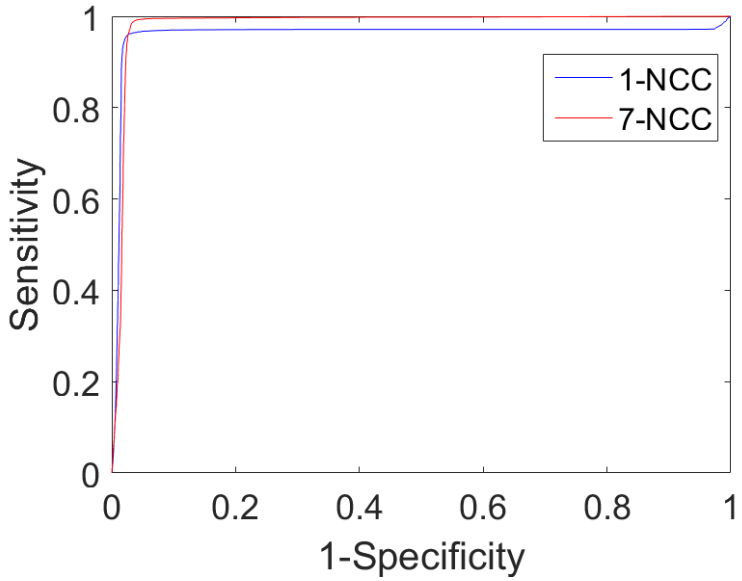


Figure 2.4: ROC curve for the simulation data.

In order to find the accuracy of the classifier, a 10-fold cross validation is performed for using one NCC value and seven proposed features (Table 2.1). The results show that accuracy of the classifier is substantially increased from about 94% to 98%. Receiver Operating Characteristic (ROC) [62] curves for the simulation data set are also shown in Fig. 2.4. It shows that the area under the curve for using one NCC value and the proposed method are respectively 0.9606 and 0.9826, a substantial improvement achieved by using the NCC profile.

2.3.3 Phantom Data

An Antares Siemens system (Issaquah, WA) at the center frequency of 6.67 MHz with a VF10-5 linear array at a sampling rate of 40 MHz was used to acquire RF data at John Hopkins University. RF data is collected from freehand palpation of a CIRS breast elastography phantom (CIRS, Norfolk, VA) with a lesion three times stiffer than the surrounding tissue. This data is available online

Table 2.2: Classification accuracy using one NCC value and the proposed method for the CIRS phantom data.

Data set	1-NCC	7-Features
Phantom data	82.2	96.7

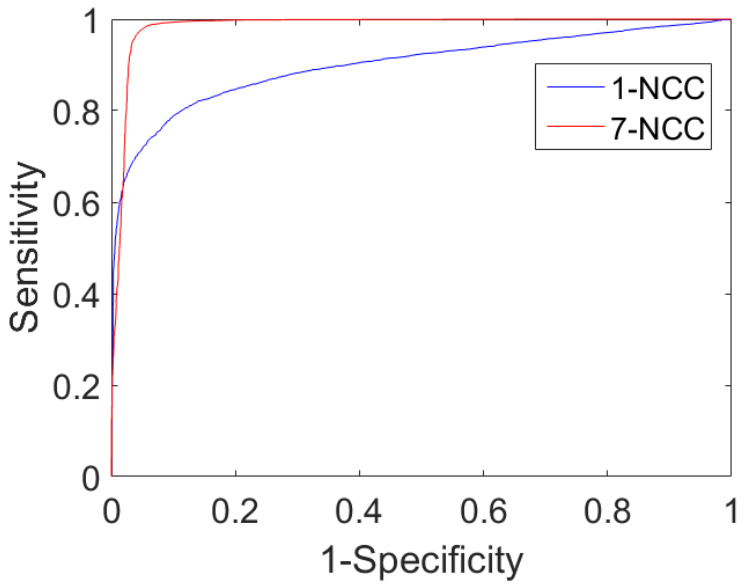


Figure 2.5: ROC curve for the CIRS phantom data.

along with publication [60].

The DPAM method [60] is utilized to calculate the displacement field. In Table 2.2, the overall accuracy of the proposed classifier by using one NCC value and seven NCC values is shown. The substantial improvement from 82.2% to 96.7% is achieved by using the seven features. The ROC curves are shown in Fig. 2.5. The areas under curve for the one NCC method and the proposed method are, respectively, 0.8994 and 0.9846, showing substantial improvement.

2.3.4 *In-Vivo* Data

The RF data is collected by ablation therapy of three patients with liver cancer using an Antares Siemens (Issaquah, WA) ultrasound machine and A VF10-5 linear array at the center frequency of 6.67 MHz with a sampling rate of 40 MHz for the RF data acquisition. The displacement matrix

Table 2.3: Classification accuracy using one NCC value and proposed seven features for three patient data.

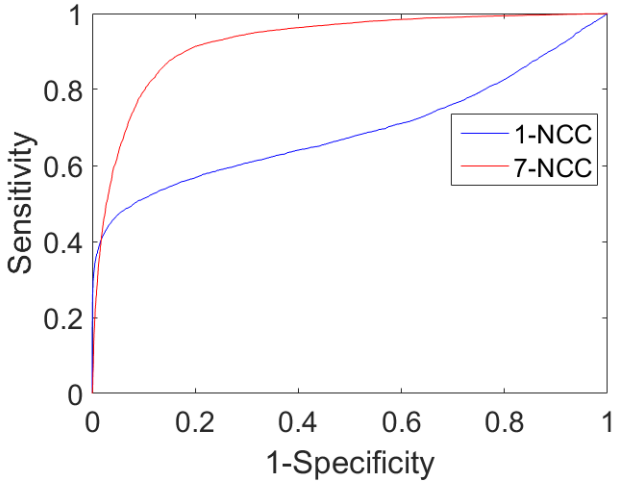
Data set	1-NCC	7-Features
Patient data 1	67.6	88.6
Patient data 2	79.3	92.2
Patient data 3	90.6	98.2
Average	79.2	93.0

for all the three patients is calculated by the DPAM method [60]. This data is available online along with the publication [60].

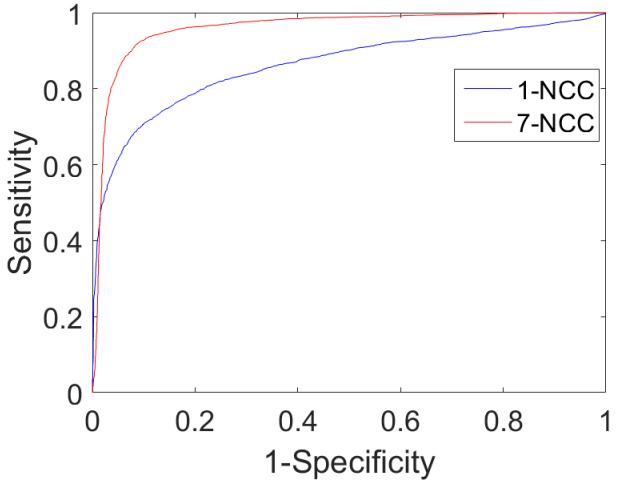
The overall classification accuracy using one NCC value and seven NCC values are depicted in Table 2.3 for all the three patients. The accuracy is improved by more than 13% on the average by using the proposed method. Also, for patient 1 with the worst available data set in terms of quality of RF data, it shows a very high improvement of about 21%. Finally, the ROC curves for all the three patients are depicted in Fig. 2.6. The area under the ROC curve is increased from 0.6906 to 0.9255 for patient 1, from 0.8624 to 0.9598 for patient 2 and from 0.9610 to 0.9926 for patient 3. The significant improvement in the results is obtained by exploiting the NCC profile around the estimated displacement field.

2.3.5 Accuracy Map in Region of Interest (ROI)

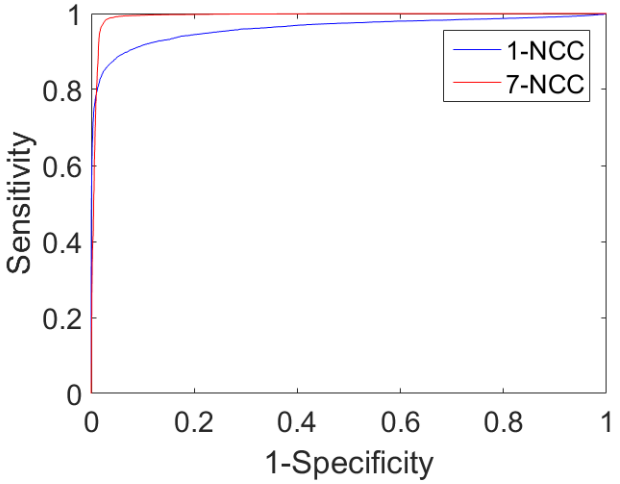
The proposed method has been quantitatively validated in the previous sections. To visualize the performance of the proposed method, we show the map of the accuracy assessment. In Figs. 2.7 (a) and (b), the red boxes in the figures indicate the tumor region in the B-mode ultrasound strain images for patient 3. The displacement field is calculated using DPAM and is visually checked. Therefore, all samples belong to the true set and the classifier should label those pixels as the true displacement estimate. Figs. 2.8 (a) and (b) show the results of the proposed method using one NCC and seven NCC values. Blue samples denote successfully classified regions and yellow samples show the unsuccessfully classified samples. In a second experiment, all the samples are either peak-hopping samples or are corrupted by the jitter error. Therefore, the algorithm should classify all the samples



(a) patient 1

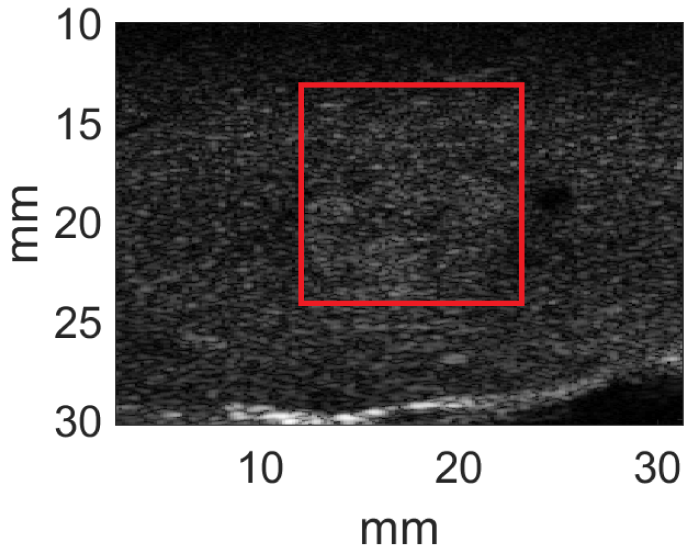


(b) patient 2

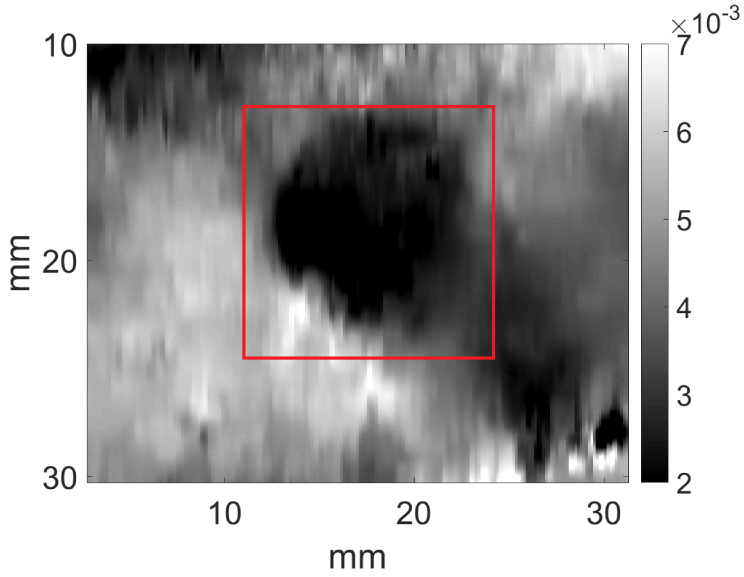


(c) patient 3

Figure 2.6: ROC curves for three patients.

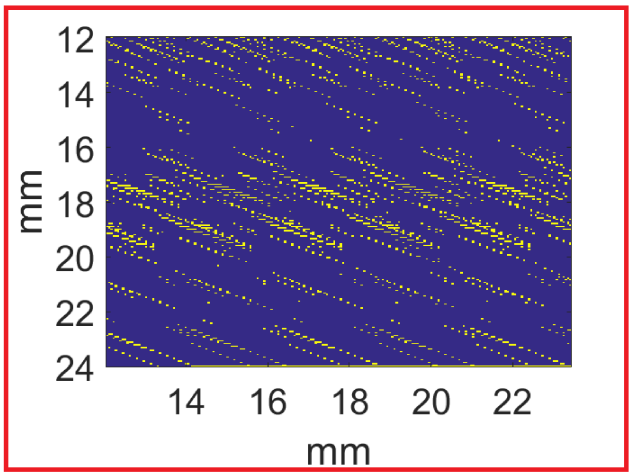


(a) Bmode Image

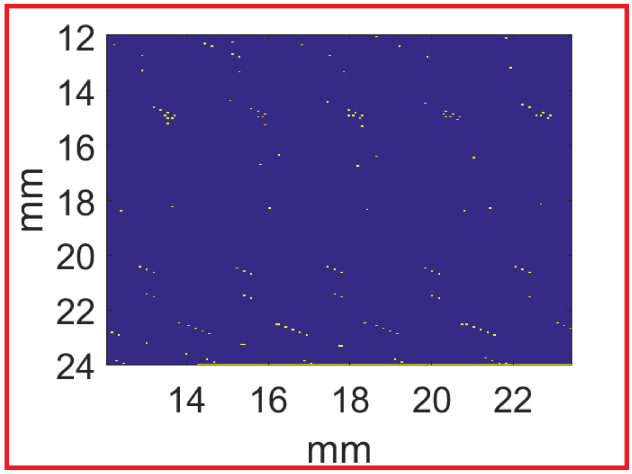


(b) Strain Image

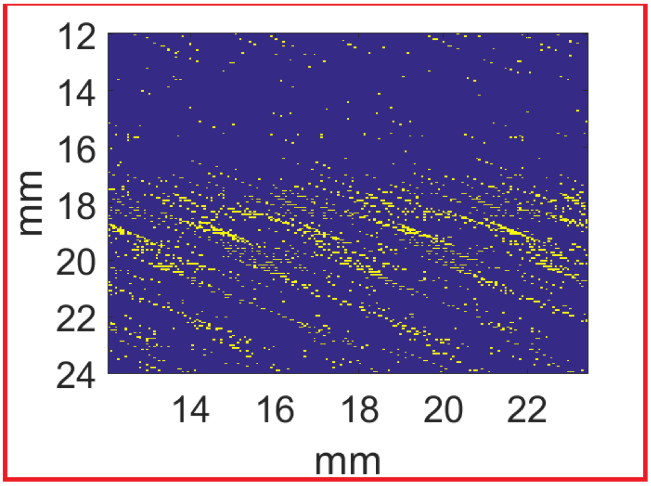
Figure 2.7: Accuracy map in tumor region for patient 3 is shown in red boxes.



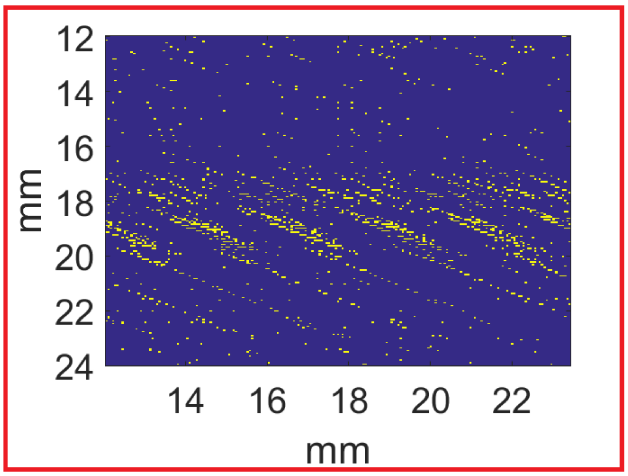
(a) TP1



(b) TP7



(c) FP1



(d) FP7

Figure 2.8: Accuracy map in tumor region for patient 3 is shown in red boxes for true and false positive cases.

as incorrect displacement. Figs. 2.8 (c) and (d) show the result of the proposed method for using one NCC and seven NCC values, respectively. Again, blue and yellow respectively represent correct and incorrect classification. The results of this figure clearly demonstrate that the proposed method substantially improves the results.

The accuracy maps are created for the True Positive (TP) and False Positive (FP) cases around the tumor. The tumor region is chosen as region of interest to locally show the performance of the classifiers using one NCC value or the seven proposed features. It is important to note that the samples in this region (ie the vicinity of the tumor) are not used for training and are only used in the testing stage. In the TP case, all the samples in the tumor region have correct displacement estimates, which are obtained using DPAM. For the FP case, all the samples of the validation set have incorrect displacement estimates in the form of either peak-hopping or jitter error. These incorrect displacements are generated by the procedure described in Section 2.1.3.

2.4 Summary

In this chapter, we have shown that accuracy assessment of TDEs in ultrasound elastography could be performed in a supervised approach using a binary classifier. Then, we have illustrated that using the set of seven proposed features substantially improves the accuracy of the supervised model. In fact, five features are utilized to recognize the jitter error and two features to distinguish the peak-hopping samples. All of these are achieved by using the information of the NCC profile around the estimated time delays.

Chapter 3

Automatic Accuracy Assessment of Ultrasound Elastography Using Correlation Profile and Prior Information of Displacement Continuity

3.1 introduction

The proposed method in the previous chapter takes the advantage of using *NCC profile* with 7 valuable features and also looks at accuracy assessment in a supervised approach. In this chapter, we build on our previous work by utilizing continuity features in axial and lateral directions as useful information. We show that these features improve the sensitivity and specificity of the classifier. After extracting the continuity features in addition to features proposed in our previous work, we train a linear Support Vector Machine (SVM) on available data sets to show the significant improvement of utilizing the proposed features. Also, we use Global Time-Delay Estimation in Ultrasound Elastography (GLUE) [63] as a novel elastography algorithm to obtain silver standard ground truth

for training the model which means that the proposed supervised approach can be re-trained by any accurate elastography methods in future.

3.2 Method and Material

Let I_1 and I_2 be pre-and post-compressed images in Fig. 3.1, and the displacement field obtained by an elastography algorithm corresponds to samples in I_1 and I_2 images. For each sample (i,j) in the pre-compressed image, a and l show axial and lateral displacement values, respectively in Fig. 3.1. Accuracy of TDE is evaluated in a supervised approach using information of NCC profile in [54] and seven valuable features are proposed for training the model.

Continuity (i.e. lack of rupture) in real tissue implies that TDEs should not change significantly in a small region in tissue. Therefore, for each sample (i,j) in I_1, a window of size 3 by 3 is taken into account in displacement field so that variance of those nine TDE values in axial and lateral directions are taken as continuity features according to Fig. 3.2. This means in a case that peak-hopping sample happens, TDE is estimated far from the correct TDE and these variances will be increased incredibly in which peak-hopping samples will be simply classified as incorrect TDE.

3.2.1 Supervised Learning

In this chapter, Support Vector Machine (SVM) is trained for the supervised quality assessment of time-delay estimates. SVM is a supervised classification method and is utilized as binary classifier to categorize correct and incorrect estimated time-delays. Therefore, it requires training an testing samples categorized as true and false classes which are formed according to the scheme that is thoroughly described in the previous chapter and in [54].

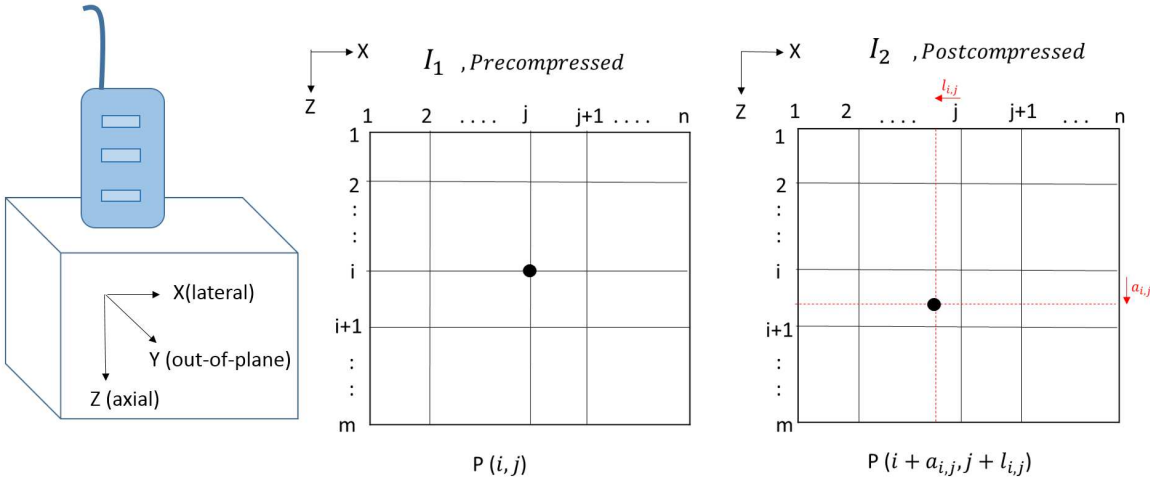


Figure 3.1: Displacement between pre-and post compressed images. I_1 and I_2 are pre- and post-compressed images, respectively. Z, X, Y are axial, lateral and out-of-plane directions, respectively. The coordinate system is attached to the ultrasound probe. The sample P (i,j) in I_1 has been moved by ($a_{i,j}$,$l_{i,j}$) in I_2.

3.2.2 Training Sets

For phantom and *in-vivo* data, the ground truth TDE is not available. In [54], a method based on DPAM of a regularized cost function was used. Although the method was accurate enough, we want to show that the proposed scheme would work even in case of using another accurate method in obtaining ground truth of TDE for patient data. Therefore, we utilize the new and more reliable proposed elastography algorithm in [63] to find the correct TDE as ground truth. The method known as Global Ultrasound Elastography (GLUE), finds all TDE values for all RF frames simultaneously by using a non-linear cost function and is optimized in an efficient way. In addition, the quality of strain images are visually checked before using GLUE TDEs to ensure that GLUE is successful in providing accurate displacement fields [55].

3.2.3 Classification

The main purpose of this work is to add continuity features to those seven valuable features discussed in [54] to improve the quality of supervised classification. Therefore, for each sample (i,j)

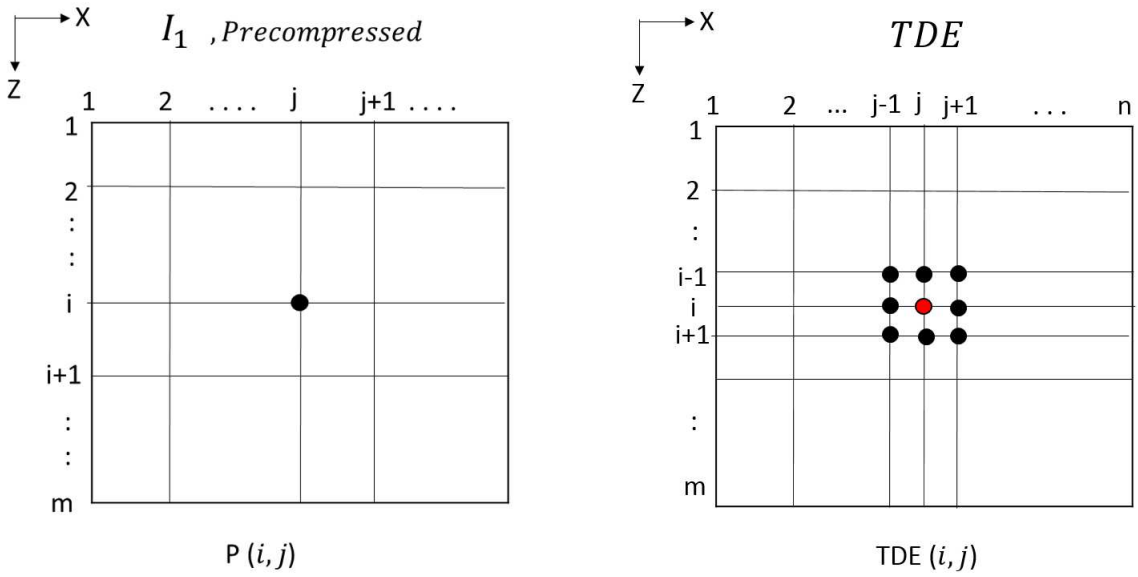


Figure 3.2: Pre-compressed image on the left and the corresponding TDE between I_1 and I_2 on the right. For sample P (i,j) in I_1, eight neighboring TDE values are shown as black points on the right image.

in I_1, nine features are calculated to train and validate the proposed model. In Fig. 3.3, flowchart of the proposed method is shown to better illustrate different stages of the supervised accuracy assessment.

3.3 Results

This work is implemented in MATLAB and is tested on phantom and three clinical data sets [55]. This approach makes use of all samples of I_1 in training and testing. This means that for each sample (i,j) in I_1, windows of size 51 by 1 are considered for calculating the NCC values to obtain seven features discussed in [54] and then variances of corresponding samples in displacement field are calculated. The training procedure described in this work is similar to the model that is proposed in the previous chapter. For training, true and false classes are needed. Therefore, true class is formed by using GLUE method and false class is constructed by finding peak-hopping samples and adding an uniform noise to the rest of the samples in TDEs in both axial and lateral directions. The uniform noise is in the range of [0.4 0.6] sample in the axial and [-0.1 0.1] sample in the lateral direction

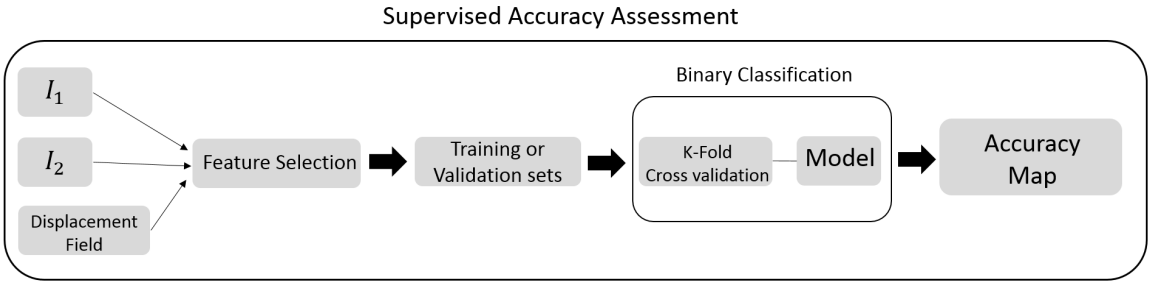


Figure 3.3: Flowchart of the proposed method. I_1 and I_2 are pre- and post-compressed images, respectively. Displacement field is available from a displacement estimation algorithm. In this flowchart, feature selection section involves extracting 9 features instead of 7 features proposed in Fig 2.3 and Chapter II.

meaning that there is always a minimum 0.4 sample error in the axial direction for all samples in false class.

In order to find the accuracy of the classifier for each data set, 10-fold cross validation is performed for using one NCC value and seven proposed features. In 10-fold cross validation, the original data set is partitioned randomly into 10 equal sized subsets. Each time, one subset is used as the validation set to test the accuracy of the classifier, and the remaining 9 subsets are utilized as the training set. This will ensure that the algorithm is not trained and tested on the same data. In the final step, the averaged accuracy of all 10 experiments is calculated as the final accuracy of the classifier.

3.3.1 Data Sets

The phantom and three *In-Vivo* data sets are the same data sets that are used in chapter II. The RF data was acquired at Johns Hopkins University using an Antares Siemens (Issaquah, WA) ultrasound machine and a VF10-5 linear array at the center frequency of 6.67 MHz with a sampling rate of 40 MHz. All three *in-vivo* data sets are obtained from ablation therapy of the patients with liver cancer. Phantom data is collected from freehand palpation of a CIRS breast elastography phantom (CIRS, Norfolk, VA) with a lesion three times stiffer than the surrounding tissue. This data is available online along with the publication [60]. As mentioned before, the displacement field for phantom

Table 3.1: Classification accuracy using one NCC value, 7 features, and 9 features for phantom and the three patient data using SVM classifier.

Data set	1-NCC	7-Features	9-Features
Phantom data	77.26	95.11	95.25
Patient data 1	66.7	82.7	84.1
Patient data 2	69.2	83.9	85.2
Patient data 3	81.2	91.5	95.6

and all the three patients is calculated utilizing the GLUE method [63].

The overall classification accuracy using one NCC value, seven and nine features are shown in Table 3.1 for phantom data and all the three patients. The accuracy is improved by more than 14% on the average by using the seven features and more than 16 % on the average by using the nine features. In first and second patient data, peak-hopping samples in forming the false class involves 2% of all incorrect samples and rest of them are jitter samples. However, we added more peak-hopping samples to the third patient's false class (10% of the false class) to show the power of the two proposed features. Although using the nine features for the third patient is improving the results, it has the disadvantage of requiring a larger training sample (i.e. peak-hopping samples), which comes with increased computational cost. Last but not least, the Receiver Operating Characteristic (ROC) curves for phantom and all the three patients are depicted in Fig. 3.4 to 3.7. The area under the ROC curve is increased from 0.8548 to 0.9821 for phantom data, from 0.6932 to 0.8965 for patient 1, from 0.7477 to 0.9008 for patient 2 and from 0.8777 to 0.9707 for patient 3 by using the nine features. The improvements achieved by utilizing the information of the NCC profile around the time-delay estimate and locally continuity properties of the displacement field due to continuity of the displacement field.

3.3.2 Accuracy Map in Region of Interest (ROI)

Quantitative validation of the proposed model and features has been done in the previous sections. Herein, the accuracy map of the scheme is visualized for patient 3. The red boxes in Figs. 3.8 (a)

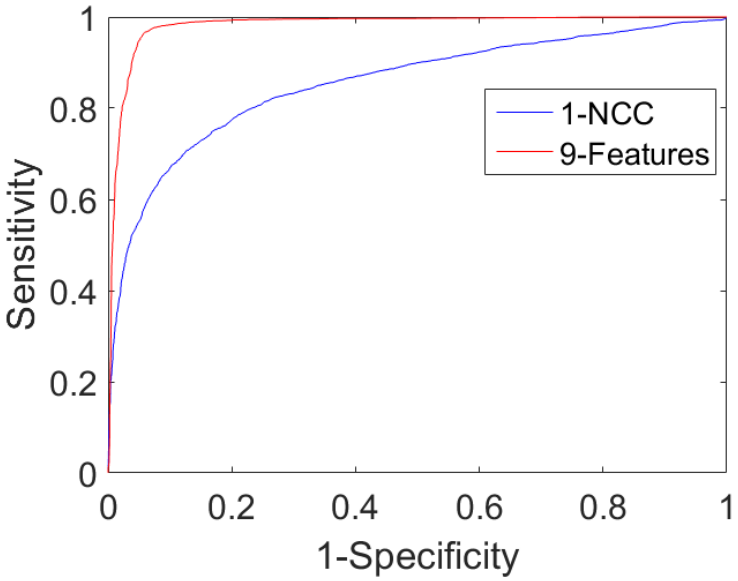


Figure 3.4: ROC curve for phantom data.

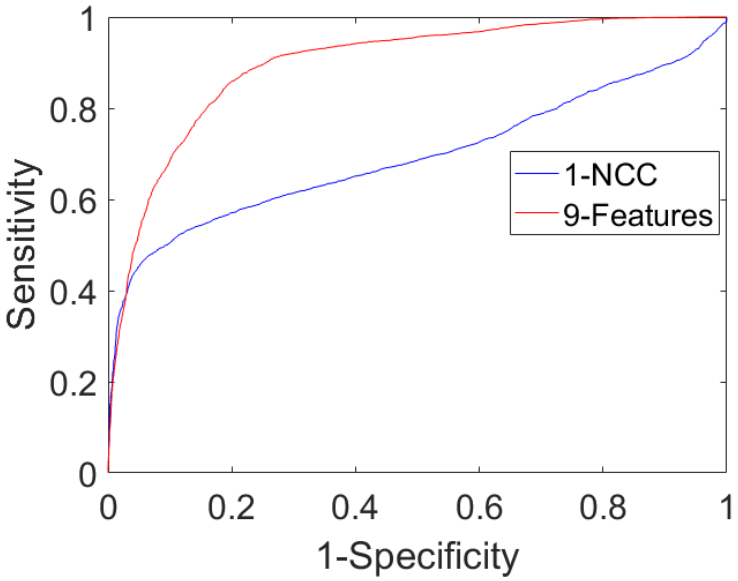


Figure 3.5: ROC curve for patient 1.

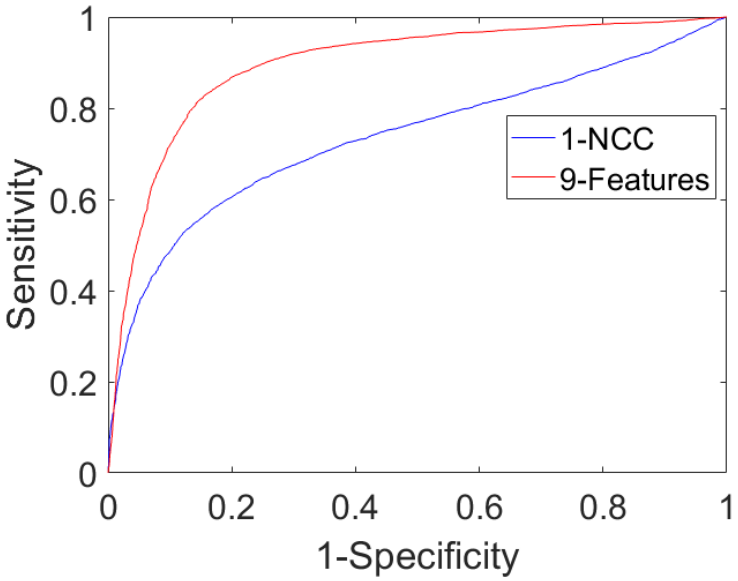


Figure 3.6: ROC curve for patient 2.

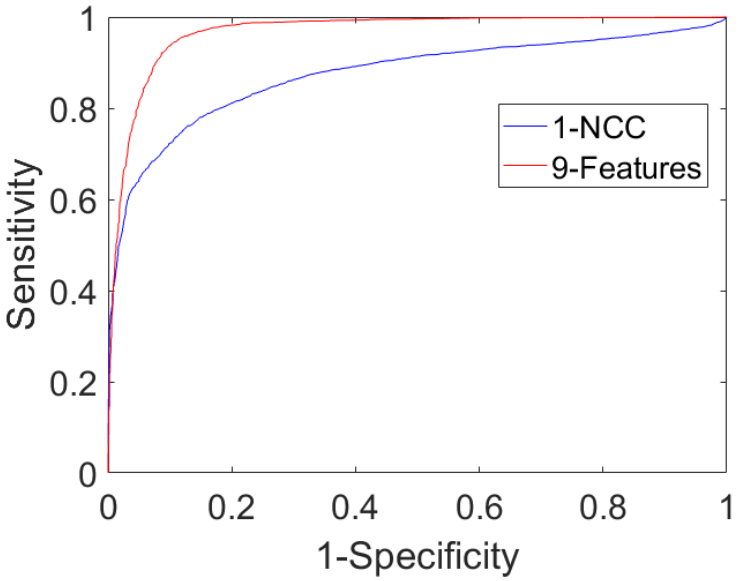
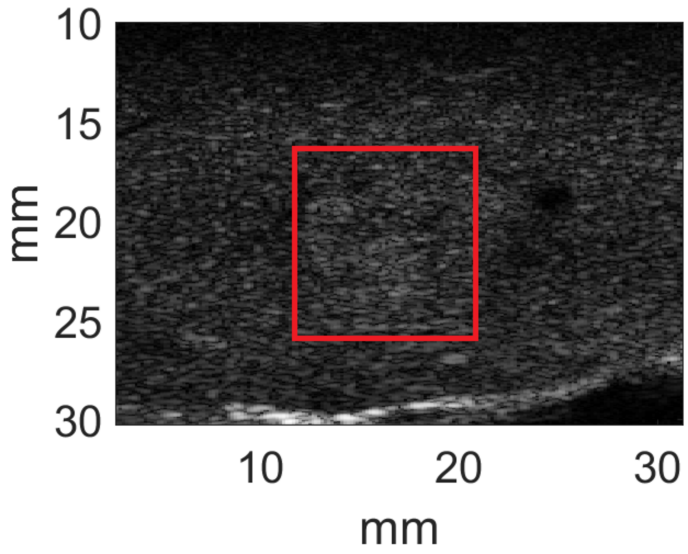
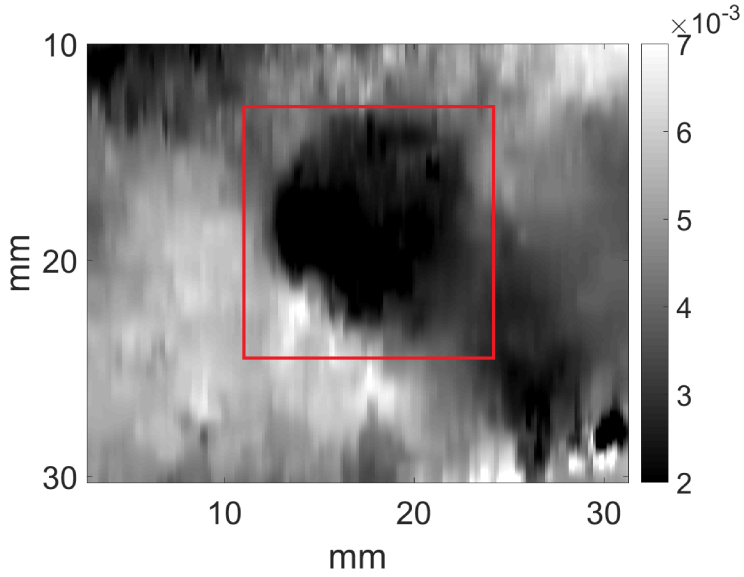


Figure 3.7: ROC curve for patient 3.

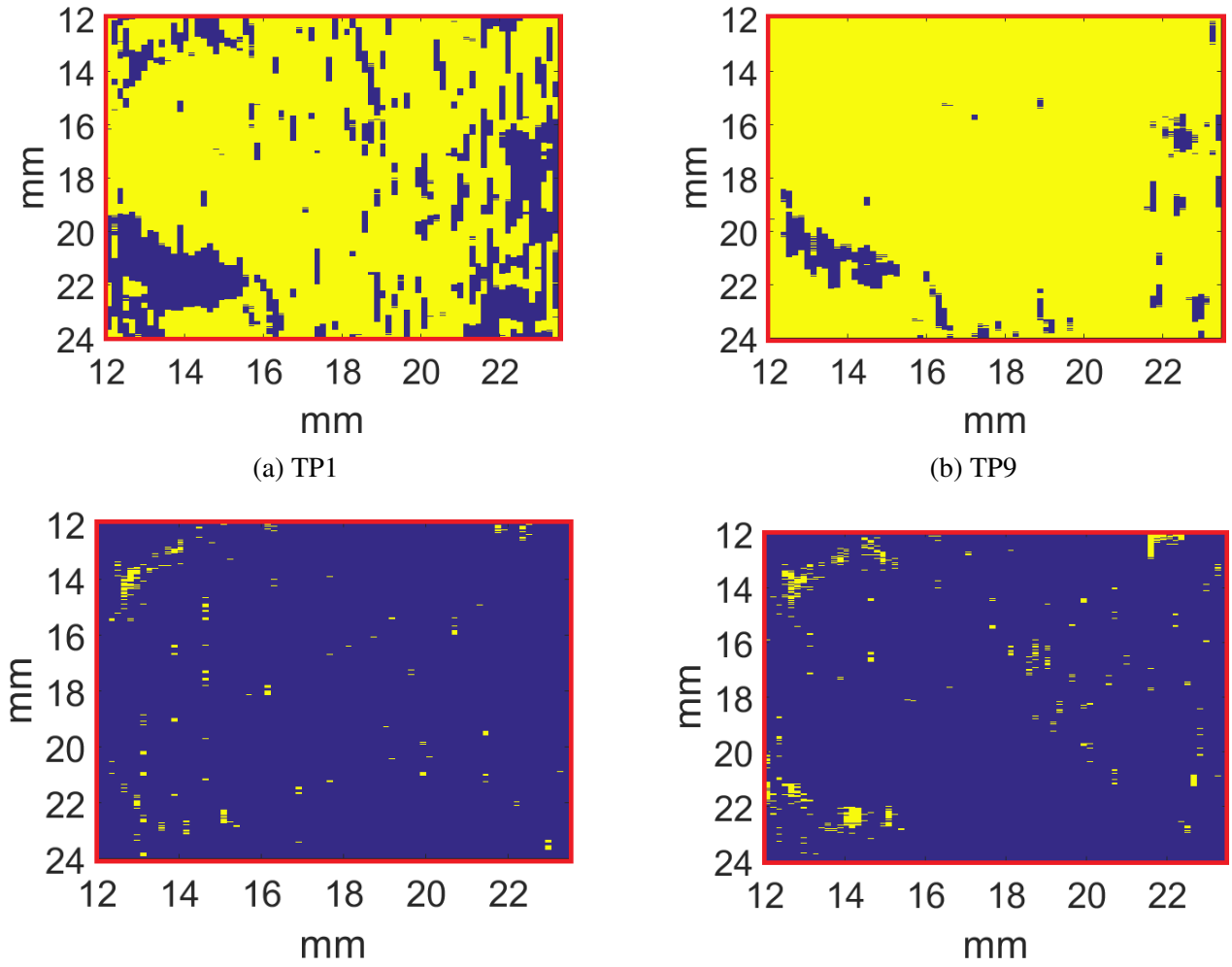


(a) Bmode Image



(b) Strain Image-glue

Figure 3.8: Accuracy map in tumor region for patient 3 is shown in red boxes.



(a) TP1

(b) TP9

(c) FP1

(d) FP9

Figure 3.9: Accuracy map in tumor region for patient 3 is shown in red boxes for true and false positive cases.

and (b), show tumor region in B-mode ultrasound and strain image, respectively. To visualize the performance of the proposed method, two experiments are performed. First, the displacement field for the tumor region is obtained by GLUE and is visually checked. Therefore, the binary classifier should recognize them as true displacement estimate. In Figs. 3.9 (a) and (b), results of using one NCC and nine features are shown. Yellow samples indicate correct classified regions and blue samples show incorrect classified samples. Second, all samples in tumor region are peak-hopping or corrupted by noise as jitter samples. Therefore, the proposed method should classify them as incorrect estimates. Figs. 3.9 (c) and (d) show the result of using the method for using one NCC and nine features values, respectively. Correct and incorrect classifications are indicated in yellow and blue, respectively. The results distinctly demonstrate that using the nine features outperforms utilizing only one NCC value.

3.4 Summary

In this chapter, continuity properties of the tissue enabled us to introduce two more features in addition to the seven proposed ones in the previous chapter to better recognize incorrect estimated TDEs. The two features were the variance of TDEs in axial and lateral directions which improved the accuracy of the classifiers. Also, we used GLUE as a new method to get the ground truth TDE for training and testing our models on available data sets. The results showed that the proposed model is not biased to any specific elastography algorithms in a sense that it could be trained by other elastography methods in future.

Chapter 4

Performance Evaluation of Well-known Classifiers and Importance Measurement of the Proposed Features

4.1 Introduction

In this chapter, we evaluate the performance of several well-established classifiers in accuracy assessment of time delay estimation in Python. Binary classifiers such as Support Vector Machine (SVM) [59], Random Forest [56], Artificial Neural Network (ANN) [64], Adaboost [65], Naive Bayes [66], Decision Tree [67], K-Nearest Neighbor (KNN) [68], Linear Discriminant Analysis (LDA) [69] are used to find the best model for accuracy assessment of TDE [70]. We later show that Random Forest outperforms all other classifiers using the nine proposed features in the previous chapter.

Table 4.1: Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the simulation data. The variance of the classifier is shown in the parentheses beside the mean value.

Classifier	1-NCC	7-Features	9-Features
Random Forest	96.36 (0.01)	97.67 (0)	**99.31 (0)**
SVM	95.36 (0.01)	97.8 (0)	97.82 (0)
Adaboost	96.36 (0.01)	97.9 (0)	97.82 (0)
ANN	96.71 (0.01)	97.14 (0)	97.09 (0)
Naive Bayes	72.21 (0.22)	98.09 (0)	96.18 (0.02)
Decision Tree	96.26 (0.01)	97.29 (0)	99.18 (0)
KNN	96.71 (0.01)	97.14 (0)	97.09 (0)
LDA	67.85 (0.3)	97.9 (0)	97.91 (0)

4.2 Accuracy of Classifiers

The predictive models are trained and tested by the seven and nine proposed features in [54] and [55]. The nine proposed features include the seven features discussed in Chapter II and the two continuity features described in Chapter III. We evaluate the performance of the classifiers using simulation, phantom and the three patient data sets. FEM and DPAM are used to obtain the correct displacement for training the supervised model. We used DPAM for phantom and patient data sets instead of GLUE since it provided better classification accuracy in all data sets [54, 55]. Also, 10-fold cross validation is used to avoid training and testing on the same data. Tables 4.1 to 4.5 show the mean and variance of the accuracy of the classifiers using 1-NCC, 7 features and 9 features for simulation, phantom and three patient data sets, respectively. The results show that the accuracy increases as we add more features, and Random Forest is outperforming compared to other classifiers in all five data sets.

Table 4.2: Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the phantom data. The variance of the classifier is shown in the parentheses beside the mean value.

Classifier	1-NCC	7-Features	9-Features
Random Forest	74.2 (0.15)	94.89 (0.05)	**96.52 (0.03)**
SVM	79.54 (0.17)	95.68 (0.04)	96.5 (0.03)
Adaboost	82.04 (0.2)	93.35 (0.07)	95.4 (0.04)
ANN	80.05 (0.2)	95.27 (0.04)	96.01 (0.03)
Naive Bayes	73.83 (0.2)	94.95 (0.03)	76.21 (0.25)
Decision Tree	74.27 (0.16)	92.39 (0.05)	95.42 (0.04)
KNN	81.04 (0.2)	95.5 (0.04)	96.18 (0.03)
LDA	74.5 (0.2)	95.4 (0.03)	96.47 (0.03)

Table 4.3: Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 1. The variance of the classifier is shown in the parentheses beside the mean value.

Classifier	1-NCC	7-Features	9-Features
Random Forest	56.34 (0.17)	77.67 (0.22)	**92.52 (0.05)**
SVM	58.2 (0.1)	81.1 (0.18)	89.71 (0.17)
Adaboost	57.84 (0.32)	81.49 (0.21)	89.53 (0.09)
ANN	58.18 (0.3)	77.18 (0.23)	92.41 (0.05)
Naive Bayes	34.4 (0.2)	79.14 (0.21)	77.44 (0.2)
Decision Tree	57.2 (0.17)	76.51 (0.2)	90.14 (0.06)
KNN	56.43 (0.25)	77.65 (0.22)	83.54 (0.15)
LDA	25.6 (0.17)	81.77 (0.17)	81.82 (0.17)

Table 4.4: Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 2. The variance of the classifier is shown in the parentheses beside the mean value.

Classifier	1-NCC	7-Features	9-Features
Random Forest	70.72 (0.04)	88.65 (0.04)	**94.16 (0.03)**
SVM	75.02 (0.06)	90 (0.04)	90.27 (0.04)
Adaboost	78.67 (0.1)	89.75 (0.04)	89.47 (0.04)
ANN	77.45 (0.08)	88.55 (0.4)	88.91 (0.04)
Naive Bayes	65.46 (0.2)	88.48 (0.06)	57.52 (0.41)
Decision Tree	70.6 (0.06)	84.01 (0.04)	91.24 (0.04)
KNN	77.6 (0.08)	89.39 (0.04)	89.53 (0.04)
LDA	66.57 (0.15)	89.29 (0.04)	89.36 (0.04)

Table 4.5: Mean and variance of the accuracy of the classifiers using one NCC value, 7 features, and 9 features for the patient 3. The variance of the classifier is shown in the parentheses beside the mean value.

Classifier	1-NCC	7-Features	9-Features
Random Forest	87.3 (0.17)	97.37 (0.22)	**99.99 (0.02)**
SVM	90.13(0.06)	97.58 (0.01)	97.86 (0.01)
Adaboost	90.71 (0.07)	97.29 (0.02)	99.99 (0.00)
ANN	90.72 (0.07)	97.31 (0.02)	97.69 (0.01)
Naive Bayes	86.08 (0.09)	97.24 (0.1)	73.69 (0.26)
Decision Tree	87.38 (0.07)	96.03 (0.02)	99.99 (0.00)
KNN	90.62 (0.07)	97.55 (0.01)	97.78 (0.01)
LDA	85.65 (0.09)	97.5 (0.01)	97.56 (0.01)

4.3 Feature Importance Measurement

Feature importance measurement for all five available data sets is shown in Figs. 4.1 to 4.5. The feature indices from 0 to 8 represent NCC_1, NCC_2, NCC_3, NCC_4, NCC_5, varinace and skewness of NCC profile, variance of TDE in axial and lateral directions, respectively. These nine feature are well-described in chapter II and III. The feature importance is measured for Random Forest classifier using the Mean Decrease Impurity method described in [56, 71]. The importance of the features for simulation data is depicted in Fig. 4.1, it shows that NCC_1, NCC_1, NCC_3, skewness of NCC profile are the most informative features for simulation data. Fig. 4.2 also shows that NCC_1, NCC_2, NCC_3, varinace and skewness of NCC profile are the most important features for phantom data. According to Figs. 4.3 to 4.5, we could realize that NCC_1, NCC_2, NCC_3, varinace of NCC profile, and variance of TDE in lateral direction are the most informative features in training the supervised method for accuracy assessment of TDE for three patient data sets. Another point is that, NCC_1, NCC_2, NCC_3 values play an important role in distinguishability of the random forest classifier in all five data sets. This is intuitive since axial displacement field has usually higher quality and importance than lateral displacement field due to the structure of imaging systems and ultrasound probes. That is why a 3-D super resolution technique is proposed in [72] to obtain higher resolution displacement fields in other directions rather than only in the axial direction.

4.4 Summary

In this chapter, we evaluated the performance of the several well-known classifiers in the literature to assess the accuracy of the TDEs. We showed that Random Forest outperforms other classifiers in all available data sets using the nine proposed features in the previous chapter. Then, we measured the importance of the nine features for the Random Forest classifier using the Mean Decrease Impurity method. The results on three patient data sets showed that NCC_1, NCC_2, NCC_3, varinace of NCC profile, and variance of TDE in lateral direction are the most effective features in training the Random Forest classifier.

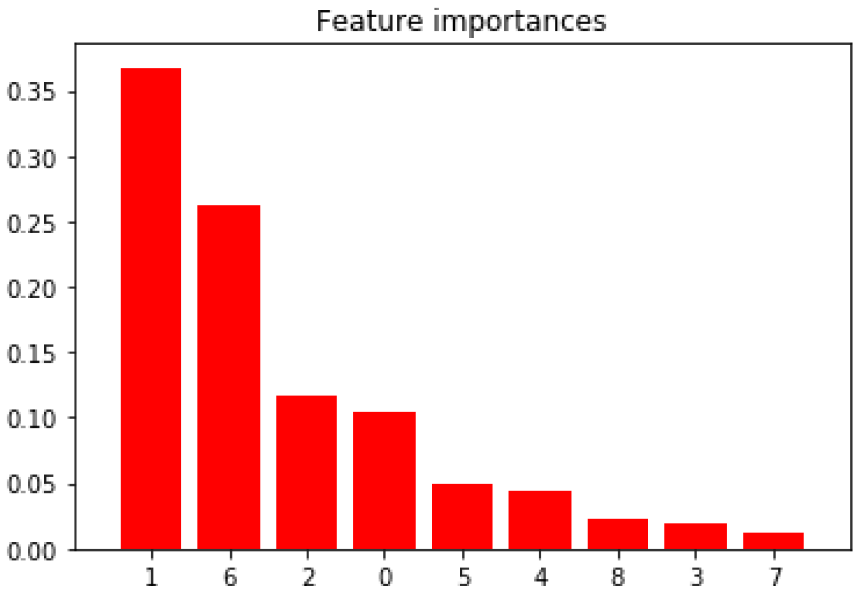


Figure 4.1: Feature importance measurement for simulation data.

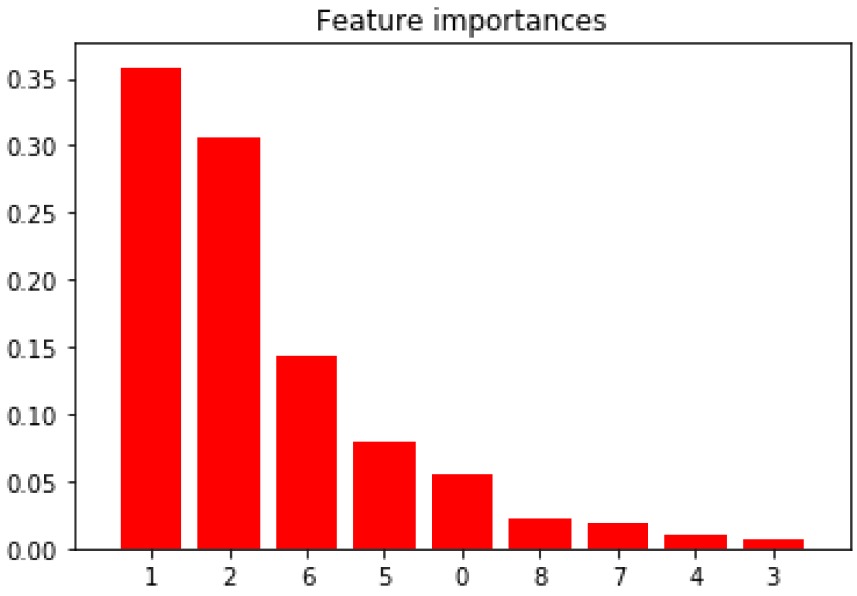


Figure 4.2: Feature importance measurement for phantom data.

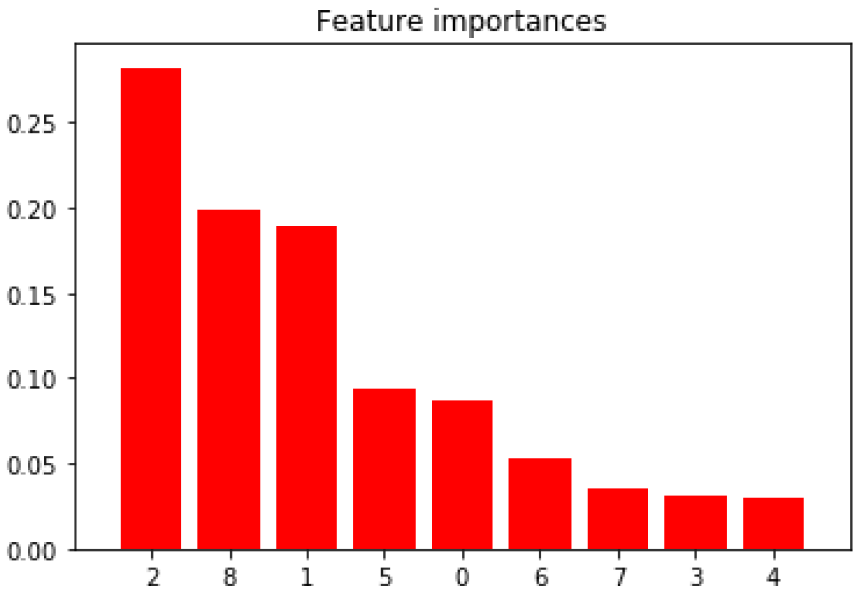


Figure 4.3: Feature importance measurement for patient 1.

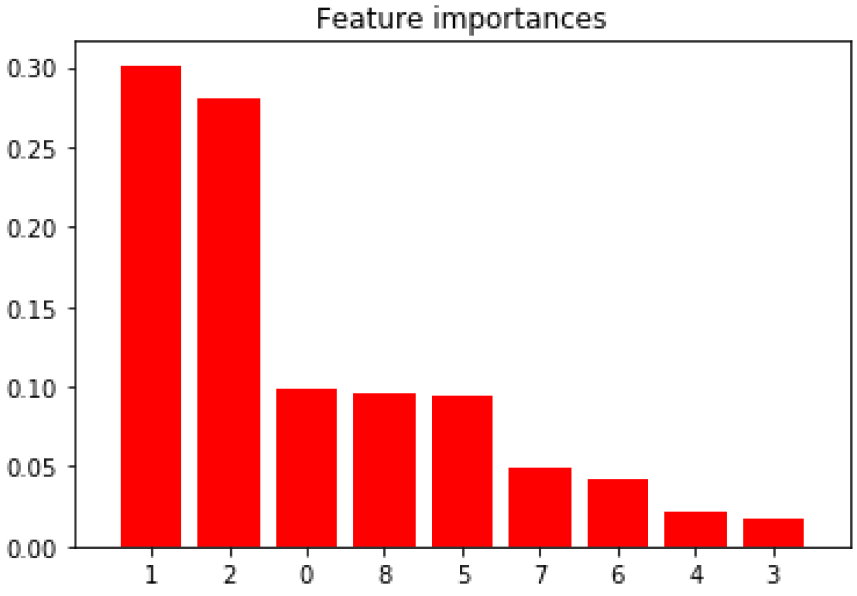


Figure 4.4: Feature importance measurement for patient 2.

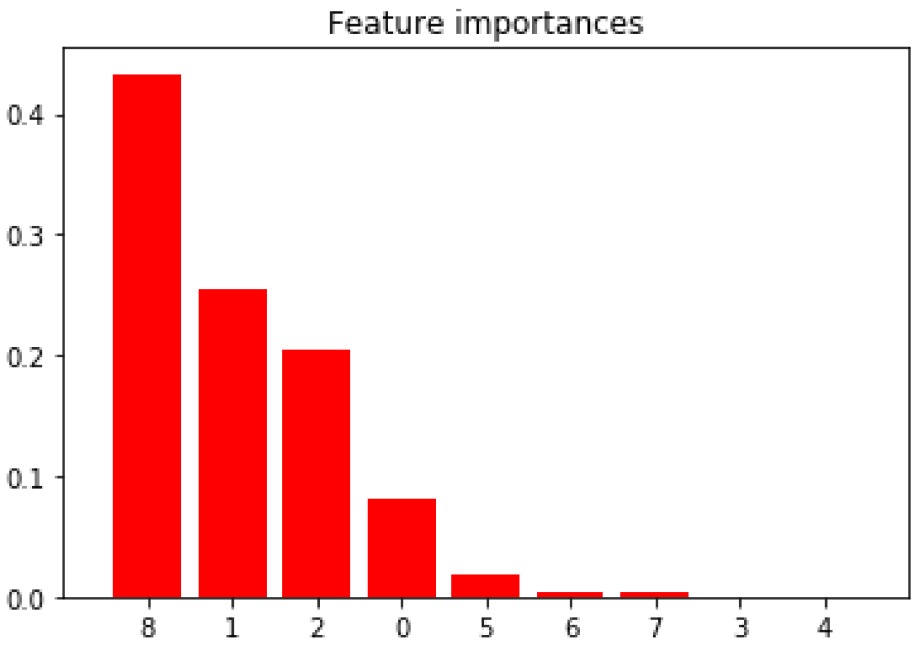


Figure 4.5: Feature importance measurement for patient 3.

Chapter 5

Discussion and Conclusion

5.1 Discussion

The information around TDE is always informative in discriminating the correct versus incorrect displacement estimates. Since estimation of the displacement is generally more difficult close to boundaries between different organs or tissue types, it is more likely to find an incorrect displacement in these regions. Therefore, our method is more invaluable in these regions wherein incorrect displacement estimates may lead to incorrect diagnosis or surgical planning.

Although there is no gold standard for displacement fields for phantom and *in-vivo* data sets, time-delay estimates are examined visually to be certain about their accuracy. In this thesis, TDEs have been obtained by DPAM or GLUE, which are a promising method to obtain accurate displacement fields. In order to show that the estimated TDEs are accurate, we here changed the axial regularization weight of DPAM by 20% in five equal steps (20 to 24). The mean squared error (MSE) for those displacement values is on average less than 0.02 sample (0.0004 mm) for all data sets (Table 5.1). The small MSE of the displacement fields quantitatively shows that the displacement estimates have a low variance. It is to be noted that the 0.02 sample variance is substantially smaller than the 0.4 to 0.6 samples that are added to the correct displacement field to generate jitter.

Table 5.1: The Mean Squared Error (MSE) of displacement values obtained using DPAM for all data sets.

Phantom	Patient1	Patient2	Patient3
0.0204	0.0172	0.0054	0.0039

In order to improve the accuracy of the classifier, seven and nine features have been proposed and utilized instead of using only one NCC value at the estimated time-delay estimate. The results have shown that using the proposed seven or nine features improves the performance of the classifier substantially. However, one might consider additional features that could improve the performance of the classifier further. Future work would utilize a convolutional neural network (CNN) [73] to investigate the possibility of having a more efficient model in terms of accuracy, training and testing time. A disadvantage of CNN is that it needs a very large training database that was not available for this work.

The proposed automatic quality assessment of TDE is a novel approach that can play an important role in industrial and academic applications of ultrasound imaging specifically in quasi-static elastography. This method can be applied as a promising approach to quantitatively compare the performance of TDE algorithms, since it is more reliable than solely using the value of NCC at the estimated displacement. Moreover, this method is suitable for real-time applications, and therefore, can be used to train sonographers to obtain higher quality strain images by displaying the accuracy map instantaneously. Unsupervised classification methods, such as k-nearest neighbors (KNN) [70] have an advantage over the proposed method in that they do not require training data. However, they need to search for nearest neighbors in high-dimensional feature spaces and are usually not suitable for real-time applications. In contrast, the proposed supervised technique based on SVM and Random Forest runs in real-time, and therefore, provides clinically a more relevant solution.

We have used DPAM [60] and GLUE [63] in Chapters II and III, respectively, to get silver standard TDEs for training and testing the proposed model. Although both the algorithms are pioneer in providing accurate time delay estimates, we have used only DPAM in Chapter IV in view of its

higher classification accuracy for all the data sets. This means that having more accurate TDEs results in more accurate models, therefore, we could better compare the performance of the state of art classifiers.

5.2 Conclusion

In Chapters II and III, we have presented a novel method for accuracy assessment of TDE using NCC profile and continuity property of TDE around the estimated displacement value, which can be used to mask out the erroneous regions of the strain image. It can further reduce the user-dependence of strain imaging and help train the sonographer. Our technique is based on SVM classifications, a nonlinear classifier that often substantially outperforms linear classifiers. Training our SVM classifier is usually computationally expensive, but it can be performed offline. Once trained, our classifier is computationally efficient and can classify the accuracy of TDE in real-time. The performance of the proposed method has been validated through simulation, phantom and *in-vivo* data. Furthermore, we here used DPAM [60] and GLUE [63] in this thesis to obtain accurate TDEs for training and testing our proposed model. However, any other reliable elastography algorithms could also be used in our proposed method meaning that the model could be trained with any accurate elastography method in future. In Chapter IV, we have compared the performance of the state of the art classifiers using the proposed nine features in accuracy assessment of TDEs. We have shown that Random Forest outperforms in all the available data sets. We have then measured the importance of the nine features, which were used in training the Random Forest classifiers using the Mean Decrease Impurity method and have illustrated that the features NCC_1, NCC_2, NCC_3, the variance of the NCC profile, and the variance of TDE in lateral direction are the most informative features. This is intuitive because axial displacement field always has higher importance and resolution than lateral displacement in ultrasound elastography.

5.3 Scope for Further Investigation

Future work might include the use of more sophisticated algorithms for binary classifier in order to develop a model that is more accurate and efficient in terms of accuracy, computational cost, and training and testing time. For instance, convolutional neural network (CNN) [73] could be investigated to address the accuracy assessment issue. Moreover, optimization of the proposed model using the nine features or proposing new features to achieve higher classification accuracy along with computational cost might also be investigated. Collecting and using data from more patients could also increase the reliability of using such supervised methods.

References

[1] Tomy Varghese and Jonathan Ophir. A theoretical framework for performance characterization of elastography: The strain filter. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 44(1):164–172, 1997.

[2] Andrew Gee, Joel Lindop, Graham Treece, Richard Prager, and Susan Freeman. Stable, intelligible ultrasonic strain imaging. *Ultrasound*, 16(4):187–192, 2008.

[3] Jonathan Ophir, Ignacio Cespedes, Hm Ponnekanti, Y Yazdi, and Xin Li. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrasonic Imaging*, 13(2):111–134, 1991.

[4] T. Sugimoto, S. Ueha, and K. Itoh. Tissue hardness measurement using the radiation force of focused ultrasound. In *IEEE Symposium on Ultrasonics*, pages 1377–1380 vol.3, Dec 1990.

[5] Matthew O'Donnell, Andrei R Skovoroda, Benjamin M Shapo, and Stanislav Y Emelianov. Internal displacement and strain imaging using ultrasonic speckle tracking. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 41(3):314–325, 1994.

[6] M. O'Donnell, A. R. Skovoroda, B. M. Shapo, and S. Y. Emelianov. Internal displacement and strain imaging using ultrasonic speckle tracking. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 41(3):314–325, May 1994.

[7] S. K. Alam, J. Ophir, and E. E. Konofagou. An adaptive strain estimator for elastography.

IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 45(2):461–472, March 1998.

[8] A. Pesavento, C. Perrey, M. Krueger, and H. Ermert. A time-efficient and accurate strain estimation concept for ultrasonic elastography using iterative phase zero estimation. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 46(5):1057–1067, Sept 1999.

[9] Tsuyoshi Shiina, Kathryn R Nightingale, Mark L Palmeri, Timothy J Hall, Jeffrey C Bamber, Richard G Barr, Laurent Castera, Byung Ihn Choi, Yi-Hong Chou, David Cosgrove, et al. Wfumb guidelines and recommendations for clinical use of ultrasound elastography: Part 1: basic principles and terminology. *Ultrasound in Medicine & Biology*, 41(5):1126–1147, 2015.

[10] Richard G Barr, Kazutaka Nakashima, Dominique Amy, David Cosgrove, Andre Farrokh, Fritz Schafer, Jeffrey C Bamber, Laurent Castera, Byung Ihn Choi, Yi-Hong Chou, et al. Wfumb guidelines and recommendations for clinical use of ultrasound elastography: Part 2: breast. *Ultrasound in Medicine & Biology*, 41(5):1148–1160, 2015.

[11] Giovanna Ferraioli, Carlo Filice, Laurent Castera, Byung Ihn Choi, Ioan Sporea, Stephanie R Wilson, David Cosgrove, Christoph F Dietrich, Dominique Amy, Jeffrey C Bamber, et al. Wfumb guidelines and recommendations for clinical use of ultrasound elastography: Part 3: liver. *Ultrasound in Medicine & Biology*, 41(5):1161–1179, 2015.

[12] David Cosgrove, Richard Barr, Joerg Bojunga, Vito Cantisani, Maria Cristina Chammas, Manjiri Dighe, Sudhir Vinayak, Jun-Mei Xu, and Christoph F Dietrich. Wfumb guidelines and recommendations on the clinical use of ultrasound elastography: Part 4. thyroid. *Ultrasound in Medicine & Biology*, 43(1):4–26, 2017.

[13] Richard G Barr, David Cosgrove, Marko Brock, Vito Cantisani, Jean Michel Correas, Arnoud W Postema, Georg Salomon, Masakazu Tsutsumi, Hui-Xiong Xu, and Christoph F Dietrich. Wfumb guidelines and recommendations on the clinical use of ultrasound elastography: Part 5. prostate. *Ultrasound in Medicine & Biology*, 43(1):27–48, 2017.

[14] Mickael Tanter, Jeremy Bercoff, Alexandra Athanasiou, Thomas Deffieux, Jean-Luc Gennisson, Gabriel Montaldo, Marie Muller, Anne Tardivon, and Mathias Fink. Quantitative assessment of breast lesion viscoelasticity: initial clinical results using supersonic shear imaging. *Ultrasound in Medicine & Biology*, 34(9):1373–1386, 2008.

[15] Mark L Palmeri and Kathryn R Nightingale. Acoustic radiation force-based elasticity imaging methods. *Interface Focus*, page rsfs20110023, 2011.

[16] Timothy J Hall, Paul E Barboneg, Assad A Oberai, Jingfeng Jiang, Jean-Francois Dord, Sevan Goenezen, and Ted G Fisher. Recent results in nonlinear strain and modulus imaging. *Current medical imaging reviews*, 7(4):313–327, 2011.

[17] Graham Treece, Joel Lindop, Lujie Chen, James Housden, Richard Prager, and Andrew Gee. Real-time quasi-static ultrasound elastography. *Interface Focus*, 1(4):540–552, 2011.

[18] MM Doyley. Model-based elastography: a survey of approaches to the inverse elasticity problem. *Physics in Medicine and Biology*, 57(3):R35, 2012.

[19] Seyed Reza Mousavi, Ali Sadeghi-Naini, Gregory J Czarnota, and Abbas Samani. Towards clinical prostate ultrasound elastography using full inversion approach. *Medical Physics*, 41(3), 2014.

[20] Olalekan A Babaniyi, Assad A Oberai, and Paul E Barbone. Recovering vector displacement estimates in quasistatic elastography using sparse relaxation of the momentum equation. *Inverse Problems in Science and Engineering*, 25(3):326–362, 2017.

[21] Reza Zahiri-Azar and Septimiu E Salcudean. Motion estimation in ultrasound images using time domain cross correlation with prior estimates. *IEEE Transactions on Biomedical Engineering*, 53(10):1990–2000, 2006.

[22] Richard GP Lopata, Maartje M Nillesen, Hendrik HG Hansen, Inge H Gerrits, Johan M Thijssen, and Chris L De Korte. Performance evaluation of methods for two-dimensional displacement and strain estimation using ultrasound radio frequency data. *Ultrasound in medicine &*

biology, 35(5):796–812, 2009.

[23] Adib Nahiyan and Md Kamrul Hasan. Hybrid algorithm for elastography to visualize both solid and fluid-filled lesions. *Ultrasound in medicine & biology*, 41(4):1058–1078, 2015.

[24] Xiaochang Pan, Ke Liu, Jinghua Shao, Jing Gao, Lingyun Huang, Jing Bai, and Jianwen Luo. Performance comparison of rigid and affine models for motion estimation using ultrasound radio-frequency signals. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 62(11):1928–1943, 2015.

[25] Andrey Kuzmin, Aaron M Zakrzewski, Brian W Anthony, and Victor Lempitsky. Multi-frame elastography using a handheld force-controlled ultrasound probe. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 62(8):1486–1500, 2015.

[26] Roch Listz Maurice and Michel Bertrand. Lagrangian speckle model and tissue-motion estimation-theory [ultrasonography]. *IEEE transactions on medical imaging*, 18(7):593–603, 1999.

[27] Claire Pellot-Barakat, Frédérique Frouin, Michael F Insana, and Alain Herment. Ultrasound elastography based on multiscale estimations of regularized displacement fields. *IEEE transactions on medical imaging*, 23(2):153–163, 2004.

[28] Elisabeth Brusseau, Jan Kybic, Jean-François Déprez, and Olivier Basset. 2-d locally regularized tissue strain estimation from radio-frequency ultrasound images: Theoretical developments and results on experimental data. *IEEE Transactions on Medical Imaging*, 27(2):145–160, 2008.

[29] Hassan Rivaz, Emad Boctor, Pezhman Foroughi, Richard Zellars, Gabor Fichtinger, and Gregory Hager. Ultrasound elastography: a dynamic programming approach. *IEEE Transactions on Medical Imaging*, 27(10):1373–1377, 2008.

[30] Jérémy Bercoff, Mickael Tanter, and Mathias Fink. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE transactions on ultrasonics, ferroelectrics, and*

frequency control, 51(4):396–409, 2004.

[31] Pengfei Song, Heng Zhao, Armando Manduca, Matthew W Urban, James F Greenleaf, and Shigao Chen. Comb-push ultrasound shear elastography (cuse): a novel method for two-dimensional shear elasticity imaging of soft tissues. *IEEE transactions on medical imaging*, 31(9):1821–1832, 2012.

[32] Jeffrey Bamber, D Cosgrove, CF Dietrich, J Fromageau, J Bojunga, F Calliada, V Cantisani, J-M Correas, M DâĂŹonofrio, EE Drakonaki, et al. Efsumb guidelines and recommendations on the clinical use of ultrasound elastography. part 1: Basic principles and technology. *Ultraschall in der Medizin-European Journal of Ultrasound*, 34(02):169–184, 2013.

[33] Armen P Sarvazyan, Oleg V Rudenko, Scott D Swanson, J Brian Fowlkes, and Stanislav Y Emelianov. Shear wave elasticity imaging: a new ultrasonic technology of medical diagnostics. *Ultrasound in medicine & biology*, 24(9):1419–1435, 1998.

[34] Kathryn Nightingale, Stephen McAleavey, and Gregg Trahey. Shear-wave generation using acoustic radiation force: in vivo and ex vivo results. *Ultrasound in medicine & biology*, 29(12):1715–1723, 2003.

[35] W. F. Walker and G. E. Trahey. A fundamental limit on delay estimation using partially correlated speckle signals. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42(2):301–308, March 1995.

[36] Lujie Chen, Graham M Treece, Joel E Lindop, Andrew H Gee, and Richard W Prager. A quality-guided displacement tracking algorithm for ultrasonic elasticity imaging. *Medical Image Analysis*, 13(2):286–296, 2009.

[37] Yael Petrank, Lingyun Huang, and Matthew O'Donnell. Reduced peak-hopping artifacts in ultrasonic strain estimation using the viterbi algorithm. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 56(7):1359–1367, 2009.

[38] Michael H Wang, Mark L Palmeri, Veronica M Rotemberg, Ned C Rouze, and Kathryn R Nightingale. Improving the robustness of time-of-flight based shear wave speed reconstruction methods using ransac in human liver in vivo. *Ultrasound in Medicine & Biology*, 36(5):802–813, 2010.

[39] Claire Pellot-Barakat, Muriel Lefort, Linda Chami, Mickaël Labit, Frédérique Frouin, and Olivier Lucidarme. Automatic assessment of shear wave elastography quality and measurement reliability in the liver. *Ultrasound in medicine & biology*, 41(4):936–943, 2015.

[40] G Clifford Carter. *Coherence and time delay estimation: an applied tutorial for research, development, test, and evaluation engineers*. IEEE, 1993.

[41] William F Walker and Gregg E Trahey. A fundamental limit on delay estimation using partially correlated speckle signals. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42(2):301–308, 1995.

[42] A Weiss and Ehud Weinstein. Fundamental limitations in passive time delay estimation–part i: Narrow-band systems. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(2):472–486, 1983.

[43] I Cespedes, M Insana, and J Ophir. Theoretical bounds on strain estimation in elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42(5):969–972, 1995.

[44] Yeun-Chung Chang, Min-Chun Yang, Chiun-Sheng Huang, Shao-Chien Chang, Guan-Ying Huang, Woo Kyung Moon, and Ruey-Feng Chang. Automatic selection of representative slice from cine-loops of real-time sonoelastography for classifying solid breast masses. *Ultrasound in Medicine & Biology*, 37(5):709–718, 2011.

[45] Shao-Chien Chang, Yan-Wei Lee, Yi-Chen Lai, Chui-Mei Tiu, Hsin-Kai Wang, Hong-Jen Chiou, Yu-Wei Hsu, Yi-Hong Chou, and Ruey-Feng Chang. Automatic slice selection and diagnosis of breast strain elastography. *Medical Physics*, 41(10):102902, 2014.

[46] Elyas Shaswary, Yuan Xu, and Jahan Tavakkoli. Performance study of a new time-delay estimation algorithm in ultrasonic echo signals and ultrasound elastography. *Ultrasonics*, 69:11–18, 2016.

[47] Rongmin Xia, Guozhi Tao, and Arun Kumar Thittai. Dynamic frame pairing in real-time freehand elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(6):979–985, 2014.

[48] Faouzi Kallel and Jonathan Ophir. A least-squares strain estimator for elastography. *Ultrasonic Imaging*, 19(3):195–208, 1997.

[49] Graham M Treece, Joel E Lindop, Andrew H Gee, and Richard W Prager. Uniform precision ultrasound strain imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(11):2420–2436, 2009.

[50] Joel E Lindop, Graham M Treece, Andrew H Gee, and Richard W Prager. An intelligent interface for freehand strain imaging. *Ultrasound in Medicine & Biology*, 34(7):1117–1128, 2008.

[51] Hossein Khodadadi, Amir G Aghdam, and Hassan Rivaz. Edge-preserving ultrasonic strain imaging with uniform precision. In *Proc. 37th IEEE International Conference on Engineering, Medicine and Biology (EMBC)*, pages 3835–3838, 2015.

[52] Jingfeng Jiang, Timothy J Hall, and Amy M Sommer. A novel performance descriptor for ultrasonic strain imaging: a preliminary study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(6):1088–1102, 2006.

[53] Bhaskara Rao Chintada, Adhitya Vikraman Subramani, Bagyam Raghavan, and Arun Kumar Thittai. A novel elastographic frame quality indicator and its use in automatic representative-frame selection from a cine loop. *Ultrasound in Medicine & Biology*, 2016.

[54] M. Ghasemi Amidabadi, M. O. Ahmad, and H. Rivaz. Supervised classification of the accuracy of the time delay estimation in ultrasound elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(1):21–29, Jan 2018.

[55] M. Ghasemi Amidabadi, M. O. Ahmad, and H. Rivaz. Automatic accuracy assessment of ultrasound elastography using correlation profile and prior information of displacement continuity. *IEEE GlobalSIP*, November 2017.

[56] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[57] Hassan Rivaz, Emad M Boctor, Michael A Choti, and Gregory D Hager. Ultrasound elastography using multiple images. *Medical Image Analysis*, 18(2):314–329, 2014.

[58] Y Raghavender Rao, Nikhil Prathapani, and E Nagabhooshanam. Application of normalized cross correlation to image registration. *International Journal of Research in Engineering and Technology*, 3(05):12–16, 2014.

[59] Vladimir Vapnik, Steven E Golowich, Alex Smola, et al. Support vector method for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems*, pages 281–287, 1997.

[60] Hassan Rivaz, Emad M Boctor, Michael A Choti, and Gregory D Hager. Real-time regularized ultrasound elastography. *IEEE Transactions on Medical Imaging*, 30(4):928–945, 2011.

[61] Jørgen Arendt Jensen. Field: A program for simulating ultrasound systems. In *10tH Nordicbaltic conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1: 351–353*. Citeseer, 1996.

[62] Tom Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[63] H. S. Hashemi and H. Rivaz. Global time-delay estimation in ultrasound elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(10):1625–1636, Oct

2017.

[64] Christopher M Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.

[65] Yoav Freund and Robert E Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.

[66] Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine learning*, 29(2-3):131–163, 1997.

[67] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. Classification and regression trees. 1984.

[68] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[69] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of human genetics*, 7(2):179–188, 1936.

[70] Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[71] Zahra Karimaghaloo, Hassan Rivaz, Douglas L Arnold, D Louis Collins, and Tal Arbel. Adaptive voxel, texture and temporal conditional random fields for detection of gad-enhancing multiple sclerosis lesions in brain mri. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 543–550. Springer, 2013.

[72] Kirsten Christensen-Jeffries, Jemma Brown, Paul Aljabar, Mengxing Tang, Christopher Dunsby, and Robert J Eckersley. 3-d in vitro acoustic super-resolution and super-resolved velocity mapping using microbubbles. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 64(10):1478–1486, 2017.

[73] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.