

Classification of Individual Finger Flexions Using Ultrasound Radiofrequency Signals

by

Alexander Fernandes

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Biomedical Engineering

Carleton University
Ottawa, Ontario

© 2020, Alexander Fernandes

Abstract

The objective of this thesis is to develop a wearable human machine interface (HMI) to classify individual finger flexions using ultrasound transducers (UTs). To study the effect that lateral resolution has on the finger classification performance, 127 ultrasound radiofrequency (RF) signals are acquired from a 40 mm width linear array probe. The acquired signals were laterally averaged to show an estimated number of 4 – 8 reconstructed RF signals can maintain high accuracies to that of full resolution. The pattern classification pipeline of this study is verified to make accurate finger predictions on 100 ms time intervals. The proposed spatial feature extraction method novel to ultrasound-based studies is 10 times faster to execute than conventional methods. Finally, three 280 μ m thin wearable UTs are attached to the forearm achieving a classification accuracy of $97.5 \pm 2.9\%$. The results of this thesis provide the guidelines to design for wearable HMI applications.

Acknowledgements

I would first like to thank both of my thesis supervisors Dr. Yuu Ono and Dr. Eranga Ukwatta. I would like to thank Dr. Yuu Ono for his passion, enthusiasm, and immense guidance. I would like to thank him for answering all my endless questions from the moment I contacted him in September 2017, starting the fourth year of my Bachelor of Engineering degree. I would also like to thank him for his recommendation to take his graduate medical imaging course, prior to starting my Master of Applied Science degree. I would also like to thank Dr. Eranga Ukwatta for the time and efforts he spent helping me tackle my machine learning problems with his knowledge and unique perspectives. From their guidance, I have learned so much from them.

I would also like to acknowledge Carleton University, National Sciences and Engineering Research Council of Canada (NSERC), Ontario Student Assistance Program (OSAP), Dr. Walter and Mary Chudobiak, Vector Institute, and the Canadian Medicine and Biological Engineering Society (CMBES) for the opportunity and financial support they provided for my thesis research studies.

I am grateful to all my graduate student colleagues for their time in helping me with my research experiments, moral support, and encouragement throughout these years. I am grateful to all the professors, students, and friends I made at Carleton University for making the opportunity to study here worth it.

Finally, I must express my gratitude to my brother, mother, and father, for providing me with their unfailing love, support, and encouragement keeping me steady throughout this journey. This thesis would not have been possible without them.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	viii
List of Figures.....	ix
List of Abbreviations	xiv
List of Symbols	xv
Chapter 1: Introduction	1
1.1 Overview	1
1.2 Problem Statement.....	4
1.3 Objective.....	5
1.4 Thesis Contributions.....	7
1.5 Research Publications.....	9
1.6 Thesis Organization.....	10
Chapter 2: Technical Background and Literature Review.....	12
2.1 Technical Background.....	12
2.1.1 Anatomy for Finger Control.....	12
2.1.1.1 Flexor Muscles	13
2.1.1.2 Cross Section of the Forearm	15
2.1.2 Ultrasound Data and Image Acquisition	17
2.1.2.1 A-Mode	19
2.1.2.2 B-Mode.....	21
2.1.2.3 M-Mode.....	23
2.2 Literature Review on Human Machine Interfaces and Prosthetic Applications.....	25

2.2.1	Neuromuscular Sensor Strategies.....	25
2.2.1.1	Brain Computer Interfaces	25
2.2.1.2	Peripheral Nerve Electrodes	27
2.2.1.3	Surface Electromyography	27
2.2.1.4	Commercial Prosthetics.....	28
2.2.2	Current Research Using Ultrasound for Human Machine Interfaces.....	29
2.2.2.1	Clinical Ultrasound Probe Studies.....	29
2.2.2.2	Multiple Single-Element Ultrasound Transducer Studies	31

Chapter 3: Experimental Emulation Method to Study the Effect of Lateral Spatial Resolution on Individual Finger Flexion Classification 33

3.1	Introduction	33
3.2	Data Acquisition of a Clinical Ultrasound Probe	33
3.2.1	Probe Placement on Forearm	34
3.2.2	Ultrasound Data Format.....	35
3.2.3	Lateral Resolution Reconstruction Procedure	38
3.2.4	Individual Finger Flexion Recording Procedure	40
3.3	Pattern Classification Pipeline.....	42
3.3.1	Spatial Feature Extraction	44
3.3.1.1	The DWT-MAV Method.....	45
3.3.1.2	The ENV-LR Method.....	50
3.3.2	Temporal Feature Extraction.....	52
3.3.3	Procedure for Continuous Classification.....	52
3.3.4	Feature Selection.....	53
3.3.5	Performance Evaluation	54

Chapter 4: Results and Discussions of the Effect of Reduced Lateral Resolution for Individual Finger Flexion Classification..... 55

4.1	Introduction	55
4.2	Preliminary Study on Reducing Lateral Spatial Resolution	55
4.2.1	Correlation Between Classification Accuracy and Lateral Resolution	56
4.2.2	Classification Accuracy and F1 Score at a Reduced Lateral Resolution	57
4.2.3	Discussion	59
4.3	Detailed Study on Lateral Resolution and Individual Finger Flexion Classification ...	60
4.3.1	Linear Discriminant Analysis Projections of the Extracted Features.....	60
4.3.2	Individual Finger Flexion F1 Scores at Reduced Resolutions	63
4.3.3	Classification Accuracy vs Resolution.....	67
4.3.4	Computation Time vs Resolution.....	69
4.3.5	Classification Accuracy vs Computation Time.....	71
4.3.6	Discussion	73

Chapter 5: Classification of Individual Finger Flexions Using Wearable Ultrasonic Sensors **75**

5.1	Introduction	75
5.2	Materials and Methods	75
5.2.1	Wearable Ultrasonic Sensor System and Experimental Setup.....	76
5.2.1.1	Sensor Design and Experimental Configuration	76
5.2.1.2	Ultrasound Data Acquisition Setup.....	78
5.2.2	Data Acquisition Procedure	80
5.2.2.1	Data Format of Wearable Ultrasonic Sensor	80
5.2.2.2	Finger Flexion Recording Procedure.....	82
5.2.3	Modifications Made to Pattern Classification Pipeline.....	85
5.2.3.1	Frequency Range for Discrete Wavelet Transform.....	85
5.2.3.2	Machine Learning Classifiers.....	87
5.2.3.3	Procedure for Continuous Classification.....	88

5.3	Results	89
5.4	Discussion.....	99
Chapter 6: Conclusions and Future Works		101
6.1	Conclusions	101
6.2	Limitations.....	105
6.3	Future Works.....	106
Appendices.....		111
Appendix A Signal Processing and Machine Learning References		111
A.1	Review on the Wavelet Transform.....	111
A.2	Linear Discriminant Analysis.....	114
Appendix B Acquisition for Four Ultrasound Transducers.....		115
B.1	Hardware Wiring Management up to Four Channels	115
B.2	Python Code: readATB.py	116
References		122

List of Tables

Table 2-1 Action performed by the flexor muscles [16].....	15
Table 4-1 Precision and recall metrics for $N = 4$ reconstructed RF signals.....	58
Table 4-2 F1 scores for classifying each finger flexion using 10 independent subject arm experiments with comparison of DWT-MAV and ENV-LR methods.	64
Table 4-3 Average computation time to execute a single prediction using different spatial feature extraction methods.....	70
Table 5-1 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 1$ and $\sigma = 100$ when predicting for 100 ms sequential time intervals.	91
Table 5-2 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 2$ and $\sigma = 50$ when predicting for 100 ms sequential time intervals.	92
Table 5-3 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 4$ and $\sigma = 25$ when predicting for 100 ms sequential time intervals.	93
Table 5-4 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 5$ and $\sigma = 20$ when predicting for 100 ms sequential time intervals.	94
Table 5-5 CI Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 10$ and $\sigma = 10$ when predicting for 100 ms sequential time intervals.....	95
Table 5-6 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 25$ and $\sigma = 4$ when predicting for 100 ms sequential time intervals.	96
Table 5-7 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 50$ and $\sigma = 2$ when predicting for 100 ms sequential time intervals.	97
Table 5-8 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 100$ and $\sigma = 1$ when predicting for 100 ms sequential time intervals.	98

List of Figures

Figure 2-1 a) Superficial and b) deep muscles from the anterior side of the left forearm. Image is reproduced from “Anatomy of the Human Body” by Henry Gray and the image is licensed under the Creative Commons Attribution 4.0 International license. Image is free access online at <https://www.bartleby.com/107/125.html> [15]. 14

Figure 2-2 Transverse cross section of the forearm with the anterior side on top. Image is reproduced from “Anatomy of the Human Body” by Henry Gray and the image is licensed under the Creative Commons Attribution 4.0 International license. Image is free access online at <https://www.bartleby.com/107/illus421.html> [17]. 16

Figure 2-3 Ultrasound imaging system. Image is reproduced from “Medical Ultrasound Scanner” by D. W. Rickey and the image is licensed under Creative Commons Attribution-Share Alike 2.5 Generic license. Image is free access online at <https://en.wikipedia.org/wiki/File:AlokaPhoto2006a.jpg> [18]. 18

Figure 2-4 A hand-held ultrasound probe. Image is reproduced from “Medical Ultrasound linear array Probe/scan head/transducer” by D. W. Rickey and the image is licensed under Creative Commons Attribution-Share Alike 2.5 Generic license. Image is free access online at <https://commons.wikimedia.org/wiki/File:UltrasoundProbe2006a.jpg> [19]. 18

Figure 2-5 a) Depiction of a single element transducer with the ultrasonic pulse propagating along the scan line with echoes reflecting at tissue/muscle/bone boundaries. b) An ultrasonic RF signal obtained with a pulse echo of the forearm..... 20

Figure 2-6 a) Linear array of ultrasonic transducers acquiring multiple scan lines of RF signals. b) Envelope of an RF signal. c) B-mode image of tissue in the forearm..... 22

Figure 2-7 a) B-mode image with the 120th scan line indicated in red for b) M-mode image of the 120th scan line during a pinky finger flexion recording..... 24

Figure 2-8 Example setup of an EEG recording. Image is reproduced from “EEG Cap” by D. Myers at Dalhousie Life Sciences Center and the image is licensed under the public domain by its author, Thuglas at English Wikipedia. Image is free access online at https://commons.wikimedia.org/wiki/File:EEG_cap.jpg [31]..... 26

Figure 3-1 Placement of the ultrasound probe on the anterior side of the forearm, 5 cm away from the wrist. 34

Figure 3-2 a) Experimental setup to acquire ultrasound RF signals during an individual finger flexion using a linear array ultrasonic probe. b) corresponding B-mode ultrasound image of the forearm. The field of view is 40-mm in width and 35-mm in depth. 36

Figure 3-3 Schematic representation of ultrasound data acquisition coordinate and structure: (a) ultrasound probe orientation, (b) data structure of the acquired 127 A-mode RF signal, $r(i, j, k)$, in B-mode frame, and (c) N channels of reconstructed RF signals, $u(i, j, k)$, at the reduced lateral spatial resolution. 37

Figure 3-4 Block diagram of the pattern classification pipeline used to classify individual finger flexions from ultrasound RF signal data. 43

Figure 3-5 Discrete wavelet transform to compute the detail coefficients of ultrasound RF signal data using Mallat’s algorithm [61]. 46

Figure 3-6 a) ultrasound RF signal, b) frequency bandwidths of the DWT filters at the levels from 1 to 6 using the Daubechies 3 mother wavelet (dotted lines) and the frequency spectrum (solid line) of the RF signal, and c) wavelet detail coefficients for the levels 1, 2, and 3 computed with the RF signal. 49

Figure 3-7 Procedure of ENV-LR method for spatial feature extraction adapted from [10]: a) RF signal, b) envelope using the Hilbert transform, and c) linear fit (solid line) using linear regression of the envelope (circles) in the given segment region denoted by dashed lines in b).	51
Figure 4-1 Classification accuracies verses number of reconstructed RF signals for subject 1 left arm.....	56
Figure 4-2 Confusion matrix for $N = 4$ reconstructed rf signals of subject 1 left arm.....	58
Figure 4-3 LDA projections of feature extracted from subject 3 right arm for lateral resolutions $N =$ a) 1, b) 12, and c) 127.....	62
Figure 4-4 Comparison between the spatial feature extraction methods DWT-MAV and ENV-LR for F1 Scores of a) thumb, b) index, c) middle, d) ring, and e) pinky finger flexions.....	67
Figure 4-5 Classification accuracy for varying number of reconstructed RF signals: a) DWT-MAV method, and b) ENV-LR method with 10 segments.....	68
Figure 4-6 Comparison between the spatial feature extraction methods DWT-MAV and ENV-LR for a) resolution verses classification accuracy, b) resolution verses computation time, and c) computation time verses classification accuracy.	72
Figure 5-1 One Wearable Ultrasonic Sensor I constructed for this thesis. a) front side with active area, b) backside with lined paper backing, and c) full transducer soldered to a coaxial cable.....	77
Figure 5-2 Three WUSs placed on my forearm 5 cm away from my wrist using ultrasound couplant gel. The WUSs were secured by a 5 cm width self-adhering elastic bandage.	77

Figure 5-3 An overview of the designed WUS system. Ultrasound acquisition setup for three WUSs.	79
Figure 5-4 A single pulse echo acquired RF signal with a region of interest ranging between sample points 320 to 1520.	81
Figure 5-5 M-mode images of each WUS during acquisition of five repeated motions in a 6 second trial recording of a a) thumb, b) index, c) middle, d) ring, and e) pinky finger flexion procedure.	84
Figure 5-6 a) DWT detail coefficients computed for the first 6 levels and b) frequency spectrum (solid line) of the RF signal in the ROI and the bandwidth frequency spectrums for the first 6 levels of the DWT (dashed lines).	86
Figure 5-7 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 1$ and $\sigma = 100$ when predicting for 100 ms sequential time intervals.	91
Figure 5-8 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 2$ and $\sigma = 50$ when predicting for 100 ms sequential time intervals.	92
Figure 5-9 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 4$ and $\sigma = 25$ when predicting for 100 ms sequential time intervals.	93
Figure 5-10 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 5$ and $\sigma = 20$ when predicting for 100 ms sequential time intervals.	94
Figure 5-11 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 10$ and $\sigma = 10$ when predicting for 100 ms sequential time intervals.	95
Figure 5-12 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 25$ and $\sigma = 4$ when predicting for 100 ms sequential time intervals.	96

Figure 5-13 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 50$ and $\sigma = 2$ when predicting for 100 ms sequential time intervals.	97
Figure 5-14 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 100$ and $\sigma = 1$ when predicting for 100 ms sequential time intervals.	98
Figure A-1 Visualization of the Short Time Fourier Transform (STFT) and Wavelet Transform (WT). a) Frequency modulations of a Gaussian window used in the STFT. a) Time-frequency tiling of the STFT. c) Shifting and scaling of a mother wavelet used in the WT. d) Time-frequency tiling of the WT. Image is reproduced from “Wavelets and Subband Coding” in Figure 2.12, by Martin Vetterli and Jelena Kovačević, and the image is licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License [69].	112
Figure A-2 Mallat’s filter bank implementation of the discrete wavelet transform [61].	113
Figure B-1 Full hardware wiring diagram setup for using up to four UTs, pulsar/receiver, function generator, oscilloscope, and analog/digital converter (waveform digitizer). ...	115
Figure B-2 readATB.py Python code.	121

List of Abbreviations

Abbreviation	Description
BCI	brain computer interface
DWT	discrete wavelet transform
DWT-MAV	discrete wavelet transform – mean absolute value
ENV-LR	envelope – linear regression
HMI	human machine interface
MAV	mean absolute value
PCP	pattern classification pipeline
PVDF	polyvinylidene difluoride
RF	radiofrequency
sEMG	surface electromyography
UT	ultrasound transducer
WUS	wearable ultrasonic sensor

List of Symbols

<i>Symbol</i>	Description
α	Threshold of family-wise error rate
A_l	Analysis coefficient at level l
D	Ultrasound transducer width
D_f	Fresnel distance
D_l	Detail coefficient at level l
e	Envelope of RF signal
l	Level of the recursive discrete wavelet transform procedure
λ	Ultrasound wavelength in soft tissue
M	Length of the spatial feature vector S
N	Number of reconstructed RF signals
<i>Prob</i>	Probability function
r	Acquired ultrasound RF signal
σ	Number of probabilities averaged into a single probability
S	Spatial feature vector
τ	Number of frames averaged into a single frame
T	Temporal feature vector
u	Reconstructed ultrasound RF signal
ω	Finger (one of thumb, index, middle, ring, pinky)

Chapter 1: Introduction

1.1 Overview

A human machine interface (HMI) is a controller that connects the user to operate a virtual system or robotic machine. Wearable HMIs are typically used in applications requiring repetitive tasks. For example, a wearable HMI is used to control upper limb robotic prosthetics, aiding amputees to regain hand functionality. Most commercial upper limb prosthetics use surface electromyography (sEMG) electrodes as the non-invasive wearable HMI sensor strategy. This sensor strategy uses electrodes to detect the electrical activity of the muscles from the surface of the forearm. However, the challenges with using sEMG includes signal attenuation based on depth and a low signal to noise ratio [1]. These challenges with sEMG make it difficult to discriminate between the different muscle groups within the forearm. For this reason, the task to distinguish between all five finger movements makes it difficult to achieve individual finger control for commercial sEMG prosthetics.

For the studies conducted in this thesis, individual finger flexion control can be referred to as the ability to recognize which of the five fingers is undergoing the flexion motion every 100 ms in time. This means that every 100 ms in time, a prediction is made between one of the five fingers depending on the motion occurring at that time. For example, if it takes approximately 1 second to perform one complete finger flexion, then ten predictions would be made across the full duration of the finger flexion. If a 90% classification accuracy is reported under these conditions, then that means over the 1 second period, nine out of ten predictions were correct. This “optimal” delay time of 100–125 ms, was experimentally investigated as an appropriate time to improve the

classification accuracy without influencing the performance of a prosthetic controller [2]. Even though it is usually better to have the fastest computation speed available, it is generally accepted that prediction speeds less than 300 ms are satisfactory [3]. Executing predictions on every 100 ms time intervals will give the effect of having smooth and continuous classification in HMI applications. Currently, most ultrasound-based studies referred to in this thesis, conducted experiments on discrete hand positions or complete finger flexions. The goal of these referred studies often make a prediction when the hand is in specific position or to classify a finger flexion over the 0.5–2 second duration it takes to be completed (rather than making predictions in time intervals less than 300 ms).

Prosthetic controllers can use either invasive or noninvasive HMI strategies to control the robotic device. Researchers used invasive methods to achieve individual finger flexion classification by spatially resolving the different muscle contractions within the forearm. Davis *et al.* recruited two amputees to surgically implant an array of 96 microelectrodes to the median and ulnar nerves to manipulate a virtual prosthetic limb with individual finger flexion control [4]. However, the problem with using an invasive method includes damaging of the nerve due to a foreign body reaction and inflammation after prolong use [5, 6].

Even though it is possible for the noninvasive sEMG strategy to classify individual finger flexion, it is more challenging with the high number of electrodes required. Celadon *et al.* found that reducing the spatial resolution of a high-density, 8×24 cm, 192-channel electrode grid (8 by 24 electrode matrix) down to 24 channels (3 by 8 electrode matrix) could maintain a high classification accuracy when discriminating between four finger flexions (index, middle, ring, pinky) [7]. However most commercial

upper limb prosthetics do not implement full finger flexion control due to only using one to three sEMG electrodes [8].

The other noninvasive sensor strategy, ultrasound imaging, has proven to outperform sEMG methods when performing individual finger flexion classification. Sikdar *et al.* stated that “it can avoid many of the limitations of EMG because it can spatially resolve individual muscles deep inside tissue and visualize dynamic activity of different functional compartments” [9]. Rather than using the electrical activity from the muscles with sEMG, ultrasound imaging can be used to dynamically observe the deformation patterns of the tissue within the forearm.

Huang *et al.* compared sEMG with B-mode ultrasound imaging to discriminate between 14 discrete hand positions and observed that using ultrasound imaging achieved a classification accuracy of 96% in comparison to 90% with sEMG [10]. Then in 2017, researchers at Georgia Institute of Technology designed a prosthetic arm using B-mode ultrasound called the “Luke Skywalker’s bionic hand” [11]. This prosthetic arm gave the ability to have control over individual finger flexions, “allowing the amputee to play the piano for the first time since the accident” [12].

It became apparent that attaching a large, bulky B-mode ultrasound imaging probe to an amputee’s forearm was cumbersome. In 2018, Yang *et al.* used four A-mode ultrasonic transducers to achieve 95% accuracy in classifying 11 different hand positions in real-time [13]. Then in the following year Yan *et al.* of the same research lab constructed four lightweight wearable and flexible ultrasound sensors made of polyvinylidene difluoride (PVDF) piezoelectric film to achieve a 98% classification accuracy on the same 11 discrete hand positions [14]. The experiments they studied were

on the detection of discrete hand positions in real time, which is useful for American Sign Language applications. However, to maintain precise control over prosthetic devices and other HMI controlled systems, the ability to continuously classify between individual finger flexions needed to be experimentally tested.

1.2 Problem Statement

Ultrasound in comparison to sEMG methods, has proven to have better accuracy when discriminating between individual finger flexions and discrete hand gestures. The main advantage of ultrasound is it can noninvasively provide high spatial precision when observing the muscle movements occurring within the forearm. Multiple single-element ultrasound transducers (UTs) have been proven to achieve a high classification accuracy for discriminating between discrete hand gestures. However, the effect that lateral resolution has on classification accuracy has not been studied. More specifically, the lateral sampling interval between each UTs and number of UTs within a confined width has yet to be investigated. The optimal controller delay for prosthetic application is 100–125 ms [2]. However, an upper bound is generally accepted as having predictions be less than 300 ms [3]. Given the optimal controller delay criteria, I form the overarching thesis question:

Can multiple single-element UTs (low lateral resolution method) achieve a similar performance to a B-mode ultrasound imaging probe (high lateral resolution method), when classifying for individual finger flexions every 100 ms in time?

The two criteria for investigating the effect that lateral spatial resolution has on classifying individual finger flexions includes:

- 1) Determine an approximate minimum lateral spatial resolution that can maintain a high classification accuracy.
- 2) At low lateral spatial resolutions, predictions should be fast to execute on 100 ms consecutive time intervals to replicate dynamic finger flexion prediction. With the goal to have the total computation time for data collection and prediction to be between 100 ms and 125 ms.

The results of this study will provide guidance for designing a multiple single-element wearable UT system for wearable HMI applications such as upper limb prosthetics.

1.3 Objective

The overarching objective of this thesis research is to use multiple single-element wearable ultrasonic sensors (WUS) to continuously classify individual finger flexions for use in HMI applications. To achieve this overarching objective, this thesis is divided into the following objectives.

The first objective is to investigate classification accuracy as a function of lateral spatial resolution. By using a B-mode ultrasound imaging probe to acquire 127 lateral ultrasound radiofrequency (RF) signals, a reduction of the lateral resolution is used to emulate data acquired from multiple single-element UTs. This objective is conducted in a preliminary study by acquiring data from a single subject arm. This preliminary study is used to provide the framework to reduce lateral resolutions of ultrasound data conducted in this thesis.

The second objective is to verify the performance of the proposed pattern classification pipeline (PCP) when classifying individual finger flexions. The PCP is verified to operate for the low lateral resolution ultrasound data. The PCP takes ultrasound RF signal recordings of individual finger flexions and performs finger classification every 100 ms in time on the recordings. The PCP consists of the following steps: preprocessing, spatial feature extraction, temporal feature extraction, 100 ms time grouping of labeled data, feature selection (automatic reduction of features), and classification. Five subjects (10 independent arm experiments) were recruited to verify the results of the preliminary study conducted in the first objective. The classification accuracy and F1 scores are used to study which reduced lateral resolutions can achieve similar accuracies to that of the full 127 lateral resolution. Each experiment only used training and test data from the same subject arm. Three-fold cross validation was used to ensure training data was not used as test data. Three trials were recorded for each finger type. Each fold in the cross validation used test data coming from the same trial number for finger flexion. This was to ensure that predictions would occur continuously on the same finger flexion without using training data as test data.

The third objective is to make the PCP execute finger predictions faster than the conventional method. When reducing the lateral resolution, the two criterion requirements are: 1) the performance in discriminating between each finger maintains a similar accuracy to using the full resolution and 2) the data collection and computation time required to make a prediction speed be within 100–125 ms. This time criteria to make predictions on 100–125 ms time intervals is based on the optimal controller delay for prosthetic devices [2]. However, as long as prediction speeds are less than 300 ms it is

generally accepted as satisfactory [3]. A spatial feature extraction technique novel to ultrasound-based HMI strategies is proposed and compared to a method commonly used in previous studies. The novel feature extraction method uses the discrete wavelet transform (DWT) to preprocess the ultrasound RF signals. In comparison, the commonly used methods perform preprocessing by obtaining the signal envelope. The DWT is used to preprocess the ultrasound RF signal into multiple time-frequencies in comparison to only using the amplitude from the envelope.

The final objective is to construct a multiple single-element WUS system to continuously classify individual finger flexions. For this study, the proposed PCP is used on the ultrasound RF signals acquired from the WUS system. The first, second, and third objectives optimize the performance of the PCP using emulated low lateral resolution ultrasound data. This fourth objective verifies the performance of the PCP using low lateral resolution ultrasound data acquired from a multiple single-element WUS system.

The completion of these objectives can be used to guide the design of an ergonomic sensor strategy for wearable HMI applications.

1.4 Thesis Contributions

The following are the main contributions of this thesis.

- Using a B-mode linear array ultrasound probe, I investigated classification accuracy as a function of lateral resolution. I am the first to study the effect that lateral resolution has on individual finger flexion classification for ultrasound-based studies. An averaging technique is proposed to reconstruct the lateral resolution of the acquired ultrasound data from a B-mode linear array ultrasound

probe. The purpose is to emulate acquiring ultrasound data from multiple single-element UTs. Using data acquired from 10 subject arms, an estimated number of 4 – 8 single-element UTs can maintain a similar classification accuracy to using the full-resolution of a bulky B-mode ultrasound imaging probe.

- I developed a PCP to continuously classify individual finger flexions every 100 ms using reduced lateral resolution ultrasound data. The designed PCP includes preprocessing of the ultrasound data, spatial feature extraction, temporal feature extraction, 100 ms time grouping, classification, and performance validation. Most research studies use the pixel values of B-mode images or the signal envelope of the ultrasound RF signals. Other ultrasound-based studies only perform classification of a complete finger flexion (make a prediction after a finger flexion was completed) or discrete hand gestures (make a prediction once the hand is in a position such as a fist). The PCP takes raw ultrasound RF signal data and performs a classification every 100 ms in time throughout the full recording duration. This PCP can be used in future studies to classify for more than the proposed five finger flexions by adding different label recording data. For example, a simultaneous finger flexion could be added, resulting in a sixth motion to classify (including the other five finger flexions).
- I developed a faster method that preprocesses and extracts spatial features from the ultrasound RF signals. This method is novel to ultrasound-based studies for individual finger flexion classification. This method uses the DWT to obtain relevant spatial information from multiple time-frequency regions in the ultrasound RF signal data. This proposed method is compared to a previous

spatial feature extraction method. The proposed method was at least 10 times faster to execute than the previous method for emulated B-mode ultrasound data at lower lateral resolutions.

- A multiple single-element WUS system is designed to classify individual finger flexions. The PCP is used with the WUS system to continuously classify individual finger flexions. The WUSs are constructed using a flexible PVPF piezoelectric polymer film. After including protection and insulation layers, the total thickness of each WUS is 280 μ m. A study is conducted by attaching three WUSs to the wrist of a single subject's arm. Ultrasound RF signals are recorded during the individual finger flexions. A classification accuracy of $97.53 \pm 2.89\%$ (using 10-fold cross validation) is achieved when continuously predicting individual finger flexions on consecutive 100 ms time intervals. This preliminary study provides the necessary guidelines for the design of an ergonomic sensor strategy for wearable HMI applications.

1.5 Research Publications

Conference proceedings

- A. Fernandes, Y. Ono, and E. Ukwatta, "Effect of lateral resolution on classifying individual finger flexions," in *Canadian Medicine and Biological Engineering Conference (CMBEC)*, Ottawa, Ontario, Canada, May 2019. (Based on the results presented in Chapter 4). Awarded first place student paper presentation prize.
- A. Fernandes, Y. Ono, and U. Ukwatta, "Flexible and wearable ultrasonic sensors and method for classifying individual finger flexions," in *IEEE International*

Conference on Flexible and Printable Sensors and Systems (FLEPS), Manchester, UK, August 2020. (Based on the results presented in Chapter 5).

Manuscript for Journal

- A. Fernandes, Y. Ono, and E. Ukwatta, “Evaluation of finger flexion classification at reduced lateral spatial resolutions of ultrasound,” *under preparation*. (Based on the methods presented in Chapter 3 and results presented in Chapter 4).

1.6 Thesis Organization

This thesis research includes: a technical background and detailed literature review; an experimental procedure and method to emulate the lateral resolution reduction for ultrasound RF signals captured by a clinical ultrasound imaging probe; discussion on the effect that lateral spatial resolution has on the classification of individual finger flexions; design of a WUS system for classification of individual finger flexions.

The following descriptions outline the contents discussed in each chapter.

Chapter 1 introduces and discusses the objectives presented in this thesis research.

Chapter 2 provides a technical background overview on the anatomy of the forearm and the methods of ultrasound data acquisition. In addition, a detailed literature review on the neuromuscular sensor strategies, overview of commercial prosthetics, and ultrasound sensor strategies for HMI applications.

Chapter 3 outlines the methodology to investigate the effect of lateral spatial resolution of ultrasound data on classification accuracy of individual finger flexions. The presented

PCP is used throughout this thesis study for the classification of individual finger flexions using ultrasound RF signals. The methods included are data acquisition of a clinical ultrasound probe, lateral resolution reconstruction procedure, and the proposed PCP for classifying 100 ms consecutive time intervals of finger flexions.

Chapter 4 presents the results and discussions of the effect that lateral spatial resolution of ultrasound data has on individual finger flexion classification. First, the preliminary study conducted with a single subject arm is presented, and then the detailed study results with 10 subject arms are discussed. Finally, the performance evaluation of the methods proposed in Chapter 3 are evaluated using classification metrics and computation time.

Chapter 5 presents the preliminary study for using a multiple single-element WUS system constructed using PVDF piezoelectric film. The study validates the methods proposed in Chapter 3 for finger flexion classification using flexible WUSs proposed for future HMI applications. The preliminary study discusses the overview of the sensor design, hardware acquisition setup, and modifications to optimize the PCP using the WUS system. It includes the results and discussions on the performance of continually classifying 100 ms time intervals of individual finger flexions on a single subject arm.

Chapter 6 concludes this thesis by summarizing the results obtained and provides the recommendation for future work.

Chapter 2: Technical Background and Literature Review

In this thesis work, ultrasound is used to monitor the tissue motions associated with muscle contractions within the forearm caused by finger flexion activity. This chapter is split into two sections: section 2.1 provides the necessary overview of the concepts used in this thesis research and section 2.2 provides a detailed literature review of previous research studies related to HMI applications.

Section 2.1 provides an overview on the anatomy of skeletal muscles used to control the fingers within the forearm. Then, the methods for ultrasound acquisition are presented.

Section 2.2 provides an overview on the neuromuscular sensor strategies, which has been the most used strategy for HMI applications such as prosthetic devices. Then, a literature review is presented on research studies involving ultrasound-based sensor strategies.

2.1 Technical Background

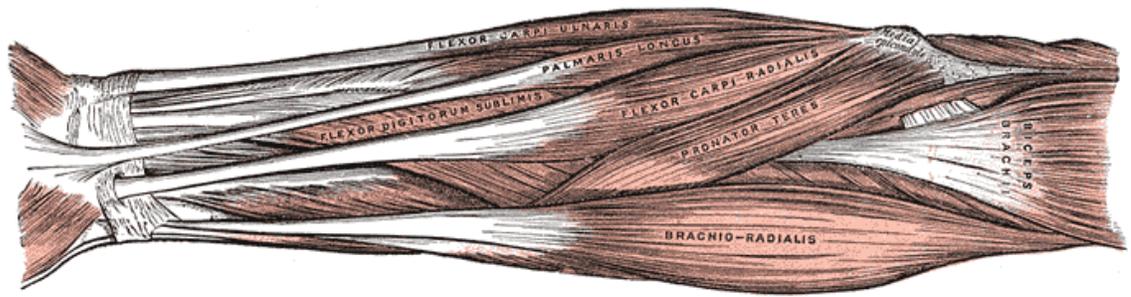
2.1.1 Anatomy for Finger Control

The hand is a naturally engineered appendage we use to interact with things daily. The functionality from the fingers includes strength and fine control provided by the muscles within the hand, wrist, and forearm regions. This gives the ability to have a powerful grasp to support the weight of the entire body. While at the same time, manipulation of small objects can be achieved with precise finger control. Muscles that contribute to finger movement comes from the muscles within the hand and forearm. The flexor muscle is used to bend the finger digits towards the palm of the hand. One end of a flexor

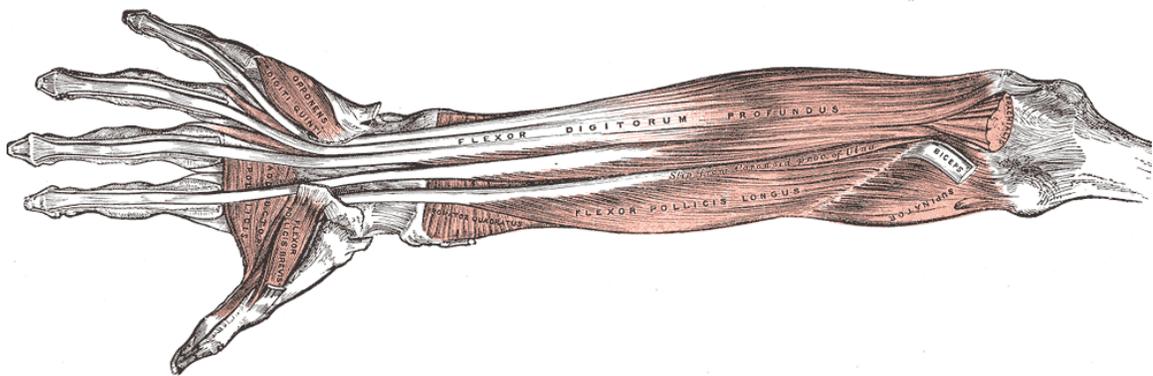
muscle is attached from the forearm to one of the five finger bones. Therefore, monitoring of the motion of flexor muscles from the forearm can be used to identify which finger is being flexed.

2.1.1.1 Flexor Muscles

The flexor muscles are located on the anterior side of the forearm. Their general function is to bend the fingers and wrist. These muscles can be divided into two depth compartments: superficial and deep regions.



(a)



(b)

Figure 2-1 a) Superficial and b) deep muscles from the anterior side of the left forearm. Image is reproduced from “Anatomy of the Human Body” by Henry Gray and the image is licensed under the Creative Commons Attribution 4.0 International license. Image is free access online at <https://www.bartleby.com/107/125.html> [15].

The flexor muscles in the superficial compartment includes the flexor carpi radialis, flexor carpi ulnaris, flexor digitorum sublimis, flexor digitorum profundus, and flexor pollicis longus as shown in Figure 2-1. Their actions are described in Table 2-1.

Table 2-1 Action performed by the flexor muscles [16].

Compartment	Flexor	Action
Superficial	Flexor carpi radialis	Flex and abduct wrist. Assists in pronation of the hand.
	Flexor carpi ulnaris	Flex and abduct the wrist.
	Flexor digitorum sublimis	Flexes the index, middle, ring, pinky fingers, and wrist. Attached to the middle bones of the fingers.
Deep	Flexor digitorum profundus	Flexes the index, middle, ring, and pinky fingers. Attached to all three bones of the fingers.
	Flexor pollicis longus	Flexes from the tip of the thumb.

2.1.1.2 Cross Section of the Forearm

The muscles in the forearm work together to control the movement of the hand and fingers. In general, muscles in the radial side maneuver the thumb. Muscles from the ulnar side maneuver the pinky. Muscles in between maneuver the index, middle, ring fingers.

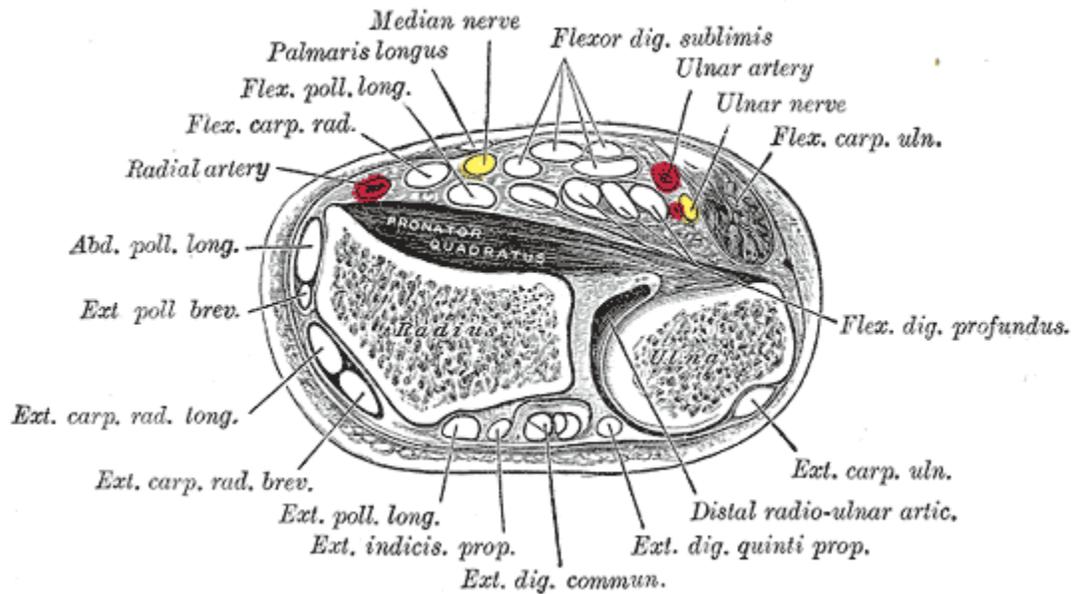


Figure 2-2 Transverse cross section of the forearm with the anterior side on top. Image is reproduced from “Anatomy of the Human Body” by Henry Gray and the image is licensed under the Creative Commons Attribution 4.0 International license. Image is free access online at <https://www.bartleby.com/107/illus421.html> [17].

Figure 2-2 shows the transverse cross section of the distal ends of the ulna (right bone region) and radius (left bone region), which is slightly before the wrist region on the forearm. The experiments conducted in this thesis research uses ultrasound to observe the forearm from the anterior side (the palm side of the hand). The anterior side of the forearm includes the superficial and deep flexor muscles. The superficial flexor muscles include the flexor carpi radialis, palmaris longus, flexor digitorum sublimis, and flexor carpi ulnaris. The deep flexor muscles include the flexor digitorum profundus, flexor pollicis longus, and pronator quadratus. The non-flexor muscles present in the anterior region includes the radial artery, ulnar artery, median nerve, and ulnar nerve. The muscles not observed in ultrasound experiments conducted in this thesis include:

abductor pollicis longus, extensor pollicis brevis, extensor carpi radialis longus, extensor carpi radialis brevis, extensor pollicis longus, extensor indicis proprius, extensor digitorum communis, extensor digiti quinti proprius, and extensor carpi ulnaris. These muscles generally contribute to the extension of the fingers.

2.1.2 Ultrasound Data and Image Acquisition

Ultrasound is a mechanical pressure wave generated by a transducer at frequencies higher than human auditory perception (frequencies above 20 kHz). Ultrasound transducers (UTs) emit pressure waves into human tissue at approximately 1540 m/s and the spatial information encoded in the echoes are used to generate a depth image. Typical ultrasound imaging systems, as shown in Figure 2-3, are an all in one system used to generate/receive the ultrasound pressure waves, store acquired ultrasound recordings, and render the ultrasound data into images on a screen. A clinical ultrasound probe shown in Figure 2-4 is designed to be hand-held by the operator. For this reason, clinical ultrasound probes are bulky and not ergonomic for HMI applications. Ultrasound imaging is used to noninvasively observe the inside of the human body by having the probe contact a skin surface coated with acoustic gel.



Figure 2-3 Ultrasound imaging system. Image is reproduced from “Medical Ultrasound Scanner” by D. W. Rickey and the image is licensed under Creative Commons Attribution-Share Alike 2.5 Generic license. Image is free access online at <https://en.wikipedia.org/wiki/File:AlokaPhoto2006a.jpg> [18].



Figure 2-4 A hand-held ultrasound probe. Image is reproduced from “Medical Ultrasound linear array Probe/scan head/transducer” by D. W. Rickey and the image is licensed under Creative Commons Attribution-Share Alike 2.5 Generic license. Image is free access online at <https://commons.wikimedia.org/wiki/File:UltrasoundProbe2006a.jpg> [19].

2.1.2.1 A-Mode

An A-mode ultrasound signal is generated from an ultrasonic transducer (UT). A single UT is used to convert a high frequency electrical signal into ultrasonic pressure waves. The frequency range is selected to be 1 MHz to 10 MHz for medical applications involving human tissue [20]. This range has been well established and experimentally tested by the medical community where low kHz frequencies will produce poor resolution while high frequencies above 10 MHz will have be susceptible to signal attenuation. With a single element UT, acquiring an ultrasonic pressure wave at these radiofrequency (RF) ranges produces an ultrasound RF signal. An RF signal contains spatial information along a depth direction (direction following into the body). Figure 2-5 shows an ultrasound RF signal acquired along the depth direction of the forearm. A single ultrasonic pulse is generated by a UT at the surface of the skin. Then the echoes that occur within the forearm propagate back towards the UT along the scan line. The sharp peaks in amplitude of the sinusoidal signal indicate a change in the tissue/muscle/bone boundary occurring at different depths.

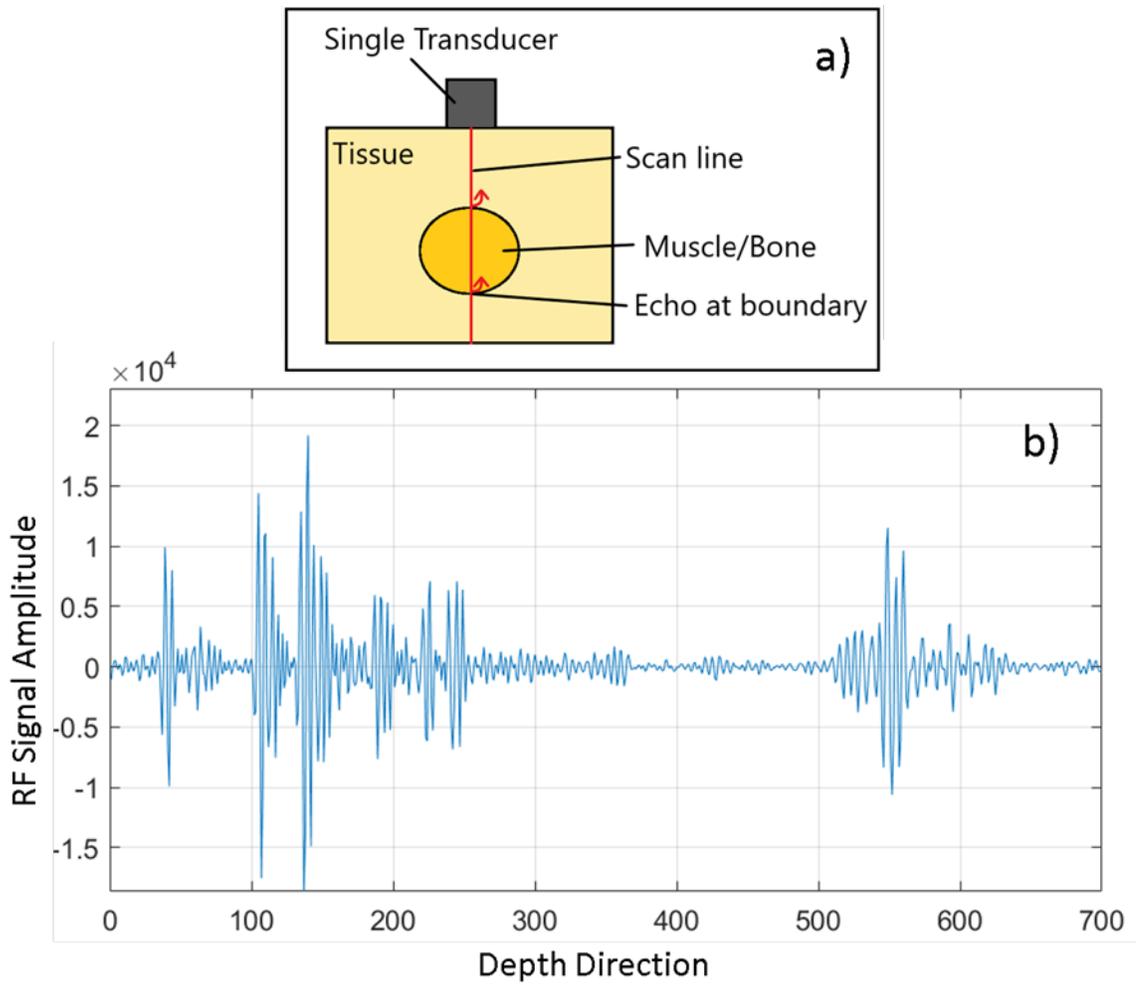


Figure 2-5 a) Depiction of a single element transducer with the ultrasonic pulse propagating along the scan line with echoes reflecting at tissue/muscle/bone boundaries. b) An ultrasonic RF signal obtained with a pulse echo of the forearm.

2.1.2.2 B-Mode

A B-Mode ultrasound image is created from multiple scan lines of A-mode ultrasound RF signals [20]. A 2-D image can be constructed using the envelope of the RF signals. In Figure 2-6 a), a linear array of ultrasonic transducers (UTs) are used to record multiple scan lines of A-mode RF signals using a beamforming technique. The beamforming technique uses multiple UTs to generate each single RF signal with a focus point. For each A-mode RF signal, the signal envelope is obtained as shown in Figure 2-6 b). Then, the envelope signals are used to display a two-dimensional image. Figure 2-6 c) presents a B-mode image of the forearm in which one can observe the contrast at the tissue boundaries within the forearm.

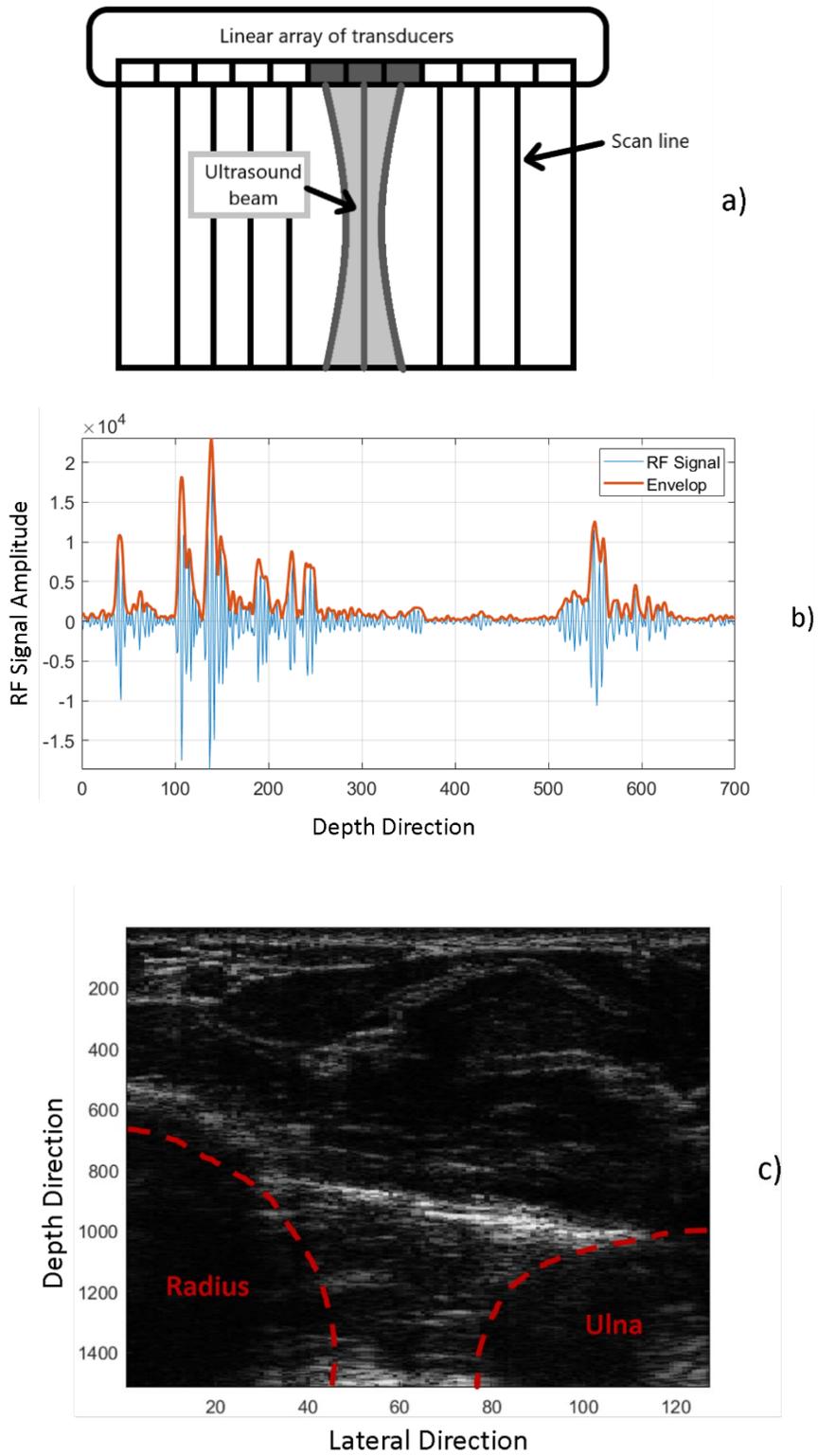


Figure 2-6 a) Linear array of ultrasonic transducers acquiring multiple scan lines of RF signals. b) Envelope of an RF signal. c) B-mode image of tissue in the forearm.

2.1.2.3 M-Mode

A M-mode image is created by taking the signal envelope of a single A-mode RF signal over time. The M-mode image is a representation of the spatially fixed A-mode signal varying over time. A sequence of M-mode frames generates an M-mode image with the frame rate controlled by the pulse repetition frequency of the UT. As shown in Figure 2-7 a) the 120th lateral scan line, highlighted in red, is shown on the B-mode image. In Figure 2-7 b) the M-mode image of the scan line is displayed over time for a recording of a single pinky finger flexion.

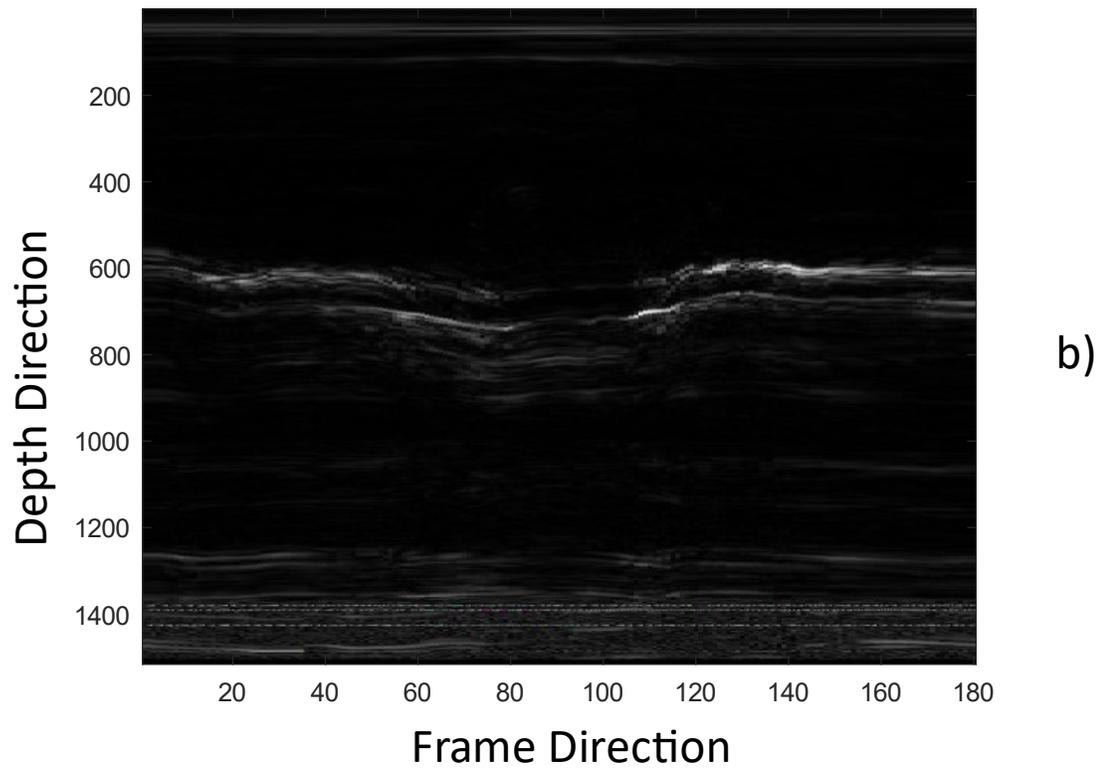
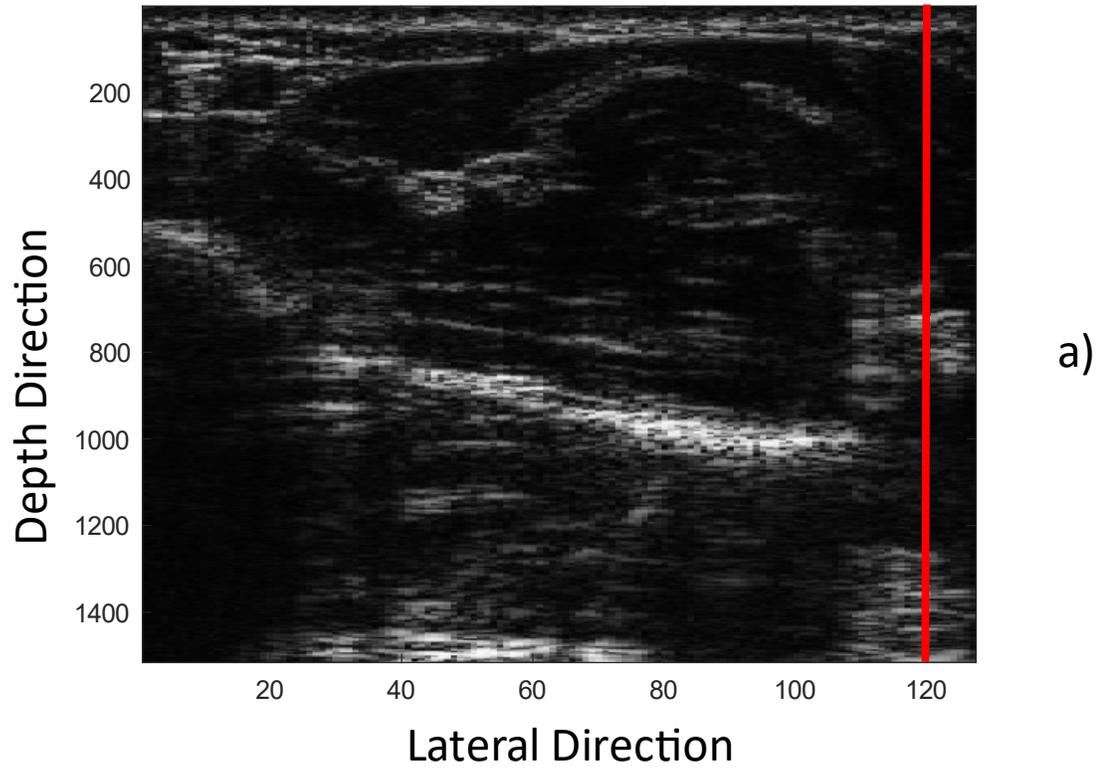


Figure 2-7 a) B-mode image with the 120th scan line indicated in red for b) M-mode image of the 120th scan line during a pinky finger flexion recording.

2.2 Literature Review on Human Machine Interfaces and Prosthetic Applications

2.2.1 Neuromuscular Sensor Strategies

Neuromuscular sensors are the most common sensor control strategy for wearable HMI applications. A neuromuscular sensor is used to record the electrical activity associated with the muscle motions. This relationship between the body's intended control signals via the nervous system and the observed electromyography signals have been researched and applied to myoelectric control since the 1950's [21, 22]. Types of neuromuscular sensors can be categorized based on the sensor's target location: central nervous system, peripheral nervous system, or muscle groups [23].

2.2.1.1 Brain Computer Interfaces

A brain computer interface (BCI) is an HMI that uses the electrical signals from the central nervous system [23]. The BCI method is commonly used for paralyzed subjects who have lost the electrical connection to other parts of their nervous system. This method involves recording electrical activity from the brain using electroencephalography (EEG) or electrocorticography (ECoG) methods. The EEG method records electrical activity of the brain by placing electrodes onto the scalp of the head. The ECoG method records electrical activity of the brain by surgically placing electrodes onto the surface of the brain organ (directly onto the gray matter). EEG is the safer, non-invasive method using surface electrodes to sense brain electrical activity outside the skull. Figure 2-8 shows an example an EEG setup. The EEG method has been used for various HMI applications, such as predictive spelling programs and detecting a hand grasp function of a tetraplegic patient [24, 25, 26]. Although it is the safest BCI

method, the thick skull results in poor signal quality and low spatial resolution [27]. The ECoG method surgically implants an electrode array onto the brain surface allowing for better spatial resolutions than EEG. Although ECoG is invasive, the advantage it has over EEG is that it has proven to achieve individual finger flexion control for upper limb prosthetic devices [28, 29, 30].

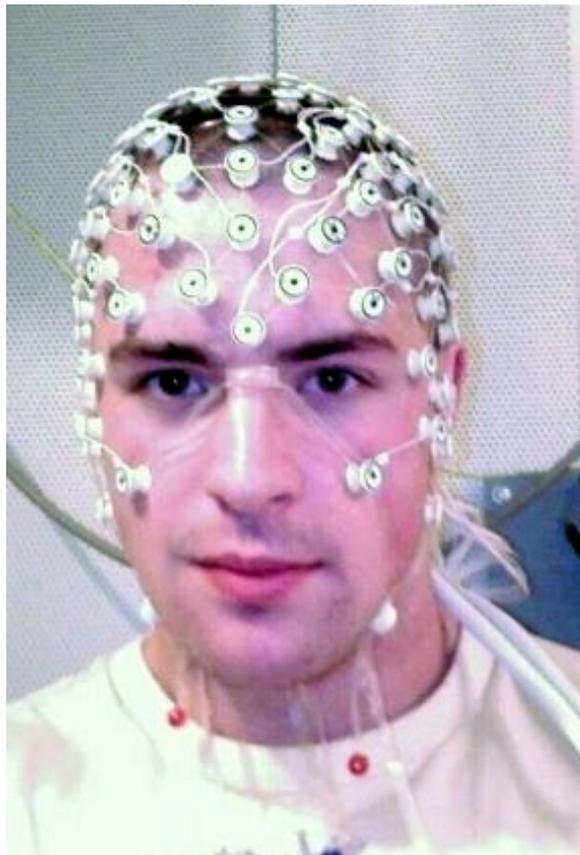


Figure 2-8 Example setup of an EEG recording. Image is reproduced from “EEG Cap” by D. Myers at Dalhousie Life Sciences Center and the image is licensed under the public domain by its author, Thuglas at English Wikipedia. Image is free access online at https://commons.wikimedia.org/wiki/File:EEG_cap.jpg [31].

2.2.1.2 Peripheral Nerve Electrodes

Peripheral nerve electrodes are surgically implanted to record signals directly from the nerves that stimulate the targeted muscles. The methods of interfacing include extraneural or intraneural electrodes. Extraneural electrodes wrap around on the surface of the nerve while intraneural electrodes penetrate the nerve [23]. An extraneural cuff electrode may be used to attach around a nerve bundle for long term use, allowing to remain functional of up to 11 years [32]. An example of an intraneural electrode is the Utah Slanted Electrode Array (array of 96 microelectrodes). A study using this intraneural electrode provided individual finger control for a virtual robotic hand [4]. The issues with peripheral nerve electrodes include damaging of the nerve due to an inflammation after prolonged use [5, 6].

2.2.1.3 Surface Electromyography

Surface electromyography (sEMG) is the most common HMI method for upper limb prosthetic control due to the noninvasive and adaptable placement of the electrodes [33]. In prosthetic devices, sEMG electrodes are placed on the skin surface near the flexor and extensor muscles. Research on individual finger flexion classification has been achieved by using many electrodes or a high-density electrode array sensor. Tenore *et al.* used 32 sEMG electrodes to classify 12 finger flexions and extensions (ten individual and two combined finger flexion/extensions) with a classification accuracy of 98% [34]. Celadon *et al.* investigated the effect of reducing the spatial resolution of a high-density, 8×24 cm, 192-channel electrode grid (8 by 24 electrode matrix) when classifying four finger flexions (index, middle, ring, pinky) [7]. They found that the resolution could be reduced

to down to 24 channels (3 by 8 electrode matrix) while maintaining a high classification accuracy.

The problems with using sEMG, are the occurrence of low signal to noise ratio and signal attenuation along the depth direction [1]. This makes it difficult to spatially resolve muscle activity at varying depth regions within the forearm. While high accuracies can be achieved using a high density electrode array of sEMG signals, most commercial upper limb prosthetics use one to three sEMG electrodes [8]. With a low number of electrodes, the amplitude of the sEMG signal is used by a proportional controller in commercial prosthetics to perform hand opening/closing, grip control, and wrist rotation movements [1].

2.2.1.4 Commercial Prosthetics

The loss of an upper limb has a dramatic impact on the daily life of amputees. It is estimated that one in 190 Americans live with some form of limb loss, which will double by 2050 [35]. About 1 in 5 patients may abandon prosthetic devices due to lack of comfort and functionality [36]. This gives incentive to improve the design of prosthetic devices to be ergonomic if ultrasound is to be used as the HMI.

For industrial upper limb prosthetics, non-invasive sensors use surface electromyography (sEMG) to monitor the associated muscle contractions. Generally, the prosthetic device provides functionality ranging from hand gestures to grip motions. Complete control over all finger flexions individually is usually limited due to lack of electrodes. One commercial prosthetic hand “i-digits” by Touch Bionics uses sEMG to sense general muscle contractions from the forearm. The “i-digits” can perform up to 32

discrete hand gestures and grip motions [37]. The 3D printed “Hero Arm” prosthetic by Open Bionics uses sEMG signals for a proportional controller to perform up to 6 different grips [38]. The more powerful “bebionic hand” by Ottobock provides 14 different grip patterns using individual motors for each finger controlled by a microprocessor [39]. Each of the described upper limb prostheses use sEMG to carry out a specific hand operation. These control schemes involve a proportional controller which relies on the amplitude of the sEMG signals to carry out the hand operation. For example, the “bebionic hand” pinch grip uses the thumb and index finger to open and close in proportion to the amplitudes of the sEMG signals recorded.

2.2.2 Current Research Using Ultrasound for Human Machine Interfaces

Recently, ultrasound has been investigated as an HMI to spatially resolve anatomical muscle activity in deep layered tissue regions. The ultrasound methods include using a clinical ultrasound probe or a multiple single-element UT system. A clinical ultrasound array probe obtains high lateral resolution ultrasound data. A multiple single-element UT system obtains low lateral resolution ultrasound data. This lateral resolution is in correspondence with the lateral sampling interval between each UT in the given systems.

2.2.2.1 Clinical Ultrasound Probe Studies

In the past decade, ultrasound technology has been investigated to distinguish finger movements in prosthetic applications. Shi *et al.* used B-mode ultrasound imaging with the optical flow algorithm to visualize the muscle deformation activity corresponding to each finger [40]. However, as this study was a preliminary study, classification

accuracies of the finger motions were not reported. Castellini *et al.* continued the findings from Shi *et al* and extracted spatial features using two-dimensional linear regression. Three feature values were obtained at every spatial location. Each spatial location included an intercept with two slopes occurring along the row and column directions of the B-mode ultrasound image [41]. The authors showed ultrasound imaging can be used to predict finger positions and flexion force by comparing their method with the ground truth obtained by a force sensor glove. Sikdar *et al.* used B-mode ultrasound imaging to classify complete finger flexions (each prediction is made over the full flexion motion) to obtain a 98% classification accuracy [9]. They also showed that the finger flexion speed can be correlated to the echogenicity of the ultrasound signal. Sikdar *et al.* suggested that B-mode ultrasound imaging has potential to be superior to sEMG sensors, as it can spatially resolve the individual muscle activity occurring at deeper layered tissues. Common problems with using sEMG comes from the inability to detect muscle innervations occurring deep below the surface tissue due to signal attenuation and low signal to noise ratio [1]. Huang et al. compared sEMG with B-mode ultrasound imaging by classifying 14 finger motions and observed that ultrasound-based methods yielded higher accuracy than that of sEMG (96% vs. 90%) [10]. Moreover, studies with amputee subjects proved that ultrasound imaging could be used to aid trans-radial amputees who lost functionality of the hand [42, 43, 44]. Further investigation using B-mode imaging has shown greater accuracies than those of sEMG [10, 45, 46, 47, 48, 49, 50]. The high spatial resolution and image processing techniques used in B-mode imaging methods can achieve very high classification accuracies. However, the high resolution may not be necessary or practical for wearable HMI applications.

2.2.2.2 Multiple Single-Element Ultrasound Transducer Studies

Instead of using a high-resolution probe, multiple single-element UTs could be used to obtain muscle motion information. B-mode ultrasound imaging involves a bulky ultrasound probe, is expensive, and not practical as a wearable device. By using multiple single-element UTs, each transducer could be designed for wearable HMI applications and record the necessary spatial information.

Li *et al.* designed a four-channel UT wearable prototype to distinguish between five finger flexions and a rest state to obtain a classification accuracy of 96% [51]. Their classification accuracy reports on classifying over a 3 second period for the full finger flexion (or rest state) and not classifying every 100 ms in time. Their feature extraction method was adapted by the two-dimensional (2D) linear regression method used by Castellini *et al.* for the ultrasound RF signals [41]. Design of a wearable multichannel UT prototype sparked further researchers to pursue lower resolution designs for the purpose of removing the bulky clinical B-mode ultrasonic imaging probe with a more cost-effective solution [52, 53].

Yang *et al.* designed a wearable 4-channel UT system to classify between 11 different hand gestures [13]. Each UT was 14mm in diameter and 18mm in height. They were able to achieve an offline recognition accuracy of 98.8% for 11 discrete hand positions and real-time recognition accuracy of 95.4% for 10 discrete hand positions. For the real-time testing, the selection time required to make each real-time prediction was 0.243 ± 0.127 seconds and motion completion time was 1.231 ± 0.239 seconds. With selection time referring to the time required to go from the rest position to being in one of the other 10 hand positions. Once the hand is in one of the other 10 hand positions, the

completion time refers to the time required to have the current hand position be predicted 10 times. Their results show that it took 1–2 seconds to confirm if the hand was in one of the 10 discrete hand positions. Then Yan *et al.* of the same research lab constructed wearable ultrasonic sensors (WUSs) with a thickness of 500 μ m, using polyvinylidene difluoride (PVDF) piezoelectric film [14]. They conducted experiments by placing four WUS around the anterior and posterior sides of the forearm. They achieved an offline classification accuracy of 97.6% when predicting for the same 11 discrete hand positions (same positions that were conducted in their previous study). Based on the results conducted in both studies, their system would be optimal for American Sign Language applications. Since the objective in their studies were to confirm if the discrete hand position was achieved, further experiments needed to be conducted on the dynamic classification of individual finger flexions.

Chapter 3: Experimental Emulation Method to Study the Effect of Lateral Spatial Resolution on Individual Finger Flexion Classification

3.1 Introduction

In this chapter, I outline the methodologies used to investigate the effect that lateral spatial resolution has on ultrasound data when classifying individual finger flexions. By using a 127-scan line B-mode ultrasound imaging probe, a reduction of the lateral resolution using an averaging technique is used to emulate ultrasound RF signals from multiple single-element UTs. Using ultrasound RF signals, the proposed pattern classification pipeline (PCP) is used to classify individual finger flexions every 100 ms throughout the recordings. The results and discussion of these methods are presented in Chapter 4. The proposed PCP used to classify individual finger flexion data from ultrasound RF signals is then used with a designed WUS system in Chapter 5.

3.2 Data Acquisition of a Clinical Ultrasound Probe

In this section, the placement of the probe on the forearm is discussed. Then the data format of the ultrasound signals acquired from a 127-scan line B-mode ultrasound imaging probe is described. Next, the procedure to reduce the lateral resolution of this acquired ultrasound data is presented using an averaging technique. Finally, the procedure for the subjects to perform the finger flexions is discussed.

3.2.1 Probe Placement on Forearm

For this study, the placement of the B-mode ultrasound imaging probe on the anterior side of the forearm was chosen to be similar to previous studies [41, 54]. McIntosh *et al.* conducted a study to compare ultrasound probe placement between transverse, longitudinal, diagonal, wrist and posterior as illustrated in Figure 2 of the manuscript “EchoFlex: Hand Gesture Recognition using Ultrasound Imaging” [46]. In their study, 12 participants were recruited to perform 10 discrete hand gestures. Using a one-way mixed analysis of variance, they found no significant difference when comparing between diagonal with wrist and longitudinal with transversal. They concluded that the diagonal and wrist position were the best positions. Based on their results, I decided to conduct the experiments by placing the ultrasound probe 5 cm from the wrist on the anterior side of the forearm as shown in Figure 3-1 to capture activity of the flexor muscles.



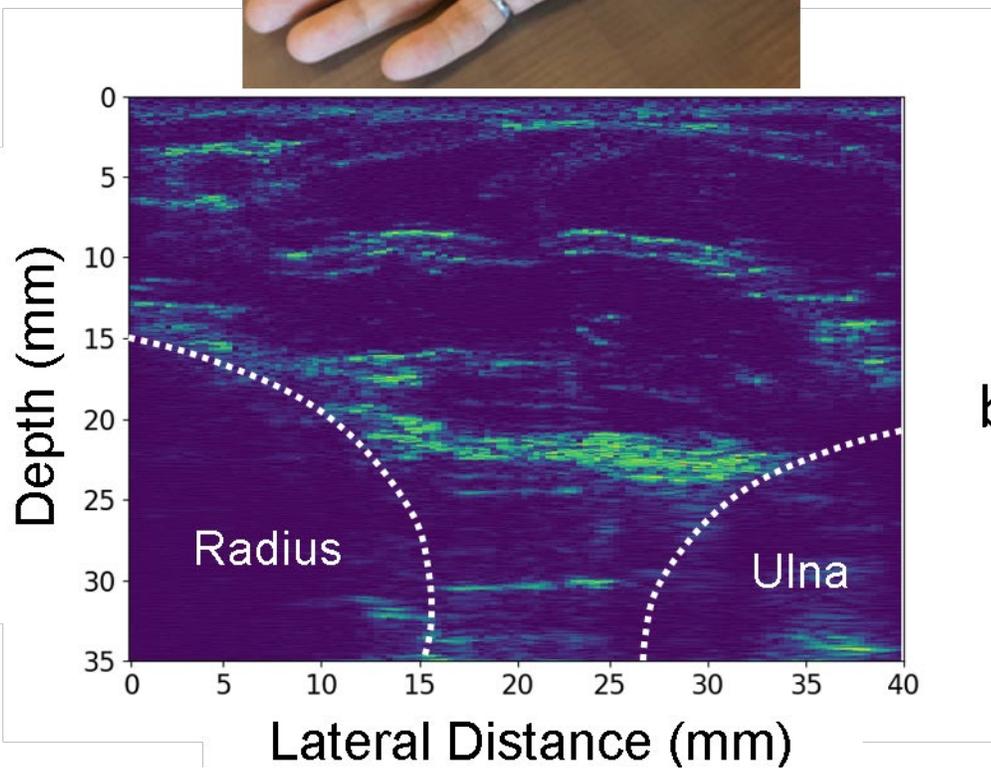
Figure 3-1 Placement of the ultrasound probe on the anterior side of the forearm, 5 cm away from the wrist.

3.2.2 Ultrasound Data Format

The clinical ultrasonic imaging system (Model PICUS from ESOATE Europe, Maastricht, Netherlands) was used to record the ultrasound data with a 40-mm linear array probe (Model: L-40). As seen in Figure 3-2 a), the probing surface was attached on the anterior side of the forearm with ultrasound gel couplant. The linear array probe was carefully positioned to observe the muscle contractions associated with the finger flexion being performed. Figure 3-2 b) shows a B-mode ultrasound image of the forearm region including flexor muscles located above the radius and ulna bones.



a)



b)

Figure 3-2 a) Experimental setup to acquire ultrasound RF signals during an individual finger flexion using a linear array ultrasonic probe. b) corresponding B-mode ultrasound image of the forearm. The field of view is 40-mm in width and 35-mm in depth.

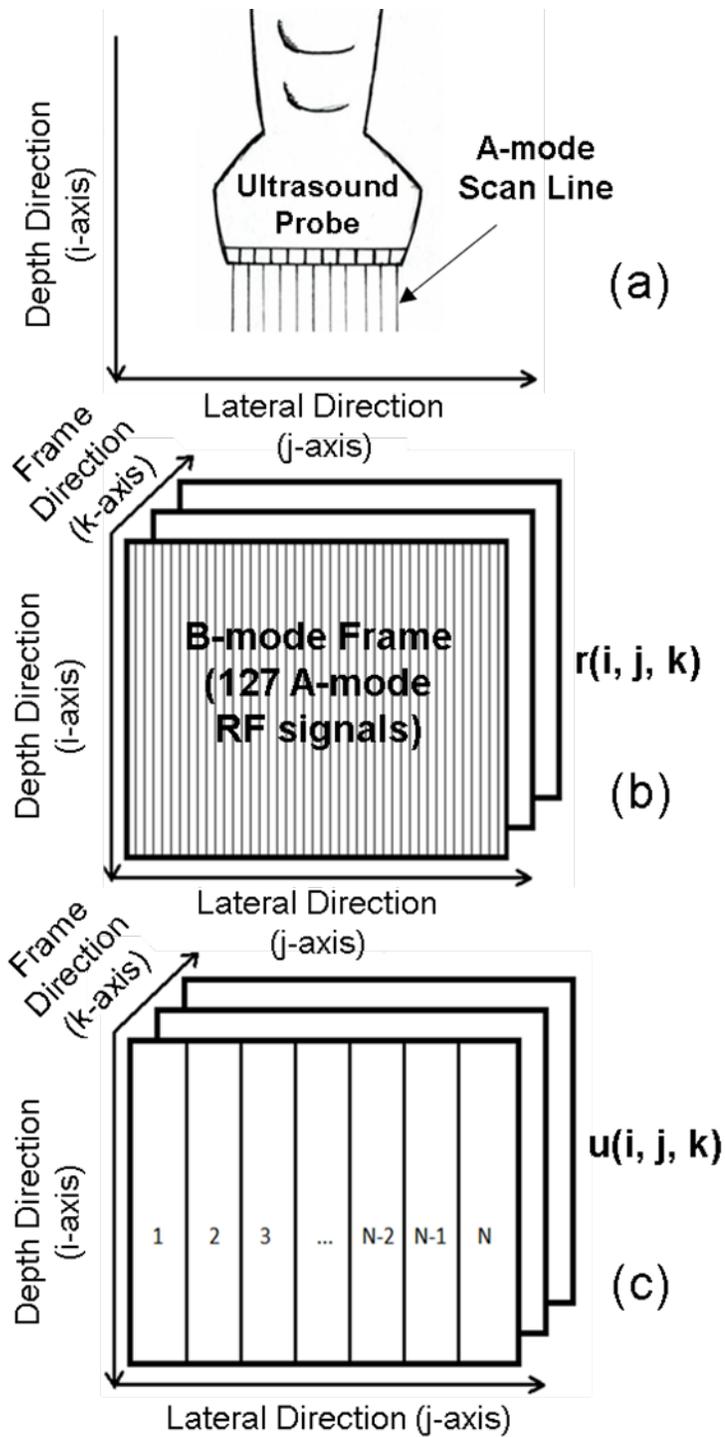


Figure 3-3 Schematic representation of ultrasound data acquisition coordinate and structure: (a) ultrasound probe orientation, (b) data structure of the acquired 127 A-mode RF signal, $r(i, j, k)$, in B-mode frame, and (c) N channels of reconstructed RF signals, $u(i, j, k)$, at the reduced lateral spatial resolution.

Figure 3-3 depicts the data acquisition coordinate and ultrasound data structure. As shown in Figure 3-3 a), the depth direction (i-axis) follows the direction into the forearm and the lateral direction (j-axis) follows along the surface scan-line direction. Figure 3-3 b) presents the acquired ultrasound data structure of the ultrasound system and probe used in this thesis research. The frame direction (k-axis) corresponds to the M-mode frame of the pulse-echo sequence. The ultrasound data acquired successive B-mode frames during the individual finger flexions with an M-mode frame rate (pulse repetition frequency) of 30 frames per second. Each B-mode frame consisted of 127 A-mode scan lines ($j = 1, 2, \dots, 127$) of ultrasound RF signals. The lateral spatial sampling interval between two consecutive A-mode scan lines was $315 \mu\text{m}$ per RF signal. The sampling frequency for each A-mode scan line (depth samples along the i-axis) was 33.3 MHz. The total depth consists of 1516 samples ($i = 1, 2, \dots, 1516$) for a total of a 35-mm depth range. Assuming an average ultrasound speed of 1540 m/s in soft tissues, the sampling interval along the depth direction is $23.1 \mu\text{m}$ per sample.

3.2.3 Lateral Resolution Reconstruction Procedure

In this thesis lateral spatial resolution will be referred to as the sampling interval between the number of ultrasound RF signals within a confined space. To emulate acquired ultrasound RF signal data at reduced lateral spatial resolutions, the acquired 127 A-mode RF signals are reconstructed into the reduced number, $N (< 127)$, of A-mode RF signals, as shown in Figure 3-3 c). Thus, we refer to lateral spatial resolution in this paper as the number of UTs that would be confined within a fixed 40 mm width. The data is obtained by averaging the acquired RF signals along the B-mode scan line (lateral) direction. This

simulation approach to reduce the lateral resolution is based on the fact that when the tissue depth range of interest is within the near field (Fresnel) region, the ultrasound beam size could be approximated the same as the UT size [20]. The center ultrasound frequency of the linear array probe was 6.6 MHz. This results in the near field region up to the depth of 35 mm when the UT size is 2.9 mm by 2.9 mm using $D_f = D^2/\lambda$. Where D_f is the Fresnel distance, D is the UT width, and λ is the ultrasound wavelength in the soft tissue [20]. Therefore, under our measurement conditions, the maximum measurement depth of 35 mm was within the near field region when the averaged channel width is greater than 2.9 mm (i.e., $N \leq 14 \approx 40 \text{ mm}/2.9\text{mm}$) which is the UT size range of interest. The data is reconstructed by averaging the acquired RF signals along the lateral direction (j-axis). As shown in Figure 3-3 b), the individual 127 acquired RF signals are represented by $r(i, j, k)$. Then, the reconstructed RF signal, $u(i, j, k)$, is given in (3.1).

$$u(i, j, k) = \sum_{n=(j-1)\lfloor \frac{127}{N} \rfloor + 1}^{j\lfloor \frac{127}{N} \rfloor} \frac{r(i, n, k)}{\lfloor \frac{127}{N} \rfloor} \quad (3.1)$$

Where $\lfloor \cdot \rfloor$ indicates a floored integer division in (3.1). After data reconstruction in u , the lateral direction (j-axis) ranges from integers 1 to N . This reconstruction of the RF signals emulates ultrasound data being acquired by N pieces of UTs in parallel. This emulates a reduced lateral resolution using UT with larger widths. Therefore, the smaller N values result in the lower lateral spatial resolution. For example, when $N = 4$, every consecutive 31 RF signals are averaged into a single RF signal. The reconstructed data emulates the data being acquired from a 4-channel UT system. In the emulated system, each UT

corresponds to having an approximate width of 10 mm. The original lateral resolution acquires the ultrasound data at 330 μm per RF signal ($N = 127$).

I believe this averaging method is a better method than selecting individual RF signals at the approximated center of the emulate UTs. This is because when emulating for UTs of varying sizes, the width of the UT will also come into effect. Hence for the purpose of the emulation study specifically using a B-mode linear array probe, averaging amongst consecutive RF signals would be more effective than individual RF signals when N is reduced.

3.2.4 Individual Finger Flexion Recording Procedure

This study was approved by Carleton University's Research Ethics Board (CUREB-B #10496). Five healthy human subjects consisting of males and females aged between 23-27 were recruited to perform the experiments on both their right and left forearms. Ten independent arm experiments were conducted with the help of five subjects. For each experiment, three trial recordings of individual finger flexions were performed for each finger. Thus, for each experiment, 15 data acquisitions (5 fingers x 3 trials) were conducted for each arm. For each trial recording, a tick sound is made every 1 second from a metronome. This was used to regulate the speed of each finger flexion consistently. First, the subjects had their hand in a comfortable rest state where all the fingers were set to be naturally extended (hand position with palm open), and then performed the finger flexion. The flexion had the fingertip touch the palm of the hand within 3 seconds, then return to the original rest state in another 3 seconds. This 6 second recording is represented as a single finger flexion trial recording. The subject was asked

to continuously perform the motions until three completed finger flexion trials were recorded for each finger. Because the finger flexions were classified on a per arm basis, each subject was asked to continuously replicate each finger flexion with the same pattern to the best of their ability. For each trial recording, 180 frames ($k = 1, 2, \dots, 180$) were acquired over 6 seconds. With the desire to predict every 100 ms in time, this means that 60 predictions would occur over one finger flexion recording trial. Three trials of recordings were obtained for each finger. In total, 2700 frames ($= 180$ frames per trial $\times 3$ trials per finger $\times 5$ fingers) were recorded for each arm experiment. Recording at a frame rate of 30 frames per second, a total of 90 seconds of finger flexion recording were used for each arm experiment.

3.3 Pattern Classification Pipeline

The procedure for classifying the individual finger flexions using acquired ultrasound RF signal data are outlined in Figure 3-4. Spatial features are extracted from the ultrasound RF signal data, which contain spatial information within the forearm to potentially discriminate between the five finger flexions. I chose two comparable spatial feature extraction methods and compared their classification performance. The spatial feature extraction method proposed in this thesis was adopted from sEMG studies, which involves signal preprocessing with the DWT and spatial feature extraction from the MAV [55, 56, 57, 58, 59, 60, 61]. The other spatial feature extraction method uses the envelope of the signal and extracts spatial features with linear regression [13, 14, 41, 51]. Temporal features are extracted by obtaining the instantaneous change in spatial features over time. These temporal features correspond to the change in tissue motion over time associated with the matching finger flexion. The number of the extracted spatial and temporal features are then reduced by the feature selection process, to only accept tissue regions with muscle activity that helps to discriminate between the type of individual finger flexion. Finally, every 100 ms of the finger flexion recordings are classified using linear discriminant analysis (LDA). LDA was used as the machine learning classifier because it is fast to compute and has been used in previous sEMG and ultrasound related studies [13, 14, 52, 53, 62].

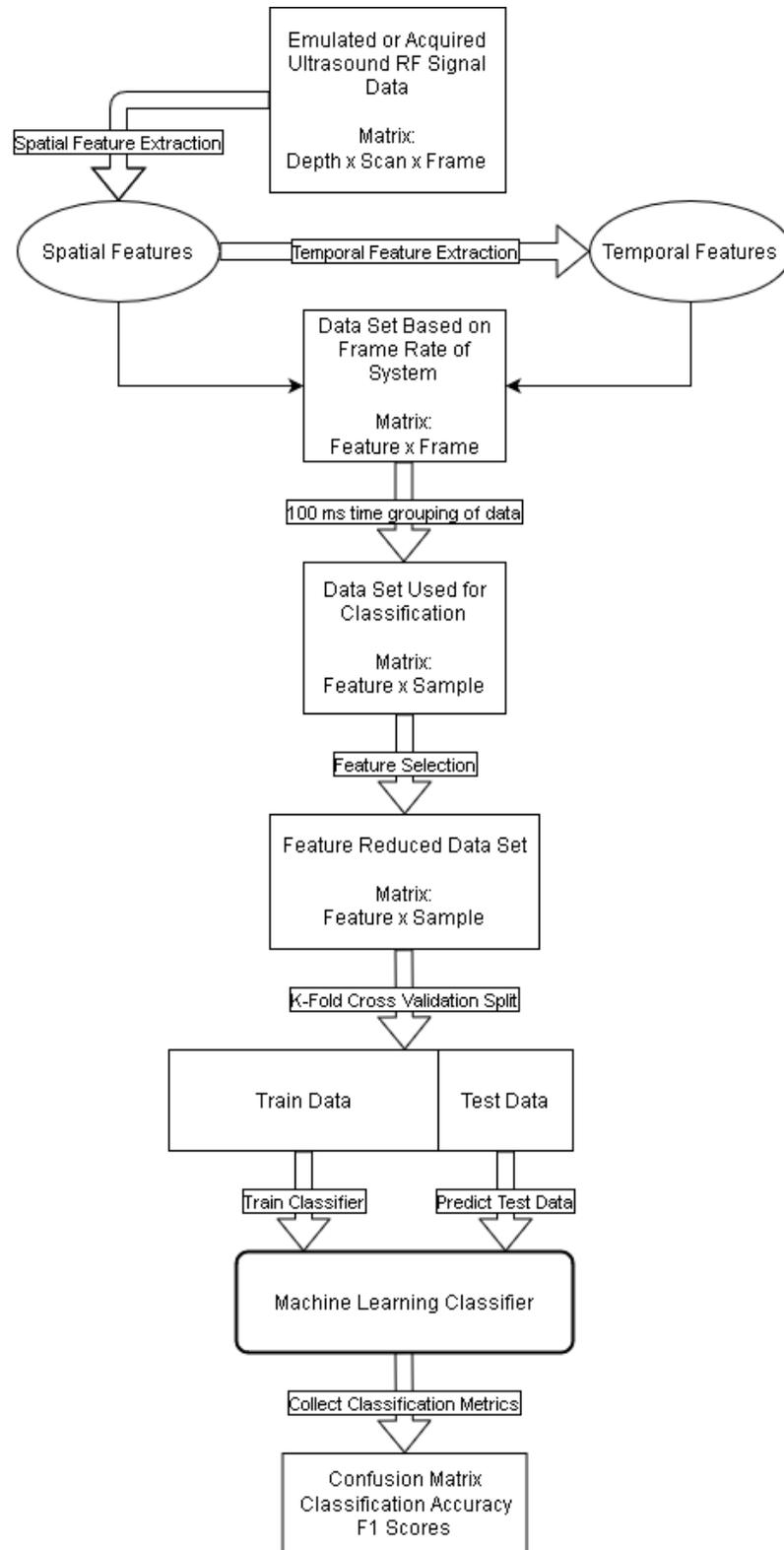


Figure 3-4 Block diagram of the pattern classification pipeline used to classify individual finger flexions from ultrasound RF signal data.

3.3.1 Spatial Feature Extraction

Two methods of spatial feature extraction were employed on the ultrasound RF signal data of $r(i, j, k)$, with i-axis as the depth direction, j-axis as the lateral direction, and k-axis as the frame direction as described previously in Figure 3-3 b). The first method preprocessed the reconstructed RF signals with the DWT. The RF signals were processed into multiple time-frequency resolution levels, computed as detail coefficients. Then, the mean absolute value (MAV) function was applied to the detail coefficients. This approach was adapted from methods commonly used in sEMG signal analysis using pattern recognition [55, 56, 57, 58, 59, 60, 61]. The MAVs were used as the spatial features of the ultrasound data. This approach will be referred as “DWT-MAV method” hereafter. The second method was described by Yang *et al.* for A-mode RF signals [13]. The signal envelope of the ultrasound RF signals were computed, then linear regression (LR) was applied across segmented regions of the signal envelope. The slopes and intercepts of the estimated linear function were used as a spatial feature. This approach is called as “ENV-LR method” hereafter. This ENV-LR method has been commonly used amongst ultrasound-based research studies with only minor modifications between the studies [13, 14, 41, 51].

3.3.1.1 The DWT-MAV Method

The DWT was performed using Mallat's filter bank implementation [63]. As shown in Figure 3-5, the DWT performed on r was conducted by recursively computing the detail coefficients D_l from the previous level analysis coefficients A_l at each level l . The A_l and D_l are computed along the depth direction (i -axis) of r , by applying the high pass filter, h , and the low pass filter, g . The ultrasound RF signals were downsampled by two at each level l . This resulted in the depth sampling interval to effectively double after each level.

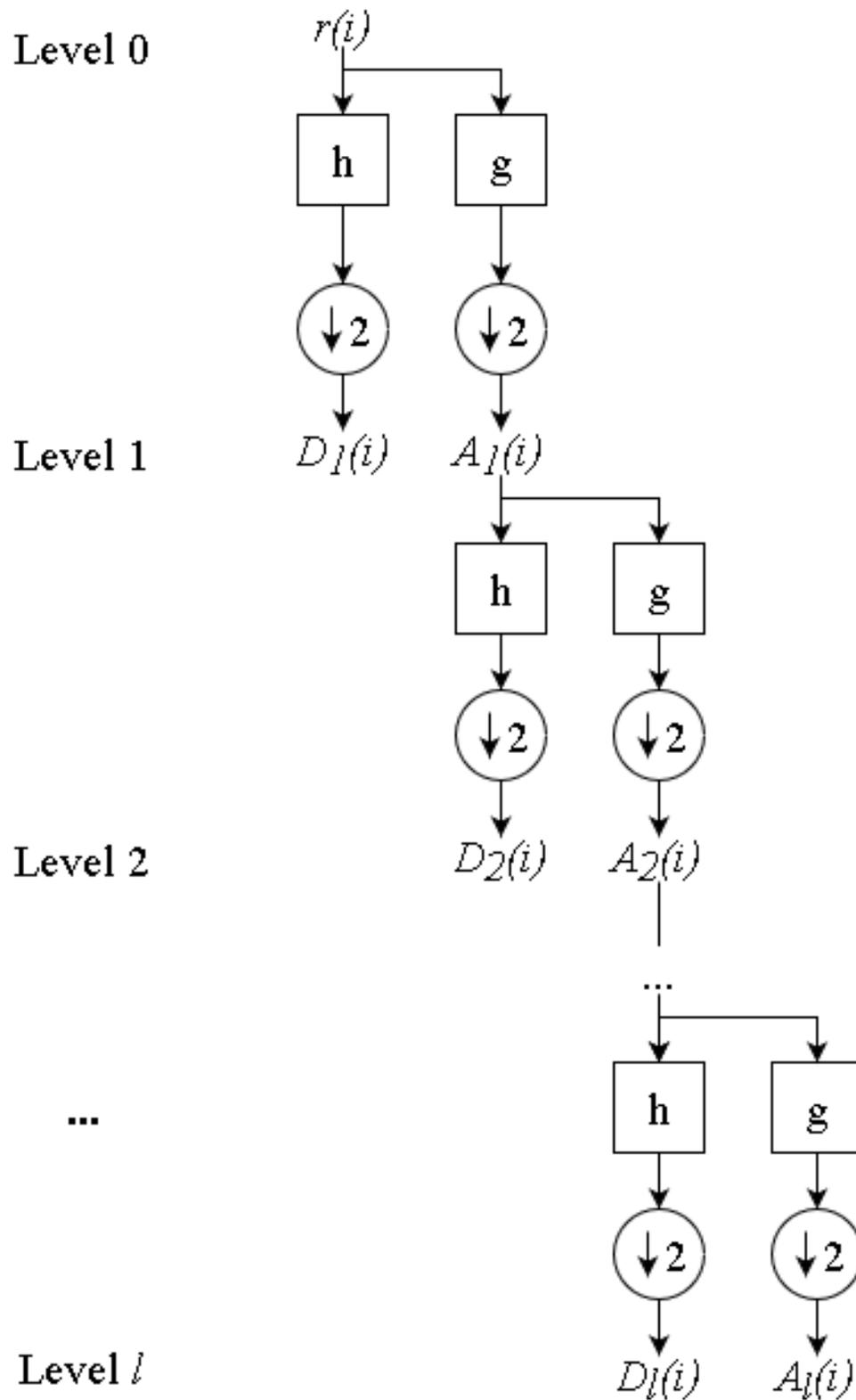


Figure 3-5 Discrete wavelet transform to compute the detail coefficients of ultrasound RF signal data using Mallat's algorithm [61].

The one-dimensional DWT equations to compute the A_l and D_l using Mallat's algorithm are expressed by (3.2) and (3.3), respectively [63].

$$A_l(i, j, k) = \frac{1}{\sqrt{2}} \sum_n g(n) A_{l-1}(2i - n, j, k) \quad (3.2)$$

where $A_0(i, j, k) = r(i, j, k)$

$$D_l(i, j, k) = \frac{1}{\sqrt{2}} \sum_n h(n) A_{l-1}(2i - n, j, k) \quad (3.3)$$

The DWT was performed using PyWavelets software package for Python [64]. In a preliminary test between using Haar (same as Daubechies 1), Daubechies 2, 3, 4, and 5 mother wavelets, the results showed no significant difference in overall accuracy between the wavelet type. One reason why I think there would be no difference is because the frequency ranges of the levels are relatively the same regardless of the wavelet. As a result, the Daubechies 3 wavelet was used because it visually had the smoothest frequency response to aid in the visualization of the level bandwidths as shown in Figure 3-6 b).

Figure 3-6 a) shows an example of an ultrasound RF signal for the illustration of spatial feature extraction. The large signal amplitudes are observed where ultrasound was reflected from tissues boundaries. The frequency spectrum obtained by the FFT of the entire RF signal in Figure 3-6 a) is presented in Figure 3-6 b), by the solid line. The bandwidths of the DWT filters at the levels from 1 to 6 using the Daubechies 3 mother wavelet. Most of the frequency components of the ultrasound RF signals are depicted within the range from 5 MHz to 7.5 MHz. In Figure 3-6 b) the frequency range at levels 4

and higher are shown to visualize the lack of frequency components from the RF signal for these levels. Therefore, the spatial features were extracted only from the first three levels of the D_l ($l = 1, 2$ and 3).

The MAV was calculated from the D_l at each level of l using (3.4). Therefore, the spatial feature vector S for the frame number (k -axis) is given by (3.5), where m is an integer ($1 \leq m \leq M$). The total number of spatial features results in M being equal to N number of reconstructed RF signals multiplied by the number of MAV coefficients per reconstructed RF signal. The length of the MAV window was experimentally chosen to be 16 to have the best performance.

$$MAV_m(k) = \frac{1}{16} \sum_{v=0}^{15} |D_l(16i - v, j, k)| \quad (3.4)$$

$$S_{\text{DWT-MAV}}(k) = [MAV_1(k), MAV_2(k), \dots, MAV_M(k)] \quad (3.5)$$

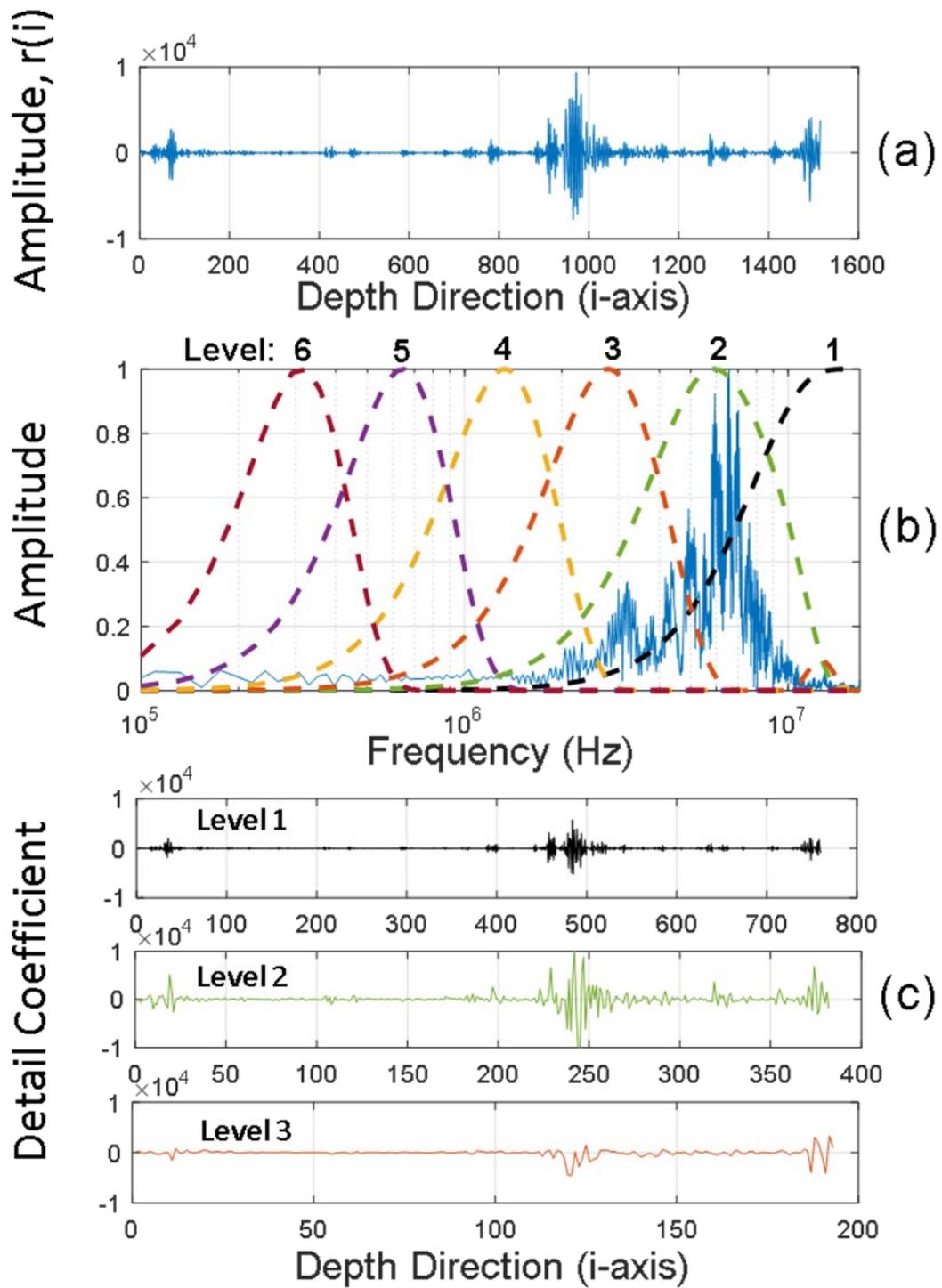


Figure 3-6 a) ultrasound RF signal, b) frequency bandwidths of the DWT filters at the levels from 1 to 6 using the Daubechies 3 mother wavelet (dotted lines) and the frequency spectrum (solid line) of the RF signal, and c) wavelet detail coefficients for the levels 1, 2, and 3 computed with the RF signal.

3.3.1.2 The ENV-LR Method

Figure 3-7 b) shows the signal envelope e , of the RF signal r in Figure 3-7 a), which is the same RF signal r in Figure 3-6 a). The signal envelope e was obtained by the Hilbert transform and log compression after filtering out the DC offset and noise in the RF signal. The signal envelope e was then segmented into regions of the equal size without overlapping along the depth. Figure 3-7 c) shows an example of the segmented region having the length of 151 sample points in e , denoted by the circle marks. For each segmented region in e , a line function having the slope a and intercept b was estimated by linear regression (LR). The linear function obtained with the presented 151 sample points is indicated by the solid line in Figure 3-7 c).

These two parameters of a and b were used as the spatial features for finger flexion classification. Therefore, S obtained by the ENV-LR method is given by (3.6) for the frame number (k -axis), where m is an integer ($1 \leq m \leq M$) and M is equal to N number of reconstructed RF signals multiplied by 10 segments per RF signal. The total number of the spatial features used by the ENV-LR method was $2 \times M$ to account for both slope and intercept values. The placeholder variable M discussed in this section has a different value than the M discussed in 3.3.1.1 for the DWT-MAV spatial feature extraction method.

$$S_{\text{ENV-LR}}(k) = [a_1(k), b_1(k), a_2(k), b_2(k), \dots, a_M(k), b_M(k)] \quad (3.6)$$

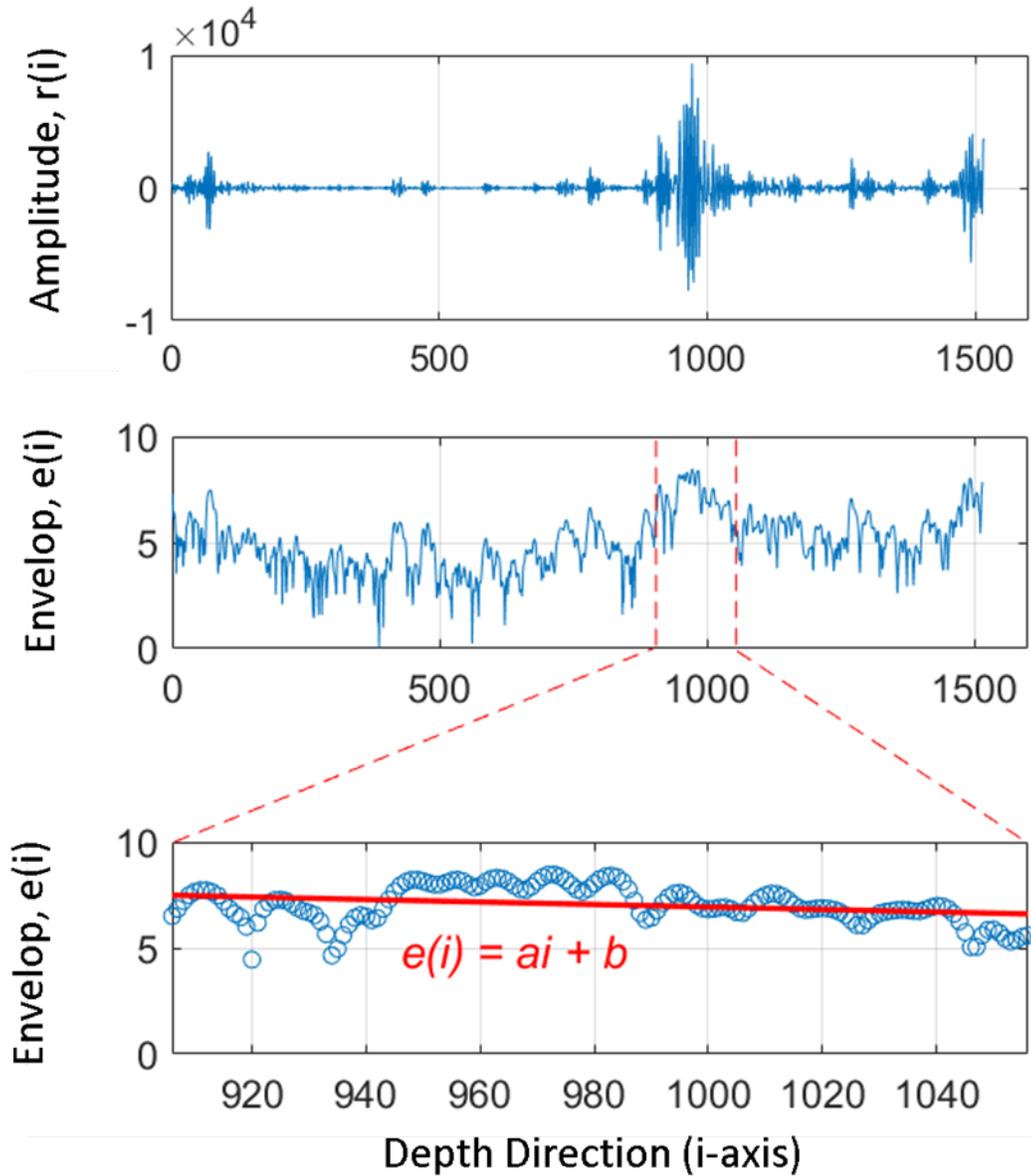


Figure 3-7 Procedure of ENV-LR method for spatial feature extraction adapted from [10]: a) RF signal, b) envelope using the Hilbert transform, and c) linear fit (solid line) using linear regression of the envelope (circles) in the given segment region denoted by dashed lines in b).

3.3.2 Temporal Feature Extraction

A temporal feature T represents the dynamic motion of the muscles captured in time series of the frames along the k -axis. The temporal feature T is obtained by taking the difference between the spatial features of two consecutive frames by the DWT-MAV or ENV-LR methods ($S = S_{ENV-LR}$ or $S_{DWT-MAV}$) as expressed by (3.7). The DWT-MAV and ENV-LR are conducted separately as independent experiments for comparison. The temporal feature extraction obtains the difference between frames of the spatial features, depending on which spatial feature extraction method is being used in the PCP.

$$T(k) = S(k + 1) - S(k) \quad (3.7)$$

3.3.3 Procedure for Continuous Classification

At this point, the ultrasound data has been converted into a matrix of feature \times frame. Meaning that up to this point, spatial S and temporal T features have been extracted for each frame. To make the finger flexion classification occur every 100 ms in time, the frames obtaining S and T features were reduced as follows. First, the minimum and maximum values for each feature in the S and T feature vectors were retrieved along every 3 frames (k -index), which corresponds to the time interval of 100 ms. Three frames were chosen because the frame rate acquired with the ultrasound probe was 30 frames per second (33.3 ms per frame). This procedure was modified for the WUS system in Chapter 5 because the ultrasound RF signals are acquired with a different frame rate. The time interval of 100 ms was chosen based on the lower bound of the optimal controller delay for prosthetic devices [2]. Therefore, 2700 frames resulted in 900 labeled samples to be

used in the classification step. The matrix of feature \times frame is then converted into feature \times sample. With sample represented as a 100 ms time interval of the ultrasound recordings.

3.3.4 Feature Selection

To further reduce redundant features that are similar to other features, feature selection was performed using one-way ANOVA F-test [65]. The ANOVA F-test is used to measure the linear dependency that each feature has from each other feature when used to discriminate between finger type. In this feature selection process, all the features are given a score of a p-value calculated from the F-statistic measure. This p-value is then used to accept the features based on the family-wise error rate criterion given by (3.8), where the threshold α was set to 0.05 (95% confidence interval that the feature is relevant). Due to the family-wise error rate, the rejection of the p-value, controlled by α , scales with the number of features. The family-wise error rate criterion is used to account for the large number of features as N increases because a p-value statistic is being calculated for each feature. The importance of using the family-wise error rate (scaled by the α value) is because as N varies from 1 to 127, the number of features per sample increases drastically (number of spatial features and temporal features increases above 100,000 for the higher resolutions). This feature selection process was performed using the Scikit-learn Python library [66].

$$p < \frac{\alpha}{\text{total number of features}} \quad (3.8)$$

3.3.5 Performance Evaluation

Classification of the individual finger flexions were conducted for each arm using three-fold cross-validation. Ten independent experiments were performed between subject arm (both arms from five subjects). This is because the design and control of a prosthetic arm could be unique to the chosen arm of the subject. Among the three trials of the finger flexions (900 labeled samples) for each arm, the features obtained from the two trials (600 samples) were used for the training data set, and the remaining trial (300 samples) was used for the testing. This is repeated three times for each trial in a three-fold cross-validation procedure. The performance of classifying the individual finger flexions for each arm was calculated. The classification accuracy of each cross-validation fold is calculated by (3.9). At each resolution of N , a total of 3 classification accuracies are computed for each of the ten independent subject arm experiments. Also, the F1-score for each individual finger flexion was calculated by (3.10). This is calculated as the number of true positive TP , true negative TN , false positive FP , and false negative FN predictions made by the classifier for each finger flexion.

$$\text{Classification Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (3.9)$$

$$F1_{\text{finger}} = \frac{2 \times TP}{2 \times TP + FP + FN} \times 100\% \quad (3.10)$$

Chapter 4: Results and Discussions of the Effect of Reduced Lateral Resolution for Individual Finger Flexion Classification

4.1 Introduction

This chapter presents the results and discussions of the methods outlined in Chapter 3. First, the results of a preliminary study using ultrasound data acquired from a single subject arm are presented. The preliminary study investigates the correlation between classification accuracy and lateral resolution. Then a detailed study using ultrasound data acquired from ten subject arms are presented. Using ten arm data sets, ten independent experiments are conducted to validate the results of the preliminary study. The results on the PCP proposed in Chapter 3 are presented in this chapter. The design of the PCP that performed best (for the emulated low lateral resolutions) is then used with a low lateral resolution multichannel WUS system in chapter 5.

4.2 Preliminary Study on Reducing Lateral Spatial Resolution

In this preliminary study, only the left arm of subject one is used. For the pattern classification pipeline, the “ENV-LR” method is used as the feature extraction method (see section 3.3.1.2). The ENV-LR method was chosen because this method has been used in several previous studies with only minor modification made to the method [13, 14, 41, 51]. This purpose of this preliminary study is to investigate classification accuracy as a function of lateral resolution. Then the results of emulating the ultrasound data as a low lateral resolution multichannel UT system is presented.

4.2.1 Correlation Between Classification Accuracy and Lateral Resolution

The classification accuracy is defined to be the percentage of the total number of correct finger predictions made throughout the full finger flexion recording. With 90 seconds of finger flexion recording occurring for the single arm experiment and predictions made on every 100 ms in time, 900 samples in total are classified. The number of reconstructed RF signals N was varied from 1 to 127 and resulted in the following asymptotic trend in Figure 4-1 with a coefficient of determination to be 0.9332.

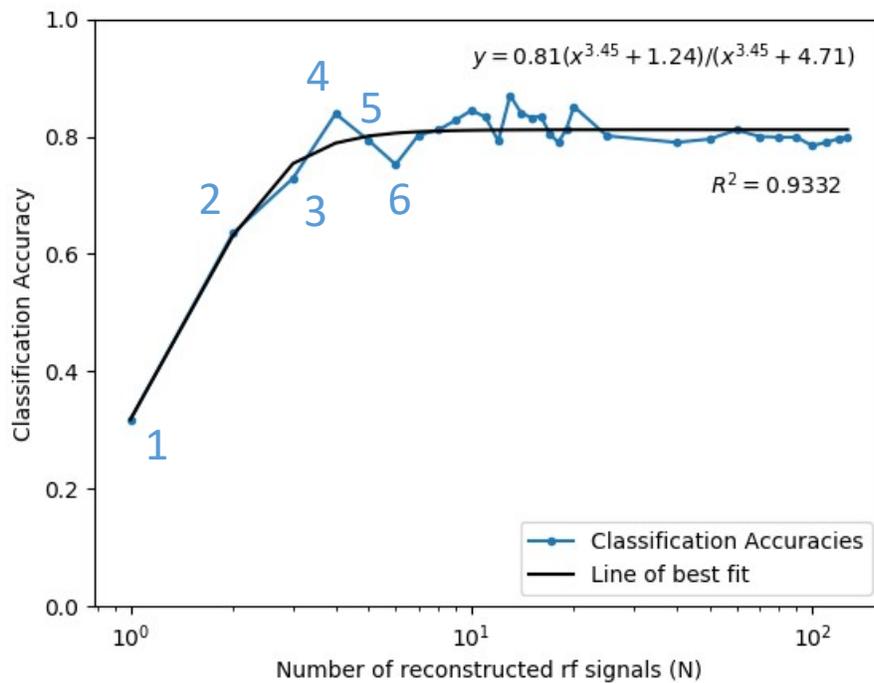


Figure 4-1 Classification accuracies versus number of reconstructed RF signals for subject 1 left arm

4.2.2 Classification Accuracy and F1 Score at a Reduced Lateral Resolution

Looking at the specific example of when $N = 4$ shown by the confusion matrix in Figure 4-2, the effective lateral sampling interval between each reconstructed RF signal becomes 10 mm and results in a classification accuracy of 83.89%. The precision and recall metrics are calculated in Table 4-1, which shows that the ring finger has the lowest calculated F1 score at 77.01% and the pinky finger to have the highest calculated F1 score at 91.35%.

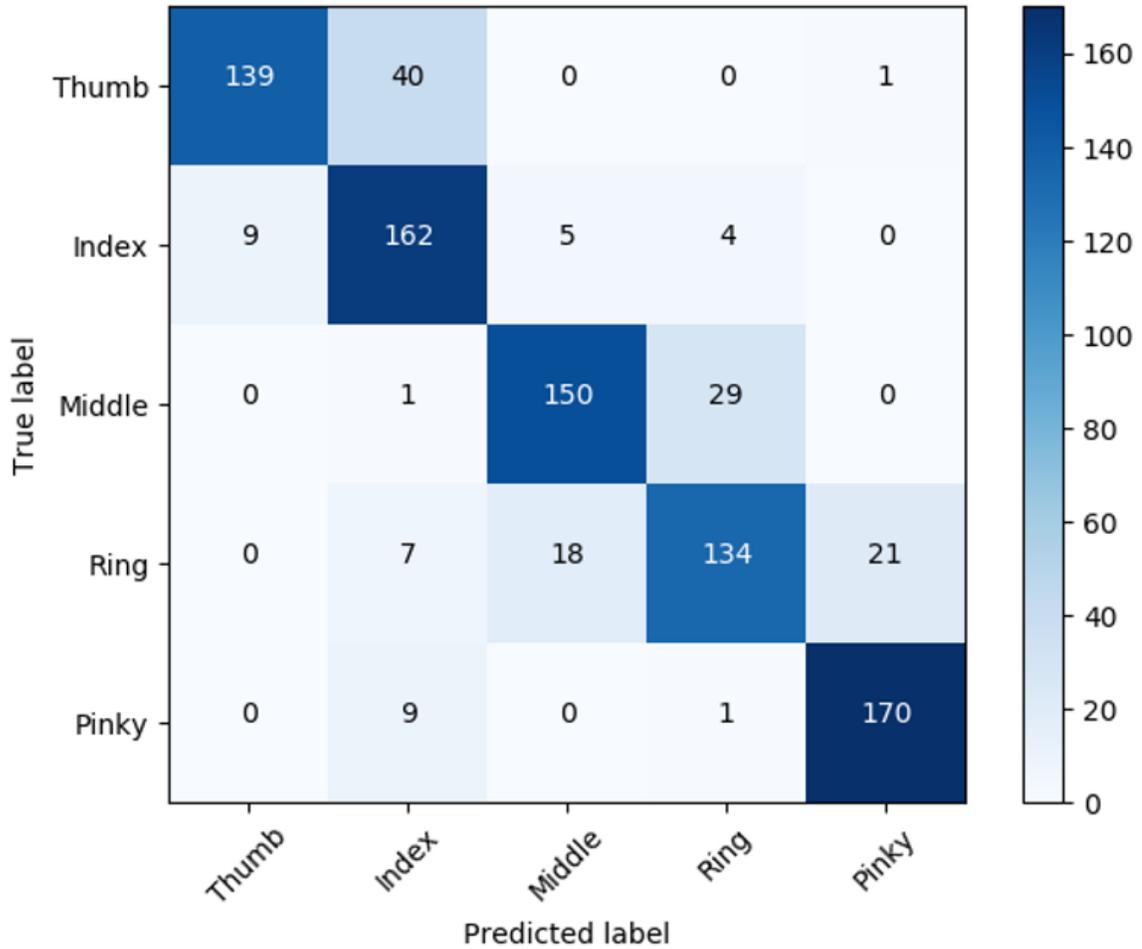


Figure 4-2 Confusion matrix for $N = 4$ reconstructed rf signals of subject 1 left arm.

Table 4-1 Precision and recall metrics for $N = 4$ reconstructed RF signals

Finger	Precision	Recall	F ₁ Score
	$\frac{TP}{(TP + FP)}$	$\frac{TP}{(TP + FN)}$	$2 \times \frac{Precision \times Recall}{Precision + Recall}$
Thumb	93.92%	77.22%	84.74%
Index	73.97%	90.00%	81.22%
Middle	86.71%	83.33%	84.97%
Ring	79.76%	74.44%	77.01%
Pinky	88.54%	94.44%	91.35%

4.2.3 Discussion

The results present the correlation between classification accuracy and lateral resolution. This correlation indicated that increasing the number of reconstructed RF signals N plateaus as an asymptotic function. As the number of reconstructed RF signals N is increased (across a 40 mm width of the ultrasound probing surface), the classification accuracy plateaus at around $N = 4$. This is represented as having a 10 mm lateral sampling interval between the four RF signals. This suggests that the full lateral resolution may not be needed to predict individual finger motions. When the effective lateral resolution is reduced to a single RF signal (no spatial variation in the lateral axis) the classification accuracy drops statistically in comparison to the full resolution. This gives evidence that using more than one UTs is beneficial.

These results show motivation that a multichannel single element UT system could be used in HMI applications instead of a complex bulky ultrasonic imaging probe. The results suggest that a multiple single element UTs could achieve similar accuracies of using a linear array probe. Implementation of multiple WUS [67] has a potential to perform individual finger flexions in a less costly and more practical manner than a conventional clinical ultrasound probe and imaging system.

4.3 Detailed Study on Lateral Resolution and Individual Finger Flexion

Classification

In this section, the performance of the proposed PCP is verified using ten independent subject arm experiments. This chapter studies the effect of reducing the lateral resolution of ultrasound RF signals using all methods outlined in Chapter 3. First the extracted features of the PCP are visualized using linear discriminant analysis (LDA) projections. In the sections following the LDA projections, the two spatial feature extraction methods DWT-MAV and ENV-LR are compared. The methods will be compared based on classification performance and computation speed at reduced lateral resolutions. The better spatial feature extraction method will be used in the PCP for the WUS system experiment in Chapter 5.

4.3.1 Linear Discriminant Analysis Projections of the Extracted Features

The features extracted from the PCP are visualized using LDA projections. The features are extracted using the DWT-MAV method from one subject arm. For these projections, the DWT-MAV method was chosen over the ENV-LR method to save computation time. The purpose of these projections is not to compare the DWT-MAV method with ENV-LR method but to visualize the effect that lateral resolution has on the feature space. By examining the feature space, a better understanding can be presented for how lateral resolution affects classification accuracy.

The third trial is used as test data and are projected onto the two-dimensional LDA component projection. Figure 4-3 shows the LDA projections at different lateral resolutions using the subject 3 right arm data set. The coloured regions are the decision

boundaries using the first and second trial data as training data. As shown in Figure 4-3 a), b), and c), it is apparent that each of the fingers form clusters together on the projected space within the decision region boundaries for the test data. The thumb and pinky finger test data completely falls onto the proper decision boundary region. However, several of the middle, index and ring finger test data points are misclassified as other middle, index and ring fingers. One reason why the index, middle and ring fingers may be misclassified is because these three fingers share the same flexor muscles for the hand grasp motion. At the lowest resolution of $N = 1$ in Figure 4-3 a), the test data points have the greatest spread amongst presented resolutions. At the full resolution of $N = 127$ in Figure 4-3 c), the test data points have the least spread amongst the presented resolutions. At the resolution $N = 12$ in Figure 4-3 b) the spread of the data points are less than $N = 1$ but greater than $N = 127$. The trend to be inferred, is that when the lateral resolution is reduced, the variance of the test data points increases. At the resolution $N = 12$ in Figure 4-3 b), it visualizes that most of the test data points lie within the corresponding decision regions. Meaning correct classifications have occurred for these test data points. There exists a trade-off between reducing lateral resolution and increasing the variance in test data points. The desired lateral resolution would want to be minimized without losing the discriminability between finger flexions (based on a maximum variance in test data points). This shows evidence that high resolution is not required to effectively discriminate between the finger flexions.

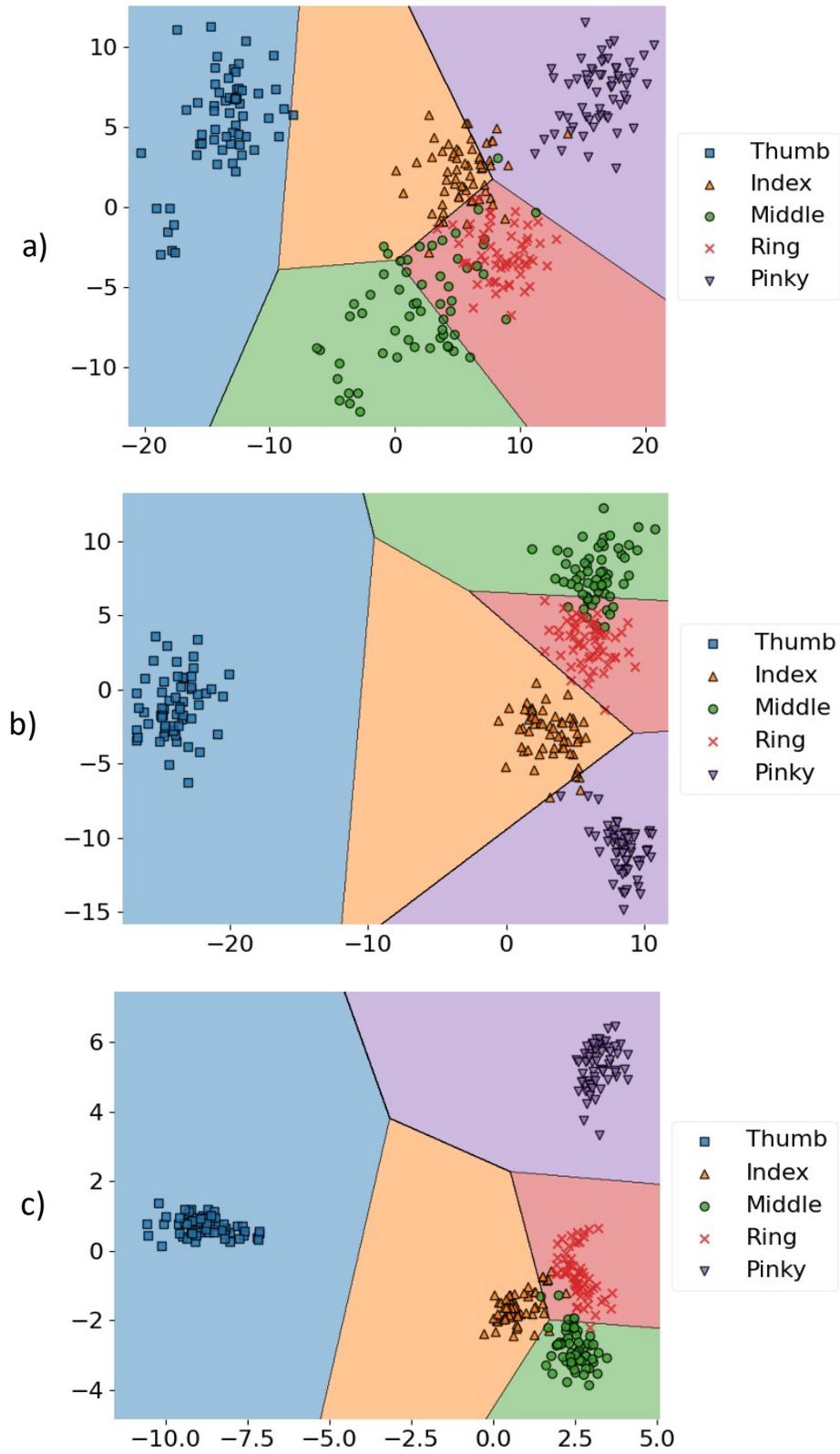


Figure 4-3 LDA projections of feature extracted from subject 3 right arm for lateral resolutions $N =$
a) 1, b) 12, and c) 127.

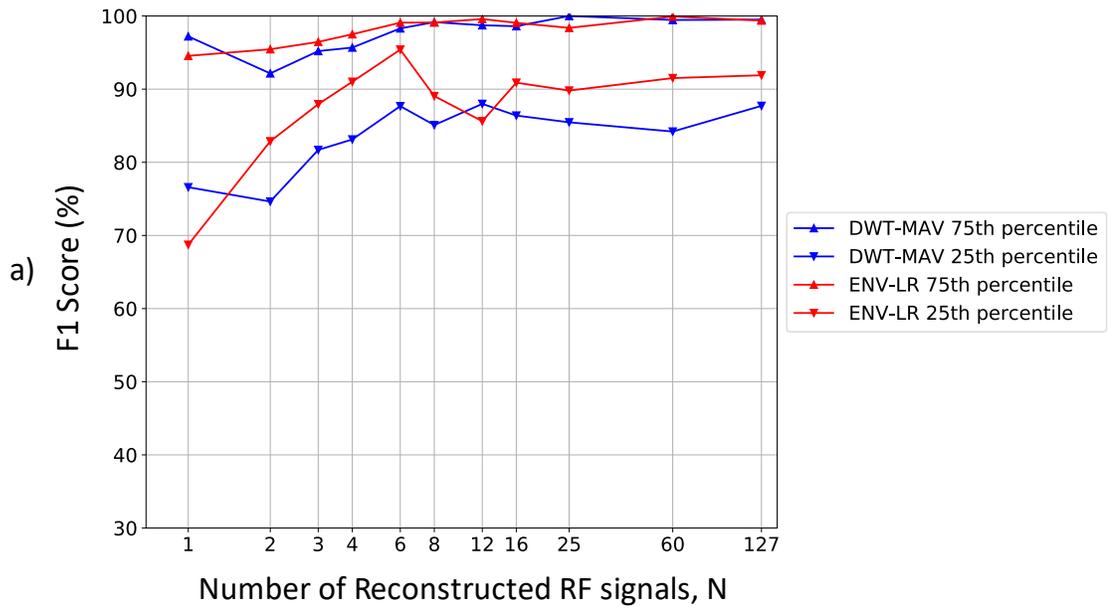
4.3.2 Individual Finger Flexion F1 Scores at Reduced Resolutions

Ten independent arm data experiments were conducted. Because the experiments do not follow a normal distribution when observed together (i.e. mean and standard deviation amongst the 10 experiments) we use the median and interquartile range (25th percentile and 75th percentile). Table 4-2 presents the 25th percentile, median, and 75th percentile values of the F1 scores for classifying each finger flexion. The results from Table 4-2 are plotted in Figure 4-4 to visualize the 25th and 75th percentiles of the F1 scores for each finger flexion. The F1 scores for each finger flexion remain lowest at $N = 1$ for both the DWT-MAV and ENV-LR methods. When comparing the F1 scores between the DWT-MAV and ENV-LR methods, they maintain similar classification performance. For both methods, the 75th percentile F1 scores are always above the 25th percentile F1 scores of the opposite method at each resolution. Therefore, I cannot conclude that the DWT-MAV method is superior to the ENV-LR method when comparing the accuracy of individual finger flexion predictions.

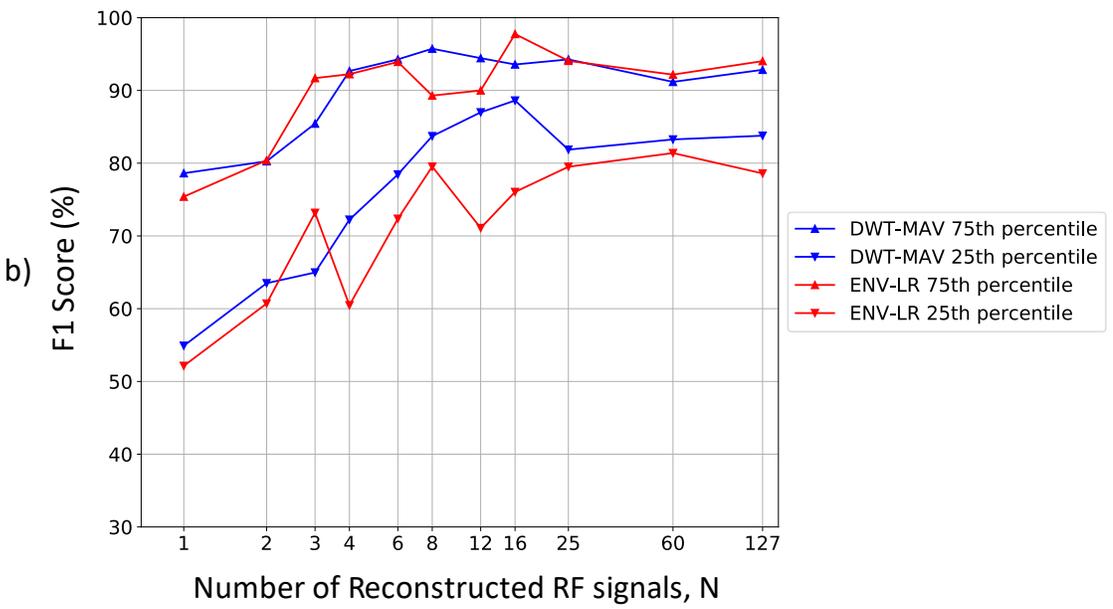
Table 4-2 F1 scores for classifying each finger flexion using 10 independent subject arm experiments with comparison of DWT-MAV and ENV-LR methods.

Resolution N	Method	F1 Score				
		25 th Percentile (%), Median (%), 75 th Percentile (%)				
		Thumb	Index	Middle	Ring	Pinky
1	DWT-MAV	76.6, 87.3, 97.2	54.9, 71.8, 78.6	48.3, 60.2, 75.2	50.3, 67.2, 76.9	65.1, 79.5, 87.0
	ENV-LR	68.7, 80.8, 94.6	52.1, 64.4, 75.4	59.0, 62.9, 70.6	56.3, 63.7, 69.5	73.4, 76.8, 82.3
2	DWT-MAV	74.6, 85.3, 92.2	63.5, 68.7, 80.3	59.0, 63.5, 73.8	67.3, 71.7, 77.7	73.0, 83.4, 88.7
	ENV-LR	82.9, 92.8, 95.5	60.7, 69.4, 80.4	52.1, 69.1, 78.8	66.0, 76.3, 79.2	77.9, 89.0, 95.2
3	DWT-MAV	81.7, 93.1, 95.2	65.0, 78.3, 85.4	57.5, 64.6, 79.9	59.6, 73.9, 85.2	79.9, 85.6, 95.1
	ENV-LR	87.9, 95.6, 96.5	73.2, 75.0, 91.7	64.8, 70.1, 88.6	71.5, 81.2, 89.5	86.1, 94.0, 96.1
4	DWT-MAV	83.1, 91.8, 95.7	72.2, 87.0, 92.7	65.9, 79.1, 91.7	66.1, 84.7, 89.4	82.7, 90.2, 92.3
	ENV-LR	91.0, 93.8, 97.5	60.5, 84.9, 92.2	62.9, 76.5, 88.5	75.2, 81.4, 85.5	79.8, 87.0, 95.0
6	DWT-MAV	87.7, 92.2, 98.3	78.5, 88.7, 94.3	71.8, 81.3, 92.2	71.7, 87.2, 90.7	79.6, 91.8, 95.7
	ENV-LR	95.4, 97.2, 99.1	72.4, 86.9, 93.9	67.8, 77.6, 93.5	73.9, 84.2, 89.3	81.2, 90.9, 98.2
8	DWT-MAV	85.1, 96.0, 99.2	83.7, 93.6, 95.7	68.8, 91.8, 95.8	79.7, 90.8, 93.0	80.2, 91.8, 98.9
	ENV-LR	89.1, 91.6, 99.1	79.5, 85.1, 89.3	67.2, 77.5, 89.9	67.3, 81.0, 84.8	79.9, 87.5, 96.5
12	DWT-MAV	88.0, 97.0, 98.7	87.0, 91.3, 94.4	71.1, 82.9, 92.6	75.7, 85.4, 93.0	84.2, 92.5, 99.4
	ENV-LR	85.6, 88.9, 99.6	71.1, 89.1, 90.0	58.9, 78.2, 87.1	69.0, 81.9, 93.2	81.8, 92.4, 97.3
16	DWT-MAV	86.4, 96.4, 98.6	88.6, 91.3, 93.6	78.6, 80.2, 90.2	82.5, 85.5, 89.5	79.4, 91.6, 98.4
	ENV-LR	90.9, 93.4, 99.1	76.0, 89.4, 97.7	67.1, 77.2, 89.4	64.7, 81.7, 87.6	85.3, 93.8, 99.0
25	DWT-MAV	85.5, 96.6, 100	81.9, 91.1, 94.3	77.2, 83.4, 90.0	73.1, 84.9, 91.1	78.4, 90.3, 99.3
	ENV-LR	89.8, 93.8, 98.4	79.5, 88.8, 94.1	75.5, 80.7, 91.5	75.9, 83.1, 93.9	79.3, 94.9, 98.2
60	DWT-MAV	84.2, 96.3, 99.5	83.3, 89.5, 91.2	77.6, 81.7, 90.5	79.9, 86.2, 90.5	82.4, 87.4, 99.3
	ENV-LR	91.5, 96.0, 99.9	81.4, 91.2, 92.2	74.0, 81.3, 89.4	79.7, 85.0, 92.7	79.1, 94.5, 96.7
127	DWT-MAV	87.7, 96.4, 99.5	83.8, 91.1, 92.8	78.5, 84.8, 92.6	82.9, 86.1, 92.1	82.6, 86.0, 99.5
	ENV-LR	91.9, 93.4, 99.4	78.6, 90.4, 94.0	74.3, 81.9, 86.4	79.7, 85.7, 92.9	78.5, 93.3, 96.6

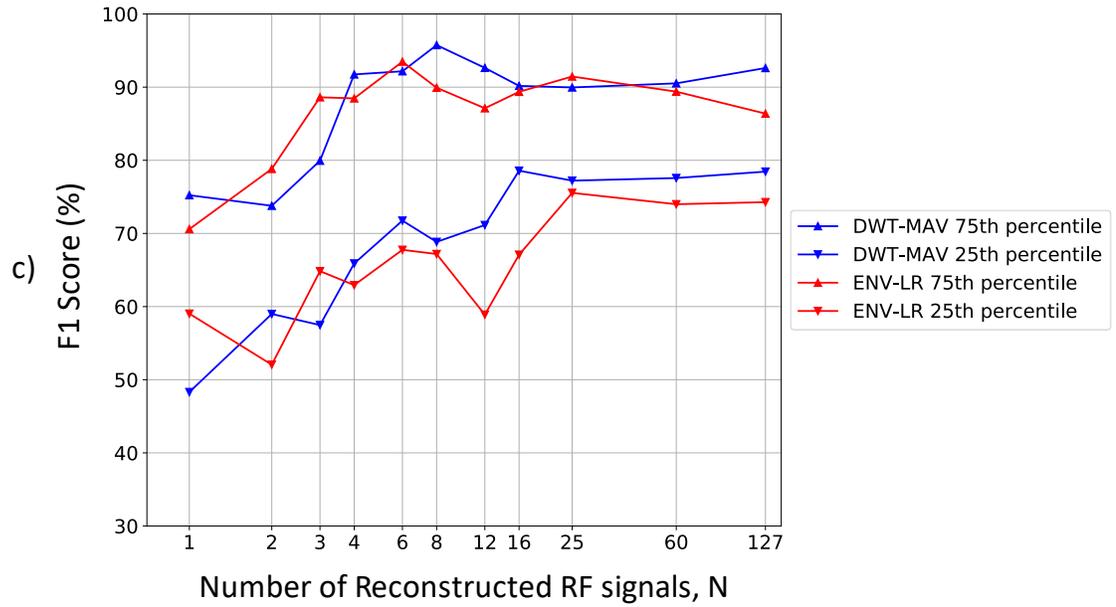
Thumb Finger Flexion



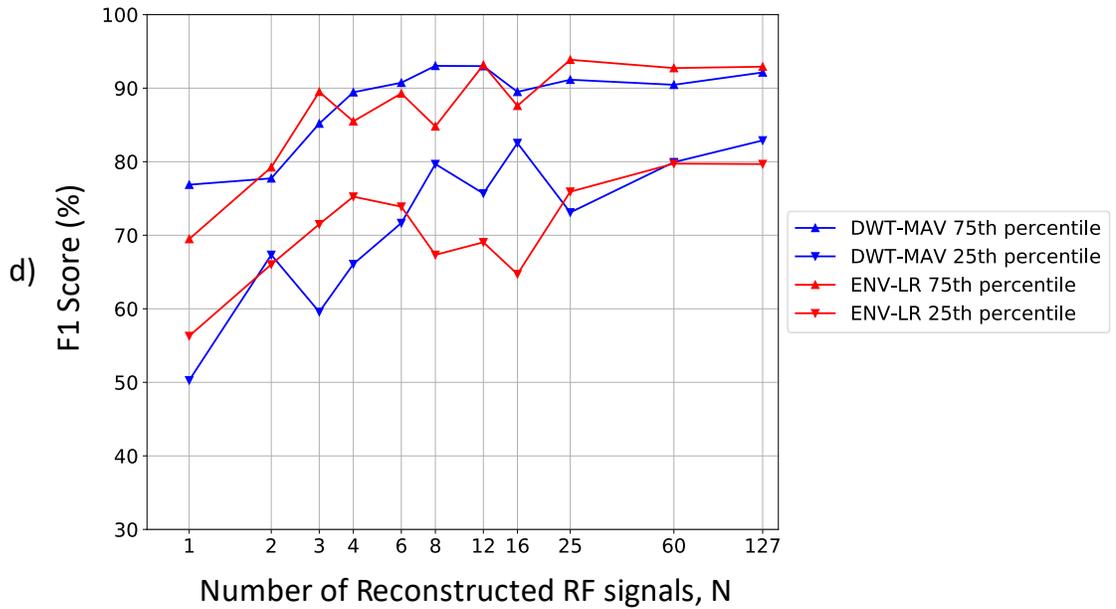
Index Finger Flexion



Middle Finger Flexion



Ring Finger Flexion



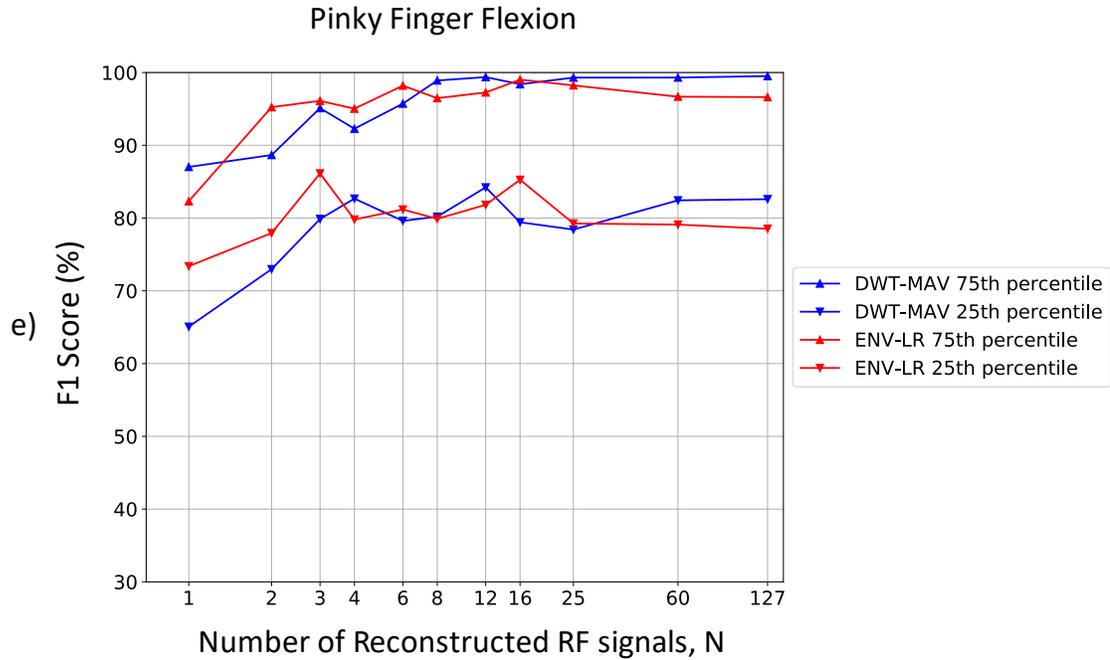


Figure 4-4 Comparison between the spatial feature extraction methods DWT-MAV and ENV-LR for F1 Scores of a) thumb, b) index, c) middle, d) ring, and e) pinky finger flexions.

4.3.3 Classification Accuracy vs Resolution

Figure 4-5 a) shows the results for the classification accuracy of all the finger flexions for the DWT-MAV method as a function of N . It is noted that N indicates the lateral spatial resolution of ultrasound data (number of reconstructed ultrasound RF signals). For both methods (DWT-MAV and ENV-LR), the median value of the classification accuracy decreases to the lowest point for $N = 1$. When $4 \leq N \leq 127$, the median classification accuracy fluctuates between 86% and 92% for the DWT-MAV method and between 85% and 90% for the ENV-LR method. The results of this reconstructed ultrasound data suggest that a multichannel 4-UT system ($N = 4$) with the lateral resolution of 10 mm would have the potential to achieve the same classification accuracy as a linear array probe with 330 μm space between each UT lateral resolution ($N = 127$).

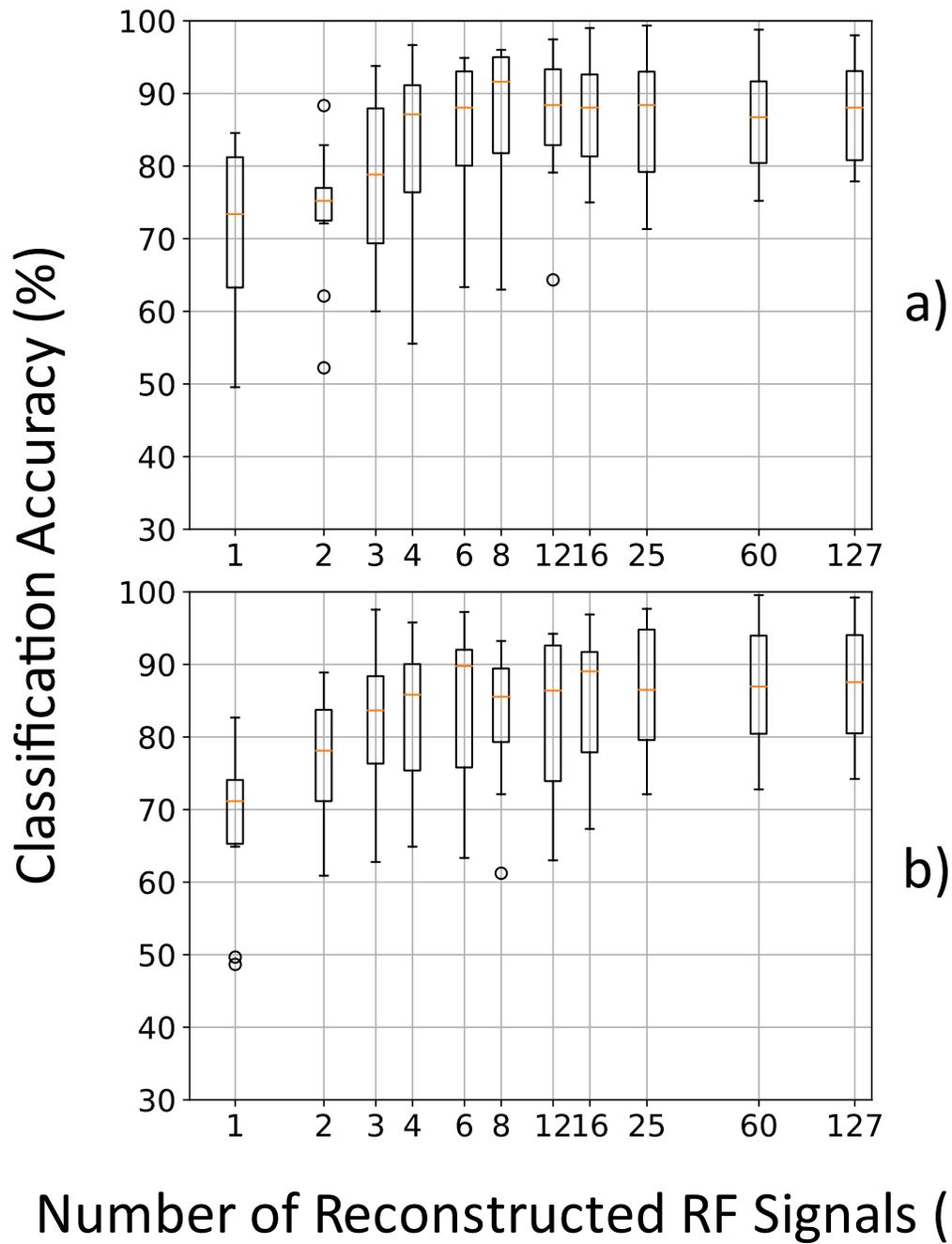


Figure 4-5 Classification accuracy for varying number of reconstructed RF signals: a) DWT-MAV method, and b) ENV-LR method with 10 segments.

4.3.4 Computation Time vs Resolution

Table 4-3 shows the mean computation time with the standard deviation to perform a single prediction on the 100 ms time intervals of data. The computation time required to make a single prediction is presented with a mean and standard deviation. These experiments were conducted using an Intel core i7-6700 processor (4 cores) with a clock speed of 3.4 GHz. For the ENV-LR method, 10 segments resulted to have the highest classification accuracy for the fastest computation time. The DWT-MAV method is more than 10 times faster than the ENV-LR method with 10 segments. In addition, for the ENV-LR method, the computation time increases drastically as N increases. For the ENV-LR method, 10 segments were experimentally chosen to produce the highest classification accuracy with fast computation speeds. The table also shows when having five times more segments at 50 segments, the processing time triples. Since the data acquisition is every 100 ms, the total computation time results in 100 ms plus the processing time. In Table 4-3, the processing times that make the total computation time do not meet the optimal delay criteria of 100–125 ms is in bold and the processing times that make the total computation time exceed the accepted 300 ms time are highlighted.

Table 4-3 Average computation time to execute a single prediction using different spatial feature extraction methods.

Resolution N	Computation Time (ms)		
	DWT-MAV	ENV-LR (10 segments)	ENV-LR (50 segments)
1	1.63 ± 0.33	6.48 ± 0.68	17.53 ± 1.10
2	1.78 ± 0.51	11.70 ± 1.60	33.90 ± 1.40
3	2.14 ± 0.55	17.07 ± 1.09	50.20 ± 1.40
4	2.57 ± 0.47	22.53 ± 1.31	66.66 ± 1.50
6	3.43 ± 0.58	33.41 ± 1.46	99.41 ± 1.61
8	4.31 ± 0.53	44.40 ± 1.49	133.46 ± 2.22
12	5.97 ± 0.55	66.16 ± 1.54	198.73 ± 1.80
16	7.68 ± 0.58	88.34 ± 1.57	264.82 ± 2.68
25	11.41 ± 0.52	137.48 ± 1.74	412.77 ± 2.79
60	26.33 ± 0.41	328.20 ± 2.30	991.00 ± 6.04
127	54.69 ± 0.58	694.72 ± 4.29	2097.18 ± 11.85

4.3.5 Classification Accuracy vs Computation Time

The results of the classification accuracy vs lateral resolution from Figure 4-5 is reproduced in Figure 4-6 a) to compare the DWT-MAV and ENV-LR methods. Based on the results in Figure 4-6 a), either method cannot be confirmed to have a significantly higher classification accuracy. In Figure 4-6 b), the results of Table 4-3 are visualized to show that the DWT-MAV method is significantly more computationally faster than the ENV-LR method. These computation times for the given resolution N are then shown with the corresponding classification accuracies in Figure 4-6 c). Comparing classification accuracy vs computation time (both of which are a function of resolution) demonstrates how the DWT-MAV method can achieve the classification accuracy with lower computation times. The conclusion is that while the classification accuracies are comparable for both the methods, the DWT-MAV method significantly outperforms ENV-LR in execution speed.

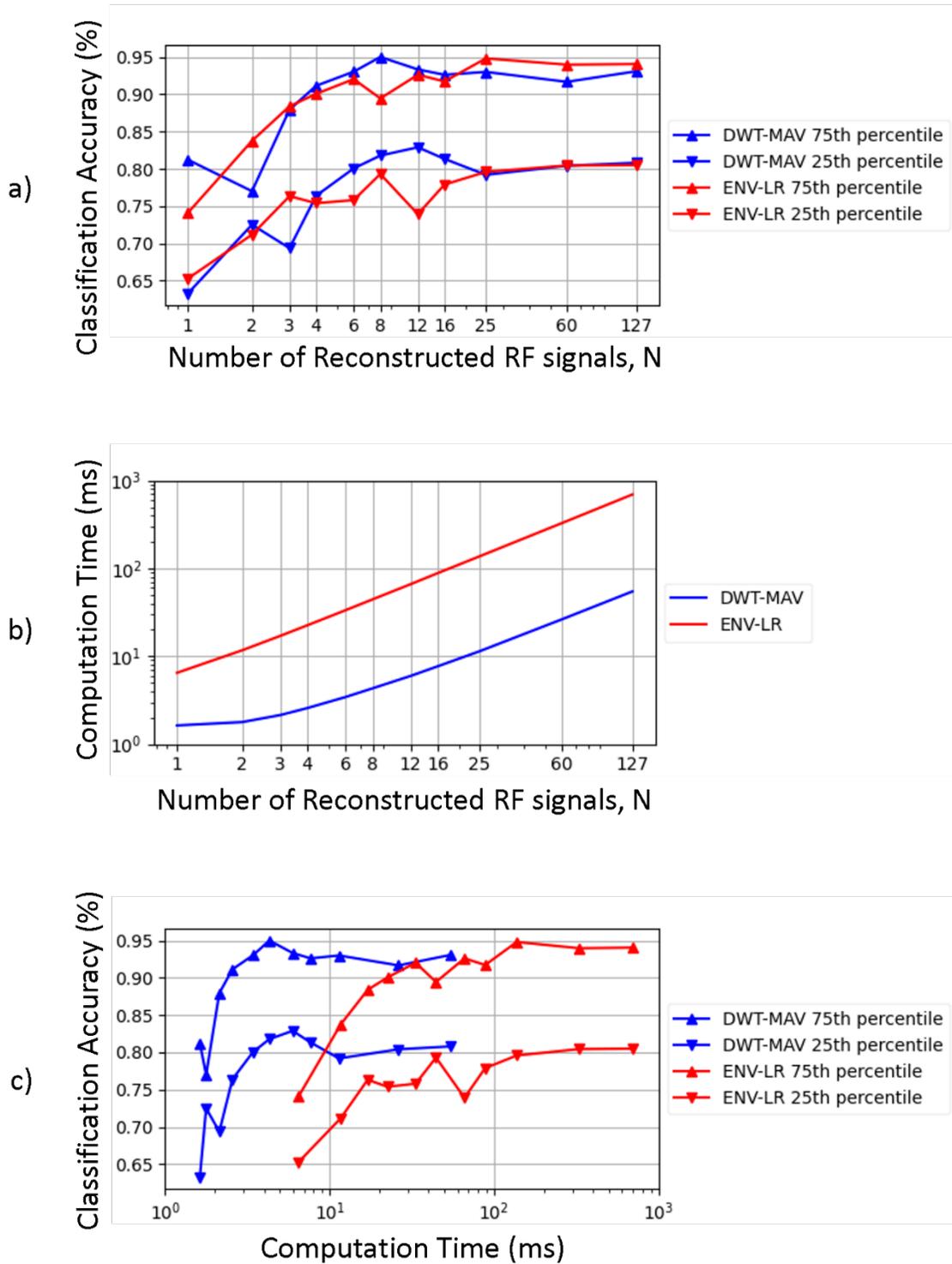


Figure 4-6 Comparison between the spatial feature extraction methods DWT-MAV and ENV-LR for a) resolution versus classification accuracy, b) resolution versus computation time, and c) computation time versus classification accuracy.

4.3.6 Discussion

The DWT-MAV method yielded similar classification accuracies and F1 scores to that of the ENV-LR method throughout all lateral resolutions. However, the DWT-MAV method proves to be more computationally efficient. The DWT-MAV method is at least 10 times faster than the ENV-LR method of 10 segments and 30 times faster than the ENV-LR method of 50 segments. It required 2.57 ms at $N = 4$ by the DWT-MAV method. But for the ENV-LR method with 10 segments it required 22.53 ms and 66.66 ms with 50 segments. The experiments used Intel core i7-6700 3.4 GHz processor. This means that even when using powerful hardware (in comparison to what would be used in wearable systems), the ENV-LR method struggles to have a total data acquisition and processing time within 100–125 ms. Therefore, the PCP should use the DWT-MAV spatial feature extraction method over ENV-LR. This faster PCP could be used in a real-time and wearable system requiring lower computation power hardware. By emulating for 4 to 8 UTs (i.e. $4 \leq N \leq 8$), the required computation time to make a prediction on the 100 ms groupings of time is 4 to 8 ms using the DWT-MAV method and 22 to 44 ms using the ENV-LR method (10 segments) as shown in Table 4-3. When using hardware with lower processing power, the DWT-MAV method can be expected to greatly outperform the ENV-LR method in speed. The ENV-LR method might not be able to perform computations faster than 25 ms if the number of depth segments is chosen too high.

The results also indicate that high resolution is not required to achieve high accuracy for individual finger flexion classification. This verifies the results conducted in the preliminary study. The results indicate that a multiple single element UT system with 4 to 8 UTs could be used as to achieve the same classification performance (with faster

computation speeds) to that of a bulky conventional clinical ultrasound linear array probe. The design of a multiple single element UT system would be able to use a low-cost wearable system. Thus, the lower classification performance among the index, middle and ring fingers could be improved further using the multiple single element UT system. The results from this detailed study suggest that that a multiple single element UT system can be used in HMI applications that would perform optimally with 100–125 ms between successive prediction time intervals [2]. Further research using wearable multiple single element UTs constructed out of PVDF material is the next step [67, 68]. This is performed as a preliminary study in Chapter 5.

Chapter 5: Classification of Individual Finger Flexions Using Wearable Ultrasonic Sensors

5.1 Introduction

This chapter is a preliminary study involving wearable ultrasonic sensors (WUS) to classify individual finger flexions. In this chapter, an ergonomic WUS system was designed to classify individual finger flexions for future use in HMI applications. The PCP described in Chapter 3 is used on ultrasound RF signals acquired by the WUS system to classify the individual finger flexions. Based on the superior results of the DWT-MAV spatial feature extraction method from Chapter 4, this method was chosen over the ENV-LR method. The goal of this chapter is to verify the accuracy of the PCP using a low lateral resolution multiple WUS system. The results of this preliminary study will give motivation for future studies when designing a portable WUS system for HMI applications.

5.2 Materials and Methods

Due to the lightweight, thin, and flexible design of the WUSs, these sensors can noninvasively monitor muscle movement without restriction [67]. The study conducted by Yan *et al.* proved that these WUSs could perform discrete hand gesture recognition for wearable HMI applications [14]. The purpose of my study is to determine if these WUSs can perform individual finger flexion classification. In this section, the design and setup of the WUS system is described. Then the procedure to record individual finger flexions from the WUS system is discussed. Then finally the modifications made to the PCP from Chapter 3 are described in this chapter to operate with the WUS system.

5.2.1 Wearable Ultrasonic Sensor System and Experimental Setup

5.2.1.1 Sensor Design and Experimental Configuration

This study was approved by Carleton University's Research Ethics Board. The experiment was conducted with one healthy human subject (male, age 23). This feasibility study was conducted to verify the proof of concept of the WUSs and method for classifying individual finger flexions. For this study, I constructed three WUSs using 110- μm thick PVDF piezoelectric polymeric film coated on both sides with silver ink. The design and material properties of these sensors is based on the study conducted to monitor skeletal muscles by AlMohimeed *et al.* [67]. I constructed the electrode size to have an active ultrasonic area of 20 mm \times 20 mm as shown in Figure 5-1 a). I decided to use an active area of 20 mm \times 20 mm such that they could be placed in parallel across the anterior side of my forearm as shown in Figure 5-2. The total thickness of the WUS including a protection and insulation layers was 280 μm . Lined paper was used to provide an air pocket as the backing layer shown in Figure 5-1 b). I soldered each WUS to a coaxial cable shown in Figure 5-1 c) to aid in noise reduction. The echo of the ultrasound center frequency was measured to be 5.1 MHz for the RF signals acquired with a human subject.

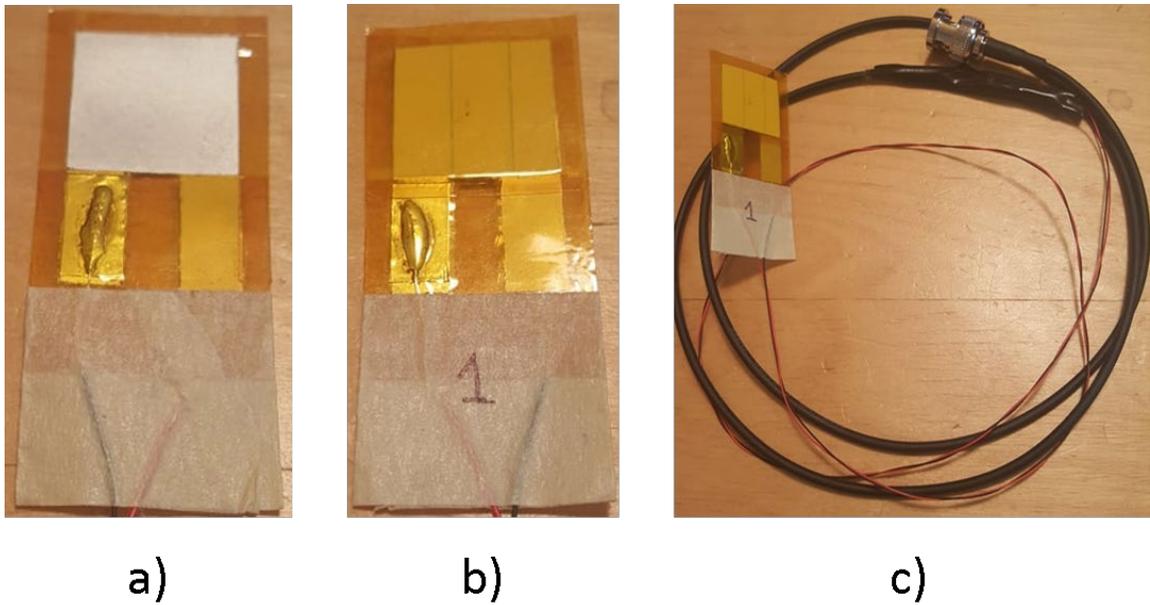


Figure 5-1 One Wearable Ultrasonic Sensor I constructed for this thesis. a) front side with active area, b) backside with lined paper backing, and c) full transducer soldered to a coaxial cable.



Figure 5-2 Three WUSs placed on my forearm 5 cm away from my wrist using ultrasound couplant gel. The WUSs were secured by a 5 cm width self-adhering elastic bandage.

As seen in Figure 5-2, the three WUSs were attached on my forearm using ultrasound couplant gel. The ultrasound couplant gel is used as a matching layer between the WUS and my forearm. Each WUS was 5 cm away from the wrist and secured using self-adhering elastic bandage. In the previous linear array probe experiments, the flexor muscles were monitored at this location as discussed in Chapter 3. Therefore, the placement of these WUSs were chosen to replicate the location of the experiments using the linear array probe.

5.2.1.2 Ultrasound Data Acquisition Setup

The hardware was setup using the configuration shown in Figure 5-3. Each of the WUSs were controlled by an ultrasonic pulsar/receiver (P/R) (two identical models: DPR300, JSR Ultrasonics, Pittsford, NY and one model: 5900PR, Olympus Panametrics NDT, Waltham, MA) to transmit and receive the ultrasonic RF signals in pulse-echo mode. The received RF signals were acquired in M-mode and digitized by a multiple-channel A/D converter (model ATS 9440, Alazartech, Montreal, QC) controlled by a computer. The three P/Rs were driven by the common trigger signals to synchronize the ultrasonic data acquired.

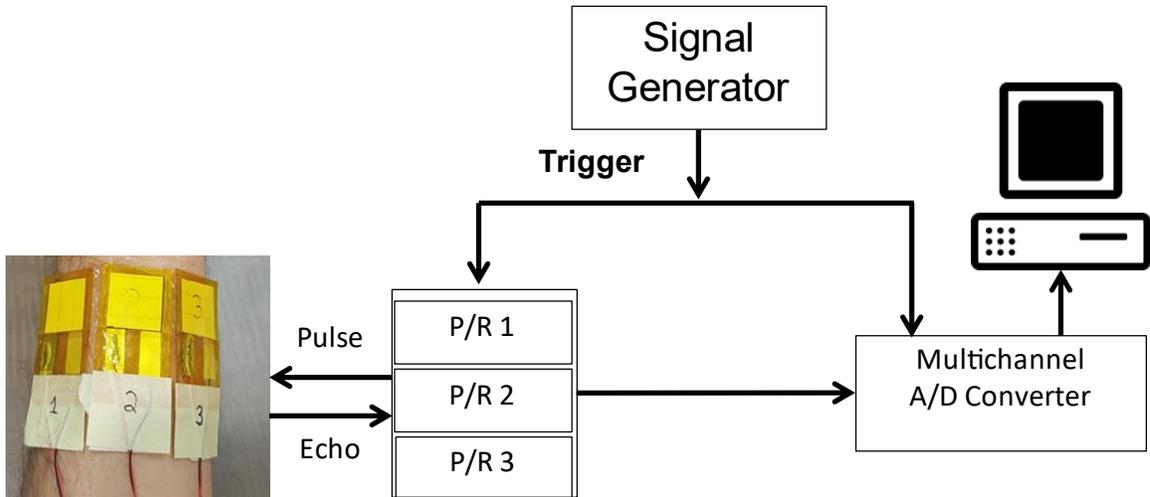


Figure 5-3 An overview of the designed WUS system. Ultrasound acquisition setup for three WUSs.

5.2.2 Data Acquisition Procedure

5.2.2.1 Data Format of Wearable Ultrasonic Sensor

The received RF signals were sampled at the maximum sampling rate of 125 MHz (every 8 ns), offering the highest depth resolution provided by the system. Thus, assuming an ultrasound speed of 1540 m/s for soft tissues, the depth resolution was 6.16 $\mu\text{m}/\text{sample}$. A total of 2432 depth sample points (full depth range of 14.98 mm) was used when recording for the individual finger flexions. The M-mode frame rate (pulse repetition frequency) was set to 1000 frames per second. This frequency was chosen because it was experimentally tested as the maximum pulse repetition frequency that the WUS system could handle. A maximum pulse repetition frequency was desired to obtain the highest time (frame rate) resolution for the recordings. For each acquired RF signal, the DC component was removed. Sample points outside of the region of interest (ROI) were manually excluded to include only the regions of tissue motion activity as shown in Figure 5-4. This ROI was chosen across all the WUS recordings to remove the large pulse in the shallow region. While for the deeper regions, no motion activity was observed to be present and cropped out. For future studies, the ROI will need to be reselected through observation. The first recommendation is to crop out the large pulse since the large amplitude will amplify any effects of noise appearing in the extracted spatial features. The second recommendation is to trim the depth region with no motion activity occurring to save on computation time.

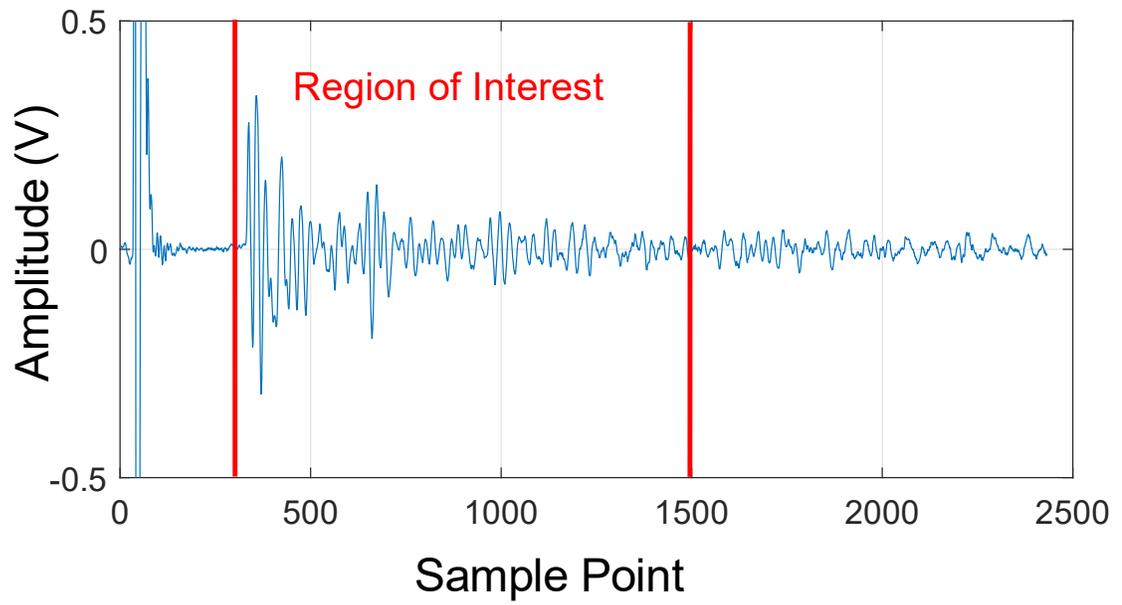
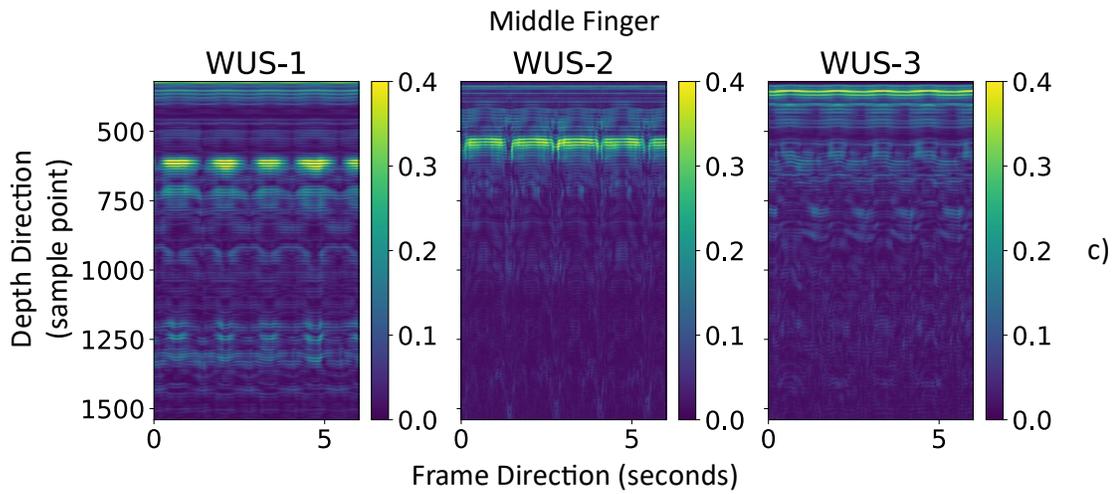
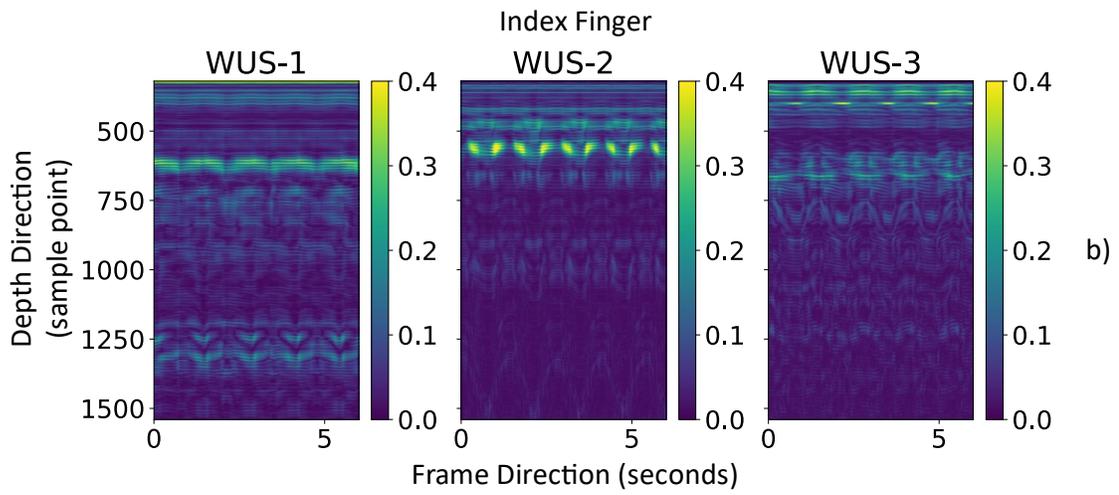
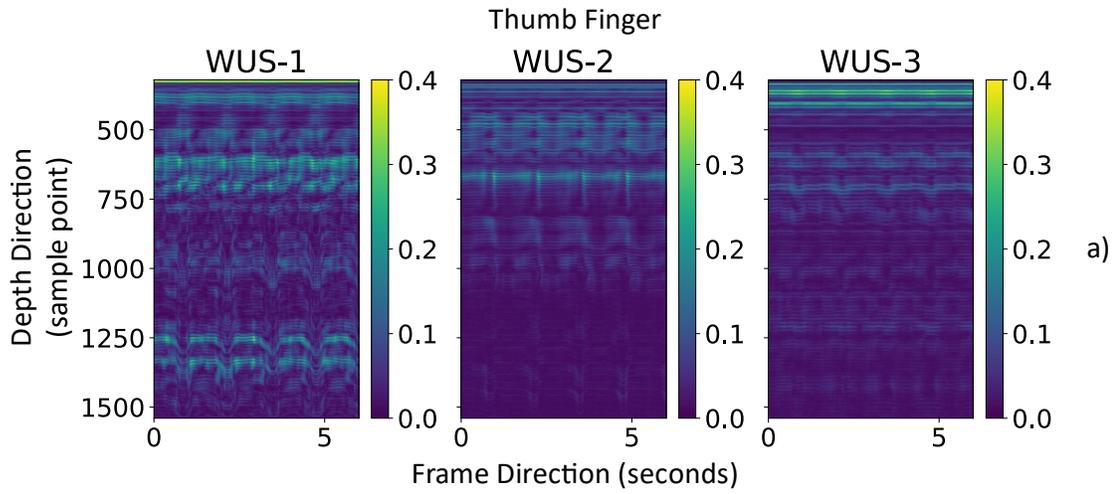


Figure 5-4 A single pulse echo acquired RF signal with a region of interest ranging between sample points 320 to 1520.

5.2.2.2 Finger Flexion Recording Procedure

During the data acquisition, individual finger flexions of the thumb, index, middle, ring and pinky were recorded. The tissue motions beneath the WUSs were caused by contraction/relaxation of the flexor muscles (digitorum profundus, digitorum superficialis and/or pollicis longus) influenced the acquired RF signals to fluctuate. These fluctuations could be used to distinguish between the individual finger flexions. Each finger flexion recording consisted of repeatedly flexing the same finger. By only flexing one finger in each recording, it made it easier to determine the finger flexion type when collecting for labeled training and testing data. For each individual finger flexion recording, the selected finger continuously performed a complete 180° flexion and then relaxation at a rate of 45 cycles per minute over a 6 second period (totaling to approximately 4 or 5 flexions of each finger in a single recording). For each of the five fingers, 10 trial recordings were repeated, totaling to 5 minutes of recording (50 finger flexion trial recordings, with each trial lasting 6 seconds with approximately five repeated flexions). An example of one trial recording for each finger flexion is shown in Figure 5-5. The trial recordings are displayed as M-mode images of the acquired ultrasonic RF signals. The envelope of the acquired RF signals was obtained by the Hilbert transform and presented in an M-mode image for each WUS. The tissue motions can be observed as the repetitive muscle contraction/relaxation motion occurring every 1.5 seconds in the M-mode images. Where one complete finger flexion will occur approximately every 1.5 seconds across the 6 second recording. One example of muscle activity that is visualized in Figure 5-5 can be seen around the 1250 sample point of WUS-1. Around this region the flexor muscle moves in a different pattern throughout each finger flexion.



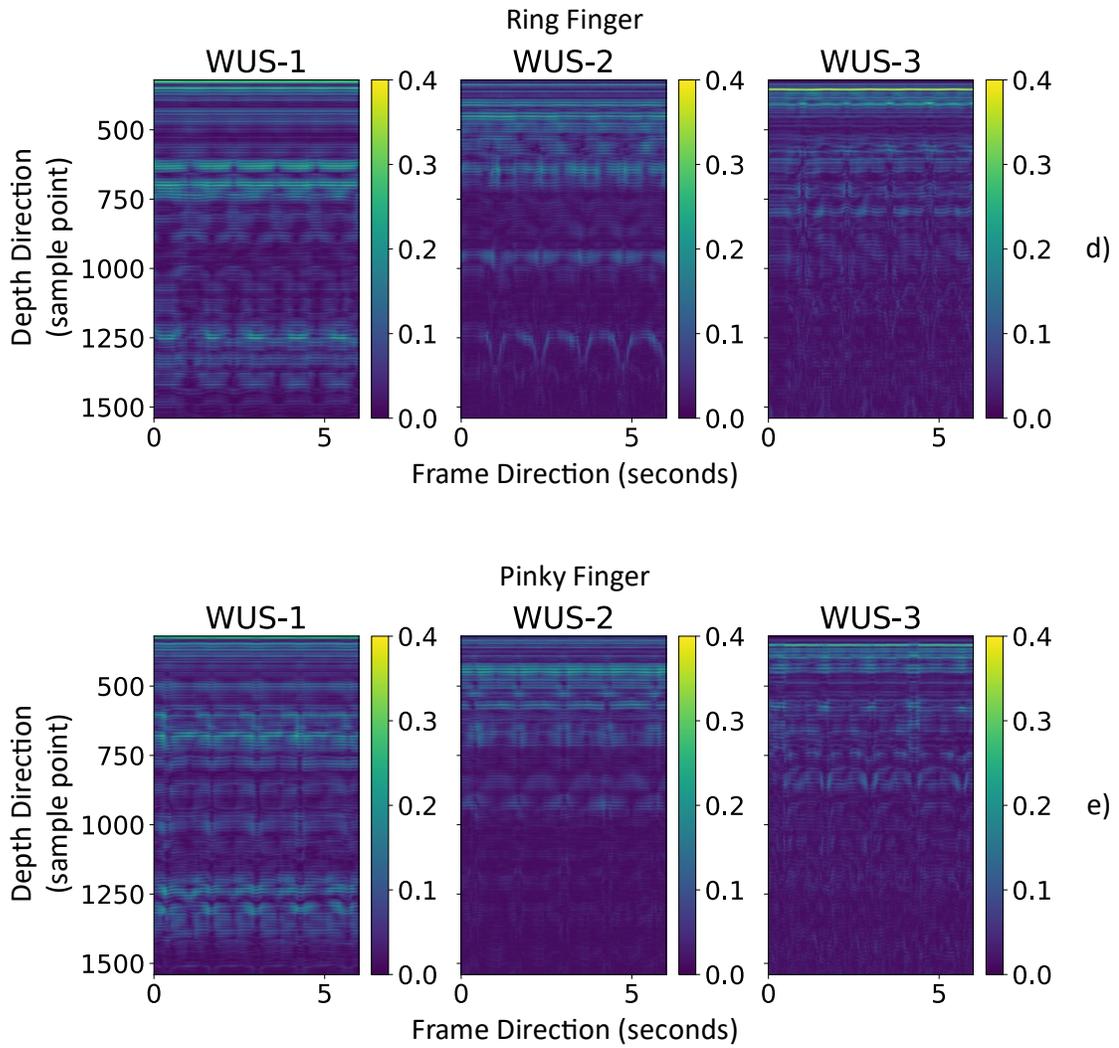


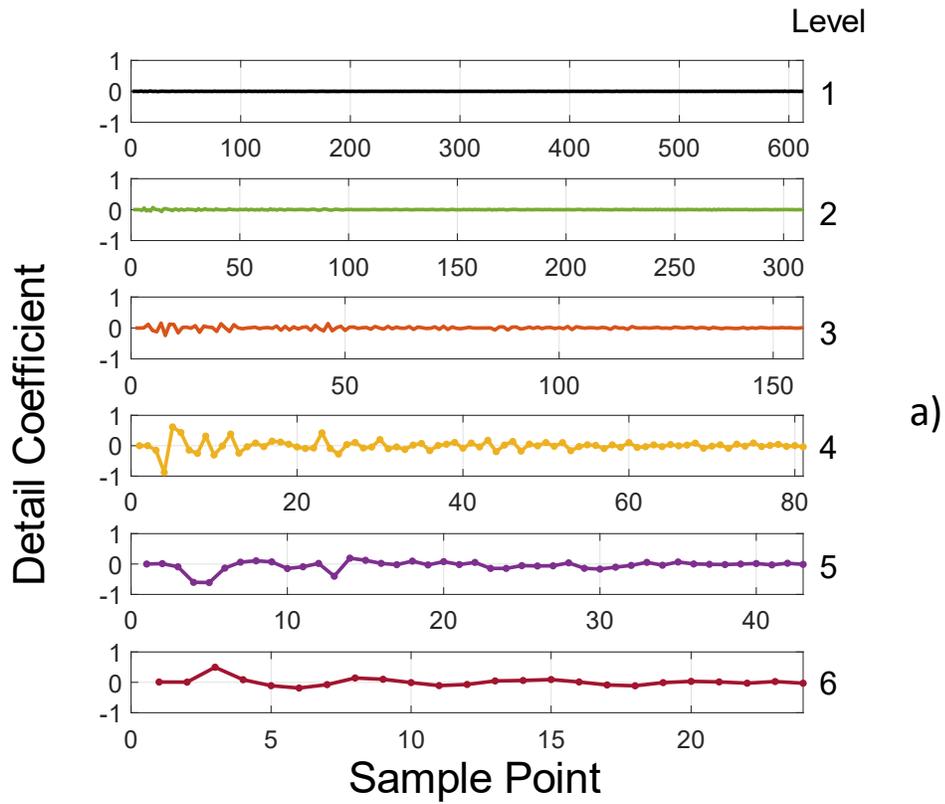
Figure 5-5 M-mode images of each WUS during acquisition of five repeated motions in a 6 second trial recording of a) thumb, b) index, c) middle, d) ring, and e) pinky finger flexion procedure.

5.2.3 Modifications Made to Pattern Classification Pipeline

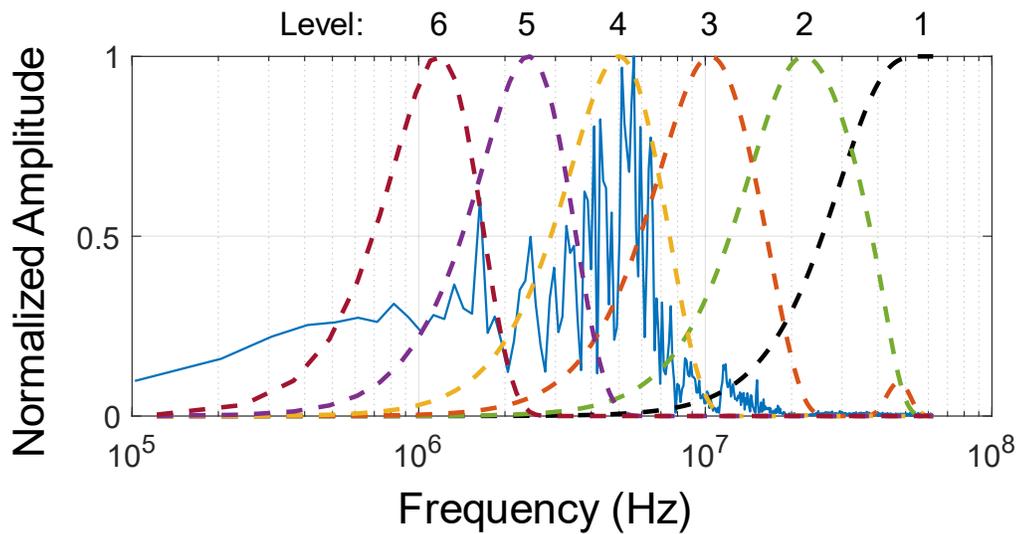
For this study, the pattern classification pipeline as described in Chapter 3 has been modified based on the results of Chapter 4 and WUS system in this chapter. For this machine learning problem, a total of 300,000 labeled frames of data with 426 spatial and temporal features were used for classification. The detail coefficients from levels 4, 5, and 6 have been chosen based on observing the frequency spectrum of the WUS system. These levels were used instead of 1, 2, 3 because the depth sampling rate for the WUS system is 125 MHz in comparison to 33.3 MHz for the clinical linear array probe system. Each prediction outcome is represented across a 100 ms time interval. This means that throughout a total of 5 minutes of finger flexion recordings, a prediction was made every 100 ms in time. In other words, 3,000 finger predictions were made in this preliminary study for each independent experiment. Then a multilayered perceptron (MLP) and K Nearest Neighbours (KNN) classification methods have been included alongside the LDA. The labeled sample data with their corresponding features were partitioned into training and testing data. The testing data used labeled samples belonging from one of the 10 recorded trials while the training data used samples belonging to other 9 trials (10-fold cross validation). This was repeated 10 times such that each trial became the testing data.

5.2.3.1 Frequency Range for Discrete Wavelet Transform

I applied the DWT to the RF signals in the ROI along the depth to compute the first 6 DWT detail coefficients as shown in Figure 5-6 a). The RF signal in Figure 5-4 was used to produce Figure 5-6. The analysis coefficients A_l and detail coefficients D_l for levels $l = 1, 2, 3, 4, 5, 6$ were computed for the DWT-MAV method.



a)



b)

Figure 5-6 a) DWT detail coefficients computed for the first 6 levels and b) frequency spectrum (solid line) of the RF signal in the ROI and the bandwidth frequency spectrums for the first 6 levels of the DWT (dashed lines).

In Figure 5-6 b), by observing the frequency range of the acquired RF signals (solid line) and the frequency spectrum of each DWT level (dashed lines), the levels from 1 to 3 are discarded. These levels were discarded due to lack of frequency content within these ranges. This eliminates redundancy in both the number of computed depth samples and number of extracted features. This is the same reason that levels higher than 6 were not computed.

5.2.3.2 Machine Learning Classifiers

Three classification methods were used to evaluate the performance of the WUS system and PCP. If the results are successful across all three classification methods, it would suggest that the features extracted from the WUS system are ‘good’ at discriminating between the finger type. As opposed to the classification method is the reason why the WUS system and PCP is ‘good’ at discriminating between the finger type.

The classifiers used were linear discriminant analysis (LDA), multilayer perceptron (MLP) and K nearest neighbours (KNN). The architecture of the MLP (one hidden layer of 15 nodes), was chosen from the ultrasound-based study for discrete gesture classification which gave the highest classification accuracy [46]. This architecture is also quick to train and execute. The KNN classifier was chosen based on the study by Sikdar *et al.* in which they were able to quickly classify individual finger flexions with high accuracy [9]. The performance of the WUS system and feature extraction methods is evaluated by using these three machine learning classifiers. Achieving high classification across all machine learning classifiers gives an indication that the extracted features are useful in discriminating between finger flexions.

5.2.3.3 Procedure for Continuous Classification

To follow the optimal controller delay for prosthetic devices, a prediction was made every 100 ms for all finger flexion recordings [2]. With the M-mode frame rate set to 1000 frames per second, this allows for frame averaging and improving the classification accuracy. In comparison to the linear array probe experiments fixed with a frame rate of 30 frames per second, this allowed for more room to optimize the 100 ms time grouping procedure. Prior to feature extraction, the number of M-mode frames (τ) were averaged into a single data sample to reduce the noise variance occurring along the frame direction. After extracting features for the averaged frames, the classifier assigns a probability to each finger (thumb, index, middle, ring, pinky). A consecutive number of probabilities (σ) were then averaged to make the prediction of the 100 ms interval of time. Then, the maximum probability across the five fingers would be used on final finger decision.

For example, when $\tau = 10$, 10 frames are averaged into a single frame before being input into the PCP (before the spatial feature extraction step). With an acquisition frame rate of 1000 frames per second, each averaged frame would represent a 10 ms time grouping. Then the PCP assigns a probability for each of the five fingers for the averaged frame. These probabilities are calculated using scikit learn [66] for each classifier to transform the output into a probability for each finger instead of a decision. LDA computes based on the softmax of the linear discriminant function's decisions corresponding to each finger, MLP computes based on the output of the network for each corresponding finger, and KNN computes based on the estimated densities amongst the 5 nearest neighbours of each sample related to each corresponding finger. The probability of each finger $Prob^{10ms}$, is represented as the probability of the finger over a 10 ms time

interval. When $\sigma = 10$, a final probability $Prob^{100ms}$ is computed by averaging ten sequential $Prob^{10ms}$ to represent the probability for each finger over a 100 ms time interval. The ω variable is represented as one of the fingers: thumb, index, middle, ring or pinky. $Prob(\omega)$ represents the probability that the PCP believes it is that finger. The finger ω with the highest probability was chosen as the final decision $\omega_{decision}$ as described by (6.1) and (6.2).

$$Prob^{100ms}(\omega) = \frac{1}{10} \sum_{n=1}^{10} Prob_n^{10ms}(\omega) \quad (6.1)$$

$$\omega_{decision} = \omega \operatorname{argmax} \left(Prob^{100ms}(\omega) \right) \quad (6.2)$$

5.3 Results

Every 100 ms in time throughout the finger flexion recordings a prediction was made. The classification accuracies and F1 scores with a mean and standard deviation following 10-fold cross validation (based on trial number). The confusion matrices in Figure 5-7 through Figure 5-14 and classification metrics in Table 5-1 through 5-8 shows the results of the LDA, MLP and KNN classifiers with different values of τ and σ .

For the LDA classifier, $\tau = 1$ and $\sigma = 100$ produced the highest classification accuracy at $94.60 \pm 7.91\%$. For the MLP classifier, $\tau = 1$ and $\sigma = 100$ produced the highest classification accuracy at $94.47 \pm 8.89\%$. For the KNN classifier, $\tau = 4$ and $\sigma = 25$ produced the highest classification accuracy at $97.53 \pm 2.89\%$.

The experimental parameters that gave the greatest individual finger flexion F1 scores was using the KNN classifier with $\tau = 4$ and $\sigma = 25$. This resulted in the lowest F1

score to be the index finger at $94.73 \pm 1.67\%$. The mean for each F1 score ranged between $94.73 - 98.11\%$, with each mean having a standard deviation of less than 3%.

The parameters that gave the greatest individual finger flexion performance across all classification methods was with $\tau = 10$ and $\sigma = 10$. This resulted in the lowest F1 score (given these parameters) to be the ring finger of the LDA classifier at $82.67 \pm 8.70\%$. The mean for each F1 score ranged between $82.67 - 100.00\%$ across the three classifiers, with each mean having a standard deviation of less than 9%.

Therefore, $\tau = 1$ and $\sigma = 100$ gave the best results for the LDA and MLP classifiers. $\tau = 4$ and $\sigma = 25$ gave the best result for the KNN classifier. Finally, the optimum parameters were $\tau = 10$ and $\sigma = 10$, which gave the best result when taking all three classifiers into account. This suggests that the features extracted from the WUS system using $\tau = 10$ and $\sigma = 10$, has the best discriminability across all three classifiers. This shows that averaging every 10 frames yields the best features. Then when τ increases above 10, (while σ decreases below 10 to maintain the 100 ms time interval), the performance across all three classifiers decreases.

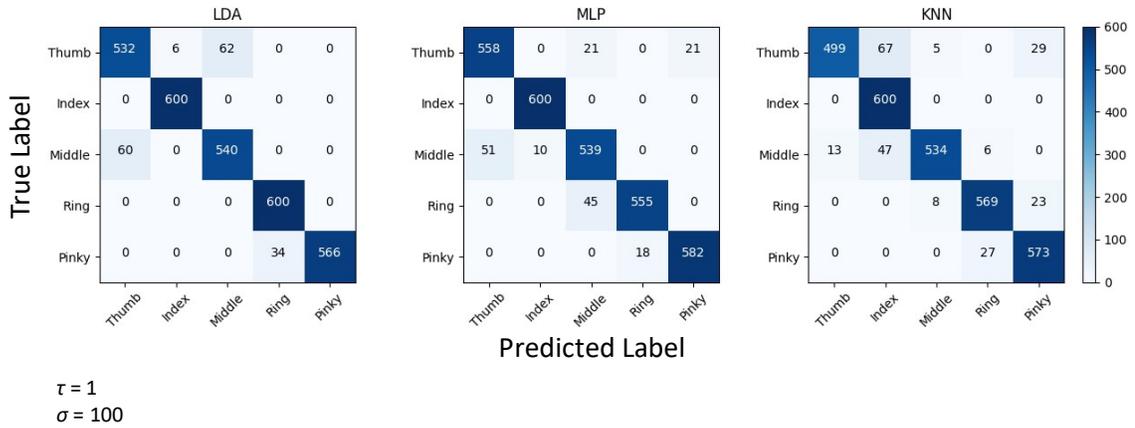


Figure 5-7 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 1$ and $\sigma = 100$ when predicting for 100 ms sequential time intervals.

Table 5-1 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 1$ and $\sigma = 100$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	94.60 ± 7.91	86.04 ± 8.04	99.70 ± 0.25	88.96 ± 4.91	93.09 ± 5.42	90.77 ± 9.76
MLP	94.47 ± 8.89	82.51 ± 9.46	97.66 ± 2.01	71.08 ± 25.49	90.12 ± 4.10	92.01 ± 4.93
KNN	92.50 ± 13.96	69.42 ± 25.68	81.16 ± 11.26	77.03 ± 26.78	90.71 ± 3.55	86.35 ± 7.05

Lowest F1 Score: KNN, thumb finger, 69.42 ± 25.68%

Highest F1 Score: LDA, index finger, 99.70 ± 0.25%

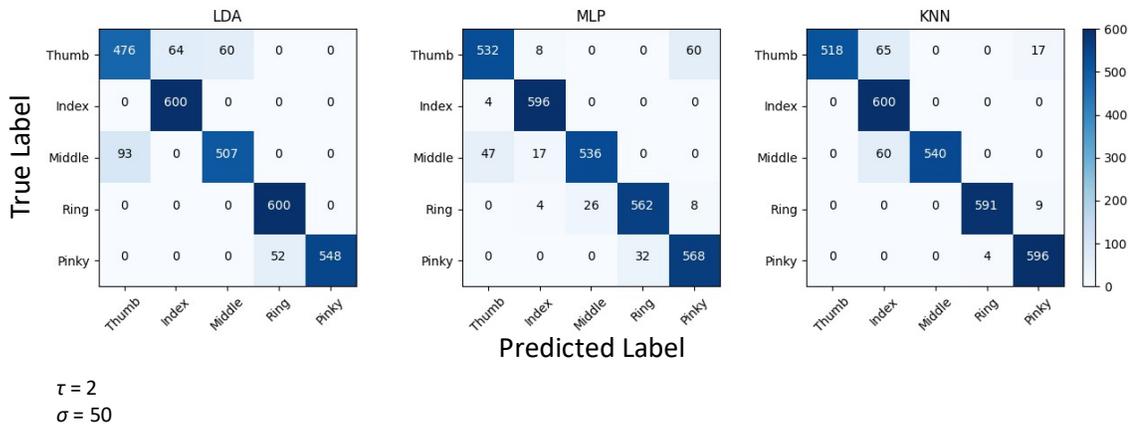


Figure 5-8 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 2$ and $\sigma = 50$ when predicting for 100 ms sequential time intervals.

Table 5-2 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 2$ and $\sigma = 50$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	91.03 ± 8.38	77.98 ± 8.53	96.25 ± 3.22	84.72 ± 6.08	91.01 ± 6.71	86.61 ± 15.07
MLP	93.13 ± 10.26	78.84 ± 11.43	94.89 ± 3.95	74.97 ± 26.43	89.55 ± 3.74	83.40 ± 8.00
KNN	94.83 ± 12.19	74.06 ± 27.22	79.47 ± 12.09	78.67 ± 27.45	98.40 ± 0.66	95.78 ± 1.64

Lowest F1 Score: KNN, thumb finger, 74.06 ± 27.22%

Highest F1 Score: KNN, ring finger, 98.40 ± 0.66%

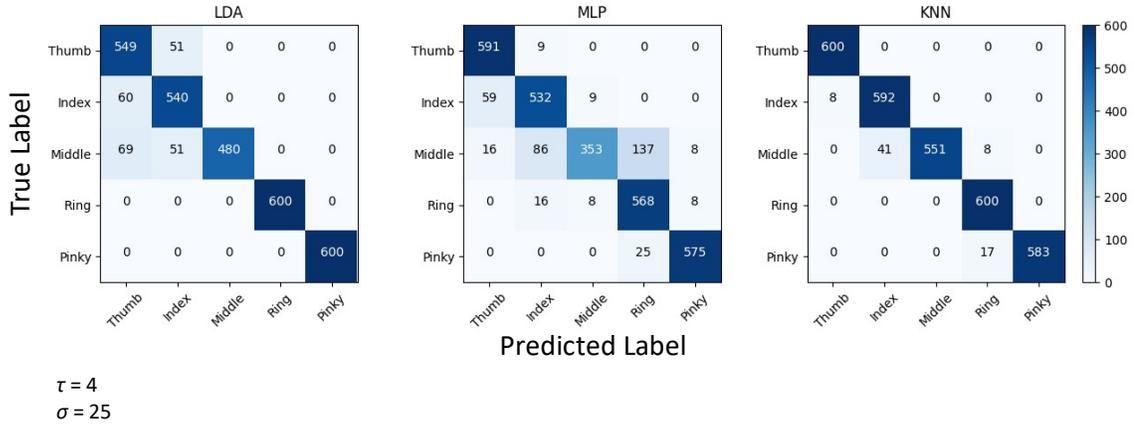


Figure 5-9 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 4$ and $\sigma = 25$ when predicting for 100 ms sequential time intervals.

Table 5-3 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 4$ and $\sigma = 25$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	92.30 ± 12.55	80.60 ± 5.48	70.71 ± 24.42	91.09 ± 6.15	100.00 ± 0.00	100.00 ± 0.00
MLP	87.30 ± 9.87	86.20 ± 8.63	71.10 ± 19.49	67.07 ± 13.57	89.37 ± 5.14	96.41 ± 1.27
KNN	97.53 ± 2.89	98.11 ± 1.64	94.73 ± 1.67	96.33 ± 2.57	97.31 ± 1.73	97.15 ± 2.04

Lowest F1 Score: MLP, middle finger, 67.07 ± 13.57%

Highest F1 Score: LDA, ring and pinky fingers, 100.00 ± 0.00%

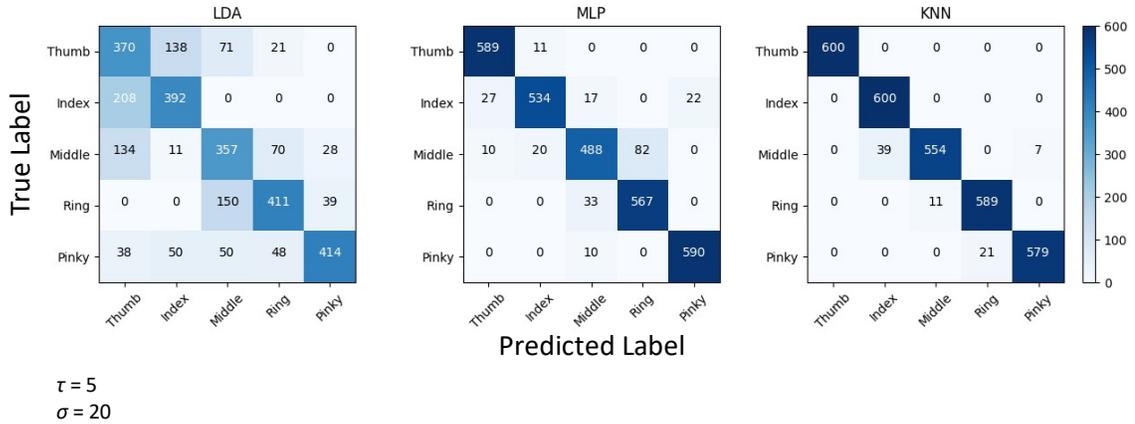


Figure 5-10 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 5$ and $\sigma = 20$ when predicting for 100 ms sequential time intervals.

Table 5-4 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 5$ and $\sigma = 20$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	64.80 ± 22.32	55.21 ± 8.56	57.58 ± 20.25	59.40 ± 6.27	61.27 ± 7.61	59.58 ± 23.27
MLP	92.27 ± 5.52	93.04 ± 2.43	84.45 ± 7.98	84.05 ± 2.15	92.08 ± 1.81	97.06 ± 2.00
KNN	97.40 ± 3.43	100.00 ± 0.00	96.83 ± 2.24	95.36 ± 3.21	96.18 ± 1.96	95.96 ± 2.43

Lowest F1 Score: LDA, thumb finger, 55.21 ± 8.56%

Highest F1 Score: KNN, index finger, 100.00 ± 0.00%

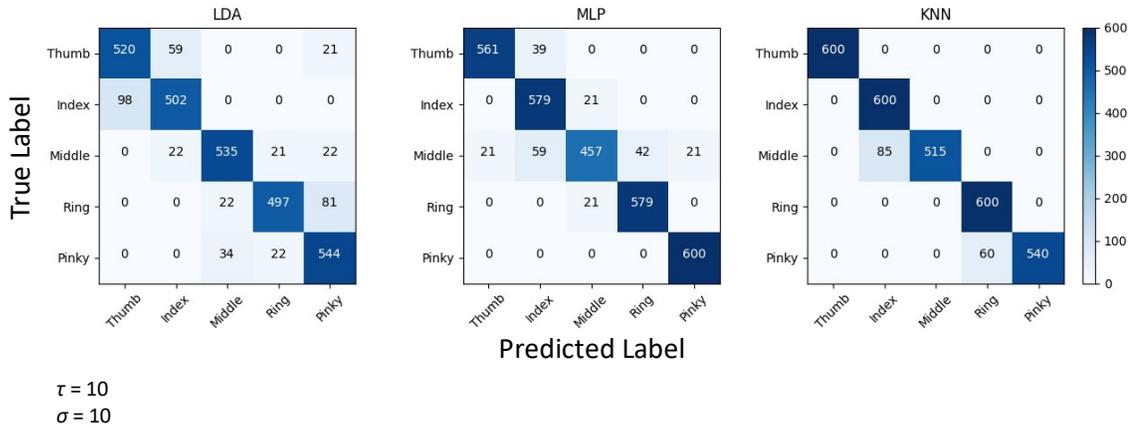


Figure 5-11 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 10$ and $\sigma = 10$ when predicting for 100 ms sequential time intervals.

Table 5-5 Cl Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 10$ and $\sigma = 10$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	86.60 ± 10.11	86.79 ± 5.33	86.28 ± 5.28	88.79 ± 1.46	82.67 ± 8.70	82.78 ± 2.72
MLP	92.53 ± 7.62	95.13 ± 3.37	90.23 ± 5.87	85.82 ± 7.87	95.74 ± 2.30	99.18 ± 1.02
KNN	95.17 ± 6.85	100.00 ± 0.00	94.03 ± 4.05	92.77 ± 4.98	91.52 ± 5.30	88.67 ± 8.96

Lowest F1 Score: LDA, ring finger, 82.67 ± 8.70%

Highest F1 Score: KNN, thumb finger, 100.00 ± 0.00%

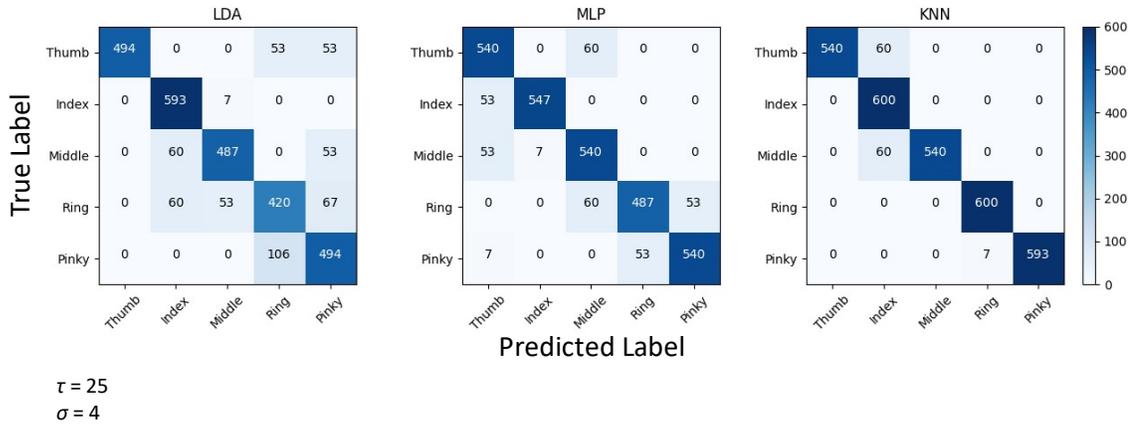


Figure 5-12 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 25$ and $\sigma = 4$ when predicting for 100 ms sequential time intervals.

Table 5-6 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 25$ and $\sigma = 4$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	82.93 ± 17.41	89.12 ± 7.21	78.27 ± 13.41	66.11 ± 24.05	47.63 ± 25.26	74.98 ± 10.86
MLP	88.47 ± 15.49	70.37 ± 14.70	81.02 ± 22.00	64.67 ± 25.28	79.36 ± 11.68	91.48 ± 3.15
KNN	95.77 ± 12.70	78.67 ± 27.45	79.80 ± 12.25	78.67 ± 27.45	98.34 ± 1.44	98.24 ± 1.64

Lowest F1 Score: LDA, ring finger, 47.63 ± 25.26%

Highest F1 Score: KNN, ring finger, 98.34 ± 1.44%

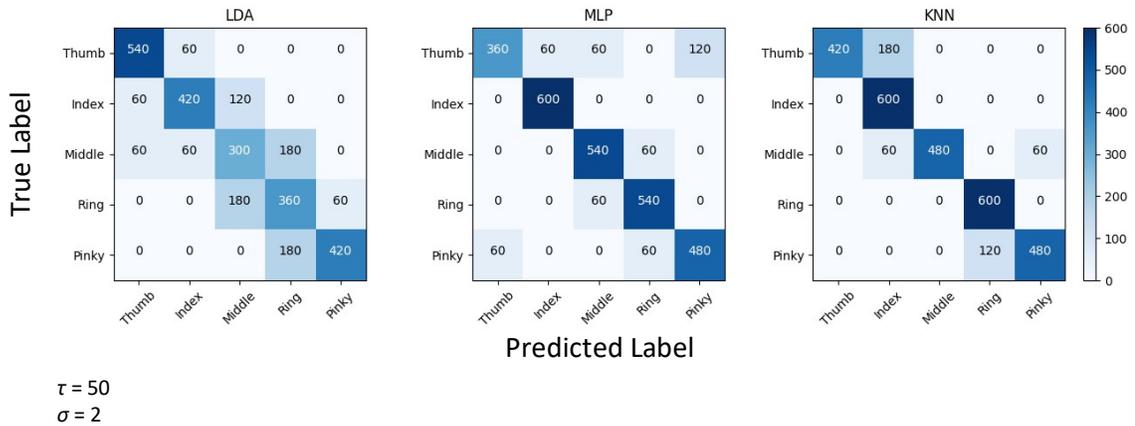


Figure 5-13 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 50$ and $\sigma = 2$ when predicting for 100 ms sequential time intervals.

Table 5-7 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 50$ and $\sigma = 2$ when predicting for 100 ms sequential time intervals.

Classifier	Classification	Thumb F1	Index F1	Middle F1	Ring F1	Pinky F1
	Accuracy (%)	Score (%)	Score (%)	Score (%)	Score (%)	Score (%)
LDA	68.00 ± 22.27	88.99 ± 4.91	54.21 ± 29.29	40.94 ± 14.70	25.56 ± 22.70	62.29 ± 22.66
MLP	84.00 ± 14.97	68.14 ± 6.85	99.52 ± 1.43	69.03 ± 23.86	79.95 ± 7.70	68.16 ± 24.77
KNN	86.00 ± 20.10	64.91 ± 22.91	73.38 ± 9.14	73.28 ± 25.23	87.28 ± 7.91	73.55 ± 25.13

Lowest F1 Score: LDA, ring finger, 25.56 ± 22.70%

Highest F1 Score: MLP, index finger, 99.52 ± 1.43%

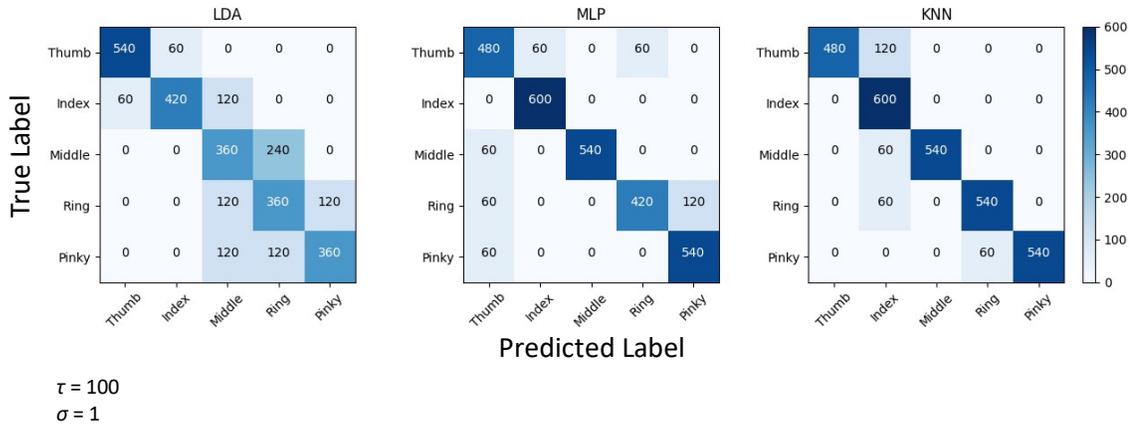


Figure 5-14 Confusion matrices for LDA, MLP and KNN. Results generated using, $\tau = 100$ and $\sigma = 1$ when predicting for 100 ms sequential time intervals.

Table 5-8 Classification metrics for LDA, MLP and KNN. Results generated using , $\tau = 100$ and $\sigma = 1$ when predicting for 100 ms sequential time intervals.

Classifier	Classification Accuracy (%)	Thumb F1 Score (%)	Index F1 Score (%)	Middle F1 Score (%)	Ring F1 Score (%)	Pinky F1 Score (%)
LDA	68.00 ± 24.00	91.00 ± 5.16	59.85 ± 31.45	49.13 ± 17.75	26.80 ± 23.42	52.40 ± 19.40
MLP	86.00 ± 23.75	48.65 ± 27.13	88.19 ± 8.46	78.67 ± 27.45	60.92 ± 21.73	68.09 ± 24.80
KNN	90.00 ± 18.44	71.38 ± 24.93	69.98 ± 12.62	78.67 ± 27.45	77.67 ± 26.95	98.89 ± 2.23

Lowest F1 Score: LDA, ring finger, 26.80 ± 23.42%

Highest F1 Score: KNN, pinky finger, 98.89 ± 2.23%

5.4 Discussion

Individual finger flexions were classified using the designed WUS system. Ultrasound RF signals were acquired in pulse-echo mode from three WUSs, attached on the forearm. Averaging of every 10 frames produced the best classification performance amongst the LDA, MLP, and KNN classifiers. Based on these results (Figure 5-11 and Table 5-5), the mean F1 scores for each finger ranged between 82% to 100% across all three classifiers. Each mean F1 score had a standard deviation less than 9%. By achieving a minimum F1 score of $82.67 \pm 8.70\%$ (LDA ring finger), it shows that averaging of every 10 frames yielded the best features when discriminating between finger type. This means that the WUS system achieved sequential F1 score prediction rates above 80%. With prediction occurring every 100 ms in time throughout the total 5 minutes of recording (3,000 predictions). The significance of these results shows that the WUS system has the potential to be used to in HMI applications requiring high accuracy while making a prediction every 100 ms in time.

The parameters of the pattern classification pipeline that gave the best classification performance was using the KNN classifier when averaging every 4 frames. This resulted in the highest classification accuracy of $97.53 \pm 2.89\%$. The mean F1 scores ranged between 94.73 – 98.11% with each standard deviation of the mean being less than 3%. This gives the optimal parameters for the PCP when used with the WUS system and proves that extremely high classification accuracies can be achieved for low resolution systems. Sikdar *et al.* used a B-mode ultrasound probe to achieve a classification accuracy of 98.33%. However, in their study they classified complete finger flexions occurring over a 1 second time interval (i.e. their accuracy did not report for predicting

every 100 ms in time but predicted for when a complete finger flexion occurred).

Regardless of this difference, the results of this study are comparable to that of a high-resolution B-mode ultrasound imaging probe. The WUS system with the proposed classification method has the capability of detecting which finger is being flexed every 100 ms in time with high accuracy. This strongly shows the feasibility and potential for these sensors to be used for HMI applications such as prosthetic devices.

In future studies, more subjects should be recruited to verify these classification accuracies, along with predicting when in the resting position. It should be noted that when predicting for the resting position, it will require more work to label the occurrence in each recording. The optimal number and placement of the WUSs should be investigated to obtain the best classification accuracy. In addition, future studies should perform the same hand gestures and motions made in previous ultrasound transducer experiments [13, 14, 53, 52, 51] and sEMG studies [7, 8, 69, 34].

Chapter 6: Conclusions and Future Works

6.1 Conclusions

In this thesis research, I first investigated the effect that lateral spatial resolution has on the classification performance of individual finger flexions. This was performed by laterally reconstructing ultrasound RF signals acquired from a linear array probe.

Secondly, I verified the performance of the proposed pattern classification pipeline at reduced lateral resolutions to classify individual finger flexions every 100 ms in time. By conducting 10 independent arm experiments, the results suggest that artificially reducing the lateral resolution to 4 – 8 ultrasound RF signals could accomplish the same accuracy as using 127 RF signals (along a 40 mm width). Thirdly, the computation speed of the pattern classification pipeline was improved to execute 10 times faster. This was due to the contribution of my DWT-MAV spatial feature extraction method which is novel to ultrasound-based studies. Then finally, I designed a WUS system to classify individual finger flexions with the pattern classification method. Multiple independent experiments were conducted to optimize the parameters of the pattern classification pipeline.

Optimum parameters of the pattern classification pipeline were obtained by conducting three simultaneous independent experiments changing only the final classification step to be one of LDA, MLP or KNN classifier. Under optimum parameters, the mean F1 scores for each finger achieved a range to be between 82% to 100% across all three classifiers. Each mean F1 score had a standard deviation of less than 9%. The significance of these results indicate that the features extracted from the WUS system and optimized pattern classification pipeline make it very easy to discriminate between finger flexion type. This shows that these sensors can be used to achieve high accuracy while meeting the optimal

delay criteria for wearable HMI applications such as prosthetic devices [2]. The results of this research provided guidance to design an ergonomic WUS system for HMI applications.

The first objective of this thesis research before designing the WUS system was to investigate the effect of reducing the lateral spatial resolution when classifying individual finger flexions. Using a 127-scan line B-mode ultrasound imaging probe, the lateral resolution was reduced to emulate an ultrasonic recording from multiple single element UTs. In the preliminary study, data was acquired from only one subject arm. The results showed that the classification accuracy increased asymptotically as a function of lateral resolution (coefficient of determination to be 0.9332). The results showed that the classification accuracy would increase significantly from 1 RF signal but would reach the asymptotic peak at around 4 RF signals. These results showed that the full lateral resolution of an ultrasonic imaging probe is not required to predict individual finger flexions with high accuracy. This suggests that a more complex bulky ultrasonic imaging probe could be replaced with a multiple single element UTs for HMI applications.

The second objective was to verify the performance of the proposed pattern classification pipeline to successfully classify individual finger flexions when reducing the lateral resolution of ultrasound data. Using 10 subject arm data sets, the results suggest that the ultrasound data with artificially reduced lateral resolution could accomplish classification accuracies and F1 scores comparable to a high-resolution clinical ultrasound probe. The experiment consisted of 10 subject arm data sets (ten independent classification experiments). The result of the classification accuracies was confirmed to be maintained when reducing the lateral resolution down to 4 RF signals

before a significant drop in accuracy occurs. With the ultrasound data originally acquired at 330 μm per RF signal (127 columns of RF signals), the classification accuracies were reported have a lower and upper quartile range of 80–93%. Then when reducing down to 5 mm RF signal (8 columns of RF signals), 6.7 mm per RF signal (6 columns of RF signals) and 10 mm per RF signal (4 columns of RF signals) under our experimental conditions the median classification accuracies fluctuate between 85–93%. However, with 4 reconstructed RF signals, the lower quartile range would drop to a classification accuracy of 75% (this means that that 4 reconstructed RF signals is not significantly different but that lower bound has decreased slightly). While using 40 mm per RF signal (averaging into a single reconstructed RF signal across the full width) the median classification accuracy drops down to 71%. This means that there is a significant drop in classification performance if no lateral resolution (a single RF signal) is present.

The third objective was to make the proposed pattern classification pipeline be faster than the conventional method. Within the pattern classification pipeline, the bottleneck came from extracting the spatial features. This was investigated by comparing two spatial feature extraction methods. The proposed DWT-MAV method was adopted from sEMG methods and is novel to ultrasound-based studies for finger classification. The DWT-MAV method provided a similar classification accuracy to the ENV-LR method. The novel DWT-MAV spatial feature extraction method preprocessed the ultrasound RF data into varying time-frequency bands. In comparison to the ENV-LR method which obtained the signal envelope before extracting spatial information. This DWT-MAV spatial feature extraction method was designed with the intent of being computationally efficient by extracting spatial features from selective frequency ranges.

The DWT would preprocess the ultrasound RF signal data into varying time-frequency levels and MAV spatial features would only be applied to the time-frequency levels containing frequency ranges with sufficient energy. These frequency ranges were determined by taking the Fourier transform and visualizing the RF signal's power spectrum. The main advantage that DWT-MAV gives, is it can make the entire PCP execute predictions under 5 ms for resolutions of $N \leq 8$ (Intel i7-6700 processor, 4 cores, 3.4 GHz clock frequency). In comparison, the ENV-LR method results in the pipeline to execute predictions up to 45 ms for resolutions of $N \leq 8$. Although 45 ms time would result in having a sufficient processing time plus data acquisition to be less than 300 ms, this processing time could be expected to scale beyond 300 ms under less powerful hardware conditions. Therefore, the faster DWT-MAV method is recommended, being approximately 10 times faster than the ENV-LR method across reconstructed resolutions of $4 \leq N \leq 8$. The DWT-MAV method would be superior to the ENV-LR method when the hardware is limited for wearable and portable devices. The significance of having a more computationally efficient feature extraction method will allow for additional processing time when using low-cost signal processing hardware for wearable HMI applications.

The final and main goal of this thesis research was to design a WUS system to continuously classifying individual finger flexions. This was achieved by using the proposed PCP, which was verified through emulation, to operate for low lateral spatial resolution ultrasound systems (as outlined by the three previous goals). Three flexible, lightweight, and thin WUSs were constructed from PVDF piezoelectric polymer film. The PVDF film had a thickness of $110\mu\text{m}$ and was coated with silver ink. Each WUS has

an active sensing array of 20 mm x 20 mm and a total thickness of 280 μ m. The three WUSs were placed on the anterior side of the forearm 5 cm away from the wrist. The highest classification accuracies observed for each machine learning classifier were $97.53 \pm 2.89\%$ for the KNN classifier, $94.47 \pm 8.89\%$ for the MLP classifier (one hidden layer of 15 nodes) and $92.30 \pm 12.55\%$ for the LDA classifier. High classification accuracies were achieved under optimal parameters by performing multiple independent experiments in this preliminary study. The features extracted from the WUS system can be used to discriminate the finger type regardless of the machine learning classifier given the optimum parameters. The significance of designing this high accuracy WUS system is that it can be used as the ergonomic sensor strategy for HMI applications.

6.2 Limitations

For future experiments, it should be noted that the experiments involving the ultrasound probe were limited to observe the flexor muscles with parallel ultrasound RF signal columns. Despite this limitation, the results of the experiments still strongly suggest that lower lateral resolutions can achieve high classification accuracies to that of a high-resolution probe. Another limitation in this study is that only one in five finger flexions were classified at a time. No experiments were conducted on classifying for simultaneous finger flexions. However, because different spatial regions can correspond to different finger flexions, there is evidence that simultaneous finger flexion could be achieved. Another limitation of this study was to perfect the classification accuracy in lab conditions. Even though the experiments were conducted under optimal lab conditions, they were necessary because I am the first to investigate research on the effect of lateral

resolution for individual finger flexion classification. Future studies should record experiments with motion artifacts, sensor shifting, sensor detachment, and randomized motion recordings. A lot of the work relating to these future experiments have been conducted for sEMG based studies and can be applied for ultrasound studies. Ultrasound-based studies could involve a sliding window analysis, randomizing the order of movements and accounting for the transition period from rest state to finger motion as conducted in relating sEMG studies [70, 71]. This would replicate a more practical usage of HMI applications. Collecting data from amputated subjects would also be significant for validating the WUSs performance for upper limb prosthetics.

6.3 Future Works

As stated throughout this thesis research, the main goal was to develop a wearable ultrasound-based HMI to perform continuous prediction of individual finger flexions. The significance of these results shows that multiple single-element UTs could be used as a replacement to a bulky B-mode ultrasound imaging probe. This allows for more ergonomic methods without losing the performance required for HMI applications such as upper limb prosthetics. Implementation of multiple WUS [67] has the potential to perform individual finger flexions in a less costly and more practical manner than a convention clinical ultrasound probe and imaging system.

The research completed was a first step in the direction of designing a wearable and flexible ultrasound-based system for HMI such as prosthetic applications. The future work suggestions are outlined by the following categories.

1) Sensor Development

For the design of the wearable and flexible transducers, the ultrasonic properties such as the composition and thickness of the piezoelectric material can be evaluated. This would allow to study optimal acoustic impedance and resonant frequency design parameters for the WUS. The development of the sensors may also include the size of the active area and method of housing. The method of housing could consider the placement of the piezoelectric layer, matching layer and backing layer. Multiple active areas could also be constructed such that an array of sensors would be constructed as one wearable unit. The material of the ultrasound couplant gel (used as the matching layer) could also be investigated to last longer for use throughout the day.

2) Data Acquisition Hardware

The hardware acquisition system could be evaluated by comparing the results for the number of multichannel transducers, pulse-echo, through transmission, pulse amplitude, repetition frequency and sampling rate. The resolution and focus position of the ultrasound beam would be key factors to investigate when spatially discriminating between muscle regions. Lab environment testing could be performed by varying the voltage levels of the pulsar-receivers, pulse echo frequency, and sampling rate. The motivation of the lab testing would be used as a guideline when determining the battery powered, portable and low-cost hardware. A portable system could be designed with the requirement of being battery powered, using low power pulsar-receivers driven with a mobile device. A suggestion could be to use a Bluetooth ultrasonic pulsar-receivers to

connect to a mobile application to drive the ultrasonic pulses and perform the required preprocessing, feature extraction and classification algorithms.

3) Classification of the Finger Movements

Akhlaghi *et al.* estimated the finger flexion angle by comparing B-mode ultrasound imaging with markers placed on hand and fingers [45]. Estimation of the finger's bending angle could be tested for the WUS using a similar setup to monitor the fingers. The relationship between classification accuracy and speed of the flexions/extensions could also be studied.

4) Simulation Experiments

Simulation experiments could be performed using a soft tissue-mimicking phantom. The simulation experiments could study the ultrasonic properties, design parameters and sensor layout of a virtual sensor. The soft tissue-mimicking phantom could consider replicating multiple tissue regions to simulate different flexor muscles. The aim of the simulation study would be to discriminate between the movements of the different tissue regions.

5) Signal Analysis and Classification Methods

Various algorithms could be compared to determine optimal feature extraction when using low-cost processing hardware. This thesis investigated preprocessing of the ultrasound RF signals using Mallat's filter bank implementation of the DWT. The lifting scheme could be used to implement the DWT (as opposed to filter banks used in this

thesis), which would further reduce the computation time required to preprocess the analysis and detail coefficients. The model's performance between various feature extraction methods and classifiers could be evaluated in an offline and online setting. Transfer learning strategies could be performed to improve the adaptability of the ultrasound system when training between different subjects. Yang *et al.* investigated using adaptive learning strategies to improve inter-subject sEMG-based hand movement classification [72]. The adaptive learning strategies could be applied to ultrasound-based acquisition and account for the differences in the feature space that occurs between different subjects. To be specific, the occurrence of drift in sEMG patterns as shown in Figure 6 of Yang *et al.* "Improving the functionality, robustness, and adaptability of myoelectric control for dexterous motion restoration" [72], would be expected to also occur for ultrasound data when analyzed between subjects. In this thesis research, the temporal features were based on the difference between frames of the spatial features. This is equivalent to using a constant coefficient, first order difference equation on the spatial features along the frame axis. The temporal feature extraction method could compare using a varying number of consecutive frames (a higher order difference equation). McIntosh *et al.* extracted features using the optical flow algorithm on B-mode ultrasound to achieve continuous prediction of the finger flexion angle [46]. A temporal feature extraction method could be investigated based on a one-dimensional implementation of the optical flow algorithm when used with the WUS system.

6) In vivo Experiments

The location of the WUS placed on the forearm (both anterior and posterior) could be compared to investigate which regions could produce the highest classification accuracy and be least susceptible to motion artifact noise.

Appendices

Appendix A Signal Processing and Machine Learning References

A.1 Review on the Wavelet Transform

The wavelet transform is a useful mathematical tool to analyze signals in both time and frequency domains. The continuous wavelet transform CWT_s of a continuous signal $s(t)$ is defined as follows in (A.1) [73].

$$CWT_s(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (\text{A.1})$$

The scaling and shifting parameters, a and b , are used to modify the mother wavelet $\psi(t)$, which acts as an impulse response of a given band pass filter. The scaling and shifting of the mother wavelet is used to measure the similarity of the continuous signal with varying frequency ranges at different time intervals. Figure A-1 shows a comparison with the Short Time Fourier Transform (STFT) and the Wavelet Transform (WT). While both transforms can be used to analyze the frequencies of a time varying signal, the WT has constant-Q frequency resolution (i.e., constant frequency resolution in % of frequency considered) compared to the STFT requiring to adjust the window length.

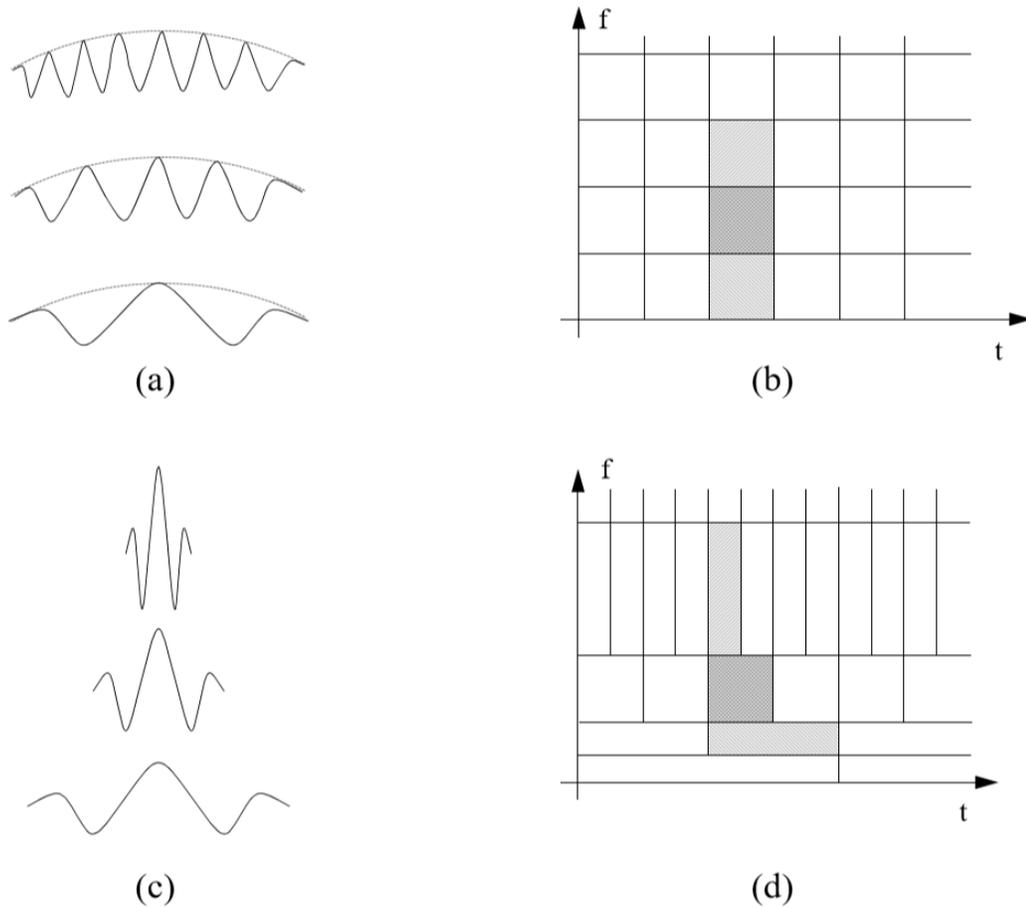


Figure A-1 Visualization of the Short Time Fourier Transform (STFT) and Wavelet Transform (WT). a) Frequency modulations of a Gaussian window used in the STFT. a) Time-frequency tiling of the STFT. c) Shifting and scaling of a mother wavelet used in the WT. d) Time-frequency tiling of the WT. Image is reproduced from “Wavelets and Subband Coding” in Figure 2.12, by Martin Vetterli and Jelena Kovačević, and the image is licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License [69].

A practical implementation of the wavelet transform is with the Discrete Wavelet Transform (DWT). The DWT can be obtained with Mallat’s filter bank implementation by passing the discrete signal s through a series of filters and downsampling as shown in Figure A-2.

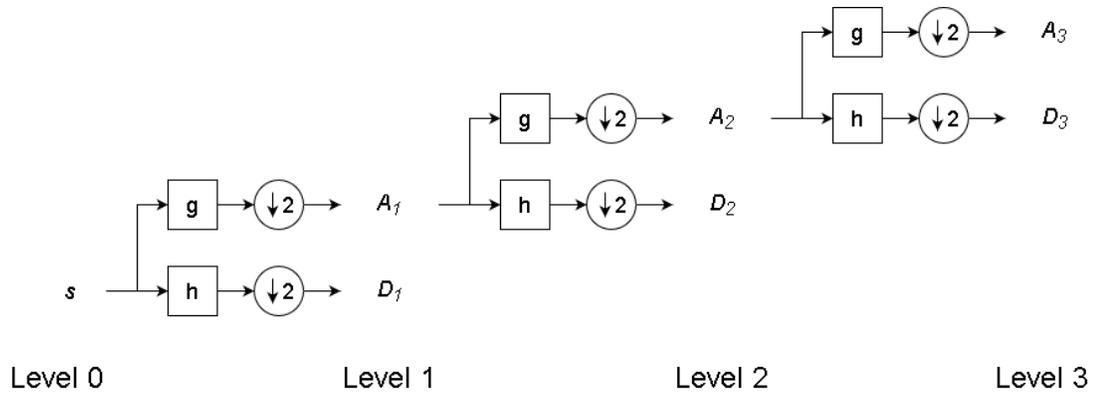


Figure A-2 Mallat's filter bank implementation of the discrete wavelet transform [61].

The approximation coefficients A_l and detail coefficients D_l are computed at each level l . The computed detail coefficients result in the discrete realization of the wavelet transform with scaling and shifting in powers of two. These detail coefficients are recursively computed with the approximation coefficients from the previous level by performing a convolution with high pass filter g and low pass filter h as follows in (A.2), (A.3) and (A.4).

$$A_0[n] = s[n] \quad (\text{A.2})$$

$$A_l[n] = \frac{1}{\sqrt{2}} \sum_k g[k] A_{l-1}[2n - k] \quad (\text{A.3})$$

$$D_l[n] = \frac{1}{\sqrt{2}} \sum_k h[k] A_{l-1}[2n - k] \quad (\text{A.4})$$

The filters g and h are chosen based on the mother wavelet $\psi(t)$ to correspond in a discrete realization of the wavelet transform computed as the detail coefficients.

The DWT has been commonly used in classifying hand and finger movements by preprocessing the sEMG signals into multiple time-frequency representations for feature extraction [55, 56, 57, 58, 59, 60, 61].

A.2 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is used in this thesis to help discriminate between individual finger flexions. LDA is a dimensionality reduction technique that projects the original data onto a lower dimensional space, then linear functions are used to separate the projection space into class regions. The LDA projection is achieved by determining the between class variance S_B , within class variance S_W , then constructing a projection matrix \mathbf{w} that maximizes S_B and minimizes S_W [74]. With the cost function $J(\mathbf{w})$ of the projection matrix w , it is maximizing in the following equation in (A.5), when applied to the data \mathbf{x} in (A.6).

$$J(\mathbf{w}) = \frac{\mathbf{w}^t S_B \mathbf{w}}{\mathbf{w}^t S_W \mathbf{w}} \quad (\text{A.5})$$

$$\mathbf{y} = \mathbf{w}^t \mathbf{x} \quad (\text{A.6})$$

The dimensionality of the data matrix $\mathbf{x} \in R^{d \times c}$ is reduced to $\mathbf{y} \in R^{k \times c}$ with d as the number of features, c as the number of classes, and k as the reduced dimension of features (where $d - k$ features are removed). LDA is a computationally efficient and robust classification method that has been commonly used involving sEMG and ultrasound data for human machine interface applications [13, 14, 52, 53, 62].

Appendix B Acquisition for Four Ultrasound Transducers

B.1 Hardware Wiring Management up to Four Channels

The methods to record ultrasound RF signal data is contained within this appendix section B. Figure B-1 shows the wiring setup for recording up to 4 UTs operating in pulse-echo mode for the ATS 9940 Waveform Digitizer by AlazarTech.

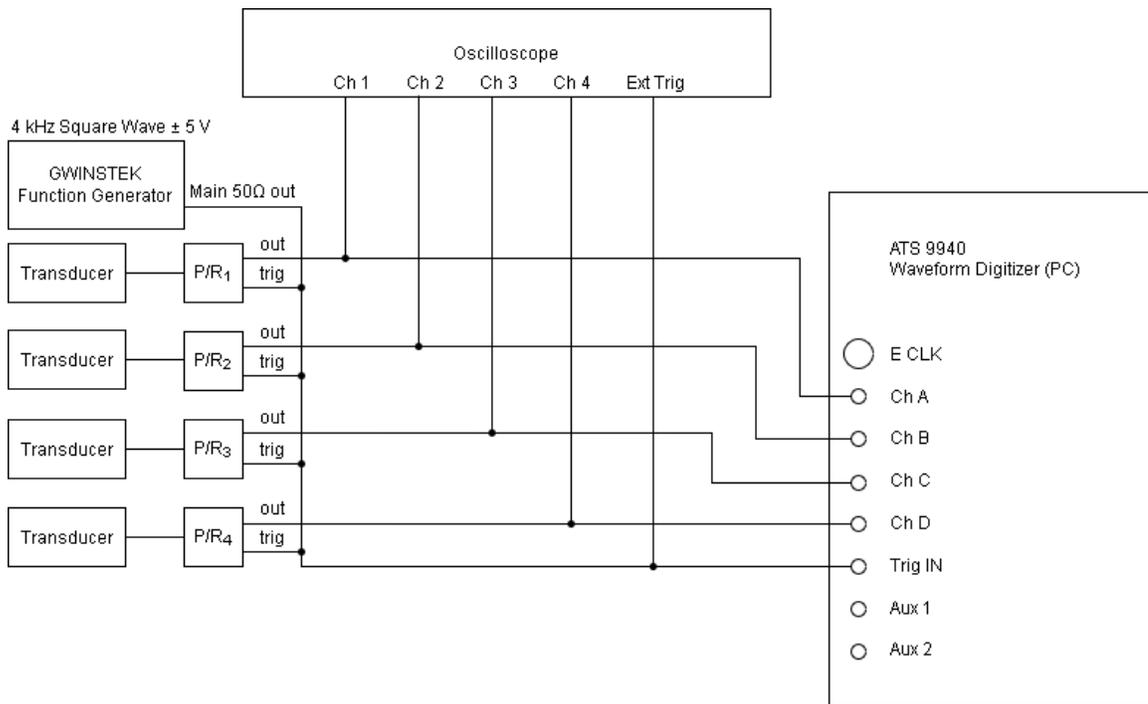


Figure B-1 Full hardware wiring diagram setup for using up to four UTs, pulsar/receiver, function generator, oscilloscope, and analog/digital converter (waveform digitizer).

B.2 Python Code: readATB.py

The following python file readATB.py in Figure B-2 was written and created based on the Alazar DSO User Guide version 1.1.72 [75]. The following python code is used to read the AlazarTech ATB binary files into usable python variables. ATB files are recorded using the AlazarDSO software program. Using the electronic copy of this thesis, the code can be copied directly into a python file named “readATB.py”. Then follow the instructions given in the read_atb function as an example on how to read an ATB file. The NumPy Python library [76] is required to run the readATB code and Matplotlib [77] is required to plot the ATB file into an image.

```

"""readATB

This python module is used to read AlazarTech ATB files which are in binary.
Created based on AlazarDsoGuide-1.1.72.pdf section 4.10 - User Manual which can be found at
https://www.alazartech.com/Product/AlazarDSO/

read_atb is the main method used by other python scripts and modules and the only method you should call unless manually
reading the binary file is necessary.
"""

__author__ = 'Alexander Fernandes'
__version__ = '0.1'

import numpy as np
from typing import Tuple, IO

class FileInfo(object):
    """A container to hold dynamic information when reading a file.

    This class is used to dynamically hold instantaneous variables when reading information from a file. It is
    specifically used to organize data being read by a file using the fread function. The only variable it will always
    have is position, which is used to determine the location it is pointing to in the file.
    """

    def __init__(self):
        self.position = 0 # intended as private variable for purposes of manually reading a binary file in units of bytes

def fread(fid: IO, nelements: int, dtype: np.dtype, finfo: FileInfo) -> np.ndarray:
    """Read specified size of data from file.

    Read desired amount of bytes based on nelements and dtype as element size into a numpy array.
    This function is intended to only be used by read_atb unless you know what you are doing.

    :param fid: file object
    :param nelements: number of dtype elements to read
    :param dtype: size of a single element to read
    :param finfo: FileInfo class that holds external file information
    :type fid: IO
    :type nelements: int
    :type dtype: numpy.dtype
    :type finfo: object

    :return: array of elements read from the file from the last position it read from
    :rtype: numpy.ndarray

    """
    if dtype is np.str:
        dt = np.uint8 # WARNING: assuming 8-bit ASCII for np.str!
    else:
        dt = dtype

    data_array = np.fromfile(fid, dt, nelements)
    data_array.shape = (nelements,)
    if nelements == 1:
        data_array = data_array[0]

    finfo.position += np.dtype(dtype).itemsize * nelements

    return data_array

class CAPTURE_FILE_HEADER(object):
    """Generic class to organize the binary data into easily accessible public instantiated variables

    An ATB file starts with a CAPTURE_FILE_HEADER structure.
    m_dwType
        |Specifies the file type. It must be 'ATSb'.

    m_dwSize
        |Specifies the size in bytes of the CAPTURE_FILE_HEADER structure.

    m_nBitsPerSample
    m_nBytesPerSample
    m_uSamplesPerRecord
    m_dSamplesPerSecond
    m_dSecondsPerRecord
    m_uTrigDelay_samples
    m_nPreTrig_samples
    m_nRecordsPerCapture
        |Specifies the acquisition properties.

    m_nCapturesPerFile
        |Specifies the number of captures in the file.

    m_nChannels
        |Specifies the total number of CHANNEL_FILE_HEADER structures that follow.

    m_cComment[32]
        |Specifies a user-defined comment for this acquisition.
    """

```

```

m_dwFlags
| Specifies attributes of the file. These attributes may include:
| Value      Name      Meaning
| 0x00000001 CFHF_SAMPLE_INTERLEAVED If this flag is set, samples in a record are channel
|                                     interleaved (i.e. A-B-A-B). If not set, samples in a
|                                     record are channel contiguous (i.e. A-A-B-B).
| 0x00000002 CFHF_CONTINUOUS   If set, each record is a transfer in a continuous
|                                     acquisition.
| 0x00000004 CFHF_STREAM       This flag indicated how the sample data is
|                                     organized in the file.
| If this flag is not set, then the records are arranged as described in section 4.10.3.1 "Normalized records"

m_dwDataOffset_bytes
| Specifies the offset in bytes from the start of file of the sample data.

m_dwStreamFlags
| When the CFHF_STREAM bit is set, then this field contains flags that define how the
| sample data is organized in the file.

m_nTimestampMultiplier
| This value specifies a board specific multiplier used in the conversion of timestamp
| values from clock ticks to seconds.

"""

def __init__(self):
    self.m_dwType = None # type: str -> 4 characters
    self.m_dwSize = None # type: np.uint32
    self.m_nBitsPerSample = None # type: np.int32
    self.m_nBytesPerSample = None # type: np.int32
    self.m_uSamplesPerRecord = None # type: np.uint32
    self.m_dSamplesPerSecond = None # type: np.double
    self.m_dSecondsPerRecord = None # type: np.double
    self.m_nTrigDelay_samples = None # type: np.int32
    self.m_uPreTrig_samples = None # type: np.uint32
    self.m_nRecordsPerCapture = None # type: np.int32
    self.m_nCapturesPerFile = None # type: np.int32
    self.m_nChannels = None # type: np.int32
    self.m_cComment = None # type: str -> 32 characters
    self.m_dwFlags = None # type: np.uint32
    self.m_dwDataOffset_bytes = None # type: np.uint32
    self.m_dwStreamFlags = None # type: np.uint32
    self.m_nTimestampMultiplier = None # type: np.uint32

class CHANNEL_FILE_HEADER_ARRAY_ELEMENT(object):
    """Generic class to organize the binary data into easily accessible public instantiated variables

    A CHANNEL_FILE_HEADER structure specifies the properties of one input channel.
    A CHANNEL_FILE_HEADER_ARRAY structure follows the CAPTURE_FILE_HEADER structure in an ATB file.
    Note that the m_nChannels field from the CAPTURE_FILE_HEADER defines the number of channel file header
    structures in the file.

    m_bDisabled
    | Specifies if the channel is disabled. If the channel is disabled, then no records from the
    | channel are written to the file.

    m_dInputRange_volts
    | Specifies the input range of the sample data in volts

    m_dProbeAttenuation
    | Specifies the attenuation of the probes used during the acquisition.

    """

    def __init__(self):
        """Class structured as an array according to the number of channels specified

        This class is intended to be an array element encapsulated in a list (i.e. channel 1 indexed as 0)

        """
        self.m_bDisabled = None # type: bool
        self.m_dInputRange_volts = None # type: np.double
        self.m_dProbeAttenuation = None # type: np.double

def read_atb(filename: str):
    """Read AlazarTech ATB binary files

    \n
    ATB file format structure:
    1. CAPTURE_FILE_HEADER
    2. CHANNEL_FILE_HEADER array
    3. Binary sample data

    \n
    Since the binary file is arranged in this order, read_atb will return them as a Tuple\n
    \n
    :return: capture_file_header, channel_file_header_array, sample_data, frame_time_stamps
    :rtype: object, object, np.ndarray, np.array or None

    - capture_file_header: class that holds all header information about the binary file\n
    - channel_file_header_array: array of classes containing header information specific to each channel (zero-indexed)\n
    - sample_data: the signal data stored as a 3D array: (depth, channel, frame)\n

```

```

- frame_time_stamps: if included in the binary file, 1D array of the slow-time timestamp for each frame\n
\n
:Example:\n
>>> atb_directory = "<Directory Location>"
>>> capture_file_header, channel_file_header_array, sample_data, frame_time_stamps = read_atb(atb_directory + "filename.atb")
>>> import matplotlib.pyplot as plt
>>> plt.figure(1)
>>> plt.plot(sample_data[:, 0, 0]) # plot a frame from the first channel
>>> plt.figure(2)
>>> plt.imshow(sample_data[:, 0, :], aspect='auto') # plot first channel in m-mode
\n
**Authors**
- Alexander Fernandes

"""

finfo = FileInfo()
capture_file_header = CAPTURE_FILE_HEADER()
channel_file_header_array = []
sample_data = None
time_stamps = None

fid = open(filename, 'r')

## C-type variables to python numpy types (based on experimenting and based on the documentation)
# MS-DTYP | Numpy-Type (size)
# DWORD  | np.uint32
# int     | np.int32
# UINT    | np.uint32
# double  | np.float64
# WCHAR   | np.uint16 <- UTF-16
# BOOL    | np.uint32

#
# Capture File Header
#

# Specifies the file type. It must be 'ATSb' for ATS binary type
capture_file_header.m_dwType = ''.join([chr(fread(fid, 1, np.ubyte, finfo)) for _ in range(4)])

if capture_file_header.m_dwType == 'ATSb':
    capture_file_header.m_dwSize = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_nBitsPerSample = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_nBytesPerSample = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_uSamplesPerRecord = fread(fid, 1, np.uint32, finfo)
    m_dwReserved_0 = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_dSamplesPerSecond = fread(fid, 1, np.float64, finfo)
    capture_file_header.m_dSecondsPerRecord = fread(fid, 1, np.float64, finfo)
    capture_file_header.m_nTrigDelay_samples = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_uPreTrig_samples = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_nRecordsPerCapture = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_nCapturesPerFile = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_nChannels = fread(fid, 1, np.int32, finfo)
    capture_file_header.m_cComment = ''.join([chr(fread(fid, 1, np.uint16, finfo)) for c in range(32)])
    capture_file_header.m_dwFlags = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_dwDataOffset_bytes = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_dwStreamFlags = fread(fid, 1, np.uint32, finfo)
    capture_file_header.m_nTimestampMultiplier = fread(fid, 1, np.uint32, finfo)
    m_dwReserved = fread(fid, 13, np.uint32, finfo)

#
# Capture File Header Array
#

for channel in range(capture_file_header.m_nChannels):
    channel_file_header_array_element = CHANNEL_FILE_HEADER_ARRAY_ELEMENT
    channel_file_header_array_element.m_bDisabled = bool(fread(fid, 1, np.uint32, finfo))
    m_dwReserved_0 = fread(fid, 1, np.uint32, finfo)
    channel_file_header_array_element.m_dInputRange_volts = fread(fid, 1, np.float64, finfo)
    channel_file_header_array_element.m_dProbeAttenuation = fread(fid, 1, np.float64, finfo)
    m_dwReserved = fread(fid, 4, np.uint32, finfo)

    channel_file_header_array.append(channel_file_header_array_element)

# Advance to m_dwDataOffset_bytes position in file
while (finfo.position < capture_file_header.m_dwDataOffset_bytes):
    fread(fid, 1, np.uint8, finfo)

#
# Sample Data
#

uFramesTotal = int(capture_file_header.m_nRecordsPerCapture * capture_file_header.m_nCapturesPerFile)

# select sample size in bytes using numpy data types
sample_dtype = np.byte
if capture_file_header.m_nBytesPerSample == 1:
    sample_dtype = np.uint8
elif capture_file_header.m_nBytesPerSample == 2:
    sample_dtype = np.uint16
elif capture_file_header.m_nBytesPerSample == 4:
    sample_dtype = np.uint32

```

```

elif capture_file_header.m_nBytesPerSample == 8:
    sample_dtype = np.uint64
elif capture_file_header.m_nBytesPerSample == 16:
    sample_dtype = np.uint128

bInterleaved = False
bContinuous = False

if capture_file_header.m_dwFlags & 0x00000001: # CFHF_SAMPLE_INTERLEAVED
    bInterleaved = True
if capture_file_header.m_dwFlags & 0x00000002: # CFHF_CONTINUOUS
    bContinuous = True

if capture_file_header.m_dwFlags == 0x00000004: # CFHF_STREAM
    # TODO: Implement Stream. Follow "Raw records" section 4.10.3.2

    sample_data = np.zeros([capture_file_header.m_uSamplesPerRecord,
                           capture_file_header.m_nChannels, # first index used as time stamp values
                           uFramesTotal],
                           sample_dtype)

    bCBF_AUTODMA = False
    bCBF_NO_PRETRIGGER = False
    bCBF_SAMPLE_INTERLEAVED = False
    bCBF_HAVE_HEADER = False
    bCBF_ADC_CALIBRATION_V_1 = False
    if capture_file_header.m_dwStreamFlags & 0x00000001:
        bCBF_AUTODMA = True
    if capture_file_header.m_dwStreamFlags & 0x00000002:
        bCBF_NO_PRETRIGGER = True
    if capture_file_header.m_dwStreamFlags & 0x00000010:
        bCBF_SAMPLE_INTERLEAVED = True
    if capture_file_header.m_dwStreamFlags & 0x00000004:
        bCBF_HAVE_HEADER = True
    if capture_file_header.m_dwStreamFlags & 0x00000008:
        bCBF_ADC_CALIBRATION_V_1 = True

    # 4.10.3.2.1 AutoDMA from FIFO
    if bCBF_AUTODMA and bCBF_SAMPLE_INTERLEAVED:
        pass # TODO

    # 4.10.3.2.2 AutoDMA no-pretrigger mode
    elif bCBF_AUTODMA and not bCBF_SAMPLE_INTERLEAVED and bCBF_NO_PRETRIGGER:
        pass # TODO

    # 4.10.3.2.3 AutoDMA pre-trigger mode
    elif bCBF_AUTODMA and not bCBF_SAMPLE_INTERLEAVED and not bCBF_NO_PRETRIGGER:
        # In this mode sample data is organized into records where a record from Ch A (if enabled) is followed
        # by a record from Ch B (if enabled). In total, there are m_nRecordsPerCapture for Ch A (if enabled),
        # and m_nRecordsPerCapture records for Ch B (if enabled).
        for frame in range(uFramesTotal):
            if bCBF_HAVE_HEADER:
                # TODO
                pass
            for channel in range(int(capture_file_header.m_nChannels)):
                sample_data[:, channel, frame] = fread(fid,
                                                       int(capture_file_header.m_uSamplesPerRecord),
                                                       sample_dtype,
                                                       finfo)

else:
    # Normalized records

    # [depth, (timestamp, channel A, channel B, ..), frames]
    sample_data = np.zeros([capture_file_header.m_uSamplesPerRecord,
                           capture_file_header.m_nChannels, # first index used as time stamp values
                           uFramesTotal],
                           sample_dtype)

    if not bContinuous:
        time_stamps = np.zeros([uFramesTotal], np.float64)
    if bInterleaved:
        # Channel Interleaved

        # Note: timestamp only included if CFHF_CONTINUOUS flag is false
        # Capture 0 Record 0 Timestamp -> (only 1 value)
        # Capture 0 Record 0 Ch A & B & ... samples -> (m_nSamplesPerRecord * m_nChannels long)
        # Capture 0 Record 1 Timestamp
        # Capture 0 Record 1 Ch A & B & ... samples
        # ...
        # Capture 0 Record N-1 Timestamp
        # Capture 0 Record N-1 Ch A & B & ... samples
        # Capture 1 Record 0 Timestamp
        # Capture 1 Record 0 Ch A & B & ... samples
        # ...

        # TODO: implement interleaved
        pass

    else:
        # Channel Contiguous

```

```

# Note: timestamp only included if CFHF_CONTINUOUS flag is false
# Capture 0 Record 0 Timestamp -> (only 1 value)
# Capture 0 Record 0 Ch A samples -> (m_nSamplesPerRecord long)
# Capture 0 Record 0 Ch B samples
# Capture 0 Record 1 Timestamp
# Capture 0 Record 1 Ch A samples
# Capture 0 Record 1 Ch B samples
# ...
# Capture 0 Record N-1 Timestamp
# Capture 0 Record N-1 Ch A samples
# Capture 0 Record N-1 Ch B samples
# Capture 1 Record 0 Timestamp
# Capture 1 Record 0 Ch A samples
# Capture 1 Record 0 Ch B samples
# ...

for frame in range(uFramesTotal):
    if not bContinuous:
        time_stamps[frame] = fread(fid, 1, np.float64, finfo)
        for channel in range(int(capture_file_header.m_nChannels)):
            sample_data[:, channel, frame] = fread(fid,
                                                    int(capture_file_header.m_uSamplesPerRecord),
                                                    sample_dtype,
                                                    finfo)
    else:
        print("The file type is not 'ATsB'!\n\n")

fid.close()

return (capture_file_header, channel_file_header_array, sample_data, time_stamps)

```

Figure B-2 readATB.py Python code.

References

- [1] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Giamann, H. Dietl and O. C. Aszmann, "The extration of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797-809, 2014.
- [2] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 15, no. 1, pp. 111-118, 2007.
- [3] K. Englehart, B. Hudgins and A. D. C. Chan, "Continuous multifunction myoelectric control using pattern recognition," *Technology and Disability*, vol. 15, no. 2, pp. 95-103, 2003.
- [4] T. Davis, H. Wark, D. Hutchinson, D. Warren, K. O'neill, T. Scheinblum, G. Clark, R. Normann and B. Greger, "Restoring motor control and sensory feedback in people with upper extremity amputations using arrays of 96 microelectrodes implanted in the median and ulnar nerves," *Journal of neural engineering*, vol. 13, no. 3, pp. 1-15, 2016.
- [5] A. Branner, R. B. Stein, E. Fernandez, Y. Aoyagi and R. A. Normann, "Long-term stimulation and recording with a penetrating microelectrode array in cat sciatic nerve," *IEEE transactions on biomedical engineering*, vol. 51, no. 1, pp. 146-157, 2004.

- [6] M. B. Christensen, S. M. Pearce, N. M. Ledbetter, D. J. Warren, G. A. Clark and P. A. Tresco, "The foreign body response to the Utah Slant Electrode Array in the cat sciatic nerve," *Acta biomaterialia*, vol. 10, no. 11, pp. 4650-4660, 2014.
- [7] N. Celadon, S. Došen, I. Binder, P. Ariano and D. Farina, "Proportional estimation of finger movements from high-density surface electromyography," *Journal of NeuroEngineering and Rehabilitation*, vol. 13, no. 73, pp. 1-19, 2016.
- [8] M. Atzori and H. Muller, "Control capabilities of myoelectric robotic prostheses by hand amputees: a scientific research and market overview," *Frontiers in systems neuroscience*, vol. 9, no. 162, pp. 1-7, 2015.
- [9] S. Sikdar, H. Rangwala, E. B. Eastlake, I. A. Hunt, A. J. Nelson, J. Devanathan, A. Shin and J. J. Pancrazio, "Novel Method for Predicting Dexterous Individual Finger Movements by Imaging Muscle Activity Using a Wearable Ultrasonic System," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 69-76, 2013.
- [10] Y. Huang, X. Yang, Y. Li, D. Zhou, K. He and H. Liu, "Ultrasound-based sensing models for finger motion classification," *IEEE journal of biomedical and health informatics*, vol. 22, no. 5, pp. 1395-1405, 2017.
- [11] E. Ackerman, "'Skywalker' Prosthetic Hand Uses Ultrasound for Finger-Level Control," *IEEE Spectrum*, 12 December 2017. [Online]. Available: <https://spectrum.ieee.org/the-human-os/biomedical/devices/skywalker-prosthetic-hand-uses-ultrasound-sensors-for-fingerlevel-control>. [Accessed 10 July 2020].

- [12] J. Maderer, "The Force is Strong: Amputee Controls Individual Prosthetic Fingers," Georgia Institute of Technology, 11 December 2017. [Online]. Available: <http://www.news.gatech.edu/2017/12/11/force-strong-amputee-controls-individual-prosthetic-fingers>. [Accessed 10 July 2020].
- [13] X. Yang, X. Sun, D. Zhou, Y. Li and H. Liu, "Towards wearable A-mode ultrasound sensing for real-time finger motion recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 6, pp. 1199-1208, 2018.
- [14] J. Yan, X. Yang, X. Sun, Z. Chen and H. Liu, "A Lightweight Ultrasound Probe for Wearable Human–Machine Interfaces," *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5895-5903, 2019.
- [15] H. Gray, "Anatomy of the human body: chapter 4.7e. the muscles and fasciæ of the forearm," 1918. [Online]. Available: <https://www.bartleby.com/107/illus415.html>. [Accessed 11 May 2020].
- [16] H. Gray, "Anatomy of the Human Body," Philadelphia: Lea & Febiger, 1918; Bartleby.com, 2000. [Online]. Available: www.bartleby.com/107/. [Accessed 26 July 2020].
- [17] H. Gray, "Anatomy of the human body: chapter 4.7f. the muscles and fisciæ of the hand," 1918. [Online]. Available: <https://www.bartleby.com/107/illus421.html>. [Accessed 11 May 2020].

- [18] D. W. Rickey, "Medical Ultrasound Scanner," Wikipedia, 15 February 2006. [Online]. Available: <https://en.wikipedia.org/wiki/File:AlokaPhoto2006a.jpg>. [Accessed 21 July 2020].
- [19] D. W. Rickey, "Medical Ultrasound linear array Probe/scan head/transducer," Wikimedia Commons, 15 February 2006. [Online]. Available: <https://commons.wikimedia.org/wiki/File:UltrasoundProbe2006a.jpg>. [Accessed 21 July 2020].
- [20] J. L. Prince and J. M. Links, *Medical Imaging Signals and Systems*, Upper Saddle River, New Jersey, United States of America: Pearson, 2015.
- [21] C. K. Battye, A. Nightingale and J. Whillis, "The use of myo-electric currents in the operation of prostheses," *The Journal of bone and joint surgery*, vol. 37, no. 3, pp. 506-510, 1955.
- [22] N. Berger and C. Huppert, "The use of electrical and mechanical muscular forces for the control of an electric prosthesis," *Berger, Norman, and Curtis R. Huppert. "The use of electrical and mechanical muscular forces for the control of an electrical prosthesis." The American journal of occupational therapy: official publication of the American Occupational Therapy Association*, vol. 6, no. 3, pp. 110-114, 1952.
- [23] C. G. Y. Ngan, R. M. I. Kapsa and P. F. M. Choong, "Strategies for neural control of prosthetic limbs: from electrode interfacing to 3D printing," *Materials*, vol. 12, no. 12, pp. 1-14, 2019.

- [24] N. Birbaumer, T. Hinterberger, A. Kubler and N. Neumann, "The thought-translation device (TTD): neurobehavioral mechanisms and clinical outcome," *IEEE transactions on Neural Systems and rehabilitation engineering*, vol. 11, no. 2, pp. 120-123, 2003.
- [25] C. Neuper, G. R. Müller-Putz, R. Scherer and G. Pfurtscheller, "Motor imagery and EEG-based control of spelling devices and neuroprostheses," *Progress in brain research*, vol. 159, pp. 393-409, 2006.
- [26] G. Pfurtscheller, G. R. Müller-Putz, J. Pfurtscheller and R. Rupp, "EEG-based asynchronous BCI controls functional electrical stimulation in a tetraplegic patient," *EURASIP Journal on Advances in Signal Processing*, vol. 19, pp. 3152-3155, 2005.
- [27] B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq and F. Vidal, "Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view," *International Journal of Psychophysiology*, vol. 97, no. 3, pp. 210-220, 2015.
- [28] T. Yanagisawa, M. Hirata, Y. Saitoh, H. Kishima, K. Matsushita, T. Goto, R. Fukuma, H. Yokoi, Y. Kamitani and T. Yoshimine, "Electrocorticographic control of a prosthetic arm in paralyzed patients," *Annals of neurology*, vol. 71, no. 3, pp. 353-361, 2012.
- [29] S. Morishita, K. Sato, H. Watanabe, Y. Nishimura, T. Isa, R. Kato, T. Nakamura and H. Yokoi, "Brain-machine interface to control a prosthetic arm with monkey ECoGs during periodic movements," *Frontiers in neuroscience*, vol. 8, no. 417, pp. 1-9, 2014.

- [30] G. Hotson, D. McMullen, M. S. Fifer, M. S. Johannes, K. D. Katyal, M. P. Para, R. Armiger, W. S. Anderson, N. V. Thakor, B. A. Wester and N. E. Crone, "Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject," *Journal of neural engineering*, vol. 13, no. 2, pp. 1-13, 2016.
- [31] D. Myers, "EEG Cap," Wikimedia Commons, 12 June 2010. [Online]. Available: https://commons.wikimedia.org/wiki/File:EEG_cap.jpg. [Accessed 22 July 2020].
- [32] B. P. Christie, M. Freeberg, W. D. Memberg, G. J. Pinault, H. A. Hoyen, D. J. Tyler, R. Triolo and J. "Long-term stability of stimulating spiral nerve cuff electrodes on human peripheral nerves," *Journal of neuroengineering and rehabilitation* , vol. 14, no. 70, pp. 1-12, 2017.
- [33] Y. Fang, N. Hettiarachchi, D. Zhou and H. Liu, "Multi-modal sensing techniques for interfacing hand prostheses: a review," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6065-6076, 2015.
- [34] F. Tenore, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings and N. V. Thakor, "Towards the control of individual fingers of a prosthetic hand using surface EMG signals," in *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, France, 2007.
- [35] K. Ziegler-Graham, E. J. Mackenzie, P. L. Ephraim, T. G. Trivison and R. Brookmeyer, "Estimating the prevalence of limb loss in the united states: 2005 to 2050," *Archives of Physical Medicine and Rehabilitation*, vol. 89, no. 3, pp. 422-429, 2008.

- [36] E. Biddis and T. Chau, "Upper-limb prosthetics: critical factors in device abandonment: a survey of the last 25 years," *American Journal of Physical Medicine & Rehabilitation*, vol. 31, no. 3, pp. 236-257, 2007.
- [37] Touch Bionics Ltd, "i-digits User Manual," 20 November 2018. [Online]. Available: <http://oc4assets.azurewebsites.net/library/39600>. [Accessed 14 May 2020].
- [38] Open Bionics, "Hero Arm User Manual version 4.0," 24 October 2019. [Online]. Available: https://openbionics.com/wp-content/uploads/2020/02/d100161_04-00_Hero-Arm-User-Manual.pdf. [Accessed 19 May 2020].
- [39] Ottobock, "bebionic user guide," 2018. [Online]. Available: https://www.ottobockus.com/media/local-media/prosthetics/upper-limb/files/14112_bebionic_user_guide_lo.pdf. [Accessed 21 May 2020].
- [40] J. Shi, S. Hu, Z. Liu, J. Guo, Y. Zhou and Y. Zheng, "Recognition of finger flexion from ultrasound image with optical flow: a preliminary study," in *Biomedical Engineering and Computer Science (ICBECS)*, Wuhan, China, 2010.
- [41] C. Castellini and G. Passig, "Ultrasound image features of the wrist are linearly related to finger positions," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Francisco, CA, USA, 2011.
- [42] C. A. Baker, N. Akhlaghi, H. Rangwala, J. Kosecka and S. Sikdar, "Real-time, ultrasound-based control of a virtual hand by a trans-radial amputee," in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, 2016.

- [43] S. Patwardhan, A. S. Dhawan, B. Mukherjee, M. Alzamani, W. M. Joiner and S. Sikdar, "Evaluation of the role of proprioception during proportional position control using sonomyography: applications in prosthetic control," in *IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, Toronto, ON, Canada, 2019.
- [44] N. Hettiarachchi, Z. Ju and H. Liu, "A new wearable ultrasound muscle activity sensing system for dexterous prosthetic control," in *IEEE International Conference on Systems, Man, and Cybernetics*, Kowloon, China, 2015.
- [45] N. Akhlaghi, C. A. Baker, M. Lahlou, H. Zafar, K. G. Murthy, H. S. Rangwala, J. Kosecka, W. M. Joiner, J. J. Pancrazio and S. Sikdar, "Real-time classification of hand motions using ultrasound imaging of forearm muscles," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1687-1698, 2016.
- [46] J. McIntosh, A. Marzo, M. Fraser and C. Phillips, "Echoflex: Hand gesture recognition using ultrasound imaging," in *CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017.
- [47] V. Ortenzi, S. Tarantino, C. Castellini and C. Cipriani, "Ultrasound imaging for hand prosthesis control: a comparative study of features and classification methods," in *IEEE International Conference on Rehabilitation Robotics (ICORR)*, Singapore, Singapore, 2015.
- [48] J. Shi, J. Y. Guo, S. X. Hu and Y. P. Zheng, "Recognition of finger flexion motion from ultrasound image: a feasibility study," *Ultrasound in medicine & biology*, vol. 38, no. 10, pp. 1695-1704, 2012.

- [49] D. Sierra González and C. Castellini, "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees," *Frontiers in neurorobotics*, vol. 7, no. 17, pp. 1-11, 2013.
- [50] C. Castellini, "State of the art and perspectives of ultrasound imaging as a human-machine interface," in *Neuro-Robotics*, vol. 2, Dordrecht Heidelberg, Springer, 2014, pp. 37-58.
- [51] Y. Li, K. He, X. Sun and H. Liu, "Human-machine interface based on multi-channel single-element ultrasound transducers: a preliminary study," in *IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Munich, Germany, 2016.
- [52] J. He, H. Luo, J. Jia, J. T. Yeow and N. Jiang, "Wrist and finger gesture recognition with single-element ultrasound signals: a comparison with single-channel surface electromyogram," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 5, p. 1277–1284, 2018.
- [53] X. Yang, J. Yan, Y. Fang, D. Zhou and H. Liu, "Simultaneous prediction of wrist/hand motion via wearable ultrasound sensing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 970-977, 2020.
- [54] C. Castellini, G. Passig and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 788-797, 2012.
- [55] G. Purushothaman and R. Vikas, "Identification of a feature selection based pattern recognition scheme for finger movement recognition from multichannel EMG

- signals," *Australasian physical & engineering sciences in medicine*, vol. 41, no. 2, pp. 549-559, 2018.
- [56] A. Lolure and V. R. Thool, "Wavelet transform based EMG feature extraction and evaluation using scatter graphs," in *International Conference on Industrial Instrumentation and Control (ICIC)*, Pune, India, 2015.
- [57] I. Kuzborskij, A. Gijsberts and B. Caputo, "On the challenge of classifying 52 hand movements from surface electromyography," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 2012.
- [58] S. Abbaspour, M. Lindén, H. Gholamhosseini, A. Naber and M. Ortiz-Catalan, "Evaluation of surface EMG-based recognition algorithms for decoding hand movements," *Medical & Biological Engineering & Computing*, vol. 58, no. 1, p. 83–100, 2020.
- [59] E. Yavuz and C. Eyupoglu, "A cepstrum analysis-based classification method for hand movement surface EMG signals," *Medical & biological engineering & computing*, vol. 57, no. 10, pp. 2179-2201, 2019.
- [60] J. Too, A. R. Abdullah, N. Mohd Saad, N. Mohd Ali and H. Musa, "A detail study of wavelet families for EMG pattern recognition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, pp. 4221-4229, 2018.
- [61] N. Burhan and R. Ghazali, "Feature extraction of surface electromyography (sEMG) and signal processing technique in wavelet transform: A review," in *IEEE*

International Conference on Automatic Control and Intelligent Systems (I2CACIS),
Selangor, Malaysia, 2016.

- [62] S. Negi, Y. Kumar and V. M. Mishra, "Feature extraction and classification for EMG signals using linear discriminant analysis," in *2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Fall)*, Bareilly, India, 2016.
- [63] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2091-2110, 1989.
- [64] G. Lee, R. Gommers, F. Waselwski, K. Wohlfahrt and A. O'Leary, "PyWavelets: a python package for wavelet analysis," *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, 2019.
- [65] N. O. Elssied, O. Ibrahim and A. H. Osman, "A novel feature selection based on one-way anova f-test for e-mail spam classification," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 7, no. 3, pp. 625-638, 2014.
- [66] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg and J. Vanderplas, "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, vol. 12, no. 85, pp. 2825-2830, 2011.
- [67] I. AlMohimeed, H. Turkistani and Y. Ono, "Development of wearable and flexible ultrasonic sensor for skeletal muscle monitoring," in *IEEE Ultrasonics Symposium*, Prague, Czech Republic, 2013.

- [68] I. AlMohimeed, M. Agarwal and Y. Ono, "Wearable Ultrasonic Sensor Using Double-Layer PVDF Films for Monitoring Tissue Motion," in *IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, Quebec City, QC, Canada, 2018.
- [69] A. J. Young, L. H. Smith, E. J. Rouse and L. J. Hargrove, "Classification of Simultaneous Movements Using Surface EMG Pattern Recognition," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 5, pp. 1250 - 1258, 2013.
- [70] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 50, no. 7, pp. 848-854, 2003.
- [71] A. D. Chan and G. C. Green, "Myoelectric control development toolbox," in *CMBES Proceedings*, Toronto, 2007.
- [72] D. Yang, Y. Gu, N. Thakor and L. H, "Improving the functionality, robustness, and adaptability of myoelectric control for dexterous motion restoration," *Experimental Brain Research*, vol. 237, no. 2, pp. 291-311, 2019.
- [73] M. Vetterli and J. Kovačević, "Time-Frequency Representations," in *Wavelets and Subband Coding*, Englewood Cliffs, New Jersey, United States of America, Prentice Hall PTR, 2007, pp. 81-83.
- [74] R. O. Duda, P. E. Hart and D. G. Stork, "Multiple Discriminant Analysis," in *Pattern Classification*, New York, New York, United States of America, John Wiley & Sons, 2001, pp. 121-124.

- [75] I. Alazar Technologies, "AlazarDSO User Guide," 3 December 2015. [Online].
Available: www.alazartech.com. [Accessed 26 July 2020].
- [76] S. van der Walt, S. C. Colbert and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22-30, 2011.
- [77] J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, 2007.