

**Imaging Biomarkers for Duchenne Muscular  
Dystrophy**

by

Sisir Koppaka

M.Tech, Indian Institute of Technology Kharagpur, India  
B.Tech (Hons.), Indian Institute of Technology Kharagpur, India

Submitted to the School of Engineering  
in partial fulfillment of the requirements for the degree of  
Master of Science in Computation for Design and Optimization

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

**Signature redacted**

Author .....

School of Engineering

May 27, 2015

**Signature redacted**

Certified by.....

Brian W. Anthony

Principal Research Scientist, Department of Mechanical Engineering

Thesis Supervisor

**Signature redacted**

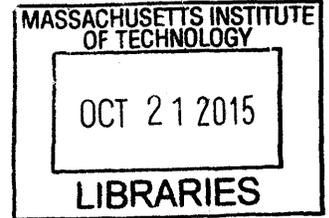
Accepted by .....

Nicolas G. Hadjiconstantinou

Professor, Department of Mechanical Engineering

Co-Director, Computation for Design and Optimization Program

**ARCHIVES**





# Imaging Biomarkers for Duchenne Muscular Dystrophy

by

Sisir Koppaka

Submitted to the School of Engineering  
on May 27, 2015, in partial fulfillment of the  
requirements for the degree of  
Master of Science in Computation for Design and Optimization

## Abstract

Duchenne muscular dystrophy (DMD) is the most common muscular dystrophy of childhood and affects 1 in 3600 male births. The disease is caused by mutations in the dystrophin gene leading to progressive muscle weakness which ultimately results in death due to respiratory and cardiac failure. Accurate, practical, and painless tests to diagnose DMD and measure disease progression are needed in order to test the effectiveness of new therapies. Current clinical outcome measures such as the six-minute walk test and North Star Ambulatory Assessment (NSAA) can be subjective and limited by the patient's degree of effort and cannot be accurately performed in the very young or severely affected older patients. We propose the use of image-based biomarkers with suitable machine learning algorithms instead. We find that force-controlled (precise acquisition at a certain force) and force-correlated (acquisition over a force sweep) ultrasound helps to reduce variability in the imaging process. We show that there is a high degree of inter-operator and intra-operator reliability with this integrated hardware-software setup. We also discuss how other imaging biomarkers, segmentation algorithms to target specific subregions, and better machine learning techniques may provide a boost to the performance reported. Optimizing the ultrasound image acquisition process by maximizing the peak discriminatory power of the images vis-à-vis force applied at the contact force is also discussed. The techniques presented here have the potential for providing a reliable and non-invasive method to discriminate, and eventually track the progression of DMD in patients.

Thesis Supervisor: Brian W. Anthony

Title: Principal Research Scientist, Department of Mechanical Engineering



## Acknowledgments

I would like to thank my advisor, Dr. Brian W. Anthony, for his guidance, and mentorship. Brian's kindness, patience, and inspiration has been invaluable to me.

I would like to thank Matthew W. Gilbertson, whose fantastic force-controlled probe was used for this research. I received extraordinary support from our collaborators at Beth Israel Deaconess Medical Center - Dr. Seward Rutkove, Dr Jim Wu, Irina Shkylar, and Adam Pacheck. I would also like to thank Dr. Anthony Samir and Dr. Manish Dhyani from Massachusetts General Hospital. I owe a lot to the help I received from Shih-Yu Sun, Aaron Zakrzewski, Ina Kundu, Kristi Oki, John Haeseon Lee and Kai E. Thomenius for their support. I am indebted to Aaron Dentinger from GE Research for his help.

I would like to thank Arun Mallya for his support through this time. I am very grateful to Shaily Jaisinghani for her presence in my life. I would like to thank my parents, Suryanarayana and Suseela, and my brother Harsha. They infuse the fun in my life away from work.



# Contents

<b>1</b>	<b>Introduction</b>	<b>15</b>
1.1	Related Work . . . . .	15
1.2	Contributions . . . . .	16
1.3	Thesis Outline . . . . .	17
<b>2</b>	<b>Background</b>	<b>19</b>
2.1	Duchenne Muscular Dystrophy . . . . .	19
2.2	Quantitative Ultrasound . . . . .	21
2.3	Hardware . . . . .	22
2.4	Summary . . . . .	22
<b>3</b>	<b>Imaging Biomarkers</b>	<b>25</b>
3.1	Motivation . . . . .	25
3.2	Methodology . . . . .	26
3.2.1	Patient selection . . . . .	26
3.2.2	Ultrasound Examinations . . . . .	28
3.2.3	US Image and Data Analysis . . . . .	28
3.2.4	Statistical Analysis . . . . .	31
3.3	Results . . . . .	31
3.3.1	Patient Demographics . . . . .	31
3.3.2	US Image and Data Analysis . . . . .	31
3.4	Discussion . . . . .	32
3.5	Summary . . . . .	34

<b>4</b>	<b>Force-controlled Ultrasound</b>	<b>37</b>
4.1	Motivation . . . . .	38
4.2	Methodology . . . . .	38
4.2.1	Data Acquisition . . . . .	39
4.2.2	Edge Detection . . . . .	39
4.2.3	Statistical Sieving . . . . .	39
4.2.4	Decision Tree . . . . .	41
4.2.5	Performance Analysis . . . . .	42
4.3	Results and Discussions . . . . .	42
4.3.1	Clinical Data . . . . .	42
4.3.2	Results . . . . .	43
4.4	Summary . . . . .	47
<b>5</b>	<b>Force-correlated Ultrasound</b>	<b>49</b>
5.1	Motivation . . . . .	49
5.2	Methodology . . . . .	51
5.2.1	Data Acquisition . . . . .	51
5.2.2	Image Set Selection . . . . .	51
5.2.3	Diagnostic Fidelity Enhancement . . . . .	53
5.2.4	Biomarker Quantification . . . . .	55
5.2.5	Clustering . . . . .	55
5.2.6	Performance Analysis . . . . .	56
5.3	Results and Discussion . . . . .	56
5.3.1	Clinical Study . . . . .	56
5.3.2	Results . . . . .	56
5.4	Summary . . . . .	58
<b>6</b>	<b>Conclusion</b>	<b>61</b>
6.1	Contributions . . . . .	62
6.2	Future work . . . . .	63
6.2.1	Imaging Biomarkers . . . . .	63

6.2.2	Optimal Ultrasound Image Acquisition . . . . .	65
6.2.3	Progression Tracking . . . . .	67
6.2.4	Segmentation . . . . .	67
6.2.5	Better Learning Methods . . . . .	70
6.3	Summary . . . . .	73



# List of Figures

2-1	Ultrasound B-Mode image and edge threshold-detected images for a DMD subject (A,C) and Control subject (B,D) respectively. . . . .	20
2-2	Photos of our force-controlling system. The force controlled ultrasound system. Images courtesy Matthew W. Gilbertson. . . . .	23
3-1	Edge detection analysis of gray scale ultrasound image for a Control and a DMD subject. The above edges are depicted at a sensitivity threshold of 0.40. . . . .	27
3-2	Muscle tracing and region of interest used in analysis. (a) Gray scale US image of the quadriceps muscles in a 6 year old control patient. (b) Overlay represents the traced muscle area (white line) excluding the subcutaneous tissue/skin, bone, and deeper tissues. Dotted line and double arrows denote the upper 1/3 of the muscle area used for analysis.	29
3-3	Binary edge detection images at different thresholds. Ultrasound images of the quadriceps muscles in DMD (a) and control subjects (f). The corresponding binary ED images detected at sensitivity thresholds of 0.05 (b and g), 0.1 (c and h), 0.2 (d and i), and 0.4 (e and j) are shown. More edges are detected at the lower thresholds than the higher thresholds. The sensitivity threshold of 0.05 provided the best discrimination between DMD and controls. . . . .	30
4-1	Ultrasound B-Mode image and edge threshold-detected images . . . . .	38

4-2	The force-controlled ultrasound probe maintains a programmable contact force between probe and patient and is used to capture images at a repeatable 2 N of contact force. An LED bar graph positioned near the top of the probe depicts the position of the probe and helps the user keep the device centered within its range of motion. Image courtesy Matthew W. Gilbertson. . . . .	40
4-3	In this figure, ultrasound scans from DMD (a), and control (b) subjects are shown with their corresponding binary maps ((c) and (d) respectively) generated via the Canny edge detection algorithm at a threshold of 0.20. It can be easily seen that when the edges are counted and normalized by area, (d) is expected to have a higher number than (c). We take advantage of this behavior. . . . .	41
4-4	Illustration of a decision tree for the two classes of DMD and Controls based on the edge detection parameter for a particular muscle group. This was visualized in MATLAB. . . . .	42
4-5	This graph shows the number of statistically significant parameters at different levels of significance for each muscle-group. The Quadriceps and Biceps tend to show relatively higher number of significant parameters consistently at multiple levels of significance. . . . .	44
4-6	In this bar chart, we see all four performance metrics for the muscle groups Biceps(B), Deltoids (D), Forearm(F), Medial Gastrocnemius(MG), Quadriceps(Q) and Tibialis Anterior(TA). . . . .	45
4-7	Trends for Accuracy, Precision, Sensitivity and Specificity for different depths of the images. . . . .	46
5-1	(a) Ultrasound scans of the Quadriceps muscle in a subject from the control group (b) The same muscle in a subject with Duchenne Muscular Dystrophy (DMD). DMD-affected muscles exhibit high fat infiltration.	50
5-2	The force-controlled ultrasound probe. Device visualization courtesy Matthew W. Gilbertson. . . . .	52

5-3	Box filter applied to the original image (a) to obtain (b). Multi-scale box-filter image stack from a single frame (c). Pixel-wise standard deviation image of the multiscale stack (d), an example of a control subject EF B-Mode, and (e) which is a DMD subject. (c), (d) and (e) are shown in color to better highlight the enhancement of discriminating features and suppression of irrelevant textures with respect to DMD and Controls. . . . .	54
5-4	Adjusted Rand Indices (ARIs) of the four variations . . . . .	57
6-1	Example of smarter edges, with appropriate morphological operations to discretize and clarify. . . . .	64
6-2	Statistically significant edge thresholds (out of 99) at different forces for Biceps and Quadriceps. Multiple levels of significance are shown. . . . .	66
6-3	Statistically significant edge thresholds (out of 99) at different forces for Deltoid and Medial Gastrocnemius. Multiple levels of significance are shown. . . . .	68
6-4	Image-based Segmentation . . . . .	69
6-5	Hypothesis for inferring muscle thickness on Quadriceps, given Image-based Segmentation . . . . .	71
6-6	Force-axis Segmentation . . . . .	72



# Chapter 1

## Introduction

Ultrasound is a versatile imaging modality. While freehand ability imparts a degree of convenience, sonographer and patient motion-induced image variability serves as a hindrance to developing standardized diagnostic applications. One of the applications that we foresee is the use of ultrasound to assess and track Duchenne Muscular Dystrophy (DMD), and other muscular disorders. DMD is a life-threatening disease that affects male children. It is caused by a genetic mutation, which generates a complex sequence of events in muscle cells, which eventually undergo fibrosis and are replaced by adipose and connective tissue. This results in fatty infiltrations in muscle and eventual death. DMD is evaluated today by using tests like the six-minute walk [26] test, which can easily be confounded by other factors such as the mood of the child, recent sleep, medicine and diet. Therefore, there is a need to develop a non-invasive, highly reliable, and repeatable method by which DMD can be quantitatively evaluated. This would be of enormous interest to the study of DMD-related drugs where measurement of drug efficiency for a particular patient, or set of patients, can be characterized in a less-confounding way.

### 1.1 Related Work

Characterized by progressive disability leading to death, Duchenne Muscular Dystrophy (DMD) remains one of the most common and devastating neuromuscular

disorders of childhood [11]. Although a variety of promising new treatment strategies are in development, outcome measures for clinical trials remain limited for the most part to a set of functional measures, such as the six-minute walk test (6MWT), a measure of the distance that a patient can quickly walk in a period of 6 minutes [10]. Clinical measures that provide high repeatability and sensitivity while still correlating strongly to disease status would find wider use; quantitative ultrasound (QUS) with image-based biomarkers is a technique that could potentially serve in this role [36, 19]. In comparison to traditional QUS techniques where echointensity of the muscle is quantified, we propose the use of image-based biomarkers in tandem with machine learning techniques to quantitatively assess DMD.

We use special force-controlled hardware developed in our group [12]. In current clinical practice, a suitable ultrasound image is acquired by the sonographer based on qualitative assessment of the image. The acquired image, deemed to be sufficiently suitable for further analysis, is often discriminatory in terms of its acquisition position, contact force, and orientation, among different sonographers. This is often a variation accountable to training and experience. Instead, the force-controlled hardware we use for all the clinical data acquisitions performed for this thesis allowed for ultrasound image acquisition at a consistent force, or force sweep pattern. This eliminates the influence of the sonographer in terms of acquisition and provides for a quantitatively reliable and deterministic acquisition process.

## 1.2 Contributions

In this thesis, the clinical utility of force-controlled acquisition of ultrasound images is studied. We study this in the context of Duchenne Muscular Dystrophy.

The contributions include:

- Evaluation of imaging biomarkers that are of relevance to Duchenne Muscular Dystrophy.
- Novel frameworks to study the utility of ultrasound images acquired at pre-

cise forces and across a force sweep for rapid and automated discrimination of Duchenne Muscular Dystrophy

- Analysis of results from a clinical study using these frameworks.
- Discussion of extensions to our work in five directions that could be of significant potential for tracking progression of muscle disorders. This includes using other types of imaging biomarkers, optimizing the ultrasound image acquisition in terms of maximizing the discriminatory power of the images acquired, progression tracking, targeting specific regions by segmenting out images prior to our frameworks and using better machine learning techniques.

### 1.3 Thesis Outline

The remainder of this thesis is organized as follows:

- In Chapter 2, we discuss the background of our work. This includes a discussion of the three topics at which this work intersects - Duchenne Muscular Dystrophy, Quantitative Ultrasound and novel controlled-acquisition hardware for Ultrasound developed in our group.
- In Chapter 3, we introduce imaging biomarkers, and develop a methodology to analyze the relevance of a particular imaging biomarker (edge count). We also discuss variance maps, as an imaging biomarker.
- In Chapter 4, we describe our system for force-controlled ultrasound acquisition. The performance of this system for force-controlled clinical ultrasound imaging is analyzed.
- In Chapter 5, we describe our algorithms for force-correlated ultrasound image analysis. We utilize a set of ultrasound images acquired across a range of forces, for clinical analysis. Clinical results are presented.
- In Chapter 6, we discuss our learnings and insights, as well as opportunities for future work. We describe future work that should be explored in five directions.



# Chapter 2

## Background

In this chapter, we briefly review prior research relevant to our thesis. Specifically, we focus on material relevant to Duchenne Muscular Dystrophy, Quantitative Ultrasound and our in-house force-controlled probe hardware.

### 2.1 Duchenne Muscular Dystrophy

Duchenne muscular dystrophy (DMD) is one of the most common muscular dystrophy of childhood. It affects nearly 1 in 3600 male births [11]. The disease is caused by mutations in the dystrophin gene leading to progressive muscle weakness which ultimately results in death due to respiratory and cardiac failure [16, 5]. Accurate, practical, and painless tests to diagnose DMD and measure disease progression are needed in order to test the effectiveness of new therapies [7]. Current clinical outcome measures such as the six-minute walk test and North Star Ambulatory Assessment (NSAA) can be subjective and limited by the patient's degree of effort and cannot be accurately performed in the very young or severely affected older patients [25, 24, 26].

In Fig. 2-1, we show the ultrasound images of a DMD subject in (A) and a control subject in (B) respectively. The corresponding images identified by an imaging biomarker, the edge detection method, are shown in (C) and (D). In all the images, the subcutaneous fat, and the muscle are highlighted in green and red polygonal annotations respectively. For the control subject, we are able to identify the bone as

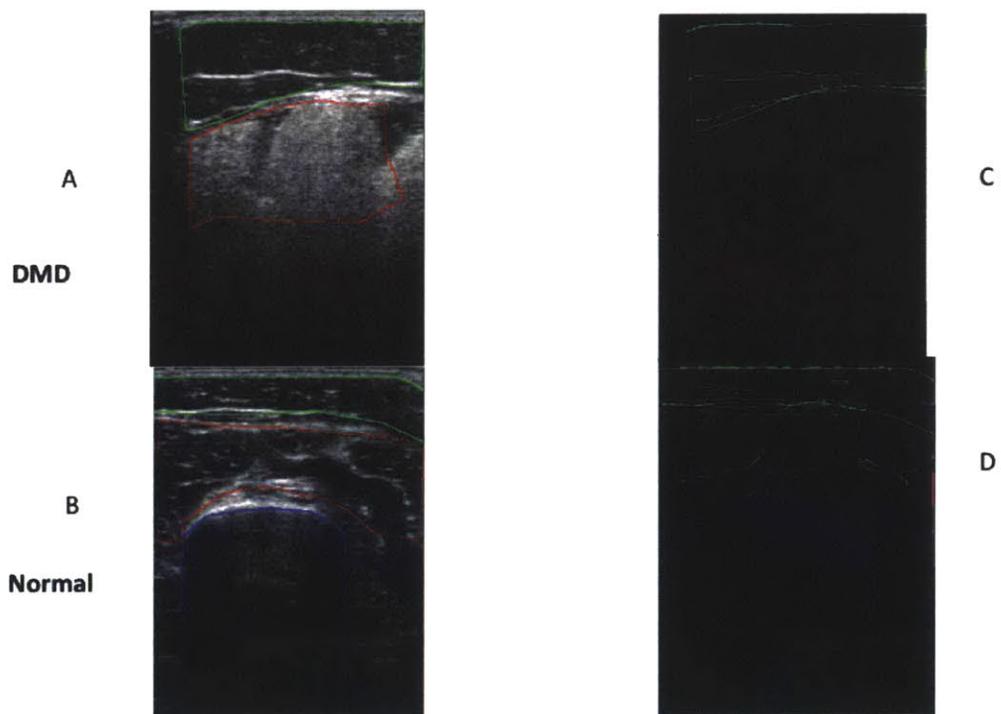


Figure 2-1: Ultrasound B-Mode image and edge threshold-detected images for a DMD subject (A,C) and Control subject (B,D) respectively.

well, which is highlighted in the blue polygonal region. Note that we can't see the bone in the DMD subject.

## 2.2 Quantitative Ultrasound

Ultrasound is one of the most popular non-invasive imaging modalities. However, unlike CT or MRI, ultrasound probe contact force deforms tissue and alters image appearance. Sonographer experience is known to be a critical factor for improving diagnosis capability. Techniques such as elastography have taken advantage of pre- and post-compression behavior to evaluate properties of tissues, and their diagnostic uses [27].

Quantitative ultrasound (QUS), assisted by image-based biomarkers, could be a painless, easy to use, and reliable test that is not affected by patient effort. Recent studies have shown that gray scale luminosity (GSL) can distinguish diseased from healthy muscle in several childhood neuromuscular disorders [35, 36, 19, 3, 4]. In many neuromuscular disorders, including DMD, intramuscular fibrosis and fatty infiltration occurs and will increase the echointensity of the muscle on gray scale ultrasound images [29]. The echointensity of the muscle can be quantified using many commercially available software programs which convert the ultrasound image into a distribution of gray scale values. Patients with DMD will have higher muscle echointensity when compared to normal controls.

However, the current approach of QUS in using the gray scale luminosity (GSL) has certain drawbacks. For example, the GSL is known to vary with the manufacturer of the ultrasound, and with the settings at which the ultrasound image was acquired (gain, etc.). In literature, certain approaches to deal with this, such as calibration the back scatter [36], have been discussed. The approach pursued in this thesis, that of imaging biomarkers, should be viewed as pursuing the same goal in a complementary manner, in that, the imaging biomarker is intended to provide a robust and settings-invariant assessment of the image.

## 2.3 Hardware

All ultrasound examinations were performed with a portable Terason t3000 system (Teratech Inc., Burlington, MA) with a 10 MHz probe. This probe was equipped with the controlled-acquisition hardware developed by Gilbertson [12]. This device is a handheld electro-mechanically actuated ultrasound probe, capable of applying a programmable force to the contact surface. The programmable force may be a constant force, which the device achieves by adjusting for the operator's hand movement, or a force sweep across a force range. In later chapters, we will continue to further describe the operation of this device, as relevant to the line of inquiry being pursued.

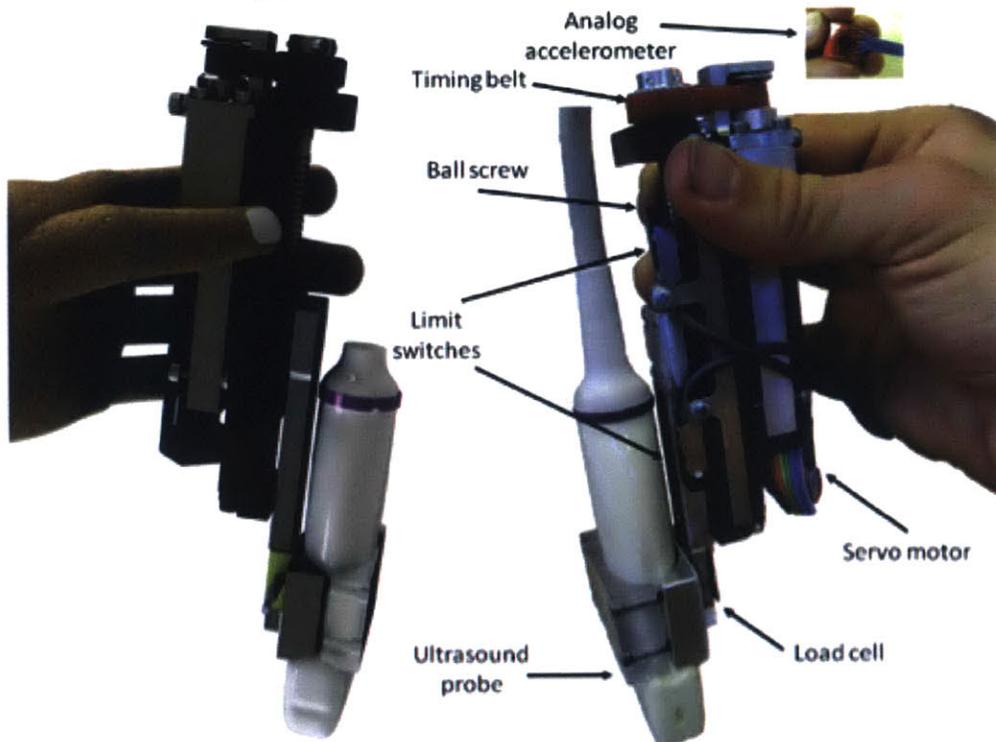
In Fig. 2-2, we show Gilbertson's [12] force controlled acquisition setup in (a) and the force controlling device in (b).

## 2.4 Summary

In this chapter, we briefly reviewed relevant prior research on Duchenne Muscular Dystrophy and Quantitative Ultrasound techniques for disease assessment. We also introduced the handheld force-controlled acquisition hardware.



(a) The force controlled acquisition setup



(b) The probe system. The force-controlling device is shown here without the ultrasound probe or protective cover.

Figure 2-2: Photos of our force-controlling system. The force controlled ultrasound system. Images courtesy Matthew W. Gilbertson.



# Chapter 3

## Imaging Biomarkers

In this chapter, we discuss ultrasound imaging biomarkers, and study one specifically, based on edge detection and enumeration, that we use in further chapters.

### 3.1 Motivation

The increased echogenicity of DMD muscle tends to obscure muscle fascia structure. Therefore, we would like to investigate the utility of edge detection image analysis to assess patients with Duchenne muscular dystrophy (DMD).

Edge detection (ED) is an image processing tool which can highlight textural and high frequency intensity changes within an image (Fig. 3-1). Pixel boundaries where the image brightness changes abruptly are edges. Edges within an image can be used to highlight the boundaries in an image using an algorithm developed by Canny [9]. By adjusting the threshold for the detection of edges, one can assess textural differences across different size structures within the image. Lower Canny threshold distinguish small structural differences whereas higher thresholds distinguish larger structures. In skeletal muscle, fibrous tissue is present within the muscle and between muscles forming intramuscular and intermuscular septa, respectively. As the muscle becomes impacted by DMD, the echogenic septa become more difficult to visualize as the entire muscle becomes more echogenic due to fibrosis and fatty infiltration [29]. The increase in muscle echogenicity will cause the normal structural pattern of the

muscle to become less discernible and should be quantifiable by an increase in the number of edges in the ultrasound image.

In Fig. 3-1, we show the edge detection analysis of typical gray scale ultrasound images from a control and a DMD subject. In (a), we see the gray scale ultrasound image of the quadriceps muscle control subject and in (b), we see a binary image using edge detection Canny analysis of (a). (c) and (d) are the gray scale US image of the quadriceps muscle in a DMD subject and the corresponding binary edge detection image. We note that the echogenic intermuscular and intramuscular septa on the gray scale image appear as bright lines and curves on the binary image. There is also a difference in the edges (corresponding to underlying muscular fascia and other image features) that are visible at a particular threshold for each of the subgroups.

It is our hypothesis that patients with DMD with have significantly more edges in their muscles when compared to normal boys due to increase muscle echogenicity and poorer visualization of the muscle septa. This is because of the higher frequency of grayscale echogenicity variation in the DMD patients as compared to the control subjects. The purpose of this study is to investigate the ability of quantitative ultrasound, using edge detection analysis, to distinguish patients with Duchenne muscular dystrophy (DMD) from normal subjects in the hope that it could serve as a useful measure of disease status and the effect of therapy.

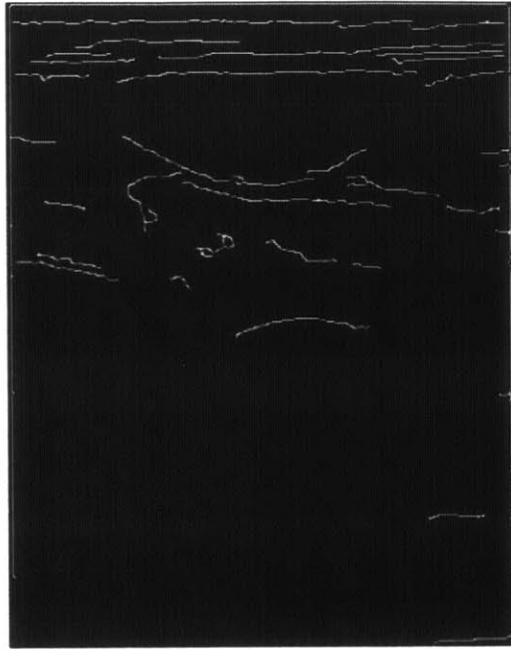
## **3.2 Methodology**

### **3.2.1 Patient selection**

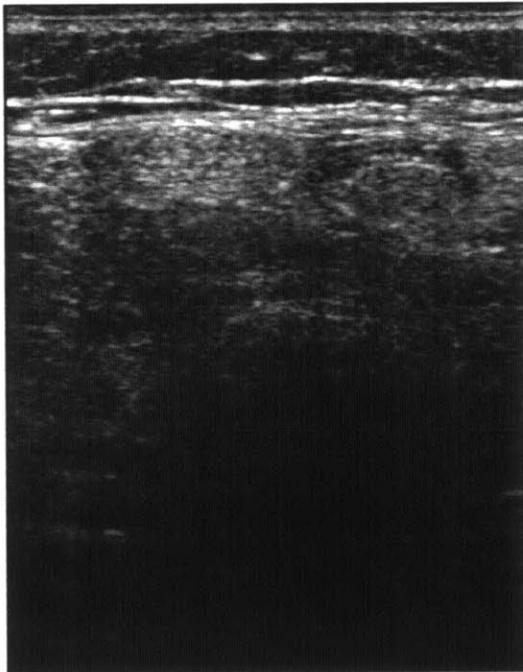
The institutional review board of Boston Children's Hospital approved this prospective study. Informed written consent and verbal assent were obtained, respectively, from parents and children. Boys with DMD were recruited and enrolled in the study through a neuromuscular disorders clinic. They had genetic mutations and clinical presentation consistent with DMD. Boys with DMD were excluded if they were involved in an ongoing clinical therapeutic trial or if they had another neuromuscular



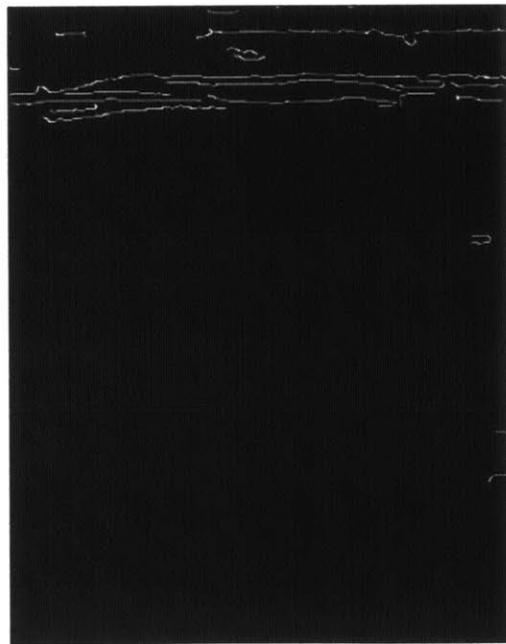
(a)



(b)



(c)



(d)

Figure 3-1: Edge detection analysis of gray scale ultrasound image for a Control and a DMD subject. The above edges are depicted at a sensitivity threshold of 0.40.

or other medical condition that substantially impacted health. Healthy subjects were recruited by advertisement and via family members and did not have a history of neuromuscular disease or other disease that would substantially impact health.

### **3.2.2 Ultrasound Examinations**

Transverse US images of six muscles (biceps brachii, deltoid, wrist flexors, quadriceps, tibialis anterior and medial gastrocnemius) on the patient's dominant side were obtained. Dominance was determined by asking the child or parent; and when unknown, the child was given a ball to throw in order to assess his dominant side. US settings (gain, compression, time gain compensation, and depth) were kept constant for all image acquisitions similar to past studies [35, 37, 30]. Research assistants, trained by a musculoskeletal radiologist, obtained all US images. US images were obtained with the subject seated with the knee bent at 90° and the arm extended at mid-chest height with the elbow straight and supported by the examiner or a pillow.

### **3.2.3 US Image and Data Analysis**

US images were exported from the Terason software to MATLAB as tagged image file format files. The muscle of interest was outlined using a polygonal-region tracing tool by a single musculoskeletal radiologist. The area of the muscle, measured in number of pixels, was calculated from the traced image. The pixel pitch was held constant by choosing the same physical parameters for image length and width during acquisition. Only the upper one third of the traced muscle area was used for analysis (Fig. 3-2), as described by Jansen et al. [19], since there is attenuation of sound waves in the deeper tissue making analysis of deeper structures in the image less reliable. Edge detection (ED) values were quantified by using the Canny edge detection algorithm [9] on MATLAB. The Canny edge detection algorithm detects edges in the image, and produces a binary map of the edges. The number of edges, measured in edge pixels highlighted by the Canny edge detection algorithm, present in each muscle region (upper one third) was then divided by the number of pixels in that same area

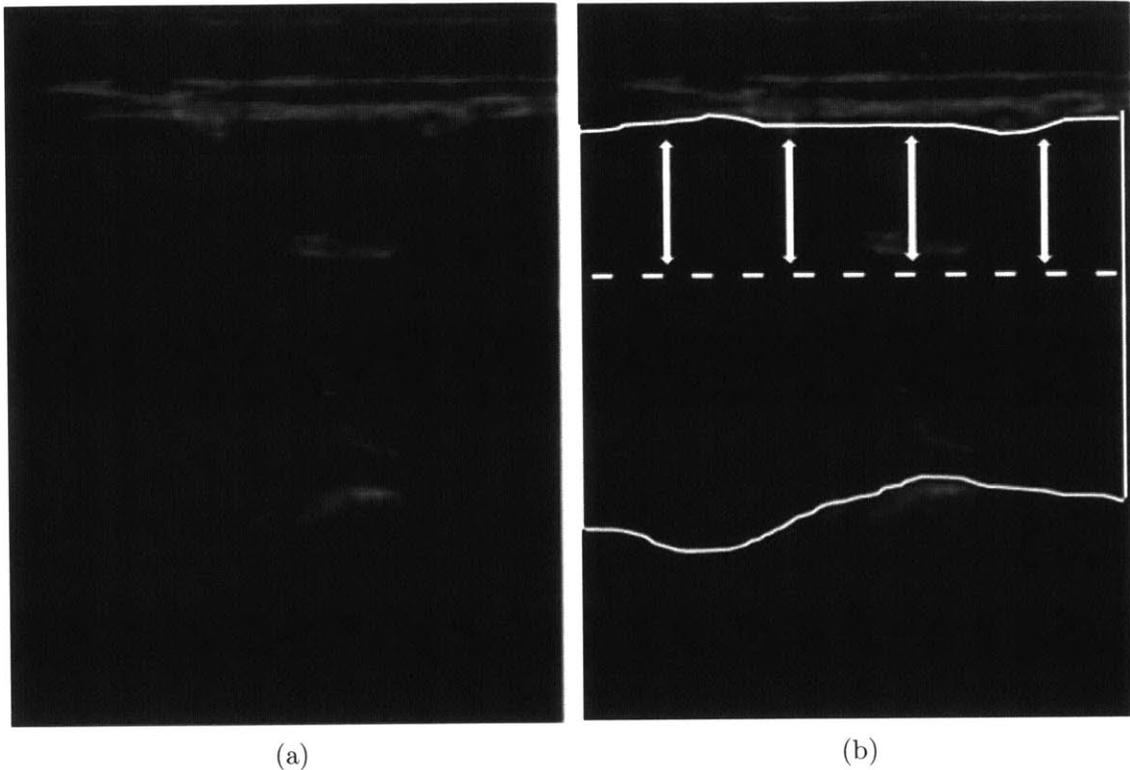


Figure 3-2: Muscle tracing and region of interest used in analysis. (a) Gray scale US image of the quadriceps muscles in a 6 year old control patient. (b) Overlay represents the traced muscle area (white line) excluding the subcutaneous tissue/skin, bone, and deeper tissues. Dotted line and double arrows denote the upper 1/3 of the muscle area used for analysis.

of muscle to arrive at the ED value (unitless value). This gives a normalized measure of the number of edges in the muscle region being evaluated.

The ED values for all six muscle groups were determined at multiple thresholds of sensitivity using the Canny algorithm. The optimal threshold to distinguish DMD from normal was also determined. The detector takes a threshold parameter for sensitivity, which can vary from 0.01 to 0.99. We generated a Canny binary map for 99 thresholds between 0.01 to 0.99 at 0.01 intervals (Fig. 3-3). This allows us to capture the number of edges detected per muscle area, and the behavior of each muscle as a function of edge thresholds.

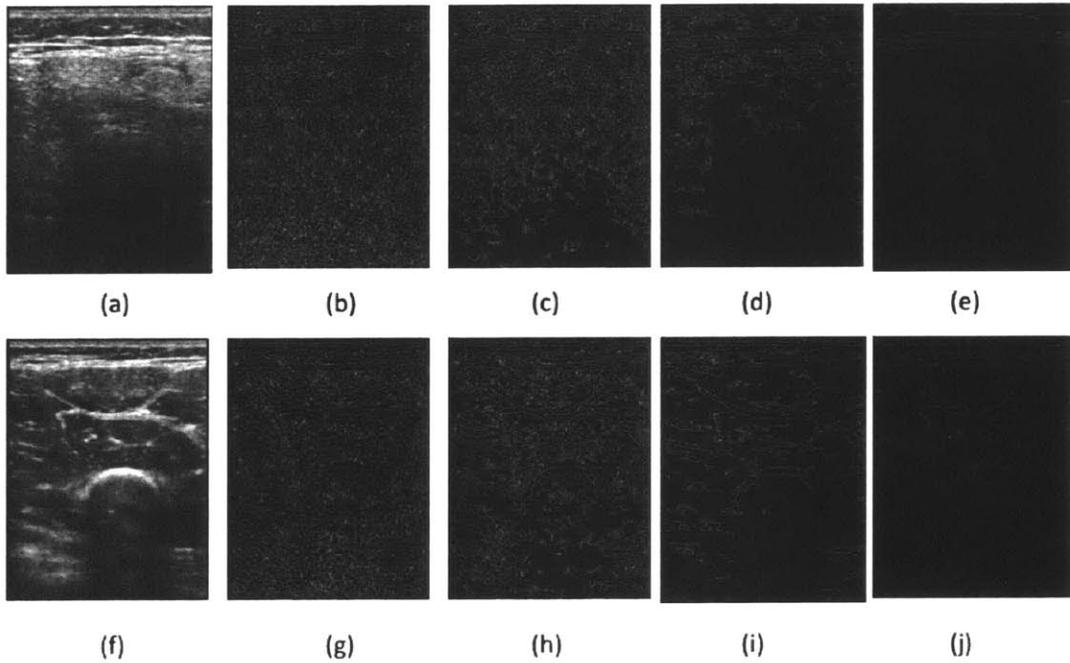


Figure 3-3: Binary edge detection images at different thresholds. Ultrasound images of the quadriceps muscles in DMD (a) and control subjects (f). The corresponding binary ED images detected at sensitivity thresholds of 0.05 (b and g), 0.1 (c and h), 0.2 (d and i), and 0.4 (e and j) are shown. More edges are detected at the lower thresholds than the higher thresholds. The sensitivity threshold of 0.05 provided the best discrimination between DMD and controls.

### **3.2.4 Statistical Analysis**

The receiver operating curve (ROC) was constructed by plotting the true positive rate against the false positive rate for various thresholds. The area under the receiver operating curve (AUC) values of ED thresholds were generated for each muscle and the average of all six muscles using MedCal. Mann-Whitney test was used to determine differences in edge detection values between DMD boys and normal subjects with a p-value of 0.05 considered significant.

## **3.3 Results**

We now discuss the results of our investigation into the Edge Detection (ED) biomarker.

### **3.3.1 Patient Demographics**

The average age of the DMD group was 8.8 years (range 3.0-14.3 years) and 8.7 years (range 3.4-13.5 years) for the control group. Of the 19 boys with DMD, 9 were on corticosteroid therapy and 11 were in the 4-10 age group. Of the 21 controls 12 were in the 4-10 age group.

### **3.3.2 US Image and Data Analysis**

For edge detection, amongst the 99 Canny sensitivity thresholds tested (0.01-0.99), a Canny sensitivity threshold of 0.05, was the optimal threshold to distinguish DMD patients from normal. We define the optimal threshold as the one that maximizes the difference between the two subject groups. Thus, this threshold was used for all subsequent analyses. Using the average of the 6 muscles, edge detection was excellent at distinguishing DMD from normal with an AUC=0.96. For the individual muscles, edge detection was best at distinguishing DMD from normal in the gastrocnemius (AUC=0.97) and poorest in the anterior tibialis (AUC=0.80), as listed in the table below.

Muscle	ED (Control)	ED (DMD)	AUC	95% CI	Std. Err	p-value
Biceps United	0.213	0.234	0.91	0.78-0.98	0.048	<0.0001
Deltoid	0.207	0.223	0.86	0.71-0.95	0.072	<0.0001
Wrist flexors	0.204	0.223	0.92	0.78-0.98	0.051	<0.0001
Quadriceps	0.215	0.231	0.92	0.79-0.98	0.055	<0.0001
Gastrocnemius	0.207	0.234	0.97	0.86-1.00	0.020	<0.0001
Tibialis Anterior	0.212	0.226	0.80	0.64-0.91	0.074	<0.0001
All 6 muscles	0.210	0.228	0.96	0.84-1.00	0.029	<0.0001

For each of the 6 muscles, there were more edges detected in the DMD group when compared to normal controls. This was also true when all 6 muscle groups were averaged together: DMD had an ED value = 0.228, while normal boys had an ED value = 0.210,  $p < 0.0001$ .

### 3.4 Discussion

There is a need for accurate biomarkers capable of measuring disease progression over time and evaluating drug efficacy [8]. In this prospective study, our results suggest that quantitative ultrasound analysis using edge detection is a potentially useful biomarker as this technique was capable of distinguishing between boys with DMD and controls with high accuracy. In this study, we found, as hypothesized, that the muscle of patients with DMD had more edges on the US images when compared to controls. With disease progression, the muscles in DMD patients become infiltrated with echogenic fibrous tissue and fat which obscures the similarly echogenic intramuscular and intermuscular septa [29].

Edge detection is a fundamental technique in the field of image processing and has been in use for decades, especially for feature detection; however, its use in medicine is relatively limited [9, 38, 13]. Edge detection has been used effectively in clinical medicine to assess the size of cardiac chambers, prostate gland size, and intima-media thickness in blood vessels from ultrasound images [38, 1, 14, 23, 28]. Abrupt edge changes may indicate the boundaries of anatomic structures and are used to determine their area or volume.

Edge detection can also assess textural variations within the object. It is this

value that we believe is most advantageous when analyzing ultrasound images of muscles. By using edge detection, we were able to distinguish DMD patients from normal similar to other quantitative ultrasound studies which have relied primarily on variations in muscle echointensity [19, 37, 30]. In these echointensity studies, a region of interest (ROI) box is placed inside the muscle, or tracing of the muscle of interest is performed and gray scale echointensity values are generated [19, 37, 30]. In general, diseased muscle has higher echointensity than normal muscle due to deposition of fibrous tissue and fat, and this has been shown to occur for a variety of diseases including DMD [35, 19, 33]. It is unclear if the increase in edges seen on this study and increase echointensity values seen in past studies of DMD patients are detecting corresponding structural changes in the muscle.

Edge detection has the ability to selectively evaluate different components of muscle which is not entirely possible with echointensity analysis. In edge detection, lower Canny sensitivity thresholds correspond to edges of smaller dimensions (length) and are more representative of smaller components of muscle such as the muscle fascicles and intramuscular septa, whereas the higher sensitivity thresholds correspond to longer edge dimensions and larger structures such as the intermuscular fascia. By adjusting sensitivity thresholds, one can assess differences in the various structural components of the muscle. In this study, edge detection was best at the lower Canny thresholds for distinguishing DMD from normal patients, suggesting that the muscle changes can be attributed to the smaller-sized edges and therefore the small components of muscle.

Edge detection analysis was able to distinguish between DMD and controls for each of the six muscles and with the average of the 6 muscles. Moreover, edge detection performed best when evaluating the gastrocnemius. This result is supported by past imaging studies using CT and MRI which have shown that muscle atrophy with fatty infiltration is most pronounced in the posterior as opposed to the anterior calf muscles [32, 34, 2]. In fact, the anterior compartment muscles are often normal on imaging even with long-standing disease [2, 34]. Clinically, calf pseudohypertrophy is a characteristic finding in DMD, where despite circumferential enlargement of the

calf from fat deposition, muscle weakness is present [7].

A few limitations deserve mention. First, this is a cross-sectional study and the sample size is small with roughly 20 patients. Future longitudinal studies with larger number of subjects will assess if edge detection has the ability to monitor disease progression and efficacy of therapies. Another limitation is the inability to change US parameters during image acquisition. Altering the depth, gain, or focus could affect the appearance of the final image and the number of edges present. However, the settings were identical for all muscles and patients and this technique has been shown to be effective for several neuromuscular disorders and research groups [35, 19, 30, 29, 19]. Using raw frequency or backscatter data could potentially correct for this issue [36, 33].

In conclusion, quantitative ultrasound using edge detection analysis was able to distinguish patients with DMD from healthy controls with excellent accuracy. Future studies are needed to determine if ED by itself or in combination with other tests can improve the assessment of disease progression or drug efficacy.

### 3.5 Summary

Following institutional IRB-approval, ultrasound exams with fixed technical parameters were performed unilaterally in 6 muscle groups (biceps, deltoid, wrist flexors, quadriceps, medial gastrocnemius, and tibialis anterior) in 19 boys with DMD and 21 age-matched controls. The muscles of interest were outlined using a tracing tool (MATLAB), and the upper one-third of the muscle was used for analysis. Edge detection (ED) values for each muscle were quantified by using the Canny edge detection algorithm and then normalizing the number of edge pixels to the area of the muscle. The ED values were extracted at multiple thresholds of sensitivity using the Canny algorithm to determine the optimal threshold (0.01-0.99) to distinguish DMD from normal. Area under the receiver operating curve (AUC) values were generated for each muscle and averaged across the six muscles.

The average age of the DMD group was 8.8 years (range 3.0-14.3y) and 8.7 years

(range 3.4-13.5y) for the control group. For edge detection, a Canny threshold of 0.05 provided the best discrimination between DMD and controls (AUC of 0.96; 95% CI 0.84-1.00). Using a Mann-Whitney test, ED values were significantly different between boys with DMD and controls ( $p < 0.0001$ ).

Quantitative ultrasound using edge detection can distinguish patients with DMD from healthy controls at low Canny thresholds where discrimination of small structures is best. ED by itself or in combination with other tests has the potential to serve as a useful biomarker of disease progression and the effect of therapy.



# Chapter 4

## Force-controlled Ultrasound

In this chapter, we discuss a technique for quantitative discrimination of Duchenne Muscular Dystrophy (DMD). Our ultrasound image data is generated with a novel force-controlled ultrasound acquisition system [12] that allows precise ultrasound image acquisition at a predetermined force. We use the texture of ultrasound images, as calculated by the Canny edge detector, as the input image feature for our analysis algorithm. After statistically sieving (selecting only those that pass through our statistical filters) through the edge detection parameters on our training set, we identify the set of parameters significant within a threshold. Decision trees are then trained on these significant parameters over a training dataset with cross-validation, and evaluated on accuracy, precision, selectivity and sensitivity on a separate test dataset. We discuss the performance of our system, by muscle groups, on data collected with our device in a clinical study. Using depth of the image as a region of interest selection mechanism, we evaluate the extent to which the performance of our system is robust to image depth. Our results support that automated assessment of Duchenne Muscular Dystrophy using force-controlled ultrasound image acquisition is possible in a reliable and robust manner.

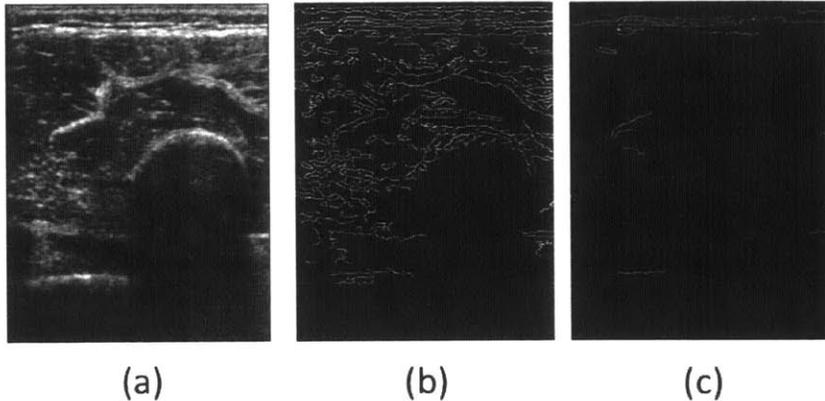


Figure 4-1: Ultrasound B-Mode image and edge threshold-detected images

## 4.1 Motivation

Gilbertson [12] developed a force-controlled probe that allows with reliable ultrasound image acquisition with low variability. We propose a method wherein we use this device that is capable of measuring the contact force applied with the ultrasound probe, and capturing the ultrasound image at a predetermined force to a high degree of precision. The device also measures the angle of contact, so that the image is captured perpendicular to the long axis of the limb or skin surface. Using edge-features related to the image texture(see Fig. 4-1), we train decision trees and test its performance on a variety of performance measures. All data for this study was collected using this controlled-force system.

In Fig 4-1, we show an ultrasound image (a), with the corresponding edges detected at two thresholds, 0.25 (b) and 0.5 (c). As can be seen, only the longer/more significant edges remain consistently visible at the higher threshold in (c). We use the information captured at each threshold, as well as across multiple edge detection thresholds to assess DMD.

## 4.2 Methodology

In this section, we briefly describe our data acquisition method, the image features on which we chose to train, the decision tree classifier that we use, and the metrics

with which we analyze system performance.

### 4.2.1 Data Acquisition

All ultrasound images were captured at 2 Newtons (N) of contact force using the force-controlled ultrasound probe [12]. We chose 2 Newtons as it is a typical contact force experienced in regular clinical study. The device, which is grasped by the user and placed in contact with the patient (Fig. 4-2), linearly actuates an ultrasound probe to maintain a programmable probe contact force between the user's hand and the patient, and accommodates up to 6 cm of relative motion. In this study, the contact force was maintained at 2.0 N with an accuracy of  $\pm 0.1N$ .

### 4.2.2 Edge Detection

We chose the Canny edge detector as the feature for the ultrasound images [9]. The Canny edge detection algorithm detects edges in the image, and produces a binary map of the edges (see Fig. 4-3). The detector takes a threshold parameter for sensitivity, which can vary from 0.01 to 0.99. We generate the Canny binary map for 99 thresholds from 0.01 to 0.99 at 0.01 intervals. This allows us to evaluate not only the number of edges detected per image, and the behavior of the image as a function of edge thresholds.

The edges detected in each image via the Canny binary map are then normalized by dividing the sum of pixels identified as part of an edge by the total size of the image in pixels. This gives us a normalized measure of the edges detected in the image.

### 4.2.3 Statistical Sieving

In order to down select from the 99 edge detection parameters, we perform a two-sample two-tailed t-test with unequal variance. In our results, we discuss the number of parameters that we found significant at different levels of significance. But the

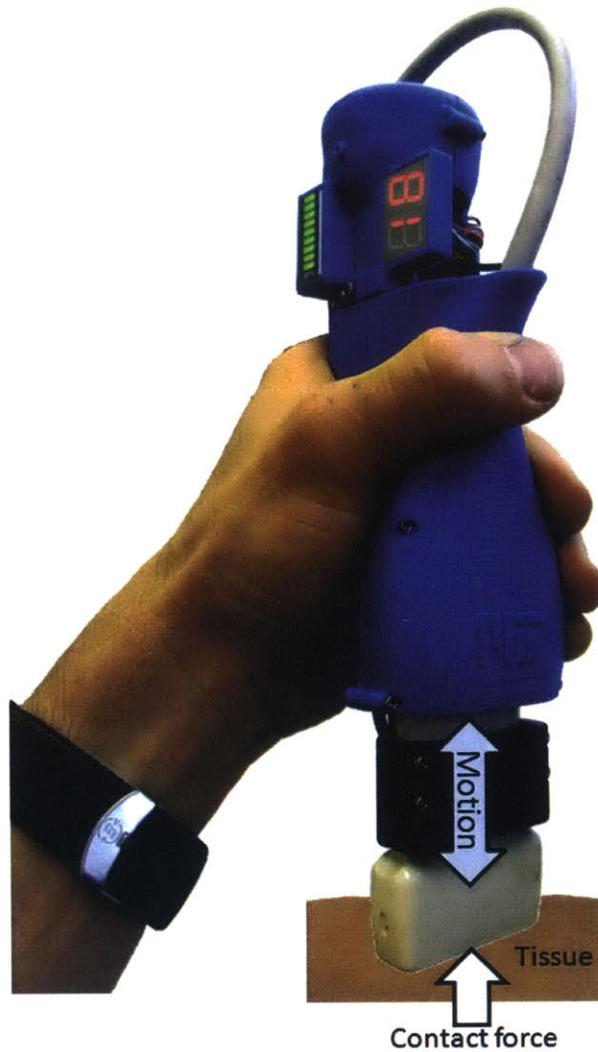


Figure 4-2: The force-controlled ultrasound probe maintains a programmable contact force between probe and patient and is used to capture images at a repeatable 2 N of contact force. An LED bar graph positioned near the top of the probe depicts the position of the probe and helps the user keep the device centered within its range of motion. Image courtesy Matthew W. Gilbertson.

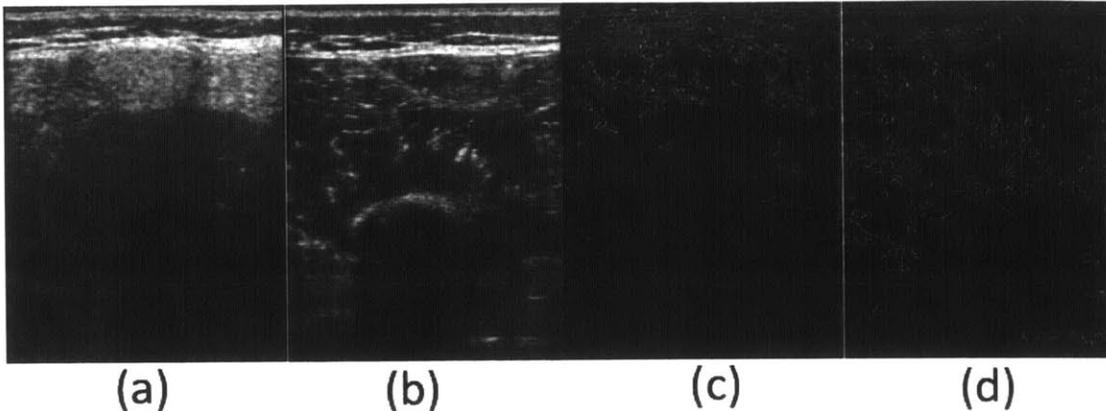


Figure 4-3: In this figure, ultrasound scans from DMD (a), and control (b) subjects are shown with their corresponding binary maps ((c) and (d) respectively) generated via the Canny edge detection algorithm at a threshold of 0.20. It can be easily seen that when the edges are counted and normalized by area, (d) is expected to have a higher number than (c). We take advantage of this behavior.

rest of our system was run with the parameters determined to be significant with a p-value  $< 0.10$ .

#### 4.2.4 Decision Tree

We use the decision tree algorithm to train on our training set with 3-fold cross-validation [6]. The best split in the tree is computed via exact search, as the number of classes is just 2 (DMD and controls). The splitting criterion used was the Gini's Diversity Index, which is given by,

$$1 - \sum_i p(i)$$

for each node. Here  $p(i)$  represents the observed fraction of classes reaching the particular node, while the summation occurs over the set of classes at the node (DMD, Control).

Decision trees have the ability to be easy to understand and visualize. Therefore, it is possible for a clinician to observe the actual methodology by which our system has used the image features to reach the conclusion. In Fig. 4-4, we depict a particular decision tree generated for one of the muscle groups for illustration.

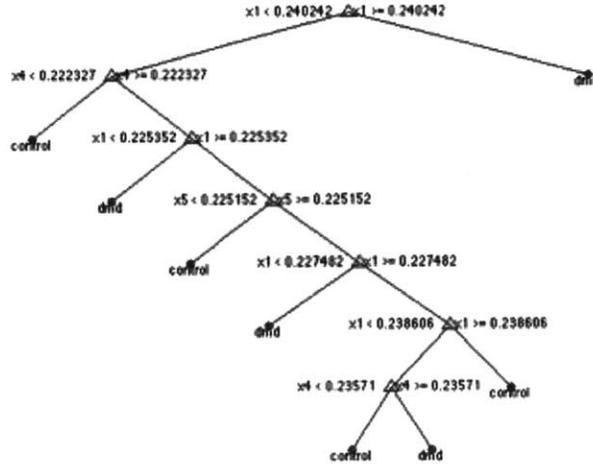


Figure 4-4: Illustration of a decision tree for the two classes of DMD and Controls based on the edge detection parameter for a particular muscle group. This was visualized in MATLAB.

## 4.2.5 Performance Analysis

The decision trees were trained on a training set. They were evaluated on a separate test set, on four major parameters, viz., accuracy, precision, sensitivity, and specificity. In addition, we also evaluated the robustness of our choice of using the entire image instead of a selective region-of-interest. We discuss this further below, including the composition of the training and test sets.

## 4.3 Results and Discussions

### 4.3.1 Clinical Data

This clinical study was approved by the Children’s Hospital Boston Institutional Review Board. All parents/subjects were required to give signed informed consent and assent. During each visit, ultrasound images were taken precisely at 2N on six muscle groups, viz., Quadriceps (Q), Biceps (B), Deltoids (D), Forearm (F), Tibialis Anterior (TA) and the Medial Gastrocnemius (MG).

The BOSTON II dataset for this work consists of a training set and a test set. The dataset was organized from the clinical study with the following goals in mind.

No individual repeats within the training set, i.e., one visit per individual. The same rule applied to the test set. The same individual could appear in both the training and test set, provided each of them was based on data collected from different visits. This ensured that there were variations induced by the uncertainty in absolute spatial positioning of the probe by the operator on the muscle group, as well as the changes due to time. Due to the nature and size of collected data and available patients, this scheme allowed us to maximize the temporal variation between the training and test datasets, while introducing in the training set as high a degree of individuals not present in the test. This reduced the chances of overfitting by individual type as much as possible under the constraints of limited data.

The training set consists of 40 unique individuals, with 18 from the DMD group and 22 from the control group. The test set consists of 10 DMD subjects and 17 control subjects.

### 4.3.2 Results

We now discuss the results of this system in assessing DMD. For brevity, we will refer to the muscles by their abbreviated name.

First, we evaluated the number of significant parameters (99 from 0.01 to 0.99 at 0.01 intervals) for Canny edge detection on the training set. We evaluated this on a per-muscle group basis. The various levels of significance considered were 0.0001, 0.001, 0.01, 0.05 and 0.1. From Fig. 4-5, we noted that Quadriceps and Biceps showed relatively higher number of significant parameters at multiple levels of significance. We have used the set of parameters with a p-value below the level of significance of 0.1 as input to the decision tree.

Decision trees were constructed on these chosen statistically relevant parameters on the training set, and then evaluated by the performance measures over the test set. We note that the ultrasound image capture being automated via force-control by the hardware device, these performance measures represent an assessment of DMD which reduces human induced variability. Our results are depicted in Fig. 4-6. Accuracy represents the proportion of predictions on the test set that were correct based on

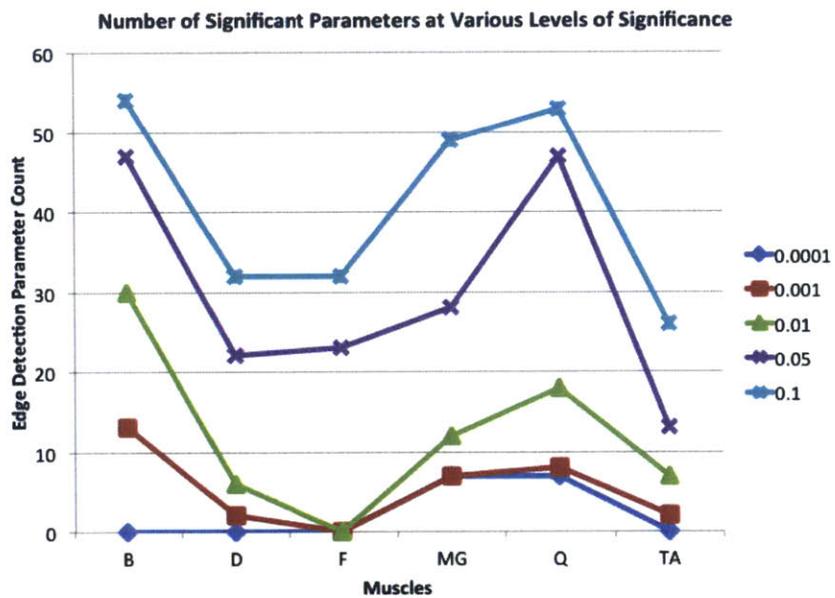


Figure 4-5: This graph shows the number of statistically significant parameters at different levels of significance for each muscle-group. The Quadriceps and Biceps tend to show relatively higher number of significant parameters consistently at multiple levels of significance.

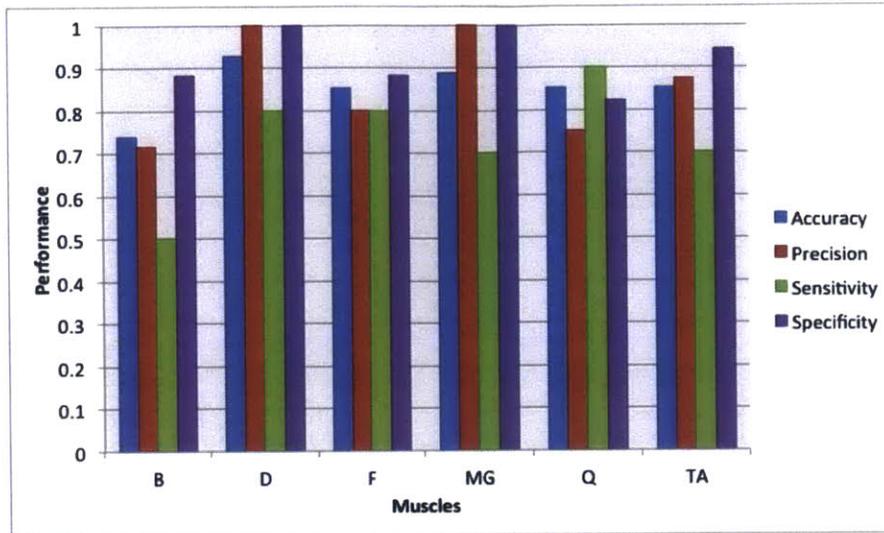


Figure 4-6: In this bar chart, we see all four performance metrics for the muscle groups Biceps(B), Deltoids (D), Forearm(F), Medial Gastrocnemius(MG), Quadriceps(Q) and Tibialis Anterior(TA).

known patient labels. The overall accuracy values for Q, TA, MG, F, D are high at 0.85-0.93, with B being slightly lower at 0.74.

Precision represents the proportion of positive DMD predictions that are correct among the study subjects eventually predicted as positive DMD. For our test set, we note that MG and D reported 1.0 on this metric, with the rest being Q(0.75), TA(0.88), F(0.8) and B(0.71). We observe that for the given image feature (edge detection), MG and D appear to be highly discriminating towards DMD, although the rest of them perform well in an absolute sense as well. This implies that if a prediction that a certain subject was DMD were to be made by using different muscle groups with our system, MG and D would have the highest reliability.

Sensitivity denotes the proportion of positive DMD labeled subjects who were predicted as positive among the actually positive DMD population. We noted sensitivities of Q(0.90), F(0.80), D(0.80), MG (0.70), B (0.50) and TA(0.70). This implies that if a study subject has DMD, the most reliable muscle groups to be used with our system for DMD assessment are Q followed by F, D and TA. Note that all the values are high in an absolute sense, but we seek to continue differentiating among them.

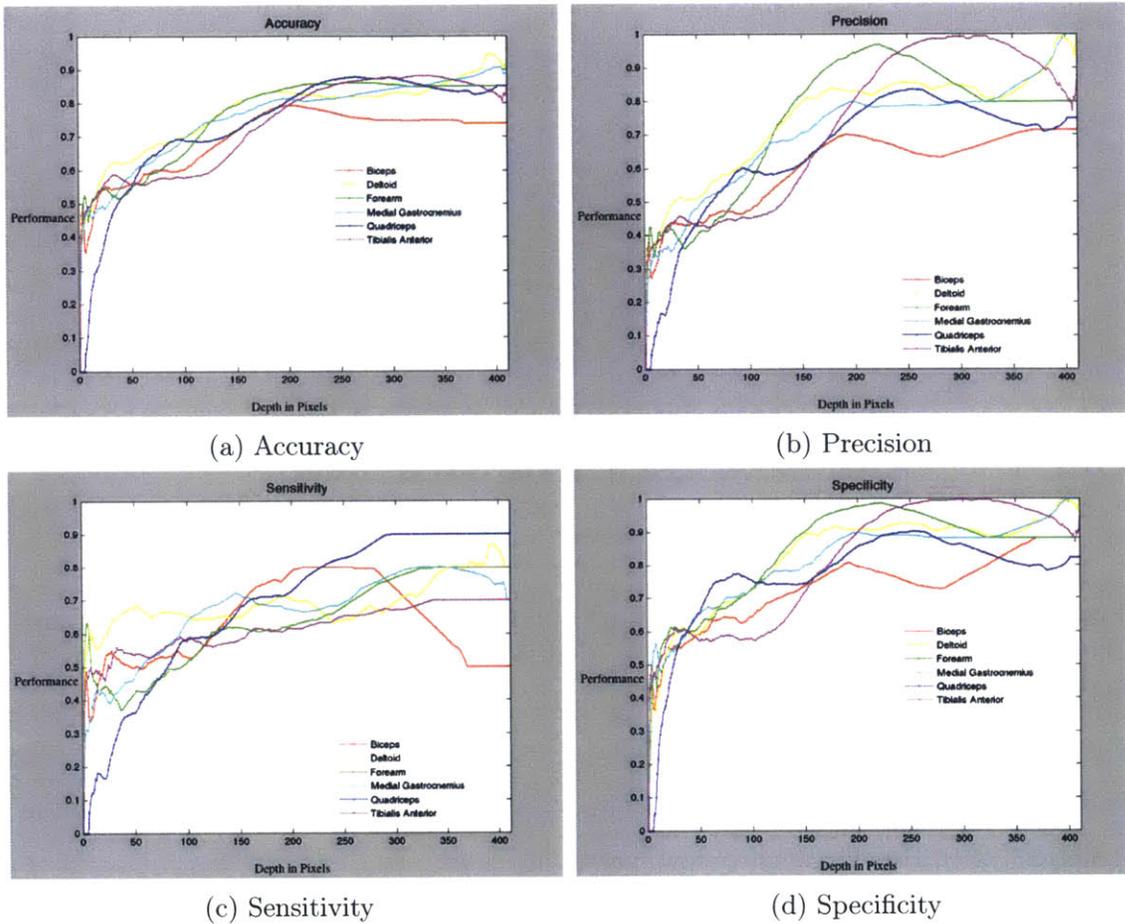


Figure 4-7: Trends for Accuracy, Precision, Sensitivity and Specificity for different depths of the images.

Specificity refers to the proportion of Control (negative) labeled subjects that were predicted as Control (negative). This refers to the behavior of our system with regard to the control subjects by muscle groups. We observed the specificities to be D(1.00), MG(1.00), TA(0.94), B(0.88), F(0.88) and Q(0.82). Therefore, D and MG are exceptionally good at predicting control subjects as controls, with the rest closely following behind.

We also evaluated how these measures would be affected by considering sub-regions of the image. To do this, we consider depth as an approximation for regions-of-interest at different parts of the image. Typically, the images contain a subcutaneous fat layer followed by muscle, and then the bone. Accuracy, Precision, Sensitivity and

Specificity were then plotted by depth-in-pixels for each of the muscle groups.

In Fig. 4-7, we see the trend for accuracy and other metrics, if our system could only view the top  $x$  pixels of the ultrasound images. We see that in general, by considering the entire image, we are not losing out on performance in any muscle group. The results were smoothed by a moving window of 20 pixels, roughly 5% of the total depth of roughly 400 pixels. Other performance metrics behaved similarly. We found that considering the entire image did not appear to hinder the performance metrics in any significant way. We care about our choice of choosing the entire image because it disassociates the need to perform segmentation and specific targeting - only as long as our trends are as described above, i.e., trends that are dependent primarily on the relevant parts of the image such as the subcutaneous fat and the muscle.

## 4.4 Summary

We presented a system that provides for precise acquisition of ultrasound images based on pre-determined contact force values, and uses such standardized images to assess DMD in patients. We note that our performance measures reflect the performance of our system with just one textural feature, that of edge detection. Using other features would be expected to show different behavior among these performance measures towards the various muscle groups. This system could be used with the appropriate muscle group choices in order to further tune the system. Many of the measures show relatively high values that could be used to reliably classify DMD vs. non-DMD in a patient. In the future, similar methods may be used to promote ultrasound as a quantitative non-invasive tool for not only classifying DMD vs. non-DMD but also tracking its progression over time.



# Chapter 5

## Force-correlated Ultrasound

In this chapter, we propose a new modality for automated diagnostic assessment of tissues in the context of Duchenne Muscular Dystrophy (DMD). In this force-correlated ultrasound imaging method, we first perform an automated extraction of a multitude of ultrasound images captured across a range of contact forces - a force video, or a force sweep. These images are then processed to enhance the diagnostic fidelity of the image with regard to DMD. We use a variance map, which computes the pixel-wise standard deviation image for a multiscale stack generated from each image. Using a biomarker quantification scheme of mean gray scale level (GSL) on the enhanced fidelity force-correlated ultrasound images, k-means clustering is then performed to discriminate the DMD subjects from the control subjects. We present our results on the use of these techniques in the diagnostic assessment of DMD on data gathered from a clinical study with our system.

### 5.1 Motivation

Our method takes advantage of spatial and force-related variations in the tissue ultrasound that we believe to be indicative of DMD or its absence. Ultrasound images from healthy and DMD subjects are shown in Fig. 5-1.



(a)



(b)

Figure 5-1: (a) Ultrasound scans of the Quadriceps muscle in a subject from the control group (b) The same muscle in a subject with Duchenne Muscular Dystrophy (DMD). DMD-affected muscles exhibit high fat infiltration.

## 5.2 Methodology

Our method takes advantage of novel hardware that acquires force-correlated ultrasound images, in tandem with mechanisms to enhance diagnostic fidelity with regard to various medical conditions like DMD. This could eventually aid in increasing the diagnostic capability, cross-operator consistency and cross-temporal reliability of quantitative measures extracted from free-hand ultrasound images.

### 5.2.1 Data Acquisition

Ultrasound images are acquired as a function of contact force using the force-controlled ultrasound probe developed by our group [12], shown in Fig 4-2. The device, grasped by the user, actuates the probe to maintain a programmable contact force between probe and patient and accommodates up to 6 cm of relative motion.

To acquire images, the user first grasps the blue probe handle and places the face of ultrasound probe in contact with the tissue. The user manually adjusts the angle of orientation of the device to ensure that it is normal to the surface of the tissue and imaging the transverse plane. When the device is properly oriented, the user holds the orientation angle steady and clicks a button to execute the force sweep. During the 8-second force sweep, the contact force varies linearly from 1.5 N to 10 N to 1.5 N while ultrasound images are recorded at a rate of 10 frames/second.

Due to the slow speed of compression with respect to the dynamics of tissue motion, it is assumed that the tissue is imaged quasi-statically. As the force varies, ultrasound images are recorded synchronously with the contact force and time. This technique enables the acquisition of ultrasound images as a function of contact force.

### 5.2.2 Image Set Selection

We select a uniformly force-distributed set of images during the force sweep in order to compare force sweeps irrespective of how the ultrasound imaging hardware may vary frame rates in response to internal settings. We collate the data points within the force sweep of 1.5 N to 10N at 1.25N to 10.25N at 0.5N intervals. Images are

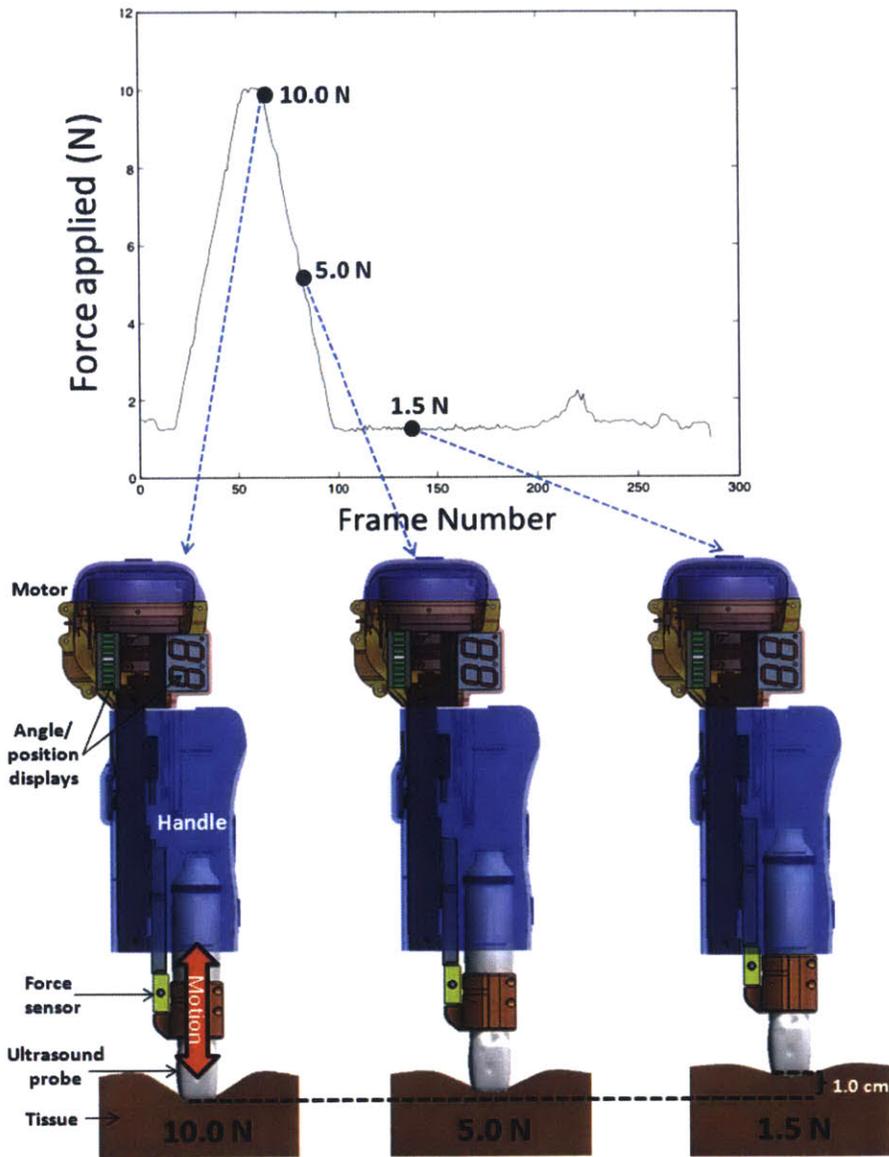


Figure 5-2: The force-controlled ultrasound probe. Device visualization courtesy Matthew W. Gilbertson.

corrected to a new discretized force index at the intervals described, based on which of the set of available points are closest.

### 5.2.3 Diagnostic Fidelity Enhancement

Typically, a radiologist would choose an appropriate image within a scan, and mark a region of interest. The native ultrasound images obtained are not tuned to highlight features indicative of DMD. Here, we take a different approach, depicted in Fig 5-3. We first apply a box filter to the original frame (a) to obtain (b). By doing this at multiple scales we obtain a stack of multiscale images for a single frame as in (c). We then compute the pixel-wise standard deviation image of the multiscale stack to obtain (d), an example of a control subject EF B-Mode, and (e) which is a DMD subject. This approach is intended to enhance the fidelity of the image with respect to DMD. This is necessary as it is not feasible to obtain human annotations of all the images at multiple forces in practice.

We now describe the design of our diagnostic fidelity enhancement measure for DMD in more detail. We use a multi-scale approach to condense and capture the spatial variation in echogenicity throughout the entire image. To do this, we apply a low-pass box filter of size  $N \times N$ , with each element equal to  $(\frac{1}{N \times N})$ . We apply this filter for  $N = 3, 5, 7, \dots, 99, 101$ , i.e., a total of 50 multi-scale images.  $N$  was chosen based on visual inspection of the feature blurring across our dataset for different filter sizes. In Fig 5-3, (b) shows a similarly filtered image generated using (a), while (c) shows an entire multiscale stack.

This stack of multiscale images is then condensed by taking the pixel-wise standard deviation, resulting in a single image that highlights the most spatially variant subregions of an image. This produces images similar to (d) and (e) in Fig 5-3, for healthy controls and DMD subjects, respectively. We refer to this method, which we apply to each image in the force sweep, as the variance map.

The design of our diagnostic fidelity measure was based on the nature of images obtained from the control subjects and the DMD subjects, and their differences. Firstly, there appears to be a greater spatial variation in echogenicity across the

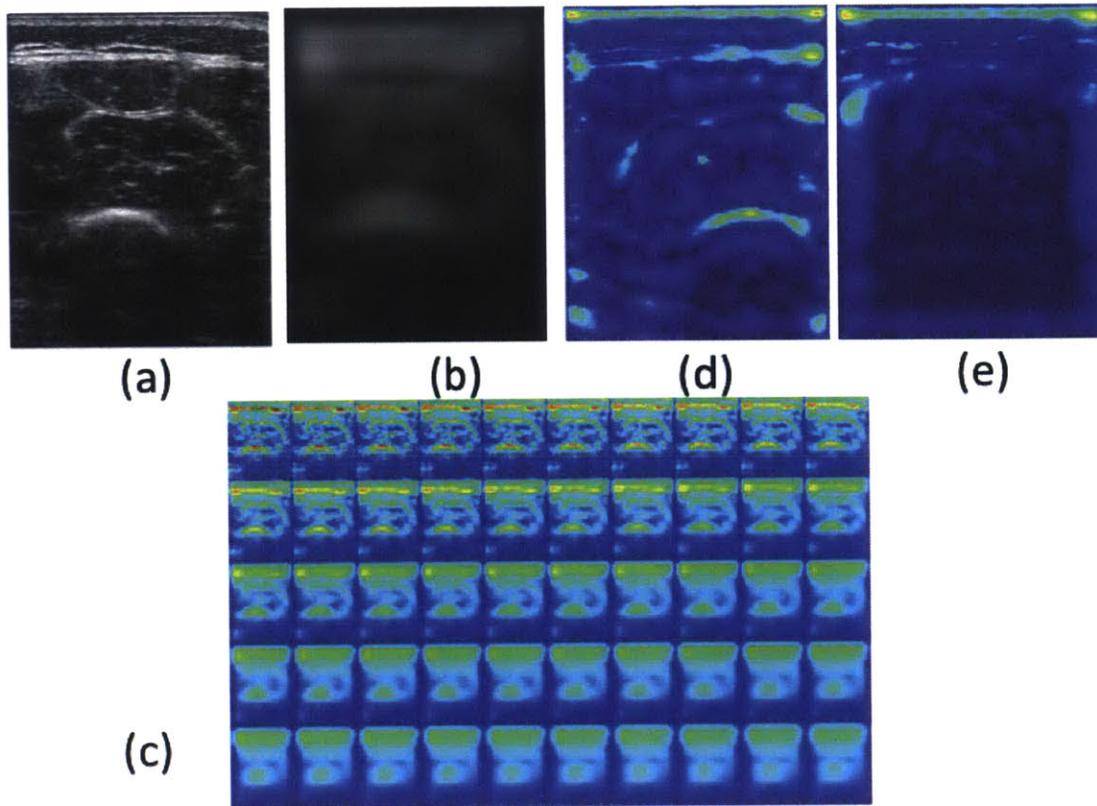


Figure 5-3: Box filter applied to the original image (a) to obtain (b). Multi-scale box-filter image stack from a single frame (c). Pixel-wise standard deviation image of the multiscale stack (d), an example of a control subject EF B-Mode, and (e) which is a DMD subject. (c), (d) and (e) are shown in color to better highlight the enhancement of discriminating features and suppression of irrelevant textures with respect to DMD and Controls.

control subjects. As a result, the visibility of various fascia is lower for DMD in general. The second aspect relates to the ultrasound attenuation at the bottom half of the image, where a bone is typically expected. While the etiology of DMD relates to the muscle, it is clear that our mechanism of observation, i.e., ultrasound, is affected by it in such a way that differentiating features can be seen in the lower part of the image due to attenuation. Therefore, we chose to incorporate this as a feature by considering the whole image as indicative of DMD, rather than drawing regions-of-interest for each image in the force sweep for only the muscle region.

#### 5.2.4 Biomarker Quantification

While the diagnostic fidelity enhancement step for the variance map generation produces an image, we still need a way to reduce this variance map to a number in order to simplify our next steps. We choose the mean gray scale luminosity (intensity) or GSL of the variance map as the biomarker quantification method. When applied to images produced by the variance map, we call this the Enhanced Fidelity B-Mode or EFB-Mode version of the mean GSL, in comparison to the normal B-Mode version.

#### 5.2.5 Clustering

Given a set of force sweeps of various muscles across many patients, we want to cluster them accurately into their respective labels of DMD or Control. We choose the k-means algorithm with 2 centers for clustering. The distance metric for the k-means algorithm was chosen as the Manhattan distance metric. Given two points  $X = (X_1, X_2, \dots, X_N)$  and  $Y = (Y_1, Y_2, \dots, Y_N)$ , the Manhattan distance metric between  $X$  and  $Y$  is defined as,

$$d(X, Y) = \sum_{i=1}^N |X_i - Y_i|$$

When this clustering is applied to the force-sweep data as quantified by the biomarker quantification method above, we refer to it as force-correlated version of our framework. We choose the mean of the biomarker quantification method across all the

images in the force sweep from 1N-10.5N as the ground truth. We call this the force-independent version.

### 5.2.6 Performance Analysis

We evaluated the performance of our system for the force-correlated and force-independent versions, with both B-Mode and the EFB-Mode processed images. We quantified the performance of our clustering using the Adjusted Rand Index (ARI) that is corrected-for-chance [18]. This measure has an expected value of 0 for random clusters, and a maximum value of 1 for completely identical clusters. It can also have a value from 0 to -1, indicating worse results for the clustering in comparison to randomized clustering. All evaluations of our system were performed on individual muscle groups to quantify muscle-specific effects.

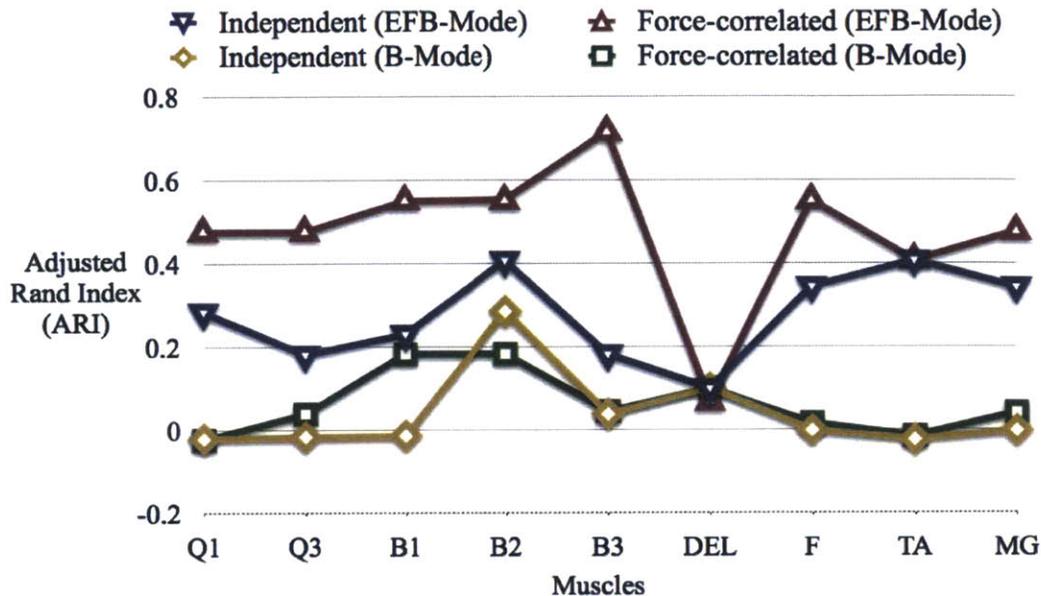
## 5.3 Results and Discussion

### 5.3.1 Clinical Study

The clinical study was approved by the Children’s Hospital Boston Institutional Review Board, and all parents/subjects were required to give signed informed consent and assent. The BOSTON I dataset comprises the force-correlated ultrasound images collected using our device on 18 DMD and 22 Control subjects. Only one visit per patient was recorded in this dataset. Each visit comprised 9 scans using our device on 6 muscle groups, viz., Quadriceps (Q), Biceps (B), Deltoids (DEL), Forearm (F), Tibialis Anterior (TA) and the Medial Gastrocnemius (MG). The Quadriceps and Biceps also had inter-rater measurements (Q1 and Q3, B1 and B3), while the Biceps had an intra-rater measurement as well (B1 and B2).

### 5.3.2 Results

The results are shown in Fig. 5-4. For brevity, we will refer to the muscles by their abbreviated names defined above. The performance of the above system was



### Mean GSL of Enhanced Fidelity B-Mode and Regular B-Mode

Figure 5-4: Adjusted Rand Indices (ARIs) of the four variations

measured using the corrected-for-chance ARI performance metric described. The independent of force ARI on the normal B-Mode data performs similar to random clustering. The force-correlated ARI on the normal B-Mode performs slightly better (Q3,B1,B3,F,TA,MG), although the improvements are exceptionally significant.

The impact of the diagnostic fidelity enhancement via variance maps can be seen in comparing the independent of force ARI on the Enhanced Fidelity B-Mode (EFB-Mode) data. Consistent increases in the ARI are seen for Q (0.25), B(0.17), F(0.35), TA(0.43), and MG(0.35). The inter-rater errors for Q and B in this case were 0.10 and 0.05. The intra-rater error for B was 0.18. This is further boosted to the highest values for each muscle group (except D) by the force-correlated ARI on the EFB-Mode data. Compared to the force-independent ARI on EFB-Mode data that were hitherto the highest, ARI increases were seen in Q (0.24), B (0.33), F(0.21),MG(0.13) and a very minor increase in TA. The overall ARIs were the highest for all muscles except D for the force-correlated on EFB-Mode case with Q(0.48), B(0.61), F(0.55), TA(0.41) and MG(0.48).

In Fig. 5-4, the Adjusted Rand Indices (ARIs) of the four variations of our system are shown, summarizing the impact of introducing DMD-specific diagnostic fidelity enhancement via variance maps, and the effect of force-correlated ultrasound imaging vis-à-vis independent of force imaging that was computed as the mean value of the biomarker quantification method (mean echogenicity) over the range of forces. The sources of variation in these values can only be due to the choices of initial cluster centers during the clustering. Due to selecting the best performing cluster among 5 internal replicates during each k-means clustering run, we noted deterministic behavior on our dataset for our system.

Consistency also improved with force-correlation, with the inter-rater error for Q being 0.00041. Inter-rater error for B rose to 0.16 from 0.05. The intra-rater error for B was 0.00043. To summarize, exceptionally low inter-rater error on Q and intra-rater error for B were seen for force-correlation on EFB-Mode data. But, the inter-rater error on B increased from 0.05 to 0.16 from independent of force to force-correlated ARIs on EFB-Mode. We believe this may be because the current system does not provide for a consistent localization in 3D on the patient's limb. Nevertheless, a very low inter-rater error on Q and intra-rater error for B were obtained. This should also be seen in the context of all the 3 B ARIs for force-correlated EFB-Mode being the highest of all the four versions considered.

The Deltoids muscle received almost no performance boost by using either EFB-Mode or force-correlation. This could be a limitation of our setup in its present form, as the same diagnostic fidelity enhancement measure or the biomarker quantification method may not work for all muscle groups. We suspect that size and geometry issues impact the practical use on the Deltoid muscle for small children by affecting the extent of usable ultrasound image within the full frame obtained.

## 5.4 Summary

This study investigated the potential for force-correlated ultrasound imaging to enable automated assessment of tissues in subjects with Duchenne Muscular Dystrophy. We

observed excellent classification improvements due to force-correlation for all muscles except Deltoids. Our system does not require training data, and can achieve good and consistent performance thanks to the performance boosts of DMD-specific fidelity enhancement and the force-correlated imaging. We believe this system holds potential for enhancing the use of Quantitative Ultrasound (QUS) as a candidate for the non-invasive measure of DMD progression, with high reliability and consistency across operators and over time.



# Chapter 6

## Conclusion

In this chapter, we review our contributions as well as provide more preliminary results that gives a context for the importance of this work. Hopefully, these lines of inquiry will be among those explored to continue this work in future.

We begin by reviewing the contributions so far, and then discuss five lines of inquiry that are of interest. Specifically, these are summarized below (and discussed further in this chapter):

- **Imaging biomarkers.** It would be of interest to evaluate the space of imaging biomarkers more exhaustively to find those that might perform better than those we considered in this thesis (to a reasonable degree of performance for DMD). Also, it is of interest to map out the optimal imaging biomarkers for each condition to which this (and future) automated discrimination and progression tracking frameworks are applied to.
- **Optimal Ultrasound Image Acquisition.** We present our preliminary results on which forces are appropriate for maximizing the discriminatory power of an imaging biomarker. This could be of great potential in providing significant cost benefits by reducing clinical time required to conduct a study, while retaining and/or improving performance of any automated framework.
- **Progression Tracking.** Being the first of its kind so far, in this thesis, we have chosen to focus on discriminatory frameworks that would be suitable for

diagnosis. We discuss some of the challenges with progression tracking.

- **Segmentation.** Throughout the entirety of this work, we have chosen to simplify the imaging biomarker sub-component by considering only biomarkers that can be applied to the entire ultrasound image while *simultaneously* showing maximal discriminatory performance given the rest of the framework. An alternate method would be to segment the image into subcutaneous fat, muscle and bone - and then hopefully use the additional information to target multiple imaging biomarkers simultaneously to better effect. We discuss a few ways we have tried in this direction.
- **Machine Learning.** We note that the results we have obtained are a *lower bound* given the simplicity of the machine learning methods we have used. Based on clinical feedback that a black box approach was not preferable, we have chosen the decision tree and k-means clustering algorithms to show that there is a respectable value addition by controlled ultrasound acquisition *despite* the simplicity of our machine learning methods.

We now summarize the thesis, and then discuss each of the above preliminary results further to illuminate possible avenues of future work.

## 6.1 Contributions

We have examined the clinical utility of force-controlled acquisition of ultrasound images for Duchenne Muscular Dystrophy.

The contributions include:

- Evaluation of imaging biomarkers that are of relevance to Duchenne Muscular Dystrophy, such as edge detection and variance maps.
- Novel frameworks to study the utility of ultrasound images acquired at precise forces and across a force sweep, for rapid and automated discrimination of Duchenne Muscular Dystrophy

- Analysis of results from a clinical study using these frameworks.
- Discussion of extensions to our work in five directions that could be of significant potential for tracking progression of muscle disorders. This includes using other types of imaging biomarkers, optimizing the ultrasound image acquisition in terms of maximizing the discriminatory power of the images acquired, progression tracking, targeting specific regions by segmenting out images prior to our frameworks and using better machine learning techniques.

## 6.2 Future work

There is good potential for the direction of this work. In particular, we describe our work in five extensions to the current work that would be good candidates for extensive exploration in future.

### 6.2.1 Imaging Biomarkers

Within this thesis, we limited ourself to two imaging biomarkers - Canny edge detection and variance maps - that we found by manually searching for those that had reasonable performance with the Duchenne Muscular Dystrophy ultrasound image datasets. A more exhaustive search for better imaging biomarkers as well as motion-estimation based methods for the force-dependent ultrasound image data would be of great potential.

In Fig 6-1, we show ultrasound images of Duchenne Muscular Dystrophy and control group subjects in (a) and (b). In (b) and (c), we show the corresponding phase congruency feature detected images [22], which is known to be illumination and contrast-invariant. This would detect features at all phase angles instead of solely based on gradients at perpendicular cutoffs. This can be discretized and clarified by simple morphological closing and opening operations to result in (e) and (f), which could potentially serve as one of the alternate imaging biomarkers to explore.

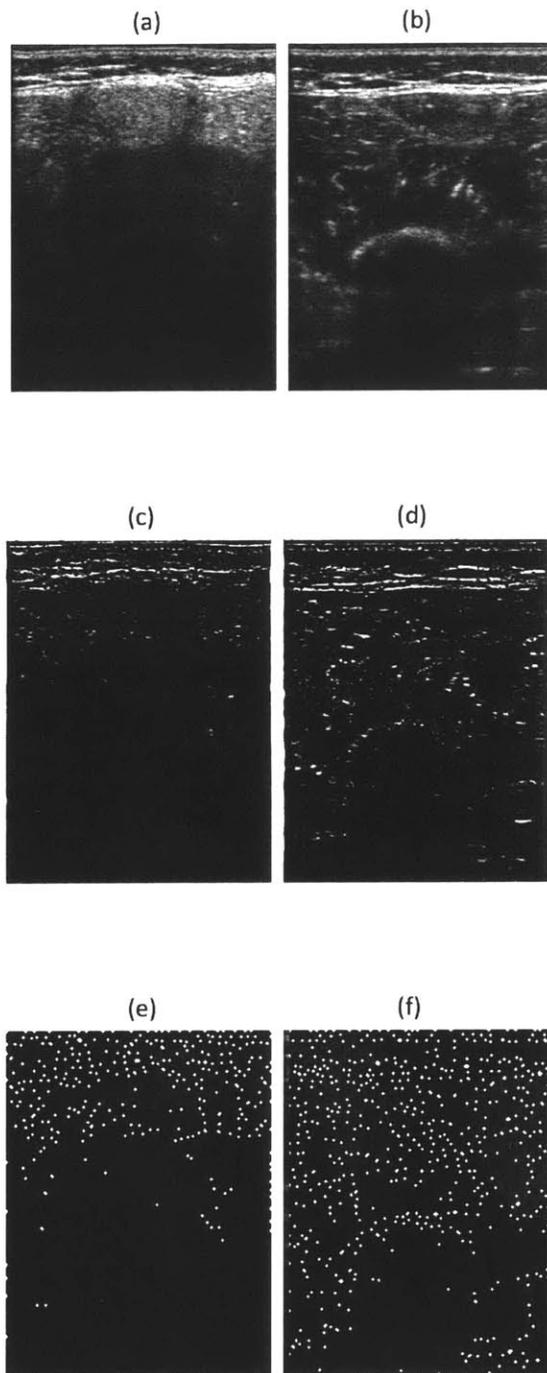


Figure 6-1: Example of smarter edges, with appropriate morphological operations to discretize and clarify.

## 6.2.2 Optimal Ultrasound Image Acquisition

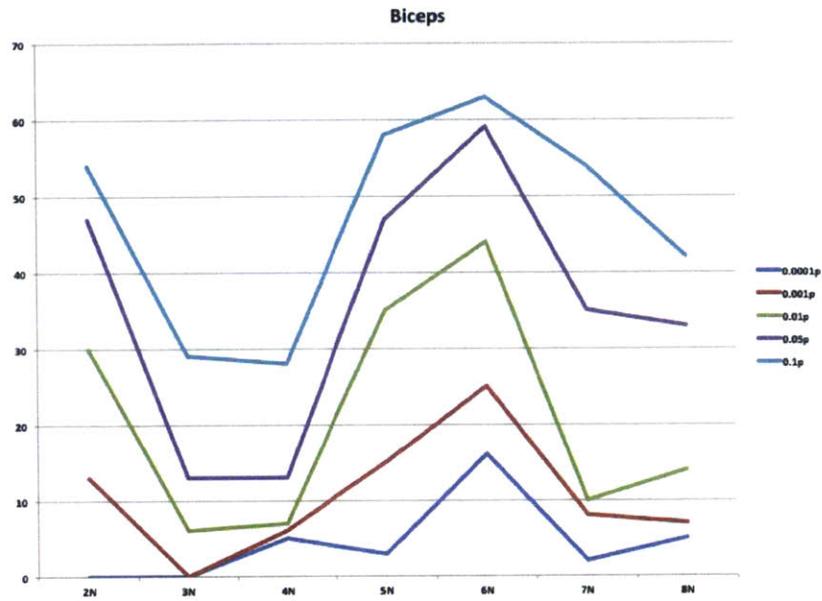
Given the ability to acquire ultrasound images at precise forces, it is useful to ask the question - what forces are optimal for discriminating ability between the control group and the Duchenne Muscular Dystrophy group. Indeed, this appears to be a useful direction of inquiry based on our preliminary results that we discuss here.

We considered the images acquired at different forces as distinct sets, and processed these through 99 thresholds of the Canny edge detection algorithm from 0.01 to 0.99 at intervals of 0.01. We then evaluated how many of these, at each force, were found to be powerful enough to be discriminating between the control group and the Duchenne Muscular Dystrophy subject group. We also attempted to measure the *power* of discrimination via the p-value of the t-test used for the discrimination.

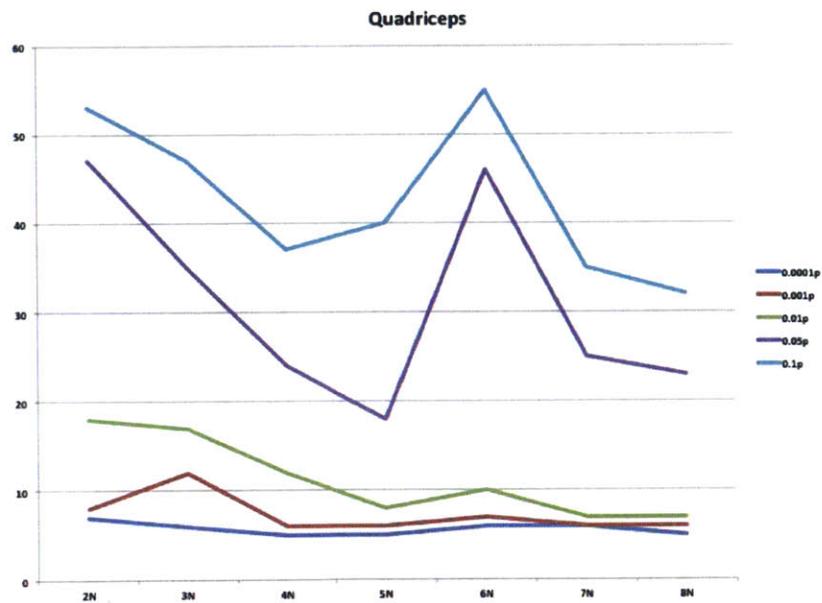
In Fig 6-2, we show this characteristic discriminatory power that varies with force, for Biceps and Quadriceps. Note that there is a lull in the discriminatory power, consistent across all levels of significance, between 3N and 4N for Biceps. There is a similar lull in discriminatory power between 4N and 5N for the Quadriceps, which is more pronounced only for the three most stringent levels of significance. There is also a clear *peak* discriminatory power at 6N for Biceps. For Quadriceps, the peak discriminatory power is at 6N for the two most stringent levels of significance and at 2N-3N for the remaining levels of significance. We are not very sure as to the physiological origins of this *peak* discriminatory power with respect to force, based on interactions with our clinical partners.

In Fig 6-3, we examine this behavior, characteristic of Deltoids and Medial Gastrocnemius. The deltoids have a lull in 4N-5N and the Medial Gastrocnemius has two lulls at 3N and at 6N. Peak discriminatory power for the Deltoid is in the lower range of the force sweep of 2N-3N, quite unlike the other muscle groups. The peak discriminatory power for the Medial Gastrocnemius is at 5N.

The significance of this kind of finding about the nature of discriminatory power of ultrasound images is enormous. Firstly, this will allow us to identify the appropriate force for maximizing utility *and* throughput (by reducing time required) for clinical



(a) Biceps



(b) Quadriceps

Figure 6-2: Statistically significant edge thresholds (out of 99) at different forces for Biceps and Quadriceps. Multiple levels of significance are shown.

data acquisition. Secondly, this tells us something about the inter-dependent behavior of a particular imaging biomarker and a muscle group. With more continuous data and extensive analysis, a complete characterization can be made leading to practical cost control for clinical trials and other purposes.

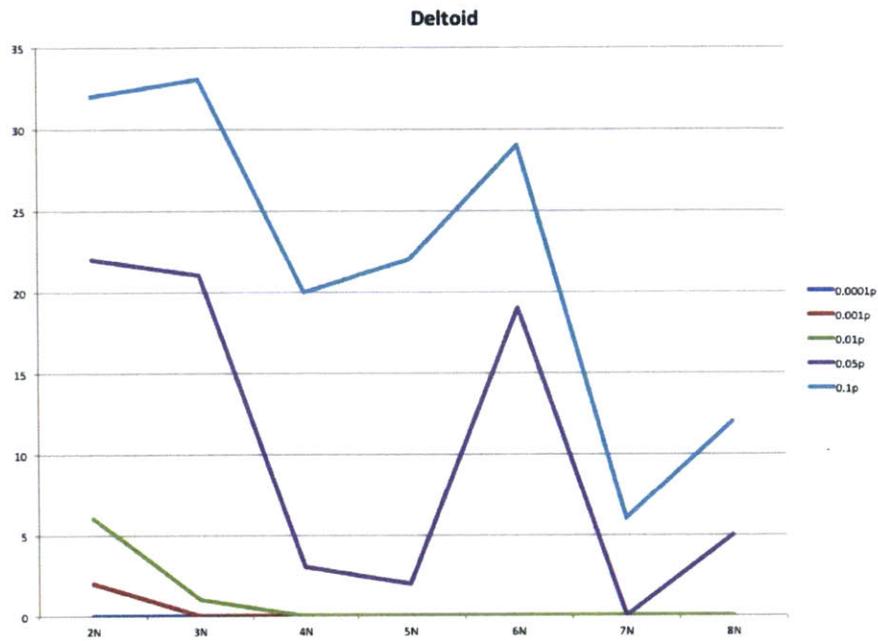
### 6.2.3 Progression Tracking

Progression tracking is of great value to the DMD case study. It allows for observing discriminating behavior between the DMD subjects and the control subjects with variance in age, and over a period of time. By also subjecting a part of the study group to a potential clinical drug, the impact of the drug can be quantified. We did try a heuristic approach where a selected feature set was chosen among the many possible feature vectors for the ultrasound image dataset, and did a PCA to form a cluster of the two groups showing differential behavior with age. However, we were limited by the fact that subspace selection continued to be an issue, with the dimension of the feature vectors necessarily required to be smaller than the dimension of the dataset (number of DMD and Control subjects) due to the PCA algorithm. We discuss a possible solution with the HOSVD algorithm below.

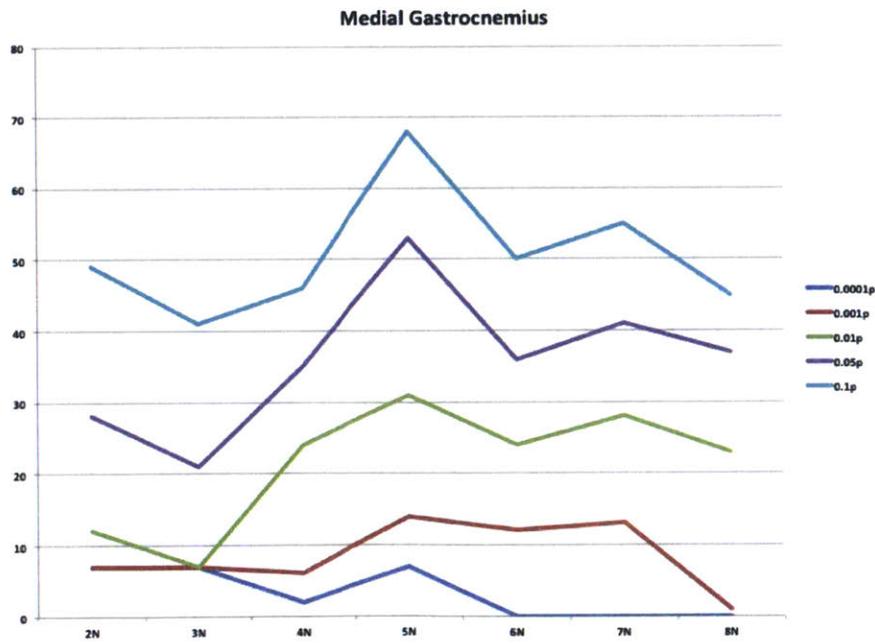
### 6.2.4 Segmentation

We made the possible argument for segmenting the image in the beginning of the chapter. We examined two ways of doing the segmentation, given our controlled-acquisition of ultrasound image data. The first is to segment each image, as one would do in any other method. The second method is to segment along the time-axis of our controlled-acquisition image sequence. This is something that wouldn't be possible in any traditional sonographer-based acquisition.

In the first way, we take each image and try to segment it. For example, in Fig 6-4, (a) represents a simple straight line segmentation of the subcutaneous fat. (b), implemented via a seed-based region growing algorithm, represents a more naturally identifiable segmentation.

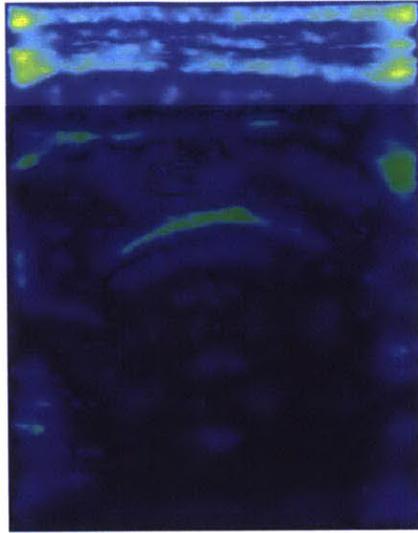


(a) Deltoid

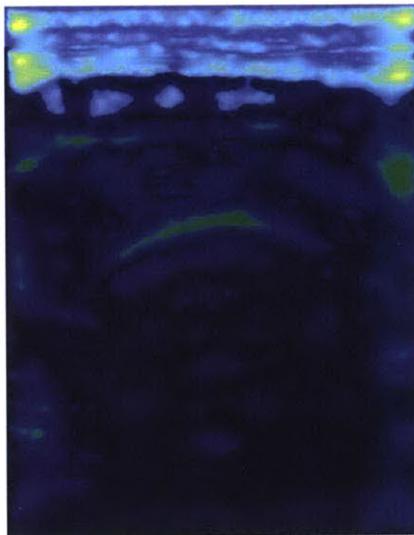


(b) Medial Gastrocnemius

Figure 6-3: Statistically significant edge thresholds (out of 99) at different forces for Deltoid and Medial Gastrocnemius. Multiple levels of significance are shown.



(a) Straight Line Segmentation



(b) Region-Growing Algorithm

Figure 6-4: Image-based Segmentation

A good use of such a segmentation in our context would be if we could measure the thickness of the muscle indirectly from the location of the bone tip and the subcutaneous fat thickness. By measuring the variation in this muscle thickness with force, we could get a proxy measure of the elasticity of the muscle tissue. Given that Duchenne Muscular Dystrophy slowly introduces fat into the muscle tissue, this could be a potential way to capture that information and use it in our framework.

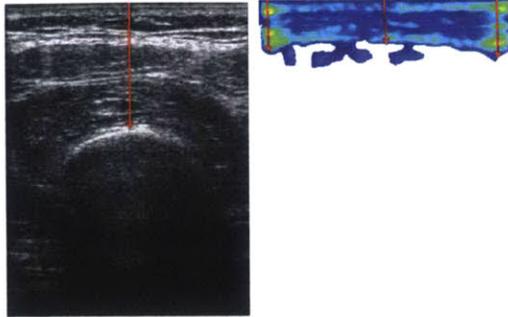
We illustrate this hypothesis in Fig 6-5. Note that the measurements described above are more similar for a control subject across visits, than in comparison to a completely different control subject.

In Fig 6-6, we visualize the force volumes acquired along the force-axis. This is better than viewing a video along time-axis, because these are equispaced along the force axis. Since these are viewed from the side, we can see the way the bone moves with a reduction in force for control subjects in (b). This could therefore be used for measuring the change in muscle thickness. Unfortunately, in (a) we cannot see the bone for the Duchenne Muscular Dystrophy patient. This is the case for most of the DMD patients. Therefore, it is our opinion that segmentation cannot be used to measure the change in muscle thickness (or other DMD specific properties, which affect only the muscle), and consequently infer properties of the muscle in our context.

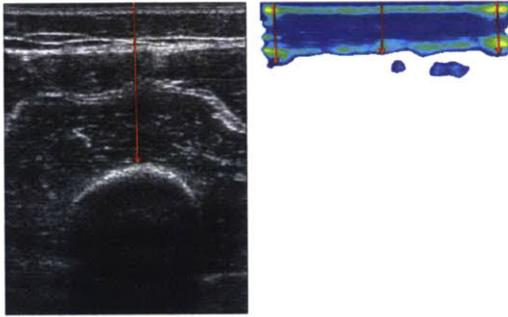
### **6.2.5 Better Learning Methods**

While our clinical feedback was against a black box approach, there is still potential for better performance than the baseline performances we have reported here. Given that the baseline performances validate the utility of controlled-acquisition ultrasound, there is immense potential for the possibilities with better methods.

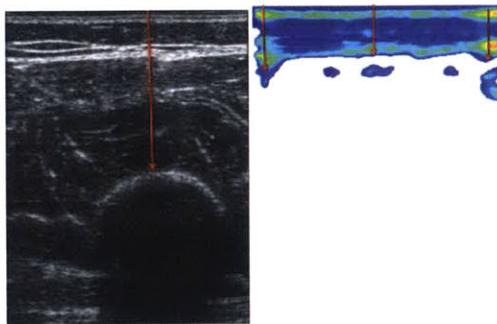
In this thesis, we have been trying to examine the utility of having controlled acquisition in ultrasound [21, 20]. This implies measuring the force, during the operation of clinical ultrasound data acquisition, and examining the utility of this additional knowledge. In order to extend it to tracking the progression of these disease conditions over time, we found that it was useful to specifically learn how to handle data that is organized by more than one dimension or type. This is a common need across



(a) A control subject



(b) A different control subject



(c) A different visit of the control subject from  
(b)

Figure 6-5: Hypothesis for inferring muscle thickness on Quadriceps, given Image-based Segmentation



(a) A DMD subject, the same subject after 6 months and a different DMD subject.



(b) A control subject, the same muscle scanned again for the same subject in the same visit, and a different control subject

Figure 6-6: Force-axis Segmentation

many problems, and a good representation we want to try is a third-order tensor (three dimensional array).

Tensor space representations for image sets have been found to be computationally efficient compared to other similar methods [15]. For example, Higher Order SVD (HOSVD) was used, with good computational performance, for the classification of handwritten digits, a standard problem in pattern recognition [31]. Representation is precise, with the training set of  $m$  number of handwritten digit images, each of size  $n \times n$ , being represented as a tensor  $\mathcal{A} \in R^{n \times n \times m}$ .

Recently, it was proven that Higher Order Singular Value Decomposition does both simultaneous subspace selection and K-means clustering, with a reasonable and a comparable accuracy to, among other things, a PCA-K-means approach [17]. Therefore, these might be good areas for prospective study.

### 6.3 Summary

Controlled acquisition of ultrasound has a large amount of untapped potential. As seen in this thesis, it can bring a great degree of cross-sonographer and cross-temporal reliability, reduce costs by speeding up acquisition thanks to fast and specifically force-targeted acquisition, and best of all, provide the potential for automation/near-automation of diagnosis and progression tracking. We look forward to seeing it in additional clinical applications in the near future.



# Bibliography

- [1] R G Aarnink, S D Pathak, J J de la Rosette, F M Debruyne, Y Kim, and H Wijkstra. Edge detection in prostatic ultrasound images using integrated edge maps. *Ultrasonics*, 36(1-5):635–42, February 1998.
- [2] Y Arai, M Osawa, and Y Fukuyama. Muscle CT scans in preclinical cases of Duchenne and Becker muscular dystrophy. *Brain & Development*, 17(2):95–103, January.
- [3] K J Bhansing, E P Hoppenreijs, A J Janssen, A van Royen-Kerkhof, M W Nijhuis-Van der Sanden, P L C M van Riel, and S Pillen. Quantitative muscle ultrasound: a potential tool for assessment of disease activity in juvenile dermatomyositis. *Scandinavian Journal of Rheumatology*, 43(4):339–41, January 2014.
- [4] Alice Bickerstaffe, Anita Beelen, Machiel J Zwarts, Frans Nollet, and Johannes P van Dijk. Quantitative muscle ultrasound and quadriceps strength in patients with post-polio syndrome. *Muscle & Nerve*, 51(1):24–9, January 2015.
- [5] B J Boland, P L Silbert, R V Groover, P C Wollan, and M D Silverstein. Skeletal, cardiac, and smooth muscle failure in Duchenne muscular dystrophy. *Pediatric Neurology*, 14(1):7–12, January 1996.
- [6] L Breiman, J Friedman, CJ Stone, and RA Olshen. Classification and Regression Trees. 1984.
- [7] Kate Bushby and Edward Connor. Clinical outcome measures for trials in Duchenne muscular dystrophy: report from International Working Group meetings. *Clinical Investigation*, 1(9):1217–1235, September 2011.
- [8] Katharine Bushby, Richard Finkel, David J Birnkrant, Laura E Case, Paula R Clemens, Linda Cripe, Ajay Kaul, Kathi Kinnett, Craig McDonald, Shree Pandya, James Poysky, Frederic Shapiro, Jean Tomezsko, and Carolyn Constantin. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurology*, 9(1):77–93, January 2010.
- [9] John Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, November 1986.

- [10] RO Crapo, R Casaburi, and AL Coates. ATS statement: guidelines for the six-minute walk test. 2002.
- [11] A E Emery. Population frequencies of inherited neuromuscular diseases—a world survey. *Neuromuscular Disorders : NMD*, 1(1):19–29, January 1991.
- [12] Matthew W. Gilbertson and Brian W. Anthony. Ergonomic control strategies for a handheld force-controlled ultrasound probe. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1284–1291. IEEE, October 2012.
- [13] M Gudmundsson, E A El-Kwae, and M R Kabuka. Edge detection in medical images using a genetic algorithm. *IEEE Transactions on Medical Imaging*, 17(3):469–74, June 1998.
- [14] A Hammoude. Edge detection in ultrasound images based on differential tissue attenuation rates. *Ultrasonic Imaging*, 21(1):31–42, January 1999.
- [15] Xiaofei He, Deng Cai, Haifeng Liu, and Jiawei Han. Image clustering with tensor representation. In *Proceedings of the 13th annual ACM International Conference on Multimedia - MULTIMEDIA '05*, page 132, New York, New York, USA, November 2005. ACM Press.
- [16] E P Hoffman, R H Brown, and L M Kunkel. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell*, 51(6):919–28, December 1987.
- [17] Heng Huang, Chris Ding, Dijun Luo, and Tao Li. Simultaneous tensor subspace selection and clustering. In *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 08*, page 327, New York, New York, USA, August 2008. ACM Press.
- [18] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, December 1985.
- [19] Merel Jansen, Nens van Alfen, Maria W.G. Nijhuis van der Sanden, Johannes P. van Dijk, Sigrid Pillen, and Imelda J.M. de Groot. Quantitative muscle ultrasound is a promising longitudinal follow-up tool in Duchenne muscular dystrophy. *Neuromuscular Disorders*, 22(4):306–317, 2012.
- [20] Sisir Koppaka, Matthew W. Gilbertson, Seward B. Rutkove, and Brian W. Anthony. Evaluating the clinical relevance of force-correlated ultrasound. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 1172–1175. IEEE, April 2014.
- [21] Sisir Koppaka, Matthew W. Gilbertson, Jim S. Wu, Seward B. Rutkove, and Brian W. Anthony. Assessing duchenne muscular dystrophy with force-controlled ultrasound. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 694–697. IEEE, April 2014.

- [22] P Kovesi. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, 1999.
- [23] Ahmed Mahmoud, Ahmed Morsy, and Eric de Groot. A new gradient-based algorithm for edge detection in ultrasonic carotid artery images. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2010:5165–8, January 2010.
- [24] E S Mazzone, S Messina, G Vasco, M Main, M Eagle, A D’Amico, L Doglio, L Politano, F Cavallaro, S Frosini, L Bello, F Magri, A Corlatti, E Zucchini, B Brancalion, F Rossi, M Ferretti, M G Motta, M R Cecio, A Berardinelli, P Alfieri, T Mongini, A Pini, G Astrea, R Battini, G Comi, E Pegoraro, L Morandi, M Pane, C Angelini, C Bruno, M Villanova, G Vita, M A Donati, E Bertini, and E Mercuri. Reliability of the North Star Ambulatory Assessment in a multicentric setting. *Neuromuscular Disorders : NMD*, 19(7):458–61, July 2009.
- [25] Elena Mazzone, Diego Martinelli, Angela Berardinelli, Sonia Messina, Adele D’Amico, Gessica Vasco, Marion Main, Luca Doglio, Luisa Politano, Filippo Cavallaro, Silvia Frosini, Luca Bello, Adelina Carlesi, Anna Maria Bonetti, Elisabetta Zucchini, Roberto De Sanctis, Marianna Scutifero, Flaviana Bianco, Francesca Rossi, Maria Chiara Motta, Annalisa Sacco, Maria Alice Donati, Tiziana Mongini, Antonella Pini, Roberta Battini, Elena Pegoraro, Marika Pane, Elisabetta Pasquini, Claudio Bruno, Giuseppe Vita, Chiara de Waure, Enrico Bertini, and Eugenio Mercuri. North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscular Disorders : NMD*, 20(11):712–6, November 2010.
- [26] Craig M McDonald, Erik K Henricson, Jay J Han, R Ted Abresch, Alina Nicorici, Gary L Elfring, Leone Atkinson, Allen Reha, Samit Hirawat, and Langdon L Miller. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle & Nerve*, 41(4):500–10, April 2010.
- [27] J Ophir, I Céspedes, H Ponnekanti, Y Yazdi, and X Li. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrasonic Imaging*, 13(2):111–34, April 1991.
- [28] S A E Peters, H M den Ruijter, M K Palmer, D E Grobbee, J R Crouse, D H O’Leary, G W Evans, J S Raichlen, L Lind, and M L Bots. Manual or semi-automated edge detection of the maximal far wall common carotid intima-media thickness: a direct comparison. *Journal of Internal Medicine*, 271(3):247–56, March 2012.
- [29] Sigrid Pillen, Ramon O Tak, Machiel J Zwarts, Martin M Y Lammens, Kiek N Verrijp, Ilse M P Arts, Jeroen A van der Laak, Peter M Hoogerbrugge, Baziel G M van Engelen, and Aad Verrips. Skeletal muscle ultrasound: correlation between fibrous tissue and echo intensity. *Ultrasound in Medicine & Biology*, 35(3):443–6, March 2009.

- [30] Seward B Rutkove, Tom R Geisbush, Aleksandar Mijailovic, Irina Shklyar, Amy Pasternak, Nicole Visyak, Jim S Wu, Craig Zaidman, and Basil T Darras. Cross-sectional evaluation of electrical impedance myography and quantitative ultrasound for the assessment of Duchenne muscular dystrophy in a clinical trial setting. *Pediatric Neurology*, 51(1):88–92, July 2014.
- [31] Berkant Savas and Lars Eldén. Handwritten digit classification using higher order singular value decomposition. *Pattern Recognition*, 40(3):993–1003, March 2007.
- [32] A Schreiber, W L Smith, V Ionasescu, H Zellweger, E A Franken, V Dunn, and J Ehrhardt. Magnetic resonance imaging of children with Duchenne muscular dystrophy. *Pediatric Radiology*, 17(6):495–7, January 1987.
- [33] Irina Shklyar, Tom R Geisbush, Aleksandar S Mijialovic, Amy Pasternak, Basil T Darras, Jim S Wu, Seward B Rutkove, and Craig M Zaidman. Quantitative muscle ultrasound in Duchenne muscular dystrophy: a comparison of techniques. *Muscle & Nerve*, 51(2):207–13, February 2015.
- [34] Martin Torriani, Elise Townsend, Bijoy J Thomas, Miriam A Bredella, Reza H Ghomi, and Brian S Tseng. Lower leg muscle involvement in Duchenne muscular dystrophy: an MR imaging and spectroscopy study. *Skeletal Radiology*, 41(4):437–45, April 2012.
- [35] Jim S Wu, Basil T Darras, and Seward B Rutkove. Assessing spinal muscular atrophy with quantitative ultrasound. *Neurology*, 75(6):526–31, August 2010.
- [36] Craig M. Zaidman, Anne M. Connolly, Elizabeth C. Malkus, Julaine M. Florence, and Alan Pestronk. Quantitative ultrasound using backscatter analysis in Duchenne and Becker muscular dystrophy. *Neuromuscular Disorders*, 20(12):805–809, 2010.
- [37] Craig M Zaidman, Jim Wu, Sarah Wilder, Basil Darras, and Seward Rutkove. Minimal training is required to reliably perform quantitative ultrasound of muscle. *Muscle & Nerve*, November 2013.
- [38] W Zwehl, R Levy, E Garcia, R V Haendchen, W Childs, S R Corday, S Meerbaum, and E Corday. Validation of a computerized edge detection algorithm for quantitative two-dimensional echocardiography. *Circulation*, 68(5):1127–35, November 1983.