

Muscle deformation correlates with output force during isometric contraction

Laura A. Hallock, Akash Velu, Amanda Schwartz, and Ruzena Bajcsy

Abstract—Despite the utility of musculoskeletal dynamics modeling, there exists no safe, noninvasive method of measuring in vivo muscle output force in real time. In this paper, we demonstrate that muscle deformation constitutes a promising, yet unexplored class of signals from which to infer such forces. Through a preliminary case study of the elbow joint, in which we examine simultaneous flexion force, surface electromyography (sEMG), and ultrasound imaging data during isometric contraction, we provide evidence that even simple measures of deformation (including cross-sectional area and thickness variation in the brachioradialis muscle) correlate well with elbow output torque to an extent comparable with standard sEMG activation measures. We then show that these simple signals, as well as the overall brachioradialis contour, can be tracked over time via optical flow techniques, enabling the use of these signals in real-time applications (e.g., assistive device control), as well as larger-scale study of deformation signals without necessitating manual annotation. To enable such future work, all modeling and tracking software described in this paper, as well as all raw and processed data, have been made available on SimTK as part of the OpenArm project (<https://simtk.org/projects/openarm>) for general research use.

I. INTRODUCTION

Despite decades of study, noninvasive, in vivo, real-time measurement of muscle forces remains an open problem in the biomechanics community. Without good models of muscle force output during natural movement, our understanding of how humans execute dexterous motions is fundamentally limited, as is our ability to safely modify this execution using assistive devices and to accurately characterize and treat musculoskeletal pathology.

Historically, muscle forces have either been computed using full-body modeling frameworks like OpenSim [1] and AnyBody [2] — which can only account for limited physiological variation and make strong optimization-based assumptions about force distribution across synergists — or using surface electromyography (sEMG), a sensing modality that measures the neurological *input* to the musculoskeletal system, not the resultant *output* forces. While these methodologies have resulted in impressive advances in motion modeling and device control, our ability to both understand and replicate dexterous motions remains severely limited.

As a complementary technology to optimization-based modeling and sEMG measurement, we propose *muscle deformation* (as measured via ultrasound) as a class of signals

that is both more directly representative of muscle output force and easier to spatially localize than sEMG. In particular, we show that 1) various simple measures of muscle deformation (cross-sectional area, thickness, etc.) are readily observable for an individual muscle, 2) these measures correlate with output joint torque during a pulsed and sustained elbow flexion task, and 3) it is feasible to track these measures over time using optical flow and other computer vision techniques.

Deformation is a complex signal class that can be parameterized in a multitude of ways, and it largely becomes useful (for example, in the context of assistive device control or human motion analysis) when it can be both extracted and interpreted in real time. Both aspects of this process — definition of the deformation signal and real-time robust observation — are challenging, as no general models of in vivo muscle deformation exist. This paper presents an exploratory analysis of several promising signal definitions, as well as proof-of-concept tracking of these signals using standard optical flow techniques. These tracking techniques are in turn essential to future analyses, because without automated methods of signal extraction, investigation of new deformation signals is often prohibitively tedious.

The precise contributions of this paper are as follows:

- a novel, quantitative description of the relationship between 2D deformation of the brachioradialis muscle and output torque at the elbow joint, consistent with simultaneous sEMG measurements of muscle activation, in a preliminary cohort of subjects;
- proof-of-concept software to track this muscle deformation over time, released as an open-source codebase and evaluated for tracking quality; and
- an open-source time series data set, including simultaneous ultrasound, sEMG, kinematic configuration, and elbow torque data, for future study and modeling.

The latter two contributions have been made available for general research use as part of the OpenArm project on SimTK (<https://simtk.org/projects/openarm>), which also hosts complementary research and data sets examining 3D (static) muscle shape under various conditions [3, 4].

II. MOTIVATIONS & RELATED WORK

Although finding a noninvasive measure of individual muscle forces remains a core challenge inhibiting our understanding of human motion [5], no single sensing modality or analysis framework has emerged as a dominant solution. In this section, we examine the biological mechanisms underlying muscle force to argue that deformation constitutes a promising candidate signal class, then motivate this paper’s

This work was supported by the NSF National Robotics Initiative (award no. 81774), Siemens Healthcare (85993), the NVIDIA Corporation GPU Grant Program, eZono AG, and the NSF Graduate Research Fellowship Program.

The authors are with the Department of Electrical Engineering and Computer Sciences and the Department of Mathematics, University of California, Berkeley, Berkeley, CA 94720, USA {lhallock, akashvelu, amanda.schwartz, bajcsy}@berkeley.edu

proof-of-concept data set as a source of initial insights into deformation's time series behavior.

A. Deformation as a Measure of Output Force

While the human musculoskeletal system is highly complex, geometrically irregular, and dominated by the physics of various nonlinear materials, the core mechanism underlying human movement is straightforward: muscles ratchet together (largely via the actin-myosin cross-bridge cycle [6], though other proteins like titin are thought to play a role as well [7]), inducing a length change along the line of action, which pulls the attached (roughly elastic) tendons, which then impart the force to the skeleton. Under the (mild) assumption that muscles are isovolumetric [8, 9], this length change by definition induces a shape change, or deformation, in the activated muscle.

Several isolated studies have established a correlation between muscle activation and shape change [10–15], including our own previous work examining the full 3D extent of the biceps brachii under static (but kinematically varied) elbow loading [3]. This prior study showed evidence that such deformation — measured as muscle cross-sectional area (CSA) or thickness changes — is readily observable via ultrasound, though the appearance of the muscle cross section varies drastically with both sensor location and kinematic configuration. This signal complexity is the result not only of the nonuniform material properties of a given muscle–tendon unit, but of its contact dynamics with surrounding structures. This complexity means that interpreting the deformation signal is not always obvious: while some locations along the arm show a reliable CSA or thickness increase corresponding to increased output force, others may show a decrease or no reliable change, and this varies based on which precise parameterization of deformation is used.

B. Brachioradialis Measurement via Ultrasound as a Proof-of-Concept

As an initial study of the force–deformation relationship, we examine deformation of the brachioradialis (one of several elbow flexors) and its relationship to elbow force output. As in previous studies [3, 4], we focus on isometric elbow flexion as a proof-of-concept motion, as it is a comparatively simple joint (with only one degree of freedom and only a few muscles) that is relevant to the upper-limb modeling cases for which this research may be especially applicable. Also as before, we measure this deformation via 2D B-mode ultrasound, a technology that is safe, portable, and provides a relatively clear image of the fascia between muscles.

To allow collection of time-varying data, we restrict ourselves to a single 2D ultrasound frame collected from the same location along the arm at all time points. We target the brachioradialis for analysis, as it is the smallest of the elbow flexors and its cross-section largely fits within the frame of a single ultrasound scan for many subjects (unlike the biceps or brachialis). Because muscle shape differs greatly across both kinematic configuration (here, elbow angle) and individual subject, we vary each of these in turn, examining the correlation of several muscle deformation signals with output force (in section III) as well as our ability to track these measures (IV).

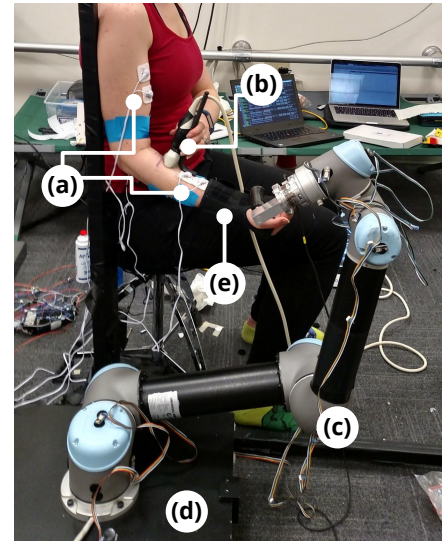


Fig. 1. Experimental setup for the collection of time series force, surface electromyography (sEMG), and ultrasound data under multiple elbow angles. Setup includes sEMG electrodes (a); ultrasound probe (b); UR5 robot (c) with attached handle, through which subject may transmit force to force plate (d); and wrist brace (e) for elbow isolation.

III. MUSCLE FORCE–DEFORMATION CORRELATION

In this section, we present preliminary data indicating that several measures of brachioradialis muscle deformation correlate with with output torque at the elbow and are consistent with simultaneous sEMG data. We first outline our subject cohort and collection procedure, then present preliminary time series data alongside qualitative and quantitative analysis of the force–deformation relationship. Lastly, we comment on how these preliminary analyses suggest future research directions.

A. Data Set Collection

As an exploratory data set, simultaneous sEMG, ultrasound, and output force data were collected from a single subject under multiple kinematic configurations during a sequence of voluntary flexion pulses, in a protocol similar to that used in [16] and illustrated in Fig. 1. To explore consistency across individuals, we collected ultrasound and force data from an additional 4 subjects at a single representative configuration. Details of this subject cohort and collection procedure are outlined below.

1) *Subject Biometric Data & Consent*: Data were collected from the right arms of 5 subjects (2 female, 3 male). Of these, 4 subjects were within age 18–24, while the last subject was age 83 and used to investigate generalizability across age groups. All subjects were healthy, with a wide variety of exercise regimens and body types. The study protocol was approved by the University of California Institutional Review Board for human protection and privacy under Protocol ID 2016-01-8261.

2) *Data Collection*: Each experimental trial consisted of a pulsed sequence of isometric elbow flexions, executed while the subject was instrumented with a lab-developed sEMG system [17] and a 3–12 MHz linear ultrasound transducer (L3–12 NGS, eZono AG, Jena, Germany) attached to its corresponding ultrasound unit (eZono 4000, eZono AG, Jena, Germany). Surface EMG electrodes were placed in

a differential configuration on both the upper and lower arm, targeting the biceps and brachioradialis, respectively. The ultrasound transducer was placed perpendicular to the lower arm (i.e., roughly perpendicular to the brachioradialis) at a consistent marked location and lightly stabilized in the sagittal plane via an L-shaped support. When exerting force, subjects pressed upward on a handle mounted to the end effector of a 6-degree-of-freedom robot arm (UR5, Universal Robots, Odense, Denmark). This robot remained static for each trial, but changing its configuration between trials served as an easy manner of repositioning the handle in space based on subject physiology and desired elbow angle during measurement. The robot was in turn mounted to a 6-channel force plate (Optima-HPS, Advanced Mechanical Technology, Inc., Watertown, MA, USA), which was used to measure the output force exerted by the subject. Additionally, each subject wore a brace to immobilize the wrist.

During data collection, the subject sat comfortably upright, feet planted, with the back of the upper arm supported and elbow flexed to the prescribed angle, as shown in Fig. 1. Subjects were instructed to press upward in the sagittal plane while only exerting force at the elbow and keeping the forearm completely supinated. For each experimental trial, the subject was guided in a series of 21 force pulses, each of 2 s duration with 1 s rest in between, via a visual display that instructed the subject to press with “low”, “medium”, or “high” intensity at each interval. The first 3 intervals were always performed in order of increasing intensity, while the remaining 18 were randomized. For select trials, this pulsed sequence was followed by a set of 3 sustained force presses (each 5 s in duration, with 2 s rest in between) in increasing order of intensity. Note that this protocol was intended primarily to generate a wide variety of force conditions, not to prescribe specific force values; none of our analyses rely on subjects following this protocol exactly, and there is high variance in the extent to which this sequence is readily observable in the collected force data.

3) *Data Scope*: Using the protocol above, simultaneous force, sEMG, and ultrasound data were collected from a single subject (denoted *Sub1*) at elbow angles of 25°, 44°, 69°, 82°, and 97°, as measured from full extension. Simultaneous force and ultrasound data were collected from an additional 4 subjects (denoted *Sub2–Sub5*) at a self-selected, comfortable elbow angle near 69° for comparison with the primary subject’s 69° trial.

B. Data Processing & Definition of Deformation Signals

Prior to our correlation analyses, we preprocessed the raw data streams in the following manner to extract force, neural activation, and deformation measures that are readily comparable. All preprocessing and analysis details have been released on the OpenArm codebase, and an exemplar of the finalized trial data can be seen in Table I.

1) *Force and sEMG Preprocessing*: To generate a single output force magnitude from the collected 6-channel (i.e., 3-channel force and 3-channel torque) force plate data, the measured wrench was first transformed into the handle’s frame of reference based on the robot’s kinematic configuration. The magnitude of the subject’s output force was then calculated

as the magnitude of the force component of the resultant wrench.

To generate muscle activation data from raw sEMG values, ambient noise was first removed with a 60 Hz notch filter. Our final activation measure is the absolute value of this denoised signal, smoothed by an exponential moving average filter. These sEMG values were collected from both the brachioradialis and the biceps brachii, both of which aid in elbow flexion. We present both in the analyses below — the first because it is more closely related to the observed brachioradialis deformation, and the second because it is a mildly cleaner signal and thus constitutes a more competitive baseline for our comparison with deformation measures.

2) *Extraction of Deformation Measures*: Given our time series data of ultrasound images, we can parameterize muscle deformation in a multitude of ways, from shifts in individual pixel values to changes in muscle dimensions. For this preliminary study, we examine the following three shape measures.

Cross-sectional area (CSA). We define *CSA* as the total area of the brachioradialis cross section visible on a given ultrasound scan.

Thickness (T). We define *T* as the maximum extent of the brachioradialis cross section measured from the surface of the skin perpendicularly away from the transducer.

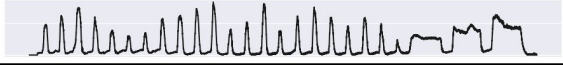
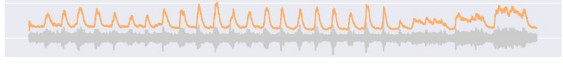




Aspect ratio (AR). We define *AR* as the ratio between the thickness *T* of the brachioradialis cross section and its maximum extent in the perpendicular dimension.

Each of these 3 metrics is calculated by first manually masking the brachioradialis cross section at each time point, then extracting each above-defined value from the mask. We note two salient sources of measurement error. First, the full extent of the brachioradialis cross section often extends beyond the measured ultrasound frame, impacting primarily the *CSA* and *AR* measures. Second, the manual masking process relies on human annotation that is inconsistent and time consuming. We are currently working to address the first through our studies of full 3D muscle deformation, in which we examine the full extent of a given muscle to pinpoint promising deformation signals visible within a single frame. The second limitation is addressed through our studies of automated cross section tracking outlined in section IV.

3) *Alignment of Data Streams*: Due to the technical peculiarities of each sensor, a fully time-synced data collection system remains under development; in this preliminary study, we align the force, sEMG, and ultrasound data streams based on the temporal location of the first force pulse, which is manually noted in each data stream, then linearly interpolate each series to attain a consistent frequency. We make no claims on the precise temporal relationships between force, activation, and deformation signals, but see this as a fascinating direction of future study, particularly with respect to the differences in timing between neurological input signals measured by sEMG and their mechanical deformation outputs.

4) *Ultrasound Drift Compensation*: Although the ultrasound probe was stabilized by a support structure during collection, the precise location of the probe relative to the brachioradialis muscle tends to drift over time, primarily due to subjects’ skin shifting and the probe rotating relative to the arm. To compensate for this drift, we fit a third-

TABLE I
EXAMPLE TIME SERIES DATA

SIGNAL	Symbol	Pearson Correlation Coefficient raw, detrended CC(·, f)	Time Series Signal
force (N)	f	1.00	
sEMG, biceps (mV raw, 0.1 mV processed)	sEMG-BIC	0.38	
sEMG, brachioradialis (mV raw, 0.1 mV processed)	sEMG-BRD	0.27	
cross-sectional area (mm ²)	CSA(-DT)	0.44, 0.48	
thickness (mm)	T(-DT)	0.59, 0.72	
aspect ratio	AR(-DT)	0.70, 0.78	

Example time series data (collected from *Sub1* at 25° elbow flexion) of elbow output force (*black*) in which the subject executed a series of prescribed force pulses and sustained exertions. Data streams include biceps (*light orange*) and brachioradialis (*dark orange*) surface electromyography (*sEMG*) data, as well as detrended ultrasound-measured cross-sectional area (*CSA*, *cyan*), thickness (*T*, *blue*) and aspect ratio (*AR*, *dark blue*) of the brachioradialis cross section. In general, *CSA*, *T*, and *AR* deformation measures correlate well with output force, especially when detrended. Moreover, these force–deformation correlations are of comparable strength to those between force and more standard *sEMG* measures. Note that raw (*gray*) and processed (*orange*) *sEMG* traces are plotted at different scales for clarity as noted, and reported correlation coefficients are for processed data.

order polynomial to each of the *CSA*, *T*, and *AR* data series, measured only at time points at which force was near zero (i.e., between prescribed pulses). We then subtract the value of this polynomial at each point in the original series to generate “detrended” data. We examine both the original deformation data streams and their detrended counterparts in the analyses below.

C. Correlation & Evaluation

In the following analyses, we use the Pearson correlation to assess the viability of using our candidate deformation measures to infer output force, alongside or as an alternative to *sEMG*. We first examine an illustrative time series, then discuss how our assertions translate across changes in kinematic configuration and across subjects.

1) *An Illustrative Time Series*: Table I shows representative pulse data from a single trial (specifically, that of *Sub1* collected at 25°) alongside each data stream’s correlation coefficient as measured against force data. In this series, and in general, our deformation measures *CSA*, *T*, and *AR* correlate comparably with *sEMG* from both the biceps and brachioradialis; in this case, in fact, they are substantially more correlated (though this varies with both elbow angle and subject, as discussed below).

This exemplar also illustrates two other trends largely consistent throughout the data. First, thickness *T* and aspect ratio *AR* are somewhat more correlated with output force than cross-sectional area *CSA*, perhaps because an increase in one dimension is often accompanied by a decrease in the other due to muscles’ isovolumetric nature. Second, detrending the data using the method detailed in section III-B.4 generally improves the observed correlation, providing evidence that sensor drift (or some biological source of drift, e.g., fatigue) indeed occurred.

2) *Deformation and Elbow Angle*: Fig. 2 illustrates the manner in which correlation strength varies with angle for each data stream on our exemplar subject, *Sub1*. For flexion angles less than approximately 90°, deformation measures remain reasonably correlated with force, with detrended data performing slightly better as in our illustrative data series. The relative performance of *CSA*, *T*, and *AR* measures, however, varies in a complex manner that is difficult to characterize but reflected qualitatively in the accompanying illustrative frames, which show significant angle-dependent variation in shape and size. This variation is perhaps most obvious in the highest-flexion 97° data series, in which the cross section changes shape entirely and our defined deformation metrics largely fail, perhaps because the belly of the muscle has shifted significantly away from the ultrasound probe location. Interestingly, *sEMG* correlations behave in the opposite manner, with the largest correlations occurring at the highest flexion angles. While this may simply be a subject-specific phenomenon, it suggests that hybrid sensor approaches — in which ultrasound-measured deformation and *sEMG*-measured activation are employed, depending on kinematic configuration — may hold more promise for robust device control than either technology alone.

3) *Multi-Subject Deformation*: Fig. 3 shows the strength of each deformation signal’s correlation with force on a representative trial (near elbow angle 69°) for each of our 5 subjects. All measures *CSA*, *T*, and *AR* are consistently correlated with a magnitude of around 0.5 or higher even though the subjects vary significantly in terms of muscle morphology, suggesting that the underlying biological mechanism generating the signal is common to all subjects. As before, detrending the data generally provides a mild performance boost if a significant trend was removed, or makes little difference if no trend was observed. (The one exception, in which detrending resulted in decreased correlation, is *Sub4*,

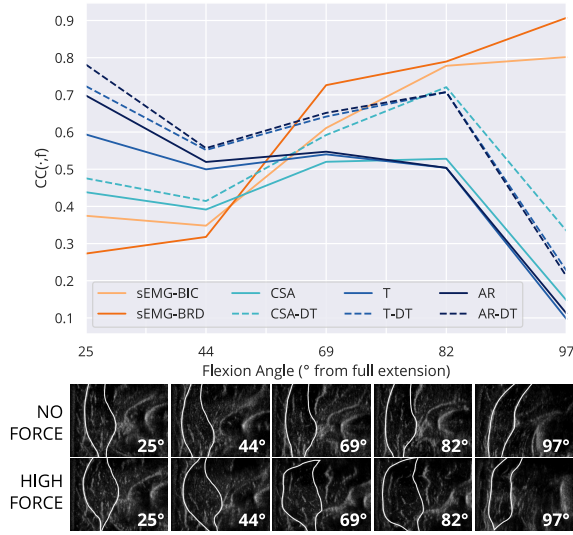


Fig. 2. *Top*: Correlation of cross-sectional area (CSA, cyan), thickness (T, blue) and aspect ratio (AR, dark blue) of the brachioradialis cross section, both raw (solid) and detrended (dashed), with elbow output force, alongside baseline force correlations with biceps (light orange) and brachioradialis (dark orange) surface electromyography (sEMG) data. Correlations were computed across flexion angles ranging from near full extension (25°) to near maximum flexion (97°). CSA, T, and AR deformation signals, especially when detrended, correlate well with elbow output force for most elbow angles, but this correlation collapses near full flexion. *Bottom*: Example ultrasound frames with annotated brachioradialis contours depicting no force (top row) and high output force (bottom row) at each examined flexion angle reflect the changing presentation of muscle deformation with changes in elbow angle.

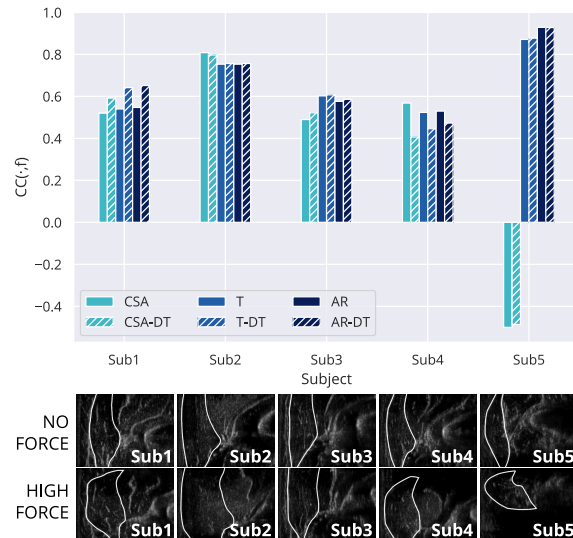


Fig. 3. *Top*: Correlation of cross-sectional area (CSA, cyan), thickness (T, blue) and aspect ratio (AR, dark blue) of the brachioradialis cross section, both raw (solid) and detrended (hashed), with elbow output force, across various subjects, collected at ~69° elbow flexion. CSA, T, and AR deformation signals, especially when detrended, correlate well with output force, though the magnitude and direction of these correlations varies widely across subjects. *Bottom*: Example ultrasound frames with annotated brachioradialis contours depicting no force (top row) and high output force (bottom row) for each subject reflect substantial morphological variation.

for whom a truly meaningful trendline could not be computed due to lack of zero-force baseline data throughout the data series.)

Perhaps the most interesting subject data is that of *Sub5*, our single elderly subject and the only subject physically small enough that virtually the entire brachioradialis cross section was visible in every data frame. Like other subjects' data, the magnitude of each deformation signal's correlation was substantial — in fact, *T* and *AR* signals were more correlated with force than for any other subject, perhaps due to the completeness of the cross-sectional image. Oddly, the subject's *CSA* data, while strongly correlated with force, was negatively correlated (i.e., *CSA* decreased with the application of force), in what could be a quirk of morphology, a function of sensor placement, or a property of aging musculature. We see this question in particular as meriting future study, as aging individuals could especially benefit from advances in muscle force measurement technology. We present this preliminary data as evidence both that measuring deformation data in elderly subjects is possible and that the data collected is largely consistent with that of younger subjects.

D. Future Directions in Force–Deformation Modeling

The sections above constitute preliminary analysis on a limited data set, which we aim to further expand with additional subjects (of varying age and ability), additional force values and durations, fully time-synchronized ultrasound, sEMG, and force data (perhaps via the methods used in [12]), and an improved (e.g., multi-channel) sEMG system for more equitable ultrasound comparison. Such enhancements will allow for expanded understanding of the results above (including the impact of age on the deformation signal) and of phenomena not yet explored (e.g., the temporal sEMG–deformation relationship, the impact of fatigue).

In addition to these data quality enhancements, we also seek an understanding of the biological mechanisms underlying our deformation measures. This is a significant analytical challenge, as shown visually in Fig. 2 and 3's illustrative frames: a single muscle cross section, without accompanying 3D shape data, is difficult or impossible to interpret, and simple deformation measures like those analyzed above barely scratch the surface of its architectural nuance. Informed by the correlations we observe in this paper, we are addressing these challenges from two complementary perspectives. First, we are constructing and analyzing full 3D muscle images to allow for improved interpretation of 2D cross-sections, including what low-dimensional deformation signals are most correlated with force and where they are best observed relative to the underlying skeleton under different kinematic configurations [3, 4]. Second, we are working to track these candidate deformation signals, as well as the movement of individual pixels along and within the muscle contour, to enable the exploration of new and higher-dimensional deformation signals; this is the subject of section IV below.

Note also that although the ultimate goal of this research is to relate measures of muscle deformation to *individual muscle forces*, for the purposes of this preliminary study, we restrict our analyses to *net joint output torque*. This research is interesting precisely because there are no readily available individual force measures with which to compare our data,

and we believe that the deformation–net-force correlations presented here, along with the insights in section II-A regarding the physiological causes of deformation, constitute a powerful case that this signal is a promising candidate for individual muscle force measurement. In the future, we aim to probe this claim both empirically (e.g., via invasive animal study that enables muscle–tendon unit isolation) and through enhanced modeling (e.g., fitting Hill-type muscle models and examining their predictive power).

IV. MUSCLE CROSS SECTION TRACKING

In this section, we present several algorithms to track both our candidate deformation signals (CSA , T , and AR , as defined in section III-B.2 above) and the overall contour of the brachioradialis muscle. We first precisely define this tracking problem and its associated performance metrics, then discuss each algorithm’s performance on the multi-angle, multi-subject data set generated and analyzed in section III.

A. Defining the Tracking Problem

Although any number of tracking schemes, both sparse and dense, could be used to extract varied measures of deformation, we focus on the problem of muscle contour tracking — more precisely, tracking the edges of the brachioradialis muscle as it shifts over time. We choose this formulation for a number of reasons. First, our candidate deformation signals CSA , T , and AR can be readily extracted from the tracked contour. Second, the manually-generated mask data used in section III’s analysis provides a ready ground truth signal for both overall tracking quality (via intersection-over-union computation, for example) and for our previously examined deformation signals of interest. Third, the tracked contour admits future extraction of more complex deformation signals (e.g., statistical shape models or masked dense optical flow measures) while isolating a given muscle of interest from the surrounding tissue.

Lastly, tracking this cross section is feasible, if challenging: the brachioradialis, like many muscles of interest, is surrounded by fascia that appear brighter than surrounding tissue on an ultrasound scan. While the fascial structure is sometimes narrow and difficult to track — a fact that informs our more sophisticated tracking schemes — the preliminary results below indicate that even standard, untuned tracking methods perform relatively well in many cases.

B. Candidate Tracking Algorithms

Tracking biological structures within ultrasound data is an active area of research, with a number of general [18] and structure-specific [11, 19–23] approaches proposed. As a proof of concept, we examine the following candidate algorithms, each of which is built on the standard iterative Lucas–Kanade method of optical flow estimation [24] as implemented in the OpenCV Python library [25].

For the latter two algorithms, we present tracking error results for two sets of parameter values, “general” (tuned via expert-informed grid search on the *Sub1* data series collected at 69° elbow flexion) and “subject-specific” (tuned using each subject’s ~69° data series). The full details of each implementation, including algorithm-specific parameter values and tuning methods, have been released with the analysis

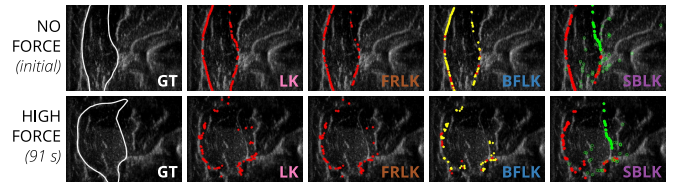


Fig. 4. Example ultrasound frames from *Sub1* 69° data series and their respective tracked contour points using each candidate algorithm, at (relaxed) initialization (*top*) and approximately 91 s into tracking during force exertion (*bottom*). *Center left*: Complete set of contour points (*red*) tracked via naive Lucas–Kanade (*LK*) generally describes ground truth muscle shape (*GT*, *left*), but fails to track more extreme deviations and exhibits significant drift error. *Center*: Refined set of contour points based on feature quality (*red*) and tracked via Lucas–Kanade (*FRLK*) describe muscle shape only mildly more accurately than naive *LK* approach. *Center right*: Refined contour points tracked after aggressive (*red*) and less aggressive (*yellow*) bilateral filtering (*BFLK*) experience slightly less drift error, but still fail to capture deviation in the top right quadrant of the contour where fascia is narrow. *Right*: Combining Lucas–Kanade tracked refined feature points (*red*) with contour points (*green, closed*) predicted based on supporter point locations (*green, open*) (*SBLK*) results in reasonable tracking of the full muscle contour as long as supporter point motion is well-correlated with contour motion.

codebase, and example frames from each tracker are shown in Fig. 4.

Naive Lucas–Kanade (LK). As a tracking baseline, we perform unmodified iterative Lucas–Kanade optical flow on every point along the brachioradialis contour, regardless of feature quality, initialized via the manually segmented contour mask in the first frame.

Feature-Refined Lucas–Kanade (FRLK). As a refinement of the *LK* procedure above, we track only the top 70% of feature points along the manually segmented contour as measured by their Shi-Tomasi corner score [26], a good indicator of the trackability of a point. In general, this refinement results in less drift-associated error from poor feature points but often fails to track narrower sections of the fascia altogether.

Bilaterally-Filtered Lucas–Kanade (BFLK). To further improve the tracking of individual contour points through noise removal without compromising edge integrity, we generate two denoised images at each time point via two separate bilateral filters [27]. The first of these filters is tuned to aggressively suppress speckle noise, but also removes narrower sections of fascia; the second generates a noisier image but retains even narrow fascia sections. Points along the initial manually segmented contour are again culled based on feature quality, this time in each of the filtered images, to generate two sets of high-quality feature points. (Points considered high quality on both filtered images are tracked on the more noise-suppressed frame.) These point sets are then tracked via standard Lucas–Kanade on their respective filtered images, and the contour at each time point is computed from the union of these points. Parameters tuned for subject-specific algorithm implementations include both the noise reduction characteristics of each bilateral filter and the fraction of highest-quality points used for tracking.

Supporter-Based Lucas–Kanade (SBLK). We further refine the *BFLK* method above in a manner similar to that described in [28]: high quality contour points (after aggressive bilateral filtering) are tracked via Lucas–Kanade, and the remainder of the contour is filled in based on the relationship of contour points to “supporter points” of high feature quality extracted from throughout the image. To avoid irregularities caused

by alternating Lucas–Kanade-tracked and supporter-tracked points — and observing that the narrowest, most featureless portion of the contour occurs in the top right quadrant of the contour — we enforce that all supporter-tracked points belong to this quadrant, and all Lucas–Kanade-tracked contour points remain outside. We then again apply two bilateral filters to each frame; using the first and most aggressive, we again obtain the top fraction of contour points by feature quality and track those that lie within the non-supporter quadrants via Lucas–Kanade. Next, using the second and less-aggressive filter, we select a set of “supporter points” of high feature quality, which we assume maintain a consistent distance to each contour point in the supporter-tracked quadrant. We then track these supporter points via Lucas–Kanade and use each supporter point to predict the new location of each contour point based on this constant-displacement assumption, in a manner similar to the “one-shot learning” method described in [29]. We calculate each contour point’s final predicted location as the dynamically weighted mean over all supporter points’ predictions, where we rely more heavily on supporter points that themselves deviate significantly over time from their initial position (to avoid relying too heavily on unmoving background feature points). Parameters tuned for subject-specific algorithm implementations include those tuned in the *BFLK* algorithm, as well as the number of supporter points, the fraction of Lucas–Kanade-tracked points, and properties of the (affine) prediction weighting function.

C. Tracking Error Metrics

In evaluating the performance of the algorithms above, we wish to characterize how well each algorithm tracks our *CSA*, *T*, and *AR* measures of interest, as well as how precisely it tracks the muscle contour overall. We thus present explicit *CSA*, *T*, and *AR* error values along with contour segmentation error, computed as Jaccard distance (i.e., one minus intersection-over-union, or IoU).

D. Preliminary Tracking Performance

In this section, we evaluate our candidate tracking algorithms on the force–deformation data set generated and examined in section III, remarking on each algorithm’s general performance and its performance variation across subjects.

1) *An Illustrative Time Series*: Table II shows aggregate tracking error values over the *Sub3* data series examined in section III. While tracking performance varies by subject and kinematic configuration, these values illustrate several observed trends. First, adding structure to specifically track narrow sections of fascia via the *BFLK* and *SBLK* algorithms often improves performance over naive Lucas–Kanade baselines (though which of these two algorithms gives this performance boost varies by subject, as reflected in Fig. 4). Second, per-subject tuning generally improves tracking performance, as discussed in detail in section IV-D.2 below. Third, thickness *T* is the easiest of our deformation measures to track, and *AR* the most challenging, suggesting that although *AR* correlates slightly more reliably with muscle output force, *T* — which also correlates reasonably well — may be a better choice of control signal.

TABLE II
EXAMPLE TRACKING ERROR

TRACKING ALGORITHM	Symbol	Tracking Error ($M \pm SD$)			
		Jaccard Distance (1-IoU)	Fractional Error		
			<i>CSA</i>	<i>T</i>	<i>AR</i>
Naive Lucas–Kanade	LK	0.17±0.05	0.37±0.22	0.11±0.08	0.29±0.13
Feature-Refined Lucas–Kanade	FRLK	0.19±0.05	0.31±0.18	0.09±0.06	0.28±0.15
Bilaterally-Filtered Lucas–Kanade, <i>general</i>	BFLK-G	0.18±0.06	0.38±0.20	0.11±0.09	0.30±0.12
Bilaterally-Filtered Lucas–Kanade, <i>tuned</i>	BFLK-T	0.16±0.04	0.24±0.13	0.10±0.07	0.31±0.12
Supporter-Based Lucas–Kanade, <i>general</i>	SBLK-G	0.20±0.06	0.30±0.22	0.09±0.05	0.29±0.09
Supporter-Based Lucas–Kanade, <i>tuned</i>	SBLK-T	0.18±0.05	0.19±0.11	0.08±0.05	0.36±0.13

Example tracking error of *LK*, *FRLK*, *BFLK*, and *SBLK* algorithms, both tuned and untuned, on *Sub3* data series, computed as both Jaccard distance (one minus intersection-over-union, or IoU) and fractional error on predicted *CSA*, *T*, and *AR* deformation signals. In general, tuned algorithms outperform untuned algorithms, more structured algorithms (*BFLK* and *SBLK*) often outperform less structured ones, and *T* is tracked with lower mean error than *CSA* or *AR*. Values are presented as mean \pm standard deviation.

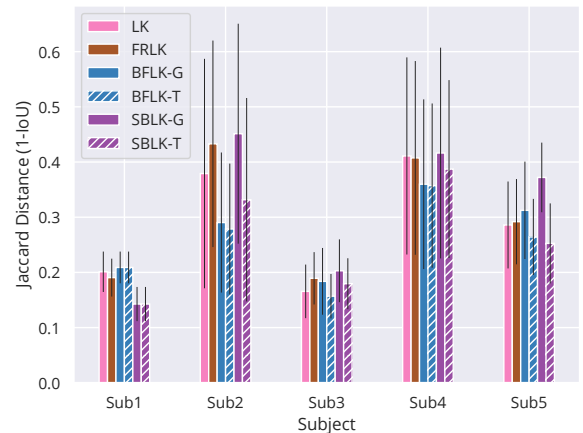


Fig. 5. Mean tracking error and standard deviation of *LK* (pink), *FRLK* (brown), *BFLK* (blue), and *SBLK* (purple) algorithms, both tuned (hashed) and untuned (solid), computed as Jaccard distance, across various subjects, collected at $\sim 69^\circ$ elbow flexion. Both the top performing algorithm and the level of tracking accuracy achieved, as well as the extent to which algorithm tuning mitigated errors, varied substantially by subject, likely due to variations in both morphology and motion quality. Note that for *Sub1*, whose tuned values formed the generic baseline for all tracking, tuned and untuned tracking algorithms are the same.

2) *Multi-Subject Tracking*: Fig. 5 shows tracking error for each proposed algorithm on the five subjects examined in section III. Overall tracking quality (both mean and standard deviation of the error) varies substantially by subject, likely due to both the morphology differences illustrated in Fig. 3 (*Sub2*, for example, had extremely narrow fascia) and differences in motion strategy (*Sub4*, for instance, moved very abruptly). The importance of subject-specific tuning varies as well; *Sub5*, whose morphology varies substantially from that of *Sub1*, showed a large decrease in mean tracking error with tuning, while *Sub3*, whose morphology is more similar, showed a much more modest improvement.

E. Future Directions in Muscle Cross Section Tracking

Both the algorithms examined and our preliminary analyses of their quality are only the first steps toward real-time muscle cross section tracking, and the algorithms we selected were tailored to our particular data set of interest only in the most general sense. True assessment of their quality will require a much more rigorous examination of time series error characteristics (e.g., whether errors occur primarily due to failure to track transient force changes or general point drift), as well as collection of data sets more expressly designed to elucidate tracking error. (The data sets here, by contrast, were primarily collected to assess correlation, and thus contain little variation in speed of force application.) In addition to addressing these concerns, we are currently examining more sophisticated model-based tracking approaches (e.g., edge tracking, applying shape priors) and evaluating the computational feasibility of deploying these approaches for real-time tracking.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have shown that several simple muscle deformation signals are correlated with output force, and that these measures can be observed and tracked over time via ultrasound. Going forward, new insights in the muscle force–deformation relationship should inform the signals we seek to track, but so too should advances in computer vision and graphics inform the candidate deformation signals we consider in the system identification process. The SimTK OpenArm project constitutes a first effort at this type of holistic approach to the muscle deformation characterization problem, and we encourage researchers in the wider biomechanics, robotics, neuroscience, and vision communities to utilize and contribute to this project.

ACKNOWLEDGMENT

The authors acknowledge the aid and advice of Dr. Gregorij Kurillo, Dr. Robert Matthew, Dr. Daisuke Kaneishi, Sarah Seko, Daniel Ho, Jeffrey Zhang, Ian McDonald, Sathvik Nair, Logan Howard, Yannan Tuo, and Vijay Govindarajan in both hardware development and data collection.

REFERENCES

- [1] S. L. Delp et al. “OpenSim: Open-source software to create and analyze dynamic simulations of movement”. In: *IEEE Tran. Biomed. Eng.* 54.11 (2007), pp. 1940–1950.
- [2] M. Damsgaard et al. “Analysis of musculoskeletal systems in the AnyBody Modeling System”. In: *Simulation Modelling Practice and Theory* 14.8 (2006), pp. 1100–1111.
- [3] L. A. Hallock, A. Kato, and R. Bajcsy. “Empirical Quantification and Modeling of Muscle Deformation: Toward Ultrasound-Driven Assistive Device Control”. In: *ICRA*. IEEE. 2018, pp. 1825–1832.
- [4] Y. Nozik et al. “OpenArm 2.0: Automated Segmentation of 3D Tissue Structures for Multi-Subject Study of Muscle Deformation Dynamics”. In: *EMBC*. IEEE. 2019.
- [5] W. Herzog. “Skeletal muscle mechanics: questions, problems and possible solutions”. In: *Journal of NeuroEngineering and Rehabilitation* 14.1 (Dec. 2017), p. 98.
- [6] A. F. Huxley. “Muscle structure and theories of contraction”. In: *Prog. Biophys. Biophys. Chem.* 7 (1957), pp. 255–318.
- [7] W. Herzog et al. *A new paradigm for muscle contraction*. June 2015.
- [8] J. Swammerdam. “The Book of Nature II”. In: *London (UK): Seyffert* (1758), pp. 122–132.
- [9] T. A. McMahon. *Muscles, Reflexes, and Locomotion*. Princeton University Press, 1984.
- [10] J. M. McMeeken et al. “The relationship between EMG and change in thickness of transversus abdominis”. In: *Clinical Biomechanics* 19.4 (2004), pp. 337–342.
- [11] V. J. Schrier et al. “Reliability of ultrasound speckle tracking with singular value decomposition for quantifying displacement in the carpal tunnel”. In: *J. Biomech.* 85 (Mar. 2019), pp. 141–147.
- [12] Q. Zhang, K. Kim, and N. Sharma. “Prediction of Ankle Dorsiflexion Moment by Combined Ultrasound Sonography and Electromyography”. In: *IEEE Tran. Neur. Sys. and Rehab. Eng.* 28.1 (Jan. 2020), pp. 318–327.
- [13] J. Shi et al. “Assessment of muscle fatigue using sonomyography: Muscle thickness change detected from ultrasound images”. In: *Medical Engineering and Physics* 29.4 (2007), pp. 472–479.
- [14] B. Juul-Kristensen et al. “Comparison of muscle sizes and moment arms of two rotator cuff muscles measured by Ultrasonography and Magnetic Resonance Imaging”. In: *European Journal of Ultrasound* 11.3 (June 2000), pp. 161–173.
- [15] J. Y. Guo et al. “Continuous monitoring of electromyography (EMG), mechanomyography (MMG), sonomyography (SMG) and torque output during ramp and step isometric contractions”. In: *Medical Engineering and Physics* 32.9 (Nov. 2010), pp. 1032–1042.
- [16] L. A. Hallock et al. *Sensor-Driven Musculoskeletal Dynamic Modeling*. Tech. rep. University of California, Berkeley, 2016, UCB/EECS–2016–66.
- [17] D. Kaneishi, R. P. Matthew, and M. Tomizuka. “Optimal Control Parameterization for Active/Passive EXoskeleton with Variable Impedance Actuator”. In: *BIOROB*. Vol. 2018-Augus. IEEE Computer Society, Oct. 2018, pp. 713–719.
- [18] S.-M. Kim et al. “The Study of Pre-Processing Algorithm for Improving Performance of Optical Flow Techniques on Ultrasound Image”. In: *Journal of the Institute of Electronics Engineers of Korea SC* 47.5 (2016), pp. 24–32.
- [19] G. Q. Zhou and Y. P. Zheng. “Automatic Fascicle Length Estimation on Muscle Ultrasound Images With an Orientation-Sensitive Segmentation”. In: *IEEE Tran. Biomed. Eng.* 62.12 (2015), pp. 2828–2836.
- [20] I. Mikic, S. Krucinski, and J. D. Thomas. “Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates”. In: *IEEE Tran. Med. Imag.* 17.2 (1998), pp. 274–284.
- [21] E. D. Angelini and O. Gerard. “Review of myocardial motion estimation methods from optical flow tracking on ultrasound data”. In: *EMBC*. 2006, pp. 1537–1540.
- [22] J. W. H. Korstanje et al. “Development and validation of ultrasound speckle tracking to quantify tendon displacement”. In: *J. Biomech.* 43.7 (May 2010), pp. 1373–1379.
- [23] D. J. Farris and G. S. Sawicki. “Human medial gastrocnemius force-velocity behavior shifts with locomotion speed and gait”. In: *PNAS* 109.3 (Jan. 2012), pp. 977–982.
- [24] B. D. Lucas and T. Kanade. “An Iterative Image Registration Technique with an Application to Stereo Vision”. In: *Proceedings of Imaging Understanding Workshop*. 1981, pp. 121–130.
- [25] G. Bradski. “The OpenCV Library”. In: *Dr. Dobb’s Journal of Software Tools* (2000).
- [26] J. Shi and C. Tomasi. “Good features to track”. In: *CVPR*. Publ by IEEE, 1994, pp. 593–600.
- [27] C. Tomasi and R. Manduchi. “Bilateral filtering for gray and color images”. In: *ICCV*. 1998, pp. 839–846.
- [28] E. Ozkan et al. “Robust motion tracking in liver from 2D ultrasound images using supporters”. In: *International Journal of Computer Assisted Radiology and Surgery* 12.6 (June 2017), pp. 941–950.
- [29] H. Grabner et al. “Tracking the invisible: Learning where the object might be”. In: *CVPR*. 2010, pp. 1285–1292.