

Received 7 October 2021; revised 7 December 2021; accepted 1 January 2022.
Date of publication 6 January 2022; date of current version 17 January 2022.

Digital Object Identifier 10.1109/JTEHM.2022.3140973

Classification Performance and Feature Space Characteristics in Individuals With Upper Limb Loss Using Sonomyography

SUSANNAH ENGDAHL¹, ANANYA DHAWAN¹, AHMED BASHATAH¹, GUOQING DIAO²,
BISWARUP MUKHERJEE¹, BRIAN MONROE³, RAHSAAN HOLLEY⁴,
AND SIDDHARTHA SIKDAR¹, (Member, IEEE)

¹Department of Bioengineering, George Mason University, Fairfax, VA 20030, USA

²Department of Biostatistics and Bioinformatics, The George Washington University, Washington, DC 20052, USA

³Hanger Clinic, Laurel, MD 20707, USA

⁴MedStar National Rehabilitation Hospital, Washington, DC 20010, USA

(Susannah Engdahl and Ananya Dhawan are co-first authors.) CORRESPONDING AUTHOR: S. SIKDAR (ssikdar@gmu.edu)

This work was supported in part by the Department of Defense under Award W81XWH-16-1-0722 and in part by the National Institutes of Health under Award U01EB027601.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board, George Mason University, under Approval No. 492701, and the Institutional Review Board, the MedStar National Rehabilitation Network, under Approval No. 2016-173.

This article has supplementary downloadable material available at <https://doi.org/10.1109/JTEHM.2022.3140973>, provided by the authors.

ABSTRACT *Objective:* Sonomyography, or ultrasound-based sensing of muscle deformation, is an emerging modality for upper limb prosthesis control. Although prior studies have shown that individuals with upper limb loss can achieve successful motion classification with sonomyography, it is important to better understand the time-course over which proficiency develops. In this study, we characterized user performance during their initial and subsequent exposures to sonomyography. *Method:* Ultrasound images corresponding to a series of hand gestures were collected from individuals with transradial limb loss under three scenarios: during their initial exposure to sonomyography (Experiment 1), during a subsequent exposure to sonomyography where they were provided biofeedback as part of a training protocol (Experiment 2), and during testing sessions held on different days (Experiment 3). User performance was characterized by offline classification accuracy, as well as metrics describing the consistency and separability of the sonomyography signal patterns in feature space. *Results:* Classification accuracy was high during initial exposure to sonomyography ($96.2 \pm 5.9\%$) and did not systematically change with the provision of biofeedback or on different days. Despite this stable classification performance, some of the feature space metrics changed. *Conclusions:* User performance was strong upon their initial exposure to sonomyography and did not improve with subsequent exposure. *Clinical Impact:* Prosthetists may be able to quickly assess if a patient will be successful with sonomyography without submitting them to an extensive training protocol, leading to earlier socket fabrication and delivery.

INDEX TERMS Upper limb, pre-prosthetic training, prosthesis control, sonomyography, feature space.

I. INTRODUCTION

Despite the enormous investment of resources in the development of new multi-articulated upper limb prosthetics, a large proportion of individuals with upper limb loss discontinue use of their prosthesis [1]–[3]. Users often experience dissatisfaction with the function and control of their prosthesis [4], [5], so it is crucial that they receive training to mitigate the challenges associated with using the device [6], [7]. Although training has been correlated

with increased prosthesis use [8] and satisfaction [9], patient access to rehabilitation and prosthetic services in the United States is frequently limited [10]. There is also a scarcity of clinicians who specialize in treating upper limb loss [11] and possess the specialized knowledge required to train patients on effective prosthesis use. The difficulty of learning to use a prosthesis may be apparent given certain limitations in the predominant method for sensing and decoding user intent, surface electromyography (EMG). EMG is limited by poor

amplitude resolution and low signal-to-noise ratio, especially with dry electrodes used in prosthesis sockets [12], [13]. There is also low specificity between muscles due to cross-talk and co-activation, especially for deep-seated muscle groups in the forearm that are responsible for finger movement [14]–[17]. Consequently, multi-articulated hands tend to rely on direct control strategies for opening and closing the terminal device in which EMG signals are recorded from an agonist-antagonist muscle pair. Other methods for controlling multi-articulated prosthetic hands rely on pattern recognition to decode user intent from EMG signal patterns. Although pattern recognition algorithms enable successful real-time grasp classification [18]–[20] and allow for control of a prosthetic hand during real-world functional tasks [21]–[24], users and therapists both report that extended periods of training are typically necessary to achieve stable performance [25], [26].

Given the challenges related to training and reliably decoding user intent with surface EMG, some researchers are pursuing an alternative approach called sonomyography (SMG), which uses ultrasound to sense muscle deformations in the residual limb. Although surface EMG provides temporal information about muscle deformation based on electrical features, ultrasound permits both temporal and spatial characterization of the deformation based on the acquired images. This spatiotemporal information is accessible even for muscles in deep-seated compartments, thus avoiding the problem of low specificity. Numerous prior studies have demonstrated clear potential for the use of SMG in controlling a prosthesis or other human-machine interface [27]–[34]. Our own work has shown that SMG is capable of accurately classifying motor intent for individuals without upper limb loss [35], [36] and individuals with upper limb loss [37], [38] in both offline and real-time settings. In particular, we have demonstrated an offline classification accuracy of 96% for five grasps in individuals with upper limb loss [38]. While these early studies have demonstrated robust classification performance is possible with SMG, the primary clinical benefits have not yet been established. To successfully translate SMG to clinical practice, it is important to better understand the clinical need that SMG can fill. In particular, it remains unknown how quickly prosthesis users can learn to use SMG and repeatably isolate control signals. Our study is focused on investigating this question.

Regardless of the specific prosthesis control strategy, the process by which naïve patients learn to use their device encompasses multiple stages [39], including pre-prosthetic and prosthetic training. The pre-prosthetic training stage involves learning how to generate the requisite control signals for operating a prosthesis, but is usually accomplished without the use of a physical prosthesis. Users are provided various sources of biofeedback, often within a virtual environment, to help them understand what happens physiologically when they perform a movement and how to modulate that activity for prosthetic control. For example, they might view a real-time representation of the EMG

signal to demonstrate how muscle contractions are linked to electrical activity. This functionality is available in several commercial products, such as Ottobock MyoBoy, Ottobock Myo Plus, and Coapt Complete ControlRoom. The prosthetic training stage includes performing functional tasks with a physical prosthesis to become proficient using it in real-world settings. In this paper, we will restrict our discussion to the pre-prosthetic training stage.

During pre-prosthetic training for EMG pattern recognition, the user must learn to produce a specialized set of EMG patterns that are sufficiently consistent and separable from each other to permit accurate gesture classification. This can be difficult since people generally do not have experience modulating EMG signal amplitudes [40]. Individuals with limb loss may be further disadvantaged by motor cortex reorganization following amputation [41], as well as muscle atrophy due to disuse of the residual limb and/or increased reliance on the intact limb [42]. Given these difficulties, it is unsurprising that first attempts to use pattern recognition are often error-prone. For example, one study reported an average initial classification accuracy for individuals with transradial limb loss to be 77.5% for nine motion classes [43]. Training over the course of multiple sessions or days appears to mitigate some of these errors for individuals with and without limb loss, regardless of whether feedback on their performance is provided [43]–[46]. These improvements are credited to changes in the EMG signal patterns such that they become more consistent and/or separable, although the correlation between performance and EMG pattern characteristics is complex and not yet fully understood [46]–[49].

Beyond the primary purpose of helping patients learn how to use a prosthesis, pre-prosthetic training serves an important clinical function by helping prosthetists to understand whether their patients are cognitively and physiologically capable of using a particular control modality. Fabricating and fitting an upper limb prosthesis is time-consuming and expensive, especially if multi-articulated prosthetic hands are included, so prosthetists must ensure patients are suited to the control modality prior to beginning the process [50]. In the case of EMG pattern recognition, patients must be able to produce consistent and separable EMG signal patterns. However, it can be difficult to demonstrate this ability quickly given the prolonged time period needed to develop proficiency, creating a burden on both the prosthetist and patient. Reduced pre-prosthetic training times and quicker assessment of a prospective user's ability to generate consistent and separable control signals may help improve a patient's ease of access to care.

In this study, we investigated whether users can quickly and repeatably isolate SMG control signals during pre-prosthetic training. We characterized user performance during their initial and subsequent exposures to SMG in order to better understand the time-course over which proficiency develops. Performance was characterized by offline classification accuracy, as well as metrics describing the consistency and separability of SMG signal patterns in feature space.

TABLE 1. Participant characteristics.

ID	Sex	Years since amputation	Prior prosthesis experience ^a	Affected limb	Experiment 1	Experiment 2	Experiment 3 ^c
Am1	M	46	Direct control (experienced)	Right	X	X	X (Session 1 – 2 datasets, Session 2 – 11 datasets)
Am2	M	Congenital	unknown	Left	X		
Am3	M	50	Direct control (experienced), Patten recognition (novice)	Left	X	X	X (Session 1 – 6 datasets, Session 2 – 11 datasets)
Am4	M	7.5	Patten recognition (experienced)	Left	X		
Am5	M	1.5	Direct control (novice), Patten recognition (novice)	Right ^b	X	X	X (Session 1 – 11 datasets, Session 2 – 4 datasets)
Am6	M	1	Direct control (novice), Patten recognition (novice)	Right	X		
Am7	M	1	Direct control (novice)	Left	X	X	
Am8	F	12	Direct control (experienced)	Right ^b	X	X	

^aNovice indicates less than two years of experience

^bParticipant was affected bilaterally

^cDifferent numbers of datasets were collected during Experiment 3 due to variation in the participants' availability and stamina

We asked three questions: 1) What is the performance during initial exposure to SMG? 2) Is biofeedback useful in helping users change their performance? 3) Is performance repeatable across multiple exposures to SMG?

II. METHODS

A. SUBJECTS

We recruited eight participants, including seven individuals with transradial limb amputation and one individual with congenital limb absence (Table 1). These individuals reported using myoelectric prostheses in their daily lives, but were naïve to the use of SMG prior to beginning the study. All individuals provided written informed consent prior to participating in this study, which was approved by the Institutional Review Boards at George Mason University (#492701, approved Oct. 24, 2013) and Med-Star National Rehabilitation Network (#2016-173, approved Sept. 21, 2016).

B. PROTOCOL

We conducted three different experiments to evaluate our research questions, with some individuals participating in only a subset of the experiments depending on their availability.

1) EXPERIMENT 1

All eight participants were included in Experiment 1. Data collection was performed using a clinical ultrasound system (Terason uSmart 3200T, Terason, Burlington, MA). A low-profile, high-frequency, linear 16HL7 ultrasound transducer was positioned on the volar aspect of participants' residual limb using a stretchable fabric cuff such that muscle deformations associated with all individual phantom finger movements were visually identifiable on the ultrasound images. Ultrasound image sequences were acquired and transferred to a PC in real-time using a USB-based video grabber

(DVI2USB 3.0, Epiphan Systems, Inc.). The captured screen was then downsampled to 100 × 140 pixels to include only the relevant ultrasound image. The acquired image frames were processed in MATLAB (MathWorks, Natick, MA) using custom algorithms.

To create a dataset for training the classifier, participants performed repeated iterations of one motion from a set of motions that they felt were intuitive to perform. Am2 and Am3 performed power grasp, wrist pronation, thumb flexion, and index finger flexion. Am6 performed wrist pronation and supination, wrist flexion and extension, power grasp, ulnar deviation, and thumb flexion. The other participants performed power grasp, wrist pronation, key grasp, tripod, and index point.

Starting from a resting position, participants followed an auditory cue and moved towards the end state of the desired motion over the course of one second, held the end state position for one second, moved back to rest over the course of one second, and remained at rest for one second. After they repeated this process five times in succession, we extracted the ultrasound image frames corresponding to the motion end state and rest. This process was repeated until all five motions were included in the dataset. Once the dataset was completed, we performed leave-one-out cross-validation with a modified 1-nearest-neighbor classifier that used Pearson's correlation coefficient as a similarity measure [36]. Our modified classifier averages the similarity measurements by class and selects the most similar class instead of selecting the most similar individual image.

2) EXPERIMENT 2

A subset of five participants was included in Experiment 2. Data collection procedures were identical to Experiment 1, except that participants instead created a series of datasets across two different phases of data collection. Both phases

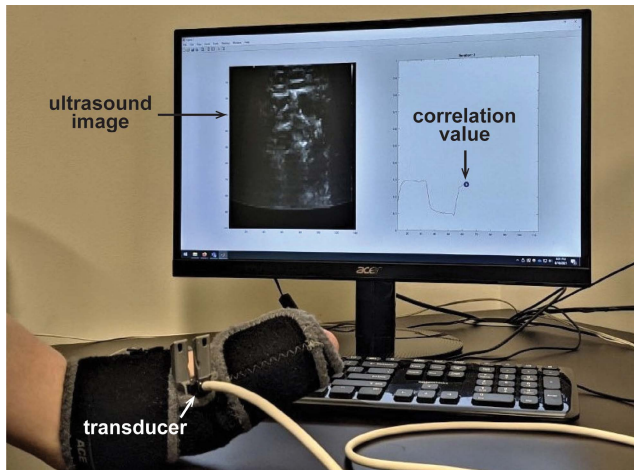


FIGURE 1. Visual display used to provide feedback to participants during Experiment 2. The Pearson correlation between the current ultrasound image and the first image recorded in the sequence is indicated by the blue circle. High correlation values indicate that current image is dissimilar from rest, while low values indicate similarity to rest. Correlation values for previous images remained on the display (red line) to help participants achieve consistency in the muscle deformation pattern.

occurred during a single day. Additionally, all participants performed the same five motions in Experiment 2 (power grasp, wrist pronation, key grasp, tripod, index point).

During the first phase (*baseline*), participants were asked to create three datasets to establish their baseline performance for that day. We did not provide any feedback to participants about their performance during this phase.

During the second phase (*feedback*), participants were asked to create three datasets while receiving verbal and visual biofeedback about their performance. Specifically, participants were allowed to view the ultrasound images in real-time so that they could understand their residual limb anatomy and how their muscles deformed when attempting each grasp. They were also asked to describe their sensations of phantom hand movement associated with each grasp and to demonstrate the movement using their intact hand. Based on their explanation, we offered suggestions on how they could make their movements more separable.

Participants were also given a visual cue to help them monitor the consistency in their muscle deformation patterns (Fig. 1). For each movement sequence, we calculated the Pearson correlation between the first ultrasound image frame (corresponding to a rest state) and the incoming image. The correlation value was inverted and graphically displayed in real-time such that high values indicated dissimilarity from rest (i.e., motion end state) and low values indicated similarity to rest. A “plateau” of high correlation values during the one-second hold period for each motion end state would indicate that the muscle deformation pattern was similar between successive images. Thus, while this visual display does not reveal which individual muscles are being contracted, it can help participants achieve consistency in the overall deformation pattern.

Finally, participants were told the results of the cross-validation after creating each dataset and were shown the associated confusion matrix to help them understand the source of any errors. They were also given suggestions on how they could alter their movements to try and improve the classification accuracy. For example, power grasp and key grasps were occasionally confused because these movements are fairly similar. In this case, participants might be instructed to increase thumb adduction during key grasp to further differentiate the motion from power grasp.

Although best practices for teaching SMG have not yet been established, we believe that the general principles used for teaching EMG pattern recognition should be applicable. Thus, most of the strategies we used for delivering biofeedback have been recommended in existing literature. This includes asking the user to move the phantom limb or intact hand to mirror the intended grasp [39], [51], showing the user raw signals to demonstrate that different grasps are associated with different muscle activation patterns [39], [51], offering suggestions on how to make movements more separable [51], and showing the user a confusion matrix to demonstrate which grasps were confused [51]. While the Pearson correlation display was a unique approach to this study, other studies have used visual displays to help users develop consistency in their muscle activation (e.g., [48]). It is also important to note that we did not follow a rigorously structured protocol for delivering identical feedback to every participant, but this format is similar to what is done in clinical settings for patients learning EMG pattern recognition. The prosthetist or therapist may use the same general techniques for all patients, such as providing a visual display to show how different movements produce distinct muscle activity patterns or asking patients to consider the movement of their phantom hands. However, training is conducted under the clinician’s supervision so that specific instructions can be customized for each patient based on their residual limb anatomy and rehabilitation goals [39], [47], [51]. Similarly, we provided the same general sources of feedback to our participants but customized the specific instructions when necessary.

3) EXPERIMENT 3

A subset of three participants was included Experiment 3. Data collection procedures were identical to Experiments 1, except that participants created a series of datasets across two data collection sessions occurring on separate days. All participants performed the full set of five motions in Experiment 2 (power grasp, wrist pronation, key grasp, tripod, index point). We intended to collect as many datasets as possible from each participant depending on their availability and stamina. Consequently, we obtained differing numbers of datasets across participants and sessions (Table 1).

C. DATA ANALYSIS

The primary outcome metric was cross-validation accuracy (1), defined as the percent of data correctly classified during

the leave-one out validation process for a given dataset i :

$$CA_i = 100 * \frac{P_{correct_i}}{P_{total_i}}, \quad (1)$$

where $P_{correct_i}$ is the correct number of predictions by the closest-class classifier and P_{total_i} is the total number of predictions (i.e., the total number of datapoints).

Cross-validation accuracy is a combined measurement of the user's ability to perform a motion and the classifier's ability to label individual motion performances. Since user performance and classifier performance are inherently linked in this metric, it is possible that a user's performance could change over time without affecting the cross-validation accuracy. For example, a user may perform the tripod grasp with very little variation for a given dataset, resulting in a high cross-validation accuracy. On the next dataset, they may perform the grasp with two different variations having slightly different levels of middle finger flexion. As long as the closest identified class for each of the variations is still tripod, the cross-validation accuracy would be unaffected. Therefore, we wanted to more appropriately understand the changes in user performance independent of the classifier performance.

To accomplish this goal, we performed a dimensionality reduction on all individual datasets collected from the participants during each experiment. The datasets contained all ultrasound images that were recorded for each grasp. Every 100×140 pixel image in the dataset was represented as points in 14,000-dimensional space such that each pixel in the image corresponded to an axis in the high dimensional space. This high dimensional space was then reduced to five dimensions through principal components analysis. Next, we defined point clusters in five-dimensional space such that each cluster was comprised of all the points in an associated motion class. We then utilized several metrics to describe the characteristics of these clusters in SMG feature space, including Within-class Distance (WD), Inter-class Distance Nearest Neighbor (IDNN), Inter-class Distance All Neighbors (IDAN), Most Separable Dimension (MSD), and Mean Semi-Principal Axis (MSA) [46]. These metrics characterize the clusters' consistency or separability. If a participant's performance of a given motion becomes more similar to other performances of the same motion, the points in that motion cluster move closer together and the consistency increases. If a participant's performance of a given motion becomes more distinct from the performances of another motion, the clusters move further apart and the separability increases. A more detailed explanation of each metric is provided below.

1) WITHIN-CLASS DISTANCE

WD is a measure of consistency between all five repetitions of the same motion

$$WD_j = \sum_{r=1}^5 \sum_{k=1}^5 \frac{dist_{kj}^{rj} * dist_{rj}^{kj}}{dist_{kj}^{rj} + dist_{rj}^{kj}} \quad (2)$$

Here, $dist_{kj}^{rj}$ is half the Mahalanobis distance in feature space between repetitions r and k of motion j , and $dist_{rj}^{kj}$ is half the Mahalanobis distance in feature space between repetitions k and r of motion j

$$dist_{kj}^{rj} = \frac{1}{2} \sqrt{((\mu_{Trj} - \mu_{Tkj})^T * S_{Trj}^{-1} * (\mu_{Trj} - \mu_{Tkj}))} \quad (3)$$

$$dist_{rj}^{kj} = \frac{1}{2} \sqrt{((\mu_{Tkj} - \mu_{Trj})^T * S_{Tkj}^{-1} * (\mu_{Tkj} - \mu_{Trj}))} \quad (4)$$

where μ_{Trj} and μ_{Tkj} represent the feature vectors from repetitions r and k , respectively. S_{Trj} and S_{Tkj} are the covariances from repetitions r and k , respectively. The total WD for each participant is defined as the mean WD over all n motions:

$$WD_{total} = \frac{1}{n} \sum_{j=1}^n WD_j. \quad (5)$$

2) INTER-CLASS DISTANCE NEAREST NEIGHBOR

IDNN is a measure of separability between different motions

$$IDNN_i = \min_{i=1, \dots, j-1, j+1, \dots, n} \frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j} \quad (6)$$

such that only the distance to a motion's nearest neighbor is included. Here, $dist_j^i$ is half the Mahalanobis distance in feature space between motions i and j , and $dist_i^j$ is half the Mahalanobis distance in feature space between motions j and i

$$dist_j^i = \frac{1}{2} \sqrt{((\mu_{Ti} - \mu_{Tj})^T * S_{Ti}^{-1} * (\mu_{Ti} - \mu_{Tj}))} \quad (7)$$

$$dist_i^j = \frac{1}{2} \sqrt{((\mu_{Tj} - \mu_{Ti})^T * S_{Tj}^{-1} * (\mu_{Tj} - \mu_{Ti}))} \quad (8)$$

where μ_{Ti} and μ_{Tj} represent the feature vectors from motions i and j , respectively. S_{Ti} and S_{Tj} are the covariances from motions i and j , respectively. The total IDNN for each participant is defined as the mean IDNN over all n motions:

$$IDNN_{total} = \frac{1}{n} \sum_{j=1}^n IDNN_j. \quad (9)$$

3) INTER-CLASS DISTANCE ALL NEIGHBORS

IDAN is a measure of separability between different motions and is similar to IDNN, except that distances to all neighbors are included

$$IDAN_i = \sum_{j=1}^n \frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j}. \quad (10)$$

The total IDAN for each participant is defined as the mean IDAN over all n motions:

$$IDAN_{total} = \frac{1}{n} \sum_{j=1}^n IDAN_j. \quad (11)$$

4) MOST SEPARABLE DIMENSION

MSD is similar to IDNN, except that only one of the five dimensions is used to determine the nearest neighbor. In this case, the dimension with the largest separability is chosen

$$MSD_i = \min_{i=1, \dots, j-1, j+1, \dots, n} \left(\max_{d=1, \dots, 5} \frac{dist_j^i * dist_i^j}{dist_j^i + dist_i^j} \right) \quad (12)$$

The total MSD for each participant is defined as the mean MSD over all n motions:

$$MSD_{total} = \frac{1}{n} \sum_{j=1}^n MSD_j. \quad (13)$$

5) MEAN SEMI-PRINCIPAL AXIS

MSA is a measure of variability for a given motion and is calculated as a geometric mean

$$MSA_k = \left(\prod_{k=1}^5 a_k \right)^{\frac{1}{5}} \quad (14)$$

where a_k is the semi-principal axes of the hyperellipsoids. The total MSA for each participant is defined as the mean MSA over all n motions:

$$MSA_{total} = \frac{1}{n} \sum_{j=1}^n MSA_j. \quad (15)$$

D. STATISTICAL ANALYSIS

For Experiment 2, we compared each outcome metric (cross-validation accuracy, WD, IDNN, IDAN, MSD, MSA) from the *baseline* phase to the means from the *feedback* phase using the following linear mixed model

$$Y_{ij} = \beta_0 + b_i + \beta_1 X_{ij} + \epsilon_{ij}, \quad (16)$$

where Y_{ij} is the outcome metric for the i th subject at the j th measurement, X_{ij} is a dichotomous variable with value 0 for the *baseline* phase and 1 for the *feedback* phase, ϵ_{ij} is the residual error, and b_i is a random intercept accounting for within-subject correlations among repeated measures. Both b_i and ϵ_{ij} are assumed to be normally distributed and independent. The baseline phase is treated as the reference level. To account for the small sample size and potential violation of the model assumptions, we used the permutation test [52] to assess significance ($\alpha = 0.05$). One-sided p-values were based on 1000 permuted samples.

To assess whether there was a change over time for the outcomes, we fit the same model as (16) but replaced the dichotomous variable X_{ij} with the normalized time from the first measurement in the *baseline* phase. It is worth noting that it is not appropriate to include both normalized time and the dichotomous feedback variable in the linear mixed models since they are highly correlated with a Pearson correlation coefficient 0.923 ($p < 1.0e-12$).

To evaluate the effect of the feature space metrics on cross-validation accuracy, we fit the following linear mixed model

$$Y_{ij} = \beta_0 + b_i + \beta_1 X_{ij1} + \beta_2 WD_{ij} + \beta_3 IDNN_{ij} + \beta_4 IDAN_{ij} + \beta_5 MSD_{ij} + \beta_6 MSA_{ij} + \epsilon_{ij}, \quad (17)$$

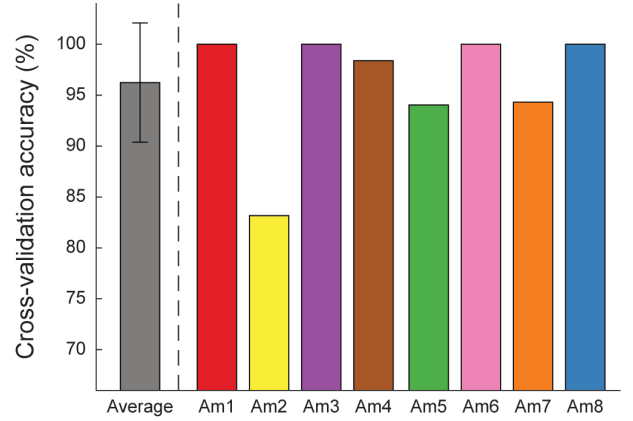


FIGURE 2. Between-subject average (grey bar) and per-subject (colored bars) cross-validation accuracy. Error bar represents standard deviation.

where X_{ij} is either a dichotomous variable with value 0 for the *baseline* phase and 1 for the *feedback* phase or the normalized time, and WD_{ij} , $IDNN_{ij}$, $IDAN_{ij}$, MSD_{ij} , MSA_{ij} are the feature space metrics. The baseline phase is treated as the reference level. Two-sided p-values were based on 1000 permuted samples.

For Experiment 3, we compared each outcome metric between sessions 1 and 2. We fit the following linear mixed model

$$Y_{ij} = \beta_0 + b_i + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \epsilon_{ij}, \quad (18)$$

where X_{ij1} is the normalized time from the first measurement in the *baseline* phase, and X_{ij2} takes value 0 for session 1 and 1 for session 2. Compared to (16), this model includes two covariates. Since the phase variable and the normalized time are highly correlated and there were very few observations for some phases in some subjects, we choose to compare the mean outcomes between session 1 and session 2 while controlling for the confounding effects of the normalized time. To account for the small sample size and potential violation of the model assumptions, we used the permutation test [52] to assess significance ($\alpha = 0.05$). Two-sided p-values were based on 1000 permuted samples.

To evaluate the effect of the feature space metrics on cross-validation accuracy, we fit the following linear mixed model

$$Y_{ij} = \beta_0 + b_i + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \beta_3 WD_{ij} + \beta_4 IDNN_{ij} + \beta_5 IDAN_{ij} + \beta_6 MSD_{ij} + \beta_7 MSA_{ij} + \epsilon_{ij}, \quad (19)$$

where X_{ij1} is the normalized time from the first measurement in the *baseline* phase, and X_{ij2} takes value 0 for session 1 and 1 for session 2. Two-sided p-values were based on 1000 permuted samples.

III. RESULTS

A. EXPERIMENT 1

The average cross-validation accuracy across subjects was $96.2 \pm 5.9\%$, with a range of 83.2% to 100% (Fig. 2). The SMG feature space metrics were moderately correlated with

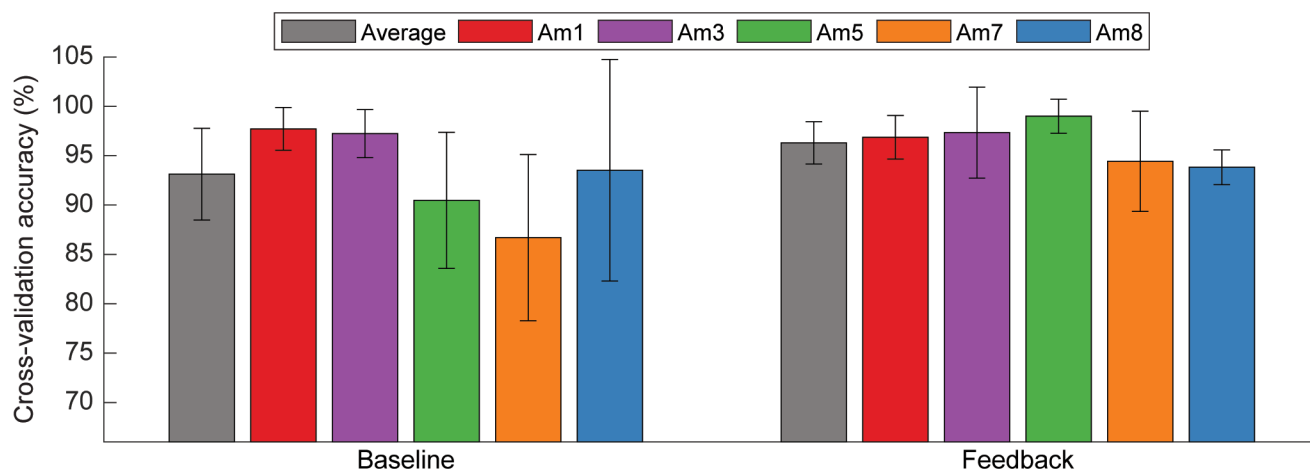


FIGURE 3. Average between-subject (grey bars) and within-subject (colored bars) cross-validation accuracy for each phase. Error bars represent standard deviation.

cross-validation accuracy ($0.47 \leq |r| \leq 0.76$; Supplementary Fig. S1), although none of the correlations were significant except for WD ($p = 0.03$). Grasp-specific metrics did not follow a consistent trend, but wrist pronation and/or supination often had the lowest MSA and highest IDNN, IDAN, and MSD among all grasps collected from individual participants (Supplementary Figs. S2-S6).

B. EXPERIMENT 2

Cross-validation accuracy exceeded 77% for all 30 datasets collected across participants. Furthermore, 19 of these datasets had a cross-validation accuracy of at least 95%. The average cross-validation accuracy was 93% or higher for both phases (*baseline*: $93.1 \pm 4.6\%$; *feedback*: $96.3 \pm 2.1\%$; Fig. 3). The average elapsed collection time was 69 ± 47 minutes (Supplementary Fig. S7), although this value is elevated by the unusually long testing time for Am7 (151 minutes). The elapsed time was much more consistent across the remaining subjects (49 ± 14 minutes). There was no significant effect of phase ($p = 0.08$) or elapsed time ($p = 0.141$) on cross-validation accuracy (Supplementary Tables S1 and S2).

Although overall cross-validation accuracy for all five grasps was generally high, the accuracy values for individual grasps reveal that there was occasional misclassification. However, visual inspection of the misclassification rates across all six datasets for each participant shows no obvious patterns over time (Supplementary Fig. S8).

There was a significant increase in IDNN ($p = 0.012$) and IDAN ($p = 0.012$) from the *baseline* phase to the *feedback* phase. Similarly, there was a significant increase in IDNN ($p = 0.032$) and IDAN ($p = 0.018$) over time. There were no significant changes in the other feature space metrics between phases or over time (Supplementary Tables S1 and S2). The overall range for each metric varied slightly between participants (Supplementary Fig. S9). Of the five feature

space metrics, only MSD had a marginally significant effect on cross-validation accuracy (Supplementary Table S3).

C. EXPERIMENT 3

The cross-validation accuracy did not change significantly between sessions (session 1: $96.9 \pm 2.7\%$, session 2: $96.7 \pm 1.4\%$; $p = 0.586$; Fig. 4). There was a significant increase in IDNN ($p = 0.013$), IDAN ($p = 0.038$), and MSA ($p < 0.001$) between sessions, but no change in the other feature space metrics (Supplementary Table S4). The overall range for each metric varied slightly between participants (Supplementary Figs. S10-S12). None of the five feature space metrics had a significant effect on cross-validation accuracy (Supplementary Table S5).

IV. DISCUSSION

The primary purpose of this study was to characterize user performance during their first and subsequent exposures to SMG. We demonstrated that participants' classification performance during their first experience with SMG was very strong ($96.2 \pm 5.9\%$). Am2 had the poorest performance of all the participants (83.2%), which may be a consequence of having congenital limb absence. It is possible that Am2 had a limited phantom hand sensations or proprioceptive sense in his residual limb, making it challenging to perform the hand gestures. Nonetheless, his performance was still stronger than some first-time users of EMG pattern recognition, who have achieved classification accuracies as low as 77.5% for nine motion classes [43]. Am2's performance may have improved with additional practice, but we could not test this hypothesis because he was unable to participate in Experiment 2.

We found that most participants were able to generate the requisite control signals on their first try, even without biofeedback. There may be different explanations, but one possibility is that SMG provides both spatial and temporal information within the ultrasound image sequence about the user's muscle deformation. This means the problem of low specificity between muscles due to cross-talk and

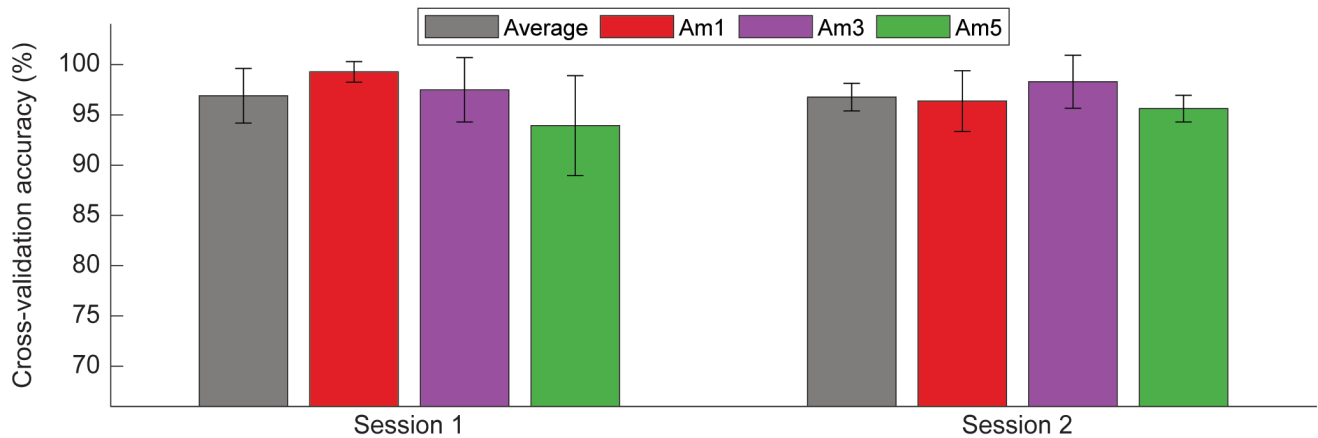


FIGURE 4. Average between-subjects (grey bars) and within-subject (colored bars) cross-validation accuracy for each session. Error bars represent standard deviation.

co-activation is avoided. In turn, users are able to generate muscle contractions that are congruent with the intended grasp, without needing to modify the contraction to make it more distinct from other grasps. These results indicate that one important clinical benefit of SMG is to enable prosthetists to quickly assess a prospective user's ability to generate consistent and separable control signals. This assessment may allow the prosthetist to determine an appropriate treatment plan without requiring the user to go through a lengthy pre-prosthetic training phase.

Although our participants' classification performance was strong even upon their first exposure to SMG, we asked whether their performance could improve further. In fact, their classification performance did not systematically change with the provision of biofeedback or between different data collection sessions. Nonetheless, it should be acknowledged that we cannot fully distinguish between the effects of repetitive practice and provision of biofeedback on classification performance during Experiment 2. Inclusion of a control group that underwent repeated exposures to SMG without biofeedback could have clarified whether these factors are dissociable, but it was infeasible to recruit an additional set of participants with limb loss. Another possibility would be to include an interaction term in the linear mixed models, but we chose not to do this because of the small sample size. However, visual inspection of the results leads us to believe that an interaction term would not have been significant even if it was included.

Most participants experienced small fluctuations in performance between datasets due to isolated misclassifications. These misclassifications may result from problems such as movement of the transducer with respect to the residual limb or minor variations in the muscle deformation patterns, which likely can be addressed with appropriate intervention (i.e., a more secure transducer mounting system or a classification algorithm less sensitive to variation than 1-nearest neighbor). While we do not believe the transient offline misclassifications observed in this study negate the potential viability of

SMG as a control modality, additional study will be required to determine how significantly real-time misclassifications impact user performance in functional settings.

The lack of an online evaluation in this study is a limitation, as it is widely acknowledged that offline classification performance is not necessarily a definitive predictor of real-time control ability [53]. A variety of virtual testing environments, such as the Target Achievement Control test [54], have been developed to characterize user control strategies beyond simple classification accuracy and we are planning to include similar approaches in future studies. Testing with a physical prosthesis will also be necessary to explore the effect of factors like changes in arm position, sensor shifting, sweating, muscle fatigue, or changes in signal characteristics over time [55], which can degrade classification accuracy and require users to retrain the classifier after some period of use. We plan to characterize these issues in future work, as this is a crucial step in demonstrating the real-world viability of SMG.

We also noted inconsistencies in how the SMG feature space metrics changed with provision of biofeedback or between different data collection sessions. For both Experiments 2 and 3, distance between clusters increased according to IDNN and IDAN, suggesting that separability between grasps improved with biofeedback or between days. MSA also increased between days for Experiment 3, which indicates greater variability in how the grasps were performed.

Interestingly, the overall classification accuracy was stable despite these alterations in how participants performed the grasps. While it might be expected that classification accuracy would increase with greater intercluster distance and decrease with greater cluster variability, these relationships were not evident in our results. Given that the feature space metrics only explained 49.8-59.5% of the variability in classification accuracy (Supplemental Tables S3 and S5), this could mean that classification accuracy is related to other feature space patterns that were not quantified here.

Alternatively, this finding may indicate that there are many ways to perform a grasp without sacrificing classification accuracy. The motor system is redundant and the same movement can be achieved through multiple muscle activation patterns. It is possible that our analysis method of using dimensionality reduction accounted for this redundancy during movement classification, but the differences in activation patterns persist in the other feature space metrics. Regardless, there is a precedent for this complex relationship between classification performance and feature space within the EMG literature, which has similarly reported discrepancies in how classification performance relates to EMG signal pattern characteristics [43]–[49]. However, most of these studies quantified real-time classification performance, rather than offline performance like in our study. Further comparison with the EMG literature should be withheld until real-time SMG performance can be studied.

Our finding that participants were immediately able to generate separable movements and could consistently repeat those movements represents a significant benefit of SMG. In comparison, people do not naturally have much experience with modulating EMG patterns [40], nor is it clear which EMG signal pattern characteristics are most relevant to classification performance [46]. It is therefore difficult to know how to effectively train users on EMG pattern recognition. As a result, pre-prosthetic training protocols can be lengthy, involving practice over multiple sessions or days with [43], [46] or without [44], [45] the provision of external feedback in order to improve classification performance. Developing training protocols and delineating relationships between signal pattern characteristics and classification performance appears to be less critical with SMG, as users seem capable of achieving successful classification without intervention.

Shortening the pre-prosthetic training phase with SMG could enable patients to devote more resources towards functional training with a physical prosthesis, which may still require involvement from a therapist. Furthermore, rapid pre-prosthetic training can simplify the process by which prosthetists evaluate whether patients are cognitively and physiologically able to use SMG. This evaluation could perhaps be completed in a single clinical visit, allowing prosthetists to proceed more quickly with socket fabrication if the patient has demonstrated an ability to use SMG or to recommend an alternative control modality if they are unable to use SMG.

A reduction in the amount of pre-prosthetic training may also help diminish barriers to prosthesis access in the United States, where few clinicians specialize in caring for people with upper limb loss or have experience with justifying a course of treatment to insurers [11]. For these reasons, it is perhaps unsurprising that one survey reported 35% of individuals with unilateral upper limb loss received no training of any kind and only 22% received more than 10 hours of training from a prosthetist or therapist [56]. Therapy is an essential component of the rehabilitation

process and the receipt of training to use a first prosthesis has been associated with increased satisfaction [9]. Thus, we hypothesize that SMG may create a potential for increased satisfaction without the need for extensive involvement from a therapist. Experiencing an early sense of accomplishment from successfully learning the control strategy may also motivate users to continue practicing with the prosthesis and could reduce the likelihood that they abandon prosthesis use. We hope to investigate this hypothesis in future studies.

One limitation of this study is that sample size was small and the variability between participants may have reduced our ability to detect statistically significant results. While having a larger sample size would be useful, it should also be acknowledged that the characteristics of individuals with upper limb loss can be very heterogeneous and generalizing beyond the study sample should be done cautiously. As an example, Am7 had poorer classification performance in Experiment 2 compared to the other participants and required over twice as much time to create each dataset (Supplementary Fig. S7). Am7 had undergone amputation about one year prior to this study and was an extremely inexperienced myoelectric prosthesis user, having owned his prosthesis for only one week. He had significant muscle atrophy in his residual limb as a result of this disuse, which may have contributed to his difficulties with generating muscle contractions and relaxing his muscles to a “resting” position between repeated grasps. While these challenges wouldn’t necessarily preclude him from becoming a proficient SMG user, it is important to consider factors like this when assessing a patient’s potential for success with SMG. Thus, evaluations should be made on a case-by-case basis.

Additionally, we did not quantify the degree to which participants changed their performance after specific feedback was given. Since the initial classification accuracy was high, there was limited room for improvement. In future studies involving functional testing with a terminal device, we will quantify the specific ways in which feedback can improve performance.

Another limitation is that we utilized a commercially-available ultrasound imaging system with an array transducer. For translation of SMG technology to practical prosthesis sockets, we anticipate utilizing single-element transducers with low power electronics. Our previous work has indicated that the classification accuracy with sparse sensing is not compromised [57]. However, this result has yet to be validated in individuals with limb loss. We are currently developing fully-integrated prototype SMG systems and additional studies are planned in the future.

Finally, it should be noted that the reported classification accuracies were obtained using a 1-nearest neighbor classifier. We purposely utilized one of the simplest classifiers in an effort to decouple user performance from classifier performance. More sophisticated classifiers commonly used for EMG pattern recognition, such as linear discriminant analysis, are expected to provide improved classification accuracy.

V. CONCLUSION

This study quantified classification performance and the associated SMG feature space for individuals with upper limb loss during their initial and subsequent exposures to SMG. We showed that participants could immediately achieve high motion classification accuracies and that their performance did not change with the provision of biofeedback or across multiple exposures to SMG. Some of the SMG feature space characteristics changed despite the stable classification accuracy, suggesting that these metrics do not fully predict classification performance. If SMG is deployed clinically in the future, the process of assessing patient suitability for SMG during pre-prosthetic training could be completed quickly, which ultimately may improve patient access to timely and appropriate care.

ACKNOWLEDGMENT

The authors would like to thank the participants for their involvement in this study. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense or National Institutes of Health.

REFERENCES

- [1] E. A. Biddiss and T. T. Chau, "Upper limb prosthesis use and abandonment: A survey of the last 25 years," *Prosthet. Orthot. Int.*, vol. 31, no. 3, pp. 236–257, 2007, doi: [10.1080/03093640600994581](https://doi.org/10.1080/03093640600994581).
- [2] K. Østlie, I. M. Lesjø, R. J. Franklin, B. Garfelt, O. H. Skjeldal, and P. Magnus, "Prosthesis rejection in acquired major upper-limb amputees: A population-based survey," *Disab. Rehabil. Assist. Technol.*, vol. 7, no. 4, pp. 294–303, 2012, doi: [10.3109/17483107.2011.635405](https://doi.org/10.3109/17483107.2011.635405).
- [3] L. V. McFarland, S. L. H. Winkler, A. W. Heinemann, M. Jones, and A. Esquenazi, "Unilateral upper-limb loss: Satisfaction and prosthetic-device use in veterans and servicemembers from Vietnam and OIF/OEF conflicts," *J. Rehabil. Res. Develop.*, vol. 47, no. 4, pp. 299–316, 2010.
- [4] P. J. Kyberd and W. Hill, "Survey of upper limb prosthesis users in Sweden, the United Kingdom and Canada," *Prosthet. Orthot. Int.*, vol. 35, no. 2, pp. 234–241, 2011, doi: [10.1177/0309364611409099](https://doi.org/10.1177/0309364611409099).
- [5] E. Biddiss, D. Beaton, and T. Chau, "Consumer design priorities for upper limb prosthetics," *Disabil. Rehabil. Assist. Technol.*, vol. 2, no. 6, pp. 346–357, 2007.
- [6] Management of Upper Extremity Amputation Rehabilitation Working Group. (2014). *VA/DoD Clinical Practice Guideline for the Management of Upper Extremity Amputation Rehabilitation*. Accessed: Jan. 1, 2015. [Online]. Available: <https://www.healthquality.va.gov/guidelines/Rehab/UEAR/VAoDoDCPGManagementofUEAR121614Corrected508.pdf>
- [7] L. M. Smurr, K. Gulick, K. Yancosek, and O. Ganz, "Managing the upper extremity amputee: A protocol for success," *J. Hand Therapy*, vol. 21, no. 2, pp. 160–176, 2008.
- [8] K. Østlie, I. M. Lesjø, R. J. Franklin, B. Garfelt, O. H. Skjeldal, and P. Magnus, "Prosthesis use in adult acquired major upper-limb amputees: Patterns of wear, prosthetic skills and the actual use of prostheses in activities of daily life," *Disab. Rehabil., Assistive Technol.*, vol. 7, no. 6, pp. 479–493, Nov. 2012.
- [9] L. Resnik, M. Borgia, A. W. Heinemann, and M. A. Clark, "Prosthesis satisfaction in a national sample of veterans with upper limb amputation," *Prosthetics Orthotics Int.*, vol. 44, no. 2, pp. 81–91, Apr. 2020, doi: [10.1177/0309364619895201](https://doi.org/10.1177/0309364619895201).
- [10] C. P. F. Pasquina, A. J. Carvalho, and T. P. Sheehan, "Ethics in rehabilitation: Access to prosthetics and quality care following amputation," *AMA J. Ethics*, vol. 17, no. 6, pp. 535–546, 2015.
- [11] *The Promise of Assistive Technology to Enhance Activity and Work Participation*, Nat. Academies Sci., Eng., Med., Nat. Academies Press, Washington, DC, USA, 2017.
- [12] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J. Electromyogr. Kinesiol.*, vol. 12, no. 1, pp. 1–16, Feb. 2002.
- [13] B. Graimann and H. Dietl, "Introduction to upper limb prosthetics," in *Introduction to Neural Engineering for Motor Rehabilitation*, D. Farina, W. Jensen, M. Akay, Eds. Hoboken, NJ, USA: Wiley, 2013, pp. 267–290, doi: [10.1002/9781118628522.ch14](https://doi.org/10.1002/9781118628522.ch14).
- [14] Y. K. Kong, M. S. Hallbeck, and M. C. Jung, "Crosstalk effect on surface electromyogram of the forearm flexors during a static grip task," *J. Electromyogr. Kinesiol.*, vol. 20, no. 6, pp. 1223–1229, 2010, doi: [10.1016/j.jelekin.2010.08.001](https://doi.org/10.1016/j.jelekin.2010.08.001).
- [15] H. van Duinen, S. C. Gandevia, and J. L. Taylor, "Voluntary activation of the different compartments of the flexor digitorum profundus," *J. Neurophysiol.*, vol. 104, no. 6, pp. 3213–3221, 2010, doi: [10.1152/jn.00470.2010](https://doi.org/10.1152/jn.00470.2010).
- [16] H. van Duinen, W. S. Yu, and S. C. Gandevia, "Limited ability to extend the digits of the human hand independently with extensor digitorum," *J. Physiol.*, vol. 587, no. 20, pp. 4799–4810, 2009, doi: [10.1113/jphysiol.2009.177964](https://doi.org/10.1113/jphysiol.2009.177964).
- [17] T. L. McIsaac and A. J. Fuglevand, "Motor-unit synchrony within and across compartments of the human flexor digitorum superficialis," *J. Neurophysiol.*, vol. 97, no. 1, pp. 550–556, 2007, doi: [10.1152/jn.01071.2006](https://doi.org/10.1152/jn.01071.2006).
- [18] C. Cipriani et al., "Online myoelectric control of a dexterous hand prosthesis by transradial amputees," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 3, pp. 260–270, Jan. 2011, doi: [10.1109/TNSRE.2011.2108667](https://doi.org/10.1109/TNSRE.2011.2108667).
- [19] N. Jiang, H. Rehbaum, I. Vujaklija, B. Graimann, and D. Farina, "Intuitive, online, simultaneous, and proportional myoelectric control over two degrees-of-freedom in upper limb amputees," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 501–510, Aug. 2014, doi: [10.1109/TNSRE.2013.2278411](https://doi.org/10.1109/TNSRE.2013.2278411).
- [20] A. M. Simon, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Target achievement control test: Evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses," *J. Rehabil. Res. Dev.*, vol. 48, no. 6, pp. 27–619, 2011.
- [21] S. Amsuess et al., "Context-dependent upper limb prosthesis control for natural and robust use," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 7, pp. 744–753, Jul. 2016, doi: [10.1109/TNSRE.2015.2454240](https://doi.org/10.1109/TNSRE.2015.2454240).
- [22] T. A. Kuiken, L. A. Miller, K. Turner, and L. J. Hargrove, "A comparison of pattern recognition control and direct control of a multiple Degree-of-Freedom transradial prosthesis," *IEEE J. Transl. Eng. Health Med.*, vol. 4, pp. 1–8, 2016, doi: [10.1109/JTEHM.2016.2616123](https://doi.org/10.1109/JTEHM.2016.2616123).
- [23] L. Resnik, H. Huang, A. Winslow, D. L. Crouch, F. Zhang, and N. Wolk, "Evaluation of EMG pattern recognition for upper limb prosthesis control: A case study in comparison with direct myoelectric control," *J. NeuroEng. Rehabil.*, vol. 15, no. 1, p. 23, 2018, doi: [10.1186/s12984-018-0361-3](https://doi.org/10.1186/s12984-018-0361-3).
- [24] A. M. Simon, K. L. Turner, L. A. Miller, L. J. Hargrove, and T. A. Kuiken, "Pattern recognition and direct control home use of a multi-articulating hand prosthesis," in *Proc. IEEE 16th Int. Conf. Rehabil. Robot. (ICORR)*, Jun. 2019, pp. 386–391, doi: [10.1109/ICORR.2019.8779539](https://doi.org/10.1109/ICORR.2019.8779539).
- [25] A. W. Franzke et al., "Users' and therapists' perceptions of myoelectric multi-function upper limb prostheses with conventional and pattern recognition control," *PLoS ONE*, vol. 14, no. 8, Aug. 2019, Art. no. e0220899, doi: [10.1371/journal.pone.0220899](https://doi.org/10.1371/journal.pone.0220899).
- [26] L. J. Resnik, F. Acluche, and S. Lieberman Klinger, "User experience of controlling the DEKA arm with EMG pattern recognition," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0203987.
- [27] Y. P. Zheng, M. M. F. Chan, J. Shi, X. Chen, and Q. H. Huang, "Sonomyography: Monitoring morphological changes of forearm muscles in actions with the feasibility for the control of powered prosthesis," *Med. Eng. Phys.*, vol. 28, no. 5, pp. 405–415, Jun. 2006, doi: [10.1016/j.medengphy.2005.07.012](https://doi.org/10.1016/j.medengphy.2005.07.012).
- [28] J. Guo, Y. Zheng, Q. Huang, and X. Chen, "Dynamic monitoring of forearm muscles using one-dimensional sonomyography system," *J. Rehabil. Res. Dev.*, vol. 45, no. 1, p. 187, 2008.
- [29] X. Chen, Y.-P. Zheng, J.-Y. Guo, and J. Shi, "Sonomyography (SMG) control for powered prosthetic hand: A study with normal subjects," *Ultrasound Med. Biol.*, vol. 36, no. 7, pp. 1076–1088, Jul. 2010, doi: [10.1016/j.ultrasmedbio.2010.04.015](https://doi.org/10.1016/j.ultrasmedbio.2010.04.015).
- [30] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 6, pp. 788–797, Jul. 2012, doi: [10.1109/TNSRE.2012.2207916](https://doi.org/10.1109/TNSRE.2012.2207916).

- [31] J. Shi, J. Y. Guo, S. X. Hu, and Y. P. Zheng, "Recognition of finger flexion motion from ultrasound image: A feasibility study," *Ultrasound Med. Biol.*, vol. 38, no. 10, pp. 1695–1704, 2012, doi: [10.1016/j.ultrasmedbio.2012.04.021](https://doi.org/10.1016/j.ultrasmedbio.2012.04.021).
- [32] X. Yang, X. Sun, D. Zhou, Y. Li, and H. Liu, "Towards wearable a-mode ultrasound sensing for real-time finger motion recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 6, pp. 1199–1208, Jun. 2018, doi: [10.1109/TNSRE.2018.2829913](https://doi.org/10.1109/TNSRE.2018.2829913).
- [33] X. Yang, Z. Chen, N. Hettiarachchi, J. Yan, and H. Liu, "A wearable ultrasound sensor for sensing muscular morphological deformations," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 51, no. 6, pp. 1–10, Jun. 2019, doi: [10.1109/TSMC.2019.2924984](https://doi.org/10.1109/TSMC.2019.2924984).
- [34] X. Yang, J. Yan, Z. Chen, H. Ding, and H. Liu, "A proportional pattern recognition control scheme for wearable A-mode ultrasound sensing," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 800–808, Jan. 2020, doi: [10.1109/TIE.2019.2898614](https://doi.org/10.1109/TIE.2019.2898614).
- [35] S. Sikdar *et al.*, "Novel method for predicting dexterous individual finger movements by imaging muscle activity using a wearable ultrasonic system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 69–76, Jan. 2014.
- [36] N. Akhlaghi *et al.*, "Real-time classification of hand motions using ultrasound imaging of forearm muscles," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 8, pp. 1687–1698, Aug. 2016, doi: [10.1109/TBME.2015.2498124](https://doi.org/10.1109/TBME.2015.2498124).
- [37] C. A. Baker, N. Akhlaghi, H. Rangwala, J. Kosecka, and S. Sikdar, "Real-time, ultrasound-based control of a virtual hand by a trans-radial amputee," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2016, pp. 3219–3222, doi: [10.1109/EMBC.2016.7591414](https://doi.org/10.1109/EMBC.2016.7591414).
- [38] A. S. Dhawan *et al.*, "Proprioceptive sonomyographic control: A novel method for intuitive and proportional control of multiple degrees-of-freedom for individuals with upper extremity limb loss," *Sci. Rep.*, vol. 9, no. 1, p. 9499, Dec. 2019, doi: [10.1038/s41598-019-45459-7](https://doi.org/10.1038/s41598-019-45459-7).
- [39] A. M. Simon, B. A. Lock, and K. A. Stubblefield, "Patient training for functional use of pattern recognition-controlled prostheses," *JPO J. Prosthetics Orthotics*, vol. 24, no. 2, pp. 56–64, 2012.
- [40] R. E. Johnson, K. P. Kording, L. J. Hargrove, and J. W. Sensinger, "EMG versus torque control of human-machine systems: Equalizing control signal variability does not equalize error or uncertainty," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 660–667, Jun. 2017, doi: [10.1109/TNSRE.2016.2598095](https://doi.org/10.1109/TNSRE.2016.2598095).
- [41] G. Di Pino, A. Maravita, L. Zollo, E. Guglielmelli, and V. Di Lazzaro, "Augmentation-related brain plasticity," *Frontiers Syst. Neurosci.*, vol. 8, p. 109, Jun. 2014, doi: [10.3389/fnsys.2014.00109](https://doi.org/10.3389/fnsys.2014.00109).
- [42] A. Chadwell *et al.*, "Upper limb activity in myoelectric prosthesis users is biased towards the intact limb and appears unrelated to goal-directed task performance," *Sci. Rep.*, vol. 8, no. 1, p. 11084, Dec. 2018, doi: [10.1038/s41598-018-29503-6](https://doi.org/10.1038/s41598-018-29503-6).
- [43] M. Powell, R. Kaliki, and N. Thakor, "User training for pattern recognition-based myoelectric prostheses: Improving phantom limb movement consistency and distinguishability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 522–532, May 2014, doi: [10.1109/TNSRE.2013.2279737](https://doi.org/10.1109/TNSRE.2013.2279737).
- [44] N. E. Bunderson and T. A. Kuiken, "Quantification of feature space changes with experience during electromyogram pattern recognition control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 3, pp. 239–246, May 2012, doi: [10.1109/TNSRE.2011.2182525](https://doi.org/10.1109/TNSRE.2011.2182525).
- [45] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: Implications for EMG pattern recognition in prosthesis control," *J. Neural Eng.*, vol. 12, no. 4, Jun. 2015, Art. no. 046005.
- [46] M. B. Kristoffersen, A. W. Franzke, C. K. van der Sluis, A. Murgia, and R. M. Bongers, "The effect of feedback during training sessions on learning pattern-recognition-based prosthesis control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2087–2096, Oct. 2019, doi: [10.1109/TNSRE.2019.2929917](https://doi.org/10.1109/TNSRE.2019.2929917).
- [47] M. B. Kristoffersen, A. W. Franzke, C. K. van der Sluis, A. Murgia, and R. M. Bongers, "Serious gaming to generate separated and consistent EMG patterns in pattern-recognition prosthesis control," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102140, doi: [10.1016/j.bspc.2020.102140](https://doi.org/10.1016/j.bspc.2020.102140).
- [48] M. B. Kristoffersen, A. W. Franzke, R. M. Bongers, M. Wand, A. Murgia, and C. K. van der Sluis, "User training for machine learning controlled upper limb prostheses: A serious game approach," *J. NeuroEngineering Rehabil.*, vol. 18, no. 1, p. 32, Feb. 2021, doi: [10.1186/s12984-021-00831-5](https://doi.org/10.1186/s12984-021-00831-5).
- [49] A. W. Franzke, M. B. Kristoffersen, V. Jayaram, C. K. van der Sluis, A. Murgia, and R. M. Bongers, "Exploring the relationship between EMG feature space characteristics and control performance in machine learning myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 21–30, 2021, doi: [10.1109/TNSRE.2020.3029873](https://doi.org/10.1109/TNSRE.2020.3029873).
- [50] D. H. Gates, S. M. Engdahl, and A. Davis, "Recommendations for the successful implementation of upper limb prosthetic technology," *Hand Clinics*, vol. 37, no. 3, pp. 457–466, Aug. 2021, doi: [10.1016/j.hcl.2021.05.007](https://doi.org/10.1016/j.hcl.2021.05.007).
- [51] M. A. Powell and N. V. Thakor, "A training strategy for learning pattern recognition control for myoelectric prostheses," *J. Prosthetics Orthotics*, vol. 25, no. 1, pp. 30–41, Jan. 2013, doi: [10.1097/JPO.0b013e31827af7c1](https://doi.org/10.1097/JPO.0b013e31827af7c1).
- [52] J. Higgins, *An Introduction to Modern Nonparametric Statistics*. Pacific Grove, CA, USA: Brooks/Cole, 2004.
- [53] M. Ortiz-Catalan, F. Rouhani, R. Branemark, and B. Hakansson, "Offline accuracy: A potentially misleading metric in myoelectric pattern recognition for prosthetic control," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Nov. 2015, pp. 1140–1143, doi: [10.1109/EMBC.2015.7318567](https://doi.org/10.1109/EMBC.2015.7318567).
- [54] S. M. Wurth and L. J. Hargrove, "A real-time comparison between direct control, sequential pattern recognition control and simultaneous pattern recognition control using a Fitts' law style assessment procedure," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, p. 91, 2014, doi: [10.1186/1743-0003-11-91](https://doi.org/10.1186/1743-0003-11-91).
- [55] I. Kyranou, S. Vijayakumar, and M. S. Erden, "Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses," *Frontiers Neurobot.*, vol. 12, p. 58, Sep. 2018, doi: [10.3389/fnbot.2018.00058](https://doi.org/10.3389/fnbot.2018.00058).
- [56] S. Kestner, "Defining the relationship between prosthetic wrist function and its use in performing work tasks and activities of daily living," *JPO J. Prosthetics Orthotics*, vol. 18, no. 3, pp. 80–86, Jul. 2006.
- [57] N. Akhlaghi *et al.*, "Sparsity analysis of a sonomyographic muscle-computer interface," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 3, pp. 688–696, Mar. 2020, doi: [10.1109/TBME.2019.2919488](https://doi.org/10.1109/TBME.2019.2919488).

• • •