AALBORG UNIVERSITY
STUDENT REPORT

Examining if Confidence Score Feedback During User Training Can Improve Users' Ability in Controlling Upper Limb Prosthetics

2nd semester Masters, Biomedical Engineering & Infomatics - Spring 2018

Project group: 18gr8408

Simon Bruun, Oliver Thomsen Damsgaard, Martin Alexander Garenfeld, Christian Korfitz Mortensen

Examining if Confidence Score Feedback During User Training Can Improve Users' Ability in Controlling Upper Limb Prosthetics

Simon Bruun*, Oliver T. Damsgaard*, Martin A. Garenfeld* and Christian K. Mortensen*
** Undergrad. Student, AAU*

Abstract

The user rejection rate of myoelectric prosthetics is currently high, due to slow and inaccurate control. Previous studies have shown user training to be an important part of overcoming the challenge of making transradial upper limb prosthetics more accurate, as the control systems depend on the user generating the same distinguishable muscle patterns when using the prosthesis. Different methods have been sought when adapting users to perform specific distinguishable movements. This study aimed to investigate whether confidence score feedback from a LDA classifier during user training could improve user performance in a Fitts' Law test compared to a control group who only received label feedback. 16 able-bodied subjects were recruited for the study; 8 subjects randomly assigned to each group. Each subject went through a three session experiment; one session per day over three consecutive days. During each session the subject received 16 minutes of user training and went subsequently through a Fitts' Law test to evaluate the performance. A significant improvement in cluster dispersion of EMG signals of separate movement was found in the control group, where the third session resulted in more dense clusters both when compared to the first session of the control group and third session of the test group. The results from the Fitts' Law test showed no significant difference between the two groups and no improvement over the three sessions for either of the groups. Overall, three sessions of user training showed to be an insufficient training period to observe any significant improvement within and between subject groups.

Keywords: surface electromyography, lower arm prosthetics, linear discriminant analysis, user training, confidence scores.

I. INTRODUCTION

The loss of any part of a limb or limb as a whole is a debilitating condition for any human. The hand is one of the most precious tools humans have and thus a loss of this would prove to be a great loss of functionality and independence. In an effort to restore some of that ability and autonomy, many patients are provided with an upper limb prosthesis. [1]

In recent years, myoelectric prosthetics have become increasingly better in performance in a clinical environment, where the number of degrees of freedom (DOF's) in prosthetic hands have increased. However, the ability to control these DOF's are limited by the need for more complex control systems, leading to a lack of functionality when used in daily life tasks. In addition to the restricted control, the discomfort of prosthetics are causing patients to reject their provided prosthetic device with a 23-45 % rejection rate. [2] Commercially available prosthetics range from passive cosmetic prosthetics to functional few DOF's cable-driven prosthetics and more advanced myoelectric prosthetics.

In recent years, several complex multi DOF prosthetic hands have been developed. Examples of these are the Vincent hand by Vincent Systems, iLimb hands from Touch Bionics, the Bebionic hands from RSL Steeper and the Michelangelo hand from Otto Bock. [3] Despite the efforts to advance and improve the functionality of prosthetics, a critical bottleneck still exist: the ability to properly control the prosthesis [4].

Most commercially available myoelectric controlled prosthetics rely on manually switching between different DOF's in the prosthesis [5]. This is a robust control scheme, but it is slow and non-intuitive in movement. In the research area of myoelectric prosthetics, newer control schemes have been developed. These control schemes are classification- or regression-based. Classification have been used for many years in research, but is to date only used in one commercially available prosthesis [6]. When using classification as a control scheme the classifier attempts to classify similar patterns in electromyography (EMG) signals based on previously acquired training data sets and real-time acquired samples [5, 7]. The regression control scheme determines the output signal for an input, based on a regression model. This provides a continuous output value, facilitating simultaneous control contrary to classification which provides only a single class output at a time. [4, 8] In both control schemes, the general challenge for users is to be able to consistently produce distinguishable muscle patterns, which enables the control system to recognize the performed movements accurately. [9]

In recent years many advancements have been made in research on system training. System training is the training of the control algorithm to enable the system to recognize the input signals from the user. This area focuses on the design of the hardware and software side of the system in EMG prosthetics. [5] The awareness in the research area shows a very single-minded approach to possible improvements of control, as system training has been the main focus area of the research [10]. Jiang et al. [10] discussed that other methods of improving prosthetic devices have been underestimated. One such implementation, which have been addressed in only a few studies, is user training [9, 11, 12]. Contrary to system training, user training focuses on the user's ability to control a prosthesis [5]. User training is a focused training of the user in learning to adapt to the control system, before the actual use of the prosthesis in daily life. This is carried out in order to train the user in performing more distinguishable movements. Here different types of feedback can be used to inform the user on how well the system recognizes the users performed movements. [9, 13]

In a 2014 study, Powell et al. [9] provided amputees with real-time visual feedback of an animated prosthesis. This feedback provided the user with visual information on how the prosthesis would move related to which contractions the user performed. Through a 10 session experiment the subject group experienced an average increase in movement completion percentage from 70.8 % to 99.0 %. Fang et al. [11] provided real-time visual feedback of subjects' performed movement in relation to the classes defined in the system. The feedback visualized a 2D map of movement cluster centroids based on a principle component analysis (PCA) of the EMG. The users were instructed in matching the centroid of an online PCA based cursor to the centroid on the 2D map corresponding to the performed movement. When subjects could match the cursor to the centroid of a movement cluster, the performed movement corresponded the best with the class of that movement. The study demonstrated steady improvement of hand motion accuracy compared to conventional label feedback. [11] Based on the findings of these studies other training schemes should be examined in order to find whether other methods could improve user performance further.

A 2013 study by Scheme et al. [14] proposed a novel approach of utilizing confidence scores from a Linear Discriminant Analysis (LDA) classifier to aid the control scheme to either accept or reject the class output. The system functioned by the principle that for each input value the likelihood of it belonging to a certain class was calculated, and used in the process of deciding in which class the input belonged, and if the likelihood was high enough to not be rejected. These likelihoods called confidence scores, were calculated from a modification of Bayes' theorem. Scheme et al. [14] showed a significant improvement in performance with the use of the rejection-capable system when compared to a non-rejection based LDA classifier. A similar approach could be used in user training by providing the confidence scores of the classification to the user as a form of visual feedback.

Thus, this study propose a novel method of providing users with feedback containing confidence scores representing how well the classification model recognizes the performed movements when using a LDA based control scheme during user training. Contrary to current feedback methods in user training this approach could enable users to better understand how the classification works based on their performed movements. Additionally, this proposal of user training could improve the user's ability to produce more distinguishable movements by showing them which classes the system recognizes. This would give the possibility for the user to modulate their EMG patterns in order to increase the confidence of the classifier.

Based on the presented possibilities in improving user training, this study will seek to investigate the use of confidence scores as visual feedback to improve the users' ability to control a transradial prosthesis. This is done under the hypothesis that exposing subjects to user training, in which confidence scores of movement class recognition is used as visual feedback, will show statistically significant improvement in performance in a Fitts' Law test compared to a control group receiving label feedback

The remainder of this paper will cover the following: Section II will cover the calculation of confidence scores. Section III will cover how the study design was structured in order to test the application of confidence scores in user training. Afterwards the three interfaces for data acquisition, user training and performance test will be presented respectively. Section IV will present the findings and in section V the findings will be discussed and analysed. Section VI concludes the study.

II. BACKGROUND

As the application of confidence scores during user training is the main focus of this study, a brief derivation of confidence scores will be presented in this section. This theoretical derivation of confidence scores from a LDA classifier is based on a study by Scheme et al. [14].

The decision rule for LDA classification is based on deciding the class with the highest probability of having produced a given input sample. LDA classification is derived from Bayes principles [15], from which the Bayes theorem expresses that the posterior probability $P(\omega_j|x)$, the probability of sample x belonging to class j, can be written as:

$$P(\omega_j|x) = \frac{P(x|\omega_j)P(\omega_j)}{P(x)} \tag{1}$$

Where $P(x|\omega_j)$ is the class conditional probability, the likelihood that a sample from class j occurs, $P(\omega_j)$ is the prior probability, the probability of class j occurring, and $P(x)$ is the normalization factor that ensures the probabilities of all class sum to 1. As $P(x)$ is common for all classes, it can be excluded, which leaves the following function:

$$g_j(x) = P(x|\omega_j)P(\omega_j) \tag{2}$$

Where $g_j(x)$ is a decision boundary. An assumption of LDA is that each class belongs to a Gaussian distribution. Thus, the class conditional probability can be written as the multivariate normal distribution:

$$P(x|\omega_j) = \frac{1}{|\Sigma_j|^{1/2}}(\frac{1}{\sqrt{2\pi}})^d \; e^{-1/2}(x-\mu_j)' \; \Sigma_j^{-1}(x-\mu_j) \quad (3)$$

Where Σ_j and μ_j are the covariance matrices and mean vector for class j and d is the number of dimensions.
It can be assumed that all classes share the same covariance matrices Σ. Σ_j can thus be replaced with the common covariance matrix Σ. Through taking the natural logarithm to remove constants, and through mathematical manipulation the function in equation (2) can be written as:

$$g_j^*(x) = \mu_j \Sigma^{-1} x' - \frac{1}{2}\mu_j \Sigma^{-1} \mu_j' - ln(P(\omega_j)) \quad (4)$$

Which can be written as the linear discriminant classifier:

$$g_j^*(x) = weight_j \cdot x' + bias_j \quad (5)$$

The likelihoods obtained from equation (5) can be used to calculate the confidence score of a sample belonging to a class j. The natural logarithmic operation used to derive $g_j^*(x)$ transformed the function to the log domain. To calculate the confidence scores the function must be transformed back to the linear domain. Additionally, the class j likelihood must be normalized regarding the sum of all class likelihoods, in order to be a value between 0 and 1, and results in the following calculation of confidence score:

$$CS_k(x) = \frac{e^{g_j^*(x)}}{\sum_{j=1}^{J} e^{g_j^*(x)}} \quad (6)$$

Where $CS_k(x)$ is the confidence score of a sample x belonging to class j. The normalization operation was included to represent the class confidence score as a percentage of the sum of all class confidence scores, in order to have $CS_k(x)$ presented as a more intuitive number for the user. The LDA classifier will be used in the control scheme. To obtain smoother control, the class with the highest average likelihood based on features from the previous three segments is chosen as output class.

III. METHODS

Subjects

In this study, 16 healthy able-bodied subjects were included (15 male and 1 female - 14 right handed and 2 left handed of mean age 25.3 $\pm$ 1.5). The subjects were recruited by contacting students at Aalborg University. Prior to the experiments the subjects received an experiment protocol, containing information on the objective of the study and steps of the experiment. To ensure full understanding and cooperation, the subjects were thoroughly instructed prior the initiation of each step during the experiment. All 16 subjects participated in the entirety of the experiment, from which no data was excluded. The subjects participated voluntarily and received no reimbursement.

Experimental Protocol

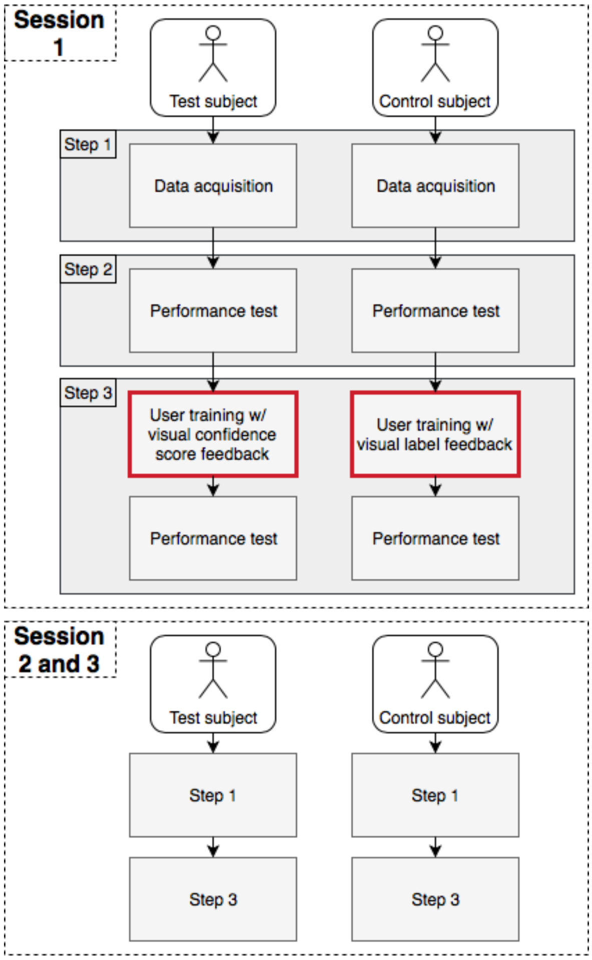


Fig. 1: Graphical illustration of the experiment showing the steps of each session for the test and control group. Highlighted is user training in step 3 which was the only step that varied between the two groups, and comprised the main area of research interest in the experiment.

Each subject underwent three sessions; one session per day over three consecutive days. The subjects were randomly allocated to either a test or control group; 8 subjects in each group. During each session EMG signals were initially acquired from the subjects and used to train the control system. The subjects then underwent user training with the purpose of learning how to adapt to the control system. Finally the subjects went through a real time performance test to evaluate their ability to operate a virtual prosthesis. In the first session, the subjects completed the performance test prior to user training. This test was used as a baseline assessment of the subject's performance. All implementations have been performed using MATLAB (2017b).
The difference between the test and control groups, and the main area of interest in the study, was in the feedback provided during user training. The test group received the estimated probabilities of each class (confidence scores), while the control group only received label feedback (the estimated class). A flowchart of the study design can be seen in figure 1.

Data Acquisition

EMG signals were recorded with the Myo armband (MYB) from Thalmic Labs - an eight channel dry stainless steel electrode armband. The MYB, which samples at 200 Hz, has a built in 50 Hz notch filter and a Bluetooth 4.0 unit which enables wireless communication with a computer. A 2^{nd} order Butterworth high-pass filter with a 10 Hz cut-off was digitally implemented to reduce movement artefacts. Due to the low sampling with no beforehand low-pass filtering, aliasing of the signal was inevitable, thus no anti-aliasing filter was implemented. Despite the low sampling rate, the MYB has shown to provide EMG recordings that can be classified with significantly similar accuracy as EMG recordings acquired with conventional EMG surface electrodes sampled at 1000 Hz [16].

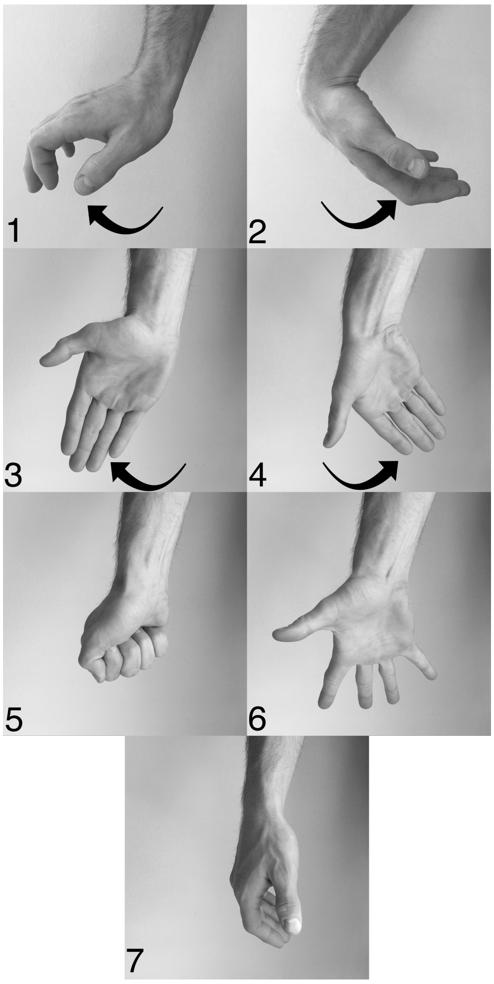


Fig. 2: Illustration of the movements performed in the experiment. 1: Wrist extension, 2: Wrist flexion, 3: Radial deviation, 4: Ulnar deviation, 5: Closed hand, 6: Opened hand, 7: Rest.

The subjects were instructed to elicit muscle contractions corresponding to the following classes of hand movements: *Wrist extension, Wrist flexion, Radial deviation, Ulnar deviation, Closed hand, Open hand and Rest*, which are illustrated in figure 2. The subjects had their dominant forearm disinfected, and were instructed in wearing the MYB at the thickest part of their forearm. To ensure the same placement of the MYB on each subject, the main electrode-channel was placed most laterally when standing in the anatomical standard position. The subjects were seated on a chair with the dominant arm hanging relaxed laterally down the torso during the whole experiment.

According to Scheme et al. [7], the use of dynamically changing contraction data in training a classification-based control scheme has shown to improve performance and tolerance to proportional control. Based on this finding, the subjects performed three repetitions of each movement, where each repetition constituted of a 2.5 second increasing ramp contraction, a 5 second steady state contraction at the peak of the increasing ramp contraction and a 2.5 second decreasing ramp contraction. To assure that each repetition was carried out correctly, the subjects were instructed in tracking a cursor, representing the EMG signal, on a trapezoidal trajectory, where the slopes corresponded to the ramp contractions and the plateau corresponded to the steady state contraction. The plateau of the trajectory differed between the three repetitions as 40 %, 50 % and 70 % of an initial recorded 15 second constant force of Maximum Voluntary Contraction (MVC). Of the recorded time only the plateau phase and the last second on the incline and first second of the decline were used to fit the classifier. To avoid muscle fatigue the subjects were given 30 seconds rest after a MVC recording and 10 seconds rest between repetitions.

Feature Extraction

Before training the classifier, features were extracted from the signal. The raw EMG signal from each MYB channel was segmented into 200 ms windows with a 50 % overlap respecting the findings of Farfán et al. [17]. Based on using the MYB for data acquisition recommendations made by Donovan et al. [18] regarding the optimal features for low bandwidth sEMG pattern recognition were taken into consideration. These features provided useful signal information even though the MYB only samples sEMG signals with 200 Hz, and in this case offering better accuracy than the Hudgins features [19] in a LDA based control scheme [18].

Four space domain (SD) features of Scaled Mean Absolute Value (SMAV), Correlation Coefficient (CC), Mean Absolute Difference Normalized (MADN), Scaled Mean Absolute Difference Raw (SMADR) were used for feature extraction. These features represent a portion of the features Donavan et al. [18] proposed, as the rest were left unused due to the intent of reducing feature redundancy. The calculation of SD features lean on the calculation and relation of other SD features. Special for the SD features is utilizing the relation between signals acquired in the different channels of the MYB. Additionally, the well known Hudgins time domain feature Waveform Length (WL) was included to cover complexity information in the time domain [20]. The calculation of the features can be found in the appendix.

Proportional Control

Classification provides the recognition of a movement as an output, but not the intensity of that movement. Therefore, to estimate the intensity, multivariate linear regression models were utilized. One regression model was trained for each movement class (six in total), where the independent variables were Mean Absolute Values (MAV) extracted from each segment in each channel of the MYB. The dependent variables were set as the signal generated when tracking the trapezoidal trajectory during the data acquisition. Thus, the proportional output value was a single value between 0 and 1. The calculation was as follows:

$$\hat{Y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_i + \varepsilon_i \qquad (7)$$

Where i is the number of MYB channels, $\hat{Y}$ is the proportional control output, X_i is the MAV feature of a segment in the i^{th} channel, α is the regression intercept, β is the regression slope and ε_i is the error term. Similarly as the classification control, the proportional control output was calculated as the average output from the three previous segments to obtain smooth control. This control scheme was used in both the user training and the performance test.

User Training

Subjects were set to train their understanding of making distinguishable hand movements, using a user training interface, where feedback corresponding to their assigned group was presented. Prior to training, subjects were informed of the importance of their efforts in relation to the experiment with the intent of encouraging a focused participation.
The user training interface contained the following feedback: an illustration of the movement needed to be performed, a horizontal bar visualizing the contraction level and a vertical bar plot visualizing which movement was being recognized by the control system. The only difference between the test and control group was the feedback received in the vertical bar plot. The test group received confidence score feedback (multiple bars) as seen in figure 3 and the control group received label feedback (single bar) as seen in figure 4.

The test group was shown the classifier confidence scores for multiple classes, which enabled the possibility of having multiple vertical plots shown. Thus, more diverse feedback was presented, which the user could utilize to correct the confidence of the performed movement. The control group had only the movement with the highest confidence shown, thereby limiting the confidence feedback to only one bar visible at a time. Thus, the control group was not informed on the exact probabilities of which movements the control system recognized.

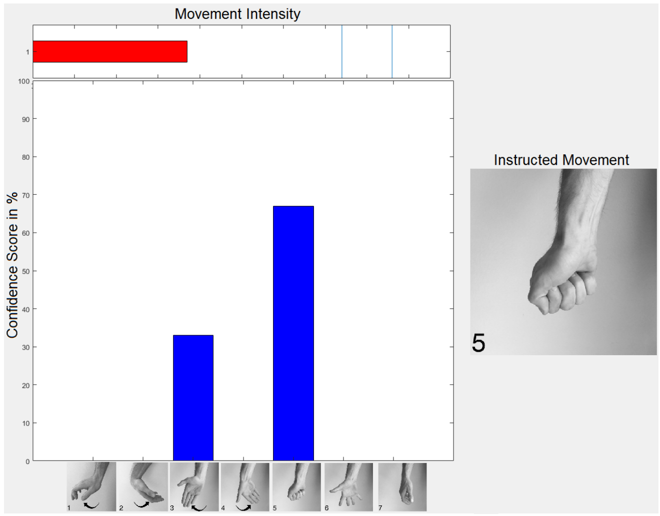


Fig. 3: Illustration of the user training interface for the test group. The vertical bar plot indicates the recognized movements, visualized by the images of each movement; a full bar corresponds to 100 % recognition confidence. The horizontal bar plot indicates the contraction level, where a full bar corresponds to the MVC. The two vertical lines in the contraction level bar plot illustrates the contraction level interval the subject must reach. The large picture of a movement on the right of the bar plot indicates which movement needs to be performed. The test group received confidence score feedback by having the possibility of multiple bars being shown.

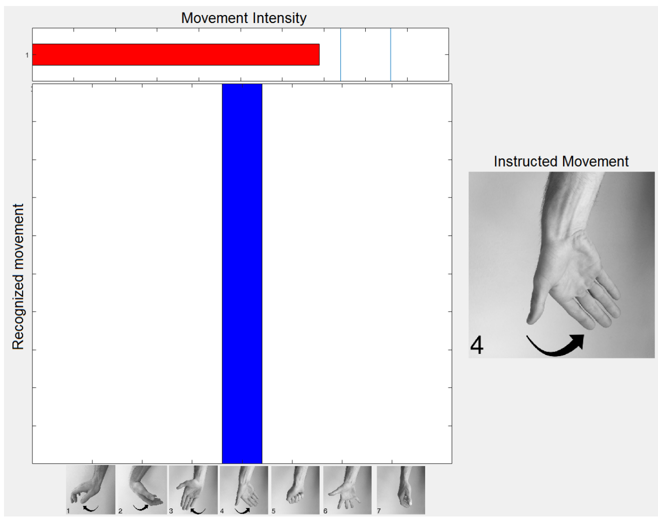


Fig. 4: Illustration of the user training interface for the control group. The same interface used for the test group was presented to the control group, except that the feedback consisted of label feedback instead. Thereby, the control group was only presented with the most certain recognized movement, shown with a single bar.

The intent of user training was to train the subject in being more aware of how to perform a movement in a way the classifier would recognize as the movement the user actually performed. Basically, the training should motivate the subjects to modulate their contractions to maximize the classifier confidence. The subjects should perform contractions such that the probability bar for the target class was maximized while the probability bars for the other classes were minimized. To motivate the subject during user training a simple task was implemented in the interface.

The subject had to perform the instructed movement and achieve a minimum of 75 % confidence for the test group and the correct class for the control group, whilst also managing to perform the movement within the contraction level interval indicated by the vertical boundaries in the horizontal bar plot. Once these requirements were met and withheld for one second, a sound was played indicating task completion. The subjects had to return to the rest class and then repeat the movement. A task completion was referred to as a repetition and was saved as a user training outcome measure. The goal was to manage as many repetitions as possible within 30 seconds, then a 10 second break was issued before moving to the next movement.

The sequence of a training session were put together in form of the subject having to perform each of the six movements in combination with four different contraction level intervals; 75-85 %, 55-65 %, 35-45 % and 15-25 % of their MVC. The instructed movements were trained in a random order and the subjects needed to perform all movements in the same contraction level interval before moving to a new interval. This resulted in a total training session time of 16 minutes.

Performance Test

A performance test was developed to evaluate the users' ability to operate a virtual prosthesis. The test was implemented as a 3D Fitts' Law target reaching test, similar to methods reported in [14, 15]. The user controlled a circular cursor in a Cartesian coordinate system, where the cursor was to be matched with the appearing targets. Extension/flexion of the wrist moved the cursor horizontally, radial/ulnar deviation moved the cursor vertically and opened/closed hand increased/decreased the size of the cursor. The cursor moved proportional to contraction intensity with a velocity between 0 and 1, where 1 corresponded to the MVC. An illustration of the Fitts' Law test interface can be see in figure 5.

To reach a target the user had to match the size and position of the target with the movable cursor, and dwell within the area for 1 second. The target would appear for 15 seconds or until it was reached, after which a new target would appear and the cursor position would be reset to origin. A total of 16 targets would appear before the test ended. The sequence of targets appearing was different between all four test session, to avoid bias of subjects remembering the sequence in which targets would appear.

Originally the Fitts' Law test had a single performance measure, *throughput* (TP) [21]. TP uses the relationship between time taken to reach a certain target in seconds (MT) and the index of difficulty (ID), and is defined as:

$$TP = \frac{1}{N}\sum_{i=1}^{N}\frac{ID_i}{MT_i} \tag{8}$$

Where i is a specific movement and N is the total number of movements. ID relates to the target's width W and distance D from origin, where W and D are unitless. The ID is calculated as:

$$ID = log_2(\frac{D}{W}+1) \tag{9}$$

According to [15], it is in practice most resourceful to use a variety of ID's in a Fitts' Law test. Based on this assumption, the target ID's seen in table 1 were calculated for this study.

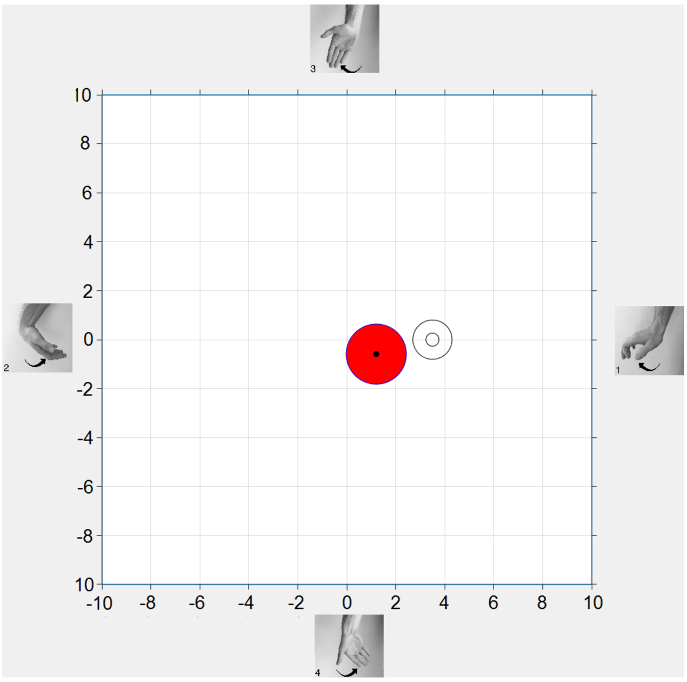


Fig. 5: The implemented interface for the modified Fitts' Law test. The user controlled the red cursor with the centred bold mark. The target consisted of a circle with a larger circle surrounding it. The user was instructed in matching the cursor with the target, where the bold mark should be positioned inside the inner circle of the target, and the outer circle of the cursor should be matched in size with the outer circle of the target. The cursor would then turn green to indicate the matching was correct, and blue when the dwell time was reached.

Further performance measures were included similar to previously reported in [14, 15]. These measures consists of *Path Efficiency*, *Overshoot*, *Stopping Distance* and *Completion Rate*.

The additional four measures were added to quantitatively assess performance of naturalness, spontaneity, and compensatory motions during control. The calculation of these features can be found in the appendix.

Tab. 1: The index of difficulty used in the Fitts' Law test.

Distance	Width	ID
28.0	0.33	6.41
24.5	0.33	6.22
22.0	0.33	6.01
18.5	0.33	5.82
16.0	0.33	5.61
13.0	0.33	5.32
12.5	0.33	5.27
9.5	0.33	4.88

Cluster Dispersion and Separability

The EMG signal for each movement class acquired from the subjects forms clusters of multidimensional data points. The lower the dispersion of the individual movement class clusters is, the more distinguishable the movements are, and the classifier will recognize the movement classes with higher accuracy. Additionally, a higher distance between cluster centroids will facilitate a higher classification accuracy as well. The dispersion of the class clusters and the distance between cluster centroids will be calculated as an outcome measure of the data used to train the classifier.
To calculate cluster dispersion, the centroid of multidimensional clusters must be calculated as in:

$$C = \frac{\sum_{n=1}^{N} a_n, b_n, \ ... \ k_n}{N} \tag{10}$$

Where C is the centroid, n is the current data point in a dimension, N is the total number of data points in a dimension and k is the number of dimensions. To calculate cluster dispersion, the Euclidean distance (ED) from data point p to the corresponding cluster q can be computed:

$$ED(p,q) = \sqrt{(p_1-q_1)^2+(p_2-q_2)^2 \ + \ ... \ +(p_k-q_k)^2} \tag{11}$$

Where p_k and q_k are the coordinates of vectors p and q respectively. This procedure is performed for all data points in a cluster, from which the average is calculated to obtain a general impression of the cluster dispersion.
To calculate the cluster separability, the ED between cluster centroids is calculated as in equation (11).

Statistical Analysis

Statistics were applied to evaluate improvements in the results obtained in the performance test, user training and data clustering. A Friedmans test was used for multiple comparison and a Tukey-Kramer correction was conducted when detecting an effect. For comparison between groups in each session, a Mann-Whitney U test was applied.

IV. RESULTS

Performance Evaluation

This section presents the results acquired from the Fitts' Law target reaching test. The test had five measures which each expresses a parameter of subjects' performance. The plotted mean and standard deviations of each measure for all subjects in each group in the performance test for each session can be seen in figure 6.
The baseline performance test showed no difference between the two groups, showing the two groups to be homogeneous at initiation. The Fitts' Law test results did not show any significant improvement over the three sessions for any of the five test measures for both the test and control group ($p > 0.05$). Similarly, there was no significant difference between the two groups' performance in any session ($p > 0.05$), meaning neither of them performed significantly better than the other group in any of the sessions.

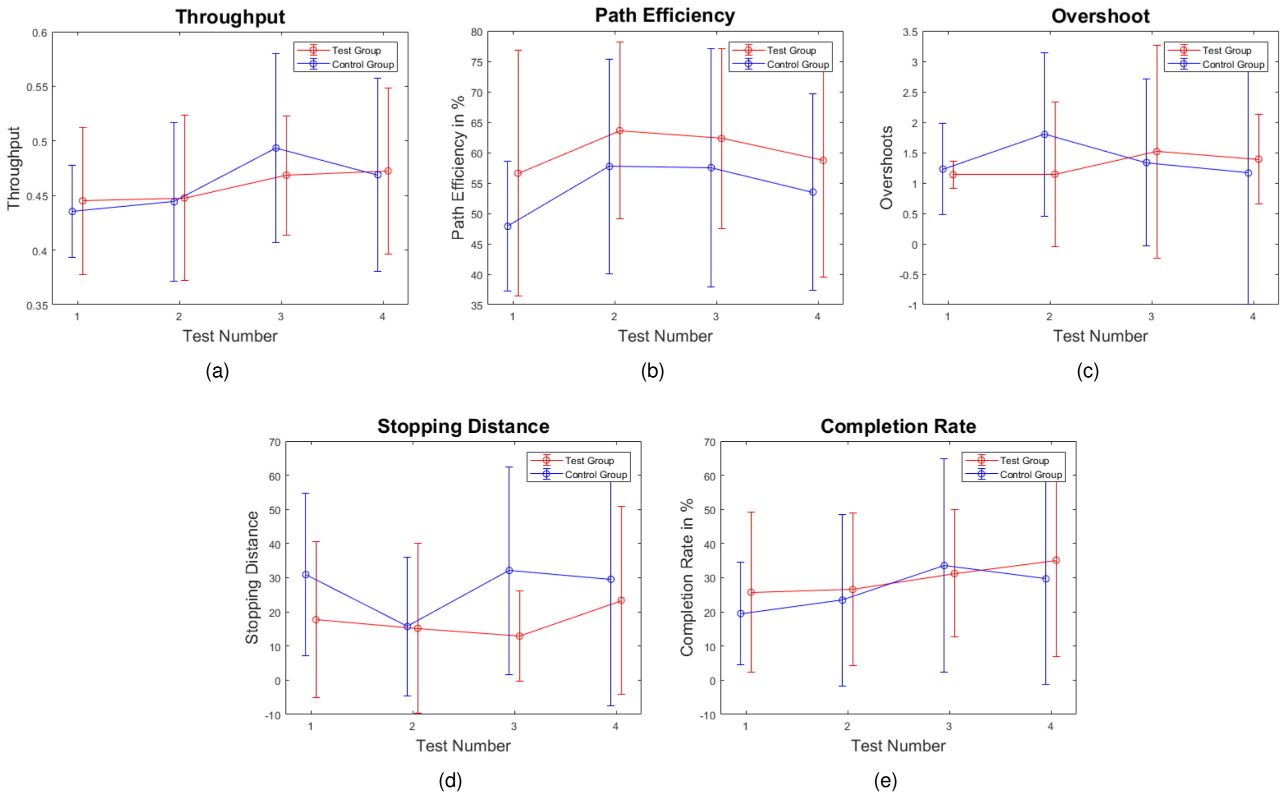


Fig. 6: Figure illustrating the five performance measures; *a) Throughput, b) Path efficiency, c) Overshoot, d) Stopping distance, e) Completion rate*, used for quantifying user performance across all four tests. Test number 1 is the acquired baseline used for assessing group homogeneity and the following numbers indicate performance test results after user training in each session. The red line indicates the progression of the test group and the blue line the progression of the control group.

User Training Evaluation

This section covers the outcome measure results obtained during user training sessions. During user training subjects were instructed to train movements in being performed, such that the control system recognized the movement as the actually performed movement.

No significant difference in the total number of repetitions was found between sessions of either group ($p > 0.05$). When comparing the total number of repetitions of each session between groups accordingly, no significant difference were found either ($p > 0.05$).

An increased ability to get repetitions in the low intensities was found for the control group ($p < 0.05$, session 1 $= 16.13 \pm 5.59$, session 3 $= 21.38 \pm 6.78$). Otherwise, no significant difference were yielded for both groups when comparing the subjects' ability to reach the three other contraction levels between sessions ($p > 0.05$). No difference was found, when comparing the two groups' ability to reach different intensities during training either ($p > 0.05$).

Comparing the ability to perform different movements during the training showed a significant improvement for the test group in ulnar deviation ($p < 0.05$, session 1 $= 11.38 \pm 4.27$, session 3 $= 16.13 \pm 2.95$) and open hand ($p < 0.05$, session 1 $= 11.25 \pm 3.85$, session 3 $= 17.88 \pm 2.46$). A significant decrease in performance was found for the control group's ability to perform flexion ($p < 0.05$, session 2 $= 16.63 \pm 2.77$, session 3 $= 11.00 \pm 3.16$). Otherwise, no significant difference between the three sessions for the two groups was found ($p > 0.05$).

A significant difference ($p < 0.05$) was found between the test and control group's ability to reach the closed hand movement, with a mean of 26.8 ± 13.5 number of repetitions for the test group and 38 ± 12.2 for the control group. No significant difference was found for any of the other movements when comparing the two groups ($p > 0.05$).

Cluster Dispersion and Separability Results

In this section, results from the data acquisition are presented. The data used for training the LDA based classifier was examined. Each movement resulted in a cluster of data points, which was examined in order to analyse the change in cluster dispersion and distance between cluster centroids.

For both groups the mean distance between the cluster centroids were calculated. The change in between cluster distances over the three sessions were tested and showed no significant difference ($p > 0.05$). Likewise, no significant difference in the development of cluster distances between the groups was found ($p > 0.05$).

The mean distance from data points to the cluster centroid was calculated. This showed no significant difference for the test group ($p > 0.05$), but a significant difference was found for the control group ($p < 0.05$). The Tukey-Kramer correction showed the significant difference was between session one and three ($p < 0.05$), where the mean for session one was 502.02 ± 274.88 arb. unit, and session three was 323.43 ± 171.13 arb. unit. The comparison between groups showed that the control group achieved a significant improvement in cluster dispersion compared to the test group in session three ($p < 0.05$), where the test group had a mean cluster dispersion of 584.34 ± 250.02 arb. unit, while the control group had 323.43 ± 171.13 arb. unit.

V. DISCUSSION

The objective of the study was to investigate if exposing subjects to user training, in which confidence scores of movement class recognition was used as visual feedback, would show statistically significant improvement in a Fitts' Law test, when compared to subjects who received label feedback.
The results showed no significant difference between the test and control group within the Fitts' Law test, in all comparisons between and within groups ($p > 0.05$). This meant that no group performed better compared to the other, and that neither of the groups managed to improve significantly during the three sessions of training and testing. The only significant difference ($p < 0.05$) between the groups were found in the user training when performing the closed hand motion, where the test group performed worse than the control group. This difference could be the result of the training type, the number of subjects or a faster learning ability within the control group. The control group improved in number of repetitions in lower intensities during user training between session 1 and 3. It was expected that this improvement would be found within the test group, as the low intensity motions where the motions from which the classifier would get most confused, and the confidence score then would provide the test subjects with insight on how to correct the movement best possible. A reason for the contradictory result might be that the subject were confused by the confidence scores and found the information excessive.
A main cause of the lacking improvement within the groups could be the result of higher ID's ($4.88 - 6.41$) compared to other studies ($1.59 - 3.46$) [14, 15]. Several subjects struggled in reaching any targets, and if the subject was unable to reach any targets, all the Fitts' Law measures except CR were unusable in statistics. This resulted in a weaker statistical test, as fewer observations were included in the comparison of the remaining measures. In addition, this lead to problems when examining the results, as it was expected that the statistical differences would primarily be found when looking at other measures than CR, as they would offer better insight into the improvement of precision when completing the test.
At the same time a high ID led to subjects becoming frustrated when they had troubles reaching targets. When observing the test it was clear that this frustration resulted in the subjects forgetting how to perform precise movements, which then led to further frustration. This factor could also have had an effect on the subjects' performance. Significant improvement in development of movement precision might also take more than three sessions, and this could also be a cause of the lacking development of the subjects. In developing the understanding of precision, there should also be a higher focus on rest, as this is a crucial part of the performance test. Some of the subjects did not understand the importance of returning to rest after a performed movement during user training, which might have been reflected in the performance test.
The above points should be taken into consideration in future studies when examining the use of confidence scores as visual feedback in user training to improve performance.

While examining the EMG data it was found that the cluster dispersion improved within the control group ($p < 0.05$) between the sessions. When applying a Tukey-Kramer correction it was found that the difference was between the first and third session ($p < 0.05$). This result shows that the control group became better at performing precise movements, as the EMG data was more closely clustered after training for the three sessions.
Furthermore, a significant difference ($p < 0.05$) was found when comparing the within cluster distance of the two groups of the third session, where the mean distance for the control group (323.43 ± 171.13 arb. unit) was close to half of the distance within the test group (584.34 ± 250.02 arb. unit). This lead to the assessment that the control group became better at performing the exact movements during data acquisition when compared to the test group.

Optimization of Study

When implementing the performance test interface, the ID's and minimum number of DOF's used to reach a target, should be lowered in order for the subjects to reach a CR of 80% to 100%, as reported in previous studies [14, 15]. This might yield a more clear indication of precision of the control, which is shown better in the remaining Fitts' Law measures. At the same time a lower ID would give the subjects a feeling of success rather than frustration when performing the test, which might encourage them to retain the interest and focus when carrying out the performance test. A problem observed during the Fitts' Law test was that subjects were affected by the cursor being reset to origin. The subjects' current movement were carried over during the transitioning between targets. A suggestion to future studies is to include a transition break of 1 second when a new target appears.
During user training the subjects should be forced to get back to rest, in order to train the ability to dwell within a target in the performance test. This requirement was not implemented in the current training interface, but the importance of learning to

rest when using classifiers should be examined in future studies.
In future testing, the number of sessions should be more than three. This was also found in [12], who similarly did not achieve significant improvement in performance following a short time user training intervention, whereas Powell et al. [9] found a steady improvement during a 9 session user training study. In relation, it would be beneficial to examine the time it takes to improve performance in order to find the minimum number of sessions necessary to achieve higher precision when performing specific hand movements.
At last a larger number of subjects could result in a better distribution within the groups, as some subjects were able to get close to 100 % CR in the first or second session, while others struggled with reaching just one target during each session.

VI. CONCLUSION

Based on the results in the experiment it was found that training the user with confidence score feedback compared to label feedback can not be linked to any significant improvement in performance evaluated through a Fitts' Law test. Furthermore, no significant improvement during a three day training period for either the control or the test group was detected. These findings are most likely due to the high index of difficulty, making it hard to draw any conclusions based on the Fitts' Law test.

Contrarily, it appears that training the user with label feedback can lead to a closer clustering of EMG data compared to training with confidence score feedback. This can be assessed, as a significant improvement was found between the first and last dataset recorded for the control group. To further support this the EMG signal of the subjects who received label feedback clustered significantly closer than the test group on the last day of testing. This shows that training based on confidence scores might not be a way to improve performance, which should be examined further by the use of Fitts' Law tests with lower ID's and a higher number of training sessions.

ACKNOWLEDGMENT

The authors would like to thank supervisors Strahinja Dosen, Jakob Lund Dideriksen and Lotte N.S. Andreasen Struijk for providing constructive feedback, and the School of Medicine and Health at Aalborg University for providing equipment and the facilities to complete this study. Additionally, the authors are very thankful for all the voluntary participants.

References

[1] Bart Peerdeman et al. "Myoelectric forearm prostheses: State of the art from a user-centered perspective". In: *The Journal of Rehabilitation Research and Development* 48.6 (2011), p. 719.

[2] Elaine Biddiss and Tom Chau. "Upper limb prosthesis use and abandonment: A survey of the last 25 years". In: *Prosthetics and Orthotics International* 31.3 (2007), pp. 236–257.

[3] Joseph T. Belter et al. "Mechanical design and performance specifications of anthropomorphic prosthetic hands: A review". In: *The Journal of Rehabilitation Research and Development* 50.5 (2013), p. 599.

[4] Han Jeong Hwang, Janne Mathias Hahne, and Klaus Robert Müller. "Real-time robustness evaluation of regression based myoelectric control against arm position change and donning/doffing". In: *PLoS ONE* 12.11 (2017), pp. 1–22.

[5] Anders Fougner et al. "Control of upper limb prostheses: Terminology and proportional myoelectric controla review". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20.5 (2012), pp. 663–677.

[6] COAPT. *Coapt engineering.* 2018.

[7] E Scheme and K Englehart. "A comparison of classification based confidence metrics for use in the design of myoelectric control systems". In: *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2015), pp. 7278–7283.

[8] J. M. Hahne et al. "Linear and Nonlinear Regression Techniques for Simultaneous and Proportional Myoelectric Control". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.2 (2014), pp. 269–279.

[9] Michael A. Powell, Rahul R. Kaliki, and Nitish V. Thakor. "User training for pattern recognition-based myoelectric prostheses: Improving phantom limb movement consistency and distinguishability". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.3 (2014), pp. 522–532.

[10] Ning Jiang et al. "Myoelectric Control of Artificial Limbs: Is There a Need to Change Focus?" In: *IEEE Signal Processing Magazine* 29.5 (2012), pp. 150–152.

[11] Yinfeng Fang et al. "Classifier-Feedback-Based User Training". In: *IEEE Transactions on Biomedical Engineering* 64.11 (2017), pp. 2575–2583.

[12] L. Pan et al. "Transcranial direct current stimulation versus user training on improving online myoelectric control for amputees". In: *Journal of Neural Engineering* 14.4 (2017), pp. 1–15.

[13] Ann M. Simon, Blair A. Lock, and Kathy A. Stubblefield. "Patient training for functional use of pattern recognition controlled prostheses". In: *Journal of Prosthetics and Orthotics* 24.2 (2013), pp. 56–64.

[14] Erik J. Scheme, Bernard S. Hudgins, and Kevin B. Englehart. "Confidence-based rejection for improved pattern recognition myoelectric control". In: *IEEE Transactions on Biomedical Engineering* 60.6 (2013), pp. 1563–1570.

[15] E Scheme and K Englehart. "Validation of a selective ensemble-based classification scheme for myoelectric control using a three dimensional Fitts' law test". In: *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 21.4 (2013), pp. 616–623.

[16] I Mendez et al. "Evaluation of the Myo Armband for the Classification of hand motions". In: *International Conference on Rehabilitation Robotics* (2017), pp. 1211–1214.

[17] Fernando D Farfán, Julio C Politti, and Carmelo J Felice. "Evaluation of EMG processing techniques using Information Theory". In: *BioMedical Engineering OnLine* 9.1 (2010), p. 72.

[18] Ian M. Donovan et al. "Simple space-domain features for low-resolution sEMG pattern recognition". In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2017), pp. 62–65.

[19] B. Hudgins, P. Parker, and R.N. Scott. "A new strategy for multifunction myoelectri control". In: *IEEE Transactions on Biomedical Engineering* 40.1 (1993), pp. 82–94.

[20] Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul. "Feature reduction and selection for EMG signal classification". In: *Expert Systems with Applications* 39 (2012), pp. 7420–7431.

[21] Paul M. Fitts. "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movements". In: *Journal of Experimental Psychology* 47.3 (1954), pp. 381–391.

[22] Richie Poulton et al. "Evaluation of Head Orientation and Neck Muscle EMG Signals as Command Inputs to a Human-Computer Interface for Individuals with High Tetreplegia". In: *IEEE Trans Neural Syst Rehabil Eng.* 360.9346 (2013), pp. 1640–1645.

[23] Ann M. Simon et al. "The Target Achievement Control Test: Evaluating real-time myoelectric pattern recognition control of a multifunctional upper-limb prosthesis". In: *Journal of Rehabilitation Research and Development* (2011), pp. 619–627.

APPENDIX

Features

In this section the equations used for calculating the features used in this project.
MAV is a feature that primarily is affected by the force produces when making a contraction. MAV is extracted for each window and calculated for each of the i^{th} channel. The extraction is expressed as:

$$MAV_i = \frac{\sum_{n=1}^{ws} |x_i[n]|}{ws} \tag{12}$$

where ws is the window size, the number of raw data points in that exact window. $x_i[n]$ is the n^{th} raw data points from the i^{th} channel.

The mean MAV across all channels, MMAV, is used to remove dependency of movement intensity. MMAV is calculated by using the MAV of all channels for the current window, and is done as following:

$$MMAV = \frac{\sum_{i=1}^{8} MAV_i}{8} \tag{13}$$

MMAV can be used to scale the MAV feature creating the SMAV feature. This feature should represent a non-dimensional relationship between channels. SMAV is simply calculated as:

$$SMAV_i = \frac{MAV_i}{MMAV} \tag{14}$$

As each of the eight EMG sensors in the MYB are located around the arm, they acquire signals from a mixture of sources. Also individual sources may affect multiple sensors depending on their size. Due to this a source measured by multiple sensors will effect their acquired signal correlation. An idea is therefore to calculate the correlation coefficient between each channel and its neighbouring channel.

$$CC_i = \frac{\sum_{n=1}^{ws} X_i[n]X_{i+1}[n]}{ws} \tag{15}$$

$X_i[n]$ is the n^{th} normalized data point from channel i. When calculating CC the data from each window is normalized by subtracting its mean value from each raw data point, and afterwards divided by their standard deviation.

Calculating CC can prove rather demanding in computational power due to the series og multiplication operations. Therefore Donovan et al. [18] proposed introducing a mean absolute difference-based feature of lower computational complexity which still characterizes the spatial relationship between channels. The MAD feature is normalized in the same way as CC, making up the MADN feature calculated as:

$$MADN_i = \frac{\sum_{n=1}^{ws} |X_i[n] - X_{i+1}[n]|}{ws} \tag{16}$$

If the normalization of the signal proves too demanding the feature can be calculated on the raw EMG signal without the normalization. This makes up the MADR feature, calculated as:

$$MADR_i = \frac{\sum_{n=1}^{ws} |x_i[n] - x_{i+1}[n]|}{ws} \tag{17}$$

As the SMAV feature the MAD feature can be scaled by MMAV to remove movement intensity dependency. SMADR is calculated for each channel by:

$$SMADR_i = \frac{MADR_i}{MMAV} \tag{18}$$

As stated in the beginning some of these features introduce redundancy, subsequently the features of SMAV, CC, MADN and SMADR are the ones used for classification. [18]

To further improve the decision foundation of the classifier it was proposed to include the time domain feature of WL calculated by:

$$WL_i = \sum_{n=1}^{N-1} |x_{i+1}[n] - x_i[n]| \tag{19}$$

WL is a measure of the signal complexity by calculating the cumulative length for each channel [20].

Fitts' Law Measures

In this section the equations for the Fitts' Law measures are presented.

Throughput (TP) which represents the trade-off between speed and accuracy. TP uses the relationship of time taken to reach a certain target in seconds (MT) and the index of difficulty (ID). This forms: [14, 21]

$$TP = \frac{1}{N}\sum_{i=1}^{N}\frac{ID_i}{MT_i} \tag{20}$$

where i is a specific movement and N is the total number of movements. ID relates to the target distance D and width W. The ID for each task, from the origin to a specific target of a certain size is calculated using [14, 21]:

$$ID = log_2(\frac{D}{W}+1) \tag{21}$$

Path Efficiency (PE) describes the quality of control by making a measure of the straightness of the cursor's path to the target, by making a ratio of the actual path distance versus the optimal path distance. This tests the users ability to continuously control the cursor position. Following the optimal path will result in a PE of 100%. PE is calculated as follows [14, 22]:

$$PE = \frac{Optimal\ Distance}{Actual\ Distance} \tag{22}$$

Overshoot (OS) is the number of times the cursor enters and then leaves the target before the dwell time inside the target is reached, across all target in the task, divided by the total number of targets. OS tests the users ability to control the velocity of the cursor accurately. A perfect OS-score of zero is reached if the cursor dwells within the target boundaries on the first try for all targets, and is calculated as the following [14, 22]:

$$OS = \frac{Total\ Number\ of\ Overshoots}{Total\ Number\ of\ Targets} \tag{23}$$

Stopping Distance (SD) describes the users ability to rest and thereby perform no movement. The SD measure is the distance moved during the dwell time across all targets, and is given as [14]:

$$SD = \sum_{i=1}^{N}(Distance\ Inside\ Target)_i \tag{24}$$

where i is a reached target and N is the total number of reached targets.

Completion Rate (CR) describes the percentage of targets reached within the total allowed time. This gives a general idea of the user's performance, and is calculated as [14, 23]:

$$CR = \frac{Number\ of\ Reached\ Targets}{Total\ Number\ of\ Targets} \tag{25}$$

WORKSHEETS

Contents

1 | Background

The background chapter serves the purpose of providing a theoretical overview of the techniques applied to deal with the novel proposal of improving users' ability to operate a myoelectric prosthesis by training the user with visual confidence score feedback. The chapter will cover the usually applied methods of pattern recognition based myoelectric prosthesis control. The idea behind myoelectric prosthetic control is to convert muscles signals (EMG signals), recorded from a user when performing a muscle contraction, into a movement performed by a prosthesis. The EMG signals are used to train a control system in recognizing a pattern in the EMG signals from different muscle contractions. The control system then decides which movement the prosthesis should perform, based on which pattern in the EMG signals that is recognized. The focus of this project is to investigate users' ability to adapt to how the control system wants the user to perform a muscle contraction for the desired movement to be performed by the prosthesis. The different processes of myoelectric prosthetic control, which the background chapter will cover, are: the mechanics of the movements the control system will be trained to recognize, the generation of EMG signals, data acquisition, data processing, feature extraction, classification and control output. The pipeline of this process can be seen in figure 1.1. Furthermore, the chapter will cover theory on how confidence scores are calculated, how user training has been used in previous studies and how real-time prosthesis control is evaluated.

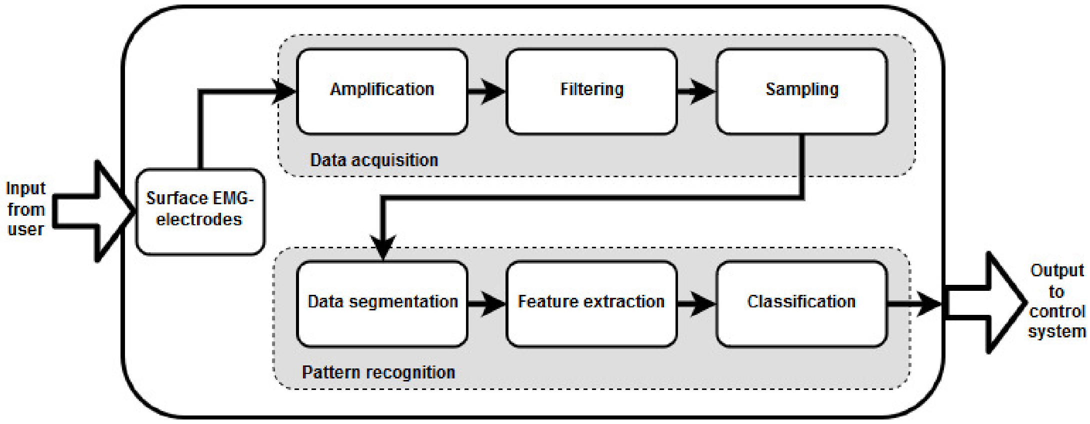


Figure 1.1: The figure shows the pipeline for pattern recognition-based myoelectric prosthetic control. The EMG signal from the user is first detected by the surface electrodes, after which it is amplified and filtered before it is sampled to process it digitally. To produce a control output the signal is subsequently segmented in windows from which features are extracted that are used to classify which movement has been made, and thus which movement should be performed by the prosthesis. The pipeline is adapted from Peerdeman et al. [1].

1.1 Choice of Movements

This project will use six movements based on three degrees of freedom (DOF) for control of a virtual interface. The movements are extension, flexion, radial and ulnar deviation, closed and opened hand as well as rest. The movements are shown in figure 1.2. These movements are distinguishable from each other and therefore very useful in classification. Training users to improve performance of these movements for use in a myoelectric prosthesis control scheme, could give a good foundation to build a classification scheme upon.

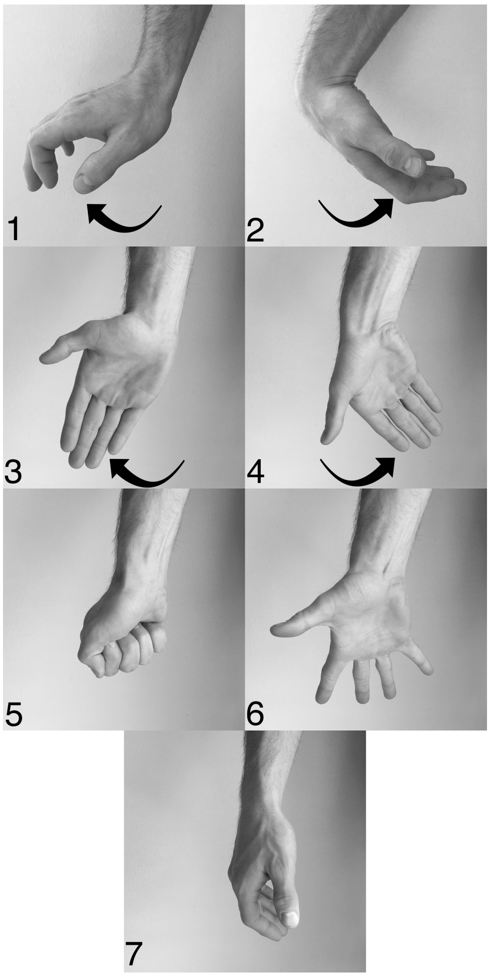


Figure 1.2: The figure shows the six hand movements used in this study as well as rest. The movements are: 1) extension, 2) flexion, 3) radial deviation, 4) ulnar deviation, 5) closed hand, 6) opened hand and 7) rest.

To perform movements of the wrist anf fingers, many muscles in the forearm are activated. All movements relevant for this study are controlled by muscles in the forearm, as well as some in the palm of the hand. However, as EMG recordings will be done from the forearm by equipment described in section 1.3.1, it is relevant to gain knowledge on the muscles in the forearm. When performing actions at the wrist the following muscles are active: flexor carpi radialis, flexor carpi ulnaris, palmaris longus, extensor carpi radialis longus, extensor carpi radialis brevis, extensor carpi ulnaris. The flexor and extensor muscles are naturally responsible for performing flexion and extension respectively at the wrist. They are however also responsible for performing radial and ulnar deviation, where the flexor carpi radialis and extensor carpi radialis brevis muscles, which are antoganistic muscles when doing flexion and extension, will work together when performing radial deviation. The extensor carpi radialis longus muscle is also responsible for performing radial deviation. The flexor and extensor carpi ulnaris muscles are responsible when performing ulnar deviation. The palmaris longus muscle is only active during flexion at the wrist. [2]

Several more muscles are further specified to perform movements of the fingers but many of these are also active during extension/flexion and radial/ulnar deviations at the wrist. Muscles responsible when opening the hand by extending the fingers are: extensor digitorum, extensor pollicis brevis, extensor pollicis longus, extensor indicis and the extensor digiti minimi muscle. Contrary, the muscles responsible for closing the hand by flexing the fingers are: flexor digitorum superficialis, flexor digitorum profundus and the flextor pollicis longus. [2]

1.2 Electromyography

This project will utilize the method of electromyography to record the muscle activation of the lower arm muscles in relation to the movements presented in section 1.1. To develop theoretical background knowledge, a short introduction of the essentials of the signal will be presented.

Electromyography is the recording of muscle activation. The amount of activity is found by measuring the electric potential, an action potential triggering a muscle contraction. The process of planning and executing a voluntary movement starts at the motor cortex in the brain, where a nerve impulse is send and travels through the spinal cord to the lower motor neuron. As seen in figure 1.3 the path from alpha motor neuron through the axon to the motor endplates is what makes up a motor unit. The alpha motor neuron originates from the spinal cord along the axon to the muscle it controls. The axon branches out to multiple muscle fibers through motor endplates innervating the muscle fibers.

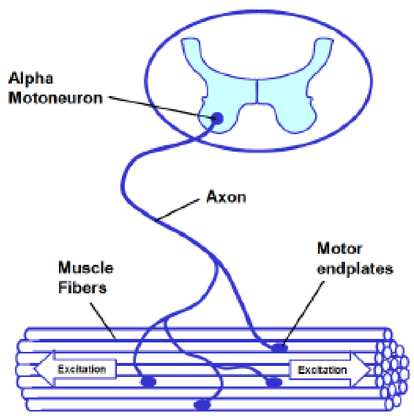


Figure 1.3: The figure illustrates the neural pathway from the alpha motor neuron to the innervated muscle fibers, making up a motor unit. [3]

A fundamental part in the application of EMG is the understanding of the excitation of muscles. Muscles contract through a series of steps of changing potentials across muscle cell membranes and rapid polarization and repolarization. However, when recording EMG it is the superposition of the spread of the motor unit action potentials (MUAPs) over the muscle membrane that is recorded. [4] Muscles are innervated by a varying number of nerves depending on the individual muscle. The MUAP is conducted to the muscle by nerves from the spinal cord, with the nerve impulses originating from the motor cortex in the brain. Muscles are not activated randomly by individual nerve fibers, but by nerves sorted into motor units. Many motor units are attached to a muscle and consist of a number of the nerves innervating the muscle. An illustration of how one motor unit attach to the muscle fibers of a muscle is illustrated in figure 1.3. When a motor unit activate, all the nerves in the motor unit is activated. This enables a controlled activation of the muscle as well as a activation that reach a higher number of muscle fibers. Motor units are also activated in an asynchronous pattern which enables different muscle fibers to be active at different times, making muscles less prone to fatigue. The force of a muscle contraction can be modulated either by motor unit recruitment or be frequency of activation. In EMG it is the sum of activity of activate motor units that is recorded. [4]

In the scenario of this project multiple EMG electrodes will record signals from many muscles in the lower forearm. This will result in some muscles being active during some movements as they contract, while other muscles will be inactive. On an EMG recording this will be visible as contracting muscles will show high activity, while others will show little to no activity. An illustration of this can be seen on figure 1.4. The figure shows the muscle activity of muscles in the forearm when performing first extension at the wrist followed by flexion. As can be seen the muscle activity is very different between the two muscles. This enables the recognition of specific movements based on several EMG recordings with several different electrode placements.

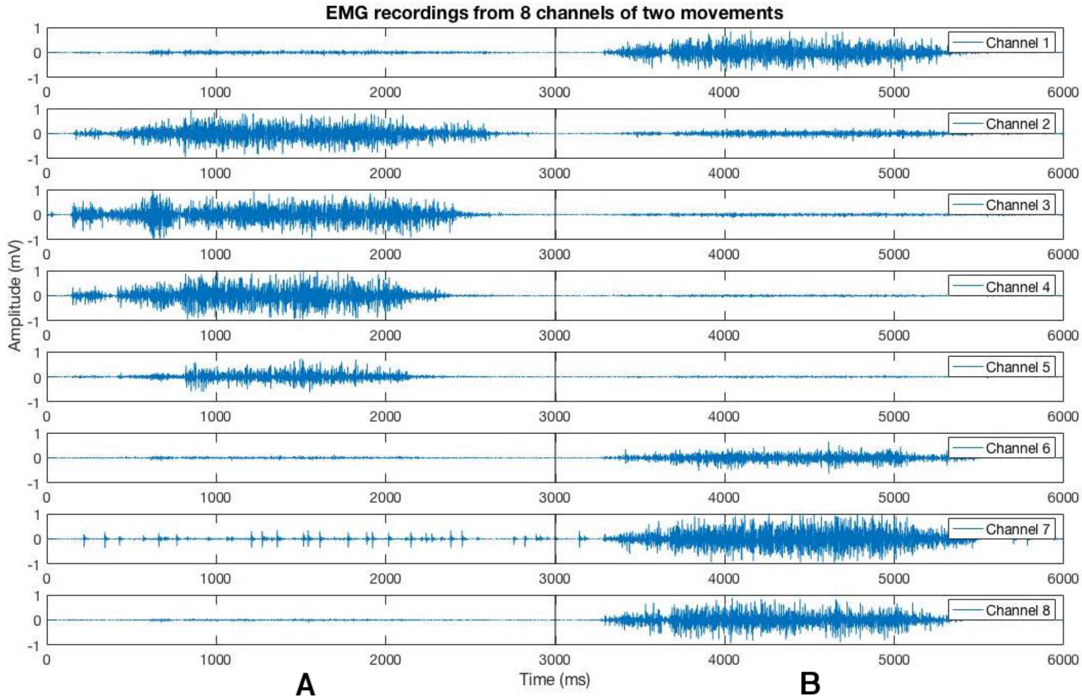


Figure 1.4: Illustration of the activity in an EMG recording of two movements: extension and flexion. Left side A) shows the activity recorded by EMG electrode channels during extension of the hand. Right side B) shows activity during flexion of the hand.

Recording EMG can be done either through the most often used surface EMG (sEMG) or by intramuscular EMG (iEMG). In iEMG a needle is inserted into the muscle measuring the MUAP directly on site. The more often used sEMG that uses electrodes to measure the sum of MUAPs on the skin surface, will be used to acquire EMG signals in this project. [4]

1.3 Data Acquisition

For a myoelectric prosthetic control system to be able to recognize hand movements it needs to be given prior information on how the movements looks like represented as a EMG signal - this is also called training the control system. Thus, EMG data needs to be acquired from the user and used to train the control system. The following section describes which types of EMG acquisition and preparation techniques that are commonly used.

As presented in section 1.2 the source of the EMG signal is the superposition of MUAPs. The energy generated in action potentials is of a very small size and is measured in microvolts. Very sensitive recording equipments is therefore key in doing electromyography. It is essential to consider the type of electrode intended to use. Electrodes come in various different sizes and shapes and are therefore very depended on the intended measurement site. Typically electrodes made of silver-impregnated plastic are used. They present characteristics by being disposable, relatively low price and by having low impedance with the skin. These can either be dry or covered with different types of gel, in order to reduce impedance and thereby noise, getting a more accurate EMG recording. Dry electrodes do not use gel, but instead rely on the skin to sweat and thereby decreasing the skin impedance. Dry electrodes should prove better to patients with sensitive skin. Different skin conditions may also effect the electrode-skin impedance. Make-up or hairy skin increase the impedance, thus the recording site should be prepared by removing hair and/or cleaning the skin with alcohol wipes. [4] The following section will introduce the choice of acquisition device used in this project.

1.3.1 Myo Armband

In this project the Myo armband (MYB) from Thalmic Labs will be used for EMG data acquisition. MYB is an electrode armband with eight dry stainless steel electrode-pairs around the inside of the armband, as depicted in figure 1.5. The advantage of dry electrodes is that they do not need to be disposed after usage as conventional gelled EMG lectrodes [4]. In addition, the MYB can communicate wirelessly to a computer via Bluetooth 4.0 [5]. Thus, making it a fast and easy usable device to use both during pilot-testing and for the final experiment. In the following section more information about the MYB will be presented.

The MYB records EMG data in a 8-bit resolution. As usual when recording EMG the higher the performed contraction is, the higher the values in the output will be. To avoid interference from power grid, the MYB is equipped with a 50 Hz notch filter from the manufacturer. However, the MYB is not able to make any further filtering, therefore this will be implemented later during signal processing described further in section 1.4. The MYB has a 200 Hz sample rate, and thus samples with a lower bandwidth than the EMG spectrum consists of, which is between 10-500 Hz [4]. Using the MYB will likely result in an aliased EMG signal and confinement in using features representing the frequency information of the signal. Besides having EMG sensors the MYB can provide position and orientation information, using its three inertial measurement units consisting of a three axis gyroscope, a three axis magnetometer and a three axis accelerometer. This inertial information is sampled at 50 Hz. [5]
When initiating the use of the armband there is two calibration phases the user must follow: the warm-up phase and the sync phase. During the warm-up phase the armband is ensuring as strong electrical connection with the muscles in the forearm as possible. This is mainly provided by light sweating on the skin under the electrodes, which improve the connection similar to electrode gel [4]. During the sync phase, the armband determines its orientation in space, position and on which arm it is placed. The MYB works most optimal when fitted tightly on the thickest part of the forearm. For users with smaller forearms a set of clips can be added for the armband to get a constrained grip. [5]

Figure 1.5: MYB from Thalmic Labs. Sensor one will always correspond to the first index in the EMG data output array, and sensor two to the second index and so forth.

When using the MYB as method for acquiring data this project will use non-specific electrode placement. The idea of the MYB is to have an easy to use device, and specific electrode placement is complicated as it requires knowledge on the specific anatomic placement of muscles in the forearm. This project has covered which muscles are active for the chosen movements, in section 1.1, in order to gain knowledge of the source of the EMG signals. When the MYB is placed correctly on the forearm the need for specific electrode placement is defeated.

1.4 Data Processing

In order to use the acquired EMG signal in myoelectric prosthetic control, it first has to be processed. Since the acquisition and most processing is done in the MYB before Bluetooth transmission, further processing of the signal is moderate. In myoelectric prosthetic control, several features are extracted from each electrode-channel for use in control, instead of using the raw EMG signal. Thus, the amount of information given to the control system is increased compared to only providing the raw EMG signal to the control system. The following two sections will briefly describe theory behind filtering and feature extraction in relation to this project.

1.4.1 Filtering

Filtering is a cornerstone in preparing an EMG signal for any kind of use. The frequency spectrum of EMG is 10 Hz to 500 Hz. [4] According to the Nyquist theorem, to achieve a loss-less representation of the signal the sampling frequency must be at least twice the maximum frequency of interest of the original signal [6]. Additionally, EMG is sensitive to movement artefacts and electrical interference. Due to these circumstances, filters are often implemented to remove these unwanted contributors [7]. General practice in filtering the EMG signal will include implementing a notch filter with very narrow width and steep slope, at frequencies 49-51 Hz or 59-61 Hz depending on the power supply. The intent is to remove any electrical interference noise. In the low frequency spectrum several recommendations (5 Hz, 10 Hz and 20 Hz) has been made for optimal corner frequency of a high-pass filter to remove movement artefacts. A low pass filter is also typically used to remove any noise and unwanted signal above 500 Hz [4].

This project will utilize a MYB for data acquisition and as mentioned in section 1.3.1 the MYB has a sample rate of 200 Hz. In relation to this project a sampling frequency of at least twice the maximum of the recorded signal is not possible [4]. This would require a sample rate of at least 1000 Hz, which cannot be achieved due to limitations in the MYB. Under other circumstances it would be astute to implement an anti-aliasing filter. However, this is not possible with the MYB since an anti-aliasing filter should be implemented before the sampling, when sampling below twice the bandwidth of the signal.

1.4.2 Feature Extraction

The input signals used for myoelectric prosthetic control are features based on the raw EMG signal. This increases the amount of information given to the control system, which facilitates more robust pattern recognition.
There are numerous feature components from an EMG signal which can be extracted either from the time-domain, frequency-domain, or time-frequency domain. Most used are features from the time- and frequency-domain. Time-domain features can be categorized in five different types based on their mathematical properties: energy information, complexity information, frequency information, prediction modelling and time-dependency. Extracting features from the frequency-domain requires a frequency transformation, calculating the spectral properties of the recorded signal, which takes up longer processing time than simply using time-domain features. Time-domain features are often chosen based on their quick and easy implementation as they do not require any transformation before extraction and are calculated based on the raw EMG signal. In addition, it is important not to choose redundant features for the classifier. [8]Extracting features for real-time prosthetic control is done by taking segments of the continuous signal, called windows. Calculation on extracting features are done in these discreet windows. This is done instead of using the instantaneous value due to the signals random nature. These windows are often overlapped to create a dense information stream for extraction. The relationship between window and overlap length is significant, when trying to determine the best representation. The window length is a matter of getting enough samples to do the calculation, but too long a window will result in delays slowing the control. Smith et al. [9]

found that the optimal window length in a classification control scheme that enables best performance ranges from 150-250 ms. Overlapping the window is a method to faster acquire windows by reusing a determined last segment of the prior window. The amount of overlap is a compromise between classification quality and responsiveness of the prosthesis; a large overlap will provide a shorter output delay, but a worse classification and vice versa [10].

1.5 Classification

For a myoelectric prosthesis to be able to distinguish between movements it needs to perform, a control scheme is needed to categorize the movements. The control scheme is trained by being given information about the EMG signal represented as the features extracted from the raw EMG. If the features between each movement are well separated the control scheme is able to recognize each distinct movement. For this purpose classification control schemes are commonly used. A classifier categorizes each movement as a class, and based on the input features it gives one output for each data window. Using a classifier thus limits the user to only performing movements which have been defined as classes. However, if trained properly a classifier can reach a low error rate for the trained movements [11]. A frequently used classification control scheme for myoelectric prosthetic control is the Linear Discriminant Analysis classifier (LDA) [11, 12, 13, 14]. The advantage of LDA is that whilst having a low computation time it still enables robust control. An assumption about the LDA is that the input needs to be Gaussian distributed, which the EMG probability properties has shown to adhere to [15, 16]. The following section provides further theoretical information about general linear classifiers and LDA.

1.5.1 Linear Classification

A linear classifier is a supervised classification method used to separate classes of data by linear decision boundaries. Each decision boundary is a hyperplane from which the distance to each feature value of a class and the center of the class is maximized. A decision boundary is defined as a linear combination of the feature values x and is given as [15]:

$$g(x) = w^t x + w_0 \tag{1.1}$$

where the weight, w, is a vector deciding the orientation of $g(x)$, and w_0 is a bias deciding the position of the hyperplane in relation to the origin. If $w_0 > 0$ the origin is on the positive side of the decision boundary, and if $w_0 < 0$ the origin is on the negative side. In the case of $w_0 = 0$ the decision boundary passes through origin. The distance from origin to the boundary is given as $\frac{w_0}{||v||}$. The position of the decision boundary is necessary to know to when separating features into regions. [15]

In a two category case the decision rule for deciding classes is to decide class ω_1 if $g(x) > 0$ and class ω_2 if $g(x) < 0$. $g(x) = 0$ then defines the decision boundary that separates the features into two decision regions R_1 for ω_1 and R_2 for ω_2. The normal vector v is orthogonal to any vector on the hyperplane, which is used to calculate the distance r from feature values x to the decision boundary [15]:

$$r = \frac{g(x)}{||v||} \tag{1.2}$$

The distance from origin and boundary to feature value x is needed to decide in which region the feature value belongs. [15] These distances are illustrated in figure 1.6.

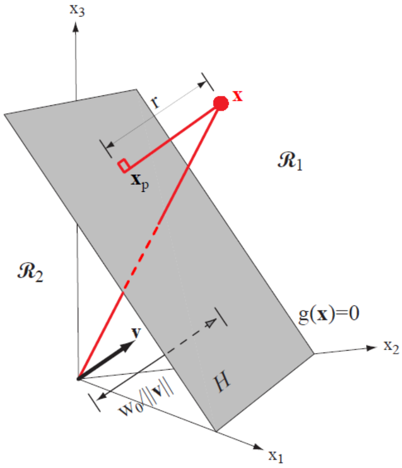


Figure 1.6: A geometric illustration of the linear decision boundary $g(x)$ that separates the feature space into two decision regions R_1 and R_2. x is the feature value, and x_p is the point on the decision boundary in which x is orthogonal projected on vector v. The distances from origin and boundary to feature value x is marked red. [15]

When feature values are to be classified into more than two classes more decision boundaries are needed. This is a multicategory case in which c numbers of boundaries are defined. When defining linear boundaries in this case any number can be chosen, but to minimize ambiguous decision regions the boundaries are defined by [15]:

$$g_i(x) = w_i^t x_i + w_{i0} \qquad i = 1, ..., c, \tag{1.3}$$

This equation follows the notation of the two-category case, with the addition of i numbers of boundaries, feature values and biases. This type of classifier is called a linear machine, dividing the feature space into c regions. A liner machine will be adopted as classification method in this project. Regions R_i and R_j, that are connected is divided by a boundary hyperplane H_{ij} defined by [15]:

$$g_i(x) = g_j(x) \tag{1.4}$$

Often regions are contiguous and will have a single boundary to separate several regions. [15] Illustrations of this case can be seen on figure 1.7.

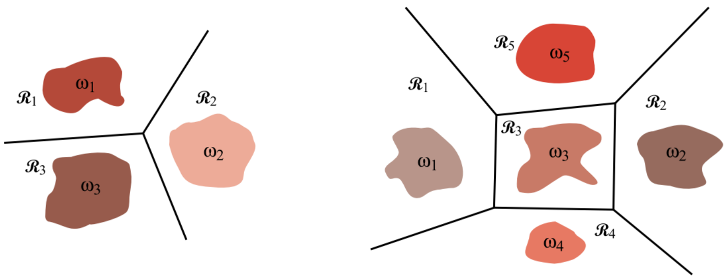


Figure 1.7: A three class (a) and five class (b) case each respectively separated by one decision boundary linear machine. [15]

When the decision boundaries $g_i(x)$ have been calculated as in equation (1.3), the input feature values can be decided upon which class they belong to by calculating the distance to the decision boundary as in equation (1.2).

1.5.2 Linear Discriminant Analysis

Despite the name, LDA is a generative model and not a discriminative model [15]. When using LDA the same procedures described in the previous section, section 1.5.1, must be completed. This however can result in classes which can show correlation. When applying LDA, classes cannot have correlation, since the data must be Gaussian distributed. Thus, when applying LDA a constraint of Gaussian distribution is put on the data. This force the class regions to have no correlation and have the same covariance matrix, where only the class means will vary. If class regions are correlated, the decision boundaries made by a discriminative linear classifier are quadratic. Thus, when using LDA and forcing the constraint of Gaussian distribution to the data enabling the decision boundaries to be linear. Additionally, with linear boundaries it is possible to calculate the likelihood of a sample value belonging to a certain class. [15, 17] The proposed method of this study is to calculate confidence scores to use as a means of providing user feedback. This will be made possible by implementing LDA as a linear classifier to enable both a classification based control scheme and to be able to calculate confidence scores to use as feedback in user training.

1.5.3 Classification Confidence Scores

The main reason for using LDA as control scheme is because it enables the calculation of confidence scores for the classes, which will be used in the user training, further described in section 1.7 to improve the prosthesis control. The principle of calculating confidence scores for input data is the basis for this project. Based on the classification of feature values by the linear classifier, confidence scores for the classes can be evaluated by computing the posterior probability of each class. Calculating the posterior probability is possible by knowing the likelihood $P(x|\omega_j)$ and the prior probability $P(\omega)$. The posterior probability for a class is a value between 0 and 1, and is calculated as follows:

$$P(\omega_j|x) = \frac{P(x|\omega_j)P(\omega)}{P(x)} \tag{1.5}$$

where ω_j represents a class and x represents a feature vector. The posterior probability is given as the product of the class conditional probability, $P(x|\omega_j)$ and the prior probability $P(\omega)$ divided by a normalization term $P(x)$ that guaranties that the posterior probabilities for all classes sums to one. $P(x|\omega_j)$ is the probability of obtaining a feature value when selecting samples randomly from a class. $P(\omega)$ is the probability of a sample from a specific class appears in its correct class, before it have actually appeared. Summation of posterior probabilities for all classes will equal 1.

1.6 Proportional Control

Classification can be used together with regression methods to provide a combination of the two in a control scheme. The output from the LDA classifier can be set to only decide upon which movement is performed, and not at which intensity the muscles used are contracting. Thus, the prosthesis can not perform any movement if only using the classification output. In statistics linear regression is often used to determine relations between variables. This notion can also be applied for myoelectric prosthetic control. While classification only provides an output on which class is recognized, a linear regression model provides a continuous output value based on the input value. If the regression model is fitted with information on different contraction levels for a given movement, control proportional to the contraction level will be achieved [18, 19, 20]. In the overall control scheme the classifier can then be used to decide which movement is performed, and a regression model can decide at which contraction level the movement is performed at. Similarly as with the classifier, regressors needs to be trained based on data acquired from the user, where the feature extracted from the raw EMG signal is used as input. This procedure is described as follows.

Different models of linear regression exist to account for different uses. When utilizing regression methods it must be considered what the input and output variables are and what type of relation these variables might have. The appropriate regression model must then be applied. Simple linear regression approximate a relation between one dependent variable Y and one independent variable X [21]:

$$Y = \alpha + \beta X + \varepsilon \tag{1.6}$$

where Y is the control output for the prosthesis, X is the feature extracted from the EMG signal, α is the predicted value of Y at $X = 0$, β is the regression coefficient in the sampled population, and ε is the error term. This model can be expanded to estimate relations between one dependent variable and several independent variables. This is called multivariate regression and expands on the equation of simple linear regression, given in equation (1.6) [21]:

$$\hat{Y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_i + \varepsilon_i \tag{1.7}$$

where i in this project corresponds to the number of channels in the MYB [21]. Since this regression model approximates the relation between several independent variables and one dependent variable, this model can be used as a control scheme in myoelectric prosthetics. Here the channel-recordings of muscle activity can be considered independent variables, and used to estimate one control output, which would be the dependent variable. [20]

1.7 User Training

As user training in relation to prosthetic control is the main focus of this project an understanding of this concept in relation to receiving a prosthetic device is of great importance. Therefore the following section will cover an introduction to the concept of user training and its importance when preparing a subject to receive a prosthetic device. In addition, some of the prior techniques of conducting user training will be presented, facilitating the possibility of assembling a user training protocol based on the most recent and cutting edge results.

When fitting an amputee with a prosthesis, the way the prosthesis is controlled is important. A lot of work lies both before and after fitting a person with a prosthesis. When developing and manufacturing a prosthesis two concepts emerge, one being system training and the other being user training. System training is training the control system to be able to recognize and differ movements based on the EMG signal being fed to the system. [22] User training on the other hand focuses on training the user in performing distinguishable movements which can be recognized by the control system. Here different types of feedback can be used to inform the user on how well the user performs a movement or how the system recognizes the users performed movements. [23, 24]

Only few studies have earlier explored the optimal way of giving visual feedback in user training [25]. In a 2014 study Powell et al. [23] provided the user with real-time visual feedback of a virtual prosthesis. This type of feedback is similar to the visual feedback a prosthesis user would receive using a normal prosthesis, although without the sensory feedback of the weight of the prosthesis. An illustration of the setup used in [23] can be seen in figure 1.8.

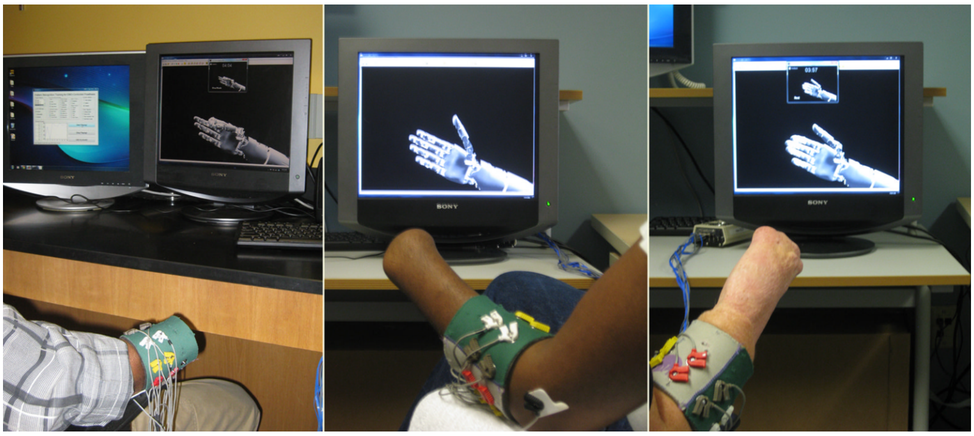


Figure 1.8: Illustration of the experimental setup used in [23]. Initially the user tries to mimic the movement, shown by a prosthesis in the interface, with their phantom limb, thus activating muscles in their residual limb corresponding to the instructed movement. The EMG produced is recorded and used to train a control system. The control system then enables the user to control the prosthesis in the interface, and receive feedback on which movement is performed.

Pan et al. [26] provided a visual feedback of an arrow to be moved on a 2D plane. The arrow was controlled by two DOF's; one controlled the horizontal position of the arrow, while the other could rotate the arrow [26]. Fang et al. [14] provided real-time visual feedback of subjects performed movement in relation to the classes defined in the system. The feedback visualized a map of clusters of different classes which subjects could match the position of a cursor to. When subjects could match the cursor to the centroid of a cluster the performed movement corresponded the best with the class of that movement. [14] An illustration of the experimental setup used in [14] can be seen in figure 1.9. All studies observed an improvement in user performance after being exposed to focused user training with visual feedback.

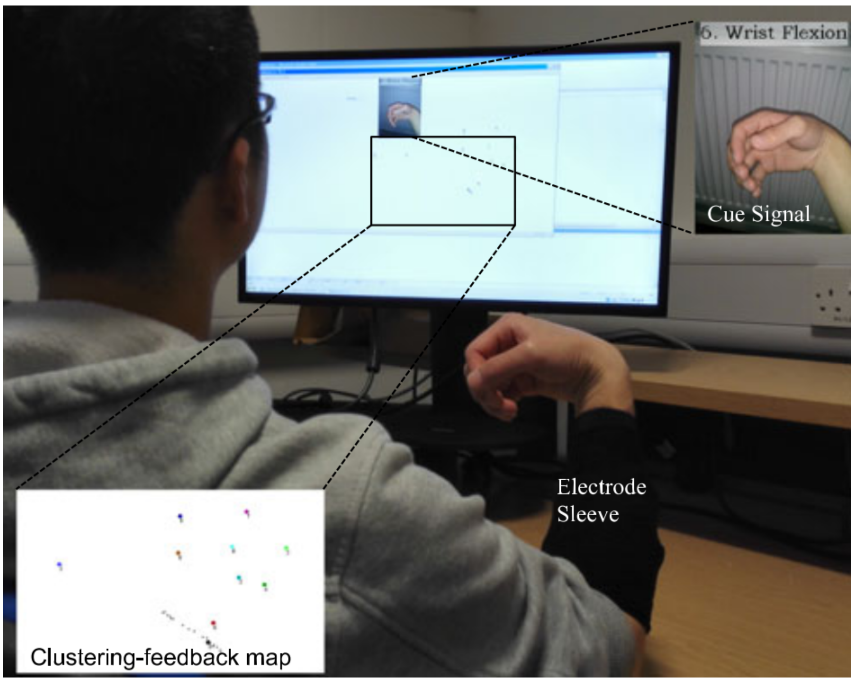


Figure 1.9: An illustration of the experimental setup used in [14]. The user received a cue on which movement should be performed. The user then had to fit a cursor, controlled by the user, to the centroid of the cluster points corresponding to the instructed movement in a clustering-feedback map.

1.8 Validating Performance

Measuring the performance of achieved prosthetic control cannot be seen as a trivial task, and different approaches can be used. The achieved performance can be measured by affixing a prosthesis on to the test subject and validate performance hereby. Often, like the current project, the subjects do not consist of actual amputees but instead healthy subject. In these cases, the performance validation is done by implementing a virtual test environment where the subjects is to control an object on the computer screen by performing movements. The following section will further elucidate the procedures of such a virtual test for validating prosthetic control.

1.8.1 Modified Fitts' Law

Fitts' Law test is a common method of quantifying performance of movements, first proposed by Paul M. Fitts in 1954 [27]. Fitts' Law states the that time required to reach a targeted area is a function of the width and distance of the target. The output of a Fitts' Law test is the throughput, as given by equation (1.8). This measure gives an idea of the trade-off between speed and accuracy. A modified Fitts' Law test designed for a virtual 2D and 3D target acquisition test has later been used by [28] and [11] respectively. Here, four additional measures were added in an real-time test, where a virtual computer cursor was used to represent the control output [11, 28]. The four additional measures, path efficiency, overshoot, stopping distance and completion rate, were made by [29] and [30]. While the throughput measure from the conventional Fitts' Law test is usable, it does not cover all aspects of the control required to complete a test. The additional four measures were added to quantitatively assess performance of naturalness, spontaneity, and compensatory motions during use. The total five proposed performance measures in assessing myoelectric control are [31]:

Throughput (TP) which represents the trade-off between speed and accuracy. TP uses the relationship of time taken to reach a certain target in seconds (MT) and the index of difficulty (ID). This forms: [11, 27]

$$TP = \frac{1}{N}\sum_{i=1}^{N}\frac{ID_i}{MT_i} \tag{1.8}$$

where i is a specific movement and N is the total number of movements. ID relates to the target distance D and width W. The ID for each target, from the origin to a specific target of a certain size is calculated using [11, 27]:

$$ID = log_2(\frac{D}{W}+1) \tag{1.9}$$

Path Efficiency (PE) describes the quality of control by making a measure of the straightness of the cursor's path to the target, by making a ratio of the actual path distance versus the optimal path distance. This tests the users' ability to continuously control the cursor position. Following the optimal path will result in a PE of 100%. PE is calculated as follows [11, 29]:

$$PE = \frac{Optimal\ Distance}{Actual\ Distance} \tag{1.10}$$

Overshoot (OS) is the number of times the cursor enters and then leaves the target before the dwell time inside the target is reached, across all target in the test, divided by the total number of targets. OS tests the users ability to control the velocity of the cursor accurately. A perfect OS-score of zero is reached if the cursor dwells within the target boundaries on the first try for all targets, and is calculated as the following [11, 29]:

$$OS = \frac{Total\ Number\ of\ Overshoots}{Total\ Number\ of\ Targets} \tag{1.11}$$

Stopping Distance (SD) describes the users ability to rest and thereby perform no movement. The SD measure is the distance moved during the dwell time across all targets, and is given as [11]:

$$SD = \sum_{i=1}^{N} (Distance\ Inside\ Target)_i \tag{1.12}$$

where i is a reached target and N is the total number of reached targets.

Completion Rate (CR) describes the percentage of targets reached within the total allowed time. This gives a general idea of the user's performance, and is calculated as [11, 30]:

$$CR = \frac{Number\ of\ Reached\ Targets}{Total\ Number\ of\ Targets} \tag{1.13}$$

1.9 Cluster Dispersion and Separability

The EMG signal for each movement class acquired from the subjects forms clusters of multidimensional data points. The lower the dispersion of the individual movement class clusters is, the more distinguishable the movements are, and the classifier will recognize the movement classes with higher accuracy. Additionally, a higher distance between cluster centroids will facilitate a higher classification accuracy further. This can be used to evaluate the users' ability to improve in performing distinguishable movements between sessions. This section describes how to calculate the cluster dispersion and separability of clusters.

1.9.1 Distance Measure

To calculate cluster dispersion, the centroid of multidimensional clusters must be calculated as in:

$$C = \frac{\sum_{n=1}^{N} a_n, b_n,\ \dots\ k_n}{N} \tag{1.14}$$

Where C is the centroid, n is the number of data point in a dimension, N is the total number of data points in a dimension and k is the number of dimensions. To calculate cluster dispersion, the Euclidean distance (ED) from data point p to the corresponding cluster q can be used, and is computed as:

$$ED(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2\ +\ \dots\ + (p_k - q_k)^2} \tag{1.15}$$

Where p_k and q_k are the coordinates of the vectors p and q respectively. This procedure is performed for all data points in a cluster, from which the average ED is calculated to obtain the dispersion of a cluster. To get a general impression of the cluster dispersion of all classes, the average of all cluster dispersions are calculated.

To calculate the cluster separability, the ED between cluster centroids is calculated. The average of the distance between all clusters are then calculated to get a general impression of the cluster separability.

2 | Methods

The background chapter covered the basic procedures associated with myoelectric prosthetic control and techniques to deal with the proposal of improving users' ability to operate a myoelectric transradial prosthesis by training the user with visual confidence score feedback. This involves descriptions of different hand movements, how the EMG signal is generated and acquired with surface EMG electrodes using the MYB and how the raw EMG signal is processed before it is segmented in windows. Furthermore, it was described how features are extracted from the segmented signal, how the feature values are used in a classification control scheme to distinguish which movement is performed, how linear regression models are used to obtain proportional control, how confidence scores can be calculated as posterior probabilities from the classification scheme, how user training previously has been used to optimize users' ability to operate a myoelectric transradial prosthesis and how the user performance is evaluated.
The information acquired in the background section will lay foundation for how the study will be designed, how the procedures will be implemented and which considerations that have been made regarding the implementation. This will be covered in the methods chapter. As the project investigates whether users' ability to operate a myoelectric prosthesis can improve after training with confidence score as visual feedback, a major focus has been put in the implementation of the user training. The chronology of the methods chapter is that the study design will be presented first, after which the implementation of the different procedures will be presented, as the procedures are implemented with regards to how the study design is formed.

2.1 Study Design

This experiment focused on training the user to improve prosthetic control on a classification-based control system. The novel approach in this study was to provide the user with visual feedback on how the system recognized the performed movements during user training, by showing the confidence scores of the movements the control system recognized. The following section will lay an overview of the structure of this experiment.
To test if myoelectric prosthetic control could be improved by using visual confidence score feedback the following research hypothesis was made:

Exposing subjects to user training, in which confidence scores of movement recognition is used as visual feedback, will show statistically significant improvement in performance in a Fitts' Law test compared to a control group receiving label feedback.

To test the hypothesis 16 subjects of mean age 25.3 ± 1.48 were recruited: 15 male and 1 female where 14 were right handed and 2 were left handed. Subjects were randomly assigned to either a control group or test group with 8 subjects in each group. The subjects enrolled were assessed to meet inclusion criteria presented in the experimental protocol for test subjects. The experimental protocol was handed out to the test participants before starting the experiment. Subjects in the test group received the experimental protocol containing specific guidance on how to perform the intended user training. The control group as well received the experimental protocol giving guidance on how to perform the user training intended for the control group. The experimental protocol containing both user training guidance schemes can be found in section A.1. During the experiment the subject was seated on a chair, with the dominant arm wearing the MYB hanging relaxed laterally down the torso as seen on figure A.7 in the experimental protocol.
The experiment was designed as a three session investigation. In each session both groups had data acquired, received user training and did a performance test through a Graphical User Interface (GUI) developed in MATLAB (2017b). During session one it was chosen to add a performance test before exposing the subjects to user training. This preliminary performance test was set to act as a baseline for each group to highlight any initiating group disparity. A graphical illustration of the stages of the study design can be seen on figure 2.1. Essential for the experiment was the difference in user training highlighted in step 3, where the groups received two different types

of visual feedback. The test group received a visual feedback of the confidence score the classifier produces, when the subject train the different types of movements. The control group received the same visual training, however this would not inform of the confidence score but instead solely show which movement the classifier thought was being performed. The sections to come will further elaborate on the implementation of each element in the experiment and how the user training differs.

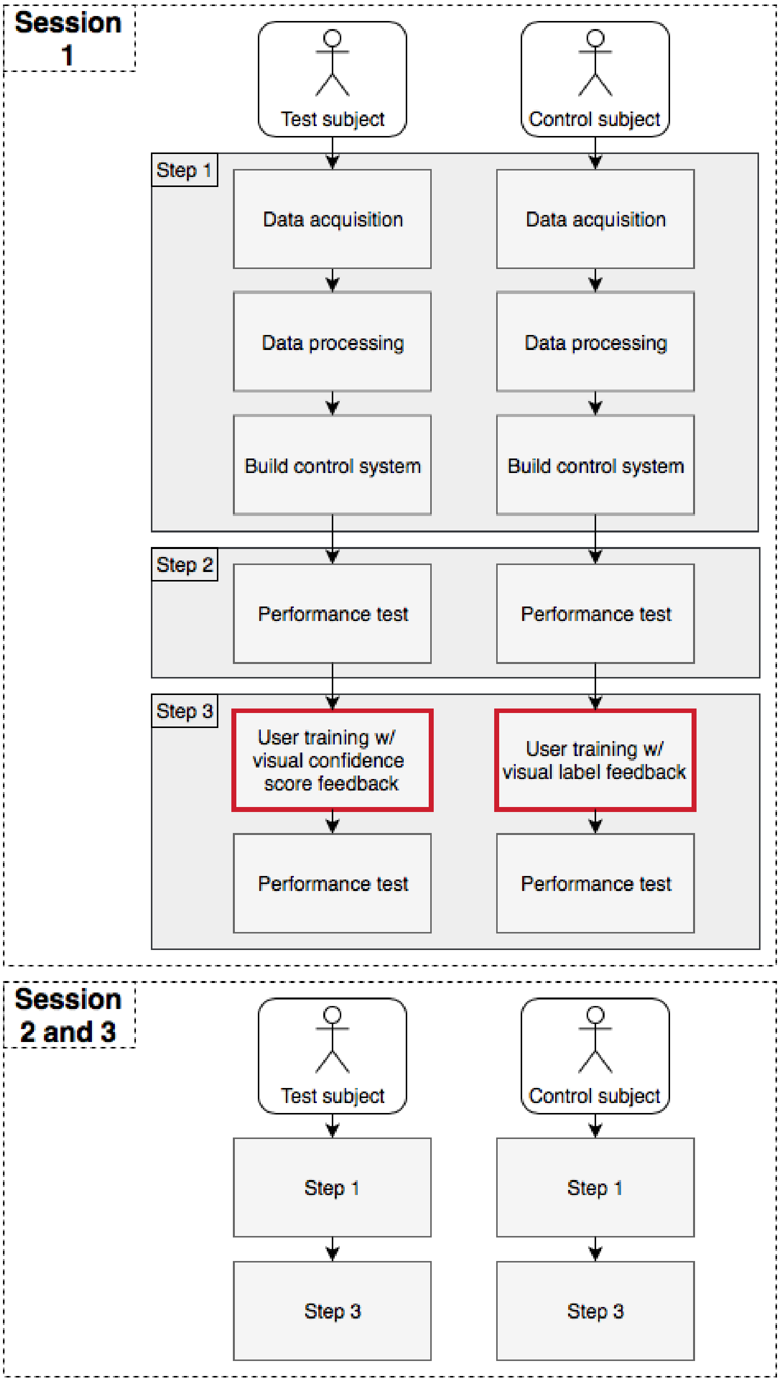


Figure 2.1: Graphical illustration of the experiment showing the steps of each session for the test and control group. Highlighted is user training in step 3 which is the only procedure that varies between the two groups, and thus the area of research interest in the experiment.

2.2 Data Acquisition

This section will clarify the method of acquiring data in this project. For data acquisition the MYB was used to record EMG signals from muscles in the forearm. The recordings were made on test subjects instructed to perform six different hand movements as introduced in section 1.1.

For acquiring data a Graphical User Interface (GUI) was designed. In the GUI the possibility to change settings for different types of recordings was implemented. The first type of recording was a baseline measurement. This recording was made in order to be able to reduce the baseline noise. This was done by subtracting the baseline from the EMG signal when the the EMG signal reached higher than the baseline. When the EMG signal was below the baseline, it was set as 0.
The second recording type was a Maximum Voluntary Contraction (MVC) which was a 15 second recording of the subject's maximum contraction of one movement that could be kept constant for 15 seconds without developing muscle fatigue. The mean MVC across all channels was set as a reference value for the following recordings.
The third type of recording was of EMG signals used to train the control system. The recordings of EMG signals were based on fractions of the MVC, which could be set using a menu in the GUI. As stated in section 2.5, three contraction levels was used: 40%, 50% and 70%. The level of contraction defined the height of the plateau of a trapezoid trajectory which would be plotted in a window in the GUI. When doing EMG recordings the subjects must perform the instructed movement to control the height of a cursor in the trapezoid plot to best match the trajectory of the trapezoid. The cursor height was calculated as the mean EMG signal across channels normalized based on the MVC. The subject only controlled the height (EMG intensity) of the cursor as the cursor would automatically move forward along the x-axis in relation with time. The recording time was 15 seconds: 2.5 seconds rest at the initiation and ending, 2.5 seconds on the trajectory incline and decline and 5 seconds on the plateau. Of the recorded time only the plateau phase and the last second of the incline and first second of the decline were used to fit the classifier. This approach provided data from a performed movement in both the transition and steady state phase. This data acquisition method was applied since the use of dynamically changing force data in training a classification-based control scheme has shown to improve performance and tolerance to proportional control [13]. During recordings the investigators evaluated whether the subject followed the trajectory well enough. Furthermore to evaluate the training data, the investigators observed a spider-plot during the acquisition, which was seen on the right side in the GUI. The spider-plot showed the amplitude output for each channel in the MYB. If the activation pattern of the channels changed dramatically, it was a sign of fluctuations in muscle activation, and thus the subject did not perform the instructed movement. If this was observed the recording was discarded and a new was be acquired. An illustration of the data acquisition GUI is shown in figure 2.2.

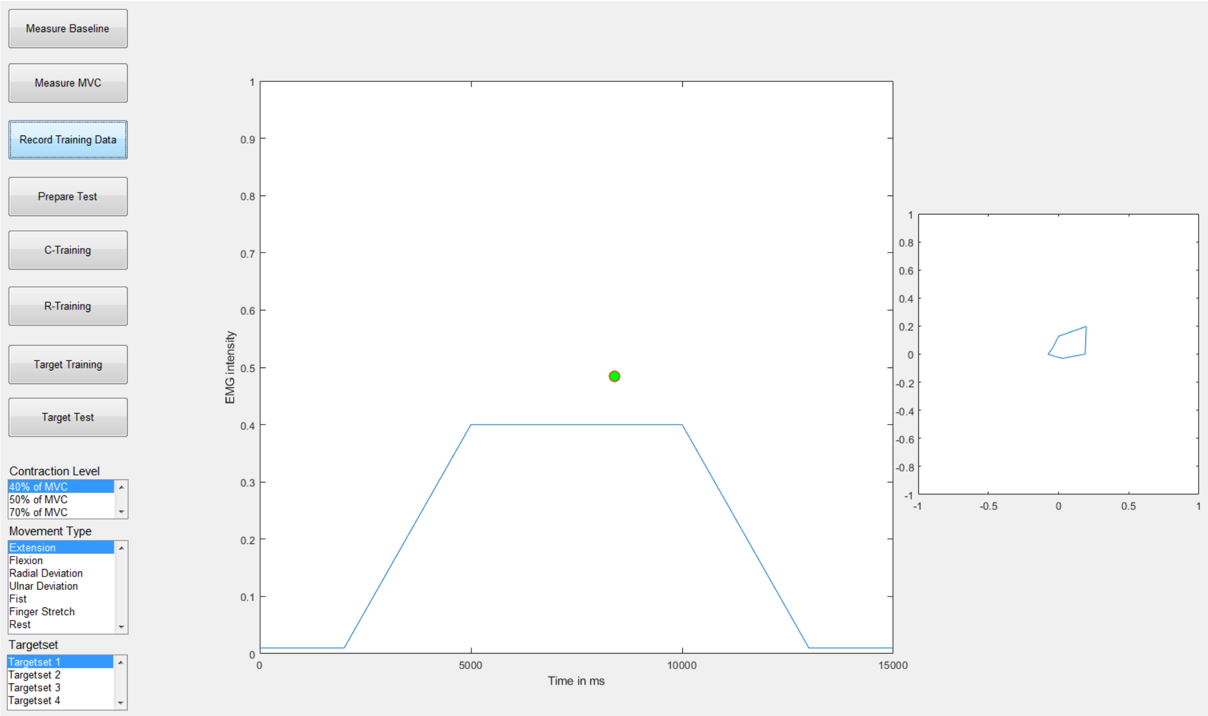


Figure 2.2: The implemented data acquisition interface. On the left is different buttons shown, where only "Measure Baseline", "Measure MVC" and "Record Training Data" is used in the data acquisition. The "Contraction Level" menu forms the trapezoidal trajectory and "Movement Type" saves the performed movement the correct label. In the center is the trapezoidal trajectory and the cursor representing the EMG signal. On the right is the spider-plot used to evaluate the quality of the performed movement.

2.3 Data Processing

The following two sections will cover the implementation of the filter used to prepare the EMG signal and the extraction of features to represent the signal. Choices behind the implemented methods builds on background knowledge acquired in section 1.4.

2.3.1 Filtering

As earlier mentioned in section 1.4.1, EMG recordings are sensitive to movement artefacts in the low-frequency spectrum, it would be resourceful to implement a high-pass filter to avoid a biased signal. In the interest of representing the signal with its true properties a 2^{nd} order Butterworth high-pass filter has been implemented with a cut-off frequency at 10 Hz. A filter steeper than second order was deselected due to a chosen trade-off between filter performance and computational performance, which is of great importance when doing real-time control. In figure 2.3 is the result of the implemented high-pass filtering. The unfiltered signal shows frequency components in low-frequency spectrum around 0-10 Hz.Furthermore is the presence of the build-in 50 Hz notch filter seen as explained in section 1.3.1.

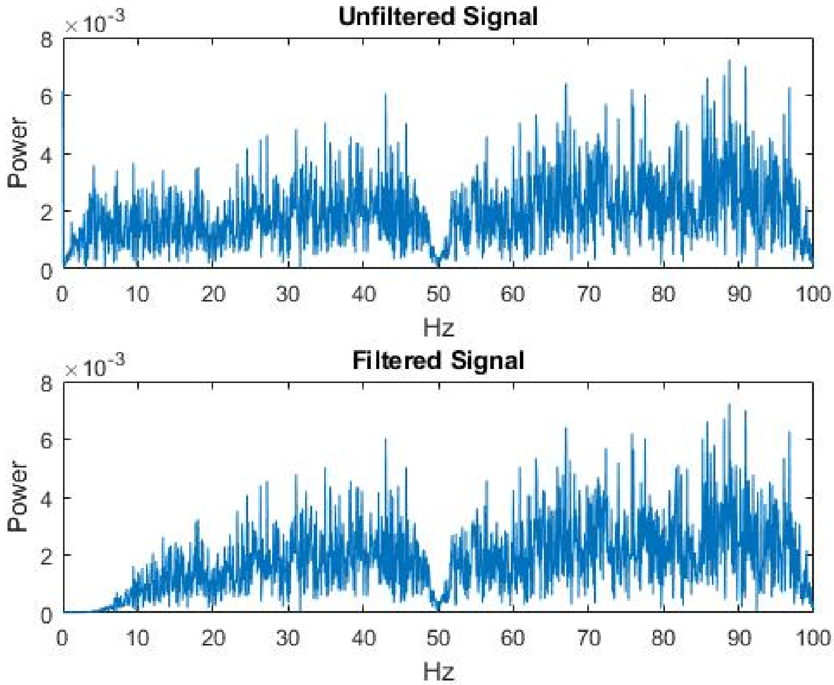


Figure 2.3: Frequency spectrum of a randomly selected EMG recording showing the difference before an after implementing the high-pass filter. The unfiltered signal shows frequency components in low-frequency spectrum around 0-10 Hz. The filtered signal shows reduction in the signal outside the cut-off frequency. Furthermore, the effect of the built-in notch can be seen as the 49-51 Hz bandwidth is attenuated.

2.3.2 Feature Extraction

In section 1.4.2 it was stated that when extracting features for real-time prosthesis control, features are extracted from the windows. This projects will for the feature extraction utilize a window size of 200 ms and a 50 % overlap for all channels. This thereby gives the possibility of calculating and updating the feature values ten times a second, thus minimizing the update delay in the real-time user training and performance test.

The features chosen to represent the information of movements contained in the signal is primarily based on recommendations from [32], where they found the optimal features for a real-time classification control scheme using the MYB. Donovan et al. [32] used so called space domain features along with the MYB and got a five percent higher accuracy than by using the well known Hudgins time domain features. Space-domain features exploit the relation between the EMG signal of each channel. A total of seven features, Mean Absolute Value (MAV), Mean Mean Absolute Value (MMAV), Scaled Mean Absolute Value (SMAV), Correlation Coefficient (CC), Mean Absolute Difference Normalized (MADN), Mean Absolute Difference Raw (MADR) and Scaled Mean Absolute Difference Raw (SMADR) were derived and the following section will explain the extraction of each. This project will use SMAV, CC, MADN and SMADR for the final classification to reduce redundancy, but all seven will be explained because some features are a combination of others. [32] Furthermore it has been chosen to extract the time domain feature of waveform length (WL) to represent frequency related information of the signal. The extraction of this feature will lastly be explained as well.
MAV is a feature that primarily is affected by the force produces when making a contraction. MAV is extracted for each window and calculated for each of the i^{th} channel. The extraction is expressed as:

$$MAV_i = \frac{\sum_{n=1}^{ws} |x_i[n]|}{ws} \tag{2.1}$$

where ws is the window size, the number of raw data points in that exact window. $x_i[n]$ is the n^{th} raw data points from the i^{th} channel.
The mean MAV across all channels, MMAV, is used to remove dependency of movement intensity. MMAV is calculated by using the MAV of all channels for the current window, and is done as following:

$$MMAV = \frac{\sum_{i=1}^{8} MAV_i}{8} \tag{2.2}$$

MMAV can be used to scale the MAV feature creating the SMAV feature. This feature should represent a non-dimensional relationship between channels. SMAV is simply calculated as:

$$SMAV_i = \frac{MAV_i}{MMAV} \tag{2.3}$$

As each of the eight EMG sensors in the MYB are located around the arm, they acquire signals from a mixture of sources. Also individual sources may affect multiple sensors depending on their size. Due to this a source measured by multiple sensors will effect their acquired signal correlation. An idea is therefore to calculate the correlation coefficient between each channel and its neighbouring channel.

$$CC_i = \frac{\sum_{n=1}^{ws} X_i[n]X_{i+1}[n]}{ws} \tag{2.4}$$

$X_i[n]$ is the n^{th} normalized data point from channel i. When calculating CC the data from each window is normalized by subtracting its mean value from each raw data point, and afterwards divided by their standard deviation.
Calculating CC can prove rather demanding in computational power due to the series og multiplication operations. Therefore Donovan et al. [32] proposed introducing a mean absolute difference-based feature of lower computational complexity which still characterizes the spatial relationship between channels. The MAD feature is normalized in the same way as CC, making up the MADN feature calculated as:

$$MADN_i = \frac{\sum_{n=1}^{ws} |X_i[n] - X_{i+1}[n]|}{ws} \tag{2.5}$$

If the normalization of the signal proves too demanding the feature can be calculated on the raw EMG signal without the normalization. This makes up the MADR feature, calculated as:

$$MADR_i = \frac{\sum_{n=1}^{ws} |x_i[n] - x_{i+1}[n]|}{ws} \tag{2.6}$$

As the SMAV feature the MAD feature can be scaled by MMAV to remove movement intensity dependency. SMADR is calculated for each channel by:

$$SMADR_i = \frac{MADR_i}{MMAV} \tag{2.7}$$

As stated in the beginning some of these features introduce redundancy, subsequently the features of SMAV, CC, MADN and SMADR are the ones used for classification. [32]
To further improve the decision foundation of the classifier it was proposed to include the time domain feature of WL calculated by:

$$WL_i = \sum_{n=1}^{N-1} |x_{i+1}[n] - x_i[n]| \tag{2.8}$$

WL is a measure of the signal complexity by calculating the cumulative length for each channel. [8]

2.4 Building the Control System

Following the data acquisition and processing, the training data was used for movement classification. The features extracted for each of the six movements and rest were used for building the classifier. In section 2.4.1 the implementation of the classification and an explanation of its output will be covered. Furthermore to be able to obtain proportional control a regression based models were made. The implementation of proportional control will be explained in section 2.4.2. An explanation of how the classifier and regression models were used in the user training and in the performance test can be found in section 2.5 and section 2.6 respectively.

2.4.1 Movement Classification

For classifying movements, LDA was used, as presented in section 1.5. The classifier was fitted with the previously acquired training data in order to build the control system. The acquired training data was assembled into matrices for each of the six movements and rest with one of the five features, containing the feature values for each of the eight channels. An example of one of these matrices can be seen in equation (2.9). This matrix contains n feature values for the feature CC for all three intensities of extension across all eight channels.

$$AllIntCC\_Ex = \begin{bmatrix} \begin{bmatrix} CCExtension40_{1,1}, CCExtension40_{1,2} \cdots CCExtension40_{1,8} \\ \vdots \qquad \ddots \qquad \vdots \\ CCExtension40_{n,1}, CCExtension40_{n,2} \cdots CCExtension40_{n,8} \end{bmatrix} \\ \begin{bmatrix} CCExtension50_{1,1}, CCExtension50_{1,2} \cdots CCExtension50_{1,8} \\ \vdots \qquad \ddots \qquad \vdots \\ CCExtension50_{n,1}, CCExtension50_{n,2} \cdots CCExtension50_{n,8} \end{bmatrix} \\ \begin{bmatrix} CCExtension70_{1,1}, CCExtension70_{1,2} \cdots CCExtension70_{1,8} \\ \vdots \qquad \ddots \qquad \vdots \\ CCExtension70_{n,1}, CCExtension70_{n,2} \cdots CCExtension70_{n,8} \end{bmatrix} \end{bmatrix} \tag{2.9}$$

The matrix consists of three sub-matrices: one for each of the intensities acquired as explained in section 2.2. The naming of the matrix is explained as that *AllInt* denotes all intensities, *CC* denotes the CC feature and *Ex* denotes the extension movement. Similar matrices were constructed for all other features for all movements named in the same fashion as the AllIntCC_Ex matrix. All these matrices were assembled into one large training matrix, TM, in a five-dimensional feature space as seen below in equation (2.10).

$$TM = \begin{bmatrix} AllIntCC\_Ex, AllIntSMAV\_Ex, AllIntSMADR\_Ex, AllIntMADN\_Ex, AllIntWL\_Ex \\ AllIntCC\_Fl, AllIntSMAV\_Fl, AllIntSMADR\_Fl, AllIntMADN\_Fl, AllIntWL\_Fl \\ AllIntCC\_Rd, AllIntSMAV\_Rd, AllIntSMADR\_Rd, AllIntMADN\_Rd, AllIntWL\_Rd \\ AllIntCC\_Ud, AllIntSMAV\_Ud, AllIntSMADR\_Ud, AllIntMADN\_Ud, AllIntWL\_Ud \\ AllIntCC\_Ch, AllIntSMAV\_Ch, AllIntSMADR\_Ch, AllIntMADN\_Ch, AllIntWL\_Ch \\ AllIntCC\_Oh, AllIntSMAV\_Oh, AllIntSMADR\_Oh, AllIntMADN\_Oh, AllIntWL\_Oh \\ AllIntCC\_Re, AllIntSMAV\_Re, AllIntSMADR\_Re, AllIntMADN\_Re, AllIntWL\_Re \end{bmatrix} \tag{2.10}$$

The classifier was trained by fitting the matrix presented in equation (2.10) with labels for each of the movements, by using the *fitcdiscr* function in MATLAB. The *fitcdiscr* function makes a LDA classifier model as described in section 1.5.2. The classifier thereby formed seven classes, one for each of the movements, with linear decision boundaries separating them. For calculating the real-time use of classification outcome and confidence scores in user training and performance test as intended, the *predict* function in MATLAB was used. The function was continuously evaluating each feature value to the different movement classes in the five dimensional feature space. Thus, the feature values were assigned to the movement class they were most likely to belong to based on the training data. The *predict* function also calculated the probability membership for the feature values to all classes giving an idea of how confident the classifier was on deciding a certain movement class and thereby indicating the correctness of the movement performed.
The classifier was only used to decide upon which movement was performed, thus not used in performing proportional control. For this purpose linear regression models were used.

2.4.2 Proportional Control

To obtain proportional control linear regression models, regressors, for each movement were made. For this purpose multivariate linear regression was used, as explained in section 1.6. To fit a regressor dependent and independent variables were needed. The independent variables were set as MAV features extracted from the raw EMG data acquired from the data acquisition as explained in section 2.2. To ensure that the regressor did not produce an output during rest, the MAV features extracted from the raw EMG from a resting position were included as independent variables in each regressor. An example of the independent variables used to fit the regressor for the extension movement is seen in equation (2.11).

$$AllIntMAV\_Ex = \begin{bmatrix} \begin{bmatrix} MAVExtension40_{1,1}, MAVExtension40_{1,2} \cdots MAVExtension40_{1,8} \\ \vdots & \ddots & \vdots \\ MAVExtension40_{n,1}, MAVExtension40_{n,2} \cdots MAVExtension40_{n,8} \end{bmatrix} \\ \begin{bmatrix} MAVExtension50_{1,1}, MAVExtension50_{1,2} \cdots MAVExtension50_{1,8} \\ \vdots & \ddots & \vdots \\ MAVExtension50_{n,1}, MAVExtension50_{n,2} \cdots MAVExtension50_{n,8} \end{bmatrix} \\ \begin{bmatrix} MAVExtension70_{1,1}, MAVExtension70_{1,2} \cdots MAVExtension70_{1,8} \\ \vdots & \ddots & \vdots \\ MAVExtension70_{n,1}, MAVExtension70_{n,2} \cdots MAVExtension70_{n,8} \end{bmatrix} \\ \begin{bmatrix} MAVRest_{1,1}, MAVRest_{1,2} \cdots MAVRest_{1,8} \\ \vdots & \ddots & \vdots \\ MAVRest_{n,1}, MAVRest_{n,2} \cdots MAVRest_{n,8} \end{bmatrix} \end{bmatrix} \tag{2.11}$$

The matrix contains four sub-matrices consisting of n feature values across all eight channels for rest and the three contraction levels: 40 %, 50 % and 70 % respectively. The naming of the matrix is explained as that *AllInt* denotes all intensities, *MAV* denotes the MAV feature and *Ex* denotes the extension movement. The independent variables matrices used to fit the regressors for the other five movements were named in the same fashion. The purpose of using data from several contraction levels was to ensure robust proportional control.
The dependent variables were similar to the independent variables. However, only a single output per window was desired. Therefore the mean of the feature values extracted from a window was calculated and scaled according to the MVC. The MVC was set as reference value of 1. The dependent variables corresponding to rest were set as 0. Thus, the dependent variables used to fit the regressor was a vector. An example of the dependent variables used to fit the regressor for the extension movement is seen in equation (2.12).

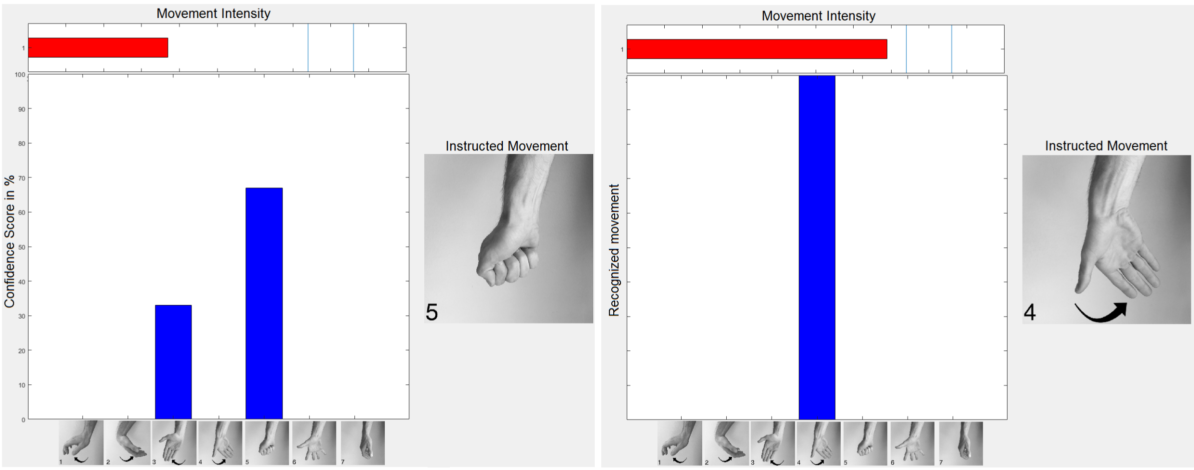
$$AllIntMAV\_ExScaled = \begin{bmatrix} \begin{bmatrix} MAVExtension40Scaled_{1,1} \\ \vdots \\ MAVExtension40Scaled_{n,1} \end{bmatrix} \\ \begin{bmatrix} MAVExtension50Scaled_{1,1} \\ \vdots \\ MAVExtension50Scaled_{n,1} \end{bmatrix} \\ \begin{bmatrix} MAVExtension70Scaled_{1,1} \\ \vdots \\ MAVExtension70Scaled_{n,1} \end{bmatrix} \\ \begin{bmatrix} 0_{1,1} \\ \vdots \\ 0_{n,1} \end{bmatrix} \end{bmatrix} \tag{2.12}$$

The output of the regressor would then be larger the more forceful a contraction the subject produced and vice

versa. The *fitlm* function in MATLAB was used to fit a regressor with the independent variables from equation (2.11) and dependent variables from equation (2.12) as input. A total of six regressors were made; one for each movement used in the experiment.
The regressors use for proportional velocity control in the control system was controlled by the classifier decided upon which movement was performed. The output movement were then fitted by the regressor and used for proportional control.

2.5 User Training

User training was the essential element of investigation in this experiment, with the movement confidence feedback as the area of interest. This section provides information on how the visual feedback was presented to the subjects in the two experiment groups during the user training, and what the objective for the subjects was.
The user training interface contained the following feedback: an illustration of the movement needed to be performed, a horizontal bar visualizing the contraction level and a vertical bar plot visualizing which movement is being recognized by the control system, as shown in figure 2.4. The difference in feedback given between subject group, lied in the information given in the vertical recognition bar plot.



(a) Test group user training interface. (b) Control group user training interface.

Figure 2.4: Illustration of the user training interface for the test group (a) and the control group (b). The vertical bar plot indicates which movement is being recognized indicated by the images of each movement and the horizontal bar plot indicates contraction level. The two vertical lines in the contraction level bar plot illustrates the contraction level interval the subject must reach. The large picture of a movement on the right of the bar plot indicates which movement needs to be performed. The difference between the feedback the two subject groups receive is the information given in the vertical recognition bar plot. The control group only sees a full bar of the movement the control system recognizes the most, whereas the test groups receives the exact recognition probabilities of all movements.

The illustration of the movement needed to be performed was shown for 30 seconds, after which an illustration indicating rest was shown for 7 seconds followed by a countdown from 3 to 1 seconds indicating the time left of the resting period. Thus, the subject needed to perform a movement for 30 seconds and rest for 10 seconds before another movement needs to be performed. The subjects needed to perform all movements in four different contraction level intervals of their maximal intensity, starting with the highest interval: 75-85 %, 55-65 %, 35-45 % and 15-25 %, visualized by the two vertical lines in the horizontal contraction level bar plot. The subjects needed to perform all movements in the same contraction level interval before moving to a new interval. The instructed movements were trained in a random order.
The horizontal bar showed the contraction level. This was calculated as the mean of the latest three intensity outputs as computed in section 2.4.2, regardless of the movement being recognized. This resulted in a 400 ms delay in the visualization of the horizontal bar at the initiation of the training of a movement, due to the windowing used in feature extraction as mentioned in section 2.3.2. However, the delay was not noticeable and the averaging

of the intensity output resulted in a smooth visualization of the vertical bar.
The vertical bar plot showed which movement(s) the control system recognized. For a movement to be recognized as an active movement, the subjects had to perform the movement with more than 15 % contraction intensity. The test group received information on the exact probabilities for the movements that were recognized. Thus, more bars could appear at the same time as seen in figure 2.4 (a). The purpose of this feedback was for the subjects to adapt to how the control system recognized the instructed movement. It gave the subjects the possibility of noticing which movements that also were recognized by the control system when performing the instructed movement. When the instructed movement was not recognized with a 100 % certainty the subject could use the information to slightly correct the performed movement until the control system recognized the instructed movement with a 100 % certainty. This bar plot was calculated as the mean of the recognition certainties calculated from the latest three feature inputs, which resulted in a smooth visualization of certainties for the movements in the bar plot.
The control group only received information on which movement was recognized the most, represented as a single full bar at the recognized movement as seen in figure 2.4 (b). Thus, the control group was not informed on the exact probabilities of which movements the control system recognized. This bar plot was calculated by finding the highest mean certainty output among all movement classes. This was done taking the mean of the last three feature inputs in each movement class..
To motivate the subjects and to train the transition to and from resting position a task was included in the user training. The subjects were instructed in performing the instructed movement with a 100 % certainty inside the instructed contraction level interval. When this was reached the horizontal bar would turn from red to green. The subjects were instructed in withholding the green colour for one second, after which the horizontal bar would turn blue and a light sound was played. After this task was completed the subjects was instructed in returning to rest and perform the task again. The objective for the subject was then to make the horizontal bar blue as many times as possible during an instructed movement of an instructed contraction level. The number of times the horizontal bar got blue during an instructed movement in an instructed contraction level was saved for later data analysis. Between each movement there was a 10 second break.
To summarize, the overall objective for the subjects during the user training was to adapt to how the control system recognized the performed movement. The user training was implemented for the subjects to possibly improve their ability to use the control system. Their ability to use the control system was then evaluated in the modified Fitts' Law test.

2.6 Performance Test

As stated in section 1.8 the theory behind Fitts' Law, this project will utilize a modified Fitts' Law test, consisting of a virtual target reaching test to evaluate the progress of subjects performance after user training. The proposed modified Fitts' Law test utilizes the performance measures: throughput, path efficiency, overshoot, stopping distance and completion rate. The following section describes how the Fitts' Law test has been implemented in this project.

2.6.1 Virtual Target Reaching Test

The virtual targets reaching test was implemented into the same GUI used for data acquisition and user training, first mentioned in section 2.2. When enabling the target reaching test in the GUI, the subject was met with the interface shown in figure 2.5. Here the subjects controlled the position of the cursor by performing movements shown on at the vertical and horizontal axes. Thus, extension of the hand would move the cursor to the right, and flexion would move the cursor to the left. Similarly, radial and ulnar deviation moved the cursor up and down respectively. This approach was used to improve the intuitiveness of the control as the direction of the cursor imitates the subjects' hand when performing the movements. Subjects controlled the size of the cursors red area by opening and closing the hand, where an open hand increases the area and a closed hand decreases the area. By being able to control the cursor on the two dimensional plane on the grid area and being able to control the size of the red cursor area, the test effectively had three dimensions.
Each target was presented by an area with a center and an outer circle. Targets existed in three sizes, both smaller and larger than the cursor, in order to test the opened and closed hand DOF. Similarly, the cursor changed size after

being rest to test this DOF. The target reaching test consisted of reaching a total of 16 targets which each appeared for 15 seconds at different distances from the origin. Thus, the maximum duration of the test was 4 minutes. The order of appearance was fixed but different for each of the four sessions but the same across subjects. Thus, individual subjects would experience the targets as appearing in a new order with each test.

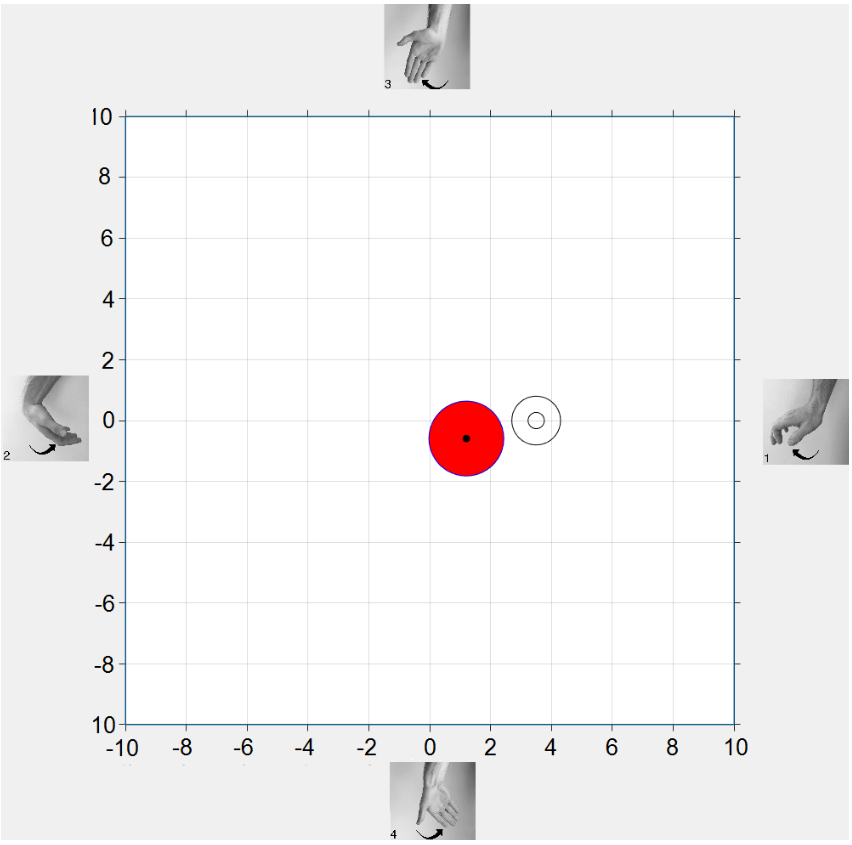


Figure 2.5: The implemented interface for the modified Fitts' Law test. The grid was the area in which the subject must reach targets with the controlled cursor. The cursor is the red circle area with the black dot in the center. Targets were shown as a circle area with an larger outer circle. Performing the movement corresponding to the axis image would move the cursor in the direction of the image. Opened/closed hand would increase/decrease the size of the cursor.

Subjects had to reach the targets inner circle with the cursor dot and expand or decrease the outer circle of the cursor to reach a size close to that of the target. A moderate size threshold was implemented to make it possible to reach targets, without a 100% accuracy of control. If a subject successfully reached a target, the cursor would change color from red to green, and if the subject dwelled in this position for a 1 second the cursor colour changed to blue, and a bell chime would sound to indicate that the target was reached. The cursor position was reset to the center of the grid area and the color of the cursor would revert back to red. If a target was not reached within 15 seconds the current target would disappear, a new target would be shown and the cursor position would be set to the center of the grid area. The approach of resetting the cursor position after each target was to equalize the path for every subject.
As mentioned earlier multiple performance measures were recorded during the target reaching test. One of these measures was throughput which depend on the index of difficulty for the targets, as stated in section 1.8.1. The ID was calculated for all targets shown in figure 2.6.

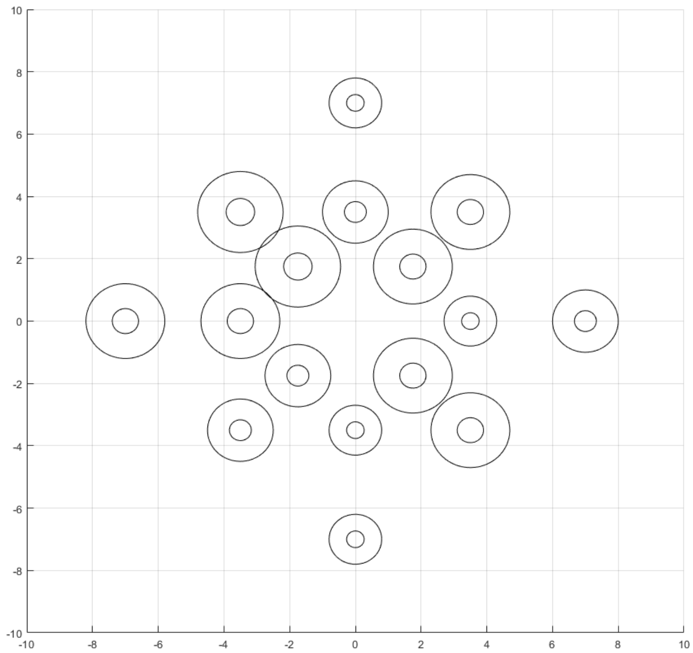


Figure 2.6: The implemented interface for the modified Fitts' Law test, where all 16 targets to be reached, are shown. Here it is seen that the smaller center circle has the same width in all targets. Furthermore, the three different target sizes are shown for the outer circle of the targets.

The configuration of targets in the test resulted in eight combinations of width and distance. The smaller circle has the same width in each target thereby possessing the same area in which the cursor center has to be within. The index of difficulty for the targets used in the modified Fitts' Law test can be seen in table 2.1

Table 2.1: The index of difficulty used in the Fitts' Law test.

Distance	Width	ID
28.0	0.33	6.41
24.5	0.33	6.22
22.0	0.33	6.01
18.5	0.33	5.82
16.0	0.33	5.61
13.0	0.33	5.32
12.5	0.33	5.27
9.5	0.33	4.88

A trace of the cursor movement throughout the whole test is recorded to decide the subjects path deviation from the optimal path to calculate the path efficiency and distances to targets. An example of a cursor trace is shown in figure 2.7.

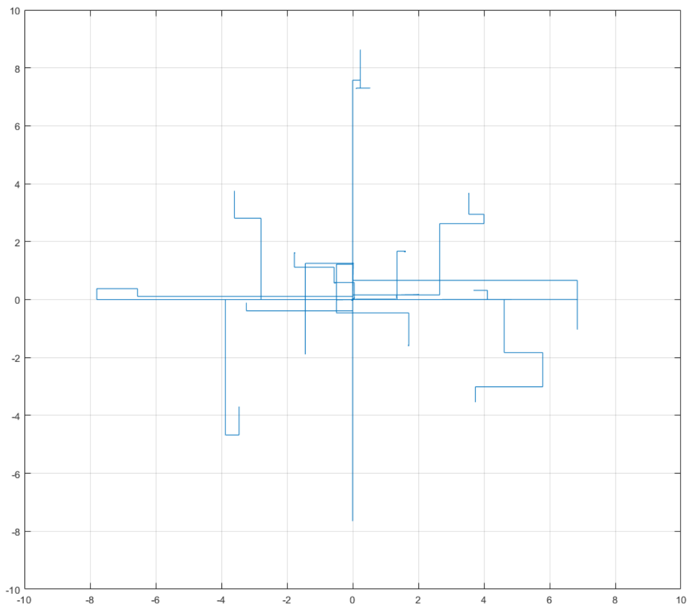


Figure 2.7: The trace of the cursor movement of an example subject after a target reaching test. This trace was not visible for the subjects.

The number of times a target is reached and exited without completing the dwell time, was recorded and used to calculate subjects' overshoot. Similarly to tracking the travelled distance of the cursor inside the grid area, the travelled distance inside of each target was also recorded to calculate the stopping distance. The number or reached targets were recorded to calculate the completion rate.
Following the completion of recording target reaching test data from all subjects the performance measures introduced in section 1.8.1 were calculated and presented in the Results in chapter 3.

3 | Results

In this chapter the results processed from collected data will be presented. All data processing have been done in accordance to earlier introduced theory and implementations of methods described in chapter 1 and chapter 2 respectively. Multiple comparison test of improvement across session of the two subject groups have been computed through a Friedman test, since the data proved to be non-Gaussian. When detecting an effect a Tukey-Kramer test was applied to correct for the problem of multiple comparison. Testing statistical differences between subject groups for each session was computed through a Mann-Witney U test.

3.1 Performance Results

This section will present the results acquired from the Fitts' Law target reaching test described in section 2.6. The test had five measures which each expresses a parameter of subjects' performance. Subjects were divided into two groups, one test group which received exact class confidence scores during user training, and a control group which only received label feedback. The plotted mean and standard deviations of each measure in the performance test for session can be seen in figure 3.1.

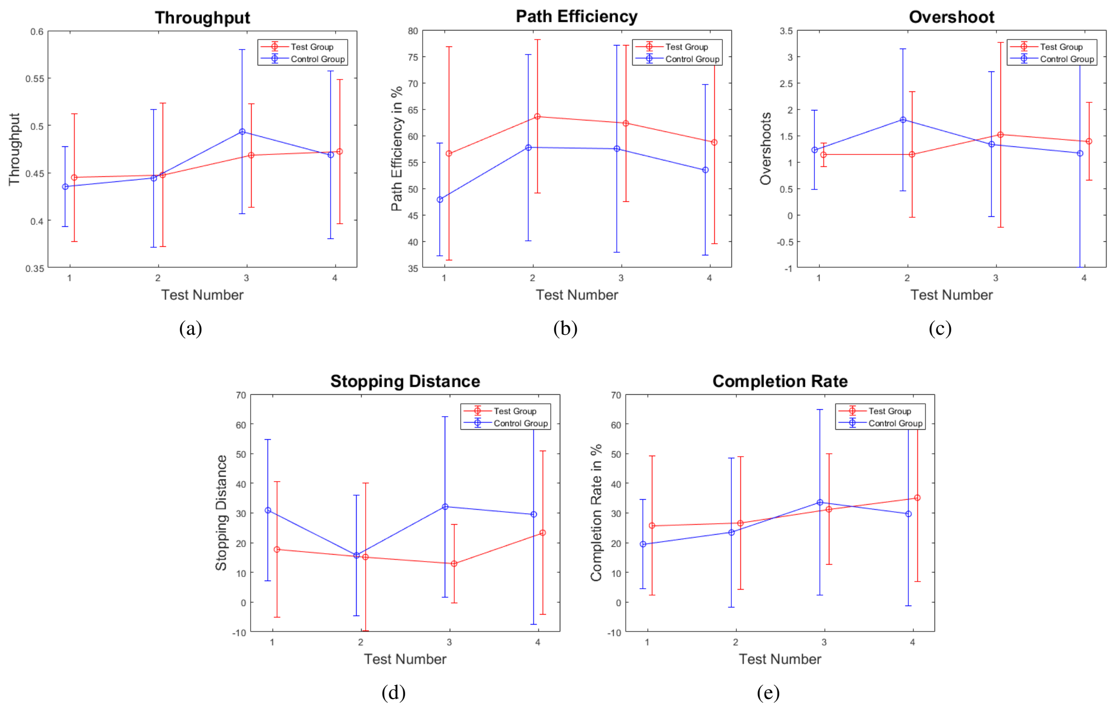


Figure 3.1: Figure illustrating the five performance measures; *a) Throughput, b) Path efficiency, c) Overshoot, d) Stopping distance, e) Completion rate*, used for quantifying user performance across all four tests. Test number 1 is the acquired baseline used for assessing group homogeneity and the following numbers indicate performance test results after user training in each session. The red line indicates the progression of the test group and the blue of the control group.

The aim of using the Fitts' Law target reaching test was to apply a method to quantitatively evaluate the performance of subjects after user training in each session. The information drawn from the measures are described in section 1.8.1. The baseline performance test showed no difference between the two group, showing the two groups to be homogeneous at initiation. The Fitts' Law test results did not show any significant improvement over the three sessions for any of the five test measures for both the test and control group ($p > 0.05$). Similarly, there was no significant difference between the two groups performance in any sessions ($p > 0.05$), meaning neither of them performed significantly better than the other group in any of the sessions.

3.2 User Training

This section covers the results acquired from measurements obtained during user training sessions. During user training subjects were instructed to train movements in being performed such that the control system recognized the movement as the actually performed movement. During this training the number of times subjects correctly performed an instructed movement to the contraction interval shown in the training interface was recorded, and will be referred to as the number of repetitions.

3.2.1 Total Completion Rate

No significant difference in the total number of repetitions was found between sessions of either group ($p > 0.05$). When comparing the total number of repetitions of each session between groups accordingly, no significant difference were found either ($p > 0.05$).

3.2.2 Contraction Levels

An increased ability to reach the low intensities was found for the control group ($p < 0.05$, session $1 = 16.13 \pm 5.59$, session $3 = 21.38 \pm 6.78$). Otherwise, similar results were yielded for both groups when comparing the subjects' ability to reach the three other contraction levels between sessions ($p > 0.05$).
No difference was found, when comparing the two groups' ability to reach different intensities during training, either ($p > 0.05$).

3.2.3 Ability to Perform Movements

Comparing the ability to perform different movements during the training showed a significant improvement for the test group in ulnar deviation ($p < 0.05$, session $1 = 11.38 \pm 4.27$, session $3 = 16.13 \pm 2.95$) and open hand ($p < 0.05$, session $1 = 11.25 \pm 3.85$, session $3 = 17.88 \pm 2.46$). A significant decrease in performance was found for the control group's ability to perform flexion ($p < 0.05$, session 2$= 16.63 \pm 2.77$, session 3$= 11.00 \pm 3.16$). Otherwise no significant difference between the three sessions for the two groups was found ($p > 0.05$). A significant difference ($p < 0.05$) was found between the test and control groups ability to reach the closed hand movement, with a mean of 26.8 ± 13.5 number of repetitions for the test group and 38 ± 12.2 for the control group. No significant difference was found for any of the other movements when comparing the two groups ($p > 0.05$).

3.3 Cluster Dispersion and Separability

In this section results from the data acquisition are presented. The data used for training the LDA based classifier was examined. Each movement resulted in a cluster of data points, which are examined in this section, in order to analyse the change in cluster dispersion and distance between clusters centroids.

3.3.1 Cluster Dispersion

The mean distance from data points to the cluster centroid was calculated. This showed no significant difference for the test group ($p > 0.05$), but a significant difference was found for the control group ($p < 0.05$). The Tukey-Kramer correction showed the significant difference was between session one and three ($p < 0.05$), where the mean for session one was 502.02 ± 274.88 arb. unit, and session three was 323.43 ± 171.13 arb. unit. The comparison between groups showed that the control group achieved a significant improvement of within cluster distances compared to the test group in session three ($p < 0.05$), where the test group had a mean distance within clusters of 584.34 ± 250.02 arb. unit, while the control group had 323.43 ± 171.13 arb. unit.

3.3.2 Cluster Separability

For both groups the mean distance between the cluster centroids were calculated. The change in between cluster distances over the three sessions showed no significant difference for both groups ($p > 0.05$). Likewise, no significant difference of cluster distances between the groups was found ($p > 0.05$).

Bibliography

[1] Bart Peerdeman et al. “Myoelectric forearm prostheses: State of the art from a user-centered perspective”. In: *The Journal of Rehabilitation Research and Development* 48.6 (2011), p. 719.

[2] Frederic H. Martini, Judi L. Nath, and Edwin F. Bartholomew. *Fundametals of Anatomy & Physiology*. 2012.

[3] Peter Konrad. “The ABC of EMG”. In: *A practical introduction to kinesiological* ... April (2005), pp. 1–60.

[4] Jeffrey R. Cram. *Cram's Introduction to Surface EMG*. Eleanor Cr. 2012.

[5] *Thalmic Labs Myo Armband*. 2013.

[6] Marco Pozzo. *Biomedical Technology and Devices Handbook*. 1949. 2004.

[7] Carlo J. De Luca et al. “Filtering the surface EMG signal: Movement artifact and baseline noise contamination”. In: *Journal of Biomechanics* 43.8 (2010), pp. 1573–1579.

[8] Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul. “Feature reduction and selection for EMG signal classification”. In: *Expert Systems with Applications* 39 (2012), pp. 7420–7431.

[9] Lauren H Smith et al. “Determining the optimal window length for pattern recognition- based myoelectric control: balancing the competing effects of classification error and controller dela”. In: *IEEE Trans Neural Syst Rehabil Eng.* 19.2 (2014), pp. 186–192.

[10] Todd R. Farrell and Richard F. Weir. “The optimal controller delay for myoelectric prostheses”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 15.1 (2007), pp. 111–118.

[11] Erik J. Scheme, Bernard S. Hudgins, and Kevin B. Englehart. “Confidence-based rejection for improved pattern recognition myoelectric control”. In: *IEEE Transactions on Biomedical Engineering* 60.6 (2013), pp. 1563–1570.

[12] K Englehart and B Hudgins. “A robust, real-time control scheme for multifunction myoelectric control”. In: *IEEE Trans Biomed Eng* 50.7 (2003), pp. 848–854.

[13] E Scheme and K Englehart. “A comparison of classification based confidence metrics for use in the design of myoelectric control systems”. In: *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2015), pp. 7278–7283.

[14] Yinfeng Fang et al. “Classifier-Feedback-Based User Training”. In: *IEEE Transactions on Biomedical Engineering* 64.11 (2017), pp. 2575–2583.

[15] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. 2nd Editio. 2001, p. 680.

[16] Kianoush Nazarpour et al. “A note on the probability distribution function of the surface electromyogram signal”. In: *Brain Research Bulletin* 90.1 (2013), pp. 88–91.

[17] R. A. Fisher. *The Use of Multiple Measurements in Taxonomic Problems*. 1936.

[18] Han Jeong Hwang, Janne Mathias Hahne, and Klaus Robert Müller. “Real-time robustness evaluation of regression based myoelectric control against arm position change and donning/doffing”. In: *PLoS ONE* 12.11 (2017), pp. 1–22.

[19] J. M. Hahne et al. “Linear and Nonlinear Regression Techniques for Simultaneous and Proportional Myoelectric Control”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.2 (2014), pp. 269–279.

[20] Simon Bruun et al. “The effect of limb position on myoelectric prosthetic control using linear regression”. In: *AAU* (2017).

[21] Zar Jerrold H. *Biostastical Analysis*. 5th ed. Pearson, 2009.

[22] Anders Fougner et al. “Control of upper limb prostheses: Terminology and proportional myoelectric controla review”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20.5 (2012), pp. 663–677.

[23] Michael A. Powell, Rahul R. Kaliki, and Nitish V. Thakor. “User training for pattern recognition-based myoelectric prostheses: Improving phantom limb movement consistency and distinguishability”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.3 (2014), pp. 522–532.

[24] Ann M. Simon, Blair A. Lock, and Kathy A. Stubblefield. “Patient training for functional use of pattern recognition controlled prostheses”. In: *Journal of Prosthetics and Orthotics* 24.2 (2013), pp. 56–64.

[25] Ning Jiang et al. "Myoelectric Control of Artificial Limbs: Is There a Need to Change Focus?" In: *IEEE Signal Processing Magazine* 29.5 (2012), pp. 150–152.

[26] L. Pan et al. "Transcranial direct current stimulation versus user training on improving online myoelectric control for amputees". In: *Journal of Neural Engineering* 14.4 (2017), pp. 1–15.

[27] Paul M. Fitts. "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movements". In: *Journal of Experimental Psychology* 47.3 (1954), pp. 381–391.

[28] Ernest N. Kamavuako, Erik J. Scheme, and Kevin B. Englehart. "On the usability of intramuscular EMG for prosthetic control: A Fitts' Law approach". In: *Journal of Electromyography and Kinesiology* 24.5 (2014), pp. 770–777.

[29] Richie Poulton et al. "Evaluation of Head Orientation and Neck Muscle EMG Signals as Command Inputs to a Human-Computer Interface for Individuals with High Tetreplegia". In: *IEEE Trans Neural Syst Rehabil Eng.* 360.9346 (2013), pp. 1640–1645.

[30] Ann M. Simon et al. "The Target Achievement Control Test: Evaluating real-time myoelectric pattern recognition control of a multifunctional upper-limb prosthesis". In: *Journal of Rehabilitation Research and Development* (2011), pp. 619–627.

[31] E Scheme and K Englehart. "Validation of a selective ensemble-based classification scheme for myoelectric control using a three dimensional Fitts' law test". In: *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 21.4 (2013), pp. 616–623.

[32] Ian M. Donovan et al. "Simple space-domain features for low-resolution sEMG pattern recognition". In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2017), pp. 62–65.

A | Appendices

A.1 Experiment Protocol

Title of the project
Examining if Confidence Score Feedback During User Training Can Improve Users' Ability in Controlling Upper Limb Prosthetics

Details on investigators
All investigators are 2nd semester biomedical engineering master students at Aalborg University.

Purpose
The purpose of this experiment is to train the subject in getting better at controlling a functional prosthesis. The subject will be training seven hand movements that is used in activating the prosthesis. In the experiment the prosthesis will be represented on a computer screen, where the subject will receive visual feedback on how the prosthesis interpret the hand movement performed by the subject. By receiving this visual feedback it is hypothesized that the user will get better at controlling the prosthesis over time.

Background
Electromyography (EMG), or muscle signals, is widely used for controlling functional lower arm prosthetics for transradial amputees. The ideal purpose of a functional prosthesis is to behave as functional as possible compared to a biological arm. Functional prosthetics that rely on pattern recognition-based control are becoming exceedingly good in performance in a clinical environment, due to highly optimized system control. However, still only one commercially available pattern recognition-based prosthesis exist. Users reject these functional prosthetics usually due to functionality issues when utilizing them in daily life tasks outside the clinical environment. Many improvements have been made in the area of system control, but another approach of improving the prosthetic control is by training the user. User training has only been explored scarcely in the research literature, thus, new techniques to improve the user's ability to control a prosthesis are yet untouched. This experiment will focus on training the user to improve prosthetic control on a fixed pattern recognition-based control system. The novel approach in this study is to provide the user with information on how well the system recognizes the performed movement during user training.

Research hypothesis
Exposing subjects to user training, in which confidence scores of movement recognition is used as feedback, will show statistically significant improvement in performance in a classification-based myoelectric prosthetic control scheme, when compared to subjects who have not had the same feedback during user training.

Session time
The experiment consist of three sessions, which are spread over three consecutive days; one session per day. Each session is estimated to have a total duration of 30-60 minutes.

Inclusion criteria
The subject needs to be:

- able bodied.
- above 18 years old.
- able to read, understand and speak Danish and/or English.
- assessed by the investigators to understand and perform the instructions given during the experiment.

Exclusion criteria
The subject must not have:

- diseases that might influence subject performance.

Experiment procedure

The experiment consists of three sessions containing different procedures as illustrated on figure A.1. The concept and chronology of each procedure is described below the illustration. During the experiment it is important that the subject is placed sitting on a chair, with the arm wearing the Myo armband (MYB) hanging relaxed down by the side of the body, as shown in figure A.7 on page 42 illustrating the experiment setup.

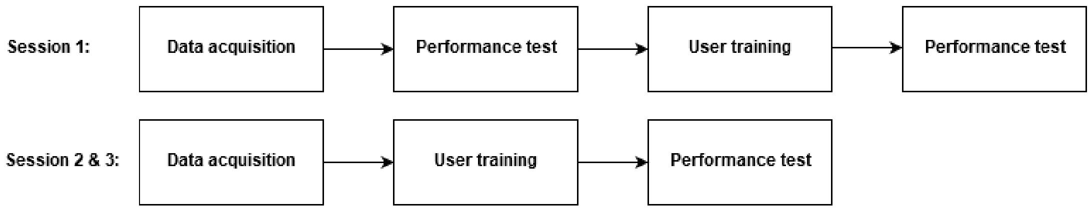


Figure A.1: Pipeline for the three sessions in the experiment and what procedures each session contains.

Data acquisition
For the myoelectric control system to be able to identify a performed movement as the movement that is actually performed, it needs information about how the movement looks when represented as a EMG signal. Thus, EMG data needs to be acquired from the forearm of the subject while the subject performs the movements that is used in the experiment, see figure A.6 on page 41. This data is fed to the control system to train the system to recognize each movement. In this experiment EMG data will be acquired from the subject with an EMG electrode armband: MYB from Thalmic Labs. The chronology of this procedure is as follows:

1. Apply MYB on dominant forearm at the thickest part.
2. Synchronize MYB by performing wrist extension until three distinct vibrations are felt from the MYB.
3. Perform 15 seconds of maximum voluntary contraction (MVC) of instructed movement. The MVC is a contraction the subject is able to withhold in a constant intensity for the 15 seconds. Following the MVC the subject will be given a 30 seconds resting period to avoid muscle fatigue.
4. Perform three 15 seconds contraction trails of respectively 40%, 50% and 70% of MVC. During these contractions the subject will control a green marker representing the EMG signal and try to follow a trapezoidal trajectory as precise as possible. The trapezoidal trajectory consists of two 2.5 second transition phases and one 5 second plateau phase. Between each trial the subject will be given a 10 seconds resting period to avoid muscle fatigue.
5. Repeat step 3-4 until training data from all four wrist movements has been recorded.

An illustration of the interface used for data acquisition is shown in figure A.2.

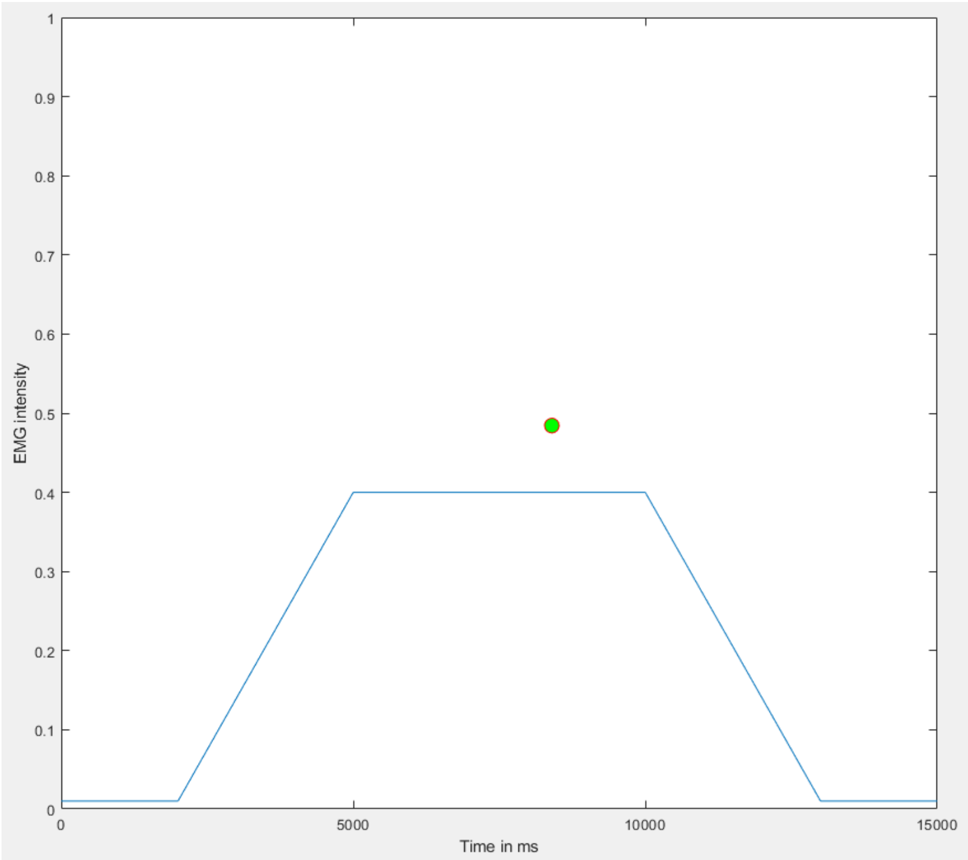


Figure A.2: Illustration of the data acquisition interface showing the trapezoidal trajectory and the green marker representing the EMG signal.

User training for Test Group

The purpose of user training is for the subject to train the movements used in the performance test. During the user training the subject will train one movement at a time at different contraction levels. When training a movement, visual feedback in form of confidence scores on how well the control system recognizes movements, is shown in percentage in a bar plot. In addition, the level of contraction is shown in a horizontal bar above the other bar plot. When performing the instructed movement at the instructed level of contraction the horizontal bar plot will appear green; if it is outside the instructed level or if the system does not recognize the performed movement, it appears red. The aim for the subject is to reach and withhold the instructed contraction level with 100 % confidence for each movement. When the subject withholds the contraction level inside the instructed contraction level for 1 seconds with a 100 % confidence the colour of the horizontal bar will appear blue, and a bell chime will appear. This indicates that the subject is performing well. After it has appeared blue, the subject must return to rest and perform the movement again and try to reach the instructed contraction level with a 100 % confidence. An additional aim for the subject is to make the horizontal bar plot appear blue as many times as possible. The chronology of this procedure is as follows:

1. Perform instructed movement at 75-85 % contraction level for 30 seconds followed by 10 seconds rest.
2. Perform step 1 for the remaining movements.
3. Repeat step 1-2 at 55-65 % contraction level.
4. Repeat step 1-6 at 35-45 % contraction level.
5. Repeat step 1-6 at 15-25 % contraction level.

An illustration of the interface used for user training is shown in figure A.3.

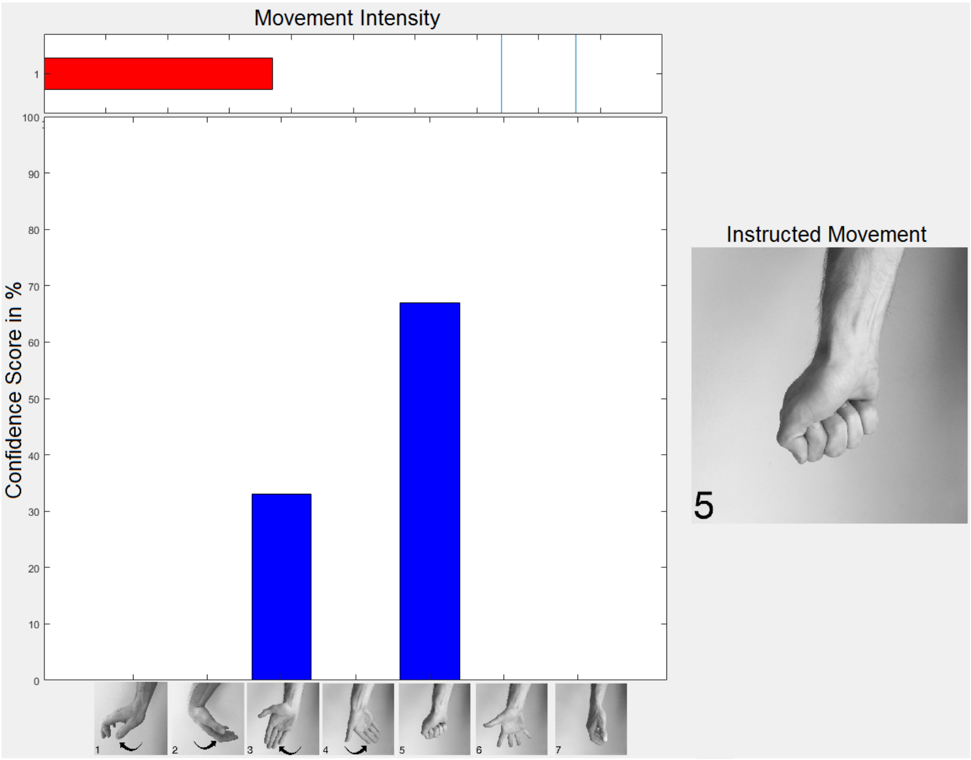


Figure A.3: Illustration of the user training interface showing the bar plot indicating the confidence score of movement recognition and horizontal bar plot indicating contraction level. The picture on the right side of the bar plot indicates which movement needs to be performed.

User Training for Control Group

The purpose of user training is for the subject to train the movements used in the performance test. During the user training the subject will train one movement at a time at different contraction levels. When training a movement, visual feedback in form of which movement the control system recognizes, is shown in a bar plot. In addition, the level of contraction is shown in a horizontal bar above the other bar plot. When performing the instructed movement at the instructed level of contraction the horizontal bar plot will appear green; if it is outside the instructed level or the control system does not recognize the performed movement, it appears red. The aim for the subject is to reach and withhold the instructed contraction level for the instructed movement while the control system recognizes it. When the subject withholds the contraction level inside the instructed contraction level for 1 seconds while the control system recognizes it the colour of the horizontal bar will appear blue, and a bell chime will appear. This indicates that the subject is performing well. After it has appeared blue, the subject must return to rest and perform the movement again and try to reach the instructed contraction level while the recognition of the control system matches the performed movement. An additional aim for the subject is to make the horizontal bar plot appear blue as many times as possible. The chronology of this procedure is as follows:

1. Perform instructed movement at 75-85 % contraction level for 30 seconds followed by 10 seconds rest.
2. Perform step 1 for the remaining movements.
3. Repeat step 1-2 at 55-65 % contraction level.
4. Repeat step 1-6 at 35-45 % contraction level.
5. Repeat step 1-6 at 15-25 % contraction level.

An illustration of the interface used for user training is shown in figure A.4.

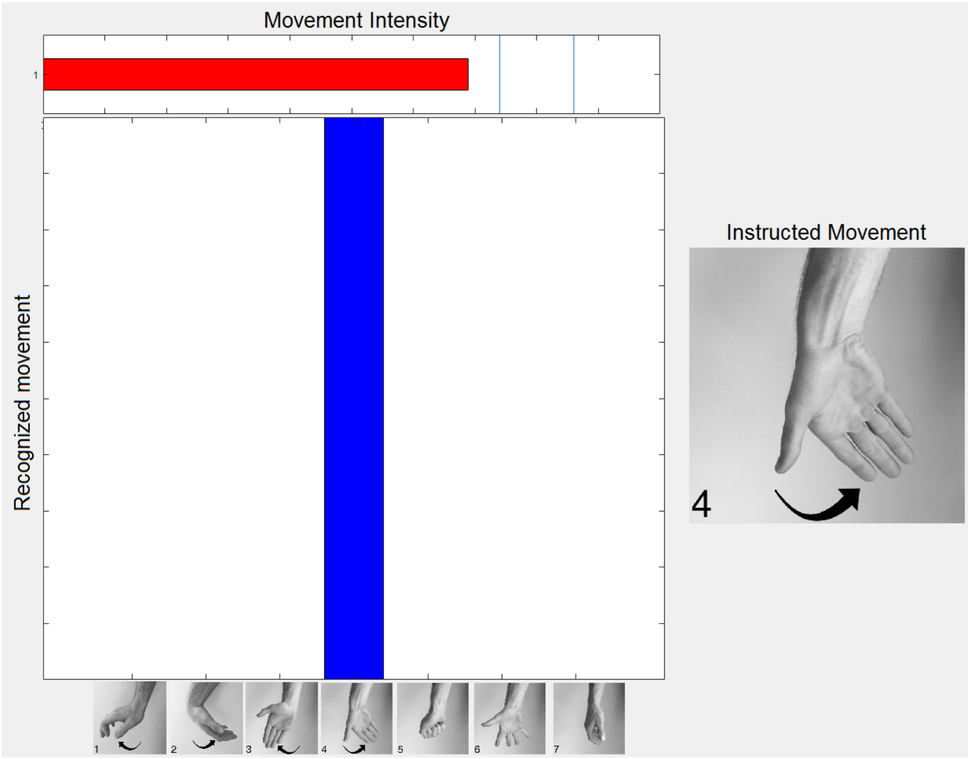


Figure A.4: Illustration of the user training interface showing the bar plot indicating which movement is being recognized and the horizontal bar plot indicating contraction level. The picture on the right side of the bar plot indicates which movement needs to be performed.

Performance test
The purpose of the performance test is to assess the subject's ability to control a prosthesis. Instead of doing a test with a real prosthesis a virtual alternative has been developed for this experiment. The prosthesis is represented as a red circular cursor with a black dot inside in a Cartesian coordinate system, which the subject can move as well as expand and shrink in size by performing the trained movements. The following bullets describe which movement corresponds to which action in the coordinate system:

- Extension moves the cursor right.
- Flexion moves the cursor left.
- Radial deviation moves the cursor up.
- Ulnar deviation moves the cursor down.
- Closed hand shrinks the cursor.
- Opened hand expands the cursor.
- Rest keeps the cursor still.

The performance test consists of a target reaching test, where the subject must reach 16 targets of different sizes and locations. A target consists of a circle with a smaller circle inside. Only one target will be visible at a time.

For the subject to reach a target and make a new appear, the subject must center the black dot of the cursor in the small circle of the target and expand/shrink the cursor to fit the size of the outer circle of the target. The cursor will appear green, when located at the correct position. The subject must dwell the cursor in a target for 1 seconds for it to be reached. When the cursor has dwelled for 1 second, it will appear blue for 1 second to indicate that the target has been reached. If a target is not reached within 15 seconds a new target will appear. When a new target appears the cursor will reset its position the origin. The aim for the subject is to reach as many target as possible as quickly as possible. The subject is only able to perform one movement at a time, as trained in the user training. Thus, no simultaneously performed movements will be recognized by the control system. The chronology of this procedure is as follows:

1. Use 2 minutes to get acquainted with the test.
2. Reach the visible target.

An illustration of the interface used for the performance test is shown in figure A.5.

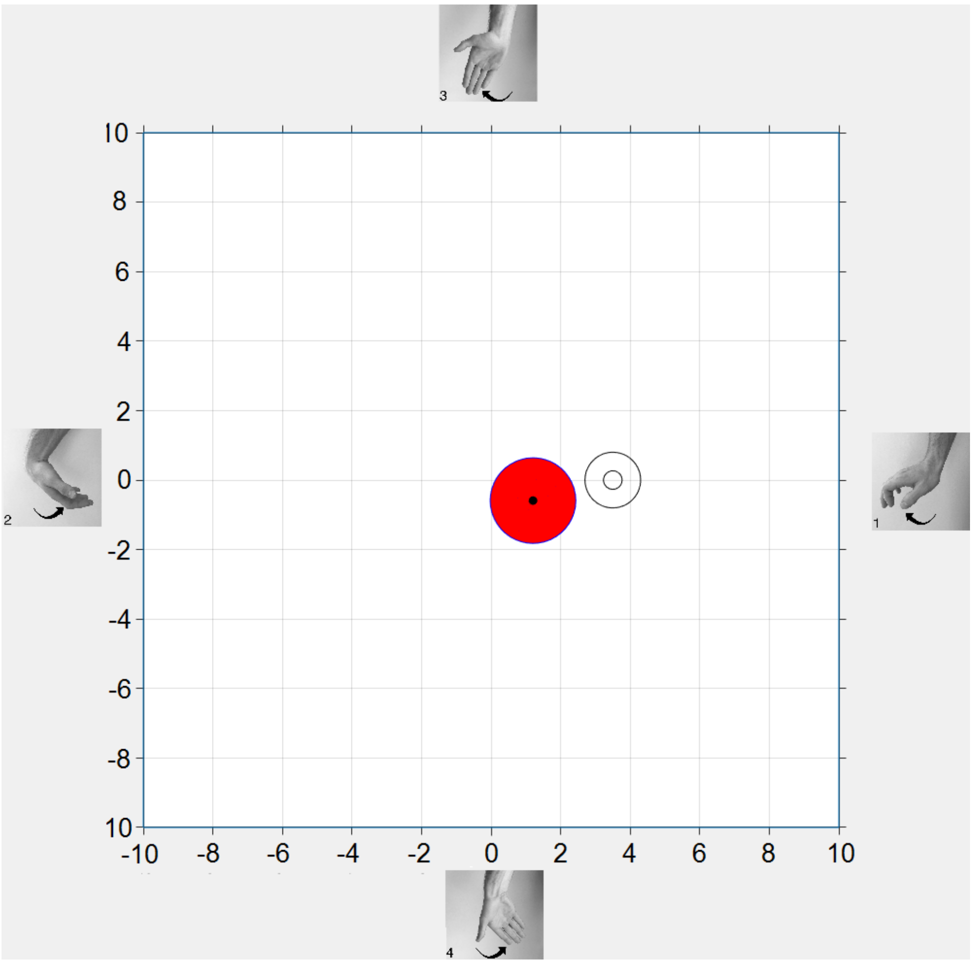


Figure A.5: Illustration of the performance test interface showing a target and the cursor representing the prosthesis output. The pictures on the axes indicate which movement must be performed to move the cursor in a certain direction.

Movements used in the experiment

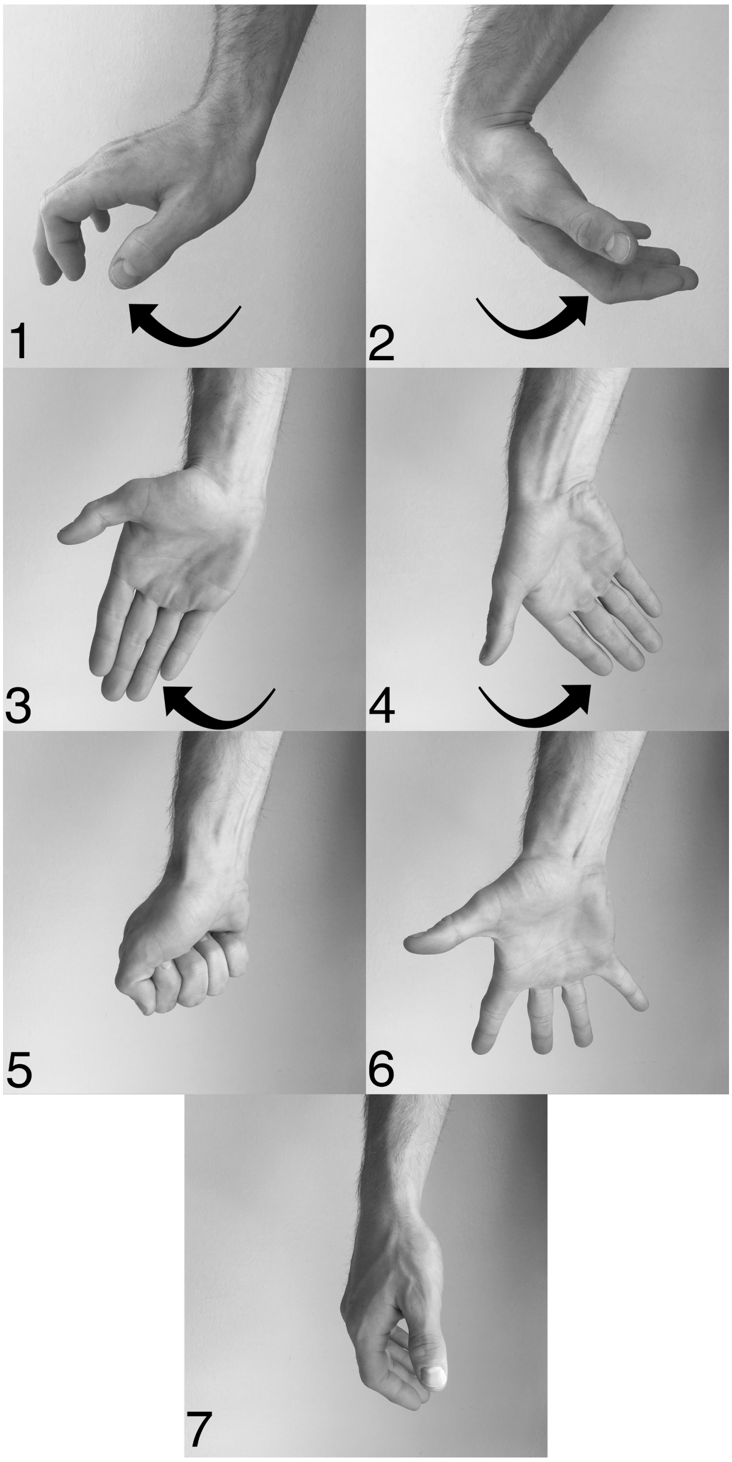


Figure A.6: Illustration of the movements used in the experiment. 1: extension, 2: flexion, 3: radial deviation, 4: ulnar deviation, 5: closed hand, 6: opened hand and 7: rest.

Experiment setup

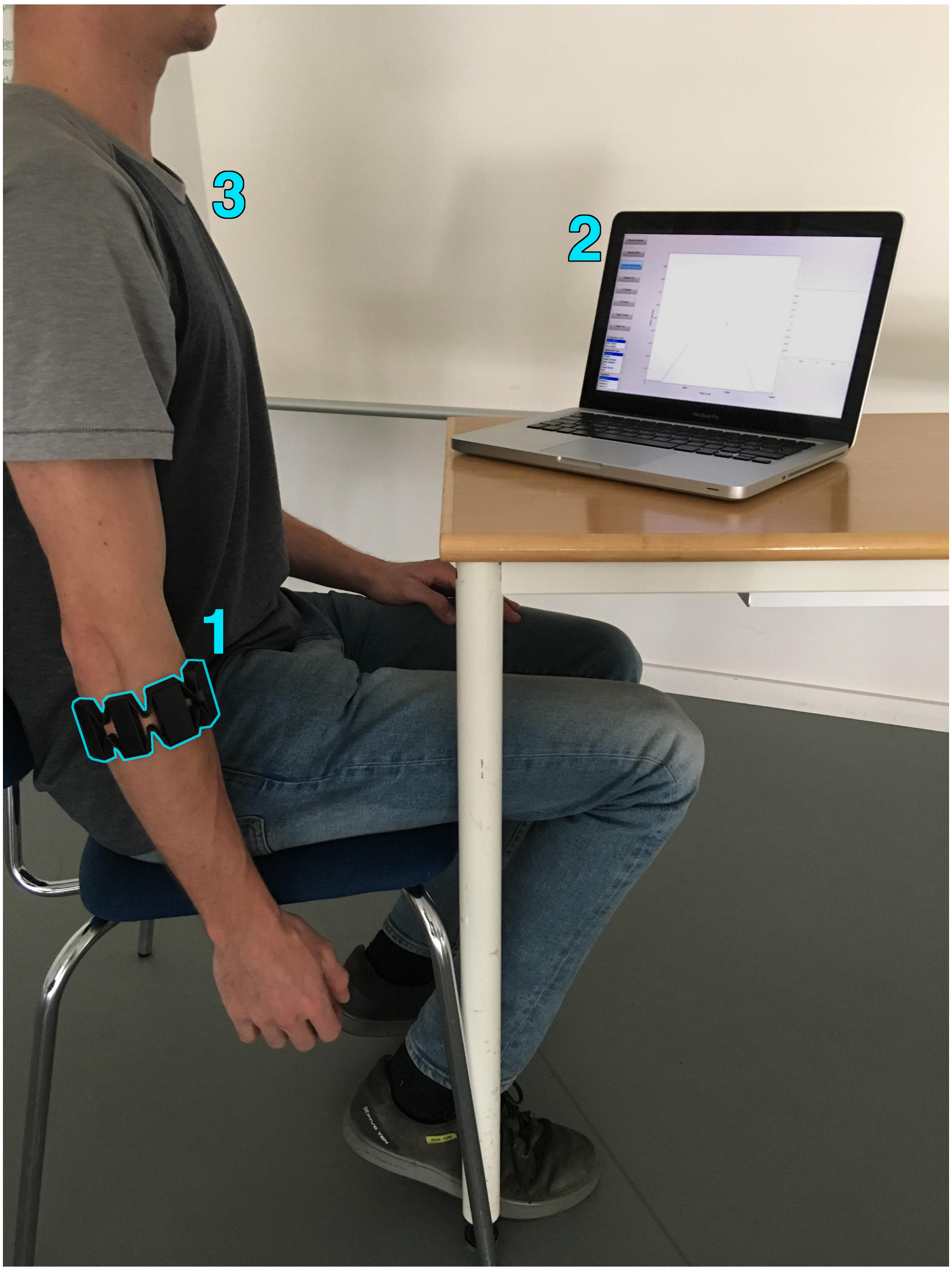


Figure A.7: Illustration of the experimental setup; 1: MYB, 2: computer with interface and 3: subject. The subject is seated facing the computer screen with the arm wearing the MYB hanging relaxed down the side of the torso.

A.2 Experiment Protocol for Investigators

Subject name:

Session number:

This protocol functions as a checklist for the investigators in the experiment "Using confidence levels of movement recognition in user training to improve prosthesis control". The checklist is used to ensure all steps in the experiment is performed correctly and that no steps will be neglected. The experiment consists of 3 session of 3-4 procedures in each session, as shown in figure A.8. The same procedures (data acquisition, user training and performance test) occur in all sessions and needs to be performed similarly each session. A checklist for each procedure is described in the sections below figure A.8.

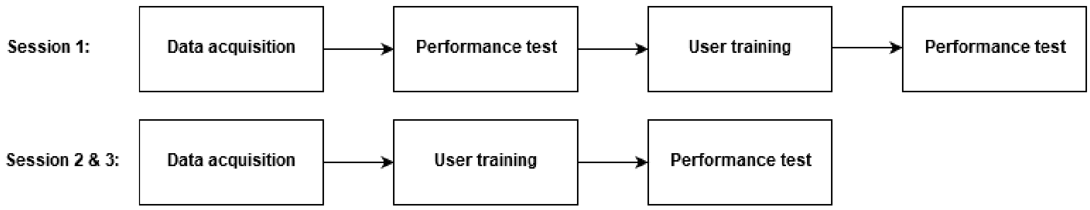


Figure A.8: Pipeline for the three sessions in the experiment and what procedures each session contains.

The instruction of the of aim the respective procedures and content and functions in the interfaces is based on the information written in the experiment protocol for test subjects. It is expected that the subject has read the experiment protocol handed out prior the experiment, but the information regarding the respective procedures is retold to verify that the subject has understood the following procedure.

Data acquisition

- ☐ Disinfect MYB with alco-swabs.
- ☐ Disinfect MYB application area of subject's dominant forearm with alco-swabs.
- ☐ Instruct subject to stand in anatomical standard position.
- ☐ Mark with a permanent marker the size of the main channel (channel with LED) of the MYB on the most lateral position of the thickest circumference of the subject's dominant forearm.
- ☐ Instruct subject in applying MYB with the main channel (channel with LED) located on the marked position. The MYB must be worn so that the LED is located as distally as possible. Add clips to tighten the MYB if necessary.
- ☐ Ensure that the main electrode-channel is placed correctly.
- ☐ Instruct subject to sit on a chair facing the screen showing the interface, with the arm wearing the MYB hanging relaxed lateral to the torso.
- ☐ Connect MYB in armband manager.
- ☐ Instruct subject in synchronizing MYB by performing extension until three distinct vibrations are felt from the MYB.
- ☐ Instruct subject in the movements about to be performed in the data acquisition.

- ☐ Instruct subject in performing an MVC; that the contraction must be steady during the 15 seconds.
- ☐ Record MVC for one movement. Observe spider plot meanwhile. If the activation pattern for the channels alters too much during the recording is to be discarded and a new must be acquired.
  - ☐ Extension
  - ☐ Flexion
  - ☐ Radial deviation
  - ☐ Ulnar deviation
  - ☐ Closed hand
  - ☐ Opened hand
- ☐ Instruct the subject in tracing the trapezoidal trajectory with the green cursor in different contraction levels of the MVC.
- ☐ Record contraction levels of MVC for one movement. Observe spider plot meanwhile. If the activation pattern for the channels alters too much during the recording is to be discarded a new must be acquired.
  - ☐ Extension: ☐ 40 %, ☐ 50 %, ☐ 70 %
  - ☐ Flexion: ☐ 40 %, ☐ 50 %, ☐ 70 %
  - ☐ Radial deviation: ☐ 40 %, ☐ 50 %, ☐ 70 %
  - ☐ Ulnar deviation: ☐ 40 %, ☐ 50 %, ☐ 70 %
  - ☐ Closed hand: ☐ 40 %, ☐ 50 %, ☐ 70 %
  - ☐ Opened hand: ☐ 40 %, ☐ 50 %, ☐ 70 %
- ☐ Build regressors for each movement and build classifier trained with all movements.

User training

- ☐ Instruct subject in aim of the user training, and explain the content and functions of the interface.
- ☐ Initiate user training.

Performance test

- ☐ Instruct subject in aim of the performance test, and explain the content and functions of the interface.
- ☐ Initiate performance test.
- ☐ Save all training data and performance measures in folder named after name of subject, session number and which experiment group the subject belongs to.

Comments: