Extraction and Classification of Multichannel Electromyographic Activation Trajectories for Hand Movement Recognition

Meena AbdelMaseeh, Tsu-Wei Chen, and Daniel W. Stashuk, *Member, IEEE*

***Abstract*—This paper proposes a system for hand movement recognition using multichannel electromyographic (EMG) signals obtained from the forearm surface. This system can be used to control prostheses or to provide inputs for a wide range of human computer interface systems. In this work, the hand movement recognition problem is formulated as a multi-class distance based classification of multi-dimensional sequences. More specifically, the extraction of multi-channel EMG activation trajectories underlying hand movements, and classifying the extracted trajectories using a metric based on multi-dimensional dynamic time warping are investigated. The developed methods are evaluated using the publicly available NINAPro database comprised of 40 different hand movements performed by 40 subjects. The average movement error rate obtained across the 40 subjects is 0.09 ± 0.047. The low error rate demonstrates the efficacy of the proposed trajectory extraction method and the discriminability of the utilized distance metric.**

***Index Terms*—EMG, hand movement recognition, multi-dimensional dynamic time warping (MD-DTW), muscle activity detection, myoelectric control.**

I. Introduction

THE HUMAN hand is so dexterous in nature that recognizing its movement poses a tremendous challenge. One approach to hand movement recognition is based on underlying electromyographic (EMG) signals acquired using surface electrodes. This approach has shown some success in controlling powered prostheses [1] and implement-free human computer interaction systems [2].

Control of powered prostheses using EMG signals, known as myoelectric control, is typically implemented using an analog circuit. A bi-directional motor is driven by differences in EMG activation-level signals (i.e., rectified and smoothed surface EMG signals) recorded from an antagonistic pair of muscles [3]. The main limitation of this approach is that the control of a two-state function (e.g., a single degree of freedom in a prosthetic device) requires acquiring EMG signals from a pair of muscles. The first attempt to overcome this limitation was to develop a three-state myoelectric controller [4], such that EMG signals from just a single muscle are used to control a two-state function by defining two activation-level signal thresholds.

Manuscript received January 08, 2015; revised March 25, 2015; accepted June 11, 2015. Date of publication June 18, 2015; date of current version June 15, 2016.

The authors are with the Systems Design Engineering, University of Waterloo, Waterloo, ON, N2L 3G1 Canada (e-mail: m2adly@uwaterloo.ca).

Digital Object Identifier 10.1109/TNSRE.2015.2447217

However to control a dexterous prosthesis or implement an intuitive human computer interaction, it is necessary to realize multifunction control in which the number of the controlled functions exceeds the number of the EMG channels [5]. This can be achieved by relying on multichannel EMG signals acquired from a set of co-activated muscles, and formulating the control problem as the multi-class classification of multidimensional sequences.

Sequence classification methods can be grouped into three main categories [6].

- Feature-based: A fixed length feature vector is extracted to represent the sequence.
- Model-based: Sequences are classified based on how well they fit an assumed model(s).
- Distance-based: A metric is defined to measure the similarity between a pair of sequences.

Three different approaches for feature-based sequence classification for myoelectric control have been used. The first approach is to represent the EMG signals corresponding to a given function using a single feature vector. An example of this approach is presented in [7]. The authors constructed a feature vector comprised of statistical moments, number of zero crossings, correlation sequences, and power spectra obtained from EMG signals recorded from two surface electrodes placed on the forearm muscles. Feature selection was used to reduce the dimensionality by selecting the most discriminative features. The second approach is presented in [8]. It relies on a feature vector describing the initial phase of a muscle contraction. A feature vector is constructed by concatenating time domain features extracted from shorter time segments within the initial phase.

The last approach of feature-based sequence classification can be referred to as continuous control [9], because a feature vector and a classification are obtained for every segment. In contrast to the previous two approaches, where a decision is only obtained when a movement is detected.

Two different approaches of model-based sequence classification for myoelectric control have been investigated. The first approach is based on a hidden Markov model [10], where each state in the model represents a different function. The second approach uses a hierarchical Bayesian model [11] as a generative probabilistic model assuming a distribution of latent intention variables inferring profiles of EMG signals.

This work investigates the usability of distance-based classification methods for hand movement recognition. Results show

the usability of distance-based classification and the possibility of performing such classifications with insignificant delay.

In this work, the possibility of segmenting the multi-channel surface EMG activation-level signals generated by muscles' contractions underlying a hand movement was investigated. These activation-level signals will be further referred to as a trajectory. Dynamic time warping (DTW) is used as a distance measure between two trajectories. DTW is chosen, because it matches sequences of different lengths and therefore there is no additional step required to transform the trajectories into an alternate representation. DTW is also indifferent to speed or non-linear distortions. Moreover, it is insensitive, to some extent, to variations in trajectories caused by inaccuracies in detecting the onset and end of a trajectory. A previous attempt to use distance-based sequence classification for myo-electric control is described in [12]. The authors normalized the data acquired from 27 intact and one amputated subjects performing 53 different movements to the same time length; and then applied nearest neighbor classification based on the normalized Euclidean distance. The reported results showed that amputees can produce stable and distinct signals for several imaginary movements.

A comprehensive survey of eight common similarity measures applied to 38 different data sets from a wide variety of application domains showed supporting evidence that DTW is the best distance measure for most domains [13]. Classification based on multidimensional DTW (MD-DTW) was shown to perform as well as or to outperform classification based on single-channel DTW [14]. In the context of hand movement recognition, DTW has been found to work quite well using video cameras [15], [14], accelerometers [16], or magnetic sensors [17].

The proposed approach was succinctly introduced by the authors in [18]. In this work, the proposed approach is evaluated using a much larger dataset comprised of 40 different movements executed by 40 subjects. The larger dataset size allows a more thorough evaluation of the proposed methods to be obtained. The rest of the paper is organized as follows: details of the proposed approach are presented in Section II; results are presented in Section III and discussed in Section IV and finally, conclusions are provided in Section V.

II. Methodology

A. System Overview

The information flow of the proposed system is shown in Fig. 1. Multi-channel EMG signals are recorded using multiple electrodes placed on the surface of the muscles expected to be involved with the hand movements of interest. Let the number of the channels to be q. The signals are then preprocessed to accentuate the electrophysiological contributions of interest relative to other signals and noise. The preprocessed signals are then rectified, smoothed and normalized to obtain a multi-dimensional activation-level signal $A(t) = \langle a_1(t), \ldots, a_i(t), \ldots, a_q(t) \rangle$.

Each element in $A(t)$ is expected to be correlated to the level of contraction of a group of muscles that are relatively close to the detection surface of the corresponding electrode. The ith channel is assumed to be active at a given sample t, if the activation-level $a_i(t)$ is significantly different from the background activity. Still, a movement is only assumed to have occurred when a number of channels are detected to be simultaneously active for a continuous period of time.

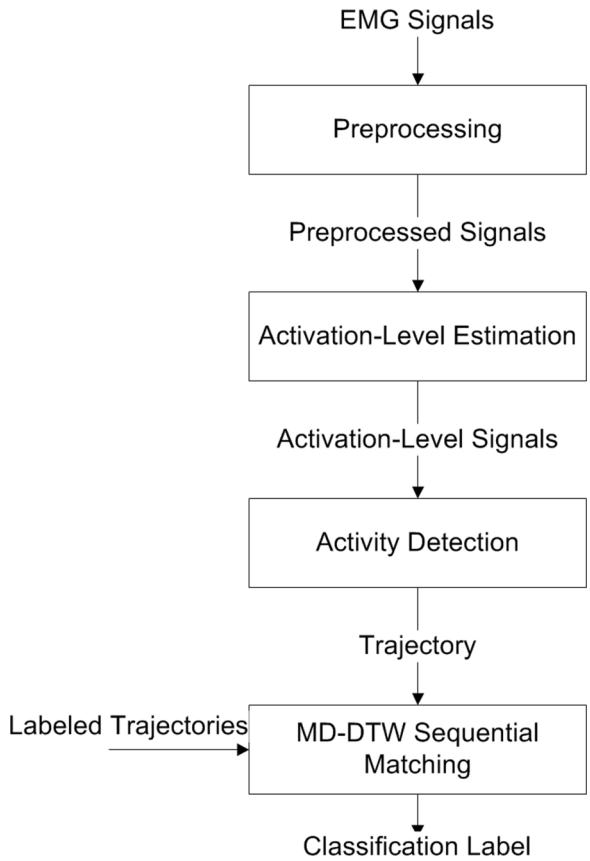


Fig. 1. Information flow of the proposed system.

Trajectory X is defined as the activation-level signals obtained from the q channels between the detected onset sample t_0 and the end sample t_f after being smoothed and decimated. For an application with m distinct hand movements, a training set $R = \{(X_1, y_1), \ldots, (X_l, y_l), \ldots, (X_n, y_n)\}$ is comprised of n labeled trajectories, where $y_l \in \{c_1, \ldots, c_m\}$ is the label of trajectory X_l. In order to classify an unlabeled trajectory, say X_k, the distance $\Psi(X_l, X_k)$ based on MD-DTW is calculated between X_l and each of n labeled trajectories. The label y_k is set to the label of the trajectory found to be closest to X_k.

B. Preprocessing

A high-pass Butterworth filter with corner frequency at 20 Hz and a 12 dB/octave roll-off is utilized for preprocessing. It was shown in [19] that using a filter with this configuration results in the best trade off between preserving EMG content and eliminating artifacts due to electrode movement. Also, high-pass filtering is expected to reduce the crosstalk [20] by attenuating contributions from muscles distant from the electrode.

C. Activation-Level Estimation

An activation-level signal (i.e., a signal correlated to the level of activation of a muscle) can be calculated using a surface

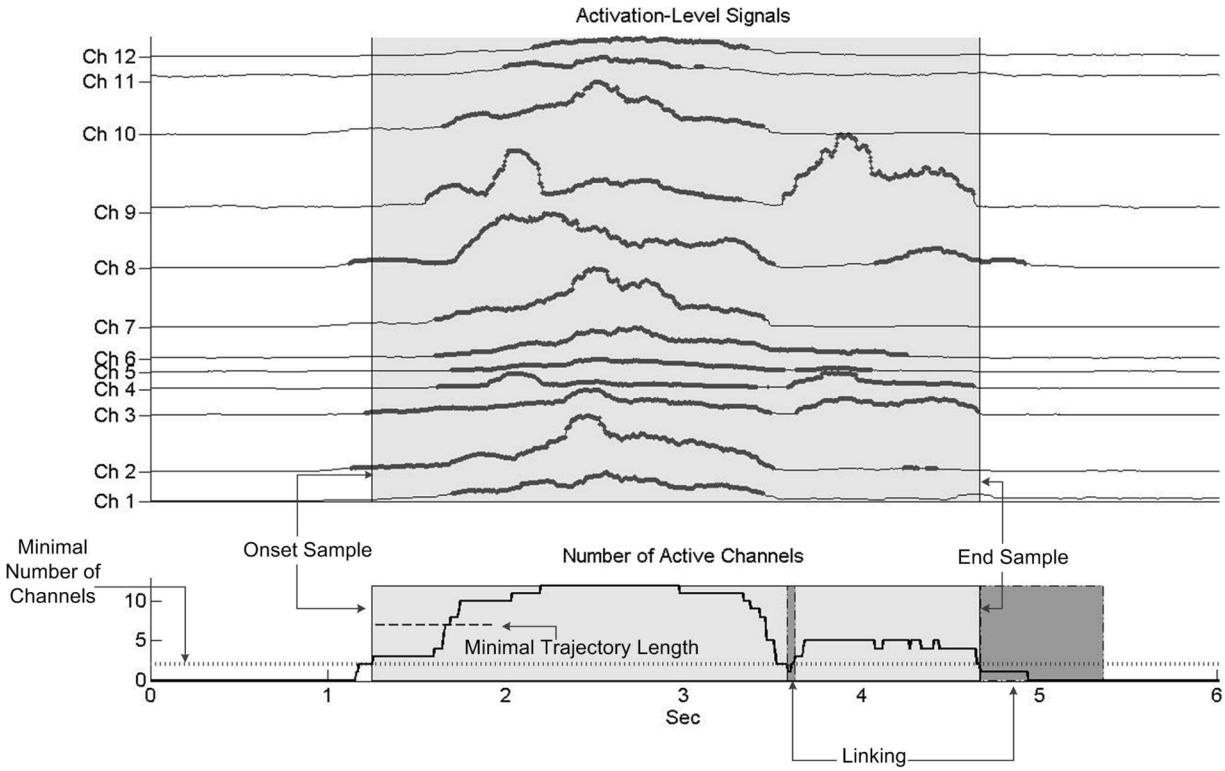


Fig. 2. Example showing the activity detection stage. The upper diagram shows the activation-level signals obtained from 12 channels. The shaded area indicates the extracted trajectory, i.e., the activation-level signals between the detected onset sample and the end sample. Note that samples that are assumed to be active are highlighted. The lower plot shows the number of active channels $u(t)$. The onset sample is detected once $u(t)$ exceeds the minimal number of channels α. $u(t)$ stays above α for a period more than the minimal trajectory length τ_{active}. At approximately 3.5 s, $u(t)$ goes below α for a period shaded by darker gray, however the end of the trajectory is not detected since $u(t)$ does not stay below α for the linking window length τ_{inactive} and the following burst of activity is linked to the trajectory. Finally, $u(t)$ goes below α and stays below it for a period τ_{inactive}, and the trajectory end sample is detected.

detected EMG signal recorded from the active muscle using the general schema of decoding, smoothing and relinearization [21]. In this work, an activation-level signal is calculated as the root mean square (rms) value of the preprocessed EMG signal, which is equivalent to decoding using squaring, smoothing by averaging over a moving window of length 64 samples (192 ms), and relinearization by taking the square root. Activation-level signals were shown to be correlated to the level of muscle contraction [22], since they are related to how many motor units are recruited and their mean firing rates.

The characteristics of the background activity are then captured by calculating the mean $\mu_{\text{rest}} = \langle \mu_{\text{rest}}(1), \ldots, \mu_{\text{rest}}(q) \rangle$ and the standard deviation $\sigma_{\text{rest}} = \langle \sigma_{\text{rest}}(1), \ldots, \sigma_{\text{rest}}(q) \rangle$ of the rms signals of training examples labeled as "rest." In order to obtain a normalized activation level-signal, the rms signal of each channel is divided by its corresponding $\mu_{\text{rest}}(i)$.

D. Activity Detection

The sample $a_i(t)$ of an activation-level signal is assumed to be "active," if $a_i(t) \geq \gamma_i \sigma_{\text{rest}}(i)/\mu_{\text{rest}}(i)$. As will be explained later, the sample activity factor γ is induced using a training set R so as to minimize confusion between examples labeled as rest and other examples.

Extraction of a trajectory (i.e., detection of its onset t_0 and end t_f) is challenging because:

- The detection of a few isolated active samples does not necessarily indicate the execution of a movement, since this activity is more likely to be due to noise or electrode movement.
- The activation trajectory of a multi-muscle activation sequence can be multi-phasic, meaning that it may include more than one burst of activity separated by rest. Therefore, the detection of rest following an activity may not be suitable for the detection of t_f

As the sample t arrives, the number of active channels $u(t)$ is estimated using a window spanning from $t-\tau$ to t. A channel is considered to be active and consequently $u(t)$ is incremented, when the ratio of active samples in the window is above the threshold δ.

As illustrated in Fig. 2, Four events need to occur in the following order for a trajectory to be detected.

1) $u(t)$ exceeds a preset minimal number of channels α.
2) $u(t)$ stays above α for τ_{active} samples, where τ_{active} represents the minimal allowable trajectory length.
3) $u(t)$ goes below α.
4) $u(t)$ stays below α for τ_{inactive} samples. τ_{inactive} is referred to as the linking window, because during this interval the algorithm is searching for later phases (bursts of activity) of the trajectory that should be linked with the earlier phases.

The sample activity factor γ_i is set for each channel to an integer between 1 and 15 that minimizes the confusion in a training set defined as the sum of 1) the number of times that a trajectory is detected during rest, 2) the number of times that

no trajectory is detected in a training set repetition that is not labeled as rest, 3) the number of times more than one trajectory is extracted from a training set repetition that is not labeled as rest.

In the first 15 iterations of the algorithm, the search starts by looking for an initial values of $\gamma_i = \bar{\gamma}$ assuming it is the same for all channels. For each of the searched values between 1 and 15, the trajectory extraction technique is applied and the confusion is calculated. $\bar{\gamma}$ is set to the middle of the range over which the confusion is minimal. The search then proceeds in a greedy manner optimizing the sample activity factor γ_i for each of the channels assuming the factors of other channels to be fixed.

Each extracted trajectory is decimated by the factor β. The decimation is preceded by low-pass filtration to diminish frequency components exceeding $F_s/2 \times \beta$ yielding a trajectory indexed by $\bar{t}$ from $\bar{t_0}$ to $\bar{t_f}$.

E. Matching Using Multidimensional Dynamic Time Warping

MD-DTW is used to align two trajectories, say X_k and X_l, diminishing the effects of different execution speeds. Let the number of decimated samples in X_k and X_l to be λ_k and λ_l, respectively. Also, let the alignment between two arbitrary samples from the two trajectories, say $X_k(:, \bar{t_{a_j}})$ and $X_l(:, \bar{t_{b_j}})$, be denoted as $\langle a_j, b_j \rangle$ and the distance associated with this alignment be the multi-dimensional Euclidean distance between the two samples $d(\langle a_j, b_j \rangle)$

$$d(\langle a_j, b_j \rangle) = \sqrt{\sum_{i=1}^{q} (X_k(i, \bar{t_{a_j}}) - X_l(i, \bar{t_{b_j}}))^2}. \quad (1)$$

MD-DTW searches for a path (i.e., a sequence of alignments between pairs of samples $P = \langle P(1) = \langle a_1, b_1 \rangle, \ldots, P(j) = \langle a_j, b_j \rangle \ldots P(|P|) = \langle a_{|P|}, b_{|P|} \rangle \rangle$), that minimizes the accumulated distance defined as

$$\rho(\langle a_{|P|}, b_{|P|} \rangle) = \sum_{j=1}^{|P|} d(\langle a_j, b_j \rangle) \quad (2)$$

where $|P|$ is the length of the path. The path is subject to the following conditions.

- Boundary: $P(1) = \langle 1, 1 \rangle$ and $P(|P|) = \langle \lambda_k, \lambda_l \rangle$.
- Monotonicity: If $P(j) = \langle a_j, b_j \rangle$ and $P(j+1) = \langle a_{j+1}, b_{j+1} \rangle$, then $a_{j+1} \geq a_j$ and $b_{j+1} \geq b_j \; \forall j$.
- Step size: If $P(j) = \langle a_j, b_j \rangle$ and $P(j+1) = \langle a_{j+1}, b_{j+1} \rangle$, then $a_{j+1} - a_j \leq 1$ and $b_{j+1} - b_j \leq 1 \; \forall j$.

The search for the path P can be obtained using dynamic programming. Let $P_{\langle a_j, b_j \rangle}$ be the optimal path between $\langle X_k(:, 1) \ldots X_k(:, \bar{t_{a_j}}) \rangle$ and $\langle X_l(:, 1) \ldots X_l(:, \bar{t_{b_j}}) \rangle$ that minimizes the accumulated distance $\rho(\langle a_j, b_j \rangle)$. The path can be obtained using the following recursive formula:

$$\rho(\langle a_j, b_j \rangle) = d\langle a_j, b_j \rangle + \min(\rho(\langle a_j - 1, b_j - 1 \rangle), \rho(\langle a_j, b_j - 1 \rangle), \rho(\langle a_j - 1, b_j \rangle)). \quad (3)$$

In this work, the distance between X_k and X_l was defined as $\psi(X_k, X_l) = \rho(\langle \lambda_k, \lambda_l \rangle)$. The distances between an unlabeled trajectory and all labeled trajectories belonging to R are calculated sequentially. The label of the unlabeled trajectory is then assigned to the label of the closest labelled trajectory.

III. Evaluation

A. Data Acquisition

In this study, the second version of the publicly available database from the Non-Invasive Adaptive Prosthetics (NINAPro) project was utilized [23]. The reliance on a publicly available database facilitates the comparison of the proposed methods to existing and future methods. In order to further enable such comparison, the code for the proposed methods is made available from https://github.com/meena-abdelmaseeh/EMG-Movement-Recognition.

The second version of the NINAPro database consists of surface EMG signals, kinematic, and force measurement signals acquired from 40 different subjects, compared to 27 subjects in the first version. However, this work only used the surface EMG signals. The EMG signals were acquired using 12 surface EMG electrodes. Eight of the electrodes were placed equidistantly to each other around the forearm. Electrodes 9 and 10 were placed on the extensor digitorum communis and the flexor digitorum superficialis muscles respectively, while electrodes 11 and 12 were placed on the biceps and triceps muscles respectively. The data was sampled at 2 KHz [23], however in this work the data was initially sub-sampled by a decimation factor of 6.

A movie displayed on a computer screen was used as a visual stimulus. Subjects were asked to repeat the movements contemporarily with the movie. Following a training stage, subjects needed to perform six consecutive repetitions of 40 different movements. Following the same protocol as the NINAPro database authors, the second and the fifth repetitions of each movement were used for testing, while the other four repetitions were used for training. In this work, parameters for trajectory extraction were tuned for each subject independently and also test trajectories were only matched to trajectories extracted from the same subject.

B. Evaluation of the Trajectory Classification Based on Multi-Dimensional Dynamic Time Warping

The main objective of the experiments described in this section was to evaluate the classification based on MD-DTW independently, i.e., assuming that the onset and end of the activation trajectories were known and perfectly labeled [12]. The assumption that we can obtain perfect labeling is unrealistic, because it is not expected that the subjects will be able to follow the video stimulus accurately and consistently. The authors of the NINAPro database therefore performed an offline relabelling of the data to constrain the trajectories to periods with increased EMG activity. They searched for an onset and an end position of the trajectory in a window starting from the video stimulus onset and ending one second after the video stimulus end by optimizing for the likelihood of the data points being sampled from two distributions modelling both rest and activation [24]. The trajectories obtained from offline relabelling were assumed in this work to be the most representative trajectories that can

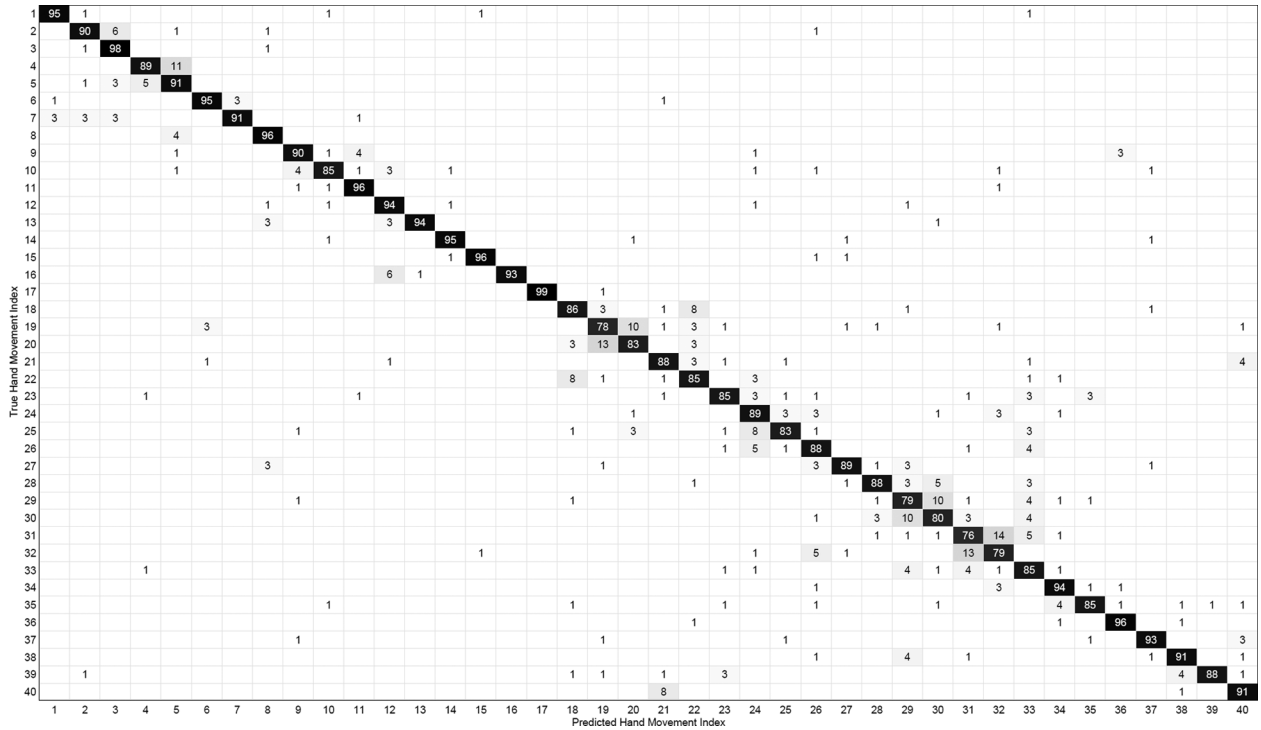


Fig. 3. Confusion matrix for trajectories obtained by offline relabelling.

be obtained, and therefore accuracies obtained from classification of these trajectories can be thought of as an upper bound for the accuracies obtained using trajectory extraction methods that do not rely on knowing the onset and the end of the video stimulus. This allowed to some extent a qualitative decoupling of the sources of the errors to be either due to trajectory extraction or classification.

Fig. 3 shows the confusion matrix for trajectories obtained from offline relabelling as classified to the nearest neighbor based on MD-DTW as a matching distance. Each row represents the testing repetitions extracted from all subjects for a given hand movement. Each row is normalized, such that its sum is 100%. From the matrix, it can be seen that the percent of correctly classified trajectories is 89%, which supports the discrimination of the proposed matching distance. The second clear observation is that there is a high variability in classification accuracy among different hand movements. While nine movements attained a classification accuracy of 95% or more, the accuracy of five movements ranged between 76% and 80%.

Fig. 3 also shows that most of the confusion occurs with movements that have indexes close to the index of the true movement. For example, all the erroneously classified trials from movement 4 were confused as movement 5. A similar observation can be seen in movements 19, 29, 30, 31, and 32. This is because the consecutive hand movements in the acquisition protocol were very similar. For instance movements 29 and 30 are precision sphere grasp and tripod grasp respectively, while movements 31 and 32 are prismatic grasp and pinch grasp respectively. It is worth noting as well that the average classification accuracy of the first 17 movements was 93.3% compared to an average accuracy of 85.9% for the remaining 23 movements. This is expected, because the first 17 movements are simple hand and wrist movements, while the other 23 are more complicated, dynamic, multi-staged grasping and functional movements.

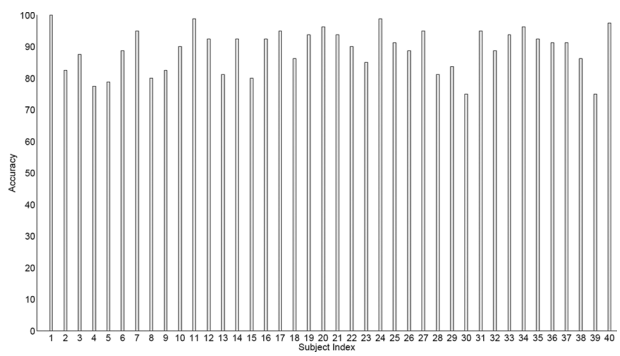


Fig. 4. Classification accuracies obtained for different subjects based on trajectories obtained from offline relabelling.

Fig. 4 shows the classification accuracies obtained for each of the subjects. It is clear there is a large variability among different subjects, while six subjects attained a performance higher than or equal to 95%, four subjects were below or equal to 80%. The variability in the obtained accuracies across different movements and different subjects suggests that an increased performance can be obtained by selecting a subset of movements specific to each of the subjects.

A greedy search was implemented on the trajectories as obtained from offline relabelling. The search discarded the movement that resulted into the lowest classification accuracy in each iteration. Since the search was applied independently to the trajectories of each subject, it was expected to select a different subset of hand movements for each subject. It can be seen from

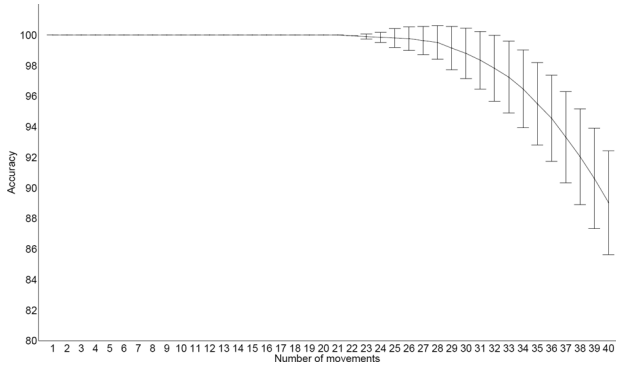


Fig. 5. Classification accuracy averaged across the 40 subjects for different number of hand movements as selected by the greedy search algorithm. The bars represent one standard deviation away from the obtained average accuracies. The results are obtained based on trajectories obtained from offline relabelling.

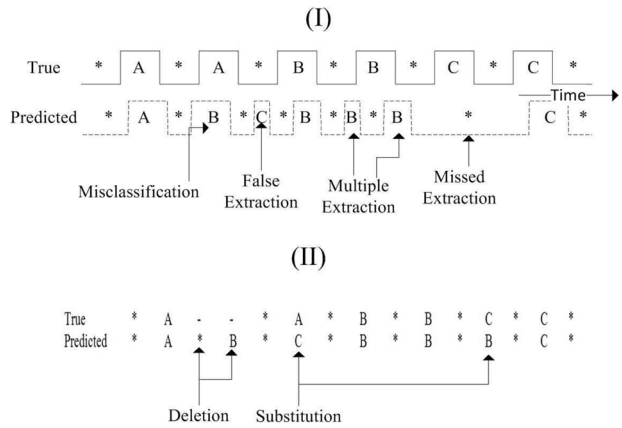


Fig. 6. An illustration of how MER counts for different types of error.(I) It is assumed that a subject is executing two repetitions of three different hand movements (A, B, C) interleaved with rest periods (*). The first timing represents the true sequence, while the second timing represents a possible output of a trajectory extraction and classification. (II) The editing done to match the two sequences based on the dynamic programming alignment used for estimation of Levenshtien distance. Since four editings (two deletions "-" relative the predicted sequence and two substitutions) are necessary, the MER is equal to $4/13$, where 13 represents the total length of the true sequence.

Fig. 5 that an accuracy of 100% can be obtained for all subjects, if the number of selected movements is limited to 21.

In these subsets of 21 hand movements for the 40 subjects, 43% of the selected movements belonged to the 17 hand and wrist movements exercise and 57% belonged to the 23 functional and grasp movements. This indicates that the search didn't favour either of the categories, even though the first category was shown to be more discriminant in Fig. 3, when considering all the movements. Instead, it relied on diversifying among different categories of movements to obtain further discrimination.

It was also observed that for 30 subjects, an accuracy of 100% was obtained for a subset of hand movements comprised of 30 movements. This suggests that in a real life myocontrol system, the number of movements that can be detected by the classifier should be tuned for each subject independently. The implemented greedy search is very far from being considered optimal for the task of movement selection, a more exhaustive search or clustering scheme based on MD-DTW as a distance can result into larger subsets of movements attaining better accuracies. The idea of simplifying the task by decreasing the number of movements to increase the classification accuracy was previously examined in [12].

C. Evaluation of Trajectory Extraction and Classification

In this section, the performance of the whole system is evaluated, i.e., the proposed activation level estimation, trajectory extraction and classification. In order to evaluate the overall performance, we need an evaluation metric that considers four different types of errors simultaneously: 1) misclassification: an extracted trajectory is classified as a wrong hand movement; 2) false extraction: one or more trajectories are extracted during the rest period; 3) multiple Trajectories Extraction: more than one trajectory is extracted during the activation corresponding to a single movement; 4) missed extraction: no trajectory is extracted during a hand movement.

It was also required that the proposed evaluation metric to be insensitive to time delays. The first reason is that it is of no consequence whether the onset of the extracted trajectory precede or succeed the onset of the labelled trajectory, because no classification is to be obtained until the end of the trajectory. Second, while delays in the detection of the end of the trajectory are important, still we, similar to what is proposed in [25], see the aforementioned four types of error and delay as competing criteria, since allowing further delay can lead to improved accuracy as shown later. This makes the traditional window-based classification accuracies [9] unsuitable for evaluation.

A slight variation of the movement error rate (MER) proposed in [26] is utilized in this section. MER is basically a normalized version of an edit distance, which counts how many insertions, deletions and substitutions are necessary for two sequences to match. In this case, the first sequence is comprised of labels indicating different movements and rest periods, while the second sequence is comprised of the extracted trajectories and their classifications. Fig. 6 illustrates, using an example, how MER accounts for the previously mentioned four types of error.

Fig. 7 shows the confusion matrix for trajectories extracted using the proposed trajectory extraction methodology. Due to missed extractions, multiple trajectory extractions and false extractions errors, there was no one-to-one mapping between the true trajectories (obtained by offline relabelling) and the extracted trajectories. A heuristic was therefore employed to establish such mapping based on the alignment resulting from the dynamic programming algorithm used for estimation of the Levenshtein distance. An example of such alignment is shown in the second part of Fig. 6. A movement span was defined relative to the aligned true sequence to include all the repetitions of a given movement, deletions, insertions, and the interleaving rests. For instance, the span of movement "A" in Fig. 6 starts at the first element and ends at the sixth. The true label of any trajectory extracted within such a span was assumed to be the same as the movement span. For instance in span of the movement "A", one extracted trajectory was matched correctly to the movement "A", while the other two were confused as movements "B" and "C".

As shown in the confusion matrix in Fig. 7, the obtained average accuracy is 86.7%, which is close to the 89% obtained for

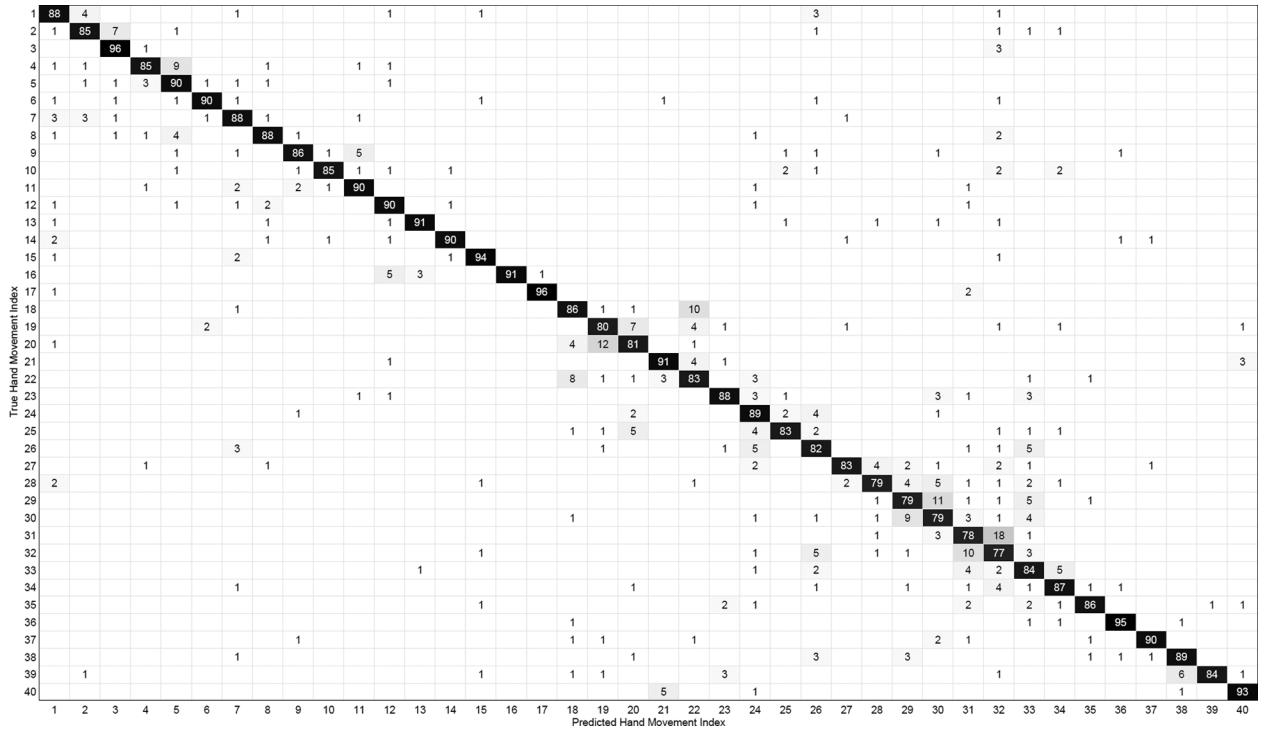


Fig. 7. Confusion matrix for trajectories extracted using the proposed trajectory extraction.

the case of trajectories as labelled using the offline relabelling, suggesting the effectiveness of the proposed trajectory extraction methods. However, one should be careful comparing these two numbers directly given the different methods used for estimation of the two accuracies. There are also other similarities between the two confusion matrices including the following. 1) High variability in accuracy among different hand movements. While 12 movements achieved an accuracy equal to or higher than 90%, the accuracy of the other five ranged between 77. 2% and 80%. 2) Most of the confusion happened with movements closely indexed to the true movement, for example, see movements 2, 4, 19, 20, 29, 30, 31, and 32. 3) There was a clear difference in average accuracy obtained for the first 17 hand and wrist movements (achieving an average accuracy of 89.7%) and the accuracies obtained for the remaining 23 functional and grasp movements (achieving an average accuracy of 85%).

Fig. 8 shows that the average MER obtained for the 40 subjects was 0.09 ± 0.047. Also, similar to the accuracies obtained for using trajectories obtained from offline relabelling, there is variability in the MER obtained for different subjects. The results show that the MER for 11 subjects was below 0.05, while for the other five subjects it ranged between 0.15 and 0.18. The results correlate well to the accuracies based on trajectories obtained from offline relabelling, the obtained Pearson correlation coefficient between these accuracies and $100 \times (1 - MER)$ was 0.68 for a linear model with a slope of 0.99. This variability in the performance obtained for different subjects and different hand movements motivates the selection of a subset of the hand movements specific to each subject.

A greedy search was implemented similar to the one used for the trajectories from offline relabelling. In each iteration of this search, the repetitions and the interleaving rest periods of

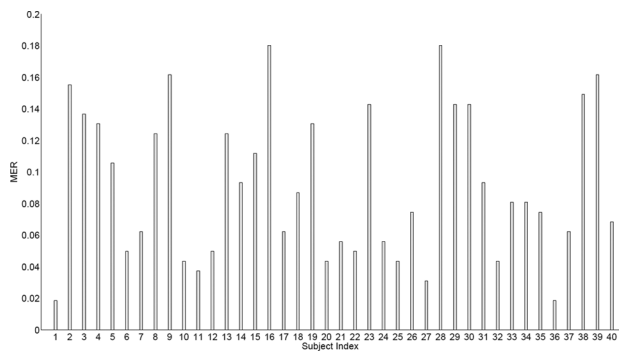


Fig. 8. Average MER obtained for different subjects based on trajectories extracted using the proposed trajectory extraction.

the movements that caused the most confusion were removed from the concatenated signal. The algorithm terminated when there was no remaining movements or if the MER increased in three following consecutive iterations. Trajectory extraction and classification were reapplied to the new constructed signal, however to speed up the evaluation, the sample activity factor δ was assumed to not change across different iterations.

As shown in Fig. 9, the average MER decreased from 0.09 to 0.03 after 20 iterations of the algorithm, meaning that each subject can perform at least 20 different movements with such a MER value. The standard deviation of the obtained accuracies across subjects remained constant at approximately 0.047 for different iterations. This could be explained to some extent by the fact that after the twentieth iteration, 26 subjects attained and remained at zero MER, while the algorithm terminated for the remaining 14 subjects due to deterioration in the obtained MER. There is still a space for improvement through

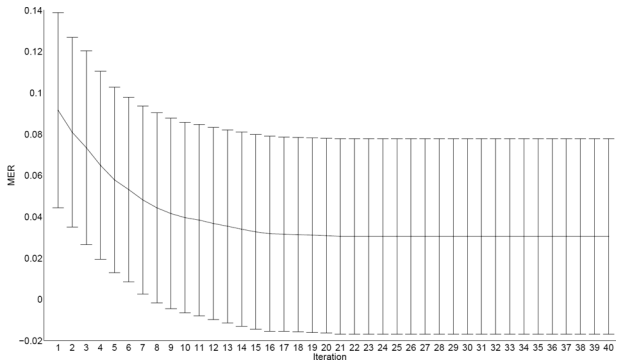


Fig. 9. MER value averaged across the 40 subjects obtained at different iterations of the greedy search algorithm used to select a subset of movements specific to each subject. Bars represent one standard deviation away from the obtained average values.

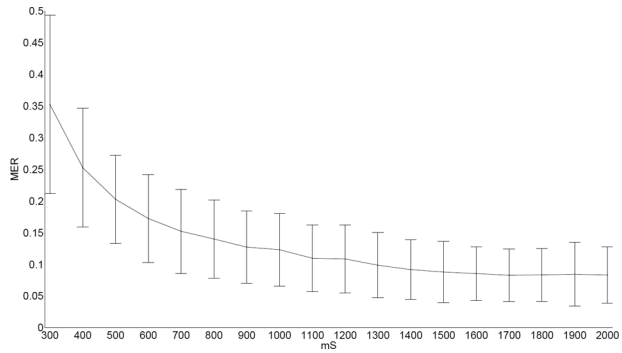


Fig. 10. Average MER obtained across subjects for different values of system lag, which is defined as the sum of the minimal trajectory length and the linking window length. Two lengths were assumed to be equal in this experiment. Bars represent one standard deviation away from the obtained average values.

using more exhaustive search algorithms that consider combinations of movements instead of a single movement in each iteration. While being more computationally expensive, such algorithms are expected to select larger subsets of movements attaining lower MER values and they would only need to be executed once during the fitting phase of the controller.

The tuning of the trajectory extraction algorithm parameters are expected to influence the system performance in terms of accuracy and responsiveness. Such tuning is to be optimal to be performed for all of the parameters simultaneously and independently for each subject based on metrics capturing the system's accuracy and responsiveness. However given the large size of the NINAPro database, it was decided to perform such tuning for some of the parameters, specifically the minimal trajectory length τ_{active} and the linking window length τ_{inactive}, using the average MER obtained across all subjects. We assumed these two parameters to have the same value and defined their sum to be the system lag. This naming of the sum as "system lag" is because the minimal trajectory length represents how quickly a subject can perform a movement to be detected by the controller, while the linking window length reflects the time the user needs to wait before executing the next movement.

As shown in Fig. 10, there is no significant improvement in the obtained average MER value for system lags above 1400 ms. Therefore, this system lag was used to obtain the results shown in Figs. 7–9. The trajectory extraction delay defined as the time difference between the trajectory end as estimated by the proposed trajectory extraction and as estimated by the offline relabelling approach was found to be -262 ± 76.7 ms. The negative sign here indicates that on average the detected trajectory end was earlier than the labelled one.

The other source of delay in the proposed method is the matching delay, i.e., the time needed to sequentially compute the MD-DTW distance between a test trajectory and all of the labelled trajectories. Since the MD-DTW has an $O(N^2)$ computational complexity, where N is the number of samples. The matching delay is expected to vary significantly for different decimation factors. Fig. 11 shows the average matching delay and the accuracy for different decimation factors for subject 1 (one of the best performing subjects in the database) and subject 30 (one of the worst performing subjects in the database). The

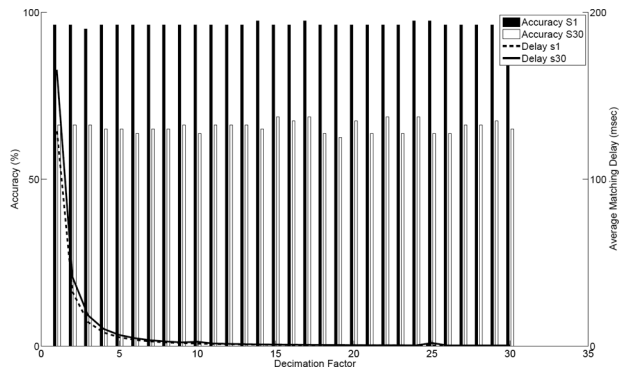


Fig. 11. Accuracy and the average matching delay obtained for different decimation factors for subject 1 (s1) and subject 30 (s30) using trajectories obtained based on the offline relabelling.

matched trajectories were obtained based on offline relabelling method. The MD-DTW was implemented in C and run on a computer having an i7–3820 processor and 32 GB of RAM. As expected, increasing the decimation resulted in a significant decrease of the matching delay for both subjects. It decreased for subject one from 165.7 ms in the case of using no decimation to 2.7 ms in the case of using a decimation factor of 10.

It was also interesting to see in Fig. 11 that increased decimation did not result in a deterioration of the performance. On the contrary, the accuracy increased slightly for some higher decimation factors in both subjects. This could be explained by the fact that the decimation was preceded by low-pass filtering, which could have resulted in the discarding of indiscriminate high frequency details from the trajectories. This motivates using a larger number of labelled examples, because the matching delay is insignificant compared to the linking window length. An increase in the number of examples is expected to increase accuracy and make the implementation of more sophisticated learning techniques feasible.

IV. Discussion

The results presented in the previous section show the ability of the MD-DTW accumulated distance to discriminate between the trajectories of different hand movements performed by the same subject. This discrimination can be attributed to the following reasons.

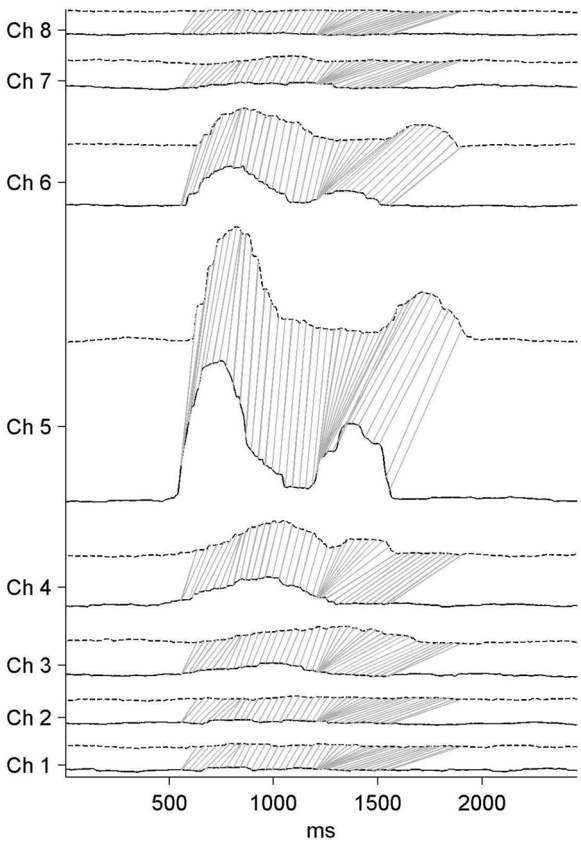


Fig. 12. Example showing how dynamic time warping accommodate to variations in the speeds of hand movement execution. Two trajectories representing the same movement as executed by the same subject are interleaved together. One trajectory is represented with dotted lines, while the other trajectory is represented with solid line. The gray lines represent every fifteenth alignment from one time instance of a trajectory to a time instance in the other trajectory. In particular, in channel 5, it can be observed that the dashed trajectory is corresponding to a slower movement. DTW is shown to be able to account for this by aligning many time instances from the trajectory corresponding to the slower movement to a single time instance in the other trajectory [18].

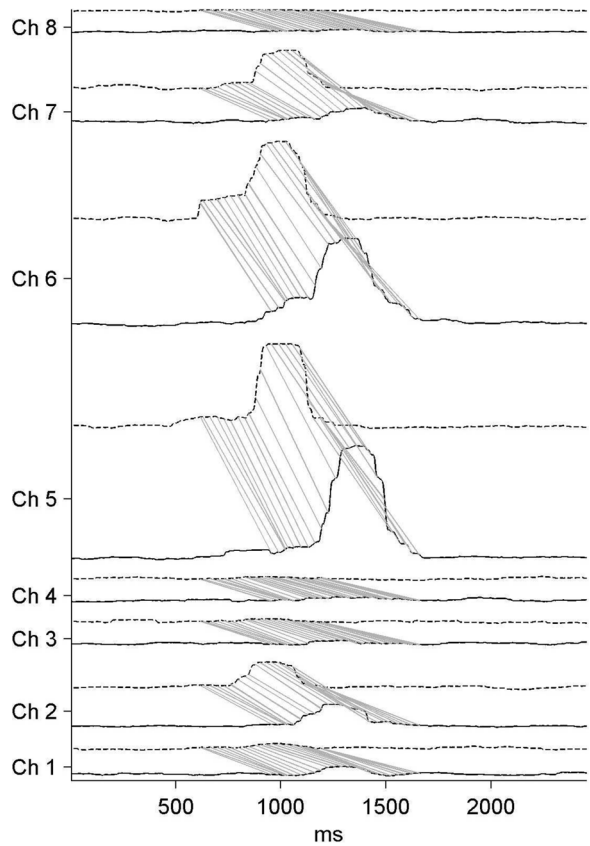


Fig. 13. Example showing the capability of MD-DTW to ignore spatially confined difference in trajectories. Two trajectories representing the same movement as executed by the same subject are interleaved together. One trajectory is represented with dotted lines, while the other trajectory is represented with the solid lines. The gray lines represent every fifteenth alignment from one time instance of a trajectory to a time instance in the other trajectory. In channel 7, the trajectories are clearly different, however the alignments are dominated by similarities in the other 7 channels.

1) MD-DTW is capable of accurately aligning trajectories of the same hand movement executed at different speeds. An example is shown in Fig. 12.
2) The reliance on multi-dimensional activation level signals and MD-DTW distance allows capturing the synergistic behaviour of co-activated muscles during the execution of a given movement.
3) The classification is based on activation-level signals that capture all the EMG changes underlying the movement. In contrast to the continuous control scheme [9] which is based on feature vectors describing a time limited window. This gives our classification an advantage in cases where the movement is multi-phasic or where more than one movement have segments with similar kinematics.
4) Activation-level signals disregard nondiscriminative aspects of EMG signals, such as the morphology of constituent motor unit potentials. This results in smoother and more discriminative trajectories.
5) MD-DTW is tolerant to temporally limited and spatially localized variations in trajectories as shown in Fig. 13, because the sequence of alignments is selected to optimize the accumulated distance over the complete lengths of the two matched trajectories for all channels.
6) MD-DTW is capable of diminishing the effects of irrelevant and/or mistakenly detected activities.

There are three likely reasons that can explain the decrease in accuracy, when utilizing trajectories extracted using the proposed method instead of using trajectories obtained using offline relabelling:

1) failure of the trajectory extraction technique to link the phases in a given movement;
2) inaccuracy of the detected onset and the end positions;
3) propagation of error, i.e., an error in trajectory extraction and/or matching to the correct hand movement results in an error in matching the following movement.

In previously published work on the Ninapro database such as [12] and [26], the delay is defined as the time elapsed between the movement initiation and the classifier outputting the correct movement label. This definition cannot be utilized to evaluate the methods presented in this work, because a classification cannot be obtained until the end of the trajectory (i.e., the changes in the electromyographic activation underlying the movement) is detected. Therefore if the delay was calculated

relative to the movement initiation, it would depend on the speed of the movement execution.

As discussed earlier in Section III-C, the use of MER as an evaluation metric accounts for different types of errors resulting from classification and trajectory extraction, and is insensitive to different sources of delay. Thus the MER values obtained in this work can be directly compared to the results presented in [26].

The authors in [26] employed a continuous control [9] scheme, where a feature vector and a classification were obtained for each overlapping time-limited segment. The authors smoothed the resulting classification sequence by applying a sliding majority-vote window. In order to investigate the trade-off between the error rate and delay, they evaluated the effect of varying the majority-vote window size (i.e., the number of segments contributing to a single majority vote). As expected, increasing the majority-vote window size resulted in a lower MER and longer prediction delay.

The red solid trace in Fig. 9(b) in [26] shows the delay as estimated from the movement initiation versus different MERs obtained by varying the majority-vote window size. The trace was obtained using surface EMG signals only and setting the segment width to 400 ms. For a majority-vote window size of 1 segment, the obtained MER was approximately 25 for an approximate delay of 320 ms from the movement initiation. Increasing the majority-vote window size to 250 segments, the MER decreased to approximately 0.08 for an approximate delay of 1250 ms from the movement initiation. Similar to our results, these results also confirm that incorporating more complete information regarding the changes in EMG signals underlying a movement is expected to decrease error rates.

As shown in Fig. 10, the average MER obtained in [26] for a majority-vote window size set to 250 segments (MER = 0.08) is comparable to the value that we obtained for a system lag of 1300 ms (MER = 0.09). For a system lag of 1300 ms, the proposed trajectory extraction algorithm predicted on average the movement end 262 ms earlier than the end as estimated by the offline relabeling. It is worth noting that the evaluation based on MER can be seen as unbalanced. This is because the number of rest periods is much more than the repetitions of any of the 40 movements. The sequence coding and alignment procedures used in MER estimation consider both rest and movement similarly, since an edit of a rest or a movement segment is penalized equally. For example, consider the first deletion in Fig. 6.

It is also important to re-emphasize here that the accuracies reported in this manuscript in Sections III-B and III-C cannot be directly compared to the window-based accuracies used in [12] and [26]. Window-based accuracy evaluates the classification of a time-limited fixed length segment. On the other hand, the accuracy used in this work evaluates the classification (and also trajectory extraction in Section III-C) for a variable length trajectory. This is not an issue for evaluation based on MER since the authors in [26] erased adjacent duplicate classifications, which ideally would result in variable length trajectories each corresponding to a different movement.

The results presented in the previous section show that it is beneficial to perform fitting of the system for each subject. In this fitting, a subset of movements are selected depending on the sought application and the attained error rates. Lower error rates can also be obtained by subject specific tuning of the preprocessing and trajectory extraction parameters including the sample activity factor, minimal trajectory length, linking window length and decimation factor. Both the fitting of the system and the tuning of its parameters can be performed using the acquired training examples in an automated manner.

The proposed trajectory extraction technique was tested by creating simulated signals obtained by concatenating the second and the fifth repetition of each movement. In order to simulate the scenario that each movement is preceded and followed by a resting period, halves of the resting durations that preceded and followed the execution of each of the movement repetitions were also added to the simulated signal. This might not represent the situation in which the subject does not rest after the execution of each movement. Still, the use of the simulated signal allows investigation of the possibility of detecting more than one movement from a streamed signal. It also allows testing the ability of the proposed trajectory extraction technique to link multiple phases of a given movement and to discard bursts of irrelevant activity.

The main limitation of the proposed work is that a classification of the movement can not be obtained until the end of the trajectory is detected. This restricts the usability of the proposed system to applications that can be controlled using discrete commands. Many human computer interface systems are well suited for this type of control. While for prosthetic devices, the proposed methods can be used to decode the intended movement and perform locking of the device. A regressor relying a signal acquired from another source can be then used to control the duration and the force of the actuation. Another limitation of the proposed classification methods is its inability to handle the case when the subject executes a movement that is not among the training examples. One possible solution is to set a threshold on the normalized matching distance (MD-DTW distance divided by the path length) above which no matching will be assumed to have occurred. However further testing and analysis are necessary to validate the applicability of this solution.

Instead of relying on the simple approach of the nearest neighbor for classification, more advanced classification kernel-based techniques such as support vector machines can use MD-DTW distance as the kernel function. Similar approaches are shown to be effective in speech recognition techniques [27]. However such techniques are not possible given the available number of training examples.

1) Future Directions:

- Combining EMG and accelerometer trajectories is expected to improve both trajectory extraction and the classification. Such a combination has already been shown to be useful in window based classification methods [26].
- Regressing characteristics of the extracted trajectories including duration, maximum amplitude and number of phases to properties of the intended control such as the speed of execution, force and other application dependent control parameters.
- Investigation of the optimal number of examples to be kept for each movement, so as to optimize the trade off between classification accuracy and classification time needed. An

alternative of keeping examples is to estimate a statistical template of examples belonging to each movement. A potential method for estimation of such a template is statistical dynamic time warping [28].

- Accelerating the classification by using a lower bound estimate of the MD-DTW distance. A lower bound estimate is to be less expensive to compute compared to the MD-DTW distance. Knowing the lower bound distance for a given matching might result in skipping the computation of the MD-DTW, if the current attained minimum distance is below the lower bound . Another approach that can be effective in accelerating the matching is to constrain the matching to a pre-specified limited region around a diagonal match, on the other hand this limits the extend of the temporal variation between two trajectories to be tolerated by the MD-DTW [29]. This speed up of the matching process could be crucial for successful implementation of the proposed system using low-power embedded systems and to scale the size of the training dataset.

V. Conclusion

This work shows the potential for using distance-based classification methods with multidimensional sequences in the context of myoelectric control. Accumulated distance based on MD-DTW alignments was shown to be an effective discriminant measure across trajectories corresponding to 40 different hand movements. An accuracy of 89% was achieved based on trajectories obtained using offline relabelling. In order to move towards the realization of myoelectric control, we proposed a trajectory extraction technique and obtained an average movement error rate of 0.09. While the preliminary results of this work focus on accuracy and computational complexity as metrics to evaluate the proposed system, other metrics specific to the target application need to be considered.

References

[1] P. Parker, K. Englehart, and B. Hudgins, "Myoelectric signal processing for control of powered limb prostheses," *J. Electromyogr. Kinesiol.*, vol. 16, no. 6, pp. 541–548, 2006.

[2] T. S. Saponas, D. S. Tan, D. Morris, J. Turner, and J. A. Landay, "Making muscle-computer interfaces more practical," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2010, pp. 851–854.

[3] R. Alter, "Bioelectric control of prosthesis," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, 1966.

[4] D. S. Dorcas and R. N. Scott, "A three-state myo-electric control," *Med. Biol. Eng.*, vol. 4, no. 4, pp. 367–370, 1966.

[5] D. Farina *et al.*, "The extraction of NEURAL information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797–809, Jul. 2014.

[6] Z. Xing, J. Pei, and E. Keogh, "A brief survey on sequence classification," *ACM SIGKDD Explorations Newslett.*, vol. 12, no. 1, pp. 40–48, 2010.

[7] G. N. Saridis and T. P. Gootee, "EMG pattern analysis and classification for a prosthetic arm," *IEEE Trans. Biomed. Eng.*, vol. 29, no. 6, pp. 403–412, Jun. 1982.

[8] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 1, pp. 82–94, Jan. 1993.

[9] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[10] A. D. C. Chan and K. B. Englehart, "Continuous myoelectric control for powered prostheses using hidden Markov models," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 121–124, Jan. 2005.

[11] H. Han and S. Jo, "Supervised hierarchical Bayesian model-based electomyographic control and analysis," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 4, pp. 1214–1224, Jul. 2014.

[12] M. Atzori, M. Baechler, and H. Müller, "Recognition of hand movements in a trans-radial amputated subject by sEMG," in *IEEE Int. Conf. Rehabil. Robot.*, Jun. 2013, pp. 1–5.

[13] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh, "Querying and mining of time series data: experimental comparison of representations and distance measures," *Proc. VLDB Endowment*, vol. 1, no. 2, pp. 1542–1552, 2008.

[14] G. Ten Holt, M. Reinders, and E. Hendriks, "Multi-dimensional dynamic time warping for gesture recognition," in *13th Annu. Conf. Adv. School Comput. Imag.*, 2007, vol. 300, pp. 23–32.

[15] A. Corradini, "Dynamic time warping for off-line recognition of a small gesture vocabulary," in *IEEE ICCV Workshop Recognit., Anal., Track. Faces Gestures Real-Time Syst.*, 2001, pp. 82–89.

[16] A. Akl and S. Valaee, "Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, & compressive sensing," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2010, pp. 2270–2273.

[17] N. Gillian, R. B. Knapp, and S. O'Modhrain, "Recognition of multivariate temporal musical gestures using n-dimensional dynamic time warping," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2011, pp. 337–342.

[18] M. AbdelMaseeh, T. Chen, and D. Stashuk, "Multifunction myoelectric control using multi-dimensional dynamic time warping," in *Proc. 36th Annu. Int. Conf. IEEE EMBS*, 2014, pp. 4366–4369.

[19] C. J. D. Luca, L. D. Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface EMG signal: Movement artifact and baseline noise contamination," *J. Biomechan.*, vol. 43, no. 8, pp. 1573–1579, 2010.

[20] D. A. Winter, A. J. Fuglevand, and S. E. Archer, "Crosstalk in surface electromyography: Theoretical and practical estimates," *J. Electromyogr. Kinesiol.*, vol. 4, no. 1, pp. 15–26, 1994.

[21] D. Farina, R. Merletti, and R. M. Enoka, "The extraction of NEURAL strategies from the surface EMG," *J. Appl. Physiol.*, vol. 96, no. 4, pp. 1486–1495, 2004.

[22] V. T. Inman, H. J. Ralston, J. B. D. C. M. Saunders, M. B. B. Feinstein, and E. W. Wright, Jr., "Relation of human electromyogram to muscular tension," *Electroencephalogr. Clin. Neurophysiol.*, vol. 4, no. 2, pp. 187–194, 1952.

[23] M. Atzori *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, 2014.

[24] I. Kuzborskij, A. Gijsberts, and B. Caputo, "On the challenge of classifying 52 hand movements from surface electromyography," in *Proc. Annu. Int. Conf. IEEE EMBS*, 2012, pp. 4931–4937.

[25] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: Balancing the competing effects of classification error and controller delay," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 2, pp. 186–192, Apr. 2011.

[26] A. Gijsberts, M. Atzori, C. Castellini, H. Muller, and B. Caputo, "Movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 735–744, Jul. 2014.

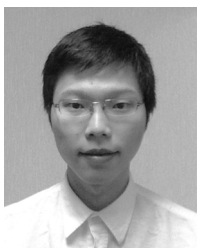
[27] H. S. K. Noma, "Dynamic time-alignment kernel in support vector machine," *Adv. Neural Inf. Process. Syst.*, vol. 14, p. 921, 2002.

[28] C. Bahlmann and H. Burkhardt, "The writer independent online handwriting recognition system FROG on hand and cluster generative statistical dynamic time warping," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 3, pp. 299–310, Mar. 2004.

[29] V. Niennattrakul, P. Ruengronghirunya, and C. A. Ratanamahatana, "Exact indexing for massive time series databases under time warping distance," *Data Min. Knowl. Discov.*, vol. 21, no. 3, pp. 509–541, 2010.

Meena AbdelMaseeh is a Ph.D. degree candidate in systems design engineering at the University of Waterloo, Canada, where he is also a member of the Center for Pattern Analysis and Machine Intelligence. He received the B.Sc. and M.Sc. degrees in systems and biomedical engineering from Cairo University, Egypt, in 2007 and 2011, respectively.

His current research focuses on development of algorithms for analysis of electromyographic and electroencephalographic signals. Applications of his work include detection of epileptic discharges, characterization of neuromuscular disorders, and myoelectric control.



Tsu-Wei Chen received the B.S. and M.S. degrees in computer engineering and systems design engineering respectively from the University of Waterloo, Waterloo, ON, Canada.

His research interests include machine learning and signal processing.

Daniel W. Stashuk (M'89) received the B.ASc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1978, and the M.Eng and Ph.D. degrees in electrical engineering from McMaster University, Hamilton, ON, Canada, in 1982 and 1985, respectively.

From 1985 to 1988 he held the position of Research Assistant Professor at Boston University NeuroMuscular Research Center and was supervisor of the Motor Unit Laboratory. Since 1988 he has been a faculty member with the Department of Systems Design Engineering at the University of Waterloo and currently has the rank of Professor. He is interested in the detection and analysis of biological signals, especially the detection and analysis of electromyographic signals for the extraction of clinically relevant information.