Segmentation and Classification of EMG Time-Series During Reach-to-Grasp Motion

Mo Han, Mehrshad Zandigohar, Mariusz P. Furmanek, Mathew Yarossi, Gunar Schirner, and Deniz Erdoğmus

Abstract—The electromyography (EMG) signals have been widely utilized in human–robot interaction for extracting user hand/arm motion instructions. A major challenge of the online interaction with robots is the reliable EMG recognition from real-time data. However, previous studies mainly focused on using steady-state EMG signals with a small number of grasp patterns to implement classification algorithms, which is insufficient to generate robust control regarding the dynamic muscular activity variation in practice. Introducing more EMG variability during training and validation could implement a better dynamic-motion detection, but only limited research focused on such grasp-movement identification, and all of those assessments on the non-static EMG classification require supervised ground-truth label of the movement status. In this study, we propose a framework for classifying EMG signals generated from continuous grasp movements with variations on dynamic arm/hand postures, using an unsupervised motion status segmentation method. We collected data from large gesture vocabularies with multiple dynamic motion phases to encode the transitions from one intent to another based on common sequences of the grasp movements. Two classifiers were constructed for identifying the motion-phase label and grasp-type label, where the dynamic motion phases were segmented and labeled in an unsupervised manner. The proposed framework was evaluated in real-time with the accuracy variation over time presented, which was shown to be efficient due to the high degree of freedom of the EMG data.

Index Terms—electromyography (EMG) signals, dynamic EMG, online grasp recognition, time-series segmentation, machine learning

I. INTRODUCTION

Human–robot interaction (HRI) technology is developed rapidly in recent years and enables growing collaborative robotics market. An intuitive interface for including human into the robotic control loop is the activity detection of hand and arm [1], which does not require users to have professional technical skills in robot control. The electromyography (EMG) signals have been widely utilized in extracting such hand/arm motion instructions, since the EMG signal can be non-invasively recorded from the skin surface and represent the electrical activity in the muscles. Different EMG-based robotic systems were proposed for estimating the hand and arm movements, shown to be efficient for controlling prosthetic devices [2–4], driving virtual hands in computer animation [5], and remotely operating robotic arms [3].

A major challenge of the online interaction with robots is the continuous and reliable recognition of EMG from real-time data streams. Previous studies mainly focused on exploiting a mapping between upper limb EMG signals and hand postures [2–7], as strategy to control the large number of the hand’s degrees of freedom. However, in these approaches only steady-state EMG signals were used to train and validate their classification algorithms, which cannot ensure the model robustness since the muscular activity varies between a static and a dynamic arm position, and the hand configuration also changes simultaneously with the arm’s motion [8]. Moreover, those studies have only considered a small number of grasp patterns, which are insufficient to generate differentiated control of various application scenarios. To improve the control effectiveness and user comfort, human intention should be detected in a more dynamic, natural and smooth manner.

Introducing more EMG variability into the model training and validation could be a feasible solution to obtain a robust online classification [9], which can improve the system applicability using the dynamic phases of transient EMG [10]. Moreover, the hand motions are always performed in concert with the arm dynamic movements - such as, during the reach-to-grasp and grasp-to-return movements, the configuration of the fingers and wrist changes simultaneously and continuously with the arm’s motion according to the shape and distance of the target object [11]. Therefore, models which can identify dynamic arm postures and varying muscular contractions could provide sufficient response time for pre-shaping the robot, and consequently improve the system usability and natural transition of grasp movement. However, only limited research focused on the detection of dynamic grasp motions [12–15], and all of those assessments on the non-static EMG classification require supervised ground-truth label of the movement status.

In this study, we propose a framework for classifying dynamic EMG signals generated from continuous grasp movements with variations on multiscale muscular contractions and dynamic arm/hand postures, using an unsupervised motion status segmentation method. We exploit the continuity of the hand formation change during the reach-to-grasp and grasp-to-return movements to increase the data variability, and decode the grasping intention of the subject in a real-time manner. We collected data from large vocabularies of gestures with multiple dynamic motion phases to encode the transitions from one intent to another based on common sequences of the grasp movements. Two classifiers were constructed for identifying the motion-phase label and grasp-type label in parallel, where the dynamic motion phases were segmented and labeled using an unsupervised method [16]...
with no ground-truth annotation of continuous movements required. The proposed dynamic-EMG identification framework was evaluated in real-time with the accuracy variation over time presented, which was shown to be efficient due to the high degree of freedom of the EMG data, so that a smarter active assistance and more intuitive interactions between human and machine could be implemented by the EMG-based system.

II. MATERIALS AND METHODS

A. Experimental Protocol and Data Collection

The utilized data [17] was collected from 5 healthy subjects (4 male and 1 female) with an average age of 26.7±3.5 years. None of the participating subjects acknowledged any known motor or psychological disorders. The subjects were all right-handed and they used only their dominant hand throughout the experiments. Prior to each collection session, the experimental procedures and instructions were fully explained to each of the participants and their consents were taken for their participation.

1) EMG Sensor Configurations: Surface EMG from $C = 12$ muscles were recorded for each subject, covering from hand to arm: intrinsic hand muscles (First Dorsal Interosseous - FDI, Abductor Pollicis Brevis - APB, Flexor Digiti Minimi - FDM), extrinsic hand muscles (Extensor Indicis - EI, Extensor Digitorum Communis - EDC, Flexor Digitorum Superficialis - FDS, Extensor Carpi Radialis - ECR, Extensor Carpi Ulnaris - ECU, Flexor Carpi Ulnaris - FCU), as well as upper arm muscles (Brachioradialis - BRD, Biceps Brachii Long Head - BIC, and Triceps Brachii Lateral Head - TRI). To palpate each of the aforementioned muscles, the experimenter followed a manual based on SENIAM recommendations [18]. After skin preparations, bipolar electrodes of Motion Lab Systems (Baton Rouge, LA, USA) were attached to each muscle, and the EMG was recorded with a sampling rate of $f = 1562.5$ Hz.

2) Experimental Protocol: The experimental protocol focused on 14 grasp types and 4 dynamic motion phases involving commonly used hand and wrist motions [19]. As shown in Fig. 1, the 14 classes were: large diameter, small diameter, medium wrap, parallel extension, distal, tip pinch, precision disk, precision sphere, fixed hook, palmar, lateral, lateral tripod, writing tripod, and open palm/rest. During each grasp task, the subject was required to complete 4 dynamic movement actions continuously including reaching (moving the object), grasping (grasping to move the object), returning (returning hand to the rest position), and resting (resting at the rest position with open palm).

Each subject participated in two collection sessions in total, involving the task to lift and move different objects from one position to the other, where in the first session the object was moved in a clockwise trajectory while the second session was in counterclockwise. The subject was allowed to rest for fifteen minutes between the two sessions. During the session, each one of the first 13 grasp types shown in Fig. 1 (not including the open-palm/rest gesture) was performed four times with four different objects, leading to 52 objects totally in each session. The subject performed 6 trials for all the 52 objects per session, where each trial was executed along its corresponding predefined path, as shown in Fig. 2. During the first trial t1, the object was moved from the initial position P0 to the position P1, followed with another five trials to move the object clockwise until it was returned to the initial position P0. The counterclockwise session was performed in a similar manner as Fig. 2 but in a different direction with respect to the initial position P0.

At the beginning of the experiment, the subject was seated facing a table and the electrodes were connected to right arm while the arm was at the rest position with an open palm, as illustrated in Fig. 2. The object center configuration was defined with 6 marks on the table. A screen was placed at the right side of the subject for showing the example picture of current grasp type to be executed. First, the subject was given 5 seconds to read the gesture figure shown on the screen, followed with an audio cue illustrating the beginning of the first trial. Each trial lasted for 4 seconds, and the object was grasped and moved along its predefined path using the designated grasp type for 6 trials without interruption, with audio cues given between different trials. Within each 4-second trial, the subject was required to complete 4 motion tasks continuously including reaching, grasping, returning, and resting. The complete timeline for grasping each object is presented in Fig. 3. Note that for each trial, the four grasp phases (reaching, grasping, returning and resting) were performed freely and naturally by the subject without limitation on the speed of each phase as long as all the 4 phases were executed within 4 seconds, in order to preserve sufficient information of the dynamic motion.

B. Data Pre-Processing

The data were filtered with a band-pass Butterworth filter between 40 Hz and 500 Hz for anti-aliasing and motion-artifact removal. After that, we constructed the root-mean-square envelopes [20] of the EMG signals, using a sliding window with the size of 150 samples. Thereafter, the resulting data for each subject were normalized based on their strength using maximum voluntary contraction (MVC) values. The MVC values for each of the subject’s muscles were collected at the beginning of their collection session, where
the subject was asked to perform the strongest isometric contraction continuously for 3 seconds.

To enable real-time inference, the processed EMG signals were further divided into sliding windows of length $T = 200$ ms, with a step size of $40$ ms between two neighboring windows.

C. Feature Extraction

Compared to other form of representations, features of EMG in time domain can be calculated based on raw EMG data without any time-frequency transformations [21], leading to a higher computation efficiency for real-time implementation. Based on this intuition, three time-domain features were adopted for this work: root mean square (RMS), mean absolute value (MAV), and variance of EMG (VAR). The RMS feature represents the square root of the average power of the EMG signal for a given period of time, which models the EMG amplitude as a Gaussian distribution. MAV feature is an average of absolute value of the EMG signal amplitude, which indicates the area under the EMG signal once it has been rectified. VAR feature is defined as the variance of EMG, which is calculated as an average of square values of the deviation of the signal. Given a pre-processed EMG time window $X \in \mathbb{R}^{C \times T}$ consisting of $C = 12$ channels (representing 12 muscles) with the window size of $T = 200$ ms, the feature extraction defines a mapping from $X$ to $Z \in \mathbb{R}^{3C \times 1}$, where $Z$ is the extracted feature in the time-domain.

D. Data Annotation

In order to approach the grasp type classification in a continuous manner and simultaneously detect the dynamic motion of the arm and hand, each EMG trial was segmented unsupervisedly into four sequences corresponding to the performed 4 motion phases, which were labeled separately according to the specific motion. As shown in Fig.3, the resulting EMG sequences within a trial were annotated in parallel by two groups of labels: motion phase label $l_1$ and grasp type label $l_2$. Note that for each trial, different grasp motion phases were performed naturally by the subject without limitation on the timing of different phases as long as the task was completed within the trial, so the length of each phase was not necessarily equal.

1) Unsupervised Segmentation of Dynamic Motion: Under specific stationary conditions, surface EMG signals can be generally modeled efficiently as zero-mean random process of Gaussian distribution [22]. Based on this assumption, the EMG trial from continuous grasp movements can be segmented into different dynamic sequences using an unsupervised method of Greedy Gaussian Segmentation (GGS) [16].

In this work, we broke down each EMG trial into four segments by assigning three breakpoints, where the segments represent the 4 phases of reach-to-grasp motion, i.e. reaching, grasping, returning and resting, respectively. The precedence of different dynamic phases is considered here, by accounting for higher probability for phase transitions that are more likely to follow one another based on our experiment protocol. For example, a “reach” movement is followed by a “grasp” action and later followed by a ”return” action, before the hand rests again. Then the local optimal segmentation boundaries of each EMG trial could be located by the GGS method. As depicted in Fig.3, four 12-channel EMG data sequences were modeled as four independent multivariate Gaussian distributions, with distinct means and variances. Lastly, the resulting segments were annotated by the motion phase label $l_1 \in \{1, 2, 3, 4\}$ representing the reaching, grasping, returning and resting phases.

2) Annotation of Dynamic Grasp Types: After the segmentation of EMG trials, in addition to the motion phase label $l_1$, the resulting EMG sequences were also annotated by another group of grasp type label $l_2 \in \{0, 1, ..., 13\}$ in parallel for classifying dynamic grasp, where $l_2 = 0$ represents the open-palm/rest gesture and $l_2 \in \{1, ..., 13\}$ corresponds to the other 13 grasp types listed in Fig.1.
Practically, during the reach-to-grasp and grasp-to-return movements, considering the shape and distance of the target object, the formation of the subject fingers may change as a result of arm motions [11]. Human hands in this regard, tend to pre-shape prior to touching the targeted object. Moreover, the hand even still continues to extend right after the grasped object is released. These behaviours are attributed to the characteristics and features of the targeted object. Therefore, in order to provide a smooth representation of the grasps, we annotated the EMG sequences of reaching, grasping and returning with the executed grasp type \( l_2 \in \{1, ..., 13\} \) corresponding to the target object, and tagged the phase of resting with the open-palm label \( l_2 = 0 \) (see Fig.3).

E. Grasp-Type Classification of Dynamic Arm Movements

We constructed two classifiers for respectively identifying the motion phase label \( l_1 \in \{1, 2, 3, 4\} \) and the grasp type label \( l_2 \in \{0, 1, ..., 13\} \), with corresponding data pairs of \((X_i, l_{1i})\)\(_{i=1}^{n}\) and \((X_i, l_{2i})\)\(_{i=1}^{n}\), where \( X_i \in \mathbb{R}^{C \times T} \) is the \( i \)th EMG window with channel number \( C = 12 \) and window length \( T = 200 \) ms of \( f = 1562.5 \) Hz sampling rate, and \( n \) is the total number of windows. For each EMG window \( X_i \in \mathbb{R}^{C \times T} \), three time-domain features of RMS, MAV and VAR were extracted as \( Z_i \in \mathbb{R}^{3C \times 1} \), leading to data pairs of \((Z_i, l_{1i})\)\(_{i=1}^{n}\) and \((Z_i, l_{2i})\)\(_{i=1}^{n}\) which were the final inputs to the motion-phase and grasp-type classifiers respectively.

To construct both classifiers, we utilized and trained the processed EMG data using the extra-trees method [23]. The extra-trees algorithm is an ensemble model, consisting of multiple randomized decision trees trained by various subsamples of the data. By averaging the overall classification over a group of decision trees, over-fitting could be efficiently prevented, leading to better accuracy and robustness. In our work, we observed that a combination of 40 trees in the extra-trees forest could provide better results compared to other parameters.

III. EXPERIMENTAL EVALUATION AND RESULTS

A. Training and Validation

We performed inter-subject training and validation for both of the 4-class motion-phase classification and the 14-class grasp-type classification. The classification analyses were implemented through a 3-fold cross-validation protocol. For each subject and object, the collected 6 EMG trials were randomly and equally divided into 3 groups (2 trials for each group), where the classifiers were independently cross-validated on each of the 3 groups after trained on the rest 2 groups. All training and testing were performed with Python3 using Scikit-learn library.

B. Results and Discussion

The performances of the two trained classifiers are shown in Table I and Fig. 5. Results in Table I demonstrate the average classification accuracies for each cross-validation fold from both motion-phase classifier and grasp-type classifier, with the mean accuracy over all folds given for each subject. The average accuracy of the 4-class motion-phase classifier for each individual subject varies between 72.9% and 77.8%, while the 14-class grasp-type classifier presents average accuracies ranging from 81.8% to 90.6%. Those results reveal that the grasp phases and types were well-predicted in general. The dynamic-EMG grasp identification showed a better performance than the motion-phase detection, due to the higher degree of freedom regarding to how subjects performed different motion phases than the grasp type, as the experiment protocol did not specify the particular speed or angle to grasp. The illustrated inter-subject variability across different validation trials may also come from the varied grasping patterns and directions for different trials of the same subject. However, this higher degree of freedom could enable more robustness and stability of the model to a wider range of user postures during the dynamic grasp activity. In addition, the training data from various subjects also present different classification performances, which could be influenced by factors such as shifting sensor locations and distinct movement patterns of different users.

In Fig. 5 we show the accuracies of motion-phase and grasp-type classifiers as functions of time in order to inspect the performance variation during different dynamic phases within a trial. Each time point in Fig. 5 represents a EMG window and the accuracies were averaged within the same window over the entire validation set from all subjects and cross-validation folds. When computing the average, timelines of all trials were aligned with the beginning of the grasping phase (marked as 0ms), and the breakpoints between motion phases were also averaged across all validation trials and indicated by the solid green lines.

Overall, the real-time accuracy of the achieved grasp-type classification shown in Fig. 5 was higher than that of the motion-phase identification. For the dynamic grasp-
type classification, the performance was stably over 80% accuracy throughout most of the trial as illustrated in Fig. 5. Compared to the motion-phase classification, the grasp-type classification relies more on the lower arm configurations which are more stable even with different grasping directions and angles. The grasp-type classification accuracy were also changing smoothly during the phase transitions, demonstrating the continuity in lower-arm configuration change during the grasp movements. The grasping phase displayed a more steady performance of grasp-type identification, since more static muscle contraction was necessary to keep holding an object so that more detectable signals could be generated. Moreover, instead of distracting the classifier from current grasp type, including the dynamic portion during reaching and returning in the training set substantially takes advantage of the continuity of grasp motions and increases the number of windows available for each class.

More performance fluctuation in time could be observed from the motion-phase classification compared to the grasp-type classification. As demonstrated in Fig. 5, the accuracies of motion-phase identification reach peaks in the middle of each motion phase, and then decrease again until the next phase. More errors occurred during phase transitions, and specifically a dramatic accuracy drop was observed between reaching and grasping phases. This accuracy drop may be caused by the fact that the subject performed contact grasps where they were touching an external object between reaching and grasping phases, which could result in a significant change in the EMG signal from the higher-arm muscles [24]. The higher arms are more in charge of the direction and angle of the limb movement, while the lower arm and hand are more responsible for the grasp activity, so that higher arm may experience more variability during a dynamic movement for the same grasp type. Therefore, a possible reason for the accuracy fluctuations of the motion-phase detection throughout the entire trial could come from varying patterns and angles that different individuals utilized in different trials, leading to distinct co-contraction responses of higher arms. In addition, the accuracy decrease during grasp-to-return movement was lower than that of reach-to-grasp movement, since releasing an object applies force in a more gradually-decreasing manner than contacting and lifting an object. There was good agreement between the two classifiers regarding their classification accuracies during resting phase, with smooth convergence to stable performances, indicating that the EMG signals from resting phase were more steady and generic for different subjects and trials.

IV. CONCLUSION

This paper presented a non-static EMG recognition method for identifying real-time hand/arm movements to generate robust control regarding the dynamic muscular activity variation in practice. We trained and validated the proposed framework using EMG signals generated from continuous grasp movements with variations on dynamic arm/hand postures, to encode the transitions from one intent to another based on common sequences of the grasp movements. We constructed two classifiers for respectively recognizing the motion-phase label and grasp-type label, where the dynamic motion phases were segmented and labeled in an unsupervised manner. Finally, the proposed method was assessed in real-time and the corresponding accuracy variation over time was presented. Results illustrated the effectiveness of the framework built with the EMG data of high degree of freedom.

REFERENCES

[1] S. Sheikholeslami, A. Moon, and E. A. Croft, “Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human–robot interaction,” The
