Online Finger Control Using High-Density EMG and Minimal Training Data for Robotic Applications

Barsotti, Michele; Dupan, Sigrid S.; Vujaklija, Ivan; Dosen, Strahinja; Frisoli, Antonio; Farina, Dario

Published in:
IEEE Robotics and Automation Letters

DOI (link to publication from Publisher):
10.1109/LRA.2018.2885753

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Abstract—A hand impairment can have a profound impact on the quality of life. This has motivated the development of dexterous prosthetic and orthotic devices. However, their control with neuromuscular interfacing remains challenging. Moreover, existing myocontrol interfaces typically require an extensive calibration. We propose a minimally supervised, online myocontrol system for proportional and simultaneous finger force estimation based on ridge regression using only individual finger tasks for training. We compare the performance of this system when using two feature sets extracted from high-density EMG recordings: EMG linear envelope (ENV) and non-linear EMG to Muscle Activation mapping (ACT). Eight intact-limb participants were tested using online target reaching tasks. On average, the subjects hit 85 ± 9% and 91 ± 11% of single finger targets with ENV and ACT features respectively. The hit rate for combined finger targets decreased to 29 ± 16% (ENV) and 53 ± 23% (ACT). The non-linear transformation (ACT) therefore improved the performance, leading to higher completion rate and more stable control, especially for the non-trained movement classes (better generalization). These results demonstrate the feasibility of proportional multiple finger control in intact subjects by regression on non-linear EMG features with a minimal training set of single finger tasks.

Index Terms—Prosthetics and Exoskeletons, Dexterous Manipulation, Electromyography (EMG), linear regression, feature selection, myoelectric control, online control

I. INTRODUCTION

Myoelectric control commonly relies on decoding human motor intent from non-invasive electromyographic signals (EMG) and on mapping EMG into control outputs, allowing for the establishment of intuitive human-machine interfaces. This control strategy has been applied for multi-functional prostheses and robotic exoskeletons [1]–[4]. With the rapid development of robotic technology, control strategies based on simultaneous and proportional control of multiple degrees of freedom (DoFs) have been proposed to mimic natural control. However, the robustness and accuracy of these controllers decrease with an increase in the number of controllable DoFs [2]. This problem is critical for hand function restoration (exoskeletons or prostheses) because of the large number of DoFs [5].

Human fingers can move dexterously and with precision in numerous ways allowing for the simultaneous activation of multiple DoFs, with different amounts of forces exerted by each finger [6]. During the past two decades, there have been attempts to enable similar articulated control of robotic hand devices [7], [8].

Both pattern recognition and regression based algorithms have been previously used for establishing online control of finger movements [9]–[15]. Cipriani et al. [11] successfully demonstrated the feasibility of a real time classification of 7 pre-defined hand-postures, including some individual fingers movements, achieving an average accuracy of the classifier of 79% in 5 amputee subjects. Khousha et al [12] succeeded in discriminating 10 individual and combined motion classes of finger movements, with an online classification accuracy of ~90%. The first session, where participants had no previous training in online control, resulted in accuracy of ~85%. They used a Support Vector Machine (SVM) classifier with non-overlapping Time Domain (TD) features selected using a Linear Discriminant Analysis.

The major limitation of classification methods is the lack of proportional activation of the recognized classes [16]. An error in classification will compromise the entire gesture due to the on/off nature of classification algorithms, leading to a frustrating situation for the user. While traditional pattern
recognition methods do not include proportional control, it is possible to extend the control algorithm with a proportional regression of the estimated forces after the class is determined [15]. Regression methods intrinsically allow for the simultaneous and proportional estimation of different DoFs. This allows the regressor to be trained on a limited data set, (e.g., single DoF), after which it can extrapolate to the movements outside the training set (e.g., DoF combinations). This could dramatically reduce the time needed for training the system, which is particularly important in rehabilitation applications, where a shorter time for the setup phase results in more time for the rehabilitation exercises.

Krasoulis et al. [17] demonstrated that non-linear regressors outperformed the linear ones when estimating movements seen by the decoder during the training task. However, when generalizing to novel movements, the performance of the two regressor types was comparable. Castellini et al. [18] showed that the accuracy of single finger estimation diminishes by including combined finger forces in the training set. It is important to note that all these studies were performed offline, and it is therefore still unknown if these differences in performance would hold when the task is performed online. Nowak and Castellini [19] illustrated that the estimation of finger grips outside of the training set is possible, in both an offline and online setting, when the training data for those movements was generated artificially by linearly combining the data of the existing classes, a procedure named linearly enhanced training (LET).

The most common features for regression approaches are based on EMG amplitude [20]. However, classical estimators of EMG amplitude (e.g. Mean Absolute Value (MAV), Root Mean Square (RMS), linear envelope (ENV)) suffer from high variability and strongly depend on the selected time window [20]. Recently, non-linear biological-inspired descriptors of EMG amplitude have been shown to outperform the classical linear estimators [21]–[24]. For instance, the so-called EMG-to-Muscle Activation (ACT) is a model-driven feature that has been successfully used in EMG-based joint kinematics reconstruction [25]. It was shown that the ACT was able to deliver better results when estimating simultaneous finger kinematics than the classic TD features (MAV, Waveform Length, Willison Amplitude and Variance) [26]. However, this has been evaluated only in an offline analysis with targeted electrode placement while using a combined motion capture and EMG data set for supervised training of the regressors.

Here, we present and test a minimally supervised online myoelectric control system for proportional and simultaneous finger force estimation. This was achieved using a reduced EMG training set suited for clinical translation, consisting of only one repetition of each individual finger flexion press and one press with all 5 fingers. Relying on a regularized linear ridge regressor driven by the ACT or ENV feature, subjects were asked to control a 5 DoF computer game. The developed method was tested through a set of dexterous tasks including both individual and multi-digit control.

The proposed experiments were designed in order to be relevant in both medical and non-medical scenarios. Two possible clinical applications were considered in particular: hand amputees controlling a prosthesis and neurological patients receiving a therapy through virtual games or robotic exoskeletons. To this aim, in the proposed experiment, no kinematic or force data was recorded and performance was evaluated using a virtual target-hitting task. In both envisaged clinical applications for amputees and neurological patients, the muscles are activated isometrically, and force or kinematics are usually not available. Additionally, both patient groups appreciate a short training session (performed with a minimal training data). The designed experimental paradigm is indeed a simplification of daily-life conditions, yet it meets the requirements for the two patient groups while still providing a relevant testing scenario for non-clinical myoelectric applications (e.g., telemanipulation).

To the best of the authors’ knowledge, no other myoelectric control scheme demonstrated the feasibility of the online simultaneous and proportional control of finger flexions trained on a small data set.

II. METHODS

In order to provide subjects with EMG driven simultaneous and proportional finger control, two ridge regressors were trained for each participant. Each regressor was tested in a separate session, using either ENV or ACT as the input feature. Following the system training, subjects were asked to complete a set of target reaching tasks in order to test and compare the proposed system in an online scenario.

A. Subjects

Eight able-bodied, right-handed subjects (age: 21 - 48 years, two females, six males) participated in the study. None of the subjects reported any history of neurological disorders. Each participant read and signed the written informed consent. The study was approved by the research ethics committee of the University Medical Center Göttingen (Nr: 32/2/16), and conformed to the Declaration of Helsinki.

B. Experimental setup

The participants sat in a comfortable position with the fingers of their right hand placed on the table so that the elbow was flexed at approximately 120°. The subject held both the hand and fingers in the same position during the entire experiment. A 26” LCD screen placed in front of the subject at a distance of 70cm displayed the visual cues. High-Density monopolar surface EMG signals were recorded using three semi-disposable, pre-gelled 8x8 electrode grids (ELSCH064NM3, OT Bioelettronica, 10 mm inter-electrode-distance) for a total of 192 electrodes. The electrode grids were placed around the forearm starting at 20% of the forearm length distally to the elbow crease, covering 8 cm longitudinally and 24 cm circumferentially (Figure 1). The edge of the first grid was placed above the ulnar bone and the other two electrodes followed medially from the top. This configuration allowed the acquisition of EMG activity of all the major forearm muscles involved in finger movements. The signals were recorded using the EMG-USB2 OT Bioelettronica amplifier with gain set to
subject-specific values of 500 or 1000, band-pass filtered at 3Hz – 900Hz and sampled at 2048Hz with a resolution of 2.44µV per least significant bit (12-bit A/D conversion). A reference electrode band was strapped around the wrist bone and the skin was cleansed with alcohol pads prior to electrode placement. The whole setup took approximately 15 min. The raw EMG acquisition, as well as the offline analysis and the online assessments, have been conducted using a PC running Microsoft Windows 7 64bit, Intel i7 1.73 GHz, 6GB RAM and Matlab 2013b.

C. Training session

The experiment consisted of a training, and a testing session. In the former, subjects were instructed to perform a set of finger presses to collect the data for the training of the ENV- and ACT-based regressors. The latter session comprised a series of target reaching tasks that subjects completed online using the regressors.

In the training session, the subjects were prompted to perform one repetition of the single finger presses, and one repetition of the five-finger press using a comfortable contraction level. This was defined as the level of force that they could maintain easily for several seconds. This contraction level was represented by filling 50% of the visual feedback bars (referred hereafter as the 50% of the Target Activation Level, TAL). Each press was preceded by a 3-s preparation phase during which the finger to be activated was indicated to the subject using the visual feedback (Figure 1). After a 1s pause, the cueing bar started moving from zero to the middle of the screen (50% of the TAL), reaching that level after 2.5 s. The 50% of the TAL had to be maintained for 4.5 s, after which the bars returned to zero in 2.5 s, where the cue stayed for 1 s. The participant was instructed to follow the cueing bar thereby replicating the trapezoidal force profile by flexing the given finger against the table. This training scheme does not require recording any generated force and is therefore clinically applicable. The corresponding regression is minimally supervised since there is no force labeling.

D. Control Framework

The online controller was designed by regressing recorded EMG features during finger flexion presses executed following the trapezoidal force cues. Since there was no force measurement, the EMG features were mapped onto the the prompts and this indirectly allowed the estimation of the relative force levels across fingers.

Linear regression provides a linear mapping $W_{ENV} \in \mathbb{R}^{D_1 \times D_2}$ between the D_1-dimensional space of input EMG feature values and the D_2-dimensional target space of finger force cues:

$$Y = W^T \cdot X$$

where $X = [x(t_1), x(t_2), \ldots, x(t_N)] \in \mathbb{R}^{D_1 \times N}$ is a matrix of feature values at N time instances and $Y = [y(t_1), y(t_2), \ldots, y(t_N)] \in \mathbb{R}^{D_2 \times N}$ contains the target cues. In this experiment, to account for all sensors, $D_1 = 192$, and $D_2 = 5$ to match all the digits. Since the labels were the force cues in only one trial of each targeted motion, the model was constrained to avoid overfitting. Therefore, we used linear regression with regularization (i.e. ridge regression):

$$W = \left(XX^T + \lambda I\right)^{-1}XY^T$$

where $\lambda$ is regularization parameter and $I \in \mathbb{R}^{D_1 \times D_1}$ is the identity matrix. Computationally heavy calculation of the pseudo inverse $(XX^T + \lambda I)^{-1}X$ is only required while establishing the regressor. Once the mapping $W$ is obtained, the control outputs (finger force estimates $Y$) are computed online by a simple matrix multiplication of $W^T$ and the newly acquired feature matrix $X$, as given in (1).

Two regressors were trained using ENV and ACT as input features (Figure 1C). Both features were calculated over 200ms windows of EMG with 50% overlap.

The ENVs were extracted by full-wave rectifying, and low-pass filtering (eight-order Butterworth digital filter, cut-off frequency 2 Hz) of the EMGs. The ACT takes into account additional physiological processes related to muscle activation (see Figure 1 for an overview of the processing steps for both regressors). First, the dynamics of neural activation $u(t)$ was modeled as [25]:

$$u(t) = (\alpha e(t - d)) - \beta_1 u(t-1) - \beta_2 (t-2)$$

where $e(t)$ is the linear envelope (as computed for ENV), $d$ is the electro-mechanical delay (EMD) and the parameters $\alpha, \beta_1$ and $\beta_2$ are the coefficients that define the second-order dynamic. In order for the given recursive filter to be stable, the parameters should satisfy the following conditions:

$$\beta_1 = \gamma_1 + \gamma_2$$

![System overview](image.png)
The online performance was quantified using the following outcome measures:

1) Completion rate: the percentage of targets the subject was able to hit before the timeout.
2) Completion time: the amount of time needed for successfully completing the task.
3) Number of dwellings: the amount of times the subject reached the target, but was unable to remain within the target window for the required dwell time of 0.5 s.

A 2×2 repeated measures factorial design was adopted for the online session ([single vs. combined fingers] × [ACT vs. ENV feature]). A 2-way ANOVA test was conducted separately for each of the three performance measures, after determining the normality of the data distribution using the Lilliefors test. The level of significance was set to p < 0.05 for the main tests, and the post-hoc comparisons were corrected using the Bonferroni method.

III. RESULTS

Figure 2 shows the summary results (mean ± standard error) grouped by the type of task (single finger vs combination), and type of feature (ACT vs. ENV).

For the completion rate, there was a significant interaction between the type of feature [ACT vs. ENV] and the type of task [single finger vs. combination] (F(1,7)=11.065, p=0.013, η²=0.796). For both features, the completion rate was significantly higher (p<0.01) in the single finger tasks compared to the combined finger tasks. The subjects achieved high completion rates during single finger tasks (91 ± 11% and 85 ± 9% for ACT and ENV, respectively), and the rates dropped to 53 ± 23% (ACT) and 29 ± 16% (ENV) during combination tasks. In the finger combination tasks, the ACT regression significantly outperformed the regression based on ENV (F(1,7)=10.573, p=0.014, η²=0.796).

The analysis of the completion time only showed a main effect of the type of task (F(1,7)=27.354, p=0.014, η²=0.793), demonstrating that the participants were faster in accomplishing the single finger tasks compared to the finger combination tasks, irrespective of the feature type.

Interestingly, the main effect of the feature type was statistically significant (F(1,7)=20.124, p=0.003, η²=0.968) for the number of dwellings. The number of dwellings was substantially higher for the ENV (3.19±1.44 dwellings) compared to the ACT (1.50±0.37 dwellings). The control was therefore more stable when using the ACT-based regression.

Figure 3 and Figure 4 report the summary results for the completion time and number of dwellings (mean ± standard error), and overall completion rate of the individual single (Figure 3) and combined (Figure 4) finger presses. The completion rate in this case indicates the percentage of subjects who successfully accomplished a specific task. The results for the single finger presses are grouped by the level of target activation (low and high force targets). Interestingly, the online control performance with ACT was good for the high force targets (91%), despite the fact that this level has not been used for the training. The performance was similar to that achieved for the low force targets (93%). Hit rate for the ENV decreased from 91% for low force targets to 80% for high force targets.
The analysis of completion times and amount of dwellings showed that the high force trials lasted longer, and included a higher number of dwellings than those at low force.

The results for the combination trials show that the task difficulty increased with the amount of fingers included in the task (Figure 4). Completion rates of the ENV-based regressor dropped from 40% for 2-finger to 18% for 3-finger combinations. None of the participants were able to hit the 5-finger trial when the control was based on the ENV. ACT-based control performed better for all the finger combinations, with completion rates of 67%, 33%, and 25% for 2-finger, 3-finger, and 5-finger combinations respectively. Completion time was similar for all the combination targets, irrespective of the number of fingers included. There was an increase in the number of dwellings for more fingers, especially in the ENV-based control.

### IV. Discussion

In this study, we showed that a limited training set, suited for clinical applications, allows for generalization of online regression outside of the trained finger presses. Proportional and simultaneous control was implemented using ridge regression. Importantly, the study has demonstrated that introducing a non-linear transformation of the linear envelope in the regression pipeline significantly improved the online control performance. The ACT based regression resulted in more stable control in all the tasks, and also improved the completion rate for the finger combination presses, therefore leading to more successful generalization (since the training set did not include the combinations). It is worth noting that the performance for the 5-finger combination task was poorer than for the other combinations of fingers, even though the 5-finger contractions were included in the training set. This could be due to the nature of the task itself. In fact, each “active” finger introduced a constraint in the task, i.e., the subject needed to increase and maintain an additional finger force within the target window. Therefore, the 5-fingers task was more difficult than any other combination task. Research into finger combination presses has shown that the maximum force produced by a finger decreases when exerted in combination with other fingers [29]. Therefore, the lower performance of the 5-finger combination might have been a result of the higher muscle activation levels needed to complete the task.

Previous offline studies have shown that linear and non-linear decoders perform similarly when predicting movements not present in the training set [17], [30]. Our results show that the regression based on the non-linear ACT feature outperformed the regression based on the linear ENV feature.
when tested on combinations. Contrary to completion rate, there was no significant difference in the completion time between the two features. Therefore, when the participants were able to hit the targets, they needed similar amount of time when using both linear and non-linear features. The higher number of dwellings with ENV demonstrated that even when the participants reached the target they were not able to stay within the target zone. Therefore, the myoelectric control was not stable with the ENV. This result might have been due to a difference in the power spectral density of the two control signals. If one of the two signals had a higher bandwidth, participants would observe a jitter in their feedback, making it more difficult to remain within the target. However, comparison of the power spectral density for both feature types showed that there was no significant difference in the mean power spectral density ($t(7)=2.0, p=0.08$). In order to analyse the difference between the bandwidth of the two regressors, the PSD of the control signals obtained in each trial have been extracted using the Welch's method and averaged over each subject and type of regressor. The lack of stable control might be due to the limited exposure to the online controller. As participants only performed 20 trials based on each feature. However, this was common to both ACT and ENV, and still the control with ACT was significantly more stable. Future research should investigate if using the controller over a longer period of time leads to a more stable control.

Interestingly, there was no significant difference in the completion rate for the single finger trials based on the feature type. Krasoulis et al. [17] previously showed a superior performance in predicting trained finger movements for the non-linear kernel ridge regression over the linear ridge regression during offline analysis. In a more recent paper Murciego et al. [31] reported how, in an offline analysis conducted over the NINA-Pro dataset, a synergy-based approach based on nonnegative matrix factorization outperformed the linear-regressor approach in estimating forces of single finger but not of fingers combination. The improvement in performance was even more substantial when they added a “classification” stage in the control scheme to determine which fingers were active. Similarly, Xiolyannis and colleagues [32] established a good predictive control using a Gaussian Process based autoregressive model according to an offline analysis conducted on six healthy subjects. However, as it has been demonstrated by Jiang et al. [2], these findings do not necessarily translate to online studies. Online control during our study might have given the participants an opportunity to reduce the errors computed by the regressor. However, as we did not include any offline tests, we cannot guarantee that this is the sole reason for our results. In addition, both features led to a good proportional control. During the training trials, participants were asked to execute comfortable presses against the table, approximating half of their maximal force. The completion rates were similar when reaching targets at almost double the trained force, but dwelling results indicated that the control at that level was more challenging, especially with the ENV-based regressor.

Caution always needs to be applied when interpreting the performance of different control algorithms. For example, the number of parameters that are fitted to the training data can influence the accuracy of the controller [33]. In this study, two parameters (the non-linear shaping factor $A$, and the regularization parameter $\lambda$) were fitted for the regressor based on the ACT feature, whereas only the regularization parameter was estimated for the ENV. The better control of the ACT-based regressor might therefore be a result of fitting more parameters to the data, and not due to the fact that the regressor used a non-linear feature. Including more fitted parameters increases the risk of overfitting to the data, especially when the amount of training data is limited [33]. This would limit the ability of the regressor to generalize to untrained finger combinations and forces. We chose not to fit all possible parameters in the EMG-to-muscle activation feature in order to avoid overfitting, and therefore the parameters characterizing the second order dynamics were fixed and taken from the literature [27]. The results demonstrated that with this strategy we did not overfit the non-linear regressor, as its performance on the non-trained presses outperformed the linear regressor, whereas there was no difference in the trained presses.

This study was a proof-of-concept, aimed at testing the feasibility of predicting untrained combinations of finger movements based on a minimal amount of training data. The chosen experimental design makes the proposed approach relevant for both medical and non-medical applications for the following reasons. The short time envisaged for the system calibration session increases the usability of the myocontrol based applications. While not crucial for able subjects, the imposed isometric conditions match those found in clinically relevant scenarios such as prosthesis control and robot-aided neuro-rehabilitation therapy. [19], [34]. In these cases, for both amputees and neurological patients, interfacing is established using EMG signals elicited during contractions in which the moment arms of the muscles remain the same. Furthermore, the proposed approach is suitable for medical applications as it eliminates the need for force measurements, and decreases the training time.

There are two main drawbacks of the study: limited amount of testing data, as all participants performed every type of press only once, and a lack of limb-impaired subjects among the tested population. Additional investigation is needed in order to conclude whether the performance of the proposed method differs when patient population is considered.

The average completion times obtained in this study, indicate that even, after accounting for the reaction, travel and dwelling time, the presented tasks were challenging enough across subjects to infer translational potential of the approach given the accomplished success rates. The fact that the subjects were able to hit the targets without any training, even when being naïve to myoelectric finger control, highlights the feasibility of the control. However, it would be interesting to study the learning aspects of the control over time.

To our knowledge, only [11] and [13] have shown finger control in amputee patients based on a regression algorithm. However, they did not investigate if they were able to predict untrained movements. Still, it has been previously shown that
both regression [35] and classification [11] based myoecontrol systems tend to perform similarly after learning the new interface, regardless of whether the user has an impairment or not.

In future work, we will focus on extending and testing the proposed system in a rehabilitation protocol for stroke and amputee patients, involving robotic devices (such as exoskeletons and prosthetics). Other possible clinical uses involve serious games for user training, and treatment of pain. Moreover, concrete applications in non-medical context, such as teleoperation, will be further explored.

REFERENCES