Controlling an Exoskeleton with EMG Signal to Assist Load Carrying: A Personalized Calibration
Benjamin Treussart, Franck Geffard, Nicolas Vignais, Frédéric Marin

To cite this version:

HAL Id: cea-02863563
https://cea.hal.science/cea-02863563
Submitted on 10 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Controlling an exoskeleton with EMG signal to assist load carrying: a personalized calibration

1st Benjamin Treussart  
SRI, LASR  
CEA-List  
Saclay, France  
benjamin.treussart@cea.fr

2nd Franck Geffard  
SRI, LASR  
CEA-List  
Saclay, France  
franck.geffard@cea.fr

3rd Nicolas Vignais  
CIAMS  
Univ. Paris-sud, Univ. Paris Saclay and Univ. d' Orleans  
Orsay and Orleans, France  
nicolas.vignais@u-psud.fr

4th Frederic Marin  
Biomecanique et Bioingenierie UMR CNRS 7338  
Universite de Technologie de Compiègne  
Compiègne, France  
frederic.marin@utc.fr

Abstract—Implementing an intuitive control law for an upper-limb exoskeleton to perform force augmentation is a challenging issue in the field of human-robot collaboration. The aim of this study is to design an innovative approach to calibrate electromyography (EMG) data in order to detect the intention to lift or put down a charge while wearing an upper-limb exoskeleton. Based on a low-cost EMG sensor bracelet placed around the arm (Myo armband, Thalmics Lab, Ontario), a subject-specific mapping procedure is implemented to discriminate motion intentions during lifting tasks with a 1-DoF upper-limb exoskeleton. The processing is divided into two main parts: (i) direction estimation with an artificial neural network, and (ii) a model-based intensity prediction. The mapping procedure has been tested on 7 healthy participants with a precision of $96.9 \pm 3.1\%$ for the classification and a RMS Error of $3.8 \pm 0.8\,N$ at the end effector. This study opens up the way for fast-deployment applications involving exoskeletons or cobots.

Index Terms—EMG, Exoskeleton, Calibration, Intention Detection

I. INTRODUCTION

Industrial workers are performing repetitive physical tasks, which expose them to musculoskeletal disorders (MSD) [1]. MSD is a major public health issue, with an impact on workers’ integrity and economics. Indeed, MSD are causing both loss in productivity and high healthcare costs. According to [2], MSD have led to the loss of 10 millions work days in France in 2012, with an average cost of 21 k€ per case of MSD. A way to prevent MSDs could be to assist the workers during forceful exertions, e.g. load carrying tasks, thus relieving the strain they endure.

Exoskeletons could become a solution for industrial load carrying. Although, exoskeletons can be back drivable (see sec. III-A), and they can help manipulating known and fixed weights [3], the way they can handle various and unknown charges in an efficient manner is yet to be achieved. The knowledge of the user’s intention could help in this matter. Surface EMG signal processing demonstrated reliability in estimating the muscle force [4], and, consequently, it may be a relevant bio-signal to capture the motion intention.

Given a context of industry-oriented application, an algorithm based on a robust calibration process has to be implemented in order to design a subject-specific intuitive-to-use intention detection methodology to control an exoskeleton. This calibration process has to be reliable and not time-consuming to be operational.
II. PREVIOUS WORKS

EMG signals have already been used to control robotic devices [5, 6]. This biocybernetic approach uses EMG signals into two different manners: discretely or continuously. The discrete-EMG method uses pattern recognition based on handcrafted features and a classifier trained with them [7] or some end-to-end neural networks [8]. The main inconvenience of this method of intention detection is its lack of flexibility. Indeed there is a trade-off between the panel of actions and the cognitive charge/reliability of the classification, which depends on the number of classes. In the case of a directional control, one class may correspond to one direction with a fixed speed [6].

On the other hand, a continuous-EMG method seems more adequate but it also presents disadvantages. First of all, it is more complex, due to the fact that the relations between the features of the EMG signals and the movement are highly non-linear [9]. In [10] a continuous 3D estimation of the position of the hand is performed with the use of 9 electrodes targeting specific muscles. Such precise requirements in the placement of the electrodes complicate the deployment and reinforce the user-dependent aspect of the method. In addition, these methods are more adapted to teleoperation, as well as prosthesis assistance rather than a situation where the torques of the robot and the user influence each other. In the current study, a mix of continuous and discrete approaches will be used to evaluate the intentions of the user. The current study is preliminary to torque control laws, and aims to relate user’s intentions to a torque. Torque values will be used rather than normalized EMG so that future implemented control laws will assure the same behavior for each user.

Finally, EMG data is individual-specific and that is why the signal processing requires to be calibrated. Different ways to calibrate EMG signal processing have been proposed. The calibration can be based on voluntary contraction to gather data for classification tasks. Usually, data from all the classes is recorded [11] but some works focus on recognizing one situation against several others, like with a 1-class classifier [12]. Another strategy is to use external load (i.e force sensors or set of weights) [4]. This approach is adapted for certain continuous evaluation, but the use of force sensors increases the cost and the complexity of the procedure and the positioning of the sensors is usually critical. Others methods use external motion capture technology to calibrate with continuous data [13], however it does not enable to estimate the torque applied by the individual. In this study, an upper-limb exoskeleton will be used to collect data with reliable labels (class or torque value). This enables to control precisely the type of movement or effort made by the user.

To summarize, the discrete-to-continuous approach presented in this study contributes from the point of view of practical implementation (using an exoskeleton both to label and as an excitation vector for identification). It is not only more flexible than a fully discrete method, but also it has much more potential of practical deployment than a fully continuous approach. Consequently, a low-cost surface EMG device will be used to foresee a practical use-case. More specific additions to the state of the art will be given in next sections of the current article.

III. SYSTEM SPECIFICATIONS

A. Exoskeleton

The exoskeleton used in this study is an under-actuated upper-body backdrivable type (Fig. 1). Each side consists of two segments (upper-arm and forearm) and four joints. Two of the joints are passive (\( \theta_1 \) and \( \theta_4 \)) and the other two are proportionally linked and powered by the same actuator (\( \theta_3 = 1.5 * \theta_2 \) [15] (Fig. 2 and Tab. I) . The end effector of the exoskeleton, that pulls the hand upward, is attached to the user’s palm of the hand by a strap. Backdrivability is "the nature that when the force is added to the output axis, the output axis is moved by this force and this motion is conveyed to the input axis and the input axis is moved by this motion in the case of actuator or power transmitting mechanism" [16].

During the interaction, human and robotic forearm dynamics are coupled, which can be modeled as follows [17]:

\[
\begin{align*}
M_r \ddot{\theta}_r + b_r \dot{\theta}_r + \tau_{FN} + Q_r(\theta_r) &= \tau_r - \tau_i \\
M_h \ddot{\theta}_h + J_h^T F_e + b_h \dot{\theta}_h + Q_h(\theta_h) &= \tau_h + \tau_i
\end{align*}
\]

Where \( \tau_i \) is the equivalent interaction torque between human and robot limbs, \( \tau_r \) - the robot torque, \( \tau_h \) - the human joint
torque and $M$ - the matrix of inertia, $b$ - the viscosity, $\tau_{FN_L}$ - the non-linear friction and $Q$ the gravity torque. $F_e$ are the external forces applied to the human and $J_h$ - the jacobian of the human. The subscripts $r$ and $h$ denote similar quantities related to robot and human systems, respectively. We are considering a quasi static interaction and we compensate the effect of gravity on the robot (Backdrivability ensures that $F_{NL}$ is low [3]), which give us:

$$\tau_h = -\tau + J_h^T F_e + Q_h(\Theta_h)$$ (2)

with :

$$\tau_r = \tau + Q_r(\Theta_r) + \tau_{FN_L}$$ (3)

B. EMG Recording

The methodology of intention detection is expected to be functional, easy to deploy, low-cost, operational in real-time and be capable of securing homogeneity between individuals. Homogeneity in our case means that, despite the fact that the features of the EMG signals vary depending on the user, our methodology should output the same value of intensity for equal intentions of different users.

The Myo-Armband (Thalmic Labs, Ontario) meets the previous specifications. The armband was positioned around the arm to capture biceps and triceps muscle activities. The Armband is composed of eight pairs of dry electrodes. We optimized the longitudinal placement to maximize the EMG signal empirically. The raw output of the armband is a zero-mean signal coded over 8 bit, it has no unit and is comprised between -126 and +127.

C. Algorithm

One could first consider using a calibration matrix as for the calibration of a 6-axis force-torque sensor. But, in this case, the non-linearity due to the model and due to EMG saturation leads to a complex and unstable parameter identification process. That is why a new algorithm is proposed here, more robust and simple, than the classical method. It is a mix of continuous and discrete methods. The discrete part enables to choose between continuous sub-parts similarly to [12]. The continuous sub-parts are simpler to calibrate than one using a fully continuous method. The continuous part brings flexibility and is designed to estimate the intention intensity. While the discrete part determines the intention direction. Moreover, this hybrid method also has the advantage of being directly exploitable for a control law. This methodology estimates the user’s intentions as a torque instead of a percentage of maximum contraction in order to have a user-independent output.

This process could be divided into two parallel blocks (Fig. 3): the first one is implemented to detect the direction of the action thanks to a classification designed in sec. IV. The second one is designed to estimate the intensity (see sec.V).

IV. DIRECTION ESTIMATION

A. Design

To detect the direction of the intention, an artificial neural network was designed for a four-class classification: rest, elbow extension i.e. triceps activation, elbow flexion i.e. biceps activation, and co-contraction. The co-contraction class has been dissociated from the others and did not play a role in the discrimination of the direction of intended movement. These four classes represent the main actions at the elbow joint (1 DoF).

The present approach is using time windows that were fed to an artificial neural network similarly to [8]. Unlike with other classification approaches like [7], [18], there was no need to handcraft features. The final architecture of the network can be seen on Fig. 5. The general idea was inspired by architectures from computer vision like the VGG net [19]. The convolutions were used to extract features from a sample and then the fully connected layers act as a classifier. In computer vision 2D convolutions are used, but in this case, for a time varying signal it is more relevant to use 1D convolutions and parallelize the processing over the 8 channels. The fact that the convolutions were parallelized helped to both reduce the number of trainable parameters (around 3000) and gain in generalization.

A time window of 30 samples was used, or 150 ms given the sampling rate of the sensors, and it was decided not to use overlapping windows to create the database from 30-second recordings in order to reduce the risk of overfitting. Our database consisted of 800 samples for each participant. A drop-out layer was used to avoid over-fitting.

The Network was implemented with Keras (Python deep learning library), the Adam optimizer was used for training [20], and the loss chosen was the categorical cross-entropy.
Fig. 5: Architecture of the neural network

(eq. (6)). Rectified Linear Units (ReLU) (eq. (4)) were used as activation function between the layers except for the last one, which was a softmax activation (eq. (5)).

\[
f_{ReLU}(x) = \begin{cases} 
0 & \text{if } x < 0 \\
 x & \text{else} 
\end{cases} \tag{4}
\]

\[
f_{softmax}(x_i) = \frac{e^{x_i}}{\sum_{c=1}^{M} e^{x_c}} \tag{5}
\]

\[
f_{loss}(p) = -\sum_{i=1}^{M} y_{o,i} \log(p_{o,i}) \tag{6}
\]

- M is the number of class
- y is a binary indicator (0 or 1) if class label i is the correct classification for observation o
- p is the predicted probability observation o is of class i

B. Training

To train the network designed in the previous section, 2 30-second-long data samples for each class were recorded. These recordings were used to create our training dataset and testing dataset (one recording per class for each).

The recording went as follows, twice, with rest between the steps:
- 30 seconds of biceps contraction at various intensities (see sec. V-C)
- 30 seconds of triceps contraction at various intensities (see sec. V-C)
- 30 seconds of rest
- 30 seconds of co-contraction at various intensities

The Table II displays the results for the 7 participants calculated with eq. (7). For each participant, only their own training data and testing data were used. An average precision of 96.9±3.1% was obtained. For this validation the placement of the electrode was the same as for the training.

\[
Prec = \frac{1}{N_{classes}} \sum_{i=1}^{N_{classes}} \frac{n_{true positive,i}}{n_{true positive,i} + n_{true negative,i}} \tag{7}
\]

Where \(n_{true positive,i}\) is the number of instances of the class i correctly estimated, \(n_{true negative,i}\) is the number of instances estimated as the class i wrongfully and \(N_{classes}\) is the number of classes.

V. INTENSITY ESTIMATION

In this section a way to extract a notion of intensity of the intention from the raw EMG signals is proposed. The value obtained in the end was expressed in a force unit, and therefore was not user-dependent (do not depend on, for example, the skin conductivity of each participant).

A. Pre processing

First, the rectified signals were summed from all the sensors (eqs. (8) and (9)). The spatial information was lost but it was already exploited by the classification part. Also, this enabled to take advantage of the cross-talk effect which increased the measure range. Indeed, when the electrode placed right upon the muscle saturated, the ones next to it were still able to sense changes in the intensity of the muscle activation.

\[
EMG_{rectif} = \text{abs}(EMG_i) \tag{8}
\]

\[
EMG_{rectif,sum} = \sum_{i=1}^{8} EMG_{rectif,i} \tag{9}
\]

Then the \(EMG_{rectif,sum}\) signal is filtered through a low-pass filter [21]. This filter is to be used in a real-time application, and, in the end, we want to use it for the command of a robotic device in a context of co-manipulation. So it has to have a low response time and little overshoot. Root mean square (RMS) and Mean filters are often used for smoothing an EMG signal [22], but there is a trade-off between response time and smoothing effect. When the size of the window is larger, the output is smoother but we introduce a greater lag. A Butterworth filter of order 2 and a 2 Hz cutoff frequency was chosen, and adapted to real-time with a sliding window approach, as it answered to our specifications.

B. Modeling

To ensure consistency among users, the model linking torque and low-pass filtered EMG presented in [23] will be used. And this would enable to design a user-independent control law despite using EMG signals.
\[
\tau = u^a \ast e^{(c-b*u)}
\]

\(\tau\) being the modeled torque, \(u\) the EMG signal and \(a,b\) and \(c\) the parameters to be determined for the biceps and triceps separately from calibration.

In the original article this model was calibrated in a controlled environment, there was one electrode placed precisely over the biceps and with the high sampling frequency (1 kHz) and the reference data was obtained by restraining the movements of the arm of the participants so that the load cell would record precisely the torque of the elbow. In this case, the use of this model is extended to two antagonist muscles, and with approximately-placed lower-sampling-rate electrodes.

C. Mapping

The first stage of the mapping aimed to obtain the torque maxima and corresponding processed EMG for the activation of the Biceps (elbow flexion or lifting gesture) and Triceps (elbow extension or putting down gesture). For this part, the arm of the exoskeleton was controlled with a PD-position-control law, with a high proportional gain, in order to have a stiff behavior and be in a quasi-static state.

We proceeded as follow, with a rest between each step:

- 5 seconds of biceps maximum contraction
- 5 seconds of triceps maximum contraction
- 6 steps of 8 seconds going gradually from 0 to 80% of biceps maximum contraction
- 6 steps of 8 seconds going gradually from 0 to 80% of triceps maximum contraction

To get the biceps maxima the participant grabbed the end of the arm and forced against the robot arm upward (Fig. 4c) as hard as he/she can and held it for a 5-second acquisition. The same approach was used for the Triceps maxima, except that the participant put his/her fist on the end of arm and pushed down (Fig. 4b). With this first stage, the value of the maximum torque the participant was able to apply on the actuator was obtained. These values were used to design torque step protocol: the robot applied levels of torques from gravity compensation up to 80% of the maximum and the user had to resist these torques. Each step lasted 8 seconds and was separated by a 1-second linear transition. This torque step protocol will ensure the precise levels of contraction are maintained by the participants.

Then, the first step for the Biceps part and the first two for the triceps part were removed from the data of the torque step protocol. Indeed these steps did not allow sufficient exertions from the participant (it was closer to a resting state). Data from the maxima of Biceps and Triceps were concatenated with their respective data from the torque step protocol. These concatenated data have been used as training data for the biceps activation and triceps activation classes (see sec. IV-B).

Finally, the parameters were estimated with a Levenberg-Marquardt non-linear regression [24]. For clarity, the results are presented for only one participant in Fig. 6. Two sets of three parameters were obtained, to characterize the activity of the biceps and the triceps. During the operating phase, the set of parameters to be used will depend on the classification phase, i.e. if it is an upward or downward movement intention.

On Fig. 7 the torque obtained using the parameters from the mapping and eq. (10) was plotted. The model was plotted up to the maximum processed EMG recorded for each participant (see sec. VI). First of all, EMG characteristics appeared quite heterogeneous between participants, which highlighted the need of a user-specific mapping. Secondly, for one of the participant, the graph is drastically decreasing at one point. This may be due to co-contraction phenomenon during the recording. A way to avoid that would be to check the level of co-contraction and asked the participant to redo the sample if it is above a certain threshold.

VI. EXPERIMENTAL TESTING

Seven participants were asked to contribute to the experiment. There were six men and one woman, all of them aged from 20 to 35 and in good health. After calibration (sec. IV and V), to test the present procedure, the following setup was proposed. During one minute, each participant performed varying torques in the 2 directions available. The exoskeleton was resisting them, like for the maxima recordings (see sec.
V-C). The fusion of the classification and intensity evaluation was realized as follow: the classes Biceps and Triceps were used to sign the intensity, positive and negative respectively. If a rest class is detected we assign the intensity the value 0. At the moment, the co-contraction class was handled in the the manner as the rest class.

On the Fig. 8a raw data obtained from one participant are presented. The class is obtained by taking the most likely output of the classifier, without threshold on the confidence score. It can be observed that the class (in green) presents important variations at the transitions. It can be explained by how the dataset was designed. Indeed, the whole spectrum of possibilities for each class were represented and at low intensities the differences became very subtle. For example, a low intensity co-contraction sample would have a signal very similar to a low intensity biceps contraction only or even a rest state. A way to stabilize the classification could be to consider one of the classes as the default state, rest in the present study, and put a threshold on the confidence score for the other to be accepted.

The processed results of 1 participant in Fig. 8b are presented. The estimated torque was used for the value and the class information to get the sign, i.e. positive for biceps class and negative for triceps, 0 otherwise, to calculate the estimated torque. An error of classification between the classes could lead to important mismatch and discontinuities of the estimation. Subsequently, the role of the classifier is crucial. One disadvantage associated to the classifier is the capacity to consider low intensity movements, e.g. areas where the classifier was hesitant on Fig. 8a resulted in small discontinuities on Fig. 8b.

In Table III, RMS errors were calculated. They were expressed in Newton at the end-effector. Indeed, in this study the actuator’s torque was used as the reference but results were converted for comparison purposes. We can see that, during the test experiment, an error of 3.8 ± 0.8 N on average was obtained, which is acceptable for applications focused on force augmentation for heavy load carrying. If we compare the results with [25], by converting their results as a force in the palm, they obtained on average 2.0 ± 0.2 N. The current error is twice as important but the use a load cell was bypassed and it did not require a precise setup of EMG electrodes.

TABLE III: Errors obtained during the validation experiment

<table>
<thead>
<tr>
<th>Person</th>
<th>RMS ( \tau_{\text{end-effector}} ) (N.m)</th>
<th>RMS ( F_{\text{end-effector}} ) (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.951</td>
<td>4.915</td>
</tr>
<tr>
<td>2</td>
<td>4.065</td>
<td>3.303</td>
</tr>
<tr>
<td>3</td>
<td>3.088</td>
<td>3.000</td>
</tr>
<tr>
<td>4</td>
<td>4.927</td>
<td>3.804</td>
</tr>
<tr>
<td>5</td>
<td>3.464</td>
<td>3.446</td>
</tr>
</tbody>
</table>

VII. CONCLUSIONS AND FUTURE WORK

In this article, an algorithm of intention detection oriented on the assistance for load carrying was presented. A new calibration method was proposed using a backdrivable exoskeleton that is time-efficient and practical. This research opens perspectives in industrial use cases. The current experiment showed that our methodology enabled to estimate the torque produced by a user.

The future of this project lays in exploring the possibility of a unified classifier which would help avoid manual subject-specific adjustment, i.e. a network trained on several participants which would be able to process data of new users. This would improve the time efficiency aspect of the proposed method. In addition, the robustness to the re-positioning of the armband between two sessions has to be formally evaluated. Moreover, the precision of our results could be improved by increasing the acquisition rate of EMG electrodes.

The calibration procedure presented in the current study has been performed in static conditions. As the load carrying assistance will be provided by the exoskeleton during dynamical tasks, it is expected that the control law will be efficient for dynamic tasks with low velocity [26]. However, with high dynamics tasks stability issues will probably need to be addressed by the control law. Other researches will consist in implementing controllers based on the output of this module of intention detection, and experimenting on the effect of the assistance of our exoskeleton in carrying various loads at various speeds. Despite the heterogeneous aspect of EMG signals, the proposed calibration method will ensure the reproduction of the same behavior of the controllers with different users.

REFERENCES
