AN ALGORITHM CALIBRATED WITH CATEGORICALLY LABELLED EMG FOR END-TO-END ESTIMATION OF CONTINUOUS HAND KINEMATICS

Alexander E. Olsson[1], Nebojša Malešević[1], Anders Björkman[2], Christian Antfolk[1]

[1]Dept. of Biomedical Engineering, Lund University, [2]Dept. of Hand Surgery, Lund University

ABSTRACT

To restore limb functionality, control of a prosthetic hand should ideally be (I) proportional, i.e. produce speeds which varies in conjunction with changes in the latent intensity of muscle contractions, and (II) simultaneous, i.e. allow for both combined and independent steering of relevant kinematic degrees of freedom (DoFs). These desiderata are not straightforwardly attainable with classificatory pattern recognition applied to surface electromyography (sEMG), which only allows for the detection of a finite set of categorically encoded gestures. To alleviate such limitations, we here introduce a related approach for myocontrol which maps sEMG envelopes directly to multiple, continuously encoded DoFs, providing proportionality and simultaneity implicitly. The proposed method, termed myoelectric representation learning (MRL), is constituted by a deep learning topology and a domain-informed model training scheme. As with conventional pattern recognition, MRL operates on sEMG exclusively and is calibrated without ground truth limb kinetics, allowing for deployment with amputee users. We demonstrate the practical viability of MRL by implementing a virtual control interface driven by a setup consisting of 8 surface electrodes and capable of decoding 2 kinematic DoFs in real-time. Experiments with 10 healthy subjects, in which the interface was used to conduct tests yielding 5 numeric performance metrics, were performed to quantify the quality of myoelectric control afforded by MRL. Comparisons with the performance obtained from of a Linear Discriminant Analysis benchmark method on an identical test revealed that MRL outperforms the former in all computed measures of control efficacy.

INTRODUCTION

Pattern recognition applied to surface electromyography (sEMG) has for a time been considered a key component in the endeavour to make intuitively controlled, multiarticulate upper limb prostheses available to transradial amputees [1]. Despite countless reports of successful application of several variations of the technology in lab environments, widespread clinical adoption remains elusive [2]. Due to the notable level of reliability and stability required for practical viability, the few commercial implementations existing currently [3] make use of linear classification algorithms applied to a robust set of handcrafted signal features [4]. Within this *gesture detection* framework, speed of motion is typically modulated separately from classification by use of the mean average value of sEMG aggregated across all available channels [5]. Albeit functional and robust, this type of approach does not allow for true *simultaneity*, here defined as the ability to separately control multiple kinematic degrees of freedom (DoF) with mutually independent speeds.

This paper introduces an alternative method for intuitive, proportional, and simultaneous myoelectric control which functions via supervised machine learning and is constituted by (I) a computationally lightweight artificial neural network (ANN) and (II) an appertaining calibration strategy. Due to its reliance on kinematically influenced signal representations arising throughout the ANN model during use, the method is termed *Myoelectric Representation Learning* (MRL).

METHODS

10 able-bodied subjects (age range 26-49 years, 5 male and 5 female) participated in the current study, which consisted of two phases: acquisition of calibration data followed by evaluation of myocontrol efficacy. The study was approved by the Regional Ethical Review Board in Lund, Sweden and all subjects gave their written consent. Data acquisition and processing were performed with custom code written for and executed in Python 3.6. All hyperparameters were selected *ad-hoc* prior to the start of experiments via empirical work on subjects not part of the current study

Data Acquisition

sEMG signals were acquired with a Myo armband (Thalmic labs, Canada) consisting of 8 equiangularly spaced dry surface electrodes. At the start of each experiment session, the armband was placed enclosing the dominant forearm of the subject at a level approximately 1/3 of the distance from the humeroradial joint to the radiocarpal joint. sEMG signals were sampled at a rate of 200 Hz and were transferred at identical rate to a host desktop computer (on which all signal processing was performed) in real-time via Bluetooth. The subject was seated comfortably in a chair, approximately 1 m from the computer screen, with elbow resting on a table; the angle and position of the elbow could be varied freely by the subject at all times.

Table 1. The recorded calibration movements and their corresponding categorical target encodings.

Movement Class	Description	Ternary Encoding $\boldsymbol{y}$
0	*Rest*	[0, 0]
1	*Wrist flexion*	[-1, 0]
2	*Wrist extension*	[1, 0]
3	*Flexion of the digits*	[0, -1]
4	*Extension of the digits*	[0, 1]
5	*Wrist flexion and Flexion of the digits*	[-1, -1]
6	*Wrist flexion and extension of the digits*	[-1, 1]
7	*Wrist extension and flexion of the digits*	[1, -1]
8	*Wrist extension and extension of the digits*	[1, 1]

The current study entailed the decoding of two DoFs: (I) wrist flexion/extension and (II) flexion/extension of all digits simultaneously. Movement instruction stimuli were encoded with a ternary scheme, where each DoF could assume the values -1 (DoF active in one direction), 0 (DoF inactive), or 1 (DoF active in the opposite direction). All of the resulting 3^2=9 combinations possible in this framework (shown in table 1) were recorded. Prior to calibration data acquisition, subjects were instructed to perform each of the 8 nonrest movements classes with maximal voluntary contraction (MVC) for 5 seconds. This step served to familiarize the subject with the movement combinations under consideration and was furthermore used to compute an MVC magnitude value specific to each subject and movement by summing the mean absolute value over all 8 sEMG channels.

Calibration data was recorded by an acquisition program which prompted the subject to perform all nonrest movements for 3 repetitions, each lasting for a duration of 5 s and separated by 3 s of rest. To aid the subject in applying a sustainable and consistent level of contraction across movements, the mean absolute value of the sEMG signal, summed over all channels of a sliding window of length 0.5 s, was mapped to the height of a bar shown in real-time on the computer screen together with a threshold set to equal 50% of the movement-specific MVC magnitude computed earlier; subjects were instructed to keep the activity level as close to the threshold as possible. Once the program concluded, recorded sEMG was, together with the concurrent movement instruction stimuli information, saved and subsequently used for calibration of two different myoelectric control methods (an example of such calibration data is provided in fig. 1).

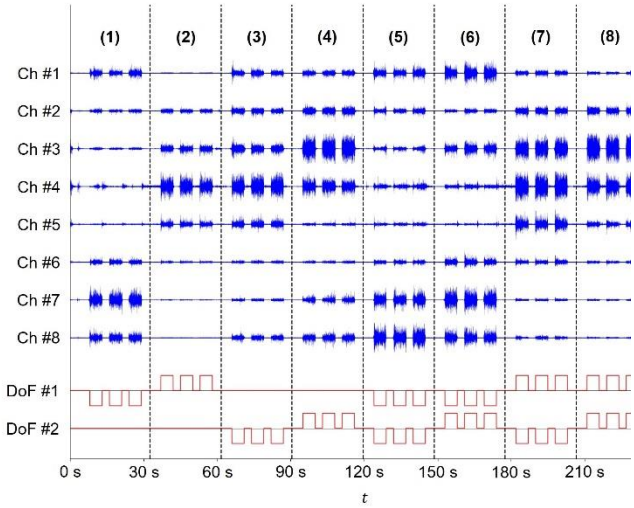

Figure 1. sEMG calibration data acquired from a single subject. **(1)** Wrist flexion. **(2)** Wrist extension. **(3)** Flexion of the digits. **(4)** Extension of the digits **(5)** Wrist flexion and flexion of the digits. **(6)** Wrist flexion and extension of the digits. **(7)** Wrist extension and flexion of the digits. **(8)** Wrist extension and extension of the digits.

Myoelectric Representation Learning

Before being applied for neural network training, the previously collected sEMG signals were subject to preprocessing in the form of an *envelope extraction* step followed by a *nonlinear rescaling* step. Envelope extraction entailed signal rectification and channel-wise lowpass digital filtering with a moving average filter of length 0.5 s (100 samples), yielding a nonnegative and unbounded signal matrix $\boldsymbol{E}^u$. Nonlinear rescaling entailed channel-wise linear rescaling, clipping and lastly transformation by the square root operator as is shown in equations 1 and 2 below.

$$E^r_{i,t} \leftarrow \frac{E^u_{i,t} - p_i^{1\%}}{p_i^{99\%} - p_i^{1\%}} \tag{1}$$

$$E^{Tr}_{i,t} \leftarrow \sqrt{max(0, min(1, E^r_{i,t}))} \tag{2}$$

$p_i^{1\%}$ and $p_i^{99\%}$ were the 1st and 99th percentile level, respectively, of the samples of the ith channel of $\boldsymbol{E}^u$. These preprocessing steps (I) guaranteed that all samples in $\boldsymbol{E}^{Tr}$, which were to be used for optimization, were constrained to the interval [0, 1] and (II) limited the impact of outlier sEMG samples on the resulting training data. The square root operator was included to bias resolution towards high levels of muscle contraction (i.e. $E^r_{i,t}$ close to 1). When the system later operated in real-time inference mode, input sEMG was identically processed using online filtering and the statistics $\boldsymbol{p}^{1\%}$ and $\boldsymbol{p}^{99\%}$ obtained from the calibration data.

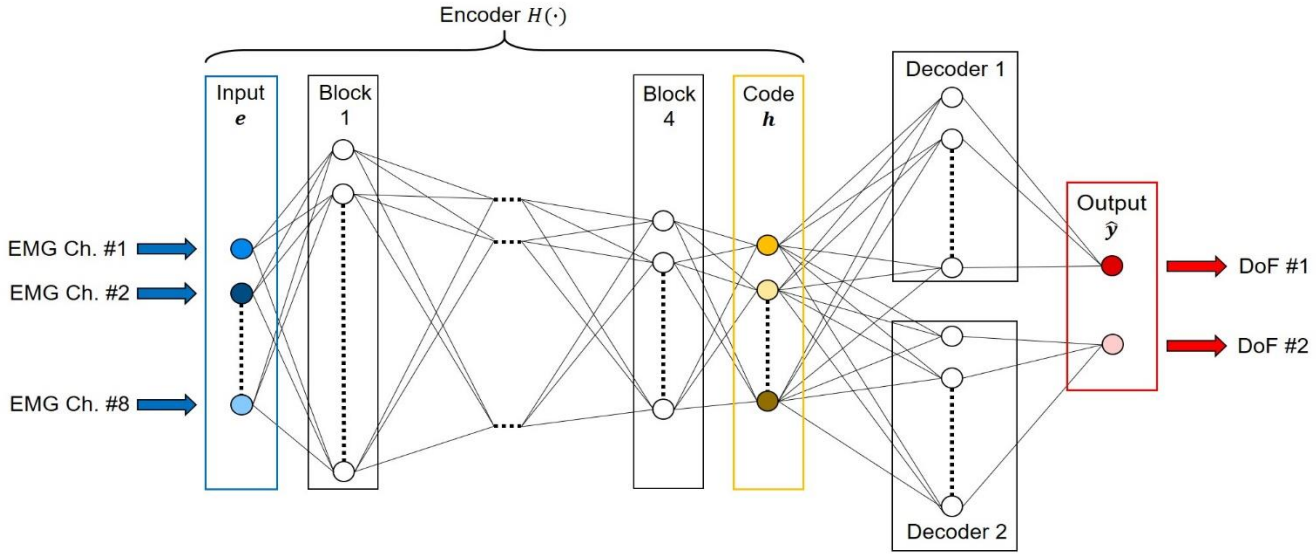

Figure 2. The regression neural network topology central to the presented approach.

The MRL topology, presented schematically in fig. 2, was composed of an *encoder subnetwork*, shared between the DoFs, and two separate *decoder subnetworks*, each specific to a DoF. The encoder network consisted of 5 fully connected *blocks*, each in turn consisting of a fully connected layer [6], a leaky ReLU activation [7], and layer normalization[8]. The number of output nodes for each encoder block was set to 128, 64, 32, 16, and 8, respectively, resulting in a *code size* of 8. Each decoder network operating on the generated signal representation consisted of one fully connected block of the same type, utilizing 128 hidden units, terminating in a fully connected layer with 1 linear output node, representing the inferred level of activity for one of the decodable DoFs.

Model training was performed via gradient descent with batch size of 4096 for 5000 iterations by the Adam algorithm [9] with $\eta = 10^{-4}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$. The loss to be iteratively minimized was given by equation 3.

$$\mathcal{L} = \mathcal{L}_i + \alpha_c \mathcal{L}_c \tag{3}$$

α_c is a hyperparameter, set to equal 10^{-2}. $\mathcal{L}_i$ is referred to as the *inference loss* and given in equation 4.

$$\mathcal{L}_i = \|\hat{\boldsymbol{y}} - \boldsymbol{y}\|_1 \tag{4}$$

The regressand $\hat{\boldsymbol{y}}$ is the 2-element vector containing the DoF-wise continuous kinematics inferred by the ANN and $\boldsymbol{y}$ is the ground truth ternary encoding of the movement instruction stimuli concurrent with the sEMG envelope regressor sample. With $\mathcal{L}_i$ minimized, the ANN produces output which matches the movement intent of the subject.

$\mathcal{L}_c$ denotes the *contractive loss* and given in equation 5.

$$\mathcal{L}_c = \frac{1}{2 \cdot 8} \sum_{i}^{8} \sum_{j}^{2} \left(\frac{\partial \hat{y}_j}{\partial e_i} \right)^2 \tag{5}$$

The term $\frac{\partial \hat{y}_j}{\partial e_i}$ denotes the gradient of the jth output DoF with regards to the ith channel of the input sEMG envelope. With a minimized $\mathcal{L}_c$, ANN output will be sensitive to variations in the level of latent muscle activity (as proxied by the sEMG envelope), i.e. control will be proportional.

Benchmark Pattern Recognition Control

To verify the conjectured advantages of MRL, a benchmark proportional pattern recognition method for myocontrol based on linear discriminant analysis was implemented. All implementation details, including feature extraction, classifier architecture, and calculation of speed, were selected to be identical of those of Method 2 introduced by Scheme et al in [5]. The method is in its entirety henceforth referred to simply as LDA.

Quantitative Method Evaluation

A real-time virtual environment was implemented to quantify myocontrol efficacy for both MRL and LDA. To counteract confounding effect from acclimation, half of subjects were selected to evaluate MRL first whereas the other half were selected to evaluate LDA first (determined randomly). The output command of the evaluated method was mapped to the velocity of a *cursor* shown on the computer screen. Detection of wrist flexion/extension translated to cursor movements left/right, and detection of flexion/extension of digits translated to cursor movements down/up. In the test, subjects were instructed to steer the cursor towards a sequence of circular *targets*, generated at 20

Table 2. Summary of real-time performance metrics.

Name (abbreviation)	Description
Completion rate (CR)	The proportion of targets which were successfully reached.
Completion time (CT)	The average time elapsed between task start and completion
Path Efficiency (PE)	The average ratio between the straight-line distance from the starting point to the target and the actual distance traversed.
Overshoot (O)	The average number of occurrences wherein the cursor leaves the target prior to the end of the dwell time
Throughput (T)	The ratio $\frac{ID}{CT}$ between index of difficulty (ID) and completion time (CT), averaged across all successfully reached targets.

locations spanning all four quadrants with 2 radii, resulting in a set of 40 targets each covering either 0.6% or 2.3% of the total screen area. The order in which targets were presented was determined randomly for each subject. An index of difficulty ID was computed for each target as in [10]. As in earlier work [11], targets were considered successfully reached after a dwell time of 0.3 s and considered failed if not successfully reached within 20 s. The 5 performance metrics introduced by Williams and Kirsch in [10] (summarized in table 2) were computed for each subject and control method at the end of experiments.

RESULTS

Linear regression showed a strong relationship between ID and CT ($R^2 = 0.89$) for MRL across all subjects and targets. The corresponding value for LDA was computed as $R^2 = 0.81$, verifying the eligibility of the Fitts's law test and by extension the validity of the throughput metric. Aggregated performance metric summary statistics of both MRL and LDA from all subjects are presented in table 3.

CONCLUSIONS

The proposed algorithm (MRL) was found to be superior to conventional pattern recognition (LDA) in the sense of outperforming the latter in all computed measures of real-time efficacy of control. These results are encouraging, but need to be replicated with a larger subject sample size (ideally including amputee subjects) as well as have their stability over longer time spans be investigated.

Table 3. Means and standard deviations of metrics.

Metric	MRL	LDA
CR	99.25 ± 1.60	98.00 ± 2.45
CT	3.68 ± 1.14	5.25 ± 1.43
PE	55.33 ± 10.83	49.93 ± 7.90
O	0.53 ± 0.19	0.61 ± 0.26
T	0.67 ± 0.15	0.51 ± 0.13

The MRL model was successful in extracting kinematics pertaining to two separate DoFs, but required calibration data of every possible movement combination. For larger numbers of DoFs, the number of movement combinations grows geometrically, leading to infeasibly long calibration data acquisition phases. Notably, this drawback is not unique to MRL, but is shared by all contemporary pattern recognition frameworks aimed at multiarticulate myocontrol

ACKNOWLEDGEMENTS

This work was supported by the Promobilia Foundation, the Crafoord Foundation, and the European Commission under the DeTOP project (LEIT-ICT-24-2015, GA #687905).

REFERENCES

[1] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *J. Rehabil. Res. Dev.*, 2011.

[2] N. Jiang, S. Dosen, K. R. Müller, and D. Farina, "Myoelectric Control of Artificial Limbs-Is There a Need to Change Focus?," *IEEE Signal Process. Mag.*, 2012.

[3] www.coaptengineering.com

[4] B. Hudgins, P. Parker, and R. N. Scott, "A New Strategy for Multifunction Myoelectric Control," *IEEE Trans. Biomed. Eng.*, 1993.

[5] E. Scheme, B. Lock, L. Hargrove, W. Hill, U. Kuruganti, and K. Englehart, "Motion normalized proportional control for improved pattern recognition-based myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 149–157, Jan. 2014.

[6] Y. Bengio, I. Goodfellow, and A. Courville, *Deep Learning*, 2015.

[7] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier Nonlinearities Improve Neural Network Acoustic Models," *arxiv preprint*, 2013.

[8] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer Normalization," *arXiv preprint*, 2016.

[9] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015*

[10] M. R. Williams and R. F. Kirsch, "Evaluation of head orientation and neck muscle EMG signals as command inputs to a human-computer interface for individuals with high tetraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 485–496, Oct. 2008.

[11] N. Jiang, H. Rehbaum, I. Vujaklija, B. Graimann, and D. Farina, "Intuitive, online, simultaneous, and proportional myoelectric control over two degrees-of-freedom in upper limb amputees," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 501–510, 2014.