EMG-BASED PREDICTION OF MULTI-DOF ACTIVATIONS USING SINGLE-DOF TRAINING: A PRELIMINARY RESULT

Claudio Castellini and Markus Nowak
Robotics and Mechatronics Center
DLR – German Aerospace Center
Weßling, Germany

INTRODUCTION

In this paper we propose an advancement to the problem of simultaneous, proportional myocontrol of hand/wrist prostheses [1,2,3]. In particular, we address the prediction of simultaneous activations of multiple degrees of freedom (DOFs) by training a machine learning method on single-DOF activations only – for example, correctly predicting simultaneous flexion of the index and thumb by training on index flexion and thumb flexion only. In myoelectric control this is a very desirable property, since training on single-DOFs only will in general not correctly predict multiple-DOF activations; on the other hand, directly gathering multiple-DOF activation data from the subject quickly becomes unfeasible as the number of DOFs grows.

So far, to the best of our knowledge, the only successful approach to this problem is represented by the application of Non-negative Matrix Factorisation to two/three DOFs of a prosthetic wrist [4]; we hereby propose an alternative approach which is able to solve the problem for single-finger activations. Surface electromyography (sEMG) data are firstly collected for single-finger forces; the data set is then augmented with artificial sEMG clusters representing multiple-DOF activations; lastly, a machine learning method is trained on the augmented data set. The augmentation procedure works by linearly combining the single-DOF sEMG clusters and is therefore called Linearly Enhanced Training (LET).

To preliminarily validate the procedure, an experiment was conducted on seven intact subjects engaged in the production of a simple, repetitive single-DOF activation pattern (thumb adduction, index flexion and little finger flexion). The results are very promising.

PROBLEM STATEMENT

Simultaneous and proportional myocontrol

Simultaneous and proportional myocontrol of a prosthetic or rehabilitation device [6] entails that a (disabled) human subject can control its *m* DOFs independently, at the same time, and in a "graded" fashion, that is, according to the desired level of activation. For instance, if each DOF can be controlled in torque, we must build a human-machine interface consisting of *d* sEMG electrodes and *m* approximant functions *f* such that $\tau \approx f(\boldsymbol{x})$ where $\tau \in \mathbb{R}$ is the required torque at the DOF and $\tau \in \mathbb{R}^d$ is the reading from the electrodes. (In this simplified framework we intentionally blur the distinction between the muscular activation and the DOF activation expressed as a torque command.) Machine learning is usually employed in the regression mode (e.g., Support Vector Regression [7,8] or Non-negative Matrix Factorisation [4]) to build the approximant functions from a set $X=\{\boldsymbol{x}_i,\boldsymbol{\tau}_i\}_{i=1}^{n}$ of (sample,target) pairs previously collected from the subject – the so-called *training set*. (Notice that in the training set one needs to have one target value per each DOF, hence $\boldsymbol{x}_i \in \mathbb{R}^d$ and $\boldsymbol{\tau}_i \in \mathbb{R}^m$.)

The usage of machine learning has the advantage of allowing *natural* control. The training set *X* is built by inducing the subject to activate one DOF (e.g., flexing a finger, pronating the wrist, etc.) and recording the corresponding sEMG values. If the input/output relationship is fairly represented by the values in *X*, then each *f* will correctly approximate the torque $\boldsymbol{\tau}$ required to control the corresponding DOF; moreover, since *X* was collected from the subject while engaged in performing the actions corresponding to the activations of each DOF, the resulting approximant will command to each DOF the intended torque – hence the term natural control, or *intent detection*.

Clearly, in order for this approach to be feasible, an appropriate sampling of the input space *for each DOF considered* is required; target values can be either gathered using a torque/force sensor, or more realistically, they can be arbitrarily set at 0 or 1 whenever, in turn, a DOF is not active or maximally active. (Recall that in general the subject is an amputee who cannot produce any reliable ground truth in principle.) Following the "realistic approach" outlined in [5], this corresponds to such a training set:

$$X=\{(X_0,\boldsymbol{\tau}_0),(X_I,\boldsymbol{\tau}_1),\ldots,(X_m,\boldsymbol{\tau}_m)\}$$

where each subset $(X_k,\boldsymbol{\tau}_k)$ corresponds to the (sample,target) pairs collected when only the th DOF is active, and all others are inactive. (The subscript 0 denotes the resting state, in which all DOFs are inactive.) For example, consider the case in which the selected DOFs are the flexion of the index (I) and little (L) finger; the corresponding training set, denoted with *sf* for single-finger activations, is

$$X_{sf} = \{(X_0,(0\ 0)),(X_I,(1\ 0)),(X_L,(0\ 1))\}$$

where X_I and X_L denote sEMG samples collected, in turn, when either the index or the little finger was maximally active. Given an appropriate machine learning method, two functions $f_I(\boldsymbol{x})$ and $f_L(\boldsymbol{x})$ trained on X_{sf} (namely, $f_I(\boldsymbol{x})$ would be trained using the first component of each $\boldsymbol{\tau}_k$ as target values, and $f_L(\boldsymbol{x})$ would be trained on the second) will return a sensible approximation of the torques required at the index and little finger *whenever either of the two fingers, or none of them, is active*.

Multi-DOF activations

Indeed, the above method will not generalise to the case in which both DOFs are active at the same time – simultaneous flexion of the index and little finger: the sEMG signal corresponding to a multi-DOF activation, call it X_{IL} in the example above, has, in general, no trivial relationship to those obtained for the single-DOF activations it is composed of. Nevertheless, being able to estimate multi-DOF activations is very desirable: e.g., while grasping, many fingers are active at the same time; while reaching with the aid of a prosthetic wrist+hand, the device must flex, pronate and grasp simultaneously.

Traditionally (see, e.g., [9]), this problem has been solved by directly gathering from the subject the sEMG signals corresponding to the required multi-DOF activation(s) – in the above case, X_{IL} would be available, and a new training set (denoted *mf* for multi-finger)

$$X_{mf} = \{(X_0,(0\ 0)),(X_I,(1\ 0)),(X_L,(0\ 1)),(X_{IL},(1\ 1))\}$$

could be used to determine the *f*s. This method will yield the expected approximants, but becomes quickly unfeasible as the number of DOFs, *m*, grows, since the number of possible combinations grows exponentially with it. (The most advanced hand prosthesis in the world at the time of writing, the *i-LIMB Ultra Revolution* by Touch Bionics, see www.touchbionics.com, has $m = 6$, which becomes 7 or 8 if a self-powered prosthetic wrist is additionally used.)

An alternative way is that of estimating X_{IL} from X_I, X_L and/or X_0; that is, trying to build a machine which will *generalise to multi-DOF activations although it has been trained on single-DOF activations only*. The only attempt so far at solving this problem, as far as we know, appears in [11] for two DOFs of the wrist plus hand opening/closing. In this work, *Non-negative Matrix Factorisation* trained on X_{sf} yields a model acting both as a linear predictor of the required activations *and* as a linear "un-mixer" of multi-DOF activation signals into single-DOF ones. Although we have no comparative results so far, we speculate that NMF will hardly generalise to the case of single fingers, for which a linear approach has been shown to produce unacceptably low prediction accuracy [5].

We rather propose to artificially augment X_{sf} in order for it to enable the desired generalisation by any machine learning method trained on it (possibly non-linear). We therefore look for a function $\mathcal{F}$ such that $\mathcal{F}(X_0, X_I, X_L)$. If such a function is available, then an "enhanced" training set $\boldsymbol{X'}$ can be built out of X_{sf},

$$X' = \left\{ \begin{array}{c} ((X_0,(00)),(X_I,(1\ 0)),(X_L,(0\ 1)), \\ (\mathcal{F}(X_0,X_I,X_L),(1\ 1)) \end{array} \right\}$$

such that training on $\boldsymbol{X'}$ will yield the required approximants. Notice that $\boldsymbol{X'}$ is built with *no* explicit knowledge of X_{IL}, avoiding the above-described exponential blowup of training time and effort.

LINEARLY ENHANCED TRAINING (LET)

A very simple idea to build such an $\mathcal{F}$ is that of considering the multi-DOF activation signal as a *linear combination* of the single-DOF signals involved in it. This hypothesis seems reasonable since both sets of motor units involved in the single-DOF activations must participate simultaneously in the multi-DOF activation, to different degrees; we will also assume that the multi-DOF activation samples lies somewhere on the vector in $\mathbb{R}^d$ bisecting the two vectors corresponding to the single-DOF activations:

$$\mathcal{F}(X_0,X_I,X_L,\alpha) = \{x \mid x = \alpha[(x_I - x_0) + (x_L - x_0)], \\ \forall x_0 \in X_0, x_I \in X_I, x_L \in X_L\}$$

where α, for which we assume , must be found by exhaustive search. This procedure adds to the original training set one cluster of linearly-built artificial sEMG samples per each DOF combination, and is therefore called Linearly Enhanced Training (LET). Notice that LET is in principle applicable to any *k*-ary combination of single-DOF activations (not only pairs), and to an arbitrary number *m* of them, in which case 2^m parameters α must be found;

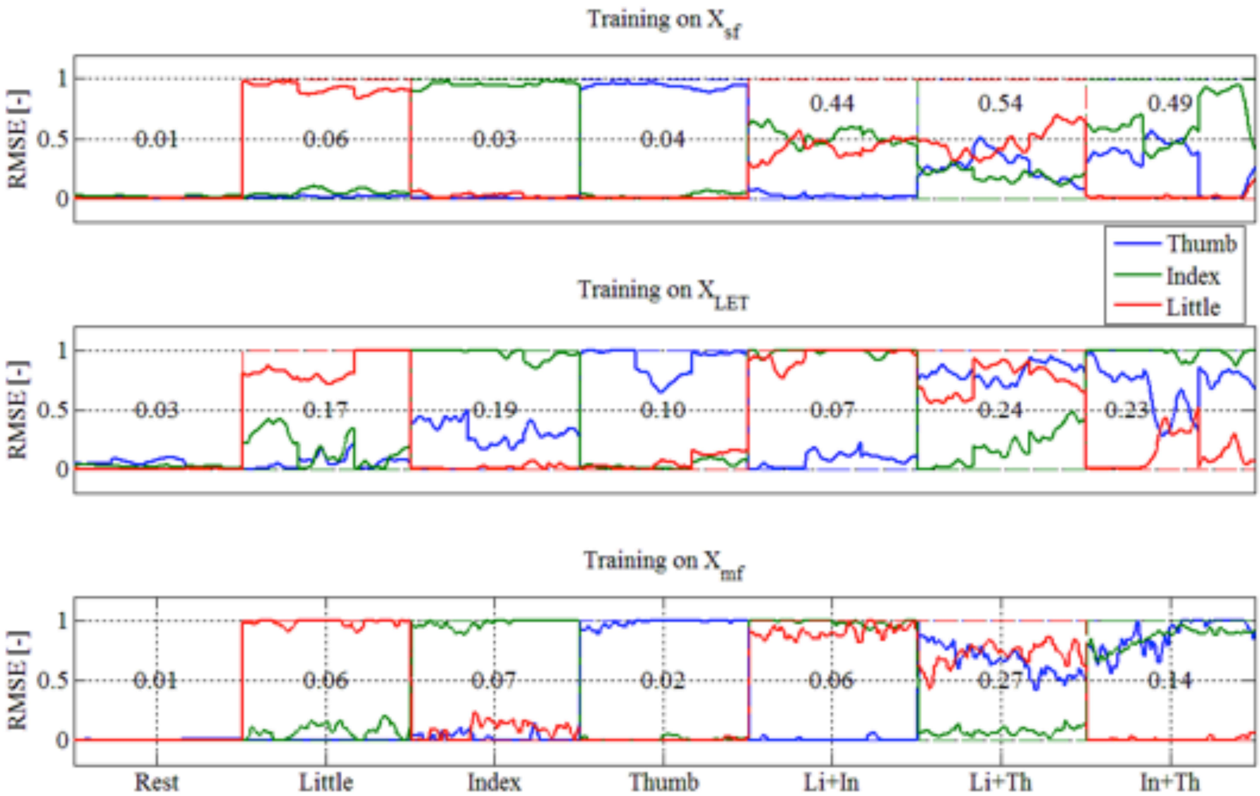

Figure 1: NRMSE for a typical subject for each DOF, DOF combination and training set.
Each block shows the error evaluated over the three repetitions of the related single- or multi-DOF activation.

moreover, it is independent of the machine learning method of choice. Notice, however, that the LET-enhanced training set X_{LET} is exponentially larger than X_{sf} (but still just as large as X_{mf}), which could be problematic in case the training heavily depends on its size.

EXPERIMENT DESCRIPTION

In order to partially validate the LET procedure, we set up a simple psychophysical experiment, stimulating human subjects to apply 3 single-DOF activations, plus all *pairs* of them, while recording their sEMG signals; we then compared the prediction accuracy of a known regression method trained, in turn, on X_{sf} and X_{LET}; for further comparison, the accuracy obtained by training on X_{mf} was also evaluated. We expected the performance obtained using X_{LET} to lie somehow in the middle between those obtained using X_{sf} and X_{mf}.

Notice that in this preliminary experiment we *do* use the explicit knowledge of X_{IL} in order to estimate X_{LET}, with the hope of finding that the required coefficients α can be treated as invariants across multi-DOF activations and subjects.

<u>Subjects</u>

Seven healthy human subjects (age 23÷42yrs, 6m/1w) were recruited for the experiment. Each subject received a thorough description of the experiment; informed written consent was obtained from all participants. Experiments with sEMG have been approved by the Ethical Committee of the DLR.

<u>Materials and methods</u>

The sEMG signal was measured using ten *MyoBock 13E200* electrodes by Otto Bock (www.ottobock.com), uniformly placed around forearm close to the elbow, using an elastic biocompatible adhesive bandage. These electrodes provide an amplified, band-pass filtered and rectified signal. To reduce noise a Butterworth filter of 1st order is applied with cut-off frequency of 1.5Hz. The sEMG data was collected at approx. 46Hz using a standard analog-to-digital conversion card connected to a Windows machine via Ethernet. (The setup closely follows that of [5] – the interested reader is referred once again to that paper.)

Experimental protocol

Each subject was comfortably seated in front of a table with a large monitor, on which two 3D hand models were shown, one acting as a visual stimulus (i.e., what the subject was required to do) and the other showing the predicted forces as finger flexions. The experiment was divided into three sessions (rest was allowed in between sessions).

The first session consisted of three repetitions of, in turn, little finger flexion, index finger flexion, thumb adduction, little and index finger, little and thumb, thumb and index – that is, three single-DOF activations and three multi-DOF ones. Data gathered during the single-DOF activations are while the union of X_{sf} and data gathered during the multi-DOF activations are X_{mf}. The collected data was used to determine the three coefficients – one for each multi-DOF activation – by minimising the Euclidean distance between the artificial samples in $\mathcal{F}$ and the “true” samples in X_{mf}. For example, for the index and little finger

$$\alpha_{IL} = \arg\min_{\alpha} ||\mathcal{F}(X_0, X_I, X_L, \alpha) - X_{IL}||^2$$

Using the αs found this way, X_{LET} was built and a non-linear, incremental regression method was then trained using, in turn, X_{sf}, X_{mf} and X_{LET}. The chosen method was Ridge Regression with Random Fourier Features [12], which we have already successfully employed in [5]. This method requires finding three hyperparameters λ, D and σ, two of which (λ and D) were set at standard values of 1 and 700 (see [5] again), whereas one optimal value of σ for each training set was found by grid search and 3-fold leave-one-repetition-out cross-validation over each related training set.

In the second and third session, the prediction was started using, in turn, the model obtained by training on X_{sf} and on X_{LET}; each subject was then again shown, using the stimulus hand, the same DOF activations as in the first session, and instructed to have the prediction hand reproduce them and to keep them stable for 3 seconds. Online testing on X_{mf} was neglected in order to keep the experiment as short as possible; rather, the performance using X_{mf} was evalauted offline by training on the first two repetitions and testing on the third.

EXPERIMENTAL RESULTS

The optimal values of α were determined to lie in the range 0.7±0.2 across all multi-DOF activations and subjects (mean plus/minus one standard deviation). The Root Mean-Squared Error (RMSE) was calculated for each of the 7 different activations (rest, 3 single-DOF, 3 multi-DOF) for each subject, for each training set and for each single- or multi-DOF activation. As it happens in [5], the ground truth is represented by the visual stimulus values, ranging from 0 to 1; the RMSE is therefore expressed in arbitrary units.

Figure 1 shows the results for one typical subject. As expected, the prediction error on single-DOF activations (and rest) is good to excellent in all three cases with a slightly worse result obtained while training on X_{LET}, whereas the error on multi-DOF activations is high when using X_{sf} (0.44, 0.54 and 0.49 for little+index, little+thumb and index+thumb in turn) and reasonably good when using X_{LET} (0.07, 0.24 and 0.23). Surprisingly, the error when using X_{mf} is on average just a little better than when using X_{LET} (0.06, 0.27 and 0.14).

Figure 2 shows the results averaged over all subjects. The trend is confirmed: with respect to training on X_{sf}, training on X_{LET} makes the error slightly worse for single-DOF activations but largely better for multi-DOF activations; and surprisingly, training on X_{mf} does not yield a considerably large improvement.

CONCLUSIONS AND DISCUSSION

The LET procedure, presented in this paper, enables in principle any machine learning method to predict multi-DOF activations using data collected during single-DOF activations only. LET works by approximating the multi-DOF activation sEMG signals, which are unfeasbile to gather directly, using a linear combination of the related single-DOF signals. In a psychophysical experiment, using the LET technique, a standard machine learning method was able to obtain prediction error values on multi-DOF activations similar to those obtained on single-DOF activations.

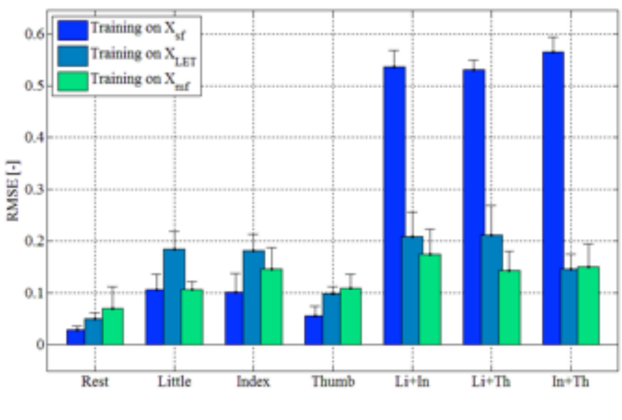

Figure 2: RMSE per training set per DOF
Mean and standard error of the mean across all subjects

We consider this a preliminary result, both because the number of subjects and DOFs considered is low, and because multi-DOF activations were explicitly used to build . However, as the values of the s lie in a quite close range for all subjects, we plan to enforce LET in a completely automatic way, and compare again its performance against the usage of and . Although initial though, this result looks promising and future work includes chacking whether it generalises to, e.g., the (combined) DOFs of the wrist, possibly in combination with a few grasping postures.

ACKNOWLEDGEMENTS

Work partially supported by the Swiss National Science Foundation Sinergia project #132700 NinaPro (Non-Invasive Adaptive Hand Prosthetics) and by the FP7 project The Hand Embodied (FP7-248587).

REFERENCES

[1] Jiang N., Dosen S., Müller K. R., Farina D. (2012) *Myoelectric control of artificial limbs – Is there a need to change focus?* IEEE Signal Processing magazine, pp. 12–15.

[2] Pilarski P. M., Dick T. B., Sutton R.S. (2013) *Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints.* Proc. IEEE Int. Conf. Rehab. Robot., pp. 1-8.

[3] Kamavuako, E.N., Rosenvang, J.C., Horup, R., Jensen, W., Farina, D. and Englehart, K.B., *Surface versus untargeted intramuscular EMG based classification of simultaneous and dynamically changing movements* IEEE Trans Neur Syst Rehab Eng, 21(6), pp. 992-998, 2013.

[4] Jiang, N.; Rehbaum, H.; Vujaklija, I.; Graimann, B.; Farina, D., *Intuitive, online, simultaneous and proportional myoelectric control over two degrees of freedom in upper limb amputees,* IEEE Trans Neur Syst Rehab Eng, in press.

[5] Gijsberts A., Bohra R., Sierra González D., Werner A., Nowak M., Caputo B., Roa M. A. and Castellini C. (2014) *Stable myoelectric control of a hand prosthesis using non-linear incremental learning.* Front. Neurorobot. 8:8.

[6] Fougner, A., Stavdahl, Ø., Kyberd, P. J., Losier, Y. G., and Parker, P. A. (2012). *Control of upper limb prostheses: terminology and proportional myoelectric control - a review.* IEEE Trans. Neural Syst. Rehabil. Eng. 20, 663–677.

[7] B. E. Boser, I. M. Guyon, and V. N. Vapnik, *A training algorithm for optimal margin classifiers,* in Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, D. Haussler, Ed. ACM press, 1992, pp. 144–152.

[8] V. N. Vapnik, *Statistical Learning Theory.* New York: John Wiley and Sons, 1998.

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

[10] Kõiva, R.; Hilsenbeck, B. & Castellini, C. *FFLS: An accurate linear device for measuring synergistic finger contractions,* Proceedings of IEEE Engineering in Medicine and Biology Society, 2012, 531-534.

[11] Jiang, N., Englehart, K.B. and Parker, P.A., *Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal,* IEEE Trans Biomed Eng, vol.56, no.4, pp.1070-1080, 2009.

[12] Rahimi, A., and Recht, B. (2008). *Random features for large-scale kernel machines.* Adv. Neural Inf. Process. Syst.20, 1177–1184.