Evaluating subsampling strategies for sEMG-based prediction of voluntary muscle contractions

Risto Köiva
Neuroinformatics Group,
Bielefeld University
D-33615 Bielefeld, Germany
e-mail: risto.koiva@uni-bielefeld.de

Barbara Hilsenbeck and Claudio Castellini
Robotics and Mechatronics Center
DLR — German Aerospace Center
D-82234 Weßling, Germany
e-mail: claudio.castellini@dlr.de

Abstract—In previous work we showed that some human Voluntary Muscle Contractions (VMCs) of high interest to the prosthetics community, namely finger flexions/extensions and thumb rotation, can be effectively predicted using muscle activation signals coming from surface electromyography (sEMG). In this paper we study the effectiveness of various subsampling strategies to limit the size of the training data set, with the aim of extending the approach to an online VMC-prediction system whose main application will be force-controlled hand prostheses. We performed an experiment in which 10 able-bodied participants flexed and extended their fingers according to a visual stimulus, while muscle activations and VMCs (represented as synergistic fingertip forces) were gathered using sEMG electrodes and a custom-built measurement device. A Support Vector Machine (SVM) was trained on a fixed-sized subset of the collected data, obtained using seven different subsampling strategies. The SVM was then tested on subsequent new data. Our experimental results show that two subsampling strategies attain a prediction error as low as 6% to 12%, which is comparable to the error values obtained in our previous work when the entire data set was used and processed offline.

I. INTRODUCTION

Over the past 30 years surface electromyography (sEMG) has received increased interest from the rehabilitation research community as an input modality for controlling advanced hand prostheses (see [1], and also the surveys in [2], [3]). Traditionally two sEMG electrodes were used to capture muscular signals that triggered movements of a simple one degree-of-freedom open/close type gripper when predefined thresholds were exceeded. However, with the appearance of polyarticular prostheses such as, e.g., Otto Bock’s Michelangelo1 and RSL Steeper’s BeBionic2 hands, more sophisticated control solutions to actuate the single fingers are being investigated. In particular, sEMG patterns are classified according to the desired hand posture, leading to a more natural form of open-loop control. This approach has, in turn, at least three disadvantages: (a) it forces the user to select from a limited set of predefined grip types; (b) the stability of the prosthesis is starkly affected by the accuracy of the classifier; and (c) it enforces no control over the force the prosthesis applies.

In order to circumvent these problems, we have recently proposed [4] a radical shift: to use Voluntary Muscle Contractions (VMCs), rather than hand position configurations, as the target values of machine-learning-based prediction. VMCs are defined as voluntary actions of muscle contraction which produce measurable effects, in our case forces. Typically, when a participant is asked to attain a specific finger force, for example to voluntarily flex the index finger to 50% of its maximum force, adjacent fingers will also change posture and produce side-effect forces [5]. Using VMCs allows us to isolate the finger of interest and associate synergistic muscle activations with it.

VMCs can be quantitatively elicited from any participant using visual stimuli, which is a standard experimental paradigm for this kind of investigation. Moreover, a set of VMCs of interest can be selected according to the degrees of freedom afforded by a particular prosthesis, therefore simplifying the control task. Finally, VMCs provide a continuous representation of the action space, independent of the participant’s anatomy (the muscles and/or motor-units involved). This means that their prediction enables force-based control over an infinite manifold of configurations. We argue that this will maximize the dexterity of control and in the future allow prosthetic wearing amputees to perform a wider variety of activities in their daily lives through enhanced control of prosthesis hardware.

In this paper we build on the results obtained in our previous work [4], extending the prediction to unseen data using a training set that is strictly limited in size. In general cases, in which the prediction model is trained online, a strategy to limit the amount of training samples is normally required to reduce the computational effort. Such models typically obtain much worse results on unseen data, largely due to the well-known issue of the sEMG signal drifting and changing over time [6], [7]. Online processing of sEMG signals for classification has yielded interesting but sub-optimal results in recent times, for instance in [8] where an average classification accuracy of 79% on amputees was reported. This is far below the results obtained for able-bodied participants in the offline setting.

In the experiment ten able-bodied participants were instructed to apply finger-force patterns associated with VMCs of interest during a data-gathering phase, while wearing sEMG electrodes around the forearm to capture muscle activations. The obtained data sets were subsampled using a number of

1www.ottobock.com
2www.bebionic.com
schemes and six Support Vector Machines (SVM) [9], [10] were trained on the reduced data set. In the final phase, VMCs were predicted (as synergistic fingertip forces) using the obtained models. The best results are obtained by ignoring samples which are too close (in force space) to previous ones. The resulting prediction errors range between 6.39% and 11.54% of the force ranges; these values are in line with those we previously obtained in a similar, but offline, experiment in [4]. We note that none of the strategies we employed emerged as a clear winner, suggesting that there exists some flexibility in choosing a subsampling strategy.

In the next section we describe the experimental setup and protocol in detail. In Sec. III we discuss the analysis methods and tested subsampling strategies. Sec. IV presents the experimental results, and finally in Sec. V we draw conclusions and outline future work.

II. Experiment Description

The goal of the experiment was to predict VMCs as synergistic combinations of fingertip forces. To this aim the participants applied finger forces and their forearm muscle activations and the forces produced by the fingertips were recorded. The setup consisted of three main parts: the sEMG electrodes, the finger force capturing sensor rig and visual stimuli presented on a monitor [Fig. 1].

A. Surface electromyography

Ten OttoBock MyoBock 13E200 sEMG electrodes were used to capture the muscle activations on the participant’s forearm. These electrodes provide an amplified, bandpass filtered and rectified signal. They were strapped around the participant’s forearm, just below the elbow, using an adjustable hook-and-loop band. Five electrodes were placed on the dorsal side, and the other five were placed on the ventral side of the forearm. All electrodes were evenly distributed without any respect to intrinsic muscles (the region of the ulna was avoided since no strong muscular signal is available there).

B. Fingertip forces

In general, a single VMC generates a simultaneous change of flexion/extension force in several fingertips, making an accurate synchronous gathering of forces from all fingers a necessity. To this end we use our custom built Finger-Force Linear Sensor (FFLS) [11]. The FFLS measures single finger flexion and extension forces of the index, middle, ring and little fingers, using an individually calibrated ±100N industrial strain gauge based force sensor for each finger. Thumb rotation and adduction/abduction forces are captured using a single radial dual-axis calibrated strain gauge sensor with a similar ±100N range. These sensors are characterized by their high signal repeatability, minimal drift over time and almost perfect linearity and non-existent hysteresis with both parameters guaranteed to deviate not more than 0.3%. Each strain gauge force sensor is connected to a dedicated industrial strain gauge amplifier, which converts the applied force into an analog voltage.

All sensors are adjustable for various hand sizes and were mounted rigidly after positional adjustments. The fingers are tightened to the sensors using hook-and-loop bands with minimal slack. Custom made gypsum casts are used for backlash-

3Finger flexion and extension refer to pushing the finger down and pulling the finger up respectively.

4Thumb rotation is an intuitive description of the thumb rotational motion, and thumb adduction/abduction refers to drawing the thumb closer or pushing the thumb further away from the palm.
Fig. 2. The stimulus and typical data collected during the training phase. (top panel) The sequence of the stimuli, six single-finger VMCs and three multi-finger VMCs. (center panel) The corresponding force data recorded by the FFLS — colors denote the force sensors and the force values are mostly negative since we defined flexion as the negative direction of the sensor axes. The six single- and three multi-finger VMCs are clearly seen; for instance, flexion of the index appears roughly between 140s and 180s as a series of five blue negative force peaks; the pinch grip (combination of index flexion and thumb adduction) appears as a series of combined blue and yellow peaks between 290s and 330s. (bottom panel) Two typical sEMG signals (red denoting an electrode placed on the dorsal side of the forearm, blue on the ventral side) corresponding to the elicited VMCs.

free (to ensure data is recorded for all movements) attachment of the participant’s thumbs to the dual-axis sensor. Both the sEMG electrodes and the FFLS are connected to a 12-bit resolution multichannel digital acquisition card that gathers data with a sampling frequency of 25Hz. This relatively low sampling rate was sufficient because the stimulated individual actions lasted for 5 to 10 seconds each and the sEMG electrodes provided a filtered signal.

C. Visual stimulus

A graphical user interface was displayed on a large monitor in front of the participant. The application was programmed to be used intuitively without prior training, and makes extensive use of wizards that led the participants’ through the different phases of the experiment, namely calibration, data gathering and prediction. The requested VMC was graphically presented using six red colored bars, one for each considered finger movement direction (flexion of the index, middle, ring and little fingers, plus thumb rotation and adduction). The height of the bars denotes the intensity of the targeted finger force. Next to the stimulus, blue colored bars show the measured fingertip forces and thus provide the participant with real-time visual feedback. During the prediction phase, the predicted finger forces are shown in an additional, third bar, colored in green.

D. Experimental protocol

The experiment consists of the calibration, data gathering and prediction phases. In the calibration phase, the idle and the maximum forces of each finger are recorded and averaged over a period of two seconds each. During the phases following the calibration, the requested force is limited to a fixed fraction of the captured maximum force. The top panel of Fig. 2 shows the nine requested VMCs during the data gathering phase. Six single-finger and three multi-finger VMCs were requested, the latter ones representing common grasp postures used in everyday life and resembling common grip patterns [12]. We concentrated on a pinch grip, an index finger pointing gesture and a power grasp. With the pinch grip (simultaneous thumb adduction and index flexion) small objects like a pen can be picked up. The index pointing gesture (flexion of the middle, ring and little finger) is used for pressing buttons and switches. Big objects, for example a bottle, are typically held using a power grasp, which is performed by flexing all fingers simultaneously.

During the data gathering phase, the stimulus increased from 0 to 80% of the participant’s maximum achievable force from the calibration phase for the single-finger VMCs and from 0 to 60% for the multi-finger VMCs, and then decreased back to 0 in a total of six seconds. Two seconds of rest were provided between VMCs in order to avoid muscle fatigue. Each VMC was repeated five times, which resulted in a total of 45 VMCs and a data gathering phase duration of ≈7 minutes [Fig. 2]. The data gathering phase was followed by a prediction phase, during which all VMCs were repeated two times (total duration, 2.8 minutes). The total experiment duration lasted ≈15 minutes, including the positional adjustment of all
sensors, the calibration, data gathering and prediction phases.

The system was tested with 10 able-bodied participants (all right-handed; 9 male, 1 female; aged 27.40 ± 4.41 years, min: 23, max: 29). The participants sat on an office chair, which was adjusted for maximum comfort, but ensuring that the armrest was kept at the same height as the FFLS. By slightly reclining the chair, we made it difficult for participants to use their body weight to generate extra force. They were also explicitly instructed not to try this. No fatigue was reported for the single-finger VMCs. However, it is interesting to note that all participants reported that it was more difficult to follow the stimulus for the multi-finger VMCs, even though the required forces were reduced to only 60% of the maximum single finger force value. Indeed, some participants did not even manage to reach the requested 60% force level simultaneously with multiple fingers. Nevertheless, it is important to note that ground truth is always available in terms of fingertip forces exerted at the fingertips; therefore, it is reasonable to assume, according to the literature [15] and using our intuition, that the relationship is a smooth one. According to this argument, samples which are too close to previously recorded ones can be neglected. For our approach, closeness is represented as an inter-sample distance in either force or sEMG space.

More formally, let $S$ be the set of sEMG/force samples collected during the data gathering phase. A subset $S' \subset S$ was built according to one of the following subsampling strategies:

1) samples were randomly picked from $S$ (random subsampling);
2) samples were randomly picked from $S$, but this time they were only selected if the forces were above 75% of the sampled maximum force during the data gathering phase (random subsampling with preselection);
3) samples picked from $S$ such that $||s_i - s_j|| > d$, $s_i, s_j \in S'$, for any $i, j = 1, \ldots, |S|$, where $d$ is a threshold that represents the minimum inter-sample distance allowed (Euclidean distance subsampling of sEMG data);
4) same as above, but the distance test is done in the output space on the force data (Euclidean distance subsampling of force data);
5) same as item 3 but the Mahalanobis distance is used instead of the Euclidean distance, i.e., given two samples $s_1, s_2, MD(s_1, s_2) = \sqrt{(s_1 - s_2)^T C^{-1} (s_1 - s_2)}$, where $C^{-1}$ is the inverse covariance matrix of $S$ (Mahalanobis distance subsampling of sEMG data);
6) same as item 4 but the Mahalanobis distance is used (Mahalanobis distance subsampling of force data);
7) sEMG samples which yield no hash collision when entered into a standard hash table of predefined size. Hashing was done using 20 equidistal bins in sEMG space (hashing).

Of course for each strategy one needs to decide the size of the resulting set, $|S'|$; in the case of random subsampling (strategies 1 and 2 above), this can be directly chosen, while in the other cases $d$ must be chosen each time in order for $|S'|$ to be close to the desired size. In both cases, in order to check the effectiveness of each strategy, we let $|S'|$ range over several values between 300 and 1500 (this upper value was limited by the CPU speed of the testing machine), in each case checking the attained error rate on the set $S - S'$ [Fig. 3]. As an error measure, we chose the Root-Mean-Square error, normalized over the range of the observed force values (NRMSE).

The best results were achieved using random subsampling with preselection and Mahalanobis distance subsampling of force data (strategies 2 and 6). Fig. 4 shows the detailed
Fig. 3. Error rates for each subsampling strategy as $|S'|$ increases.

comparison of these two strategies. For random subsampling with preselection, the test was repeated 50 times and the error was averaged (in the figure, one standard deviation is shown). The comparison of these strategies reveals no clear winner for the three VMCs shown, and this was also the case for the remaining VMCs (not depicted). In order to build the sEMG-to-VMC maps for the prediction phase of our experiment, we chose Mahalanobis distance subsampling of force data, mainly due to its non-random nature; given the same $S$, this strategy will always yield the same $S'$.

IV. RESULTS

NRMSE was used as a measure to evaluate the performance of the models obtained from $S'$ on the data collected during the prediction phase. Table I shows the results for all participants and all tested finger movements. This means that the statistics shown for each finger movement (e.g., Index) include both single and multiple VMCs.

The best accuracy was obtained by Participant #7, with error values ranging between $3.03\%$ and $9.76\%$, and the worst accuracy was obtained by Participant #8, with error values ranging between $8.23\%$ and $16.27\%$. The finger movement that was best predicted was flexion of the ring finger, with error values from $3.03\%$ to $11.79\%$, whereas the thumb adduction prediction performed worst, with values from $4.60\%$ to $16.27\%$. The fact that the thumb adduction was the hardest finger movement to predict is not surprising, given that muscles devoted to its movement lie mainly within the hand and are far from the sEMG gathering sensors. It is on the other hand comforting to see that the thumb rotation, an extremely important VMC, can be predicted with similar accuracy to the other finger VMCs ($8.90\% \pm 0.93\%$).

Fig. 5 displays a typical comparison between measured and predicted force values for the data from Participant #3. As can be seen, the prediction errors for single- and multi-finger VMCs are comparable. In some cases the prediction for the multi-finger VMCs (the peaks starting from 120 second mark) is even slightly better than for the single finger movements. This is probably due to the fact that it is impossible to simultaneously apply equal force with many fingers [5], which results in a large Mahalanobis distance through variances in the force patterns. Therefore, during the training phase, a higher percentage of data associated with multiple VMCs could be favored for selection into $S'$ to train the SVM models.

V. CONCLUSIONS AND FUTURE WORK

The existence of a stable relationship between sEMG and some VMCs of interest, namely finger flexions and extensions and thumb rotation, has already been shown in [4]. The VMCs were chosen to match the degrees of freedom of modern, polyarticulated hand prostheses. In this paper, we
extend our previous work towards having online training of SVM models using limited-sized training datasets. In the described experiment, 10 able-bodied participants were asked to apply graded VMCs according to a visual stimulus. Several subsampling strategies were applied to the obtained datasets, and models obtained from these subsets were used to predict the same VMCs obtained during a further session. Our results demonstrate that subsampling strategies can be used to obtain error rates similar to those obtained on the entire dataset [4] with values ranging from 6% to 12%. Whether this accuracy is usable in practice is still the subject of future research, but our approach is promising.

A number of factors had to be considered. Firstly, the usage of a strategy for subsampling is just one (and probably the simplest) of the possible ways to bound a training set of a machine learning method in an online setting. New data is always available, and no method can afford to work on an ever-growing dataset. This is the main motivation behind this research. Secondly, the selection of a particular subsampling strategy proved to be not as critical as we had expected. This finding needs further investigation; for example, it is not entirely clear which strategy scales up to much larger datasets, collected over hours and days. Random strategies, on the other hand, would almost certainly prove sensitive to the frequency content of the stimulus. Nonetheless, our approach is promising, revealing that online training and prediction with a limited training dataset is viable.

As far as timing is concerned, any method supposed to work online needs to enforce fast prediction, at least matching the capability of the target control system (i.e., a prosthesis) or cases in which a simulated model is given as visual output to a participant, should achieve an update rate in excess of ≈25 Hz. In our case, the prediction takes roughly 100μs with 2000 support vectors on a current typical desktop computer, which results in more than an adequate performance. We note however that the training phase, even with the reduced dataset, may take up to 10s.

A practical application of this system should take into account the limited capabilities of amputees to apply graded forces with the residual muscles of the missing limb. To this aim we are studying the possibility of training on only null and maximum forces and without using any force sensor; initial

---

**TABLE I**

Normalized Root-Mean-Square Error (NRMSE) for all participants and fingers tested, plus related statistics.

<table>
<thead>
<tr>
<th>Participant#</th>
<th>Index</th>
<th>Middle</th>
<th>Ring</th>
<th>Little</th>
<th>Th.rot.</th>
<th>Th.add.</th>
<th>mean ± std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.94</td>
<td>9.22</td>
<td>9.14</td>
<td>6.56</td>
<td>10.15</td>
<td>15.91</td>
<td>10.15 ± 3.10</td>
</tr>
<tr>
<td>2</td>
<td>6.95</td>
<td>9.27</td>
<td>8.72</td>
<td>9.18</td>
<td>8.72</td>
<td>4.97</td>
<td>7.97 ± 1.69</td>
</tr>
<tr>
<td>3</td>
<td>9.29</td>
<td>10.80</td>
<td>7.58</td>
<td>4.74</td>
<td>7.36</td>
<td>4.60</td>
<td>7.40 ± 2.45</td>
</tr>
<tr>
<td>4</td>
<td>11.16</td>
<td>7.04</td>
<td>6.51</td>
<td>7.41</td>
<td>8.43</td>
<td>11.86</td>
<td>8.74 ± 2.25</td>
</tr>
<tr>
<td>5</td>
<td>9.40</td>
<td>8.80</td>
<td>5.20</td>
<td>6.22</td>
<td>9.06</td>
<td>14.71</td>
<td>8.90 ± 3.31</td>
</tr>
<tr>
<td>6</td>
<td>10.32</td>
<td>7.79</td>
<td>6.05</td>
<td>6.19</td>
<td>8.59</td>
<td>13.87</td>
<td>8.80 ± 2.95</td>
</tr>
<tr>
<td>7</td>
<td>9.76</td>
<td>5.61</td>
<td>3.03</td>
<td>3.96</td>
<td>9.23</td>
<td>6.73</td>
<td>6.39 ± 2.73</td>
</tr>
<tr>
<td>8</td>
<td>13.07</td>
<td>11.21</td>
<td>8.23</td>
<td>9.79</td>
<td>10.74</td>
<td>16.27</td>
<td>11.54 ± 2.81</td>
</tr>
<tr>
<td>9</td>
<td>11.90</td>
<td>12.08</td>
<td>5.15</td>
<td>7.39</td>
<td>8.74</td>
<td>6.03</td>
<td>8.55 ± 2.93</td>
</tr>
<tr>
<td>10</td>
<td>7.77</td>
<td>9.70</td>
<td>11.79</td>
<td>6.42</td>
<td>8.01</td>
<td>6.25</td>
<td>8.32 ± 2.11</td>
</tr>
</tbody>
</table>

mean ± std. 9.96 ± 1.72 9.15 ± 1.87 7.14 ± 2.36 6.79 ± 1.69 8.90 ± 0.93 10.12 ± 4.58

---

Fig. 5. Recorded and predicted force values of all considered VMCs during one representative trial. All VMCs performed in the training phase are repeated two more times for testing. The data from 0 to 120 seconds refers to force values gathered during single-finger VMCs phases, and the data after 120 seconds refers to the force values captured while participants were performing multi-finger VMCs.
results reveal that this strategy is promising. Generalization of single-finger VMC to multi-finger ones is still an open issue and is being investigated as well.

We note that our choice of the NRMSE as a measure of error might not be optimal and that some alternative form of locally-weighted error (see, e.g., [15]) could be more suited to the problem. However, the final benchmark should be a physiological one, such as, e.g., the success rate of human participants engaged in a box-and-block task [16] while using a prosthetic device controlled by various algorithms. Therefore, work is underway to test our approach using a 3D hand model in order to check whether amputees can perform the exercise with the same level of performance as was achieved by able-bodied participants during this experiment. Based on previous results, e.g., [17], we are confident of success. Finally, there are plans to test the stability of our system with a robotic hand.

ACKNOWLEDGMENT

This work was partially supported by the Swiss National Science Foundation Sinergia project #132700 NINAPRO (Non-Invasive Adaptive Hand Prosthetics), by the European Commission project The Hand Embodied (FP7-248587) and by the DFG Center of Excellence EXC 277: Cognitive Interaction Technology (CITEC). We would like to thank Patrick van der Smagt of the Technische Universität München and Michael Suppa of the DLR for their continual financial, moral and logistic support. We thank Mr. Johann Buchner of the DLR, for providing us with some electro-technical parts for the experiments. Finally, we thank Jonathan Maycock for his input and proofreading.

REFERENCES


