Real-time classification of multi-modal sensory data for prosthetic hand control

Citation for published version:

Digital Object Identifier (DOI):
10.1109/BIOROB.2016.7523681

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Real-Time Classification of Multi-Modal Sensory Data for Prosthetic Hand Control

Iris Kyranou, Agamemnon Krasoulis, Mustafa Suphi Erden, Kianoush Nazarpour, Sethu Vijayakumar

Abstract— Recent work on myoelectric prosthetic control has shown that the incorporation of accelerometry information along with surface electromyography (sEMG) has the potential of improving the performance and robustness of a prosthetic device by increasing the classification accuracy. In this study, we investigated whether myoelectric control could further benefit from the use of additional sensory modalities such as gyroscopes and magnetometers. We trained a multi-class linear discriminant analysis (LDA) classifier to discriminate between six hand grip patterns and used predictions to control a robotic prosthetic hand in real-time. We recorded initial training data by using a total number of 12 sEMG sensors, each of which integrated a 9 degree-of-freedom inertial measurement unit (IMU). For classification, four different decoding schemes were used; 1) sEMG and IMU from all sensors 2) sEMG from all sensors, 3) IMU from all sensors and, finally, 4) sEMG and IMU from a nearly optimal subset of sensors. These schemes were evaluated based on offline classification accuracy on the training data, as well as with task-related metrics such as completion rates and times for a pick-and-place real-time experiment. We found that the classifier trained with all the sensory modalities and sensors (condition 1) attained the best decoding performance by achieving a 90.4% completion rate with an average completion time of 41.9 sec in real-time experiments. We also found that classifiers incorporating sEMG and IMU information outperformed on average the ones that only used sEMG signals, even when the amount of sensors used was less than half in the former case. These results suggest that using extra modalities along with sEMG might be more beneficial than including additional sEMG sensors.

I. INTRODUCTION

The majority of state-of-the-art robotic prosthetic hands are controlled by using surface electromyography (sEMG), that is electrical activity recorded non-invasively on the skin surface, which when pre-processed can be utilized to command the prosthetic hand [1,2]. Currently, amputees using prosthetic hands are trained to memorize and execute specific motion sequences which are then mapped to predefined grasp types. These correspond to a general family of human hand formations that we naturally employ whilst approaching to grasp an object. Different approaches have been proposed in the literature for grasp categorization, mainly based on observations of the way humans naturally grasp objects [3,4]. Research has identified the most commonly used grasp types [5], as well as hand synergy models that are deployed for performing grasps [6]. This information has been used to successfully control under-actuated prosthetic hands such as the Touch Bionics i-Limb1, Bebionic2, Vincent hand3, and Ottobock Michelangelo4 in a natural looking way.

Commercial prosthetic devices mostly employ on-off control strategies. Although such methods are computationally simple and perform robustly, they dramatically increase the mental effort required by the user. Furthermore, the number of different grasps that can be executed by the prosthesis is limited by the users' ability to memorize and perform the distinct motions required. Machine learning techniques have been used to analyze and map recorded sEMG activity to grasp types in a more intuitive way. A variety of classifiers have been evaluated in the literature with promising performance, often achieving offline classification accuracy higher than 97% [7-11]. Recently, pattern recognition has been employed in a commercial myoelectric control application5 allowing the decoding of wrist pronation/supination, wrist flexion/extension and hand opening/close in real-time.

A limiting factor of pattern recognition-based myoelectric control is the lack of stability and predictability of the prosthesis performance. Biddiss et al. [12] investigated the reasons of abandonment of the use of upper-extremity prostheses, including myoelectric controlled ones, and reported that 88% of patients chose to stop using prosthetic hands because they found them “too tiring and difficult to use”. Therefore, to incorporate classification methods in prosthetic hand control, prostheses have to perform robustly and predictably.

One way of achieving more robust pattern recognition-based control is by improving the decoder's classification accuracy. Recently, the impact of incorporating additional sensor signals in the decoding process other than sEMG has

* This work has been supported by EPSRC grants EP/F500385/1, EP/H1012338/1, EP/M025594/1, EP/M025977/1, EP/L016834/1 and BB/F529254/1. This work has been performed at Heriot Watt University and the University of Edinburgh under the program of Centre for Doctoral Training for Robotics and Autonomous Systems
I. Kyranou is with the Edinburgh Centre of Robotics, Heriot-Watt University, UK (e-mail: ik5@hw.ac.uk)
A. Krasoulis is with the Institute for Adaptive and Neural Computation and the Institute of Perception, Action and Behaviour, School of Informatics, University of Edinburgh, UK (e-mail: a.krasoulis@sms.ed.ac.uk)
M.S. Erden is with the School of Engineering and Physical Sciences, Heriot-Watt University, UK (e-mail: m.s.erdem@hw.ac.uk)
K. Nazarpour is with the School of Electrical and Electronic Engineering, Newcastle University, UK (e-mail: kianoushz.nazarpour@ncl.ac.uk)
S. Vijayakumar is with the Institute of Perception, Action and Behaviour, School of Informatics, University of Edinburgh, UK (e-mail: sethu.vijayakumar@ed.ac.uk)

http://www.touchbionics.com
http://bebionic.com/the-hand
http://vincentsystems.de/en
http://www.touchbio.com
http://www.living-with-michelangelo.com/gb/home
https://www.coapengineering.com
been investigated in terms of classification accuracy. Gijsberts et al. [13] and Fougner et al. [14] have observed higher offline classification performance when acceleration information measured on the users’ forearm was taken into consideration. Based on that evidence, Fougner et al. suggested that in any two-site EMG system it is preferable to add an accelerometer affixed to the forearm, rather than including a third electrode. These findings naturally raise the question of whether the use of additional sensory information, such as angular velocity and orientation, could further enhance the performance of pattern recognition-based myoelectric control. To the best of our knowledge gyroscopes and magnetometers have not been previously used in the context of upper extremity myoelectric control. Moreover, the previously mentioned studies [13,14] performed only an offline analysis, rendering essential the investigation of whether the inclusion of extra sensory modalities can prove beneficial for myoelectric control during online experiments.

In this study, we simultaneously record sEMG activity along with acceleration, rotational velocity and orientation by using inertial measurement unit (IMU) sensors. Each IMU sensor integrates a three-dimensional (3D) accelerometer, a 3D gyroscope and a 3D magnetometer measuring acceleration, angular velocity and orientation, respectively. We compared the performance of a multiclass linear discriminant analysis (LDA) classifier using solely sEMG activity, IMU information or a combination thereof. For the latter case, we also investigated the effect of using a smaller number of sEMG/IMU sensors as compared to the other three cases. The decoding performance was assessed by comparing offline classification accuracy on the initial training data, as well as real-time control performance considering task-related metrics, such as completion rates and times for a pick-and-place experiment by controlling a robotic prosthetic hand in real-time.

II. MATERIALS AND METHODS

For the purposes of this study, a real-time pick-and-place experiment was designed. A full pipeline for recording sEMG and IMU activity, data pre-processing, motion classification and prosthetic control was developed and implemented using the Robot Operating System (ROS) and custom-written software in C++. The pipeline is summarized in Fig. 1 and explained in detail in the following sections.

A. Signal Acquisition

Muscular activity and IMU information (3D acceleration, 3D angular velocity and 3D orientation) were recorded with 12 wireless Trigno™ IM sensors\(^6\). An example of the raw input data as recorded from the first sensor while one subject grasps different objects (from left to right: bottle, credit card, key).

\[\text{Fig. 2. The raw sEMG, accelerometer (ACC), gyroscope (GYR) and magnetometer (MAG) readings from one sensor while one subject grasps different objects (from left to right: bottle, credit card, key)}.\]

\[\text{TABLE I. EXTRACTED FEATURES FOR SEMG AND IMU DATA}\]

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition (per channel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Average Value (MAV)</td>
<td>[\text{MAV} = \frac{1}{N} \sum_{n=1}^{N} x_n]</td>
</tr>
<tr>
<td>Waveform Length (WL)</td>
<td>[\text{WL} = \sum_{n=1}^{N}</td>
</tr>
<tr>
<td>Logarithm of Variance (IVar)</td>
<td>[\text{IVar} = \log\left(\frac{1}{N-1} \sum_{n=1}^{N} (x_n - \bar{x})^2\right)]</td>
</tr>
<tr>
<td>Autoregressive 4th order (AR)</td>
<td>[x_n = -\sum_{i=1}^{p} a_i x_{n-i} + w_n]</td>
</tr>
<tr>
<td>Mean Value (MV)</td>
<td>[\text{MV} = N^{-1} \sum_{n=1}^{N} x_n]</td>
</tr>
</tbody>
</table>

Where \(N\) is the number of samples, \(p\) the order of the autoregressive model \((p=4)\), \(a_i\) the \(i^{th}\) autoregressive coefficient, \(x_n\) the \(n^{th}\) sample of the signal.

\[\text{6 http://www.delsys.com/products/wireless-emg/}\]
The acquisition sampling rate was 2 kHz for EMG and 128 Hz for IMU signals. A unique timestamp was associated with each data sample and the two time-series were post-synchronized by up-sampling the IMU signals to 2 kHz (i.e., interpolation).

### B. Signal pre-processing and feature extraction

To remove the artifact noise, muscular signals were band-pass filtered in the range 20 Hz to 500 Hz using a 4th order digital Butterworth filter. Subsequently, features were extracted from sEMG and IMU signals by using a sliding window approach. The length of the window was set to 256 ms with 200 ms overlap, yielding a feature sample every 56 ms.

Following the results of Scheme et al. [16], Hahne et al [17] and our previous work [18], a combination of four time-domain features were extracted from each sEMG signal, namely the mean absolute value (MAV), waveform length (WL), log-variance (IVar) and 4th-order auto-regressive model (AR) yielding four coefficients as features, hence providing a total of seven features per EMG channel (Table I). For IMU signals, we computed the mean value (MV) within the processing window, therefore each sensor contributed a total of nine features; one MV feature for each dimension of the 3D accelerometer, 3D magnetometer and 3D gyroscope. All features were mean subtracted and normalized to unit standard deviation.

The dimensionality of the feature space was defined by the input modality and the amount of sensors used in each condition (Section 2C). For instance, in the case of using solely IMU information from all 12 sensors, the input feature vector was 108-dimensional (9 features/channel × 12 channels).

### C. Classifier

To build a mapping from sEMG/IMU activity to grasp types a multi-class linear discriminant analysis (LDA) classifier was implemented. Some of the advantages of using an LDA classifier are its computational simplicity and, perhaps more importantly, its probabilistic nature. At each step, the classifier estimates a multinomial distribution over the predicted motion. The class with the highest posterior probability is then predicted and the uncertainty about the prediction is encoded in the posterior probability distribution. The number of decoded classes in our experiment was \( k = 6 \) (see Section 2D).

### D. Experimental Protocol

Eight able-bodied male subjects participated in the study. All subjects were right-handed and had no known neurological problem. Each subject participated in a single experiment that lasted on average 1.5 hours. Prior to the experimental sessions, the subjects were asked to sign an informed consent participation form. All experiments were approved by the local Ethics Committee of the School of Informatics, University of Edinburgh.

Following sEMG/IMU sensor placement (Section 2A), the subjects’ fingers were constrained in a fist formation by using elastic bandage and, finally, a prosthetic hand was mounted on the subjects’ forearms by using a custom made socket for this purpose. For this experiment, we used the
Touch Bionics Robo-limb prosthetic hand which offers the potential of individual control of 6 degrees-of-freedom (DOFs); flexion/extension of the five fingers and thumb rotation. The hand was operated by a laptop via a CAN bus connection. The full setup including the sEMG/IMU sensors and the prosthetic hand mounted on a subject’s forearm is shown in Fig. 3.

The experimental environment is depicted in Fig. 4. Participants sat on a chair in front of a computer desk. They were asked to use the prosthetic hand to grasp the objects which were located on the desk and relocate them by 50 cm in the right direction. Each trial consisted of relocating three objects and finally pressing the “space” key on a computer keyboard. Each object was labeled with a unique number (1-4) and was also associated with a specific grip type. The objects used in the experiment along with the corresponding grips are presented in Table III. The total number of classes predicted by the decoders was six, including four grips, the “open hand” motion and the “rest position”. The latter class corresponded to keeping the joints of the prosthesis fixed at their current angles (i.e. no action taken).

The experimental sessions consisted of three phases, a short preparatory phase, the training phase and, finally, the testing (i.e. real-time classification) phase.

1) Preparatory Phase

As mentioned above, the participants’ hands were constrained during the experiments to refrain participants from moving their fingers and in this way to mimic an amputee scenario as closely as possible. During the training phase, the participants were asked to perform imagery grasps of a series of objects by activating their muscles, though not being able to move their fingers. This activity might feel unnatural, thus a short preparatory phase consisting of two full trials was introduced to allow the subjects to familiarize with the experiment. During this stage no data were recorded.

2) Training Phase

During the training phase, labeled training data were collected to be used for the supervised training of decoders. At this stage, subjects performed five consecutive trials of imagery movements whilst data were collected from the sEMG/IMU sensors. Participants were instructed to reach the objects with their right hand and apply forces as they would normally, should their hands not be constrained. Simultaneously, they were asked to indicate the executed movement (class 1-5 in Table III) by pressing the corresponding number on a computer keyboard with their left hand. When no key was pressed (i.e. in between movements), the “rest pose” class was assumed. During the training phase, the prosthetic hand was switched off maintaining the rest position that is shown in Fig. 3.

3) Testing Phase

During the testing phase, participants controlled the prosthetic hand by using sEMG and/or IMU signals. Participants performed five trials for each experimental condition, as explained in Section 2C and in Table II. All trials were initiated by a “Go” cue and finished when the “space” key was pressed on the computer keyboard. The time taken to accomplish each trial was measured by the experimenter and a trial was considered successful when it was accomplished within 60 sec. This amount of time was long enough to allow for the correction of a small number of misclassifications. As misclassification we consider a pregrasp which is not the intended one that corresponds to the object the subject wants to grasp (pregrasps that correspond to different objects are listed in Table III).

The classifier yielded a grip prediction and a corresponding posterior probability every 56ms. A classification prediction led to a control action only when the posterior probability of the most probable class exceeded a pre-defined threshold. The threshold was set empirically to $\theta = 0.995$. The controller was implemented as a state machine, which means that a new action was taken only when the execution of the previous action was over. When the performed pregrasp was different to the intended one, participants were instructed to open the hand (class 5 of Table III) and try performing the movement again.

During the design of the experiment, care was taken into creating a process that would not favor one configuration over another. To account for the effect of the learning mechanisms that take place during real-time myoelectric control [20], the decoding condition order was counter-balanced across participants. The possible combinations of four conditions are 24, but we only had 8 subjects. Thus, we randomized the sequence of the conditions in every

<table>
<thead>
<tr>
<th>Class</th>
<th>Object</th>
<th>Grip</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>Rest pose</td>
</tr>
<tr>
<td>1</td>
<td>Bottle</td>
<td>Cylindrical</td>
</tr>
<tr>
<td>2</td>
<td>Card</td>
<td>Lateral</td>
</tr>
<tr>
<td>3</td>
<td>CD</td>
<td>Tripod</td>
</tr>
<tr>
<td>4</td>
<td>Keyboard key</td>
<td>Index point</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>Open pose</td>
</tr>
</tbody>
</table>

Table III. Objects used in experiments and corresponding grips
experiment in a controlled way, so as none of the previous sequence was repeated and every condition appeared exactly twice in the same position across all the experiments.

### III. Results

The performance of the four types of decoders examined in this study was evaluated in terms of offline classification accuracy on the collected training data, as well as in terms of completion rates and times in the real-time control experiment.

#### A. Offline Evaluation

The mean classification accuracies of all four conditions averaged across participants are presented in Table IV. The classification accuracy metric is defined as the ratio of correct predictions over the total number of classified instances.

The decoder that incorporated both sEMG and IMU information from all 12 electrodes (condition 1) achieved the highest training accuracy, followed by the decoder that only used IMU information (condition 3). The decoder that combined sEMG and IMU information from a reduced number of sensors (condition 4) achieved slightly lower average classification accuracy outperforming the sEMG-only based classifier (condition 2).

#### B. Real-Time Control Evaluation

To assess online myoelectric performance we employed two task-related metrics widely used in the literature [21, 22], namely the completion rate (CR) and completion time (CT). The completion rate is defined as the percentage of successful trials, whereas completion time refers to the time required to accomplish a successful trial.

Mean completion rates and times for all conditions are presented in Fig. 5. In line with offline decoding results, the highest performance was achieved by the decoder that incorporated sEMG and IMU information from all available sensors. Nevertheless, contrary to the trend observed for offline classification accuracy, the next best performance was achieved by the decoder that combined sEMG and IMU information from a selected subset of sensors. This was achieved using an average 5.38±0.92 (mean±std) selected sensors across participants, which corresponds to less than half of the initial number of sensors.

The EMG-only condition performs faster than the IMU-only condition, but presents a lower success rate than the latter. This indicates that subjects were more successful in accomplishing the tasks using information only from IMU compared to the EMG-only condition, but the tasks took on average more time to complete than those using only EMG information.

#### C. Discussion of Results

The current study sought to investigate whether myoelectric prosthetic control could benefit from the inclusion of IMU information measured by accelerometers, gyroscopes and magnetometers placed on the users’ forearm. Our work was motivated by previous studies that reported increased offline classification accuracy when accelerometry information was incorporated in the decoding process [13, 14]. Nevertheless, these studies did not validate such findings with online myoelectric control experiments. Furthermore, information from gyroscopes and magnetometers has not been previously employed in upper limb pattern recognition-based myoelectric control.

We hypothesized that by including extra sensory information recorded with IMUs the real-time performance of the prosthetic hand would improve. This hypothesis was supported by the results of both offline analysis and the online experiments. Interestingly, when sEMG information was completely discarded and only IMU features were considered, the average completion rate was 79%, which was higher than the case where solely EMG information was used (70% completion rate). It is worth noting, however, that the completion times were higher on average in the former case (Fig. 5). The significance of including additional sensory information might become more obvious when we consider the results for condition 4, where sEMG and IMU information were combined but a considerably smaller number of sensors was used. These decoders consistently outperformed solely sEMG-based decoders, even though the number of sensors in the former case was reduced to less than half. This finding is extremely important from a clinical perspective, since in real-life applications it is essential to use a minimal number of sensors.

In accordance with previous studies [21, 23], we observed that the results obtained via purely offline analysis do not

---

**TABLE IV. AVERAGE TRAINING CLASSIFICATION ACCURACY**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sensory Input</th>
<th>Number of sensors</th>
<th>Offline classification accuracy (mean±std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sEMG &amp; IMU</td>
<td>12</td>
<td>94.5±2.7</td>
</tr>
<tr>
<td>2</td>
<td>sEMG</td>
<td>12</td>
<td>82.8±6.9</td>
</tr>
<tr>
<td>3</td>
<td>IMU</td>
<td>12</td>
<td>92.8±3.8</td>
</tr>
<tr>
<td>4</td>
<td>sEMG &amp; IMU</td>
<td>4-6</td>
<td>91±3.2</td>
</tr>
</tbody>
</table>

**Fig. 5.** Top: completion rates (mean + s.e.). Bottom: summary quartile plots of completion times. Straight lines, medians; open circles, means; solid boxes, interquartile ranges; whiskers, overall ranges of non-outlier data; solid circles, outliers.
necessarily correlate with the results extracted from online experiments. For instance, by looking at offline classification accuracies (Table IV), it seems that condition 2 (inclusion of IMU information only) achieves the second highest performance only preceded by condition 1. Nevertheless, this pattern was not observed for the online experiment where condition 2 yielded the largest completion times. The main reason behind such discrepancies might lie in the different evaluation metrics used in each case, that is classification accuracy for offline analysis as opposed to task-specific metrics (e.g. completion rates and times) for online experiments. Therefore the use of different metrics in the two cases renders such direct comparisons invalid. Additionally it has been demonstrated that differences between offline and online performance can be also due to the presence of visual feedback in the latter paradigm [24].

IV. CONCLUSION

In this study, we investigated the potential benefit of using novel sensory modalities for real-time myoelectric control. We provide evidence that the inclusion of acceleration, angular velocity and orientation information can improve the online myoelectric performance, as quantified by task-related metrics such as mean completion rates and times. We also demonstrated that the inclusion of additional sensory information might allow reducing the number of required sensors without compromising the decoding performance of the classifiers. This feature is highly valuable for clinical applications. In our study, we only included able-bodied participants. Verification of our results with amputee subjects is imperative and currently seen as a future research direction.

REFERENCES