Biomedical Signal Processing and Control 46 (2018) 121–130

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier

Mahmoud Tavakoli*, Carlo Benussi, Pedro Alhais Lopes, Luis Bica Osorio, Anibal T. de Almeida

Institute of Systems and Robotics of University of Coimbra, Portugal

ARTICLE INFO

Article history:
Received 5 May 2017
Received in revised form 16 April 2018
Accepted 10 July 2018
Available online 20 July 2018

Keywords:
EMG
Gesture recognition
Human machine interfaces
Wearable armband
Support vector machine
Classification

ABSTRACT

Integration of surface EMG sensors as an input source for Human Machine Interfaces (HMIs) is getting an increasing attention due to their application in wearable devices such as armbands. For a wearable device, comfort and lightness are important factors. Therefore, in this article we focus on a minimalistic approach, in which we try to classify four gestures with only 2 EMG channels installed on the flexor and extensor muscles of the forearm. We adopted a two-channel EMG system, together with a high dimensional feature-space and a support vector machine (SVM) as a classifier. In addition, tolerance of the system for rejection of unsolicited gestures during the body movement was evaluated, and the two methods were implemented to ensure this; one based on an SVM threshold and another one based on the addition of a locking gesture. The resulting system is able to recognize up to 5 gestures (hand closing, hand opening, wrist flexion, wrist extension and double wrist flexion), presenting a classification accuracy of between 95% and 100% for a trained user and robustness against different body movements, guaranteed with the locking feature. We showed that misclassification of other gestures as the unlocking never happened for expert users.

1. Introduction

Surface EMGs have been mostly used for control of prosthetic hands. Full control of a highly articulated hands still requires a high number of EMG sensors, raising significantly the cost and complexity of the system. It is a common configuration to utilize four EMG channels, as adopted in [1] and [2], in order to be able to detect wrist and 4-fingers flexions and extensions (Fig. 1); other authors, like Yang [3] and Bugmann [4], used 6 bipolar electrodes for recognizing up to 19 and 15 hand movement respectively and controlling a highly dexterous hand. These control systems require a high number of control inputs. Some other studies reported application of eight [5] or ten [6] bipolar electrodes positioned on the forearm.

Other examples of utilizing surface EMG sensors include finger joint angle estimation using a 8 channel EMG system [7] and hand gesture recognition using a 6 channel EMG armband [8]. Also, in a recent study authors discussed the selection of best subsets of EMG electrode pairs for classification of hand movements when performing 5 hand postures at 9 different arm positions [9].

Surface EMG sensors that were previously exploited only for prosthetic devices, were considered as an important source of input for general purpose Human Machine Interfaces (HMIs) for wearable devices. For such devices, it is always interesting to received as many inputs as possible, and classify as many gestures as possible, to enrich the human control over the machine. BioSLeeve [10] is an example of a wearable device that implements 16 electrodes.

Recently, a gesture recognition armband, the Myo armband was commercialized [11]. As can be seen in Fig. 2, this armband is able to detect five gestures, which are open, wave out, wave in, fist and pinch. Myo armband embeds eight EMG channels, and is designed as a wearable armband. Yet, Myo is relatively bulky as a wearable device.

The purpose of this article is then to explore a minimalistic approach in order to achieve hand gestures. Therefore, we explore recognition of four hand gestures (open, close, wave in, wave out) with only two EMG channels.

While, the “Open” and “Close” gestures, or more generally two gestures recognition with a single EMG channel have been already explored [12,13], adding the “wave in” and “wave out” gesture provides new possibilities for the user. The goals is then to build a real-time classification algorithm for recognizing these 4 gestures

* Corresponding author.
E-mail address: mahmoud@isr.uc.pt (M. Tavakoli).

https://doi.org/10.1016/j.bspc.2018.07.010
1746-8094/

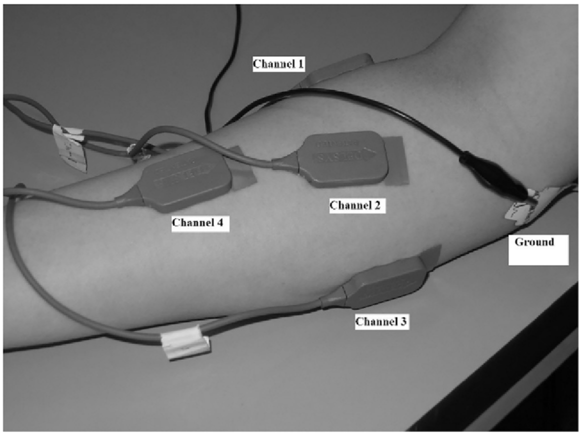
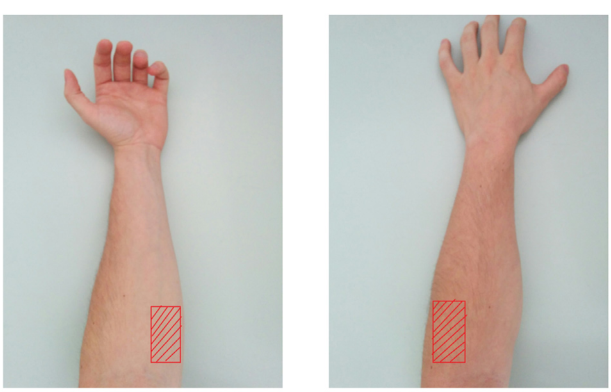


Fig. 1. 4 bipolar electrodes placement on the forearm.

Fig. 2. The Myoarmband integrates 8 EMG channels and an IMU unit and can detect up to 5 gestures.



(a) Channel 1 *(b) Channel 2*

Fig. 3. Indicative zones for electrodes positioning.

and measuring the success rate in real-time classification. Therefore the following constraints are fixed:

(1) The number of inputs is reduced to two, that is, the 4 necessary gestures must be recognized by using only two signals.
(2) The classifier must achieve good gestures recognition percentage, e.g. over 90%.
(3) The processing time for calibration must not exceed 30 s; moreover, the algorithm must be trainable with a small training set, asking so the user to perform a minimum number of gestures for the calibration.

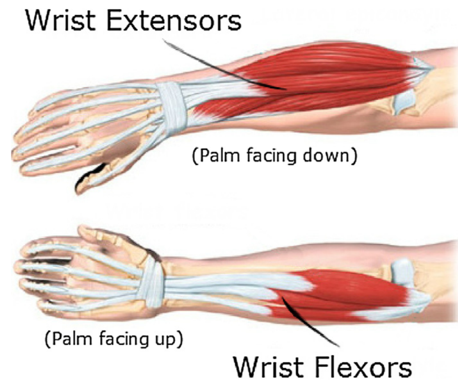


Fig. 4. The forearm anatomy, Wrist Flexor and Extensor muscles.

(4) The system response to a gesture performing must not exceed 300 ms for user comfort.
(5) The classifier should allow to be used without re-training it in every session; even when the algorithm is not re-calibrated.
(6) The system must provide robustness against other limb movements, that is, the disturbances coming from these motions must not be classified as one of the three gestures, letting the user move freely when wearing the EMG device.

In the following chapters, we describe the system design, the features of the signal that were extracted and the classifier which was selected and implemented. We then show the result of the system implementation for 7 subjects (four males and three females) that tested the system. Subjects were all healthy, with an average age of 25.6 years ($SD = 5.7$), with different levels of previous acquaintance with the system. The last constraint (tolerance to disturbances) actually found to be the most challenging for some of the beginner users, which enforced integration of a fifth gesture, as an option, for locking the system. Tests were repeated for beginner subjects and the results are presented.

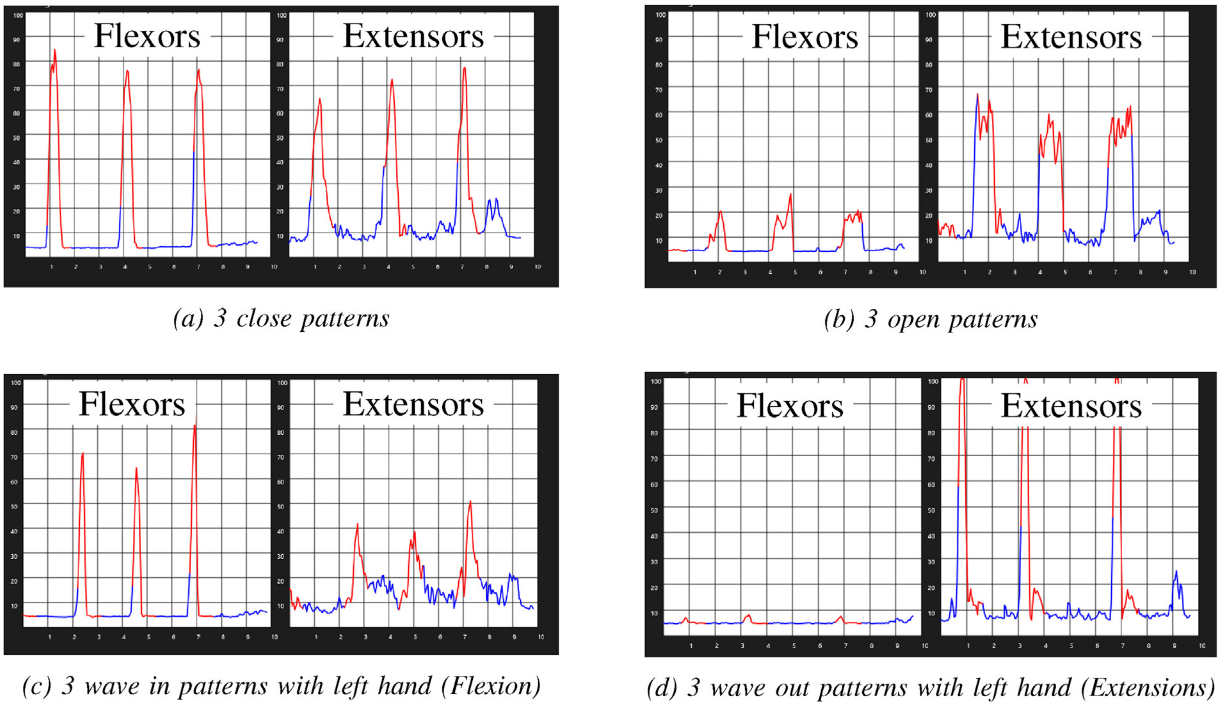
2. EMG system design and implementation

As described before, the goal of the design is to recognize hand gestures with satisfying reliability, for which we set at 90% the lower bound of correctly classified gestures. Moreover, we decided to use dry electrodes instead of gelled ones, to increase the comfort and durability of the system. Therefore we used stainless steel electrodes. The positioning of the two channels should follow the scheme shown in Fig. 3; however, we experienced during the test sessions that the optimal positioning could be considerably different for every subject, hence, the setting depicted in Fig. 3 should only serve as a starting point. Fig. 4 shows the forearm anatomy of the wrist flexor and extensor.

Each channel was then composed by 3 electrodes put in a row along the monitored muscle (which traduces in the same direction of the forearm), two of them providing the differential signal, and the middle one working as a common ground (the two resulting grounds from the 2 channels were connected together).

2.1. Hardware, pre-processing and acquisition

Before being acquired by the micro-controller, the signal is filtered and amplified. More precisely, the differential signal first passes through an instrumentation amplifier with high CMRR and unitary gain to eliminate common noise sources, such as the 50 Hz line noise; after that is filtered by a second order band-pass Sallen-Key filter with cut frequencies at 10 Hz and 500 Hz and amplified by a gain of 65 dB (1800). These operations were implemented in



(a) 3 close patterns

(b) 3 open patterns

(c) 3 wave in patterns with left hand (Flexion)

(d) 3 wave out patterns with left hand (Extensions)

Fig. 5. Gestures patterns: the figure shows the 4 gestures after filtering the noise plotted on a windows 10 application software that we developed (red line means the interval during which the features for the classification are extracted). Each picture points out the waveform envelope for the two channels coming from 3 instances of a defined class. The *y*-axis refers to V_out/V_supply, while *x*-axis labels indicate seconds. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

an in-made pre-processing board which will after be embedded in every module of the EMG armband.

The output coming from this board was then acquired by the micro-controller (ARM Cortex) with a sampling frequency of 1 kHz; depending on the muscle contraction, the signal changes significantly its frequency characteristic, and thus its variance. In order then to obtain a waveform dependent on the gesture, we took into consideration only the differences between two consecutive samples (for each channel):

$$x_i^1 = s_i^1 - s_{i-1}^1 \quad (1)$$

$$x_i^2 = s_i^2 - s_{i-1}^2 \quad (2)$$

In this way, the 4 defined gestures will present clearly different patterns, as shown in Fig. 5.

Moreover, since we select these differences as samplings describing the signal, "activity" (gesture performing) and "non-activity" patterns are clearly differentiated (Fig. 5); it is possible then to define a condition which **triggers the feature extraction block**, thus avoiding a continuous stream of classifications which could lead to "non-activity" being classified wrongly. Therefore, the following technique is used to detect when a gesture is performed:

1 For each of the two channels, the averages x_a^1, x_a^2 of the last M samplings $x_i^1 \quad x_i^2 \qquad i = 1, \ldots, M$ is evaluated and updated at each new sampling recording.
2 The differences d_i^1, d_i^2 between the current samplings x_i^1, x_i^2 and the averages x_a^1, x_a^2 are evaluated.
3 If $d_i^1 \geq t^1$ <u>or</u> $d_i^2 \geq t^2$, the features extraction is started. t^1, t^2 are predefined thresholds set by experience.

The "non-activity" patterns (i.e. noise) will not cross none of these thresholds, and cannot trigger the feature extractions.

2.2. Features extraction

Once one of the prefixed thresholds is crossed, it is necessary to extract features from the next sampling collection. This allows to reduce the dimensionality of the vector that describes one gesture, and to distinguish a particular gesture among the set of 4 gestures. Feeding the classifier with the entire array of samplings extracted from the subsequent interval is not feasible, since it would slow down significantly the classification process, and thus the overall system.

It can be seen Fig. 5 that, by taking into account both channels, the 4 gestures present different shapes; Therefore, for the distinction between them, we considered to feed a relatively low number (40) mean values for each channel, and for each gesture. In summary we decided to:

- For the first trial, fix the duration of the observation window (window during which features to feed the classifier are extracted) to maximum 0.8 s. This value was decided by analyzing the average length of a gesture performing (which can be seen in Fig. 5), and bounded superiorly for real-time constraints, to avoid slowing down the classification system.
- Extract from the samplings s_i^1, s_i^2 taken during the observation window (which result in 2 * 800 samplings) N subsequent means $f_i^1, f_2^2 \quad i = 1, \ldots, N$ to describe satisfyingly the gesture. N was fixed to 40. The feature vector sent to the classifier f is simply equivalent to the linking of f^1 and f^2, resulting in 2 * 40 = 80 features extracted from each gesture, and sent to the classifier. Each feature f_i^1, f_2^2 was then equivalent to the mean of the previous $\frac{800}{40} = 20$ samplings.

The shapes observed in Fig. 5 are thus conserved, and contemporary the dimension is reduced.

Before sending this vector to the classifier, it is necessary to normalize it for a better behavior of the classifier, which will work on "distances" between features. To achieve it, we performed the following operations:

$$Average = \frac{1}{LN}\sum_{i=1}^{L}\sum_{k=1}^{N} f_{i,k} \quad (3)$$

This evaluation of the average was carried out during the training of the classifier, where L represent the dimension of the training set, that is the number of gestures composing it, and $f_{i,k}$ is the k-th feature of the i-th gesture in this set.

After this evaluation, for each classification the current feature vector can be normalized easily:

$$f^{Norm} = \frac{f}{Average} \quad (4)$$

and then sent to the algorithm for classification.

2.3. Classifier

After addressing the feature extraction, we had to decide which classifier to adopt.

In EMG pattern recognition, different algorithms have been used to assign the feature vectors. Some could perform better than others depending on the specific system they need to control. Neural Network group of classifiers are widely used [14,15] due to its expandability and ability of bearing both simple and complex cases, however, the choice of the features and the time-constraints make this classifier excessively complex and slow for the application. Other authors preferred to utilize the Fuzzy logic approach [16,17], which allows to insert user experience in the system, and to contradict itself for patterns changing; it is stated anyway that the fuzzy classifiers do not perform well with small dimensional training sets [18], which will surely regard this design.

Since the high-dimensional features space can separate clearly and linearly the 4 classes, a simple and linear classifier is preferable to be adopted; we discarded as well the k-NN algorithm (k-nearest neighbor) frequently used for its simplicity [19], since this algorithm is too slow for our time-constraints; also as a classification algorithm, some researchers used the Support Vector Machine (SVM), trained through the Sequential Minimal Optimization (SMO) which was developed by John Platt [20]. We selected this method, due to its high speed in calibration and classification, which are required by general HMI devices as well as for controlling prosthetic hands [21]; the high dimensional feature space could cause long training times in other classifiers. Moreover, as will be seen, this algorithm performs well even with a small training set and in presence of badly performed gestures, due to its elasticity parameter.

Despite the presence of more than two classes, the binary SVM is still used; indeed, the one vs all technique is adopted (Fig. 6).

We then trained 4 SVMs, each one composed by the entire training set, but with the gestures labeled differently. For each of them, one gesture type was labeled positively, while the other 3 classes negatively.

The rigidity of the margin can also be tuned by a parameter ($C > 0$), resulting in good robustness against bad performed gestures for the training set (Fig. 7), which could otherwise modify wrongly the margin.

As a similarity function to compare two feature points, the Gaussian kernel is used, in order to bear also non-linear separable classes which could originate from large training sets or future gestures addition. The Gaussian kernel is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma}} \quad (5)$$

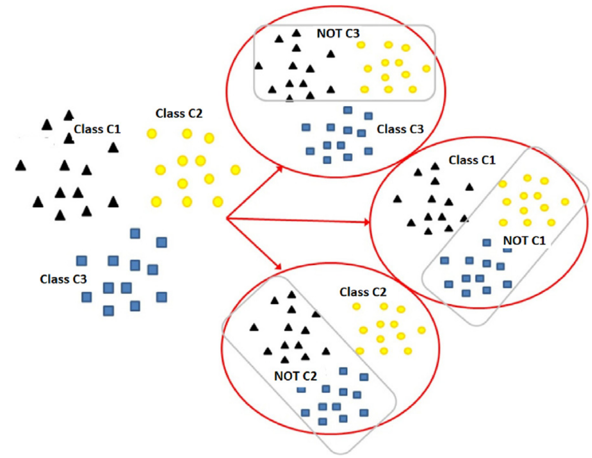


Fig. 6. One-vs-rest techniques: to bear multiple cases and use the binary SVM, the n-dimensional classification (4 in our case) is divided in n binary classifications, each one having one of the n classes compared to the rest of the training set.

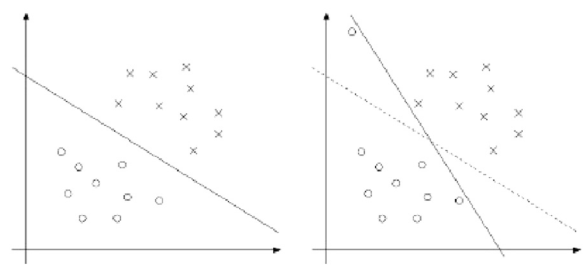


Fig. 7. Wrong instance influence: if the SVM is rigid, a performed gesture reduces significantly the margin of the hyperplane, since the algorithm searches to perfectly separate the two classes.

where $\mathbf{x}_i$, $\mathbf{x}_j$ are two feature points, and σ is the variance of the kernel.

2.4. Robustness techniques

As mentioned, it is necessary to provide to the system a good separability between the defined gestures and the other body movements (especially forearm) or line noise, in order to allow a reliable daily application to the device; to ensure this, the SVM must discard body movements from the classification. Thanks to the availability of the Raw signal and the utilization of the SVM as classifier, together with a high-dimensional feature space, we could test two techniques for the robustness achievement: Fast-Fourier Transform (FFT) and SVM threshold:

(a) Fast Fourier Transform (FFT): This technique makes use of the frequency characteristic of the EMG signal. An important source of the artifacts disturbances resides in the change of the electrode contact to the skin during the usage of the system, in particular regarding the pressure exerted from the metal layer to the arm. This causes an unbalanced amplification in the first stage of the signal pre-processing, and consequently the line noise is not eliminated by the high CMRR of the instrumentation amplifier; hence, the rectified signal rises rapidly its baseline, crossing the acquisition threshold and causing a gesture detection. However, since this classification is not caused by a hand movement, the raw signal will not present the typical frequencies of an EMG signal; in particular, above 50/60 Hz the magnitude of the Fourier Transform will have relatively low values.

The method consists then in a pre-classification, which separates line noise from the pre-defined gestures; to actuate this distinction, FFT is carried out on the samplings taken during the

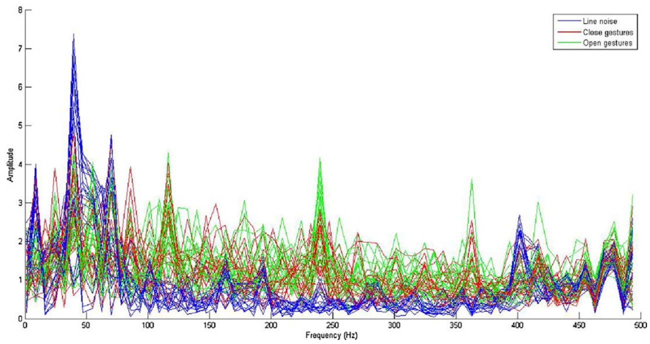


Fig. 8. FFT coefficients: The figure shows the FFT coefficients for no gestures (blue), close gestures (red), and open gestures (green), ranging from 0 to 500 Hz. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

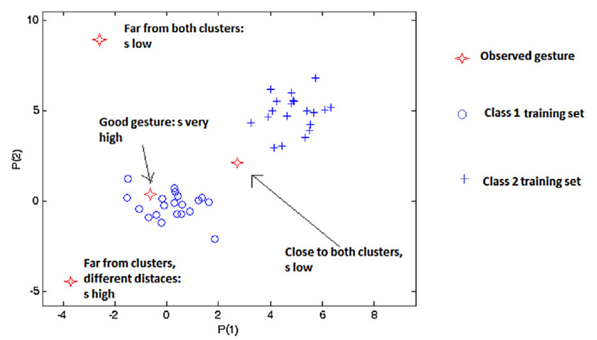


Fig. 9. SVM threshold principle: The figure shows the technique used for artifacts robustness in a binary classification. Depending on where the performed gesture lays, its s value will vary respectively.

observation windows. More in details, during the observation window, the signal is sampled with frequency $f_s = 1$ kHz for 128 ms, and after that, the FFT is operated. The features used to separate line noise from the gestures are then 16 complex coefficients of the FFT, precisely from 125 Hz to 250 Hz. As can be seen in Fig. 8, the gestures will present significantly higher values for these frequencies in relation to the line noise, and so a distinction can be made.

Although the technique ensures a clear separation from muscle contractions and line noise, it presents some important drawbacks:

- It does not provide robustness against hand or forearm movements that are not gestures.
- It is time consuming (∼100 ms).

Moreover, we noticed that for non-abrupt body movements (which do not comprehend hand movements), the rectified signals do not vary significantly, and thanks to the baseline adaptation, classifications are almost never started. Also, if the armband is tight enough, the pressure of the electrodes on the skin does not change greatly during the forearm movements.

We decided then to not adopt this technique due to the cited drawbacks, and to provide line noise robustness by designing an armband which maintains constant the pressure of the electrodes to the skin.

(b) SVM thresholds: This technique is based on the output, s, of the SVM classification, and the fact that the value of this variable increases as the performed gesture gets closer to one of the training set groups for the calibration (Fig. 9 shows the principle). It is likely that other signal patterns that are not one of the gestures will differ from all of the classes composing the training set, and so will present a low value of s; it is important then that:

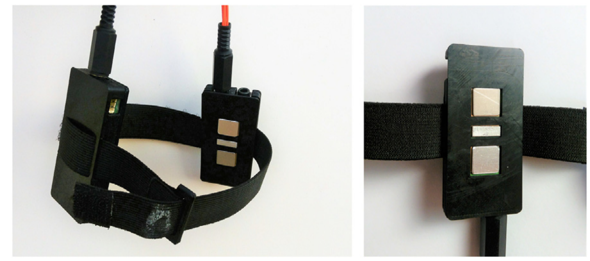


Fig. 10. Armband structure: The figure shows an elastic band with velcro attached to the 3d printed case which encloses the dry electrodes and the pre-processing board.

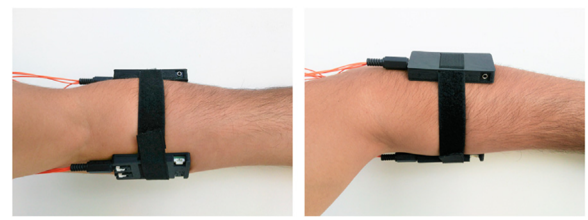


Fig. 11. Armband wearing.

- Other body movements, as well as line noise unbalance, must present different signal patterns from the ones of the predefined gestures.
- The signal patterns of the gestures must remain similar during the usage of the system, independent from the arm position, the user fatigue, the posture, etc..

The robustness is achieved by simply discarding after the classification all the gestures with s values below the prefixed thresholds.

The main challenge for reaching good results resides in the threshold settings, since we noted that each of the gestures presents different average values of s. Moreover, the s outputs are slightly dependent also on the session (that is the training set definition and the armband positioning) and, on the user; also, it is not possible to maintain fixed thresholds for every usage. The values are then regulated by an additional set of gestures (2/3 per class) given after every calibration, and the 4/5 thresholds are set equal to the lowest s outputs (respectively of every class) present in this additional set.

2.5. Setup

In order to keep the electrode's pressure constant during the utilization of the system, we designed a support structure which embeds a set of 3 electrodes, together with its preprocessing board. It is also important that the structure could be able to adapt to different arm sizes, that is, it must be elastic. For achieving these goals, a 3D printed module was designed, and wrapped around the subject's arm with a Velcro which could be tightened at will, depending on the user's arm's dimensions (Fig. 10).

Since the optimal position of the 2 channels can vary from subject to subject, it is not desired to fix their relative distance; therefore, we decided to utilize two different armbands for the system (Fig. 11), each of them embedding a single set of 3 electrodes (one channel). In this way, the positioning of each channel could be adjusted independently for finding the best setting.

To obtain accurate gesture recognition, sensors should be placed in muscles that generate clearly distinguishable signals for all the gestures. For an accurate placement of the electrodes, the muscle signal can be verified in the oscilloscope. This is a normal process that is being done by rehabilitation experts that fit prosthetic hands to amputees. However in our case, a simple manual feeling of the muscle while performing the 4 gestures found to be sufficient, and thus we did this procedure for all of the test subjects.

3. Tests

3.1. Test subjects

Seven volunteers were selected to test the system; two subjects that used the system previously several times, and at least for eight hours (here called expert user), two users who were allowed to try the system and see the signals for 2 h before the test(here called intermediate user), and the other three at their first usage(beginners). The subjects were 4 males and 3 females, all normally-limbed, with an average age of 25.6 years ($SD = 5.7$). Every test set was repeated three times for every subject, and the results were registered.

3.2. Calibration

The calibration procedure was performed only once at the beginning of the first session for each of the volunteers. In the consequent sessions performed by the same volunteer, calibration was not repeated, in order to assess session independence with a single calibration.

For the system calibration the following operations were achieved:

- Training the SVM with 10 gestures for each class.
- Setting the SVM thresholds by giving 12 additional gestures (3 for each class) and fixing the threshold values to the minimum of every 3 instances.

3.3. Test procedure

Each test was composed by two different stages.

First test was performed in order to assess the classification accuracy and the system robustness against line noise and other body movements. Subjects were asked to perform 100 gestures performing, 25 for each gesture, to evaluate how many of these gestures were classified correctly; the gestures were performed in different hand positions to simulate a daily utilization of the system. This test was performed 3 times for each subject.

Second, An additional test was performed in order to check the system tolerance to hand movements during daily activities. During 6 minutes, the user performed different activities, to check if these movements were classified correctly as "NOISE" or wrongly as one of the 4 classes. The test was repeated 3 times for all subjects. The actions performed were the following:

- Walking
- Speaking
- Climbing stairs
- Jumping
- Shoulder, arm and forearm movements
- Undefined hand gestures

4. Classification accuracy results (First test)

Tables 1–4 report the results for the 4 subjects, differentiated for each of the three sessions. In each of them, the results of the both tests are reported, by building a proper confusion matrix.

As can be seen, in all cases the accuracy for all tests was always better than 82%. It can be also seen that the expert user had excellent results of classification, always better than 95% of accuracy. In fact, there were some errors due to noise, but no misclassification between hand gestures.

For the intermediate user, the overall accuracy was over 92%, but again in all three sessions, there was only one misclassification between close and wave in.

While for the two beginners the classification accuracy was always over 82%, it can be seen that the results are improving from session 1 to session 3. Moreover the increasing accuracy from session 1 and 3 is also true for the expert and intermediate users. This shows clearly that first, the system performance is session independent, having in mind that the calibration was done only in the first session.

Second, considering the accuracy improvement over sessions for a single user, and also considering the better performance of expert and intermediate users, one can conclude that with training, users may be able to improve their results for over 90% accuracy.

Table 1
Subject 1, Expert – Male: Test 1 shows the confusion matrix for 3 sessions of the 120 performed gestures, reporting their outputs in percentage (**CL** = Close, **OP** = Open, **IN** = Wave in, **OU** = Wave out, **NS** = Noise). Test 2 shows which and how many classifications were caused by different body movements (**RN** = Random) during a 5 min sessions.

Session 1

Test 1	CL	OP	IN	OU	NS
CL	93.3%	0%	0%	0%	6.7%
OP	0%	100%	0%	0%	0%
IN	0%	0%	86.7%	0%	13.3%
OU	0%	0%	0%	100%	0%

Accuracy 95%

Test 2	CL	OP	IN	OU	NS
RN	0	1	0	1	44

Session 2

Test 1	CL	OP	IN	OU	NS
CL	96.7%	0%	0%	0%	3.3%
OP	0%	100%	0%	0%	0%
IN	0%	0%	90%	0%	10%
OU	0%	0%	0%	100%	0%

Accuracy 96.7%

Test 2	CL	OP	IN	OU	NS
RN	0	2	0	1	42

Session 3

Test 1	CL	OP	IN	OU	NS
CL	96.7%	0%	0%	0%	3.3%
OP	0%	100%	0%	0%	0%
IN	0%	0%	100%	0%	0%
OU	0%	0%	0%	100%	0%

Accuracy 99.2%

Test 2	CL	OP	IN	OU	NS
RN	0	3	0	0	27

Table 2
Subject 2, intermediate – male.

Session 1

Test 1	CL	OP	IN	OU	NS
CL	93.3%	0%	0%	0%	6.7%
OP	0%	96.7%	0%	0%	3.3%
IN	3.3%	0%	90%	0%	6.7%
OU	0%	0%	0%	86.7%	13.3%

Accuracy 91.7%

Test 2	CL	OP	IN	OU	NS
RN	10	0	0	0	1

Session 2

Test 1	CL	OP	IN	OU	NS
CL	100%	0%	0%	0%	0%
OP	0%	100%	0%	0%	0%
IN	0%	0%	90%	0%	10%
OU	0%	6.7%	0%	86.7%	6.7%

Accuracy 94.2%

Test 2	CL	OP	IN	OU	NS
RN	7	0	0	0	0

Session 3

Test 1	CL	OP	IN	OU	NS
CL	100%	0%	0%	0%	0%
OP	0%	100%	0%	0%	0%
IN	0%	0%	93.3%	0%	6.7%
OU	0%	6.7%	0%	90%	3.3%

Accuracy 95.8%

Test 2	CL	OP	IN	OU	NS
RN	4	1	0	0	0

Table 3
Subject 3, beginner – male.

Session 1

Test 1

	CL	OP	IN	OU	NS
CL	80%	0%	20%	0%	0%
OP	0%	76.7%	0%	13.3%	10%
IN	0%	0%	100%	0%	0%
OU	0%	0%	0%	86.7%	13.3%

Accuracy	85.8%

Test 2

	CL	OP	IN	OU	NS
RN	19	0	0	1	59

Session 2

Test 1

	CL	OP	IN	OU	NS
CL	93.3%	6.7%	0%	0%	0%
OP	0%	86.7%	0%	0%	13.3%
IN	0%	0%	100%	0%	0%
OU	0%	0%	0%	76.67%	23.33%

Accuracy	89.2%

Test 2

	CL	OP	IN	OU	NS
RN	10	0	0	1	34

Session 3

Test 1

	CL	OP	IN	OU	NS
CL	90%	0%	10%	0%	0%
OP	0%	86.7%	0%	13.3%	0%
IN	3.3%	0%	96.7%	0%	0%
OU	0%	0%	0%	90%	10%

Accuracy	90.8%

Test 2

	CL	OP	IN	OU	NS
RN	11	0	0	0	12

Table 4
Subject 4, beginner – female: In the 3rd session the systems wasn't placed in the exact right place and, due to that, the "Wave out" gesture was badly defined. That led to an easier recognition of the other gestures and hampered the recognition of "Wave out".

Session 1

Test 1

	CL	OP	IN	OU	NS
CL	80%	16.7%	0%	0%	3.3%
OP	0%	90%	0%	0%	10%
IN	3.3%	0%	70%	0%	26.7%
OU	6.7%	0%	0%	86.7%	6.7%

Accuracy	81.7%

Test 2

	CL	OP	IN	OU	NS
RN	76	0	0	0	12

Session 2

Test 1

	CL	OP	IN	OU	NS
CL	86.7%	6.7%	0%	0%	6.7%
OP	0%	83.3%	0%	0%	16.7%
IN	0%	0%	83.3%	3.3%	16.7%
OU	0%	16.7%	0%	76.7%	6.7%

Accuracy	82.5%

Test 2

	CL	OP	IN	OU	NS
RN	64	0	0	20	5

Session 3

Test 1

	CL	OP	IN	OU	NS
CL	93.3%	3.3%	0%	0%	3.3%
OP	0%	100%	0%	0%	0%
IN	0%	0%	96.7%	0%	3.3%
OU	20%	16.7%	0%	50%	13.3%

Accuracy	85%

Test 2

	CL	OP	IN	OU	NS
RN	55	0	0	13	7

Table 5
Overall results for test 1.

	CL	OP	IN	OU	NS
CL	91.9%	2.8%	2.5%	0%	2.8%
OP	0%	93.3%	0%	2.2%	4.4%
IN	1.1%	0%	91.4%	0%	7.8%
OU	2.8%	3.9%	0%	85.8%	8.1%

Table 6
Total percentage of correct classifications (**T** = true, correct; **F** = false, incorrect).

	T	F
TOT	90.6%	9.4%

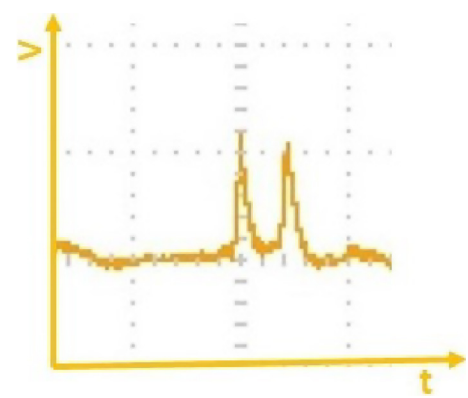


Fig. 12. Lock pattern: the lock gesture form as it appears on the oscilloscope (squares 1 V high and 1 s large) after the pre-processing block.

Overall an average accuracy of over 90% was achieved as can be seen in Table 5. The worst result was obtained in the first session of the female candidate (beginner) which achieved an accuracy of 82%, Table 4. On the opposing side, Table 1 reveals that the third session of the expert subject was the one that achieved the best accuracy, 99.2%.

5. Tolerance to noise and body movement (Second test)

Test 2 revealed that daily body movement would generate signals that could be interpreted as Gestures. In fact the expert user had very satisfactory results and had less than 3 misclassifications in each session. However the result from beginner users was not satisfactory. While it can be discussed that with training such results will improve, we decided to evaluate the possibility of adding a fifth gesture for locking/unlocking the system for two reasons:

- First, because unwanted gesture recognition might have important implications such as danger, when this system is used for controlling of machines, robots or prosthetics.
- Second, as explained earlier, for a better performance of the system against unwanted gestures, we implemented a method based on setting the SVM thresholds. While this method helped to diminish the unintentional gesture detection, it also implied that some intentional gestures would be classified as noise (as can be seen from the results of the expert and intermediate users).

Therefore, with a locking system, we can change the thresholds in favor of gesture recognition, thus we should be able to improve the accuracy of the classification and at the same time, when the system is locked, no movement will be recognized as a gesture.

5.1. Locking

In some cases the body movement was being classified as gesture even using the threshold method. To improve this, a new gesture "Lock" was added to the SVM training. During normal operation, if this gesture is detected the system becomes locked and every other gestures are ignored. If the system is locked and there is a new detection of the gesture "Lock" then it becomes unlocked. We also implemented a timer that locks the system, if the system does not receive and classify any defined gesture for a fixed duration of the time (*e.g.* 15 s). The gesture that we chose for this purpose, is a fast double wrist flexion, as can be seen in Fig. 12. When the double flexure gesture is performed, the classification of the other 4 ges-

Table 7
Overall results for test 2

	CL	OP	IN	OU	NS
RN	105	1	0	2	116

Table 8
Subject 2, intermediate – male.

Test 1 – Session 1

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Test 1 – Session 2

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Test 1 – Session 3

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Test 2

	Session 1 Detections	Session 1 Unlocks	Session 2 Detections	Session 2 Unlocks	Session 3 Detections	Session 3 Unlocks
RN	28	0	25	0	29	0

Table 9
Subject 3, beginner – male.

Test 1 – Session 1

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Test 1 – Session 2

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Test 1 – Session 3

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	100%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	100%			

Table 10
Subject 4, beginner – female.

Test 1 – Session 1

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	0%	83.3%	0%	16.7%
IN	3.3%	0%	96.7%	0%
OU	0%	0%	0%	100%
Accuracy	95%			

Test 1 – Session 2

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	13.3%	86.67%	0%	0%
IN	6.7%	0%	93.3%	0%
OU	0%	0%	0%	100%
Accuracy	95%			

Test 1 – Session 3

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	10%	90%	0%	0%
IN	0%	0%	100%	0%
OU	0%	0%	0%	100%
Accuracy	97.5%			

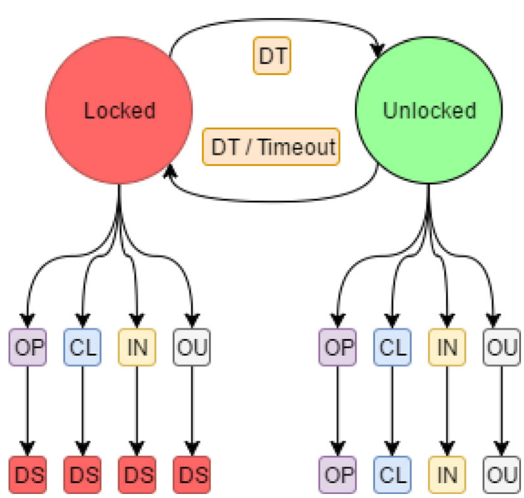


Fig. 13. High level acting algorithm of the locking command. When locked, any detected gesture is classified as noise, except the double wrist flexion, which is classified as unlock. During the unlock status, gestures are classified normally, until an unlock gesture is performed (again double wrist flexion), or no gesture is classified during 15 s (automatic system lock).

tures is locked/unlocked (dependently on the state of the system – Fig. 13).

Therefore, the system should detect five gestures, including hand closing, hand opening, wrist extension, wrist flexion and double wrist flexion. Using this new approach, we updated the algorithm and the testing sessions were repeated using the same intermediate/beginner subjects. Even though we expected that adding a new gesture could make the system more complex and reduce the classification accuracy, this was not the case. As can be seen the classification result was improved substantially. This is mainly because, with the addition of the locking gesture, the thresholds could be removed and gestures are not classified anymore as "Noise". Therefore, the accuracy of the recognition was greatly improved. Subject 2 (Intermediate) and subject 3 (beginner) were able to increase their accuracy to 100%. Subject 4 was able to achieve 97.5% as noted in Tables 8–10 respectively. Along with the improvement in accuracy, there was also a decrease in wrong classifications as can be compared in Tables 5 and 11. The average percentage of correct classifications was also improved substantially from 90.6% to 98% as can be seen in Tables 6 and 12, even though only the Intermediate/beginner subjects were considered.

To evaluate the robustness of the locking method, test 2 was changed to a new test. In the new test, the subject locked the system, and did the same movements that were previously described. We then evaluated if any of those movements are classified as double flexion for unlocking. As it was expected, using the double flexion, is very specific and hard to be misclassified. Subject 2 performed the test 2 and the results are shown in Table 8. While

Table 11
Overall results for test 1.

	CL	OP	IN	OU
CL	100%	0%	0%	0%
OP	2.6%	95.6%	0%	1.9%
IN	1.1%	0%	98.9%	0%
OU	0%	0%	0%	100%

Table 12
Total percentage of correct classifications (**T** = true, correct; **F** = false, incorrect).

	T	F
TOT	98.6%	1.4%

doing the tasks of this test the system detected several gestures. The number of detected and disregarded gestures were between 25 and 29, which is significantly higher for the same subject. This is due to elimination of the threshold. But since the system was locked, these gestures were disregarded and the system was never unlocked, showing that this procedure is very effective.As a result, the “Lock” method increased both the classification’s effectiveness and it’s immunity to undesired forearm movement.

5.2. Classification with five gestures

Finally, we performed additional tests in order to evaluate how adding a new gesture, i.e. lock, affects the classification accuracy. That is, if considering the double flexion as a gesture, how this affects the classification accuracy. Similar to previous cases a beginner (female, 22), an intermediate (male, 22) and expert (female, 22) volunteers tested the systems. None of the volunteer are same as the previous tests. One session consist of 25 gestures performed per each class. Results are shown in Table 13–15. As can be seen, even after the addition of the fifth gesture the system resulted in a classification accuracy of over 83% for the beginner, 89% for the intermediate and 95% for the expert user. Another interesting aspect to observe is the evolution of the volunteers after each session. For instance the beginner reached from 83% on the first session to 89% in the third session.

6. Discussion and conclusions

The overall results were satisfactory in terms of classification success, calibration time and session independence which were demonstrated. The real-time classification constraints were respected, since the above satisfying results were obtained with a small training set and feature space dimensions, which brought to a reasonable processing time for the calibration stage (<30 s) and a fast system response to a gesture performing (∼10 ms after completion of the gesture). In addition, we studied methods for eliminating the misclassification of unwanted gestures during the natural human activities. Two methods were applied, SVM threshold and Locking.

The SVM threshold showed excellent results with expert users, but it was also observed that the beginner results can be improved by several measures, such as training the user, and also inclusion of additional filters with sensor fusion (e.g. by addition of an IMU).

The classification accuracy for an expert user ranged from 95% to 100 % 0. An evaluation of 8 channels Myo armband showed a classification accuracy of 97 % 0 [22]. Also a recent review of classification techniques for EMG signals showed a classification accuracy ranging from 88 % 0 to 98 % 0 [23]. 98 % 0 was achieved on a pre-recorded data with a 4 channel EMG system on a single subject, who per-

Table 13
Five gesture classification results including Double Flexion(DF). female-beginner.

Test 1

Session 1

	CL	OP	IN	OU	DF
CL	84%	0%	12%	0%	4%
OP	0%	72%	0%	0%	4%
IN	20%	0%	76%	0%	4%
OU	0%	8%	0%	92%	0%
DF	0%	4%	0%	4%	92%
Accuracy	83,2%				

Session 2

	CL	OP	IN	OU	DF
CL	80%	0%	20%	0%	0%
OP	0%	92%	0%	4%	4%
IN	12%	0%	84%	0%	4%
OU	0%	8%	0%	92%	0%
DF	0%	4%	4%	0%	92%
Accuracy	88%				

Session 3

	CL	OP	IN	OU	DF
CL	68%	0%	32%	0%	0%
OP	0%	92%	0%	8%	0%
IN	8%	0%	92%	0%	0%
OU	0%	4%	0%	96%	0%
DF	0%	4%	0%	0%	96%
Accuracy	88,8%				

Table 14
Five gesture classification results including Double Flexion(DF). male-intermediate.

Test 2

Session 1

	CL	OP	IN	OU	DF
CL	84%	8%	8%	0%	0%
OP	0%	96%	0%	4%	0%
IN	8%	0%	84%	0%	8%
OU	4%	4%	0%	92%	0%
DF	0%	0%	8%	0%	92%
Accuracy	89,6%				

Session 2

	CL	OP	IN	OU	DF
CL	84%	8%	8%	0%	0%
OP	4%	96%	0%	0%	0%
IN	0%	4%	96%	0%	0%
OU	0%	8%	0%	92%	0%
DF	0%	0%	0%	0%	100%
Accuracy	93,6%				

Session 3

	CL	OP	IN	OU	DF
CL	96%	4%	0%	0%	0%
OP	0%	100%	0%	0%	0%
IN	4%	4%	88%	0%	4%
OU	0%	4%	0%	96%	0%
DF	0%	0%	0%	0%	100%
Accuracy	96%				

Table 15
Five gesture classification results including Double Flexion(DF). female-expert.

Test 3

Session 1

	CL	OP	IN	OU	DF
CL	92%	0%	8%	0%	0%
OP	0%	100%	0%	0%	0%
IN	12%	4%	84%	0%	0%
OU	0%	0%	0%	100%	0%
DF	0%	0%	0%	0%	100%
Accuracy	95,2%				

Session 2

	CL	OP	IN	OU	DF
CL	100%	0%	0%	0%	0%
OP	0%	100%	0%	0%	0%
IN	12%	0%	88%	0%	0%
OU	0%	0%	0%	96%	4%
DF	4%	0%	0%	0%	96%
Accuracy	96%				

Session 3

	CL	OP	IN	OU	DF
CL	92%	0%	8%	0%	0%
OP	0%	100%	0%	0%	0%
IN	4%	0%	96%	0%	0%
OU	0%	0%	0%	100%	0%
DF	0%	0%	0%	0%	100%
Accuracy	97,6%				

formed 121 sessions during 121 days. A direct comparison between previous works is not possible due to utilization different setups, gestures, number of subjects, and the expertise level of the subjects. Nevertheless, the classification accuracy of this study is comparable with the recent state of the art EMG classification systems, with the advantage of using a minimalistic hardware approach.

The System locking instead assured a total disturbance rejection. 0% of the other body movements were classified as an “Unlock” gesture, maintaining so the system locked and letting the user move freely. Moreover, the locking method proved to improve substantially the classification accuracy. Even though a new gesture was added, the classification accuracy was improved in all subjects, mainly due to removing the class of “Noise”, as it was implemented in the SVM threshold method. The disadvantage of the lock gesture, is evidently making the user interface less natural by adding an additional step. Therefore depending on the level of the safety required by an application, one can opt between the threshold or locking method.

EMG sensors, have been previously used in health sector for monitoring muscular activities and commanding the prosthetic hands. Recently they are being used for general purpose wearable HMI systems. Therefore taking a minimalistic approach in the hardware design, and reliability of performance are important factors for their successful application in the HMI domain. Classification of five gestures with the minimalistic two channels EMG system is a step toward this goal. Future works include further miniaturization of the hardware, to evaluate alternative electrodes, and evaluation of other classification methods such as KNN and LDA. Also, regarding the features that are fed into classifier, in this work we used an average function. However, other parameters such as standard deviation and variance can be also fed to the classifier, which is also the subject of future works.

Acknowledgements

This work was partially supported by the national funds of the Foundation of Science and Technology of Portugal through the CMU-Portugal project Stretchtronics (Nr. CMUP-ERI/TIC/0021/2014), and MATIS(CENTRO-01-0145-FEDER-000014) co-financed by the European Regional Development Fund (FEDER) through “Programa Operacional Regional do Centro” (CENTRO2020).

References

[1] G.R. Naik, D.K. Kumar, et al., Twin svm for gesture classification using the surface electromyogram, IEEE Trans. Inf. Technol. Biomed. 14 (2) (2010) 301–308.
[2] A.B. Ajiboye, et al., A heuristic fuzzy logic approach to emg pattern recognition for multifunctional prosthesis control, IEEE Trans. Neural Syst. Rehabil. Eng. 13 (3) (2005) 280–291.
[3] D. Yang, J. Zhao, Y. Gu, L. Jiang, H. Liu, Emg pattern recognition and grasping force estimation: improvement to the myocontrol of multi-dof prosthetic hands, in: IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS 2009, IEEE, 2009, pp. 516–521.
[4] A.H. Al-Timemy, G. Bugmann, J. Escudero, N. Outram, Classification of finger movements for the dexterous hand prosthesis control with surface electromyography, IEEE J. Biomed. Health Inform. 17 (3) (2013) 608–618.
[5] C. Cipriani, C. Antfolk, M. Controzzi, G. Lundborg, B. Rosén, M.C. Carrozza, F. Sebelius, Online myoelectric control of a dexterous hand prosthesis by transradial amputees, IEEE Trans. Neural Syst. Rehabil. Eng. 19 (3) (2011) 260–270.
[6] V.D.S.P.S. Bitzer, Learning emg control of a robotic hand: towards active prostheses, in: Proceedings 2006 IEEE International Conference on Robotics and Automation, ICRA 2006, IEEE, 2006, pp. 2819–2823.
[7] L. Pan, D. Zhang, J. Liu, X. Sheng, X. Zhu, Continuous estimation of finger joint angles under different static wrist motions from surface {EMG} signals, Biomed. Signal Process. Control 14 (2014) 265–271.
[8] F. Riillo, L. Quitadamo, F. Cavrini, E. Gruppioni, C. Pinto, N.C. Pastò, L. Sbernini, L. Albero, G. Saggio, Optimization of emg-based hand gesture recognition: supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees, Biomed. Signal Process. Control 14 (2014) 117–125.
[9] H.M. Al-Angari, G. Kanitz, S. Tarantino, C. Cipriani, Distance and mutual information methods for {EMG} feature and channel subset selection for classification of hand movements, Biomed. Signal Process. Control 27 (2016) 24–31.
[10] M.T. Wolf, C. Assad, A. Stoica, K. You, H. Jethani, M.T. Vernacchia, J. Fromm, Y. Iwashita, Decoding static and dynamic arm and hand gestures from the jpl biosleeve, Aerospace Conference, 2013 IEEE (2013) 1–9.
[11] Thalmic Labs, Myo gesture control armband.
[12] T. Peters, An assessment of single-channel emg sensing for gestural input, in: Dartmouth Computer Science Technical Report, 2015/767, 2014, pp. 1–14.
[13] S. Mane, R. Kambli, F. Kazi, N. Singh, Hand motion recognition from single channel surface emg using wavelet & artificial neural network, Procedia Comput. Sci. 49 (2015) 58–65.
[14] A. Soares, A. Andrade, E. Lamounier, R. Carrijo, The development of a virtual myoelectric prosthesis controlled by an emg pattern recognition system based on neural networks, J. Intell. Inf. Syst. 21 (2) (2003) 127–141.
[15] X. Zhang, X. Chen, W.-h. Wang, J.-h. Yang, V. Lantz, K.-q. Wang, Hand gesture recognition and virtual game control based on 3d accelerometer and emg sensors, Proceedings of the 14th International Conference on Intelligent User Interfaces, ACM (2009) 401–406.
[16] F.H.Y. Chan, Y.-S. Yang, F.K. Lam, Y.-T. Zhang, P.A. Parker, Fuzzy emg classification for prosthesis control, IEEE Trans. Rehabil. Eng. 8 (3) (2000) 305–311.
[17] L.-X. Wang, Adaptive fuzzy systems and control: design and stability analysis, IEEE Trans. Fuzzy Syst. 1 (2) (1993) 146–155.
[18] S. Godil, M. Shamim, S. Enam, U. Qidwai, Fuzzy logic: A “simple” solution for complexities in neurosciences? Surgical Neurology International 2 (2011) 24.
[19] A.H.M. Al-Faiz, M.Z. Ali, A. Abduladhem, A k-nearest neighbor based algorithm for human arm movements recognition using emg signals, 1st International Conference on Energy, Power and Control EPC-IQ, IEEE (2010) 159–167.
[20] J.C. Platt, Sequential minimal optimization: a fast algorithm for training support vector machines, in: Technical report msr-tr-98-14, Microsoft Research, 1998, pp. 1–21.
[21] M. Tavakoli, C. Benussi, J.L. Lourenco, Single channel surface emg control of advanced prosthetic hands: a simple, low cost and efficient approach, Expert Syst. Appl. 79 (2017) 322–332.
[22] A. Boyali, N. Hashimoto, Spectral collaborative representation based classification for hand gestures recognition on electromyography signals, Biomed. Signal Process. Control 24 (2016) 11–18.
[23] N. Nazmi, M.A. Abdul Rahman, S.-I. Yamamoto, S.A. Ahmad, H. Zamzuri, S.A. Mazlan, A review of classification techniques of emg signals during isotonic and isometric contractions, Sensors 16 (8) (2016) 1304.