

Classifying sEMG-based Hand Movements by Means of Principal Component Analysis

Milica S. Isaković, Nadica Miljković, and Mirjana B. Popović

Abstract — In order to improve surface electromyography (sEMG) based control of hand prosthesis, we applied Principal Component Analysis (PCA) for feature extraction. The sEMG data from a group of healthy subjects (downloaded from free NINAPRO database) comprised the following sets: three grasping, eight wrist, and eleven finger movements. We tested the accuracy of a simple quadratic classifier for two sets of features derived from PCA. Preliminary results suggest that the first two principal components do not guarantee successful hand movement classification. The hand movement classification accuracy significantly increased with using three instead of two features, in all three sets of movements and throughout all subjects.

Keywords — feature extraction, healthy subjects, grasp, principal component analysis, surface electromyography.

I. INTRODUCTION

THE surface electromyography (sEMG) signals have been used as input to self-powered hand prosthesis control systems for over 40 years [1]. In these systems, the sEMG signals are recorded from muscles of interest, and the user is asked to voluntarily control the sEMG amplitude in a manner that enables effective and intuitive prosthesis control. The main advantages of sEMG based prosthesis control are: noninvasive signal detection, a simple detection circuit, relatively bald signal generation, and simple adaptation to proportional control [1]. Over the past decades, the functionality requirements of the hand prosthesis have increased, leading to more complex sEMG control paradigms [1]. Novel algorithms led to increased

acquisition, processing, and classification complexity, and are often followed by a durable and exhaustive process of learning how to control the prosthesis *via* sEMG signals [2]. Moreover, the results in [3] suggest that hand prosthesis functionality has not been significantly improved over the past decades – even providing a proof that some of the novel devices performed worse than their predecessors from 1987. As a result of a relatively poor functionality (e.g. three degrees of freedom [2]), up to 50% of the amputees do not use sEMG controlled prosthesis regularly [4]. In order to increase the performance of the sEMG based control of hand prosthesis, it is essential to improve the acquisition setup, sEMG signal preprocessing, feature extraction, and movement classification. Pattern-recognition-based myoelectric control has been extensively researched and various paradigms of feature extraction (time domain features [5], the short-time Fourier transform [6], wavelet transform [6], autoregressive coefficients [7]–[8], and cepstral coefficients [8]) and classification (Bayesian classifier [9], artificial neural networks [9], support vector machine [2], [10], hidden Markov models [11], and fuzzy logic [12]) have been proposed.

NINAPRO database was created with the aim to provide the biorobotics community with a benchmarking tool for novel myoelectric hand prosthesis control methods [13]. The main goal of this database is improvement in sEMG based prosthesis control by a free and online available test set of classification algorithms and control signals. This database is a result of NINAPRO project [13] and can be downloaded from NINAWEB – the web interface of the NINAPRO project [14]. Two versions of the database were presented, as well as the various techniques used for data preprocessing, feature extraction, and hand posture classification [2], [15]–[16]. One of the techniques offered by NINAPRO project is the use of Principal Component Analysis (PCA) for feature extraction. This technique has been frequently used to extract the features from sEMG signals [17]–[19].

PCA is a multivariate statistical procedure used to reduce the dimensionality of a multidimensional data set, while retaining as much as possible of the variation present in the transformed data set termed Principal Components (PCs) [20]. PCA is used for feature extraction from multi-channel sEMG signals recorded for various hand postures and grasps in healthy subjects [2]. The results of NINAPRO project included the PCs feature classification by multi-class Least-Squares Support Vector Machine (LS-SVM) classifier with Radial Basis Function

Paper received April 3, 2015; accepted April 10, 2015. Date of publication July 15, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Branimir Reljin.

This paper is a revised and expanded version of the paper presented at the 22th Telecommunications Forum TELFOR 2014.

The work on this project was partly supported by the Ministry of education, science, and technological development, Republic of Serbia, grant No. 175016.

Corresponding author Milica S. Isaković is with the University of Belgrade - School of Electrical Engineering, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia, and with Tecnalnia Serbia Ltd., Vladetina 13/6, 11120 Belgrade, Serbia (phone: 381-62-383-585; e-mail: milicaisakovic@hotmail.com).

Nadica Miljković is with the University of Belgrade - School of Electrical Engineering, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia, and with Tecnalnia Serbia Ltd., Vladetina 13/6, 11120 Belgrade, Serbia (phone: 381-11-3218-348, e-mail: nadica.miljkovic@etf.rs).

Prof. Mirjana B. Popović is with the University of Belgrade - School of Electrical Engineering, Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia (phone: 381-11-3218-345, e-mail: mpo@etf.rs).

(RBF) kernel [2].

In this paper, we analyzed three subsets of data from the first version of the NINAPRO database (three grasping, eight wrist, and eleven finger movements) and used PCA for feature extraction as in [2]. Our goal was to examine if there is significant variability in the percentage of variance accounted for each PC among movements and subjects and ensure that the proportion of variance accounted for the first two PCs (as in [2]) isn't always sufficient for successful hand movement classification. We presented the comparison of classification accuracy when using two different sets of features (two PCs *vs.* three PCs). For classification, we used a quadratic classifier, since it is reasonably efficient and rather simple for implementation and feature comparison. We applied a piecewise quadratic classifier – a set of quadratic classifiers that provide increased flexibility for a multiclass problem such as hand movement classification [21].

II. METHODOLOGY

A. NINAPRO database

Data analyzed in this paper are publicly available from the first version of NINAPRO database. sEMG data were gathered from 27 intact subjects (20/7 male/female, 25/2 right-/left-handed, aged 28 ± 3.4 years) while performing 10 repetitions of 52 hand movements, using 10 active double-differential *MyoBock* 13E200 (Otto Bock HealthCare GmbH, Duderstadt, Germany) surface EMG electrodes. Eight uniformly spaced sEMG electrodes were placed beneath the elbow at a fixed distance from the *radio-humeral* joint, with two additional electrodes placed on the flexor and extensor muscles. The electrodes provide an amplified, bandpass-filtered in the range 0-25Hz, and rectified version of the raw sEMG signal. Acquisition setup and experimental protocol are detailed in [2], [22].

In this paper, we considered three sets of movements from NINAPRO database – three grasping movements, eight wrist movements, and eleven finger movements. Descriptions of the selected movements can be found in Table 1.

TABLE 1: DESCRIPTIONS OF SELECTED MOVEMENTS [2].

	#	Description
Grasping movements	1	Spherical power grasp
	2	Three finger sphere grasp
	3	Two finger prismatic grasp
Wrist movements	1-2	Wrist supination and pronation (rotation axis – middle finger)
	3-4	Wrist supination and pronation (rotation axis – little finger)
	5-6	Wrist flexion and extension
	7-8	Wrist radial and ulnar deviation
Finger movements	1-2	Index flexion and extension
	3-4	Middle flexion and extension
	5-6	Ring flexion and extension
	7-8	Little finger flexion and extension
	9-10	Thumb flexion and extension
	11	Flexion of ring and little finger

Signal preprocessing, PCA algorithm and the applied

classifier were implemented in MATLAB R2013a (The MathWorks, Natick, MA, USA).

B. Signal preprocessing

We performed the sEMG signal preprocessing procedure as suggested in [2]. The first step of sEMG signal preprocessing is data synchronization by linearly interpolating all data to 100 Hz. Further preprocessing consists of three steps:

- 1) Filtering – sEMG signals are low-pass filtered using zero-phased 2nd order Butterworth filter with a cutoff frequency of 1 Hz.
- 2) Segmentation – Each repetition of a movement followed with the rest is divided into three equally sized segments. The samples contained in the central segment of both movement (posture) and rest (pause) are retained only.
- 3) Averaging – Samples from central segments are averaged. As a result, each movement for each subject contains 20 samples (i.e. one per movement and the following rest repetition).

Finally, the data for individual subjects and movements are normalized, so that each sEMG signal has a zero mean and unit standard deviation. This is a necessary step for PCA application.

C. Principal Component Analysis

PCA is a statistical procedure used to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set [20]. PCA uses an orthogonal transformation to convert a set of observations of correlated variables into a new set of variables, the PCs, which are uncorrelated and ordered so that the first few retain most of the variation present in the original data set. This technique can significantly reduce a classification error in pattern-recognition-based myoelectric control [23].

As suggested in [2], we used PCA to extract features for the movement classification. The same processing technique provides visual information of the relation between different movements and sEMG signals. In order to give insights into the variability of the new data set, we calculated the percentage of variance accounted for each PC as a ratio of the corresponding eigenvalue and the total variance. This additional analysis examines variability in the percentage of variance explained by each PC among different movements and subjects, which is important information for further analysis (e.g. selection of hand movements that can be analyzed together using the same PCs, feature selection for classification, choice of number of electrodes used to record sEMG signals).

D. Classification

We designed a piecewise quadratic classifier for the classification of selected hand movements. For each subject and each set of movements (grasping, wrist, and finger movements) an individual classifier was designed, and 10 repetitions of each movement were classified using two sets of features – the first two PCs and the first three PCs.

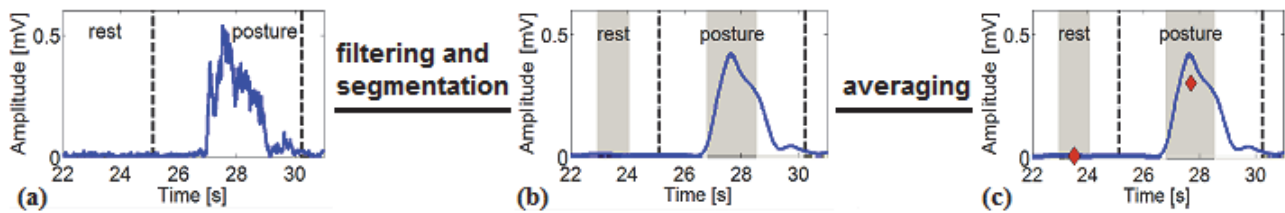


Fig. 1. Diagram of the sEMG signal preprocessing: (a) Raw sEMG signal; (b) sEMG signal after filtering and segmentation (the two distinct conditions: rest and posture are in gray central areas); (c) Filtered sEMG signal after averaging over the central segment (red diamond markers). See text for more details.

E. Statistical analysis

To evaluate the performance of the classifier, classification accuracy (percentage of correctly classified repetitions of movements and the rest) was computed for each of 27 subjects, all three sets of movements, and both sets of features. Statistically significant differences in the mean classification accuracy between two sets of features for all three sets of movements were tested using one-way ANOVA. Levene's test for equal variances was performed to assess if there were statistically significant differences in the variance of classification accuracies when two or three features were used, for all three sets of movements. The threshold for the statistical significance was adopted at $p < 0.05$.

III. RESULTS

Fig. 1 illustrates the sEMG signal preprocessing steps consisting of filtering, segmentation, and averaging. The first panel (a) shows the raw sEMG signal – one repetition of the index finger flexion and the following rest for subject no. 1 and one out of 10 electrodes. The second panel (b) shows sEMG signal after filtering and segmentation. Gray shaded areas indicate the central segments. The third panel (c) shows the result of averaging over the central segment.

The percentage of variance accounted for each out of 10 principal components for the set of three grasping movements (Table 1) and one subject (no. 4) is shown in Fig. 2. There are significant variations in the percentage of variance explained by each PC among three different grasping movements for the same subject. For example, the first PC for the grasp no. 3 explains 88% of variance, while the first two PCs for grasp no. 2 explain 85% of variance, and the first three PCs for grasp no. 1 explain 80% of variance.

The result of hand movement classification using the first two principal components as features is illustrated in Fig. 3 for the set of grasping movements in one subject (no. 4), resulting in a classification accuracy of 80%. The result of hand movement classification for the same set of movements and the same subject as in Fig. 3, but with the first three PCs used as features is given in Fig. 4. The classification accuracy in this case is 90%.

Fig. 5. demonstrates the classification accuracies averaged over all 27 subjects when considering three sets of movements (three grasping movements, eight wrist movements, and eleven finger movements, Table 1) and two sets of features (the first two PCs and the first three PCs).

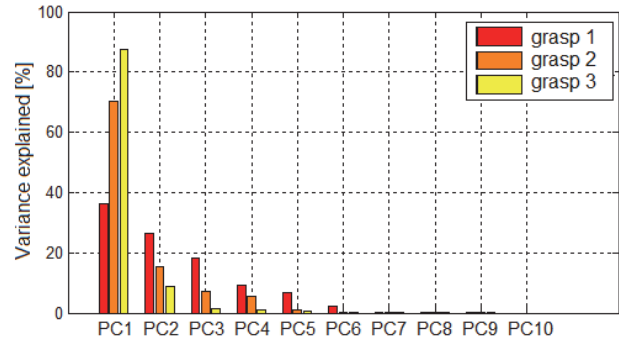


Fig. 2. Percentage of variance accounted for each PC for three grasping movements in one subject (no. 4).

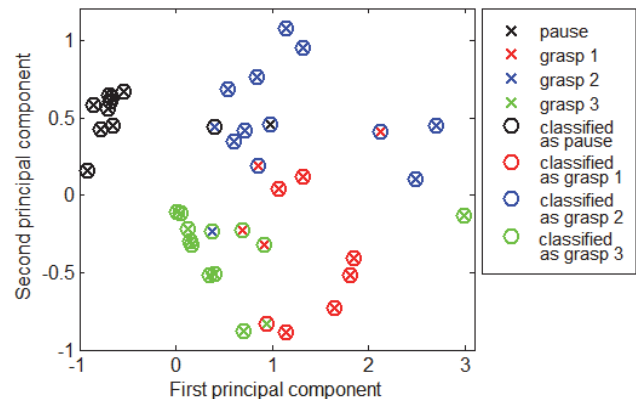


Fig. 3. Two-dimensional plot of the first two PCs of sEMG signals and the result of hand movement classification for the set of grasping movements in one subject (no. 4) using the first two PCs.

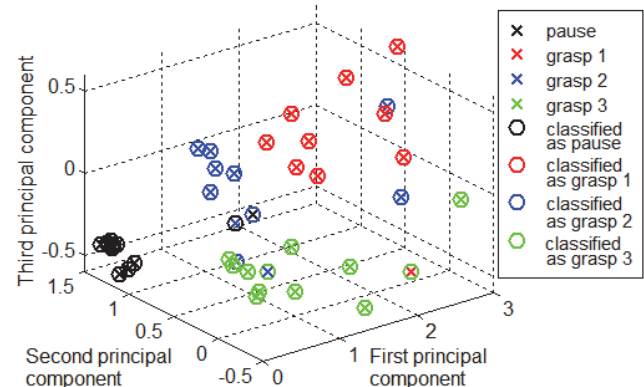


Fig. 4. Three-dimensional plot of the first three PCs of sEMG signals and the result of hand movement classification for the set of grasping movements in one subject (no. 4) using the first three PCs.

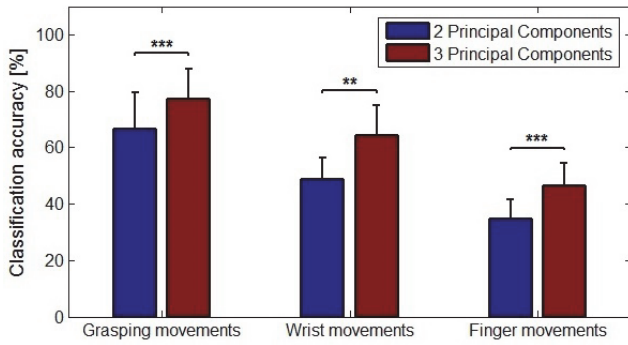


Fig. 5. The classification accuracy for the three sets of movements when using two and three PCs. Graphs report the mean and the standard deviation over all 27 subjects. Horizontal bars with asterisks indicate the statistically significant difference in mean classification accuracy between different numbers of PCs used as classification features. (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

Classification of three grasping movements results in the accuracy of $66.7 \pm 12.6\%$ when using two features, and this accuracy increases to $77.2 \pm 10.7\%$ when using three features. Classification accuracy for the set of wrist movements enhances from $48.7 \pm 7.8\%$ when using the first two PCs to $64.5 \pm 10.3\%$ when the third PC is added. In case of the set of eleven finger movements, the classification accuracy is $34.6 \pm 7.1\%$ when using the first two PCs, increasing to $46.6 \pm 7.8\%$ when using the first three PCs. Therefore, the classification accuracy averaged over all 27 subjects for all three sets of hand movements increases when using three features. The results of detailed analysis (not shown here) for each of 27 subjects and all three sets of movements showed the same trend of classification accuracy. The results of one-way ANOVA suggest that the differences in mean classification accuracy between two conditions (two PCs vs. three PCs) for all three sets of hand movements are statistically significant. The results of Levene's test for equal variances show that there are no statistically significant differences in the variance of classification accuracies when two or three features were used, for all three sets of movements.

IV. DISCUSSION AND CONCLUSIONS

Hypotheses, that there is significant variability in the percentage of variance explained by each PC among movements and subjects, and that the proportion of variance explained by the first two PCs isn't always sufficient for successful hand movement classification, are confirmed in our analysis. We have shown that mean classification accuracy is significantly increased when using the first three instead of the first two PCs of sEMG signals as features for hand movement classification, for all three sets of hand movements. Furthermore, we have shown that, for all sets of movements, there are no statistically significant differences in the variance of classification accuracy between two conditions. We assume that these variations are the consequence of great inter-subject variability, and therefore cannot be improved by solely increasing the number of classification features.

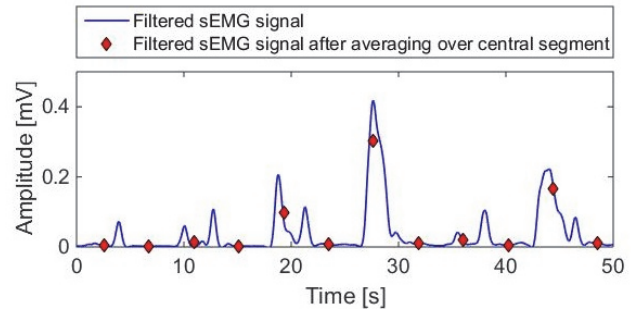


Fig. 6. Six repetitions of the index finger flexion and the following rest for subject no. 1 and one out of 10 electrodes, demonstrating intra-subject variability of sEMG signal.

The selection of sEMG signal preprocessing technique, especially movement segmentation, largely affects the final results of PCA and the classification. The suggested method of segmentation (dividing each repetition of movement into three segments and retaining only the central segment) is simple and mostly successful. However, there are repetitions of several movements that are shifted from the center of the segment, have different amplitudes or characteristic shapes, as shown in Fig. 6 in the example of index finger flexion for subject no. 1. This may occur due to electrode movements during signal recording, fatigue or subject's lack of focus. Classification error might be in such cases reduced by using one of muscle onset detection algorithms proposed in [24]. Low-pass filtering frequency affects the classification accuracy, so the optimization of the filtering frequency for each subject could lead to an improvement of the classification, as shown in [25].

An improvement of the PCA algorithm might also lead to the reduction of classification error. A novel individual PCA tuning algorithm for use prior to the pattern recognition classification of hand movements was introduced in [23]. The authors claimed that this processing technique significantly reduced classification errors for both intact-limbed and amputee subjects.

Our analysis has confirmed that the percentage of variance accounted for the third PC cannot be neglected for some movements (such as grasp no. 1 in Fig. 2). That is why one should have in mind the number of PCs used as classification features. On the other hand, the online implementation should also be taken into account when introducing a larger number of PCs to a selected classifier.

Percentage of variance accounted for the last five PCs (PC6–PC10) is less than 5% for all subject/movement combinations, indicating that recordings might be performed by a different sEMG setup (a decreased number of sEMG channels). Future work should include recording with more sEMG channels in order to select only relevant electrodes for the specific movement.

Another important issue that can affect hand movement classification accuracy is displacement of electrodes [26]. Our preliminary results suggest that an appropriate choice of the number of PCs can modify the effects of electrode array translation [27].

It is important to emphasize that, when three features are used instead of two features, the classification accuracy is increased for each out of 27 subjects and all three sets of movements. For the set of eleven finger movements, the classification error is relatively large regardless of the number of PCs used for classification, due to simplicity of a quadratic classifier, as well as to the large number of movements. Nevertheless, there is an apparent increase of the classification accuracy when using the first three PCs ($46.6 \pm 7.8\%$) compared to the first two PCs ($34.6 \pm 7.1\%$). Further, this error is approximately twice as much as error obtained for the set of grasping movements in both analyzed cases. This result implies that simpler classification methods may be more appropriate for smaller sets of movements.

Limitations of the presented study are the following:

- 1) A larger sample than the one available from the NINAPRO database should be taken into account for a more valid conclusion on the number of PCs.
- 2) sEMG recordings from transradial amputees should also be tested for the calculation of the optimal number of PCs.
- 3) Other appropriate classifiers should be examined for a final decision on dimensionality reduction when PCA feature extraction is applied. For example, the solution applied in [28] shows that the success rate of hand movement classification is significantly high when using a cascaded-structured classifier.

When considering sEMG signal as an appropriate candidate for the control of prosthesis, the selection of a suitable 1) measurement setup, 2) protocol design, 3) preprocessing method, 4) feature extraction technique, and 5) classifier design should altogether be taken into account for a valid recommendation and successful application.

REFERENCES

- [1] P. Parker, K. Englehart, and B. Hudgins, "Myoelectric signal processing for control of powered limb prosthesis," *Journal of Electromyography and Kinesiology*, vol. 16, no. 6, pp. 541-548, Dec. 2006.
- [2] M. Atzori, A. Gijsberts, S. Heynen, A.-G. Mittaz Hager, O. Deriaz, P. Van der Smagt, C. Castellini, and B. Caputo, "Building the NinaPro Database: a Resource for the Biorobotics Community," in *IEEE International Conference on Biomedical Robotics and Biomechatronics*, 2012, pp. 1258-1265.
- [3] G. Smith, R. M. Bongers, C. K. Van der Sluis, and D. H. Plettenburg, "Efficiency of voluntary opening hand and hook prosthetic devices: 24 years of development?," *Journal of Rehabilitation Research and Development*, vol. 49, no. 4, pp. 523-534, 2012.
- [4] D. J. Atkins, "Epidemiologic overview of individuals with upper-limb loss and their reported research priorities," *Journal of Prosthetics & Orthotics*, vol. 8, no. 1, pp. 2-11, 1996.
- [5] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848-854, 2003.
- [6] K. Englehart, B. Hudgins, and P. A. Parker, "A wavelet-based continuous classification scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 302-311, 2001.
- [7] D. Graupe, J. Salahi, and K. Kohn, "Multifunctional prosthesis and orthosis control via microcomputer identification of temporal pattern differences in single-site myoelectric signals," *Journal of Biomedical Engineering*, vol. 4, pp. 17-22, 1982.
- [8] W. J. Kang, J. R. Shiu, C. K. Cheng, J. S. Lai, H. W. Tsao, and T. S. Kuo, "The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition [movements classification]," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 8, pp. 777-785, 1995.
- [9] K. Englehart, B. Hudgins, P. A. Parker, and M. Stevenson, "Classification of the myoelectric signal using time-frequency based representations," *Medical Engineering & Physics*, vol. 21, pp. 431-438, 1999.
- [10] M. A. Osekoi, and H. Hu, "Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1956-1965, 2008.
- [11] A. D. C. Chan, and K. Englehart, "Continuous myoelectric control for powered prostheses using hidden Markov models," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 1, pp. 121-124, 2005.
- [12] A. B. Ajiboye, and R. F. Weir, "A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 3, pp. 280-291, 2005.
- [13] NINAPRO project web page: <http://www.idiap.ch/project/ninapro/>.
- [14] Ninaweb – the NINAPRO project database web interface: <http://ninapro.hevs.ch/>. Database was downloaded in January 2014.
- [15] M. Atzori, A. Gijsberts, I. Kuzborskij, S. Heynen, A. G. Hager, Deriaz O, Castellini C, Müller H, and Caputo B, "Characterization of a Benchmark Database for Myoelectric Movement Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 1, pp. 73-83, 2015.
- [16] M. Atzori, M. Baechler, and H. Müller, "Recognition of hand movements in a trans-radial amputated subject by sEMG," *Proc of IEEE International Conference on Rehabilitation Robotics (ICOR)*, June 24-26, 2013, Seattle, Washington, USA, pp 1-5.
- [17] G. Bosco, "Principal Component Analysis of Electromyographic Signals: An Overview," *The Open Rehabilitation Journal*, vol. 3, pp. 127-131, 2010.
- [18] N. F. Güler and S. Koçer, "Classification of EMG Signals Using PCA and FFT," *Journal of Medical Systems* vol. 29, no. 3, pp. 241-250, 2005.
- [19] D. Staudenmann, I. Kingma, A. Daffertshofer, D. F. Stegeman, and J. H. van Dieën, "Improving EMG-based muscle force estimation by using a high-density EMG grid and principal component analysis," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 4, pp. 712-719, 2006.
- [20] I.T. Jolliffe, "Principal Component Analysis," New York: Springer, Ed. 2, pp. 1-488, 2002.
- [21] K. Fukunaga, "Introduction to Statistical Pattern Recognition," Waltham, MA: Academic Press, Ed. 2, pp. 1-591, 1990.
- [22] M. Atzori, A. Gijsberts, S. Heynen, A. G. Hager, C. Castellini, B. Caputo, and H. Müller, "Experiences in the Creation of an Electromyography Database to Help Hand Amputated Persons," *Studies in Health Technology and Informatics*, vol. 180, pp. 828-832, 2012.
- [23] L. J. Hargrove, G. Li, K. B. Engelhart, and B. S. Hudgins, "Principal Component Analysis Preprocessing for Improved Classification Accuracies in Pattern-Recognition-Based Myoelectric Control," in *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1407-1414, 2009.
- [24] M. Isaković, "Muscle Onset Detection: Methods for Online and Offline Applications," Poster presentation at the *First Conference on Brain-Computer Interface from Student-to-Student Interface (BCI from SSI)*, Academic Mind, University of Belgrade - School of Electrical Engineering, March 14, 2014.
- [25] M. Atzori and H. Müller, "Electromyography Low Pass Filtering Effects on the Classification of Hand Movements in Amputated Subjects," in *SCIEI International Conference on Digital Signal Processing (ICDSP)*, Milano, Italy, November 19-21, 2014.
- [26] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placements procedures," *Journal of Electromyography and Kinesiology*, vol. 10, no. 5, pp. 361-374, Oct. 2000.
- [27] M. S. Isaković, N. Miljković, M. Štrbac, and M. B. Popović, "Classification of sEMG-based hand movements is affected by electrode array positioning," in *Second Conference on Human-Machine Interface from Student-to-Student Interface (HMI from SSI)*, Academic Mind, University of Belgrade – School of Electrical Engineering, March 20, 2015.
- [28] X. Tang, Z. Liu, C. Lv, and D. Sun, "Hand Motion Classification Using a Multi-Channel Surface Electromyography Sensor," *Sensors*, vol. 12, no. 2, pp. 1130-1147, Feb. 2012.

