Sign Language Recognition Based on EMG Signals through a Hibrid Intelligent System

Bernabé Rodríguez-Tapia[1,2], Alberto Ochoa-Zezzatti[2], Angel Israel Soto Marrufo[2],
Norma Candolfi Arballo[1], Patricia Avitia Carlos[1]

[1] Universidad Autónoma de Baja California, Escuela de Ciencias de la Ingeniería y Tecnología,
Baja California, Mexico

[2] Universidad Autónoma de Ciudad Juárez,
Departamento de Ingeniería Industrial y Manufactura, Chihuahua, Mexico

`rodriguez.bernabe@uabc.edu.mx, alberto.ochoa@uacj.mx,`
`angel.soto@uacj.mx,ncandolfi@uabc.edu.mx and`
`patricia_avitia@uabc.edu.mx`

Abstract. Non-verbal communication is an important part of everyday interactions and human-computer interaction. Vision techniques and instrumented gloves for sign language recognition are commonly used, but these are often expensive and considered invasive to the user. This research proposes the recognition of words from the American Sign Language (ASL) using the SCEPTRE database acquired by two Myoelectrical bracelets. Computational intelligence techniques were used to optimize the number of attributes using Principal Component Analysis (PCA) and a classifier based on Neural Networks (NN). The results suggest that it is possible to reduce the attributes using PCA without significantly losing the quality in classification. This allows faster processing, a convenient feature for classifiers for real-time SL recognition.

Keywords: sign language recognition, ASL, myoelectric signals, EMG.

1   Introduction

Nonverbal communication is an important part of everyday interactions, Sign Language (SL) is the native language of hearing impaired people and consists of a set of specific gestures [1]. Sign language recognition (SLR) and gesture-based control are two applications for hand gesture technologies[2], in this regard most of the present work on sign language recognition focuses on two methods: the use of a single-lens reflex camera that interprets signs through computer vision and on the other hand the use of glove-based gesture recognition [3].

The vision-based single-lens reflex camera may perform poorly under low-light conditions and captured videos/images may be considered invasive to user's privacy, as well as the detection glove is often expensive [3].

Many studies have suggested the use of surface electromyography (sEMG) signals as a method of interacting with machines. In an EMG-based interaction system, hand

gestures are captured by sEMG sensors measuring the activities of the muscular system [4].

The human-computer communication usually occurs by text or voice; trained interfaces are necessary to recognize gestures on the traditional elements of the user interface.

The aim of the present research is to decode sign language taken by two Myoelectric bracelets using the SCEPTRE database [5, 6] using computational intelligence techniques to optimize the amount of pattern through Principal Component Analysis (PCA) and a classifier based on neural networks (NN).

2   Related Work

According to the sensing technologies used to capture gestures, conventional researches on hand gesture recognition can be categorized into two main groups: data glove-based and computer vision-based techniques [7].

The work in [8] reported a system using two data gloves and three position trackers as input devices and a fuzzy decision tree as a classifier to recognize Chinese Sign Language (CSL) gestures. The average classification rate of 91.6% was achieved over a very impressive 5113-sign vocabulary in CSL. However, glove-based gesture recognition requires the user to wear a cumbersome data glove to capture hand and finger movement

The work [3] developed an impressive real-time system recognizing sentence-level American Sign Language generated by 40 words using HMMs. From a desk-mounted camera, word accuracies achieved 91.9% with a strong grammar and 74.5% without grammar, respectively. The work in [9]employed a spatiotemporal feature extraction

Unlike the approaches mentioned earlier, the accelerometer (ACC) and electromyography (EMG) sensor provide two potential technologies for gesture sensing. Previous studies indicated that the combined sensing approach could improve the performance of hand gesture recognition significantly [10]. [11] have compared the performance of ACC-based and EMG-based techniques in the detection of functional motor activities for rehabilitation and provided evidence that the system based on the combination of EMG and ACC signals can be built successfully.

The work [12] has done a study on combining accelerometer and EMG sensors to recognize sub-word level gestures for Chinese Sign Language and [4] show that combination of multiple sensors helps to increase the recognition accuracy.

Following this path, we add the EMG measurements from eight built-in pods in the Myo device to get information that can be leveraged to detect subtle finger movements and configurations, which are essential for detecting certain signs and distinguishing them from others.

3   Methodology

The global gesture recognition methodology consists of the following steps as shown in Fig. 1.
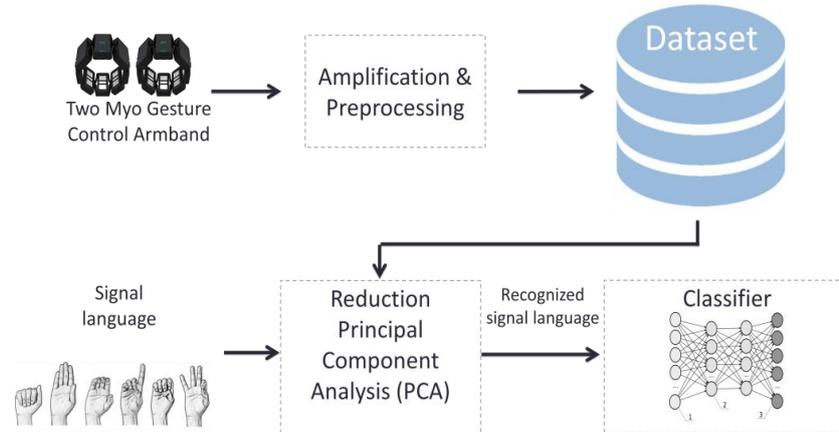
Fig. 1. Classification process of Signal Language recognition with Myo Gesture Control Armband. Adapted from Raez et al /Biological Procedures (2006) 11-35 [13].

Amplification and Preprocessing

The data set represented in Fig. 1, were taken from the database developed by [5], these data were obtained by using two MYO bracelets for some signs of American Sign Language (ASL). The data set represents a subsampling data from a Myo placed on the left hand and the other one on the right hand of each user. The data reading is made by 8 EMG sensors per bracelet at a frequency of 200Hz; it also has a nine axis Inertial Measurement Unit (IMU) including three axis gyroscop (gyp), three axis accelerometer (accl) and three axis magnetometer (o). The Myo bracelet can extract IMU data at a sampling frequency of 50Hz. The range of potentials provided by the Myo bracelet is amplified between -128 and 128 trigger units.

Fig. 2 represents the implementation of a user with wrist devices in each hand making a gesture.
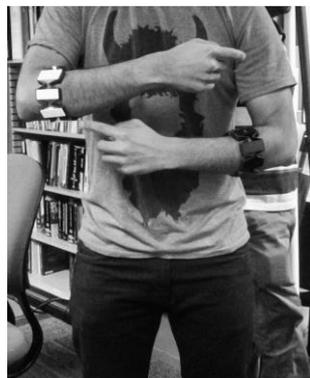
Fig. 2. Deployment: A user with a wrist-band devices on each hand performing a gesture[5].

Data Collection

The database consists of gestures taken from 10 healthy adults between the ages of 22 and 35, each of whom performs a total of 20 ASL gestures. For this research the information of 9 words made by 3 different users was selected, each user repeated the word 5 times. The data set consists of 34 attributes and 1 class. The attributes in this data set are: 16 EMG vectors, 6 accelerometer vectors, 6 gyroscope vectors and 6 position vectors. The number of classes is equivalent to the recognition of the 9 ASL words: (1) BLUE, (2) CAT, (3) COST, (4) DOLLAR, (5) ORANGE, (6) BIRD, (7) SHIRT, (8) LARGE and (9) PLEASE. Fig. 3 presents the description of the attributes and an example of the analyzed data vector.

Fig. 3. Description of attributes and data vector.

Principal Component Analysis

The PCA [14] method, a common dimensionality reduction technique, aims to find a set of orthonormal vectors in the amount of data, which can maximize the variance of the data and map the data into a lower sub-dimensional space encompassed by those vectors [15]. It is widely used to retrieve important information from a multivariable table data and to retrieve that information as a set of few new variables called main components. These new variables correspond to a linear combination of the original ones. The number of main components is less than or equal to the number of original variables.

Weka@ allows to apply a wide range of filters on the data, allowing to make transformations on them of all kinds. It carries out an analysis of the main components and transformation of the data. Dimensionality reduction is achieved by choosing enough vectors of its own to reflect some percentage of the variance in the original data, by default 0.95 (95%). Based on the code from the attribute selection scheme "Principal Components" by Mark Hall and Gabi Schmidberger [16].

Classification

A multilayer perceptron (MLP) is often used as a classifier for it is the most popular type of neural network architecture. A multi-layer perceptron has any number of inputs, has one or more hidden layers with any number of units, uses generally sigmoid activation functions in the hidden layers, has connections between the input layer and the hidden layer, between the hidden layers and between the last hidden layer and the output layer[17].

The advantage of a neural network is its ability to represent both linear and non-linear relationships, and learn these relationships directly from data being modeled. It also meets real time constraints, which are an important feature in control systems [18]

The multilayer perceptron is a classifier that uses backpropagation to learn a multi-layer perceptron to classify instances.

The network can be built by hand in WEKA@ or be configured using simple heuristics. The network parameters can be monitored and modified during the training period. The nodes in this network are all sigmoid (except when the numerical class, in which case the output nodes become linear units with no threshold) [16].
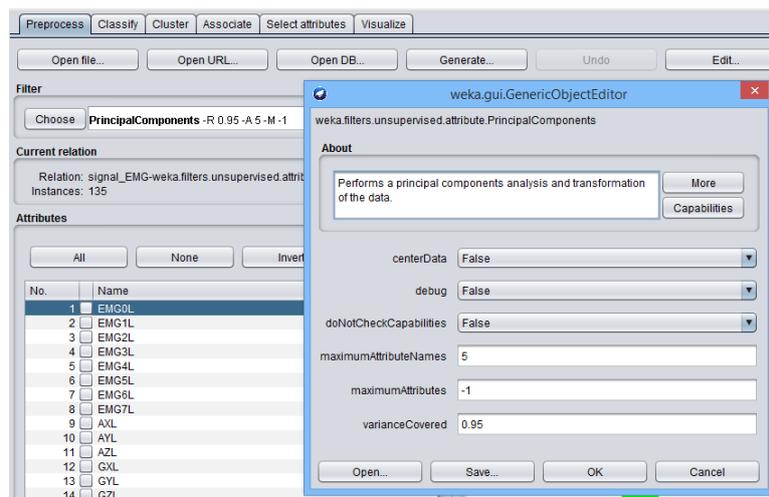
Fig. 4. Configuración WEKA@. Multilayer perceptron.

Ranking and Optimization

The classification and optimization were carried out in the WEKA@ software. A data-based of 3 users was used with 5 repetitions each for the 9 selected words.

Neural network training without optimization was performed with the 34 selected sensor attributes and using the multilayer perceptron. The training and testing was performed with the fusion "Use training set"; with this option WEKA@ will train with all available data and then will apply it again on them.

For the optimization, we used the filter "principal components" within the "unsupervised" attributes of WEKA @. The configuration of the PCA filter and the MLP neural network are described in Fig. 4 and Fig. 5.
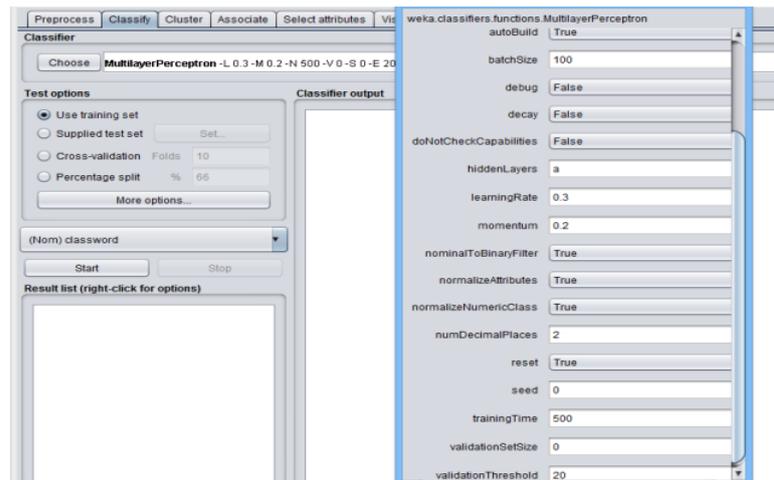
Fig. 5. Configuración WEKA@. Principal Components.

4   Evaluation and Results

The experiments were carried out with a 9-word database with 7 sensor sets (attributes): (1) emg+accl+gyp+o (34 attributes) , (2) emg+accl+gyp (29 attributes), (3) emg+accl+o (29 attributes), (4) emg+accl (23 attributes), (5) emg+gyp (23 attributes), (6) emg+o (23 attributes); (7) emg (16 attributes). The results statistics for the 3 subjects are shown in Table 1 and the confusion matrix for the best 4 classification results using only the MLP neural network is shown in Fig. 6.

Table 1. Classification and optimization results

Atributes	Multilayer Perceptron	Multilayer Perceptron with PCA
(1) All (emg, accl, gyp, o)	100%	97.77%
(2) emg+accl+gyp	99.25%	94.81%
(3) emg+accl+o	100%	99.25%
(4) emg+accl	95.55%	90.62%
(5) emg+gyp	77%	74.07%
(6) emg+o	93.44%	93.33%
(7) emg	47.4%	47.4%

```
=== Confusion Matrix ===

 a  b  c  d  e  f  g  h  i   <-- classified as
15  0  0  0  0  0  0  0  0 |  a = BLUE
 0 15  0  0  0  0  0  0  0 |  b = CAT
 0  0 15  0  0  0  0  0  0 |  c = COST
 0  0  0 15  0  0  0  0  0 |  d = DOLLAR
 0  0  0  0 15  0  0  0  0 |  e = ORANGE
 0  0  0  0  0 15  0  0  0 |  f = BIRD
 0  0  0  0  0  0 15  0  0 |  g = SHIRT
 0  0  0  0  0  0  0 15  0 |  h = LARGE
 0  0  0  0  0  0  0  0 15 |  i = PLEASE
```

(a) All sensors and emg+accl+o

```
=== Confusion Matrix ===

 a  b  c  d  e  f  g  h  i   <-- classified as
15  0  0  0  0  0  0  0  0 |  a = BLUE
 0 15  0  0  0  0  0  0  0 |  b = CAT
 0  0 15  0  0  0  0  0  0 |  c = COST
 0  0  0 14  0  0  0  1  0 |  d = DOLLAR
 0  0  0  0 15  0  0  0  0 |  e = ORANGE
 0  0  0  0  0 15  0  0  0 |  f = BIRD
 0  0  0  0  0  0 15  0  0 |  g = SHIRT
 0  0  0  0  0  0  0 15  0 |  h = LARGE
 0  0  0  0  0  0  0  0 15 |  i = PLEASE
```

(b) emg+accl+gyp

```
=== Confusion Matrix ===

 a  b  c  d  e  f  g  h  i   <-- classified as
15  0  0  0  0  0  0  0  0 |  a = BLUE
 0 15  0  0  0  0  0  0  0 |  b = CAT
 0  0 14  0  1  0  0  0  0 |  c = COST
 0  0  0 14  0  0  0  1  0 |  d = DOLLAR
 0  0  0  0 15  0  0  0  0 |  e = ORANGE
 0  0  0  0  0 15  0  0  0 |  f = BIRD
 0  0  0  0  0  0 14  1  0 |  g = SHIRT
 0  0  0  0  0  0  0 14  1 |  h = LARGE
 1  0  0  1  0  0  0  0 13 |  i = PLEASE
```

(c) emg+accl

Fig. 6. Confusion matrix for better results.

Multilayer Perceptron

◆ No. Atributes   ■ Clasifier

100%
34
99.25%
29
100%
29
95.55%
23
77%
23
93.33%
23
47.4%
16

All (emg, accl, gyp, o)   emg+accl+gyp   emg+accl+o   emg+accl   emg+gyp   emg+o   EMG
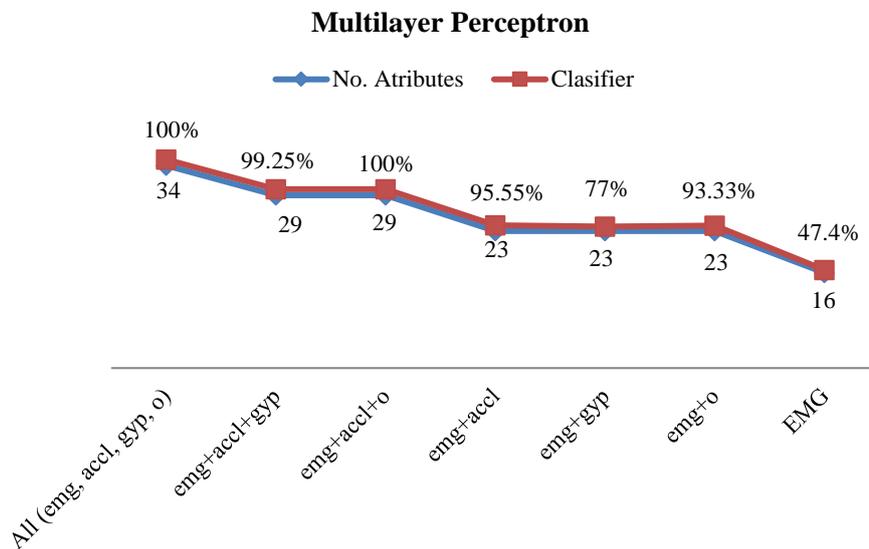
Fig. 7. % classification by no. of attributes.

In addition to using the 34 sensors of both myoelectric bracelets for MLP training and testing, the results indicate that the 2, 3 and 4 sensor sets retain sufficient characteristics to obtain a good rating. In particular, set 2 (emg+accl+o) achieves the same classification as when using all sensors.

The emg+accl data set reflects the best data set, as it decreased from 34 attributes to only 23 without the need for an optimization stage.

The Fig. 7 shows the quality of classification with respect to the number of attributes.

Applying the PCA filter to the total number of attributes showed a significant decrease of these attributes, the results for all data sets are shown in Fig. 8. The most effective option can be identified when working with the emg+accl+o sensors, since when PCA was applied it was possible to reduce from 34 to 20 attributes which achieved a classification of 99.25%, followed by this set, with only the emg+o sensors there is a total of 18 characteristics, almost half of the initials and there is a classification of up to 93.33%.

Multilayer Perceptron with PCA
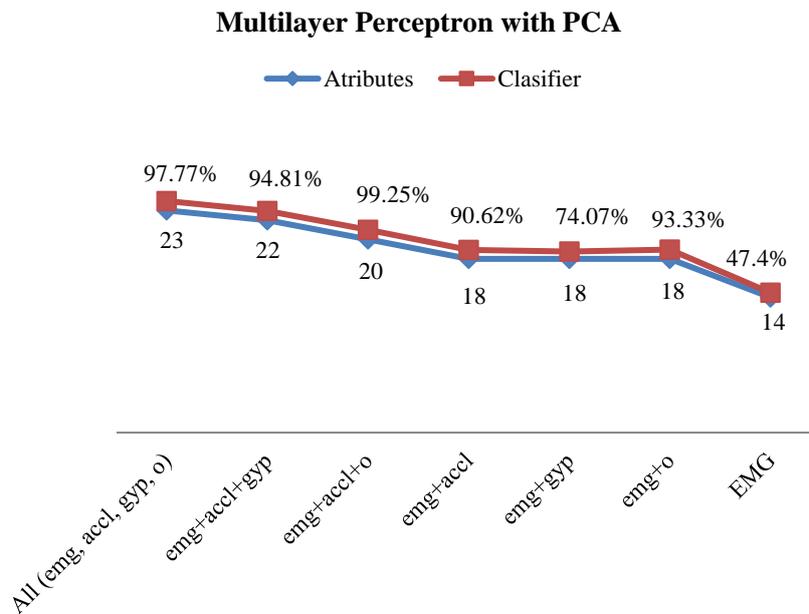
Fig. 8. % classification by no. of attributes with PCA.

5   Conclusion and Future Work

The aim of the present research was to decode the sign language taken by two Myoelectric bracelets using the SCEPTRE database. Using computational intelligence techniques, 34 to 20 attributes were optimized through the Principal Component Analysis (PCA) filter, having a 99.25% classification with a "multilayer perceptron" neural network.

For the further investigation, it is expected to use the PCA optimization method and the MLP classifier in an HMI system with the characteristics of real time operation by means of the bluetooth protocol, for the recognition of SL gestures on the traditional elements of the user interface

References

1. Kosmidou, V.E., Hadjileontiadis, L.J., Panas, S.M.: Evaluation of surface EMG features for the recognition of American Sign Language gestures. En: 2006 International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 6197-6200. IEEE, New York, NY (2006)

2. Xu Zhang, Xiang Chen, Yun Li, Lantz, V., Kongqiao Wang, Jihai Yang: A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans. 41, 1064-1076 (2011). doi:10.1109/TSMCA.2011.2116004

3. Wu, J., Tian, Z., Sun, L., Estevez, L., Jafari, R.: Real-time American Sign Language Recognition using wrist-worn motion and surface EMG sensors. En: 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN). pp. 1-6. IEEE, Cambridge, MA, USA (2015)

4. Chen, X., Zhang, X., Zhao, Z.-Y., Yang, J.-H., Lantz, V., Wang, K.-Q.: Hand Gesture Recognition Research Based on Surface EMG Sensors and 2D-accelerometers. En: 2007 11th IEEE International Symposium on Wearable Computers. pp. 1-4. IEEE, Boston, MA, USA (2007)

5. Paudyal, P., Banerjee, A., Gupta, S.K.S.: SCEPTRE: A Pervasive, Non-Invasive, and Programmable Gesture Recognition Technology. En: Proceedings of the 21st International Conference on Intelligent User Interfaces - IUI '16. pp. 282-293. ACM Press, Sonoma, California, USA (2016)

6. Paudyal, P., Lee, J., Banerjee, A., Gupta, S.K.S.: DyFAV: Dynamic Feature Selection and Voting for Real-time Recognition of Fingerspelled Alphabet using Wearables. En: Proceedings of the 22nd International Conference on Intelligent User Interfaces - IUI '17. pp. 457-467. ACM Press, Limassol, Cyprus (2017)

7. Mitra, S., Acharya, T.: Gesture Recognition: A Survey. IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews). 37, 311-324 (2007). doi:10.1109/TSMCC.2007.893280

8. Fang, G., Gao, W., Zhao, D.: Large Vocabulary Sign Language Recognition Based on Fuzzy Decision Trees. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans. 34, 305-314 (2004). doi:10.1109/TSMCA.2004.824852

9. Shanableh, T., Assaleh, K., Al-Rousan, M.: Spatio-Temporal Feature-Extraction Techniques for Isolated Gesture Recognition in Arabic Sign Language. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics). 37, 641-650 (2007). doi:10.1109/TSMCB.2006.889630

10. Brashear, H., Starner, T., Lukowicz, P., Junker, H.: Using multiple sensors for mobile sign language recognition. En: Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings. pp. 45-52. IEEE, White Plains, NY, USA (2003)

11. Sherrill, D.M., Bonato, P., De Luca, C.J.: A neural network approach to monitor motor activities. En: Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology. pp. 52-53. IEEE, Houston, TX, USA (2002)

12. Li, Y., Chen, X., Tian, J., Zhang, X., Wang, K., Yang, J.: Automatic recognition of sign language subwords based on portable accelerometer and EMG sensors. En: International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction on - ICMI-MLMI '10. p. 1. ACM Press, Beijing, China (2010)

13. Raez, M.B.I., Hussain, M.S., Mohd-Yasin, F., Reaz, M., Hussain, M.S., Mohd-Yasin, F.: Techniques of EMG signal analysis: detection, processing, classification and applications. Biological procedures online. 8, 11-35 (2006). doi:10.1251/bpo115

14. M. Turk, A. Pentland: Eigenfaces for Recognition. Journal of Cognitive Neuroscience. 3, 71-86 (1991). doi:10.1162/jocn.1991.3.1.71

15. D. Huang, W. Hu, S. Chang: Vision-Based Hand Gesture Recognition Using PCA+Gabor Filters and SVM. En: 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. pp. 1-4 (2009)

16. Witten, I.H., Frank, E., Trigg, L.E., Hall, M.A., Holmes, G., Cunningham, S.J.: Weka: Practical machine learning tools and techniques with Java implementations. (1999)

17. Bu, N.B.N., Okamoto, M., Tsuji, T.: A Hybrid Motion Classification Approach for EMG-Based Human-Robot Interfaces Using Bayesian and Neural Networks. IEEE Transactions on Robotics. 25, 502-511 (2009). doi:10.1109/TRO.2009.2019782

18. Rechy-Ramirez, E.J., Hu, H.: Bio-signal based control in assistive robots: a survey. (2016). doi:10.1016/j.dcan.2015.02.004