Computer Vision and Neural Machine Interface for Upper Limb Prostheses

Abdelrahman Hamza, Mostafa Abulwafa, Hanna Nabil, Ghadier Helmy

Abdelrahman.Hamza95@eng-st.cu.edu.eg

Systems and Biomedical Engineering

Cairo University, Egypt

***Abstract*—**This work proposes a hybrid solution that fuses different signals feed from computer vision with neural machine interface in addition to intelligent sensors to decode the user's intention. This method can achieve intelligent control of the prosthetic hand to enable transradial amputees to use a simple, yet efficient, system to grasp and move common household objects. For the Myoelectric signal (MES), we classify among four classes of movements by a method which does not require segmentation of the data, allowing a continuous stream of class decisions to be delivered. Our main conceptual novelty to make use of the MES and computer vision to classify objects with regards to the grasp pattern without explicitly identifying them and Also give the amputee more control to act naturally. We use the publicly available NINAPro database and the Amsterdam library of object images (ALOI) for empirical evaluation.

I. DEFINED PROBLEM

A transradial amputee is someone who went through a surgical separation of the radius and ulna of the lower arm (between the elbow and the wrist). In developing countries 1 in 10 is in need of an assistive device. For Egypt, there are above half a million (the largest number in the Middle East) according to National Center for Health Statistics. This number is increasing as most frequent causes of upper limb amputation such as trauma, diabetes and cancer are prevalent. There are four main aspects to be considered on attempting to substitute their loss: functionality, usability, appearance and cost. Amputees can choose between active and passive prostheses. Passive prostheses aesthetically replace a limb but cannot move. For active prostheses there are two main categories:1- Body powered prostheses, such as a hook or a prehensor. 2- Motorized prostheses, such as i-limb. They are normally controlled myoelectrically, i.e., using the bioelectrical signals from the residual muscles in the affected limbs. Motorized prostheses are more cosmetically appealing than traditional body powered prostheses, and generally superior to other types in functionality. However, its weakness lies in the fact that it is controlled by myoelectric signals. This signal weakens over time due to muscle atrophy after the amputation which makes it unreliable as a long term solution for transradial amputees. Globally, the current most promising research addresses this issue by targeted muscle reinnervation surgery. This sophisticated surgery does not seem feasible in most of developing countries due to its large expenses.

II. APPROACH

The proposed system should offer intuitive control, has an acceptable accuracy and it should be a long term solution. Our idea to solve the previously mentioned problems is taking multiple signals as inputs and processes them together to determine the final output which is the movement of the prosthetic hand. Recently there has been a preliminary research to add signals such as images of objects acquired using a camera with computer vision to help the patient move the prosthetic hand to the suitable grip. We intend to work on development in this direction using a camera and sensors instead of a surgical operation.

The information flow of the proposed system is shown in Figure 1.

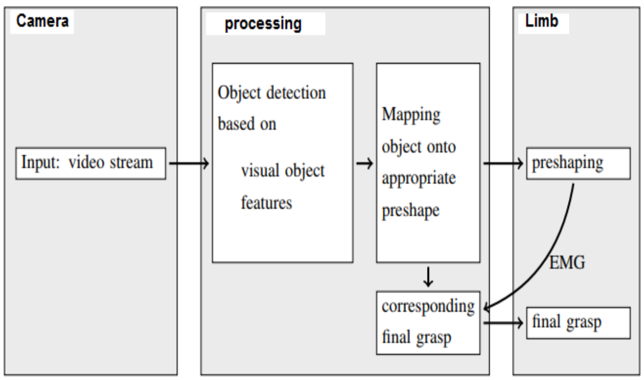

Fig. 1. Information flow of the proposed system

III. EMG ANALYSIS

A. Methodology

The construction of a continuous classifier and the acquisition of data used to evaluate it are described herein. The control problem can be defined for any set of motions. It was decided to investigate a four-class problem involving hand and wrist control, as those with below-elbow limb deficiencies represent a large proportion of prosthetic users. Twelve channels of myoelectric data were acquired using active electrodes. Pattern recognition was performed on analysis windows that may be up to 256 ms in duration. For each analysis window, a feature set was computed, and these features provided to a pattern classifier (Figure2)

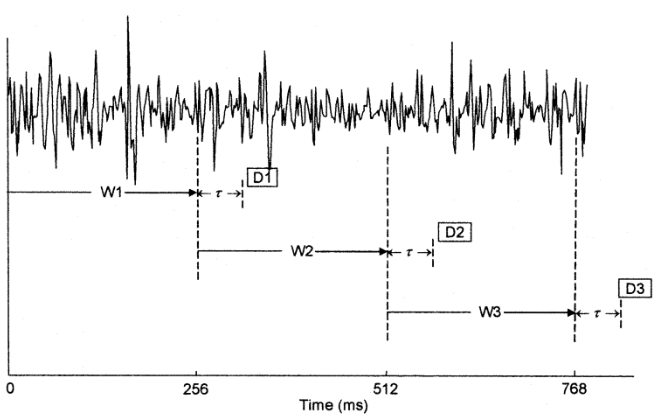

Fig. 2.Windowing of MES data in the continuous classifier. This figure cited from [2].

B. Preprocessing

A high-pass Butterworth filter with corner frequency at 20 Hz is utilized for preprocessing. This configuration is expected to eliminate artifacts due to electrode movement and reduce the crosstalk by attenuating contributions from muscles distant from the electrode. Empirically, we found that signal normalization improved the final accuracy. So, Signal amplitude of each channel is divided by its corresponding mean.

C. Features Extraction

Time-domain feature set showed the best classification accuracy compared to others [3]. Four features are selected:

1-Mean absolute value. 2-Zero crossings: A simple frequency measure can be obtained by counting the number of times the waveform crosses zero. 3-Slope sign change: A feature that may provide another measure of frequency content is the number of times the slope changes sign. 4-Waveform length: A feature that provides information on the waveform complexity in each segment.

Feature set is computed on each of twelve channels, and then concatenated to form a 48-dimensional feature vector. This feature vector was then provided to the classifier.

D. Classification

We used k-nearest neighbors (KNN) which is considered among the oldest non-parametric classification algorithms. In KNN, it is preferable not to normalize the signal in the preprocessing. We tried support vector machine algorithm (SVM) which is achieve accuracy higher than KNN. Linear support vector classification (LinearSVC) is used with kernel ‘linear’ and implemented in terms of ‘liblinear’. Also one versus reset is used as the multi class strategy with C parameter equal one. We found that signal normalization in SVM has a good impact, so we used it.

IV. COMPUTER VISION

A. Methodology

We used the advances of deep learning in computer vision research for control of hand prostheses. This method can identify the appropriate grasp type for objects according to a learned abstract representation of the object rather than the explicitly-measured object dimensions. In this way, objects are not classified based on the object category or identity, but based on the suitable grasp pattern. So this approach is conceptually different from object recognition in which object details matter. To learn this abstract representation, we use convolutional neural network (CNN) architecture.

B. DATASET AND PREPROCESSING

The ALOI dataset includes the images of 1000 common objects. Within this library, 250 objects have been photographed at a second zoom rate. We discarded these 250 objects. For each of the remaining 750 objects, the database includes 72 pictures, taken at 5_ intervals against a black background. We first selected 473 of the objects in four different classes of pinch, tripod, palmar wrist neutral and palmar wrist pronated. Other objects were either not graspable or could be picked with more than one grasp type. Then later due to imbalanced class distribution, we prefer to sample the dataset. Some instances are deleted from the over represented class, under sampling. This leads to use about 356 objects.

In preprocessing, each image I, was first converted to grey scale and was downsampled to an $N = 36$ by $M= 48$ image. Empirically, we found that image normalization, prior to the CNN setting, improved the final accuracy. Therefore, each image was normalized by subtract the mean, and then divide by the standard deviation.

C. CNN ARCHITECTURE

We used pre-trained CNN from [ ]. Figure3 shows the implemented two layers CNN architecture. We used five kernels of size 5×5 and the resultant feature maps were sub-sampled by max-pooling by a factor of two. We applied the maxpooling operation to ensure salient elements in each feature map are retained.

Many activation functions are tested like Scaled exponential linear unit, sigmoid, Tanh, rectified linear unit (ReLU) and leak ReLU. Empirically it is found that the ReLU function results in the highest performance and hence we used it in this study.

Training was carried out through back propagation using the adaptive momentum estimation, Adam algorithm, for optimizing the learned filters within each iteration. Over-fitting is avoided by using Dropout regularization which randomly shuts down some neurons in each iteration. Here, the percentage of dropped neurons is equal to 20%. We built the model by Keras package using TensorFlow backend in python.

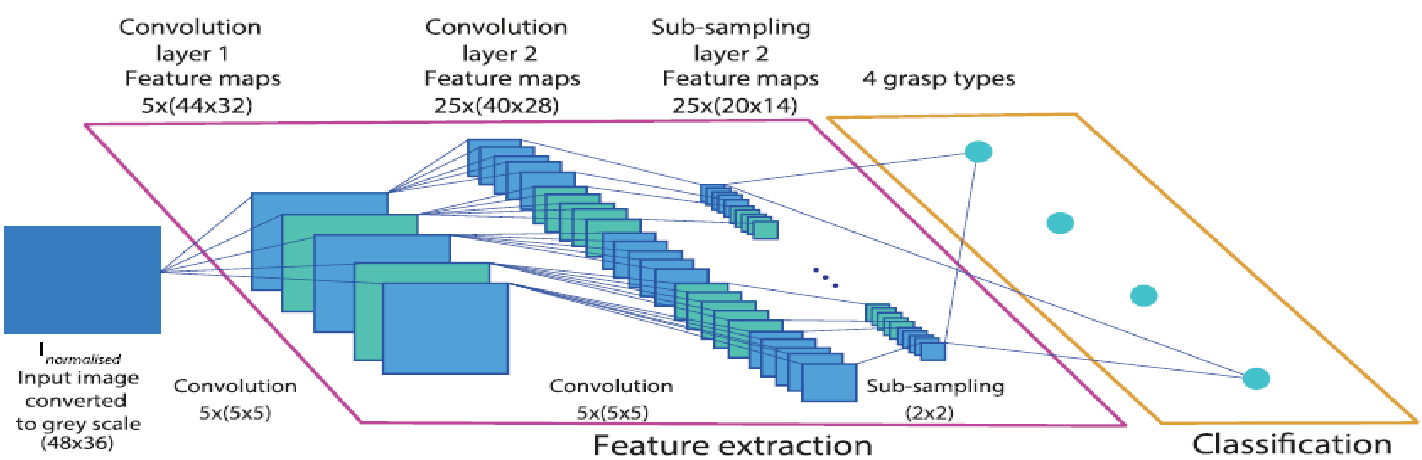

Fig. 3. Two layers CNN architecture. This figure cited from [5]

V. RESULTS

For EMG, after applying this methode we got a reasonable accuracy, around 75 % in most subjects. Figure4 show accuracy in some subjects.

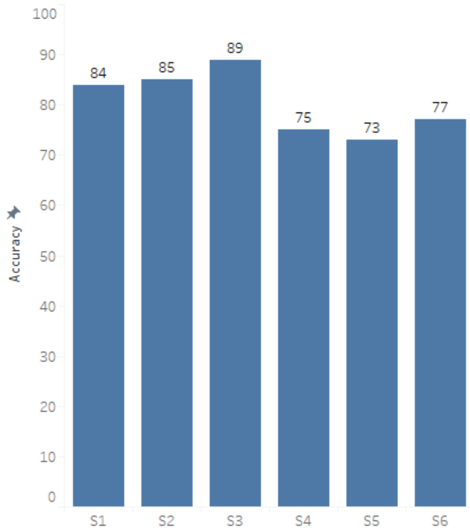

Fig. 4. EMG accuracy in some subjects

Regarding the computer vision, in Within-object cross validation (WOC)., we evaluated the ability of the proposed structure in classifying previously seen objects. The training set included 90% (65 of 72) of the views for each object in each grasp class. The remaining 10% of the views for each object were allocated to the testing set. Here we got accuracy around 90%.

In Between-object cross-validation (BOC). To be able to identify the appropriate grasps for unseen objects, we carried out the BOC test. In the BOC scheme, an object and its views were either wholly seen or unseen. Here we got accuracy around 70%. Figure5 illustrates our result.

VI. CONCLUSION

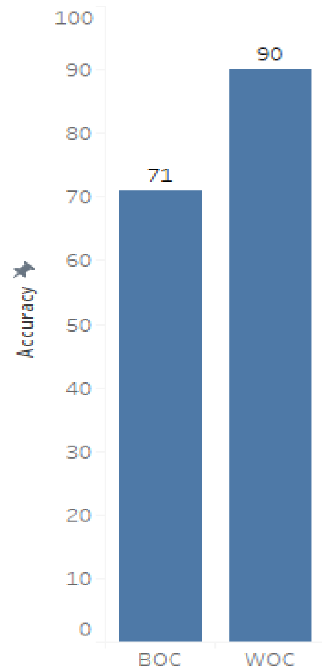

Fig. 5. Computer vision result.

REFERENCES

[1] Manfredo Atzori, Arjan Gijsberts, Claudio Castellini, Barbara Caputo, Anne-Gabrielle Mittaz Hager, Simone Elsig, Giorgio Giatsidis, Franco Bassetto & Henning Müller. “Electromyography data for non-invasive naturally-controlled robotic hand prostheses”

[2] K. Englehart and B. Hudgins, “A robust, real-time control scheme for multifunction myoelectric control,”

[3] Sungtae Shin, Reza Langari and Reza Tafreshi, “ a performance comparison of EMG classification methods for hand and finger motion’’

[4] Adenike A. Adewuyi, Levi J. Hargrove and Todd A. Kuiken “Evaluation EMG features and classifer selection for application to partial hand prothesis control ’’

[5] Ghazal Ghazaei, Ali Alameer, Patrick Degenaar, Graham Morgan and Kianoush Nazarpour “Deep learning-based artificial vision for grasp classification in myoelectric hands’’

[6] Meena AbdelMaseeh, Tsu-Wei Chen, and Daniel Stashuk “Extraction and Classification of Multichannel Electromyographic Activation Trajectories for Hand Movement Recognition’’

[7] Jan Mark Geuserbroek, Gertjan J.Burghouts and Arnold W.M. Smeulders “The Amsterdam Library for Object Images”

[8] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever and Ruslan Salakhutdinov “a simple way to prevent neural network from overfitting”