IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 52-56

Hand Gesture Recognition

Devyani Sengupta[1], Prof Sandhya Dahake[2]
[1]G.H. Raisoni Institute of Information Technology, MCA Sem-5, Nagpur, India
2. Assistant Professor, Department of computer science, GHRIIT, Nagpur, Maharashtra

Abstract: *In recent years, deep learning algorithms have become increasingly more prominent for their unparalleled ability to automatically learn discriminant features from large amount of data.*
Gesture recognition is a hot topic in computer vision and pattern recognition, which plays a vitally important role in natural human-computer interface. Although great progress has been made recently, fast and robust hand gesture recognition remains an open problem. Since the existing methods have not well balanced the performance and the efficiency simultaneously.
Keywords: *Hand gesture recognition, fast, robust*

I. Introduction

Robotics and Artificial Intelligence can be leveraged to increase the autonomy of people living with disabilities. This is accomplished in parts, by enabling users to seamlessly interact with robots to complete their daily tasks with increased independence.

Hand gesture is one of the most expressive, natural and common type of body language for conveying attitudes and emotions in human interactions. In the context of hand prosthetic control, muscle activity provides an intuitive interface on which to perform hand gesture recognition.This activity can be recorded by surface electromyography (sEMG), a non-invasive technique.widely adopted both in research and clinical settings. The sEMG signals, which are non-stationary, represent the sum of subcutaneous motor action potentials generated through muscular contraction. Artificial intelligence can then be leveraged as the bridge between sEMG signals and the prosthetic behaviour. A research paper by Hao Tang, was devoted to bridge the gap between fast and robust hand gesture recognition, simply using solo popular cue e.g., RGB, which ensures great potential in practical use. Key frames, also known as representative frames, extract the main content of a data series, which could greatly reduce the amount of processing data.The key frames of the video sequence are selected by their discriminative power and represented by the local motion features detected in them and integrated from their temporal neighbours. Carlsson and Sullivan demonstrate that specicactions can be recognized in long video sequence by matching shape information extracted from individual frames to stored prototypes representing key frames of the action.

II. Methodology

I. SEMG DATASETS

Myo Dataset

One of the major contributions of this article is to provide a new, publicly available, sEMG-based hand gesture recognitiondataset, referred to as the Myo Dataset. This dataset contains two distinct sub-datasets with the first one serving as the pertaining dataset and the second as the evaluation dataset. The former, which is comprised of 19 able-bodied participants,should be employed to build, validate and optimize classification techniques. The latter, comprised of 17 able-bodied participants, is utilized only for the final testing. To the best of our knowledge, this is the largest dataset published utilizing the commercially available Myo Armband (Thalamic Labs) and it is our hope that it will become a useful tool for the sEMG based hand gesture classification community. The data acquisition protocol was approved by the Comit´es d’ ´ Ethique de la Recherche avec des ˆetres humains de l’Universit´e Laval (approbation number: 2017-026/21-02- 2016) and informed consent was obtained from all participants.

1) sEMG Recording Hardware: The electromyographic activity of each subject’s forearm was recorded with the Myo Armband; an 8-channel, dry-electrode, low-sampling rate (200Hz), low-cost consumer-grade sEMG armband. The Myo is non-intrusive, as the dry-electrodes allow users to simply slip the bracelet on without any preparation. Comparatively, gel-based electrodes require the shaving and washing of the skin to obtain optimal contact between the subject’s skin and electrodes. Unfortunately, the convenienceof the Myo Armband comes with limitations regarding thequality and quantity of the sEMG signals that are collected.
Indeed, dry electrodes, such as the ones employed in the Myo, are less accurate and robust to motion artifact than gel-based ones. Additionally, while the recommended frequency range of sEMG signals is 5-500Hz

requiring a sampling frequency greater or equal to 1000Hz, the Myo Armband is limited to 200Hz. This information loss was shown to significantly impact the ability of various classifiers to differentiate between hand gestures. As such, robust and adequate classification techniques are needed to process the collected signals accurately.

2) Time-Window Length: For real-time control in a closed loop, input latency is an important factor to consider. A maximum latency of 300ms was first recommended in a paper. Even though more recent studies suggest that the latency should optimally be kept between 100-250ms,the performance of the classifier should take priority over speed. As is the case in one of the paper, a window size of 260ms was selected to achieve a reasonable number of samples between each prediction due to the low frequency of the Myo.

3) Labelled Data Acquisition Protocol: The seven hand/wrist gestures considered in this work are depicted in Fig. 1. For both sub-datasets, the labelled data was created by requiring the user to hold each gesture for five seconds. The data recording was manually started by a researcher only once the participant correctly held the requested gesture. Generally, five seconds was given to the user between each gesture. This rest period was not recorded and as a result, the final dataset is balanced for all classes. The recording of the full seven gestures for five seconds is referred to as a cycle, with four cycles forming a round. In the case of the pre-training dataset, a single round is available per subject. For the evaluation dataset three rounds are available with the first round utilized for training (i.e. 140s per participant)

Figure 1: The 7 hand/wrist gestures considered in the Myo Dataset.

During recording, participants were instructed to stand up and have their forearm parallel to the floor and supported by themselves. For each of them, the armband was systematically tightened to its maximum and slid up the user's forearm, until the circumference of the armband matched that of the forearm. This was done in an effort to reduce bias from the researchers, and to emulate the wide variety of armband positions that end-users. Without prior knowledge of optimal electrode placement might use (see Fig. 2). While the electrode placement was not controlled for, the orientation of the armband was always suchthat the blue light bar on the Myo was facing towards the hand of the subject. Note that this is the case for both left and right handed subjects. The raw sEMG data of the Myo is what is made available with this dataset.


Figure 2: Examples of the range of armband placements on the subject's forearm

III. Key Frames Extraction

Key frames extraction is the key technology for video abstraction, which can remove the redundant information in the video greatly. The algorithm for key frames extraction will act the reconstruction of video content. If a frame in video V can be represented by fi, where i is (1; 2; :::; n) and n is the total number of frames in video V . Hence, the key frames set SKeyframes is defined as follows:

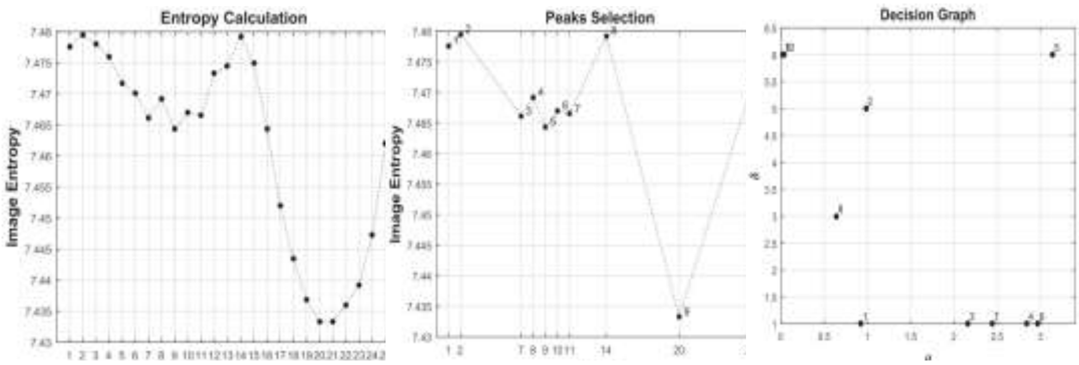
SKeyframes = fKeyframes (V); (1)

Where fKeyframes denotes the key frames extraction procedure. In this paper, a method of key frames extraction based on image entropy and density clustering is proposed, as we can see from Figure 1. Our key frames extraction methods are mainly divided into three steps, namely, 1) calculating image entropy, 2) _nding local extreme points and 3) executing density cluster. The following section would expand upon on it.

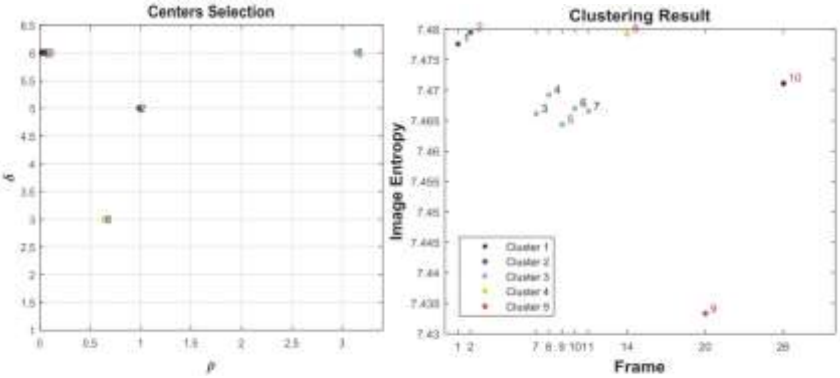
IV. Image Entropy

In this section, we try to find a proper descriptive index to evaluate each frame in a video, facilitating key frame extraction. Informative frames could better summarize the whole video where they reside, while how to quantify the information each frame contains is a hard-nut to crack. Firstly, we calculate image entropy of each frame, and then map them into a two-dimensional coordinate space, as shown in Figure 1(b). Entropy is a nice way of representing the impurity or unpredictability of a set of data since it is dependent on the context in which the measurement is taken. As for a single video frame, the grey-scale colour/intensity distribution of this frame can be seen as

(a)A hand gesture sequence sample from the Northwestern University hand gesture dataset, which contains 26 frames. The key frames obtained by our method are in green boxes, which are the 2, 9, 14, 20 and 26 frames.



(b) Calculate the image entropy. **(c)** Select the local peak points. **(d)** Calculate ρ and δ.



(e) Select the number of clustering. In this case, we choose 5 clusters.

(f) the final results of clustering. The points of 2, 5, 8, 9 and 10 are the clustering centres, therefore, the corresponding frames (2, 9, 14, 20 and 26) are the key frames of original sequence.

(g) The key frames are the 2, 9, 14, 20 and 26 frames. Now we use this sequence to replace the original sequences for the next step.

Figure 3: The framework of the proposed key frames extraction method.

p = {p1; p2;:::; pn}. For the image frames fi, their image entropy can be defined as:

E (fi) = - Xj pfi (j) logpfi (j);

Where pfi (j) denotes the probability density function of frame fi, which could be obtained by normalizing their histogram of grey-scale pixel intensities. Next we map the value E (fi) to a two-dimensional coordinate space (the E(fi) vs. i plot).

V. Performance Evaluation

In order to test the performance of the system, on the basis of how interactive the system is, recognizing the sign language, portability, efficiency and amount of power consumption. Both the systems are made to enhance the ability of using the artificial automations.

Features	Electromyographic Hand Gesture	Fast and Robust Dynamic Hand Gesture
Efficiency	Not much efficient as a person with hand disability will find difficult to use	A minute hand gesture could also be used and recognised
Power consumption	Battery is used	Needs to be charged
Interactive	Studies the contraction and relaxation of the muscles	Studies the images in the videos
Recognizing sign language	Has this ability but needs specific hand gestures	Does not need any specific hand gesture
Portability	Easily portable	Depend on the system it is mounted on

Table 1: Summary of performance evaluation

VI. Conclusion

The devices were analyzed and tested. This paper is been made with the help of research paper presented by Ulysse Cˆot´e-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Cl´ement Gosselin, Kyrre Glette, Franc¸ois Laviolettey, and Benoit Gosselinyabout Electromyographic Hand Gesture, and Hao Tang, Hong Liu, Wei Xiao, Nicu Sebe about Dynamic Hand Gesture Recognition.

References

[1]. Deep Learning for Electromyographic Hand GestureSignal Classification Using Transfer Learning Ulysse Cˆot´e-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Cl´ement Gosselin, Kyrre Glette, Franc¸ois Laviolettey, and Benoit Gosseliny

[2]. Fast and Robust Dynamic Hand Gesture Recognition via Key Frames Extraction and Feature Fusion Hao Tang, Hong Liu, Wei Xiao, Nicu Sebe