IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Survey on Gesture Identification Technique

Bincy Annamma Saji pursuing B.Tech. degree in Computer Science and Engineering from Mount Zion Institute of Science and Technology, Kerala Technological University,India.

N C Chanjal pursuing B.Tech. degree in Computer Science and Engineering from Mount Zion Institute of Science and Technology, Kerala Technological University, India.

Ruhin Mary Saji, Assistant Professor at Mount Zion Institute of Science and Technology, received B.Tech. degree in Computer Science and Engineering from Sree Buddha College Of Engineering Pattoor, Kerala university and M.Tech. degree from Rajiv Gandhi Institute of Technology,Kottayam, Kerala Technological University, India.

Abstract

Hand gesture is one of the most helpful technique for deaf and dumb people.Gesture recognition is done using IMU sensors which plays an important role in various IOTs.We can train the gestures using offline phase and can recognize these gestures during the offline phase.Without proper training,we cannot recognize these gestures.For recognizing gestures,we use Bluetooth enabled hand motion trajectory mounted on the wrist for identification.But this technique may become difficult if it willnot perform perfectly or may be damage.We use machine learning approach in which Bluetooth enabled IMU are not required.We use SSD algorithm and Inception Model.We use hands for gesture recognition and these are taken as image through camera and loaded to the tensorflow memory.Then it is compared with the trained dataset.If it matches,output is displayed on the screen.

Keywords: dynamic time warping (DTW); hand gesture recognition (HGR); inertial measurement unit (IMU); machine learning; real-time learning; restricted coulomb energy (RCE) neural network.

Introduction

Gesture is defined as the part of the body usually the hands to express some idea or meaning.Here, gestures are identified using machine learning technique. **Machine learning** (**ML**) is defined as the study of computer algorithms that improve automatically through experience. It is a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as training data to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as computer vision, where it is infeasible to develop conventional algorithms to perform the needed tasks. Machine learning is related to computational statistics that focuses on making predictions using computers. The study of mathematical optimization deliver methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on data analysis through unsupervised learning.In its application across business problems, machine learning is also referred to as predictive analytics.Dataset consists of training examples,validation examples and testing examples.Training set consists of set of input examples that the model has to be trained on by adjusting the parameters.Dataset needs to be periodically evaluated in order to train the model which is referred to as the validation set.The final evaluation that the model goes through after the training phase and the validation phase is called as a test set.

The remaining part of the paper is organized as follows. In Section II survey of all methods will be described in detail. The paper concludes with a brief summary in section III.

Literature survey

Surface electromyography (sEMG) signals have been increasingly used in pattern recognition and rehabilitation in recent years. In paper[1], a real-time hand gesture recognition model using sEMG is proposed.An armband is required which acquire the sEMG signals and apply a sliding window approach to segment the data in order to extract the features. A feedforward artificial neural network (ANN) is found and trained by the training dataset.Tthe gesture is recognized using test method and when recognized label times reach the threshold of activation times by the ANN classifier. In the experiment, they collected the real sEMG data from twelve subjects and use a set of five gestures from each subject to evaluate our model, with an average recognition rate of 98.7% and average response time of 227.76 ms, which is one-third of the gesture time.The pattern recognition system might be able to recognize a gesture before the gesture is completed.

In paper[2], Mohammed and et.al proposed a deep learning-based architecture to jointly detect and classify hand gestures. In the proposed architecture, the whole image is passed through a one-stage dense object detector to extract hand regions, which passes through a lightweight convolutional neural network (CNN) for hand gesture recognition. To evaluate the approach, they have conducted extensive experiments on four publicly available datasets for hand detection, including the Oxford, 5-signers, EgoHands, and Indian classical dance (ICD) datasets, along with the two hand gesture datasets with different gesture vocabularies for hand gesture recognition namely, the LaRED and TinyHands datasets. The experimental results demonstrate that the proposed architecture is efficient and robust. It outperforms other approaches in both the hand detection and gesture classification or identification tasks. Recent research on hand detection and gesture recognition has attracted increasing interest because of its broad range of potential applications, such as human-computer interaction, sign language recognition, hand action analysis, driver hand behavior monitoring and virtual reality. In recent years, several approaches have been proposed with the aim of developing a robust algorithm which functions in complex and cluttered environments. Although several researchers have addressed this challenging problem, a robust system is still elusive

For many applications, hand gesture recognition systems rely on biosignal data are mandatory.These systems have to be affordable, reliable as well as mobile. The hand is moved due to muscle contractions which cause motions of the forearm skin. These motions can be captured with cheap and reliable accelerometers placed around the forearm. Since accelerometers can be integrated into mobile systems easily, the possibility of a robust hand gesture recognition based on accelerometer signals is evaluated in this work. In paper[3], a neural network architecture consisting of two different patterns of recurrent neural network (RNN) cells is proposed. Experiments on three databases revealed that this relatively small network outperforms by far state-of-the-art hand gesture recognition approaches rely on multi-modal data. The combination of accelerometer data and an RNN combines a robust hand gesture classification system, i.e., the performance of the network does not vary with a lot between subjects and it is outstanding for amputees.The proposed network uses only 5 ms short windows to classify the hand gestures. Consequently, this approach allows for a quick and delay-free hand gesture detection.

Surface electromyogram (sEMG) has numerous applications. It has been widely used in various biosignal and neuro rehabilitation applications. There is an urgent need for establishing a simple robust system that can be used to identify subtle complex hand actions and gestures for the control of prosthesis and other computer- assisted devices. Earlier work identified the hand actions and gestures based on sEMG suffers from limitation that these are suitable for gross actions where there is only one prime-mover muscle included and unsuitable for small subtle and complex muscle contraction. The paper[4] presents the hand gesture identification using sEMG decomposed using semi-blind independent component analysis combined with neural network based classifier. The aim is to provide reliable and natural control for rehabilitation and human computer interaction applications. They have proposed a model based approach where the hand muscle anatomy is already known. The system was tested on 5 subjects and with the experiments repeated on different days. The system was compared with raw sEMG which is used by other researchers. The system is able to classify different hand actions by 100%. The classification of the traditional ICA and raw sEMG for the same experiments and similar features was a poor 65% and 60% in comparison respectively. This research demonstrates that sEMG can be decomposed into the individual muscle activities using the semi-blind ICA. The muscle activity after decomposition can be used to identify small and subtle hand actions and gestures accurately. Finally,the ICA source separation was validated with the mixing matrix analysis.

Availability of handy RGB-D sensors has brought about surge of gesture recognition research and applications. Among various approaches, one shot learning approach has become advantageous because it requires minimum amount of data. In paper[5], they provide a thorough review about one-shot learning gesture recognition from RGB-D data and proposed a novel spatiotemporal feature extracted from RGB-D data, namely mixed features around sparse keypoints (MFSK). In the review, they analyzed the challenges that are facing and point out some future research directions which may enlighten researchers in this field. The proposed MFSK characteristic is robust and invariant to scale, rotation and partial occlusions. To alleviate insufficiency of one shot training samples,the training samples are evaluated by artificially synthesizing versions of various temporal scales which is beneficial for coping with gestures performed at varying speed. They evaluate the proposed method on the Chalearn gesture dataset (CGD). The result shows that the approach outperforms all currently published approaches on the challenging data of CGD, such as translated, scaled and occluded subsets. When applied to the RGB-D datasets, that are not one-shot (e.g., the Cornell Activity Dataset-60 and MSR Daily Activity 3D dataset), the proposed feature produces very promising results under leave-one-out cross validation or one-shot learning.

Paper[6] describes a novel method called as Deep Dynamic Neural Networks (DDNN) for multimodal gesture recognition. A semi-supervised hierarchical dynamic framework based on a Hidden Markov Model (HMM) is proposed for simultaneous gesture segmentation and recognition where the skeleton joint information, depth and RGB images are the multimodal input observations. Unlike most traditional approaches that rely on the construction of complex handcrafted features,this approach learns high-level spatio-temporal representations using deep neural networks suited to the input modality: a Gaussian-Bernouilli Deep Belief Network (DBN) to handle the skeletal dynamics, and a 3D Convolutional Neural Network (3DCNN) to manage and fuse batches of depth and RGB images. This is achieved through modeling and learning of the emission probabilities of the HMM required to infer the gesture sequence. This purely data driven approach achieve a Jaccard index score of 0.81 in the ChaLearn LAP gesture spotting challenge. The performance is on par with a variety of state-of-the-art hand-tuned feature-based approaches and other learning-based methods, so opening the door to the use of deep learning techniques in order to further explore multimodal time series data.

Gesture spotting is an essential task for recognizing finger gestures which is used to control in-car touchless interfaces. Automated methods to achieve this task requires to detect video segments where gestures are observed, to discard natural behaviors of users' hands that may look as target gestures, and may be able to work online. In paper[7], the team addressed the challenges with a recurrent neural architecture for online finger gesture spotting. This proposed a multi-stream network merging hand and hand-location features, which help to discriminate target gestures from natural movements of the hand, since these may not happen in the same 3D spatial location. The multi-stream recurrent neural network (RNN) recurrently learns the semantic information, allowing to spot the gestures online in long untrimmed video sequences. In order to validate this method, finger gesture dataset are collected in an in-vehicle scenario of an autonomous car. 226 videos with above 2100 continuous instances were captured with a depth sensor. On this dataset, the gesture spotting approach performs state-of-the-art methods with an improvement of about 10% and 15% of recall and precision respectively. Furthermore, they have demonstrated that by combining the existing gesture classifier (a 3D Convolutional Neural Network), their proposal achieves better performance than previous hand gesture recognition methods.

Gestures play an important role in face-to-face communication and have been increasingly studied functional magnetic resonance imaging. A large amount of data has been provided to describe the neural substrates of gesture comprehension, these findings have never been quantitatively summarized and conclusion is still unclear. In paper[8],activation likelihood estimation meta-analysis investigated that the brain networks underpinning gesture comprehension while considering the impact of gesture type (co-speech gestures vs. speech-independent gestures) and task demand (implicit vs. explicit) on the brain activation of the gesture comprehension. The meta-analysis of 31 papers showed that hand actions, gestures involve a perceptual-motor network important for action recognition. As meaningful symbols, gestures involves a semantic network for conceptual processing. Finally, during the face-to-face interactions, gestures involve a network for social emotive processes. The findings indicates that gesture type and task demand influence the involvement of the brain networks during gesture comprehension. The results highlights the complexity of gesture comprehension and suggest that future research is necessary to clarify the dynamic interactions among these networks.

With the recent growth of Smart TV technology, the demand for unique and the beneficial applications motivates the study of a unique gesture-based system for a smart TV-like environment. Combining the movie recommendation, social media platform, call a friend application, weather updates, chatting app, and tourism platform into a single system regulated by natural-like gesture controller is proposed to allow ease of use and natural interaction. Gesture recognition problem solving was designed through 24 gestures of 13 static and 11 dynamic gestures that suit the environment. Dataset of a sequence of RGB and depth images were collected, preprocessed and then trained in the proposed deep learning architecture. In paper[9],a combination of three-dimensional Convolutional Neural Network (3DCNN) followed by Long Short-Term Memory (LSTM) model was used to extract the spatio-temporal features. At the end of classification, Finite State Machine (FSM) communicates the model to control the class decision results based on the application context. The results shows that the combination data of depth and RGB to hold 97.8% of accuracy rate on eight selected gestures, while the FSM has improved recognition rate from 89% to 91% in a real-time performance.

In paper[10], the work is to present a novel continuous finger gesture recognition system based on flex sensors. The system is able to carry out the accurate recognition of a sequence of gestures. Wireless smart gloves equipped with flex sensors were implemented for the collection of training and testing sets. Given the sensory data that is acquired from the smart gloves, the gated recurrent unit (GRU) algorithm was adopted for gesture spotting. During the training process for GRU, the movements associated with different fingers and transitions between two successive gestures were taken into consideration. On the basis of gesture spotting results, the maximum a posteriori (MAP) estimation was carried out for the final gesture classification. Because of the effectiveness of proposed spotting scheme, accurate gesture recognition was achieved even for the complicated transitions between successive gestures. From the experimental results, it is observed that the proposed system is an effective alternative for robust recognition of a sequence of finger gestures.

Conclusion

Gestures are trained in the dataset.Images are captured through web camera.It can also extract live videos.Gestures and dataset are loaded to tensorflow memory and SSD algorithm is applied.Output will be displayed on the screen if it matches with the trained dataset.It requires 1.5GB memory that consists of 87000 images where the image classification is done.This uses English alphabet as the basis for its gesture catalogue.It can be implemented in cloud server.It can be used in live TV news.It can be used for any communication with the people who needs hearing aid.

Acknowledgement

We are deeply expressing our sincere gratitude to our project guide Professor Ruhin Mary Saji who always supported and guided us with valuable comments. We express our immense pleasure and thankfulness to all faculty members of the Department of Computer Science & Engineering of Mount Zion Institute Of Science And Technology, Kozhuvalloor Chengannur

References

[1] Zhang Z, Yang K, Qian J, Zhang L-“Real-Time Surface EMG Pattern Recognition for Hand Gesture Based on an Artificial Neural Network,July 2019

[2]Mohammad AAQ, LV J Islam MDS-“A Deep Learning-Based End-to-End Composite System for Hand Detection and Gesture Recognition,November 2019

[3]Koch P,Dreier M,Maass M,Bohme M,Phan H,Mertins A-“A Recurrent Neural Network for Hand Gesture Recognition Based on Accelerometer Data,July 2019

[4]Naik G R,Kumar DK,Palaniswami M-“Surface EMG Based Hand Gesture Identification using Semi Blind ICA:Validation of ICA Matrix Analysis,April 2008

[5]Wan J,Guo G,LI SZ-“Explore Efficient Local Features from RGB-D Data for One-Shot Learning Gesture Recognition,August 2016

[6]WU D,Pigou L,Kindermans PJ,LE ND,Shao L,Dambre J,Odobez JM-“Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition,August 2016

[7]Benitez-Garcia G,Haris M,Tsuda Y,Ukita N-"Finger Gesture Spotting from Long Sequences Based on Multi- Stream Recurrent Neural Networks,January 2020

[8]Yang J,Andric M,Mathews MM-"The Neural Basis of Hand Gesture Comprehension:A Meta-Analysis of Functional Magnetic Resonance Imaging Studies,October 2015

[9]Hakim NL,Shih TK,Kasthuri Arachchi SP,Aditya W,Chen YC,LIN CY-"Dynamic Hand Gesture Recognition using 3DCNN and LSTM with FSM Context-Aware Model,December 2019

[10]Chuang WC,Hwang WJ,Tai TM,Huang DR,Jhang YJ-"Continous Finger Gesture Recognition Based on Flex Sensors,September 2019