Available online at http://www.advancedscientificjournal.com
http://www.krishmapublication.com
IJMASRI, Vol. 1, issue 10, pp. 346 - 350, Dec. -2021
https://doi.org/10.53633/ijmasri.2021.1.10.016

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130 **IBI IMPACT FACTOR 1.5** **DOI: 10.53633/IJMASRI**

RESEARCH ARTICLE

HAND GESTURE RECOGNITION IN REAL-TIME

[1]Priyanshi Gupta, [1]Amita Goel, [1]Nidhi Sengar and [1]Vashudha Bahl

[1]Department of Information Technology, Maharaja Agrasen Institute of Technology, PSP Area, Plot No. 1, Sector-22, Rohini, Delhi-110086 Email: priyanshi.42314803118@it.mait.ac.in, amitagoel@mait.ac.in, nidhisenger@mait.ac.in, vasudhabahl@mait.ac.in

Abstract

Hand gesture is language through which normal people can communicate with deaf and dumb people. Hand gesture recognition detects the hand pose and converts it to the corresponding alphabet or sentence. In past years it received great attention from society because of its application. It uses machine learning algorithms. Hand gesture recognition is a great application of human computer interaction. An emerging research field that is based on human centered computing aims to understand human gestures and integrate users and their social context with computer systems. One of the unique and challenging applications in this framework is to collect information about human dynamic gestures.

Keywords: Tensor Flow, Machine learning, React js, handmark model, media pipeline

Introduction

The main aim of this project is to improve the communication between normal people and deaf and dumb people and second is to increase the interaction between human and computers. The vision based technology is widely used for Human-computer interaction. In the last decades some peripheral devices (mouse , keyboard) play a significant role in human-computer interaction. But nowadays, Mobile phones are bringing up better research opportunities for human centered apps where the person is a good source of information and the phone is the first in hand sensing tool. The development of computer vision technology makes it possible to solve the interface problem from the user's point of view .This will increase the human computer interaction application to communicate with

machines. For this purpose we are using a media-pipeline model. Here we are using tensorflowJs which develops ML models in JavaScript.

Fig. 1. American Sign Language

TensorFlow

Tensorflow is an open source library .It develops ML models in JavaScript, and uses ML directly in the browser or in Node.js. It is used for machine learning algorithms.It is developed by the google brain team. It builds and trains ML models using some APIs. It trains and deploys models in the cloud, on-prem, in the browser, or on-device no matter what language you use.

Media Pipe

Media pipe is a framework which is used to build machine learning pipelines for video audio etc.It uses two models . Palm detection and hand landmark model , these two work together . Media pipe model detects 21 points which are recorded from a hand in a single frame with the help of multiple models. Palm detection And Hand Landmark model :- In Palm detection model the image which is identified and draws a box around the hand. Hand landmark model works on the image formed by palm detection and provides 2D hand key points coordinates.

Workflow

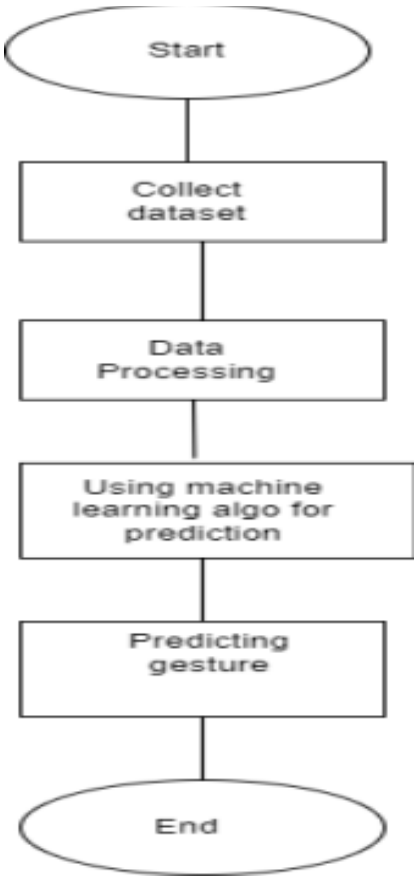
Image Processing

At first the webcam captures the hand image The image which is captured by a web camera is processed separately. In the analysis of images two steps are involved . First step is detecting the relevant image of the hand from the raw image .The second step is using these images to compute the model parameters. In this process the first step is to localize the gesture. .In the segmentation process , the image is divided into regions which are separated by boundaries.

Colour detection

In image processing applications (face detection, mask detection, gesture detection), hand colour plays an important role. It is a very difficult task because the image is sensitive to the factors - lapperance of hand, illumination, light, camera characteristics, background colours, shadows. The drawback of colour based localization is the variability of the skin colour in different lighting conditions. Now using a palm detector. Now the palm detector draws a box around this image. The hand landmark model locates 21 points 2d hand coordinates. The probability of hand presence in the input image in the form of a hand flag. . A binary flag to determine hand, i.e., left or right hand Then these points are captured by the model which is preprocessed further and passed to the keypoint classifier model which classifies the hand gestures. The 21 landmarks determination uses the same topology as both synthetic and real-world datasets were used by the model to learn the coordinates, where the relative depth w.r.t. the carpus point being learned only from synthetic images. To recover from tracking failure, a probability output mechanism is also available in the model for the case to check if a reasonably aligned hand is present or not in the provided region. If the score is lower than a fixed threshold value then the detector is re-triggered to

calculate and start tracking again. Handedness is another important requirement for interaction using hand pose. This is notably useful for a few applications where each hand is related with unique functionality. Thus a binary classification flag to predict the input hand's type, i.e., the left or right hand is available in the model.



Literature survey

Deaf and dumb Communication Interprete

The aim of this project is to create a website for deaf and dumb people through whom normal people can communicate with these people and vice versa. Hand gestures are also a way to communicate with computers. For this project we need some external device (example: - mobile phone or laptop). The camera of a mobile phone or laptop will capture the image of a hand gesture and with the help of some machine learning algorithm we will convert this gesture to the corresponding alphabet character.

Design the Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People

The main objective of this project is to make a website which converts hand gesture into text using american sign language in real time. In this paper we proposed a website foo deaf and dumb people to aid commuinaction using american sign language in which hand gesture will be converted into appropriate letter or word.

Model for hand gesture recognition

Mediapiape is a framework used to build machine learning pipelines. These pipelines are used for audio , video. It uses two models which are worked together , The first model is palm detection . In this model first we capture an image through a mobile/laptop webcam then it draws a box around the hand of the person . Second is Hand landmark , which operates the 21 points which were provided by palm detection . It draws a graph using these points. In this approach , we have recognized all 26 alphabets using hand gesture images which are taken by webcam. We have trained these 26 images . These images are preprocessed further and now we send them to the keypoint classifier model to classify the hand gesture. We are using tensorflow to convert python models into javascript.

Hand Gesture Recognition System For Dumb People

Authors presented the static hand gesture recognition system using digital image processing. For the hand gesture feature vector SIFT algorithm is used. SIFT features are calculated at fixed edges, rotation, audio addition.

Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People

In this paper, the author proposed a system to aid communication of deaf and dumb people communication using Indian sign language (ISL)

with normal people where hand gestures will be converted into an appropriate text. The main objective is to design an algorithm to convert dynamic gestures to text in real-time. Finally, after testing is done the system will be implemented on the android platform and will be available as an application for smartphone and tablet pc.

Nearest Neighbour Classification of Indian Sign Language Gestures using Kinect Camera

Zafar Ahmed Ansari and Gaurav Harit conducted extensive research into the field of accurately distinguishing the effects of Indian Sign Language. They are divided into 140 classes that include spelling numbers,letters, and other common phrases. In the database, the Kinect sensor is used. RGB 640 x 480 resolution images are taken along with their depth data.

Indian Sign Language Recognition:-

The recognition of sign symbols by Divya Deora and Nikesh Bajaj is done with PCA (Principal Component analysis). The paper also proposes recognition with neural networks. The data they acquired was using a 3 megapi xels camera and hence the quality was poor. For each sign, they took 15 images and stored them in their database. The small dataset was one of the reasons why their results were not satisfactory.

Application

.

This Hand tracking model for gesture recognition can readily be used in many day-to-day life scenarios as well as for professional purposes. After predicting the skeleton structure of the hand, we can compare gestures using some simple algorithms to compute similarities. Like, the state of each finger, e.g. if they are bent or are straight, can be determined by the angles of joints accumulated. We can also map the set of finger(s) states to a set of predefined gestures. This effective but straightforward technique allows us to predict the most basic static gestures with high efficiency. Along with static gestures recognition, it is also possible to utilize the landmark sequence to estimate various dynamic gestures.

This Hand tracking model for gesture recognition can readily be used in many day-to-day life scenarios as well as for professional purposes. After predicting the skeleton structure of the hand, we can compare gestures using some simple algorithms to compute similarities. The angles of joints accumulated, for example, can indicate the status of each finger, such as whether it is bent or straight. We may also map a collection of predefined gestures to a set of finger(s) states. This simple but powerful strategy allows us to accurately predict even the most basic static gestures. Along with static gestures recognition, it is also possible to utilize the landmark sequence to estimate various dynamic gestures.

Conclusion

In this paper, we have reviewed and discussed various end-to-end hand tracking solutions and research works as well as we've seen a suitable way to recognize and determine hand gestures using 21 landmarks. We have discussed how hands can be divided and represented as a mesh of 21 points which can be considered the main landmarks for the human hand structure. We can also differentiate among hands using the handiness binary classification. The main goal of this paper was to look at how hand poses can be determined using the CNN provided by the TensorFlow hand pose model and utilizing it for sign language detection.

References

1. Poppe. (2007). Vision-based human motion analysis: An overview. Computer Vision and Image Under- standing, 108(1-2):4–18, 2007.
2. Lukowicz,P., Ward, J.A. Unker, H. Stager, M. Tro¨ster, G. Atrash, A and Starner, T. (2004) Recognizing work- shop activity using body worn microphones and accelerometers. Proceedings of the 2nd Int Conference Pervasive Computing, pages 18–22.
3. Karantonis, D.M., Narayanan, M.R. Mathie,M. Lovell, N.H and Celler, B.G. (2006). Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on InformationTechnology in Biomedicine, 10(1):156–167.

4. Nishkam, R., Nikhil, D. Preetham, and Littman.M.L.(2005) Activity recognition from accelerometer data. In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, pages 1541–1546.
5. Bao, L and Intille, S.S. (2004). Activity recognition from user-annotated acceleration data. In editors, Pervasive Computing, pages 1–17.
6. Wu, W., Dasgupta, S., Ramirez, E.E. Peterson, C and Norman, G.J. (2012). Classification accuracies of physical activities using smartphone motion sensors. Journal of Medical Internet Research, 14(5).
7. Bin Abdullah,M.F.A., Negara, A.F.P. Sayeed, M.S. Choi, D.J and Muthu. K.S. (2012). Classification algorithms in human activity recognition using smartphones. International Journal of Computer and Information Engineering, 6:77–84.
8. Roggen, D., Calatroni, A. Rossi, M. Holleczek, T. Fo¨rster, K. Tro¨ster, G. Lukowicz, P. Bannach, D. Pirkl, G and Ferscha. A (2010). Collecting complex activity data sets in highly rich networked sensor envi- ronments. In Proceedings of the 7th International Conference on Networked Sensing Systems, 201
9. https://www.researchgate.net/publication/326972551_American_SigLanguage_Recognitin_System_ An_Optimal_Approach
10. https://www.sciencedirect.com/science/article/pii/S18770509173207 20
11. https://www.tensorflow.org/ https://reactjs.org/
12. https://www.kaggle.com/arpitjain00 7/hand-gesture-recognition-using-t ensorflow
13. https://mediapipe.dev/
14. https://cordova.apache.org/

*****