Gesture Recognition System

Diksha Verma

*Computer Information Science and Engineering University of Florida*

Gainesville, United States [dv](mailto:dverma@ufl.edu)[erma@ufl.edu](mailto:erma@ufl.edu)

***Abstract*—The prevalence of Gesture Recognition has been subtle but ubiquitous in our midst. This approach can be used to control computer input, detect sign language and for active gameplay. Our lack of understanding as well as our ignorance towards it poses a big communication barrier with such a large community. This project aims to precisely tackle this gap by creating a pattern recognition system which takes in input as the hand gestures and interpret them into gesture which will be represented by the respective emoji.**

***Index Terms*— Hand Gesture, Pattern Recognition, Machine Learning, Human Computer Interaction**

I. Introduction

With the development and growth of the technologies in the past years lead to the improved interaction between human and computer which was once restricted to the input methods like mouse and keyboard has now developed and are using new interaction methods which includes the use of biological characteristics. And this approach is currently the area of research.

With the use of facial recognition, fingerprint recognition and iris recognition, this project is using the gesture recognition method to interpret the hand pose and understand the meaning behind that gesture. As the hand pose can provide ample information as it is used in daily conversation as well and hence is useful for human computer interaction. Due to diversity and at times uncertainty of the gesture it is challenging topic for research.

II. Gesture Recognition

The hand poses are of two types which includes static and dynamic recognition, where the dynamic works involving the trajectory of the hand motion in the provided space thereby performing the tasks depending on the parameters of the trajectory obtained for example the hand gesture which are used to do up and down flip, start and projection on which they play courseware. Whereas the static recognition consists of shape of hand, interpreting the message from the hand gesture.

Initially the method for gesture recognition was using changing technology, which were supposed to be done by wearing data gloves which will then give the information of the gesture made and the location of the information to the computer. It aids to interpret the gesture and its meaning.

The data glove gives real time information precisely and then get the response from the virtual environment. It gives universal as well as direct human computer interaction mode which is beneficial due to the data being simple and fast processing speed followed by high accuracy rate.

A picture containing text, person

Description automatically generated Fig. 1. Data glove in motion

Below show is the example gesture recognition based on vision.

Diagram

Description automatically generated

Fig. 2 Vision method

In this project we are using MediaPipe Hand and Fingerpose to detect the gesture and the technologies used are Tensorflow.js and React.js which assist in making the application where you can detect the hand gesture in real time.

1. *MediaPipe Hands*

It is a finger tracking solution with greater constant change. It uses Machine Learning to check the 21 3D landmarks on the hand per frame. It uses Machine Learning pipeline which includes different models which work together. It uses two main models palm detection model and hand landmark model. The palm detection works on the full image, and it returns bounding box which is hand oriented. Whereas the hand landmark model works on small and edited images which come as an output form the palm detection model and then gives back the highly changing key-points of hand.

The images provided below are hand landmark detection output which is worked on after getting the output from palm detection.

Text

Description automatically generated with medium confidence

Fig. 3 Points from hand landmark

The fig.3 depicts all the points which are detected in your hand while going through the hand landmark detection. These points will be later used in the gesture description as well. The fig. 4 also shows the same detection method on real objects.

A picture containing colorful, different, female, set

Description automatically generated

Fig 4. Real object detected by Hand landmark detection

1. *Fingerpose*

It is a classifier which is used for the markup in the hand which are already detected by the MediaPipe Handpose. In this project it will be used to detect the gestures which corresponds to the particular emoji inside the application of source video.

The working of the fingerpose is that it initially detects the hand landmarks inside the provide input. In this project input will be in the form of a real-time video steam. After the detection are being made it will then proceed to the next step which will be working on the given landmarks. It will now identify the direction and the curl of each finger in the hand by using the marks given. The marks and the direction are judged in a particular format. The fig.5,6 show the same methodology.

Table

Description automatically generated

Fig. 5 The curl for finger

Table

Description automatically generated

Fig. 6 Name of each finger

These are the ways the landmarks for each finger are detected on the hand. In the fig8. The direction of each finger and the curl are described with in turn shows whether the finger is closed, open or curled. Also, in this direction is it pointing example up, down and diagonal.

Table

Description automatically generated

Fig. 7 Direction used for fingers

A picture containing text, person

Description automatically generated

Fig. 8 Curl defined for finger

After everything is identified it is then compared with the gesture description already given for the gestures. If the points match with the one given, and the confidence rate is more than 5 then will be give the output as that emoji during the real time video stream on the application. We are using react for the development of the application. Once we run the application it turns on the webcam and starts capturing the hand gestures by the method described above.

III. Result

The outcome of the project is shown in the fig. 9 and 10 which is detecting the hand gesture and giving the output as the emoji like that gesture.

A picture containing wall, person, indoor

Description automatically generated

Fig. 9 Victory gesture detected

A child holding a toy

Description automatically generated with low confidence

Fig.10 Thumbs Up gesture detected

IV. Conclusion

This project was able to work successfully against similar postures int different light conditions and backgrounds. Also, it was a fast detection process and allows the real time video application with low cost sensor such as PC webcam and USB camera.

Also, the study of hand gesture is not only beneficial for the deaf people which give them an ease to interact but also help in teaching the sign language and telecasting this on TV along by catching the person say and converting it into hand gesture.

V. REFERENCES

1. [Review](http://www.startasl.com/american-sign-language/) of Hand Gesture Recognition Study and Application by Yanan Xu
2. Uttam Prabhu Dessai, Amita Dessai, ”Sign Language Gesture Recog- nition System for Hearing Impaired People”, International Journal of Computer Applications (0975 – 8887), Volume 179 – No.45, 2018.
3. https://tinyurl.com/TBD-API
4. [https://www.tensorflo](http://www.tensorflow.org/tutorials)w.or[g/tutorials](http://www.tensorflow.org/tutorials)
5. Fingerpose by Prsad Pai
6. Github for MediPipe
7. MediaPipe Hands: On-device Real-time Hand tracking

Fan Zhang, Valemtin Bazarevsky