ISSN No:-2456-2165

Gesture-Based Character Recognition using SVM Classifier

Rama.S
Department of Computer Science andEngineering
SRM University
Chennai, India

Vikas. S
Department of Computer Science andEngineering
SRM University
Chennai, India

Ram Kumar. P
Department of Computer Science andEngineering
SRM University
Chennai, India

Raagul. L
Department of Computer Science andEngineering
SRM University
Chennai, India

Abstract:- Virtual Reality(VR) is a popular segment that is rapidly developing in the recent years with the advancements in compact and cost-effective GPU technologies required to render such intensive tasks in real-time. Communication in VR is usually done either through Voice or Text through a point and click method using motion controllers. It's tedious to move and point towards a specific character to type a particular word and also breaks the illusion of the VR. In this paper we have proposed a method to counter these issues, by using gesture recognition of hand movements, using accelerometer and gyroscope data, through Machine Learning. This system easily integrates into the existing platform of motion controllers as they already include all the required sensors built into them already. A prototype is also showcased using an Arduino controller and an MP-6050 (Accelerometer and Gyroscope sensor) with data transmission through Bluetooth. This paper also includes appropriate performance analysis of said method in comparison with existing methods.

Keywords:- *component; Virtual Reality; Gestures; Arduino; SVM Classifier;Sensors; Augmented Reality; Mixed Reality.*

I. INTRODUCTION

VR is a field that's gaining rapid popularity in modern day with advancements in the fields of Motion detection, Graphical processing and comparatively low-cost HMD's (Head Mounted Display). VR has the potential to encompass over everything we do with a computer. With the introduction of Mixed Reality, which is a combination of Virtual reality along with real-life, the prospect of a more ever-connected world is getting closer than ever. VR, in itself, has applications in fields ranging from Medicine, Armed Forces, Staff Training to Personal Recreation and even to the point of leading entire virtual lives, provided if the level of immersion is right. In VR, when it comes to texting or chatting, it is usually done through either voice or text chat. Text chat can be tricky or even tedious to do depending on the method used. The proposed method in this paper is to allow the user to draw the characters in the air, but in a virtual environment appear like they're drawing on a Board or even paper. These characters are recognized through Machine Learning by using Accelerometer and Gyroscope sensors which are already prebuilt into most mainstream Motion controllers used with VR like the Oculus Touch controller and the HTC Vive controller.

II. EXISTING METHODOLOGY

The most popular way to type in VR as of date is the point at the character in the keyboard on a virtual environment and press a button to select that character. The motion controller is used to point at the virtual object. The location and angle of pointing are all tracked using the inbuilt sensors and the optional tracking cameras surrounding the setup area, which is together used to accurately locate the controller in the virtual space. The accuracy of this method is also dependant on the accuracy of the tracking system itself. The drawback of this method is that it sorts of breaks the illusion of VR. Another common method for typing in VR would be voice typing, such using service like the Google Speech to Text Keyboard. The accuracy of Google's service is around 95% with it also being the threshold for speech prediction in humans. This method is also a lot faster and more convenient than the previously specified method. The drawback to this method is also that it would kind of break the illusion of VR when a specific type of action, which involves a drawing motion or typing motion, is to be elicited from the user for complete immersion. Another drawback would be the availability of the preferred language. Google Voice typing, even though it has an ever-expanding library of languages, does not provide the feature for every language that we can currently type in.

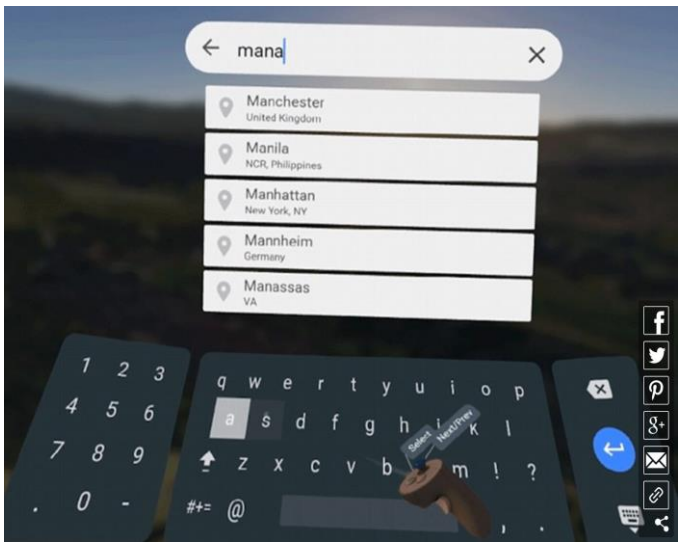

Fig 1:- Typing in Google Street View in VR

ISSN No:-2456-2165

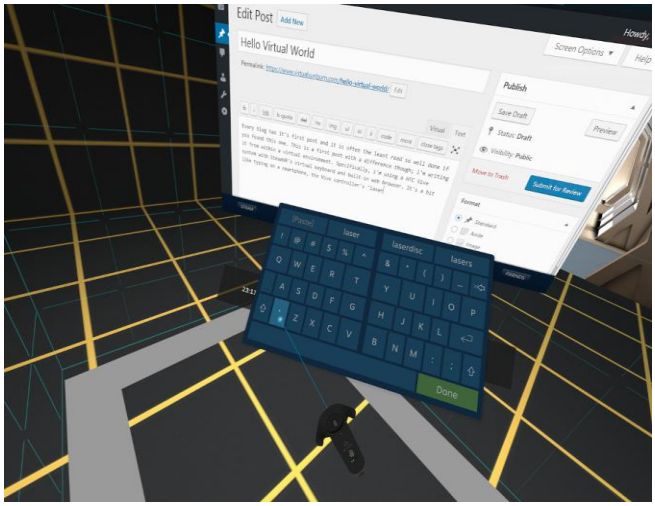

Fig 2:- Typing in Steam VR Keyboard

Previously, Hand Gestures have been observed though Accelerometer and Gyroscope sensors to do things like control robots and various other things. They have been used to detect Sign Language and even certain specific words and phrases based on their correlation with a certain action. Hand Gestures have also been observed through Image processing and Feature Extraction to the same effect.

In the paper by Bhuyan et al[3], Gyro-Accelerometer is used to control a Robotic Arm and have been met with an average accuracy of around 97% in their methods. Any hand motion detected, is reported by the dual sensors and the resulting change in data, conveys the direction of shift when compared to the default location. Using this, the same path which was detected is sent asa command to the Servo motors of the robot.

In the paper by Piawiak et al[4], A DG5 VHand Glove, which contains five flexion sensors for each finger, accelerometer and gyroscope, is used to collect the data for gesture recognition. A specific combination of Shift, Rotation and Movement of specific fingers is used to detect a specific gesture. The system was designed to detect 22 specific, commonly used gestures.

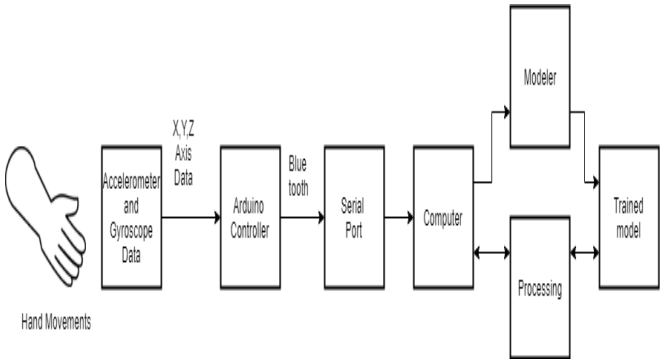

Fig 3:- Data flow

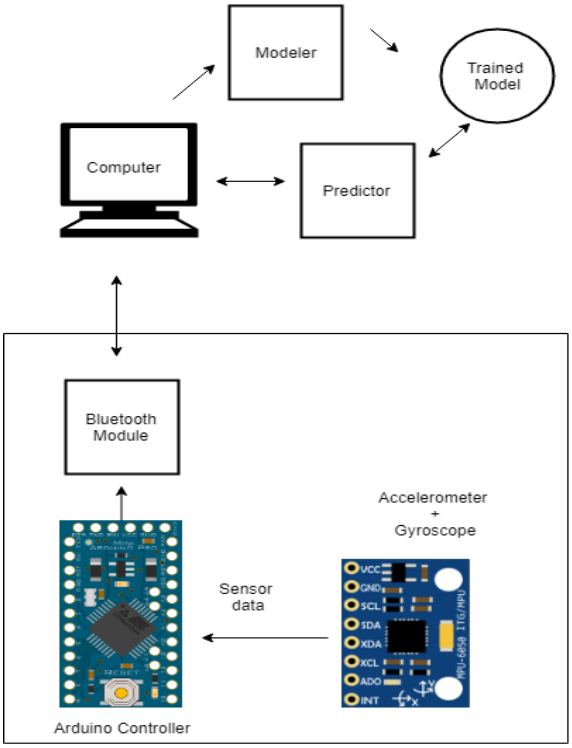

Fig 4:- Prototype Overview

III. PROPOSED SYSTEM

The proposed system is to use the existing inbuilt sensors such as the accelerometer and gyroscope to predict characters through machine learning. An SVM Classifier is used for the purposes of machine learning. The XYZ axis data from both the sensors are used to model each character. The raw axis data from the sensors are used to model for each specific character individually. The initial model is built after training multiple times with various angles and styles of the characters, so as to keep the model as accurate as possible.

The accuracy of detection increases as the number of times trained increases. Also, as the system is functionally used, the model keeps improving with more and more predictions that lie above a specific certainty factor. The proposed system would integrate easily with the existing systems of VR due to the presence of the required sensors in the existing hardware and hence would also be cost-effective in the process.

IV. IMPLEMENTATION

A prototype, for the proof of concept, of the proposed method is built using an Arduino controller. It is connected toan HC-05 Bluetooth module for Wireless connectivity. It uses an MP-6050 Accelerometer and Gyroscope sensor to provide the required data for modelling and detection. Figure 2 is used to show the overlook of the prototype and figure 3 shows the data flow across it. When the Button 1 is pressed, Transmission is initiated and the sensor data is sent through the Bluetooth COM Port. The model for classification is built using Scikit-learn's SVM classifier to classify the raw sensor data as characters and to distinguish between the different characters.

ISSN No:-2456-2165

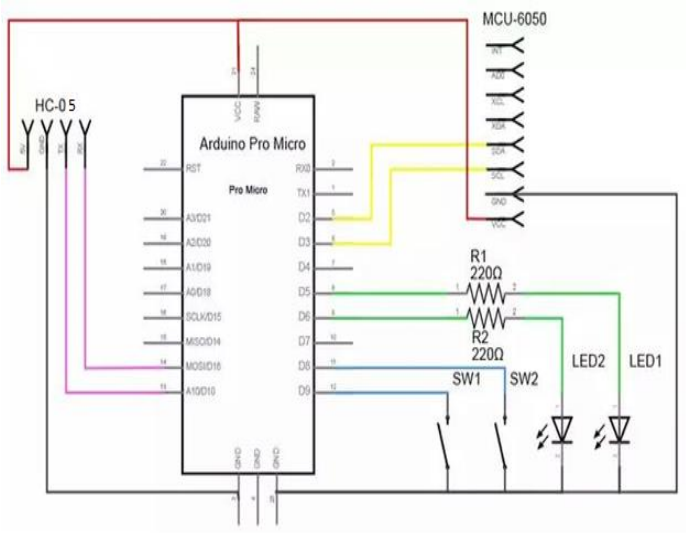

Fig 5:- Circuit diagram of Prototype

Figure 4 shows the circuit diagram of the prototype. Figure 5 and 6 shows the raw sensor data obtained while training the letter g. It also includes the characteristics of the parameters required for classifying it as the character g and f during prediction as stored in the model. The normalized x-axis, the y-axis shows the prediction certainty at every part of the gesture as per the number of samples provided.

Depending on the number of training passes per character the accuracy of prediction varies from 87% to 92%. The accuracy is measured by comparing the average amount of characters predicted correctly versus the total amount of characters provided. The complexity of the character and the likeliness of two characters tend to create some issues at times. Characters like b and h were notorious with accuracy compared to other characters and required a bit more training for accurate tracking. For these characters a little more well-defined training samples were required to have the accuracy up to the mark as the rest of the of characters. The average amount of time to train a specific character was around 120 seconds, which had a total of 40 specific trials for training.

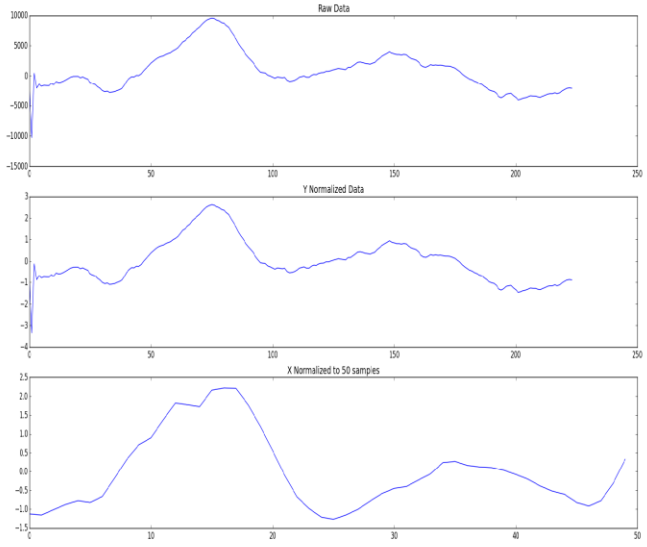

Fig 6:- Sensor Data for the character ‘g’

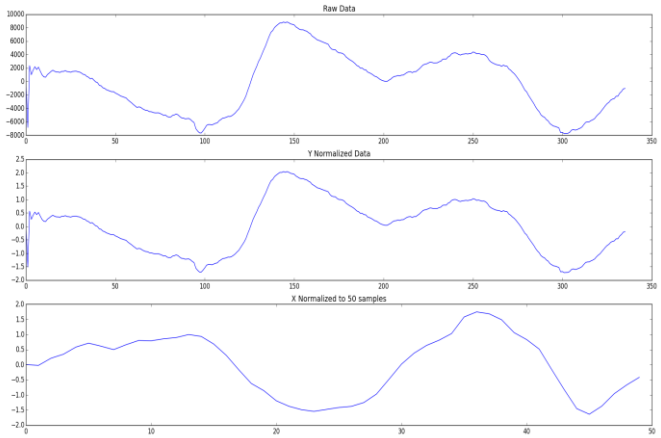

Fig 7:- Sensor Data for the character ‘f’

V. CONCLUSION AND FUTURE WORK

VR is something that is inevitable and one day should encompass our existing field of computers. With decreasing hardware cost and a steady increase in the allocation of development time towards VR applications, VR could become the next revolution for the masses. The proposed method in this paper could be very useful in VR applications that require text input, especially when they contextually require the user to perform certain hand gestures or writing actions. This method could also be used towards ALS patients or any other patients with prohibiting them from using their voice or a physical standard Keyboard.

REFERENCES

[1]. G. Sziládi, T. Ujbányi and J. Katona, "Cost-effective hand gesture computer control interface," in 7th IEEE International Conference on Cognitive Infocommunications, Wrocław, 2016.

[2]. J. Yang, Y. Wang, Z. Lv, N. Jiang and A. Steed, "Interaction with Three-Dimensional Gesture and Character Input in Virtual Reality: Recognizing Gestures in Different Directions and Improving User Input," *IEEE Consumer Electronics Magazine*, vol. 7, no. 2, pp. 64-72, 2018.

[3]. A. I. Bhuyan and T. C. Mallick, "Gyro-Accelerometer based control of a robotic Arm using AVR Microcontroller," in *The 9th International Forum on Strategic Technology*, Cox’s Bazar, Bangladesh, 2014.

[4]. P. Pławiak, T. So´snicki, M. Nied´zwiecki, Z. Tabor and K. Rzecki, "Hand Body Language Gesture Recognition Based on Signals From Specialized Glove and Machine Learning Algorithms," IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, vol. 12, no. 3, pp. 1104-1113, 2016.

[5]. J. Wang, Y. Xu, H. Liang, K. Li, H. Chen and J. Zhou, "Virtual Reality and Motion Sensing Conjoint Applications Based on Stereoscopic Display," in *Progress In Electromagnetic Research Symposium*, Shanghai, 2016.

[6]. e. a. H. Lee, "Select-and-Point: A Novel Interface for Multi-Device Connection and Control Based on Simple Hand Gestures," *CHI*, pp. 5-10, 2008.