Hand Gesture Control Quiz using Augmented Reality

Shikhar Mathur, Pawan Kumar, Rohan Kumar

DR. AKHILESH DAS GUPTA INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Guide Name:- Ms Pinky Yadav

ABSTRACT

One of the research projects in human interaction with computers is an AR-based project named Hand Gesture Control Quiz. This project works on theconcept of Augmented Reality. Augmented Reality is the interactive realization of real-world nature in which living things in the real world are developed by computer-assisted cognitive information, sometimes in a variety of senses, including visual, auditory, haptic, somatosensory and odor. This project is a hand-controlled gesture quiz where the questions will be displayed on the screen and the user must select the appropriate answer from the four options on the screen and at the end of the questions, your final score card will be displayed on the screen and the percentage achieved throughout the cycle.

INTRODUCTION

There are usually two methods of hand-detection, hardware-based, in which the user must wear the device, and the other is based on visualization using image processing and input techniques from the camera. The proposed program is based on a vision, which uses image processing and input techniques from a computer webcam. Touch-based visibility tracking. The input frame will be captured on a webcam and the systems are usually divided into categories.

Not surprisingly, all technology devices have their limitations, especially when it comes to computer devices. After a review of the different types of physical mouse, problems are identified and generalized. The following describes the common problem that the current mouse is experiencing:

• A real mouse gets wear out issues.

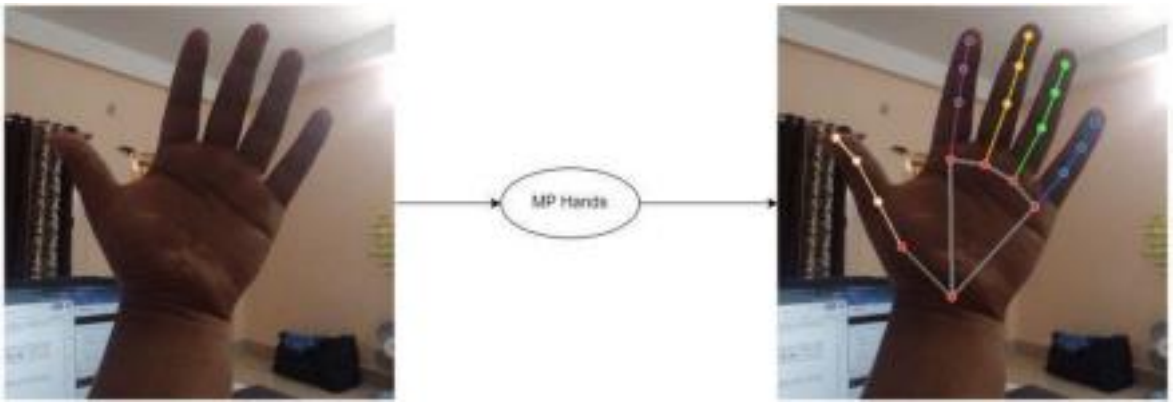
• The visible mouse needs special hardware and space to operate.

• A real mouse is not easily adapted to a variety of conditions and its performance varies from place to place.

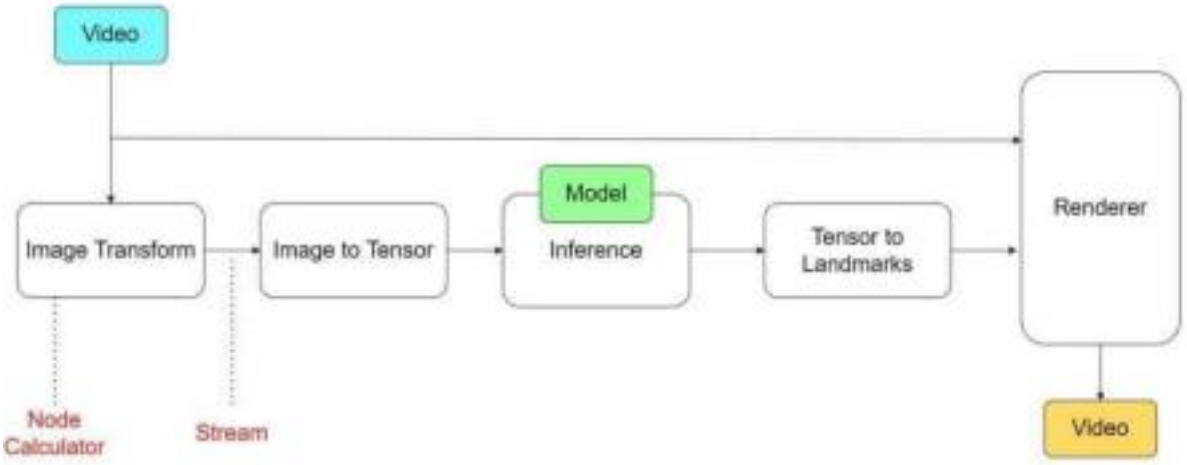
• The mouse has limited functions even in current workplaces.

• Every cordless mouse and its cordless mouse have a life span.

Considering the above issues, the applicability of virtual mouse with the help of hand gesture control becomes extremely important.



This figure represents the media-pipe working to make a visual representation of a human hand in the form of something a machine can understand.
In this specific scenario, a palm is divided into points (joints) and these joints works as the schematic points for the system to understand that the given input that the machine is getting from the live camera feed is a hand.



These are some of the counting units listed in C ++ that have been assigned certain tasks that are to be considered. Data packets (video frame or audio component) go in and out through ports inside the calculator. When you start the calculator, we announce the type of packet loading that will break through the port. Every time a graph works, the Framework uses open, process, and Close methods in calculators. Turn on start calculator; the process runs repeatedly when the package enters. The process is closed after the full graph has been applied.
The calculator, ImageTransform, captures an image in the input field and returns the converted image to 10 exit ports. On the other hand, the second calculator, ImageToTensor, takes the image as input and output the tensor.

Calculator Types In MediaPipe
1. Pre-calculated calculators are a family of image and media processing calculators. ImageTransform and ImageToTensors in the graph above fall into this category.
2. Inference calculators allow native integration with Tensorflow and Tensorflow Lite for ML detection.

3. Post-processing calculators perform post-processing ML functions such as identification, classification, and separation. TensorToLandmark is a post-processing calculation.
4. The calculators used are a family of calculators that perform final functions such as an image annotation.

SKELETON BASED RECOGNITION IMPLEMENTATION

The skeleton-based recognition specifies certain model parameters which shall improvise the detection of complex features. Where the multiple representations of skeleton data for a hand model is used for classification, it describes and elaborates upon geometric attributes and constraint and easy translated features and implements correlations of data, in order to focus on geometric and statistic features. The relatively common feature used is the joint orientation, the space between the joints, the skeletal joint location and degree of angle made between joints and trajectories and the curvature of the joints.

Hand segmentation uses depth sensor of the Kinect camera, followed by location and implementation of the fingertips using 3D connections, Euclidean distance, and geodesic distance in hand skeleton pixels to implement increased accuracy. 3D hand gesture recognition approach based on a deep learning model using parallel convolutional neural networks (CNN) processes hand skeleton joints' positions , the proposed system has a limitation and restriction where it works only with completely logical sequence. The optimal required viewpoint was estimated and the point cloud of gesture was transformed using a curve skeleton to specify the topology, then Laplacian contraction was applied to specify and target the skeleton points inside .

Where the Hungarian algorithm was further applied to calculate the match scores of skeleton point set, but the joint tracking information used by Kinect is not highly accurate which gives a result with constant vibration. A different novel method based on skeletal features was extracted from RGB recorded video of sign language, which presented difficulties in extracting highly accurate skeletal data because of occlusions. A dynamic hand gesture detection using depth and skeletal dataset for a skeleton based approach was presented in, where supervised learning (SVM) used for classification was used with a linear kernel.

Another dynamic hand gesture recognition tool using Kinect sensor depth metadata for acquiring data and segmentation was done which was further used to extract orientation feature, where the support vector machine (SVM) algorithm and HMM were being utilized for classification and recognition to evaluate and understand system performance where the SVM brought a good result than HMM in some specifications including elapsed time and average recognition rate. A hybrid method for hand segmentation was based on depth and color of data acquired by this Kinect sensor with the help of skeletal data were the further proposed in . In this particular method, the image threshold is applied to the depth frame and the super-pixel segmentation method is applied to extract the hand from the given color frame, then the two results are finally combined for robust segmentation.

LITERATURE SURVEY

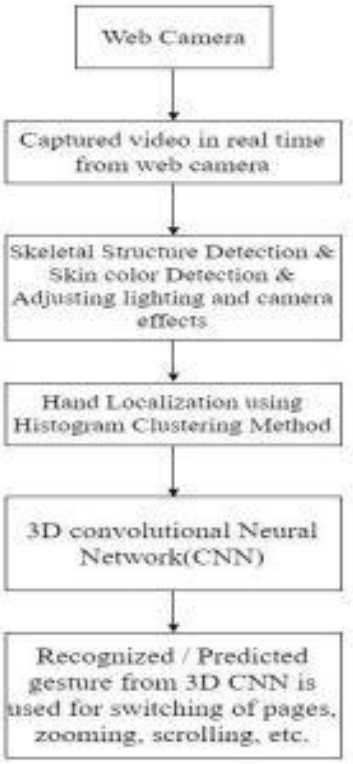
In deep learning there exists two modes of pattern recognition: one is vision based-Vision based is too dependent on the external environment. The other one is called sensor based this needs adequate light and background environment. The technique is carried out in a way such that in the data collection we use attitude sensor and data is then collected in large amounts. The next stage then is to capture the data i.e., the starting and the end points of the gesture must have an effective signal segments to be processed. Next, they use this data for the better reflection of input characteristics. At the last we classify the large amounts of data and post-process it. The conclusion we have implemented gesture recognition algorithm based on gesture sensor.

Gesture recognition is the mathematical interpretation of a human motion by a running computing device. The application is usually divided into two categories i.e. contact-based approach and vision based approach. The vision-based approach is simpler and it further employs video image processing and pattern recognition. In this paper we have trained CNN classifier to determine the shape of the hand and understand its features. In the vision-based approach we have avoided the skin color segmentation. The aim is to recognize six static and eight dynamic gestures while maintaining the accuracy and speed of the system and the recognized gestures are then used as commands to the system.

In these days the 3d pictures are being used more frequently everywhere. In this paper also we have described about 3D hand-gesture recognition (HGR). This recognition can be used in smart mobile devices/smartphones, such as head-mounted displays (HMDs) and smartphones for AR/VR applications and also, they are used in virtual 3D objects using depth sensing and hand tracking that are then used to enable the user interaction. In 3D HGR system previously these sensors have used very low power and also, they have used convolutional neural network (CNN) which can now be adopted to enhance the accuracy of the low -power stereo matching. These CNNs comprises this HGR into 6 layers. The CNN sensors can sense the skin color and applicable texture to detect the hand highly accurately. We have considered the right or left image in the CNN based stereo stage. We have selected the nearest neighbour of the image and then have coordinated the images and then have gone depth into the processing. Then we have converted the image into virtual 3D object and have further rendered the space and after doing some interactions we have obtained the final output.

Hand gesture based electronic device control is gaining more prominence nowadays. This paper represents hand gesture recognition-based Device control by using microcontrollers. The purpose of project is to design & develop a hand gesture-based system which can be easily controlled by the help of gesture recognition system by application of Augmented Reality. Experimental investigation proves stable robustness, performance and high accuracy of the proposed device controller. Then Microcontroller will become on or off the device as per gesture of a smart camera that can be defined as a vision system in which the main function is to produce a high-level understanding of the imaged scene and generate specific data to be used in an autonomous and intelligent system. The reason a smart camera is „smart“ is because there exists inside the camera a processing unit which performs application specific information processing . The main goal of the ASIP is to extract information from the captured images that are useful to an application.

PROPOSED SYSTEM AND IMPLEMENTATION



Hand gesture recognition system was engineered to record the hand movements and gestures being performed by the user and to direct the computer on the basis of the incoming information. There are many existing systems that have made use of gesture recognition using spatial modelling which is the recognition of a single gesture but there hasn't been any system that uses temporal modelling which refers to the recognition of motion of gestures. Additionally, the pre-existing systems have not been implemented in real time and use a pre-captured as an image for recognition of gestures. To overcome these limitations of the pre-existing systems, a new system was developed which was aimed at designing a vision-based hand gesture and movement recognition system with a high detection rate which can function in real time without any strict limitations or restrictions (Gloves, background etc) on the user environment. The system consists of a human computer interaction system which makes use of hand movements and gestures as input for detection and communication.

Input to the system is given from a pre-recorded video or real time through a web camera. The system then detects and recognises skin colour with the help of an adaptive algorithm in the beginning of frames. Skin colour has to be fixed for the current user based on the lighting and camera conditions. Once the skin colour is fixed, the hand is localised with the help of a histogram clustering method. Then for the detection of hand gestures in consecutive frames to distinguish the current gesture, a Machine Learning algorithm is used. These recognised gestures are then used as an input for the computer application.

The system is further divided into 3 sub systems:

• Hand and Motion Detection
A Web Camera records the hand movement and provides it as an input to OpenCV and TensorFlow Object Detector. To obtain the boundary of the hand, skin detection and edge detection are performed. This information is then sent to the 3D CNN.

• Dataset
Dataset is used for the purpose of training the 3D CNN. There are two types of datasets being used - one for hand detection and the other for motion and gesture detection. EGO dataset is used for Hand Detection and Jester Dataset is used for Motion and Gesture Detection.

• 3D CNN
CNN's are a class of deep learning neural networks used for recognising, analysing and classifying visual imagery. It consists of several layers namely : input layer, hidden layers and output layers. For better accuracy and efficiency, CNN performs back propagation. It trains and verifies recognised gestures to facilitate human-computer interaction. These human - computer interactions take place with the help of System Calls or PyAutoGUI.

CONCLUSION

Hand gesture controlled quiz addresses a problem and that problem is lack of interactivity in the classroom. With classes being held online, AR is the only way to make an online room a classroom again. Hand gesture control gives students/teachers to interact in such a way that they would not feel any barrier between their communication mode. Controlling things by hand is more natural, easier, more flexible and cheaper.
Effort into developing and maintaining reliable and robust algorithms with the help of a camera sensor has a certain characteristic and it further encounters common issues and achieve a reliable and relatively accurate result. Each technique aforementioned, however, has its own share of advantages and disadvantages and may perform depending upon situations. This hand gesture control model has also successfully implemented the concept of "Augmented Reality" and is an implementation of Machine Learning and Deep Learning.

REFERENCES

- Zhigang, F. Computer gesture input and its application in human computer interaction. Mini Micro Syst.
- Ahuja, M.K.; Singh, A. Static vision based Hand Gesture recognition using principal component analysis.
- In Proceedings of the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in
- Education (MITE), Amritsar, India, 1–2 October 2015; pp. 402–406.
- 4. Kramer, R.K.; Majidi, C.; Sahai, R.; Wood, R.J. Soft curvature sensors for joint angle proprioception.
- In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 1919–1926.

- 5. Jesperson, E.; Neuman, M.R. A thin film strain gauge angular displacement sensor for measuring finger joint angles. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New Orleans, LA, USA, 4–7 November 1988; pp. 807–vol.

- Fujiwara, E.; dos Santos, M.F.M.; Suzuki, C.K. Flexible optical fiber bending transducer for application in glove-based sensors. IEEE Sens. J. 2014, 14, 3631–3636. [CrossRef]

- Shrote, S.B.; Deshpande, M.; Deshmukh, P.; Mathapati, S. Assistive Translator for Deaf & Dumb People. Int. J. Electron. Commun. Comput. Eng. 2014, 5, 86–89.

- Gupta, H.P.; Chudgar, H.S.; Mukherjee, S.; Dutta, T.; Sharma, K. A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors. IEEE Sens. J. 2016, 16, 6425–6432. [CrossRef]

- . Lamberti, L.; Camastra, F. Real-time hand gesture recognition using a color glove. In Proceedings of the International Conference on Image Analysis and Processing, Ravenna, Italy, 14– 16 September 2011; pp. 365–373