

Augmented Reality based Gesture Controlled Virtual Board

Mohit Rakesh Taparia, Neavil Porus A
Electronics and Communication Engineering
Knowledge Institute of Technology
Salem

Jeevaa D, Gowri Shankar S
Electronics and Communication Engineering
Knowledge Institute of Technology
Salem

N. Santhiyakumari
Professor / Head
Electronics and Communication Engineering
Knowledge Institute of Technology
Salem

Abstract— In recent times, sharing various integrated devices for user-interaction or input data transmission usage have increased exponentially (Example : Laptop, Macbook). Due to the high usage of traditional Keyboard & Mouse, the physical aspect of human body has started to take a toll towards its health. Thus, this project is developed in order to lessen the distance among the actual global and the augmented surroundings to provide a blended Virtual Board for dynamic workspace and improvised interaction among the users. The focus of Human Computer Interaction is to provide intercommunication between the user and the computer by making computer easier to the user needs and allowing the user to interact with the information using natural hand gestures using gesture recognition, Augmented Reality, Computer Vision, with Virtual Keyboard, Gesture placed in Database for commands which will be initially trained using machine learning algorithm that applies the use of OpenCV libraries and Developed API's. Virtual Board can be used by folks who are specially-abled and those who can't communicate well and talk with others using this proposed virtual machine. Gesture based user interface technology is a science of tomorrow with the aim of making man-machine interface more intuitive. This technology can be used in positive ways especially by enhancing daily by recognizes the objects around us, displaying information automatically and letting us to access and use free hand gestures when needed.

Keywords:- Computer Vision, OpenCV, Virtual Keyboard, Augmented Reality, Gesture Recognition, User-Interaction)

I. INTRODUCTION

In recent years, the world is looking for safer and sophisticated methods to interact with people and gadgets[1]. Human-Computer Interaction (HCI) is an ever-evolving advancement in the development of technology as a new method of communication[2,3] between people and computers in the modern world [5,6]. Several new assistive methods, such as virtual reality, sign language recognition, speech recognition, visual analysis, brain activity, touch-free writing, steganography, anomaly detection techniques have emerged in recent years to achieve this goal.

Recently, virtual keyboard-based character writing has been widely adopted to realize the non-touch input system. As technology advances, these types of

innovations are contributing enormously to everyday tasks and making life easier [7].

This project utilizes virtual keyboard and gesture recognition as topics which come under two major computer science fields like augmented reality, HCI and Virtual System has been made with a goal of interpreting human gestures via mathematical algorithms. This algorithm allows users to use simple finger gestures to control or interact with the keyboard without physically touching them. Gesture recognition widely enables or equips humans to communicate with the machine and interact naturally without any mechanical devices or computing devices.

Using the concept of gesture recognition [4,8], it is possible to point a finger at the computer screen or web camera so that the keypad will be pressed accordingly to form meaningful sentences or words. The system was implemented with a computer vision-based approach, making use of an RGBD camera and a generic display monitor to create and immersive [9], touchless keyboard.

A. OBJECTIVE OF THE PROJECT:

The basic objective of the project is to develop a virtual gesture-controlled board using the concepts of hand gesture recognition and image processing which will function according to the methodology using hand gestures provided towards the detector, similarly using hand tracking can use keyboard functions which will be defined as per the convenience of the user [10]. This will be reducing the cost of hardware and would also make the usage of keyboards much more volatile. This approach will make tasks simpler and easier while providing a blended solution of touch free usage of gadgets and components. In addition to that, this technology will play a vital role in the ongoing booming Web3 Devices and Meta based Architectures [11].

B. PROBLEM FORMULATION:

Recent years have been marked as a sharp increase in the number of ways in which people interact with computers. The keyboard and mouse used to be the primary interfaces for controlling a computer but in today's trend, users utilize touch screens, infrared cameras [2], and accelerometers (for example, within the

iPhone) to interact with technology. In light of these changes and the proliferation of small cameras in many phones and tablets.

Human-computer interface researchers have investigated the possibility of implementing a keyboard style interface using a camera as a substitute for actual keyboard hardware. Broadly speaking, these researchers envision the following scenario: A camera observes the user's hands, which rest on a flat surface.

The camera may observe the hands from above the surface, or at an angle. The virtual keyboard's software analyzes those images in real-time to determine the sequence of keystrokes chosen by the user. These researchers envisioned several applications for this technology: in some countries, users speak many different languages, which makes producing physical keyboards for many different orthographies expensive, a camera-based keyboard can easily support many languages [3,14].

Smart-phone and tablet users may occasionally want to use a full-sized keyboard with their device, but are unwilling to carry a physical keyboard. Since most mobile devices are equipped with a camera, a camera-based keyboard will be a software-based [15] solution for this problem.

II. METHODOLOGY

When looking on the sort of approaches the computer file or source is used for decoding a tip that may well be wiped out totally in different ways. However, most of the techniques have faith in key pointers portrayed during a 3D coordinate system. Supported the relative motion of the gesture can be detected with a high accuracy, depending on the standard of the input and also the algorithm's approach. Currently, keyboards are static and their interactivity and usefulness would increase if they were created dynamic and adaptable. Numerous on-screen virtual keyboards are obtainable however it's troublesome to accommodate a full-sized keyboard.

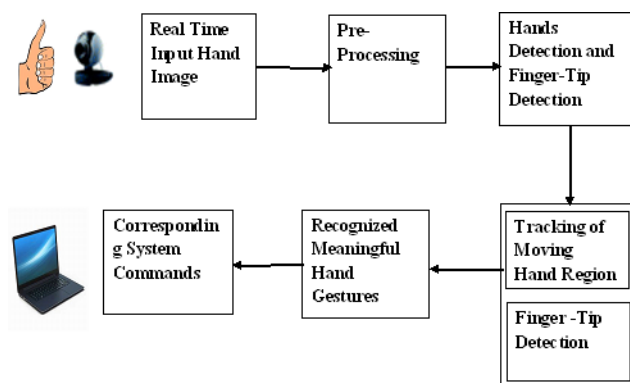


Fig -1: Block Diagram of Gesture Controlled Device

SYSTEM ARCHITECTURE

A. REAL TIME INPUT IMAGE PRE PROCESSING (CAMERA LIVE FEED:

The image analysis techniques are used to convert the user's surface-touches into keystrokes. Most state-of-the-art camera-based virtual keyboard schemes as shown in

Fig 1, are implemented using the following sequence of image analysis techniques. The system's camera captures Real Time Pre-Processing Input Hand Image Hands Detection and Finger-Tip Detection Recognized Meaningful Hand Gestures Tracking of Moving Hand Region Finger –Tip Detection. Corresponding System Commands an image "I" for processing.

The skin segmentation is applied on region "I" Then a binary image is uploaded with a region "H" representing the hand(s) in the scene. "H" is analyzed to determine the locations of the user's fingertips. The references listed above determine contours that parameterize the boundary of "H". It is associated with fingertips with the points on the contours that optimize certain geometric quantities (for example, curvature) [2]. Two of these geometric approaches to finding the fingertips and compare their performance. Some micro gestures can be formed by different combinations of fingers [4].

B. FINGER TIP DETECTION - IMAGE PROCESSING:

Step 1 : The commands are pre-trained and stored in the system.

Step 2 : Given the estimated fingertip positions, one must calculate which fingertips are touching the keyboard's keys. The shadow analysis technique is used to solve this problem. To map touch points to key presses, the keys of the virtual keyboard are described as rectangles in R2.

Step 3 : For five commands, including cursor, translation, scroll left/right, previous/next, duplicate, participants selects a micro gesture from a set of micro gestures that were similar, except when the images are either just the index finger; the index and middle finger; or it uses all five fingers.

Step 4 : The layout of the keyboard is known at compile time and the keyboard mat has control points from the stability of the mapping gets perspective correction transformation. Then a simple formula is used to convert the keyboard-space coordinates of the touch points to key presses.

Step 5 : The algorithm detects the input finger-tip and verifies with the pre defined models to identify gestures.

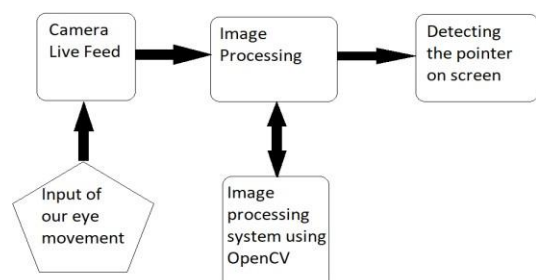


Fig -2: Block Diagram of Gesture Controlled Virtual Board

C. ALGORITHM USED FOR HANDTRACKING:

The detection of hand gestures and hand tracking is achieved when the Media-Pipe framework is used, and OpenCV library is used for computer vision. The algorithm makes use of the machine learning concepts to track and recognize the hand gestures and hand tip.

D. FINGERTIP DETECTION ALGORITHM:

The estimation of the locations of the user's fingertips (in image-space) are based on geometrical features of the contours and region. The contour is represented as - T as a sequence of positions in image-space.

Positional Sequencing is a method of sequencing keyboard that simultaneously generates information about both identity and location of keys sequences. The method involves detecting the location of sequence specific recognition, it's given in equation-1.

Position Sequencing: $\{P1 = (x1, Y1), P2 = (x2, Y2)\}$

(1)

Angle Between Two Vectors The angle between two vectors is the angle between their tails. It can be found either by using the dot product (scalar product) or the crossproduct (vector product).

The angle in between the displacement vectors does not give a very good idea of how the contour bends around the hand, so a subsequence y in which the points of the contour are spaced further apart. In experiments, the values I to assigned be every tenth point of-] is given in equation-3.

Euclidean Distance: $\{Pi \& Pi +1 = I; (0, \pi)\}$ (2)

The Euclidean distance between two points in Euclidean space is the length of a line segment between the two points. It can be calculated from the Cartesian coordinates of the points using the Pythagorean theorem, therefore occasionally being called the Pythagorean. Given a contour, one can derive several other sequences of geometrical significance and use these to locate the fingertips.

The previous processing step gives us a contour 'I' consisting of pixel locations pi such that the Euclidean distance in equation-2.

Angle of Vectors: $\{Pi + 1 - Pi \& Pi - 1 - Pi = I; \{-\pi/2, 0, \pi/2, \pi\}\}$ (3)

E. APPLICATION OF DATA PRE-PROCESSING (MEDIA-PIPE & OpenCV):

Media-Pipe is a framework which is used for applying in a machine learning pipeline, and it is an open-source framework of Google. The Media-Pipe framework is useful for cross platform development

since the framework is built using the time series data. The Media-Pipe framework is multimodal, where this framework can be applied to various audios and videos.

The Media-Pipe framework is used by the developer for building and analyzing the systems through graphs, and it has also been used for developing the systems for the application purpose. The steps involved in the system that uses Media-Pipe are carried out in the pipeline configuration[15]. The pipeline created can run in various platforms allowing scalability in mobile and desktops.

III. MEANINGFUL HAND GESTURE RECOGNITION IN APPLICATION:

The Media-Pipe framework is based on three fundamental parts; which are performance evaluation, framework for retrieving sensor data, and a collection of components which are called calculators, and it is reusable. A pipeline is a graph which consists of components called calculators, where each calculator is connected by streams in which the packets of data flow through.

Developers are able to replace or define custom calculators anywhere in the graph creating their own application. The calculators and streams combined create a data-flow diagram [16]; the graph is created with Media-Pipe where each node is a calculator and the nodes are connected by streams. Single-shot detector model is used for detecting and recognizing a hand or palm in real time.

OpenCV is a computer vision library which contains image-processing algorithms for object detection. OpenCV is a library of python programming language, and real-time computer vision applications can be developed by using the computer vision library.

The OpenCV library is used in image and video processing and also analysis such as face detection and object detection. Here it is used as a tool to recognize the gestures and hand signs provided by the keyboard to deliver accurate results [12].

The single-shot detector model is used by the Media- Pipe. First, in the hand detection module, it is first trained for a palm detection model because it is easier to train palms. Furthermore, the non maximum suppression works significantly better on small objects such as palms or fists [17]. A model of hand landmark consists of locating 21 joint or knuckle co-ordinates in the hand region.

RESULTS & DISCUSSION:

A laptop with web cam has been used to implement the virtual board and it can be used in device having a front camera and image processing and OpenCV libraries installed in it. As a rule of thumb, if the laptop does not have OpenCV and python installed in it, then this system won't work or produce the desired output. The Virtual Board Implementation Algorithm is shown in Fig-3 and it's software Design is shown in Fig-4.

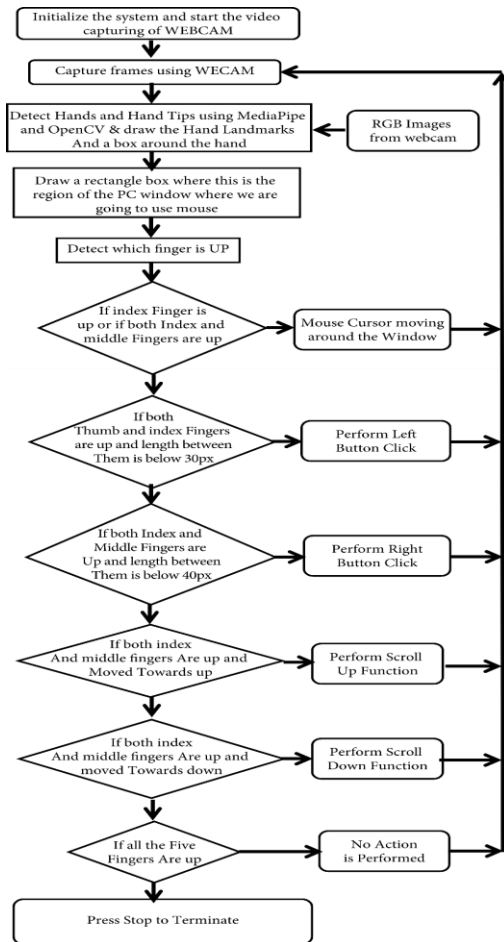


Fig 3 - Flowchart of the real-time AI virtual Keyboard system.

Virtual Keyboard has been implemented using OpenCV and Media-pipe Packages in PyCharm Software and its working has been shown in the above Fig-3.

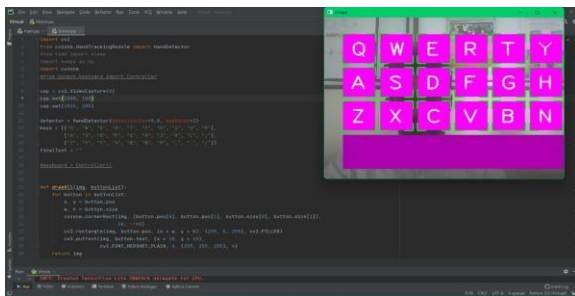


Fig-4: Software Design

The project works with 98% accuracy within a distance of 20 meters and recognizes multiple sign gestures simultaneously without any delay. Furthermore, the overall addition of displacement vectors has helped in juggling the unstable data channeling to move more effectively in terms of speed and bandwidth.

Once the setting is done the display shows a virtual key-board which is now ready to detect the keys. A proper click gesture is made for the user i.e., the fingertip first moves inwards and then outwards. The two consecutive frames are captured by the web camera

and only if the difference in the two frames is noticeable the gesture is recognized as a click [18,19]. Once a proper click is made on any key the center of that key is marked with blue. The difference in the area of the click and that of the center is calculated which make sure that the key is clicked, detected and displayed. This algorithm which is used to create this is as follows: A method which separates foreground from the background and then setup out-file for writing video.

Also, a method is used for determines the distance between two points using Euclidean distance. Now, the webcam detects the pressing of letters in the air by seeing the virtual keyboard. Webcam captures the frame and analyzes. This system is also enabled the use of swiping motion and this project recognizes swiping motion so that [20], it can click numbers, alphabets and other symbols in the keyboard. This is because our virtual keyboard is in the form of a long straight row enabled with clicking and swiping motion.

CONCLUSION

This project have been implemented with virtual keyboard on the screen without any use of external hardware. This algorithm implementation has been improved and allows web cam to quickly capture and comparing the consecutive frames and provide better results with less complexity. The keyboard would recognize a click only if there is a significant movement of the fingertip.

If the movement is quick the web cam might not be able to detect. The Virtual Keyboard detects the contour difference when the finger is placed in accordance with the letter on the keyboard and then selects and displays the letters. When the user selects a series of letters, the program analyses the coordinate's and displays the same. The user can also move the keyboard with swiping motion. From the results of the model, the proposed AI virtual keyboard system has performed better than existing part of the architecture with better accuracy and also it overcomes most of the limitations of the existing systems.

The limitations of the model in accuracy of clicking of key functions and recognition of some hand gestures are leveraged using the improvised fingertip detection algorithm, thus produces more accurate results. Since this model has good accuracy, the AI virtual keyboard can be used for real- world applications, and also, it can be used to reduce the spread of COVID-19, since this system can be used virtually using hand gestures without using the traditional physical keyboard.

REFERENCES

[1] Christy, A., Vaithyasubramanian, S., Mary, V.A., Naveen Renold, J. (2019), "Artificial intelligence based automatic decelerating vehicle control system to avoid misfortunes", International Journal of Advanced Trends in Computer Science and Engineering, Vol. 8, Issue.6, Pp.3129-3134

-
- [2] P. Chakraborty, D. Roy, M. Z. Rahman, and S. Rahman, "Eye Gaze Controlled Virtual Keyboard", International Journal of Recent Technology and Engineering (IJRTE), 2019.
- [3] H. Du, T. Oggier, F. Lustenburger and E. Charbon, "A virtual keyboard based on time-3D optical ranging Proc". British Machine Vision Conference (BMVC), Oxford. pp. 220-229, Sept. 2005.
- [4] P. Chakraborty, D. Roy, M. Z. Rahman, and S. Rahman, "Eye Gaze Controlled Virtual Keyboard", International Journal of Recent Technology and Engineering (IJRTE), 2019.
- [5] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", International Conference on Learning Representations, 2014.
- [6] T. Lee and T. Höllerer, "Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking", in Proceedings of IEEE International Symposium on Wearable Computers, pp.241-242, 2007.
- [7] G. M. Gandhi and Salvi, "Artificial Intelligence Integrated Blockchain For Training Autonomous Cars," 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2019, pp. 157-161.
- [8] S. Sadhana Rao, "Sixth Sense Technology", Proceedings of the International Conference on Communication and Computational Intelligence- 2010, pp.336-339.
- [9] Md. Shahinur Alam, Mahib Tanvir, Dip Kumar Saha, Sajal K. Das, "Two Dimensional Convolutional Neural Network Approach for Real-Time Bangla Sign Language Characters Recognition and Translation", SN Computer Science, vol. 2
- [10] Deep Learning-Based Real-Time AI Virtual Keyboard System Using Computer Vision to Avoid COVID
- [11] "Human Factors for Design of Hand Gesture Human-Machine Interaction System", H.L Wachs J.P.; Edan, Y. (SMC 2006).
- [12] "User-defined gestures for surface computing" Wobbrock, J.O.; Morris, M.R.; Wilson, A.D.(200e).
- [13] "Vision based hand gesture recognition for human computer interaction" Rautaray, S.S.; Agrawal, A. (2015). [14] "Realtime computer vision with openCV," K. Pulli, A. Baksheev, K. Korniyakov, and V.Eruhmov, (2003). [5] "Hand Gesture Recognition:
- [14] "Realtime computer vision with openCV," K. Pulli, A. Baksheev, K. Korniyakov, and V.Eruhmov, (2003).
- [15] "Hand Gesture Recognition: A Literature Review", R. Zaman, K. Noor, A. Ibraheem, (IJAI),
- [16] "Would You Do That? Understanding Social Acceptance of Gestural Interfaces" Montero, C.S.; Alexander, J.; Marshall, M.T.; Subramanian, S. (2010)
- [17] Gesture Supporting Smart Notice Board Using Augmented Reality, Smart and Innovative Trends in Next Generation Computing Technologies, June 2018
- [18] Will Gesture-Recognition Technology Point the Way? (computer.org)
- [19] David Mace, Wei Gao, and Ayse Coskun. 2013. "Accelerometer-based Hand Gesture Recognition Using Feature Weighted Naive Bayesian Classifiers and Dynamic Time Warping". In Proceedings of the Companion Publication of the 2013 International Conference on Intelligent User Interfaces Companion (IUI '13 Companion). ACM, New York, NY, USA, pp. 83-84.
- [20] "Real-time hand gesture-based interaction with objects in 3D virtual environments" Kim, J.O.; Kim, M.I Yoo. K.H. (2013).