International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 04 | Apr 2022 www.irjet.net p-ISSN: 2395-0072

VIRTUAL PAINT APPLICATION USING HAND GESTURES

Niharika M[1], Neha J[2], Mamatha Rao[3], Vidyashree K P[4]

[1,2,3] *Final year student, Dept. of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India,*
[4] *Assistant Professor, Dept. of Information Science and Engineering, Vidyavardhaka College of Engineering , Mysuru, India,*

Abstract - *It's been really tough to teach students on an online platform and make the lesson interesting during the COVID-19 pandemic. As a result, there was a need for a dust-free classroom for children. Using MediaPipe and OpenCV, this article presents a unique paint application that identifies hand movements and tracks hand joints. This program uses hand gestures to give users with an intuitive method of Human Computer Interaction (HCI). HCI's major goal is to improve human-computer interaction.*

Key Words: Hand Gesture Recognition, Human Computer Interaction, Computer Vision, Paint, MediaPipe

1. INTRODUCTION

Currently, people and machines interact mostly through direct contact methods such as the mouse, keyboard, remote control, touch screen, and other similar devices, whereas people communicate primarily through natural and intuitive non-contact methods such as sound and physical motions. Many researchers have attempted to help the computer identify other intentions and information using non-contact methods like as voice, facial expressions, physical motions, and gestures. A gesture is the most crucial aspect of mortal language, and gestures also play a significant role in mortal communication. They're considered to be the simplest way for humans and computers to communicate. Sign language recognition, robotics, and other applications fall under the category of gesture recognition.

Two methods are typically used for Gesture recognition for HCI applications. The first is based on wearable or direct physical approaches, while the second is based on computer vision and does not require any sensors to be worn. The data-glove, which is made up of sensors to capture hand motion and location, is used in the wearable or direct contact technique. The vision-based technique uses the camera to provide contactless communication between humans and machines. Computer vision cameras are simple to operate and inexpensive. However, due to differences in hand sizes, hand position, hand orientation, lighting conditions, and other factors, this approach has certain limitations.

In this paper, we introduce a virtual paint application that uses hand gestures for real-time drawing or sketching on the canvas. Hand gesture-based paint application can be implemented using cameras to capture hand movement. To accomplish activities like as tool selection, writing on the canvas, and clearing the canvas, an intangible interface is created and implemented using vision-based real-time dynamic hand gestures. The images of the hands are taken with the system's web camera and processed in real time with a single-shot detector model and media pipe, allowing the machine to communicate with its user in a fraction of a second.

2. LITERATURE REVIEW

Many methods are used for hand gesture recognition in real-time. Sayem Mohammad Siam, Jahidul Adnan Sakel, and Md. Hasanul Kabir has proposed a new method of HCI (Human-Computer Interaction), that uses marker detection and tracking technique. Instead of having a mouse or touchpad, two colored markers are worn on the tips of the fingers to generate eight hand movements to provide instructions to a desktop or laptop computer with a consumer-grade camera [1]. They have also used the "Template matching" algorithm for the detection of markers and Kalman Filter for tracking. In [2] the developed system uses a data glove-based approach to recognize real-time dynamic hand gestures. The data glove has ten soft sensors integrated in it that measure the joint angles of five fingers and are used to collect gesture data. Real-time gestures are recognized using techniques such as gesture spotting, gesture sequence simplification, and gesture recognition.

Shomi Khan, M. Elieas Ali, Sree Sourav Das have developed a system that uses a skin color detection algorithm to convert ASL (American Sign Language) into text from real-time video [3]. Because skin color and hand shape differ from person to person, detecting the hand might be challenging. The technology uses two neural networks to overcome this. The SCD (Scalable color descriptor) neural network is the first. The picture pixels are fed into the SCD neural network, which determines whether or not they are skin pixels. The second one is HGR (Hand gesture recognition) neural network to which the extracted features will be provided. The features will be extracted by two distinct algorithms namely Finding the fingertip and Pixel segmentation.

Pavitra Ramasamy and Prabhu G [4] have proposed a revolutionary technology in which the user can write the alphabet or type whatever he or she wants by merely waving

his or her finger over a colorful LED light source. Only the color of the LED is tracked to extract the movement of the finger sketching the alphabet. The color of the tracked object is changed to white, while the background is changed to black. The black and white frames are stitched together to create a single black and white image of the alphabet that the user wanted to draw.

[5] To accomplish mouse actions such as moving the mouse cursor, clicking left, and clicking right with hand gestures, an intangible interface is conceived and implemented utilizing vision-based real-time dynamic hand gestures. MATLAB is used for the implementation of the system. S. Belgamwar and S. Agrawal [6] have developed a new HCI technique that incorporates a camera, an accelerometer, a pair of Arduino microcontrollers, and Ultrasonic Distance Sensors. The primary concept behind this interface is to capture motions using Ultrasonic Distance Sensors. The distance between the hand and the distance sensor is calculated to record the gestures.

For 3D hand gesture detection, Quentin De Smedt, Hazem Wannous, and Jean-Philippe Vandeborre [8] used a skeleton-based model. They used the geometric shape of the hand to obtain an effective descriptor from the Intel Real-Sense depth camera's hand skeleton linked joints. The skeleton-based approach is better than the depth-based approach. In [9] Prajakta Vidhate, Revati Khadse, and Saina Rasal have developed a virtual paint application that uses ball-tracking technology to track the hand gestures and write on the screen. They have used a glove with a ping pong ball attached to it as a contour. [10] Ruimin Lyu, Yuefeng Ze, Wei Chen, and Fei Chen presented a customizable airbrush model that uses the Leap Motion Controller, which can track hands, to create an immersive freehand painting experience.

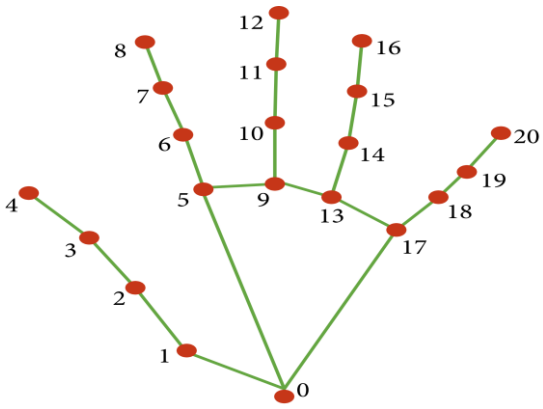
3. ALGORITHM USED FOR HAND TRACKING

Hand gesture recognition and tracking are handled by the MediaPipe framework, while computer vision is handled by the OpenCV library. To track and recognize hand movements and hand tips, the program makes use of machine learning ideas.

3.1 MediaPipe

MediaPipe is a Google open-source framework that was initially released in 2019. MediaPipe has some built-in computer vision and machine learning capabilities. A machine learning inference pipeline is implemented using MediaPipe. ML inference is the process of running real data points. The MediaPipe framework is used to solve AI challenges that mostly include video and audio streaming. MediaPipe is multimodal and platform independent. As a result, cross-platform apps are created using the framework. Face detection, multi-hand tracking, hair segmentation, object detection, and tracking are just a few of the applications that MediaPipe has to offer. MediaPipe is a framework with a high level of fidelity. Low latency performance is provided through the MediaPipe framework. It's in charge of synchronizing time-series data.

The MediaPipe framework has been used to design and analyze systems using graphs, as well as to develop systems for application purposes. In the pipeline configuration, all of the system's steps are carried out. The pipeline that was designed can run on a variety of platforms and can scale across desktops and mobile devices. Performance evaluation, sensor data retrieval, and a collection of components are all part of the MediaPipe framework. Calculators are the parts of the system. The MediaPipe framework uses a single-shot detector model for real-time detection and recognition of a hand or palm. It is first trained for the palm detection model in the hand detection module since palms are easier to train. It designates a hand landmark in the hand region, consisting of 21 joint or knuckle coordinates as shown in the Figure 1.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Fig -1: Coordinates or landmarks in the hand

3.2 OpenCV

The computer vision library OpenCV is a must-have for everyone who works with computers. It includes object detection image-processing methods. OpenCV is a python package for creating real-time computer vision applications. Image and video processing and analysis are handled by the OpenCV library.

4. ALGORITHM USED FOR HAND TRACKING

The various constraints in the system are explained in the flowchart of the Virtual Paint Application in Figure 2.

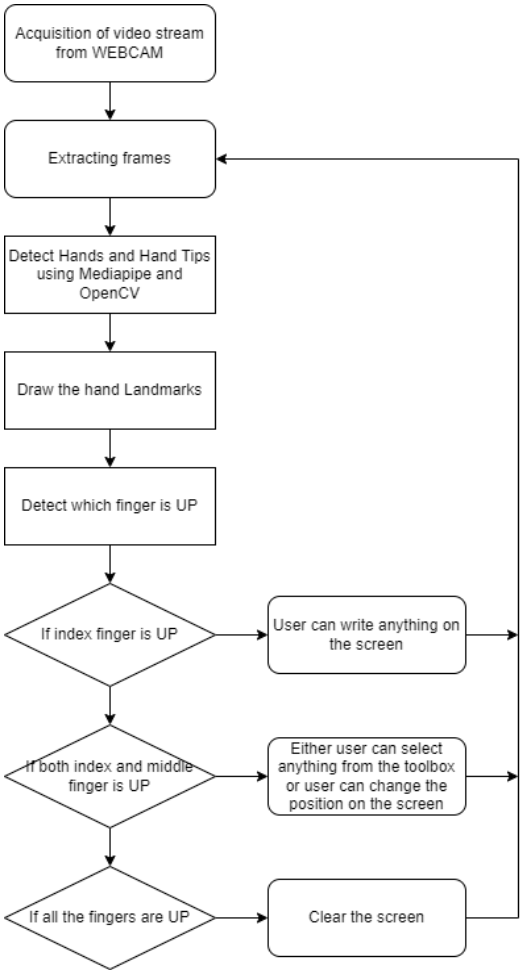


Fig -2: Flowchart of the virtual Paint Application

The virtual paint application presented is based on the frames recorded by the PC's web camera. The frames are captured by the web camera and sent to the system. The method makes use of a web camera to capture each frame till the application is finished. The video frames are transformed from BGR to RGB color to locate the hands in the video frame. The system then determines which finger is up by comparing the tip Id of the corresponding finger found via the MediaPipe to the respective coordinates of the up fingers, and then performs the appropriate function. The user can write anything on the screen if his or her index finger is raised. If both index finger and middle finger are up, the user can either change position on the screen or can select any tool provided in the toolbar of the application. If all the fingers are up except the thumb finger, the user can clear the screen. If all the fingers are up, then no action is performed on the screen.

5. CONCLUSION

The virtual paint application's fundamental goal is to deliver an AI-based tool that allows users to draw anything on screen using hand movements. This system also gives the user the option of selecting any tool from the toolbar. The user can save their completed work or see their drawing process as a replay animation with this application.

6. FUTURE WORK

This work can be further improved by experimenting with different interpolation methods such as PyGame which includes a line drawing method that could help produce smoother and cleaner lines. In the same vein, a variety of brush shapes and textures can be implemented to make this application more robust.

REFERENCES

[1] Siam, Sayem & Sakel, Jahidul & Kabir, Md. . Human Computer Interaction Using Marker Based Hand Gesture Recognition., 2016

[2] M. Lee and J. Bae, "Deep Learning Based Real-Time Recognition of Dynamic Finger Gestures Using a Data Glove," in IEEE Access, vol. 8, pp. 219923-219933, 2020, doi: 10.1109/ACCESS.2020.3039401.

[3] S. Khan, M. E. Ali, S. Das and M. M. Rahman, "Real Time Hand Gesture Recognition by Skin Color Detection for American Sign Language," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6, doi: 10.1109/EICT48899.2019.9068809.

[4] P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563.

[5] P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563.

[6] S. Belgamwar and S. Agrawal, "An Arduino Based Gesture Control System for Human-Computer Interface," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-3, doi: 10.1109/ICCUBEA.2018.8697673.

[7] Y. Wu and C. -M. Wang, "Applying hand gesture recognition and joint tracking to a TV controller using CNN and Convolutional Pose Machine," 2018 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 3086-3091, doi: 10.1109/ICPR.2018.8546209.

[8] Q. De Smedt, H. Wannous and J. Vandeborre, "Skeleton-Based Dynamic Hand Gesture Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2016, pp. 1206-1214, doi: 10.1109/CVPRW.2016.153.

[9] Prajakta Vidhate, Revati Khadse, Saina Rasal, "Virtual paint application by hand gesture recognition system", 2019 International Journal of Technical Research and Applications, 2019, pp. 36-39.

[10] R. Lyu et al., "A flexible finger-mounted airbrush model for immersive freehand painting," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, pp. 395-400, doi:10.1109/ICIS.2017.7960025.