AJEMA GLOBAL RESEARCH NETWORK

AMERICAN Journal of Engineering, Mechanics and Architecture
Volume 3, Issue 5, 2025 ISSN (E): 2993-2637

Algorithm and Software Implementation for Hand Gesture-Based Control

Abdurashidova K. T., Nurmamatov S., Akbarova M., Madaminov I.
Tashkent university of information technologies named after Muhammad al-Khwarizmi

Abstract: The use of palm motion identification in controlling digital devices has become popular due to the advancement of synthetic cognition innovation. A palm motion-governed digital pointer framework that utilizes AI algorithms to identify palm gestures and translate them into pointer movements is proposed in this paper. The framework is designed to provide an substitute platform for people who have difficulty using a conventional pointer or keyboard. The proposed framework uses a webcam to capture images of the operator's palm, which are processed by an AI procedure to identify the gestures being made. The framework is trained using a dataset of palm gestures to identify different gestures. Once the motion is recognized, it is translated into a corresponding pointer motion, which is then executed on the digital display. The framework is designed to be scalable and adaptable to different types of environments and devices. All the entry operations can be virtually governed by using dynamic/static palm gestures along with a speech assistant. In our work we make use of ML and Computer Perception algorithms to identify palm gestures and speech commands, which works without any additional hardware requirements. The model is implemented using CNN and mediapipe structure. This framework has potential applications like enabling palm-free activity of devices in hazardous environments and providing an substitute platform for hardware pointer. Overall, the palm motion-governed digital pointer framework offers a promising approach to enhance operator experience and enhance accessibility through human-computer interaction.

Keywords: Computer perception, palm motion identification, Media-pipe, digital pointer.

1. INTRODUCTION

The world is full of innovation driven factors in our day to day life. We have so many technologies, throughout the world computer technologies are growing simultaneously. They are utilized to execute various tasks which cannot be performed by humans. In fact they are ruling the human lives because they have a potential to do the tasks which cannot be done by humans. The interaction between human and computer can be done with result apparatus like pointer. The pointer is a apparatus utilized for interacting with a GUI which includes pointing, scrolling and moving etc. The hardware pointer in computers and touchpads in laptops will require a huge amount of time to execute complex tasks, incase we are carrying hardware pointer wherever we go it would be damaged sometimes.

After decades the innovation has made the pointer functionality from wired into the wireless to enhance the functionality and for the easy movements in hassle free manner. As the technologies started growing there came the speech identification technique. This identification is mainly utilized for the speech identification purpose for searching something with the help of their speech and for translation purposes but it can take time for identification to execute pointer

functions. Later the human computer interaction evolved with the eye tracking techniques for controlling the cursor of the pointer. The major drawback of this technique is that some may wear contact lens or some may have long eyelashes so it may take some time to capture their eye motion.

Different types of attempts taken by many developers for developing the models for human motion identification. Those models require expensive gloves and sensors for capturing and color cap for marking the positions of the fingertips. The technologies are still emerging, one of the vast technologies synthetic cognition is playing a major role in every sector. Synthetic cognition makes human life fast and comfortable. To overcome the problems faced in the existing approaches we are going for the latest algorithms and tools in synthetic cognition.

Palm motion governed digital pointer using synthetic cognition is a innovation that allows users to control the motion of their computer pointer using palm gestures, without the advent of a physical pointer. This innovation uses a webcam perception founded approach to monitor the movements of the operator's palm and to execute pointer functions on the computer display. The framework works by capturing video entry from a webcam pointed at the operator's palm. The computer perception algorithms then analyze the video feed to identify the operator's palm and monitor its motion. This information is given to machine learning models which have been trained to identify specific palm gestures, such as pointing or swiping, and translate them into corresponding pointer movements.

This latest super cool innovation has various advantages, including its potential to enhance accessibility for people and its ability to provide a more natural and intuitive operator experience. It can also be useful in situations where a physical pointer or touchpad is not available or practical. The use of palm gestures as a control mechanism eliminates the need for a physical pointer and provides a more intuitive and natural way of interaction with computers. This innovation has numerous applications in areas such as gaming, digital reality and accessibility quite easy for people.

2. LITERATURE REVIEW

Some work which is related to the AI digital pointer had been performed previously in that glove were utilized by the operator to identify and collect data from the framework. Later another framework utilized colored pieces of paper which are attached on hands for motion identification. But these systems are not very feasible for performing pointer operations accurately. In a glove founded approach recognizing the gloves is not viable and it might be allergic for users who have sensitive skin type. Also wearing gloves for a long time is difficult. It might sweat and outcome in skin rashes and allergic reactions. In the case of colored tips for motion identification and detection will not always give best results. Now some others have made contributions that use Google's work with the mediapipe structure. The current motion governed digital pointer uses palm gestures to execute pointer functions, in which we have control over the pointer cursor and execute certain pointer operations like left click, right click, drag and drop, volume control and brightness control etc. Efforts have been made for palm motion identification with webcam-founded detection of the palm motion platform.

[1] This review is about how a hardware-founded framework is devised. Although this model produces incredibly accurate results, many movements are challenging to execute while wearing a glove that severely limits the operator's palm's range of motion, velocity, and agility. Also wearing gloves for a long time will outcome in skin diseases and is not best suited for the users with sensitive skin type. [2] They created a machine-operator platform that uses straightforward computer perception and multimedia techniques to accomplish palm motion detection. However, a significant disadvantage is that skin pixel identification and palm segmentation from stored frames must be completed before working with motion comparison techniques. [3] They described a framework in this study for recognizing palm movements that relies on a mobile phone's webcam and a connected mobile projector as a visual feedback medium. Other mobile

applications can easily link to their structure to learn motion identification. The suggested architecture enables the quick and simple creation of research prototypes that support gestures, diverting the operator's focus away from the apparatus and towards the content.[4] A approach for performing pointer functions without any electrical equipment like sensors. It requires a webcam alone. And pointer functions like clicking and dragging files are carried out through palm gestures. The suggested model performance is low with precision and lacks more pointer functionality. [5] This study focuses on the advanced study of robots with motion controls. The first section gives an idea of the art for palm motion identification as it relates to how they are seen and captured by common video cameras. Founded on estimations of the smoothed optical flow, we extract a collection of motion features. Face detection is utilized to produce a operator-centric representation of this data, and an effective classifier is trained to differentiate. [7] In this model the palm's center is determined, and the palm's calculated radius is discovered. And using the convex hull technique, fingertip points have been determined. The palm motion is utilized to control every pointer motion. And the problem of this approach is the frame must first be saved before being processed for detection, which takes longer than what is needed in real-time. [9] The perception founded technique has been tried out in this framework. Utilized a webcam for motion identification and detection. And no external devices like sensors and gloves were utilized. Completely focuses on leveraging the YOLOv5 procedure and Synthetic Cognition (AI) to identify palm gestures and enhance HCI. [10] The framework can generate coloured masks utilizing techniques for color variation. Later pointer functions are carried out using palm gestures. This approach is difficult in its implementation.

3. ALGORITHMS AND TOOLS Utilized

For the purpose of palm and finger detection we are using the one of the effective open source library mediapipe, it is one type of the structure founded on the cross platform features which was devised by google and Opencv to execute some CV related tasks. This procedure uses machine learning related concepts for detecting the palm motion and to monitor their movements.

3.1Mediapipe

Google created the open-source MediaPipe structure to enable the development of cross-platform, real-time computer perception applications. For processing and analyzing video and audio streams, it offers a number of pre-made tools and components, such as object detection, pose estimation, palm tracking, facial identification, and more.

Developers can quickly construct intricate pipelines using MediaPipe that combine numerous algorithms and processes and execute in real-time on a variety of h/w platforms, like CPUs, GPUs, and specialized accelerators like Google's Edge TPU. Additionally, the structure has interfaces helps us interacting with other well-liked machine learning libraries, including TensorFlow and PyTorch, and it supports several programming languages, like C++, Python, and Java.

For computer perception and ML tasks, MediaPipe is a comprehensive library that offers a many of features. Here are a few of the library's main attributes and features:

1. Video and Audio Processing: MediaPipe provides tools for processing and analyzing video and audio streams in real-time. This includes functionalities such as video decoding, filtering, segmentation, and synchronization.
2. Facial Identification: MediaPipe can detect and monitor facial landmarks, including eyes, nose, mouth, and eyebrows, in real-time. This functionality is useful for applications such as facial identification, emotion detection, and augmented reality.
3. Palm Tracking: MediaPipe can monitor palm movements in real-time, allowing for palm motion identification and interaction with digital objects.

4. Object Detection: MediaPipe can detect and monitor objects in real-time using machine learning models. This functionality is useful for applications such as augmented reality, robotics, and surveillance.
5. Pose Estimation: MediaPipe can estimate the poses of human bodies in real-time, allowing for applications such as fitness tracking, sports analysis, and augmented reality.

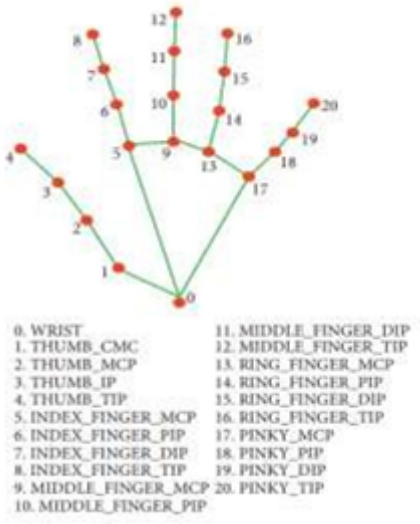

Fig -1: Palm Coordinates or Landmarks

For a variety of tasks, such as object detection, position estimation, facial identification, and more, MediaPipe offers tools for training and deploying machine learning models. All in all, MediaPipe is a potent instrument kit that gives programmers the ability to easily generate sophisticated real-time computer perception and ML applications.

3.2 Opencv

A computer perception and ML software library called OpenCV is available for free download. Its objective is to aid programmers in the development of computer perception applications. Filtering, attribute identification, object identification, tracking, and other processing operations for images and videos are all available through OpenCV. Python, Java, and MATLAB are just a few of the numerous programming languages that it has bindings for. It is written in C++.Robotics, self-driving cars, AR, medical picture analysis, and other fields are just a few of the fields where OpenCV can be employed. A wide range of algorithms and tools are included in the library, making it simple for programmers to build sophisticated computer perception applications.

The steps listed below can be utilized to broadly classify OpenCV's activity:

1. Loading and Preprocessing the Picture/Video: OpenCV can load images or videos from a variety of sources such as files, cameras, or network streams. Once the picture or video is loaded, it can be preprocessed by applying filters or transforming the picture to a different color space, such as converting a color picture to grayscale.
2. Attribute Detection and Description: OpenCV can detect and extract features from an picture or video, such as edges, corners, and blobs. These features can be utilized to identify objects or monitor their motion over time. OpenCV also provides algorithms for describing these features, which can be utilized to match them across multiple frames or images.
3. Object Detection and Identification: OpenCV can be utilized to detect and identify objects in an picture or video. This can be done using a variety of techniques, such as template matching, Haar cascades, or deep learning-founded methods.
4. Tracking: OpenCV can monitor objects in a video stream by estimating their position and

motion over time. This can be done using a variety of algorithms, such as optical flow, mean-shift, or Kalman filtering.

5. Picture and Video Result: Finally, OpenCV can be utilized to display or save the processed images or videos. This can be done by showing the images in a window, writing the video frames to a file, or streaming the video over a network.

In general, OpenCV offers a large variety of tools and techniques for working with picture and video data, making it a potent library for computer perception applications.

4. PROPOSED METHODOLOGY

4.1.The Webcam Utilized in the AI Digital Pointer Framework

The proposed framework uses web webcam for capturing images or video founded on the frames. For capturing we are using CV library Opencv which is belongs to python web webcam will start capturing the video and Opencv will generate a object of video capture. To AI founded digital framework the frames are passed from the captured web webcam.

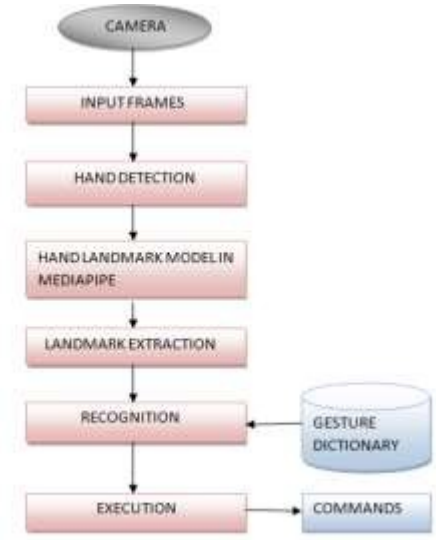

Fig -2: Flow graph of Palm Motion Identification

4.2. Capturing the Video and Processing

The capturing of the frame was done with the AI digital pointer framework until the program termination. Then the video captured has to be processed to find the hands in the frame in each set. The processing takes places is it converts the BRG images into RGB images, which can be performed with the below code,

picture = cv2.cvtColor(cv2.flip(picture, 1),

cv2.COLOR_BGR2RGB) picture. flags. writeable = False

results = hands. procedure(picture)

This code is utilized to flip the picture in the horizontal direction then the resultant picture is converted from the BRG scale to RGB scaled picture.

4.3. Rectangular Region for Moving through the Window

The windows display is marked with the rectangular region for capturing the palm motion to execute pointer action founded on the motion. when the hands are find under those rectangular area the detection begins to detect the action founded on that the pointer cursor functions will be performed. The rectangular region is drawn for the purpose of capturing the palm gestures through the web webcam which are utilized for pointer cursor operations.

Pointer Functions Depending on the Palm Gestures and Palm Tip Detection Using Computer Perception:

- For the Pointer Cursor Moving around the Computer Window.

Fig -3: Computer Window with Pointer Cursor

- To Execute Left Button Click activity

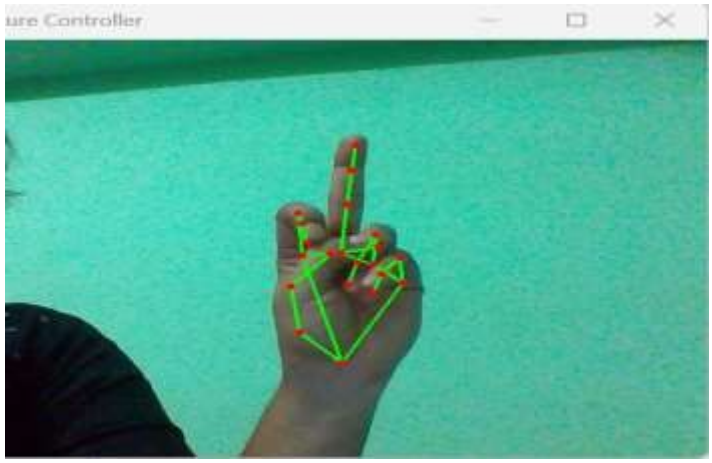

Fig -4: Pointer Activity-Left Click

- To Execute Right Button Click activity

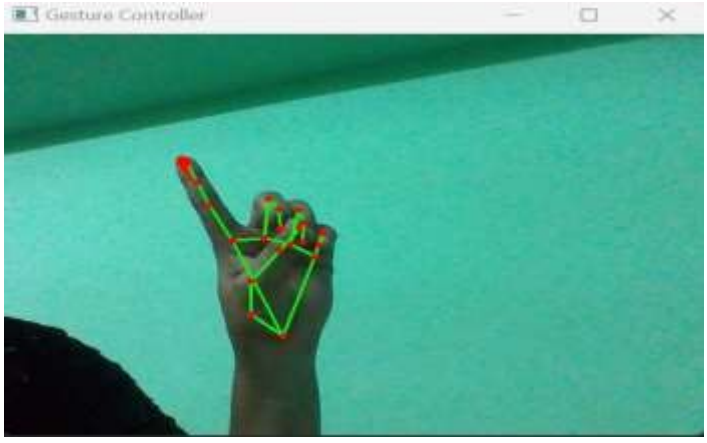

Fig -5: Pointer Activity-Right Click

- To execute a double click activity

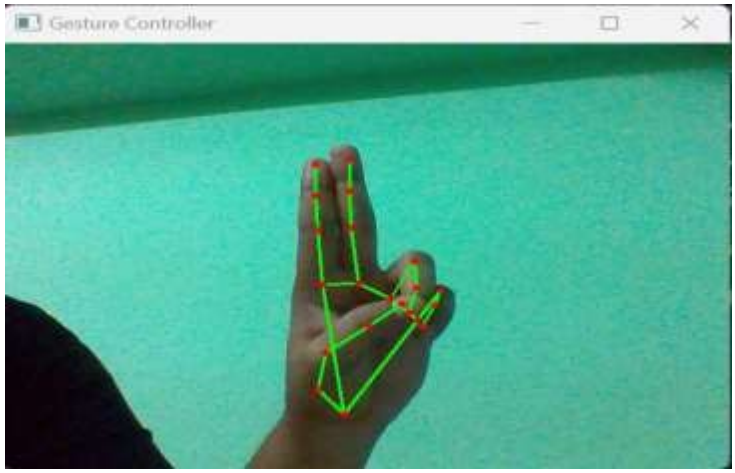

Fig -6: Pointer Activity-Double Click

- To execute scrolling activity

Fig -7: Pointer Activity-Scrolling

- To execute drag and drop activity

Fig -8: Pointer Activity- Drag and Drop

- To execute multiple item selection

Fig -9: Pointer Activity-Multiple Item Selection

- To execute volume controlling

Fig -10: Pointer Activity-Volume Control

5. Outcome AND INFERENCES

A palm-motion- governed digital pointer could provide an substitute approach for people with disabilities who may have difficulty using a conventional pointer or keyboard. This innovation can make it easier for them to interact with computers and other devices. A palm motion-governed digital pointer could also be useful for people who prefer to work or play games without being tethered to a physical pointer touchpad. This model would allow them to control their devices without the need for a physical platform.

Depending on the innovation utilized, a palm motion-governed digital pointer may offer a higher degree of precision and velocity than conventional mice or video editing. The success of this innovation will depend on the operator experience it provides. If the innovation is easy to use, reliable and provides an intuitive platform, likely to be well-received. However, if the innovation is difficult to use, unreliable, or unintuitive , users may quickly abandon it.

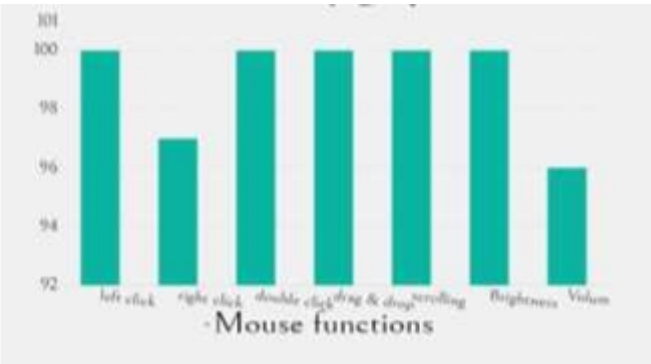

Chart -1: Precision For Pointer Operation

Overall, the palm motion-governed digital pointer using synthetic cognition has the potential to enhance the accessibility of and convenience of computer interaction for users with physical disabilities or for users who prefer an substitute to conventional entry devices.

6. CONCLUSION

AI digital pointer using palm gestures is an innovative and exciting innovation that has the potential to revolutionize the way we interact with computers. Here with the aid of a real-time webcam, we have created a framework to manage the pointer pointer and carry out its operation. It offers users a more natural, intuitive, and accessible way to control the cursor on the display, without the need for a conventional entry apparatus, a pointer.

Furthermore, with additional speech assistant support, AI digital pointer using palm gestures can further enhance the operator experience. Speech assistant which is integrated with the digital pointer framework will provide users with even more control over their devices. Users can given speech commands to do a range of tasks, such as opening applications, navigating through menus, and performing web searches, in addition to controlling the cursor on the display using palm gestures. As innovation continues to evolve, we can expect to see even more innovative solutions that enhance the operator experience and enhance accessibility for all.

7. REFERENCES

1. Quam, D.L., et.al. (1990). Motion Identification with a Dataglove. In IEEE conference on Aerospace and Electronics (pp. 755-760).
2. Guoli Wang., et.al. (2015). Optical Pointer Sensor-Founded Laser Spot Tracking for HCI entry, Proceedings of the Chinese Intelligent Systems Conference (pp. 329-340).
3. Baldauf, M., and Frohlich, p. (2013). Supporting Palm Motion Manipulation of Projected Content with mobile phones. In the European conference on computer perception (pp. 381-390).
4. Roshnee Matlani., Roshan Dadlani., Sharv Dumbre., Shruti Mishra., & Abha Tewari. (2021). Digital Pointer Palm Gestures. In the International Conference on Innovation Advancements and innovations (pp. 340-345).
5. Mayur, Yeshi., Pradeep, Kale., Bhushan, Yeshi., & Vinod Sonawane. (2016). Palm Motion Identification for Human-Computer Interaction. In the international journal of scientific development and research (pp. 9-13).
6. Shriram, S., Nagaraj, B., Jaya, J., Sankar, S., & Ajay, P. (2021). Deep Learning Founded Real-Time AI Digital Pointer Framework Using Computer Perception to Avoid COVID-19 Spread. In the Journal of Healthcare Engineering (pp. 3076-3083).
7. Steven Raj, N., Veeresh Gobbur, S., Praveen., Rahul Patil., & Veerendra Naik. (2020). Implementing Palm Motion Pointer Using OpenCV. In the International Research Journal of Engineering and Innovation (pp. 4257-4261).

8. Sneha, U., Monika, B., & Ashwini, M. (2013). Cursor Control Framework Using Palm Motion Identification. In the International Journal of Advanced Research in Computer and Communication Engineering (pp. 2278-1021).
9. Krishnamoorthi, M., Gowtham, S., Sanjeevi, K., & Revanth Vishnu, R. (2022). Digital pointer using YOLO. In the international conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (pp. 1-7).
10. Varun, K.S., Puneeth, I., & Jacob, T.p. (2019). Digital Pointer Implementation using OpenCV. In the International Conference on Trends in Electronics and Informatics (pp. 435-438).
11. Quek, F., et.al. (1994). Towards a perception founded palm motion platform, in Proceedings of Digital Reality Software and Innovation (pp. 17-31).
12. Tharsanee, R.M., Soundariya, R.s., Kumar, A.S., Karthiga, M., & Sountharrajan, S. (2021). Deep Convolutional neural network-founded picture classification for COVID-19 diagnosis. In Data Science for COVID-19 (pp. 117-145). Academic Press.
13. Newell, A., Yang, K., & Deng, J. (2016, October). Stacked hourglass networks for human pose estimation. In the European conference on computer perception (pp. 483-499). Springer, Cham.
14. Ramakrishna, V., Munoz, D., Hebert, M., Andrew Bagnell, J., & Sheikh, Y. (2014).Pose machines: Articulated pose estimation via inference machines. In the European Conference on Computer Perception (pp. 33-47). Springer, Cham.
15. Tharani, G., Gopikasri, R., Hemapriya R., & Karthiga, M. (2022). Gym Posture Identification and Feedback Generation Using Mediapipe and OpenCV. In International Journal of Advance Research and Innovative Ideas in Education (pp. 2053-2057).
16. Shibly, K.H., Kumar, S., Islam, M.A., & Iftekhar Showrav, S. (2019). Design and Development of Palm Motion Founded Digital Pointer. In the International Conference on Advances in Science, Engineering and Robotics Innovation (pp. 1-5).