iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52881

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Computer Vision based Media Control using Hand Gestures

Sangamesh Narayanpethkar[1], Abhishek Makude[2], Samrudh S[3], Yashraj Khose[4], A. D. Gujar[5]
[1, 2, 3, 4] *Student, Dept of Computer Engineering, TSSM's BSCOER Pune, Maharashtra, INDIA*
[5] *Professor, Dept of Computer Engineering, TSSM's BSCOER Pune, Maharashtra, INDIA*

Abstract: ***Hand gestures are a form of nonverbal communication that can be used in several fields such as communication between deaf-mute people, robot control, human–computer interaction (HCI), home automation and medical applications. At this time and age, working with a computer in some capacity is a common task. In most situations, the keyboard and mouse are the primary input devices. However, there are several problems associated with excessive usage of the same interaction medium, such as health problems brought on by continuous use of input devices, etc. Humans basically communicate using gestures and it is indeed one of the best ways to communicate. Gesture-based real-time gesture recognition systems received great attention in recent years because of their ability to interact with systems efficiently through human-computer interaction. This project implements computer vision and gesture recognition techniques and develops a vision based low-cost input software for controlling the media player through gestures.***

Keywords: Hand Gestures, Media Control, OpenCV, Mediapipe, Human-Computer Interface, Webcam

I. INTRODUCTION

Gesture recognition is about how various human body components, such as the head, arms, face, and fingers, move to interact with the surroundings. The design of human-computer interaction programs depends heavily on gesture recognition because they are effective and trustworthy. Common gesture recognition-based computer applications are, interacting with school children, monitoring the behavior of elderly or disabled individuals, translating sign language, and sluggish driving detection. Now for a few decades, gesture recognition has been a significant topic of research. There are different types of sensors, such as vision sensors and wearable sensors, that have been used to create gesture recognition systems. Regular cameras and the Kinect are two examples of widely used vision-based sensors. Hand gestures offer an encouraging field of research because they facilitate communication and provide a natural means of interaction to be used across a variety of applications.

At the present time, hand gesture recognition systems could be thought of as a more common and practical method of human computer interaction. Hand gestures are a form of body language that can be expressed through the position of the fingers, the center of the palm, and the shape the hand forms. There are two types of hand gestures: static and dynamic. The static gesture, as its name suggests, relates to the stable shape of the hand, whereas the dynamic gesture is made up of a sequence of hand motions like waving. It provides a new method of interacting with the virtual environment. Hand gesture recognition has great value in many applications such as sign language recognition, augmented reality (virtual reality), sign language interpreters for the disabled, and robot control.

II. RELATED WORKS

1) Gaurav Sharma, Manuj Paliwal, (2020)," The paper contributed in media control using hand gestures in which a Dynamic hand gesture is used for controlling VLC media player", IEEE 2020 18[4]: 52-57.

2) Title: Robust the Vision based Hand Gestures Interface for Operating VLC Media Player Authors: Anupam Agrawal , Siddharth Swarup Rautaray In 2010, Anupam Agrawal and Siddharth Swarup Rautaray, "The Vision based Hand Gestures Interface for Operating VLC Media Player Application "program, in that the nearest K neighbor algorithm was used see various touches. Features of VLC media player which were driven by hand gestures including play, as well pause, Full screen, pause, increase volume, and decrease capacity. Lucas Kanade Pyramidical's Optical Flow The algorithm is used to detect hand input video. The algorithm mentioned above detects movement points in the image input. Then the methods of K find a hand centre. By using this facility, the hand is the same noticed. This program uses the database it contains various hand gestures and inputs compared with this image stored and appropriately VLC media player it was controlled. The current application is not very robust recognition phase.

©IJRASET: All Rights are Reserved

3) Stella Nadar, Simran Nazareth, Kevin Paulson, NilambriNarka 'In this paper, Controlling Media Player with Hand Gestures using Convolutional Neural Network and its description and working is described thoroughly. The media control using the convolutional neural network with its advantage and disadvantage is mentioned and described.", IEEE 2021 5[2]:1-6

4) Eldose Joy, SruthyChandran, Chikku George, Abhijith A Sabu, DivyaMadhu. "Gesture Controlled Video Player – A non-tangible approach to develop a video player based on Human Hand Gestures using Convolution Neural Networks". In: International Conference in Intelligent Computing and Control Systems (ICICCS) (2018).

5) Monisha Sampath, PriyadarshiniVelraj, Vaishnavii Raghavendran and M Sumithra "Controlling media player using hand gestures with VLC media player". In this paper it is well defined how we control the media player using hand gestures, how different functionalities are worked.

III. PROBLEM DEFINITION

The main focus is to develop an interface between the system and its environment, so that it can connect with its surroundings and could identify the particular colours, and that colour is used to take input and act as a referral point so that we can interact with the system to perform some simple tasks such as controlling the video which is playing on media player and perform various functions and manipulate it.

Gestures has the potential to solve some major real-world problems like:

A. Distance Communication

It allows the users to communicate with the system even if they are far away from the system. And moreover, if high quality cameras or web cameras are taken in use then they are able to capture the gestures at some adequate distance. Users can control the media player by sitting anywhere in that place simply the gesture should be visible i.e., able to be captured by web cameras.

B. Beneficial for Person with Disability

Today in the world the unfortunate count of disabled people has reached 1 billion. This is nearly 14% of the world's population. Therefore, the gestures play an important role in solving this problem. These people can directly show the gestures and control the media player.

C. Substitute for Keyboard and Mouse

Gesture controls have the capability to change the computer hardware devices for media players like mouse and keyboard. Hence, the cost of computer hardware's can be reduced. As we know, a large portion of e-waste is generated from computer hardware. And to stop it the gesture controls can play a major role.

IV. CHALLENGES IDENTIFIED

Our solution to the problem is to create a system made of touch in which users can interact with a given touch and data will be uploaded to the proposed system to get the output. In the work program, the main problem is identified to get a very clear touch, sometimes it may be because of technical limitations like the quality of the camera was not up to standard. Plus, Camera adjustment is also important to track touch. While tracking touch movements, it was found that adequate light is also an important factor, ultimately the tracking becomes difficult when it lacks adequate light. In actual work, defining what type of object to follow sometimes makes it easier for the system to track it, so the system directly detects the preferred object without blocking the entire screen and assembling it very quickly as well. Other problems faced by the web camera are its lack of required quality, unspecified touch, different technical requirements, system errors, system limitations, memory limit, touch tracking due to insufficient brightness, unseen system touch in-screen input, and too many touches of the same type inside the screen. of tracked inputs, the user is very fast in context to provide inputs and a system that can track all the way in quick succession.

Gesture recognition should be able to recognize patterns, ex. instead of recognizing just an image of an open palm, we could recognize a waving movement and identify it e.g., as a command to close the currently used application.

V. METHODOLOGY

With Touch recognition it helps computers understand a person by its body language. This helps to create a more powerful link between people and machines, rather than just a basic user text base or graphical user interfaces (GUIs). In this project of visual acuity, the movement of the hand is studied by a computer camera.

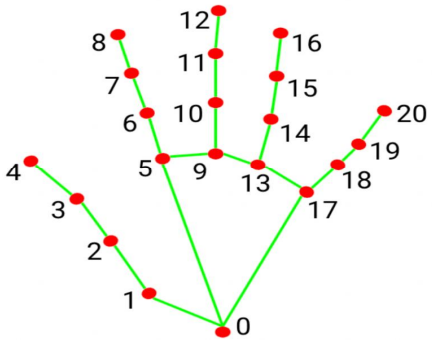
The computer then uses this data as input to handle applications. The aim of this project is to develop a visual connector that will capture the gestures of a human hand and that will control various work. It allows you to control a media player by using hand gestures and provides the user new forms of interaction with the system. They feel natural and do not require an additional device. Furthermore, they do not limit the user to a single point of input, but instead, offer various forms of interaction.

Here when we capture an image it gets converted into RGB. After that we will check whether there are multiple hands in our image. This code has an empty list where we store the list of elements of the hand, which we have detected by using the media pipe i.e. the no of points on the hand.

Now to manipulate the volume we have 2 loops 1 for id and 2 for hand landmark. The landmark information gives us x, y coordinates, and the id number is assigned to various hand points.

Then we get the height and width of the image. Then we find the central point of the image and then we finally draw all landmarks of the hand.

Now for controlling volume we need the index finger and thumb so first as discussed above the list should be empty so when our hand is detected we will assign coordinated to the thumb and index finger. Then we will draw a circle on the tip of the thumb and index finger. Then we will draw a line connecting the thumb and index finger (tip of thumb and index finger). Then we will find the distance between fingers by hypothesis and accordingly increase or decrease sound.



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 Hand Land maker Model

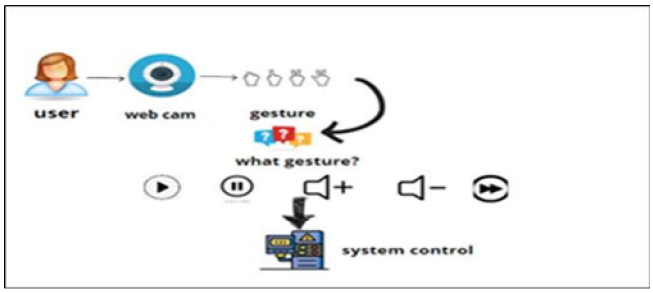


Fig. 2 Architecture

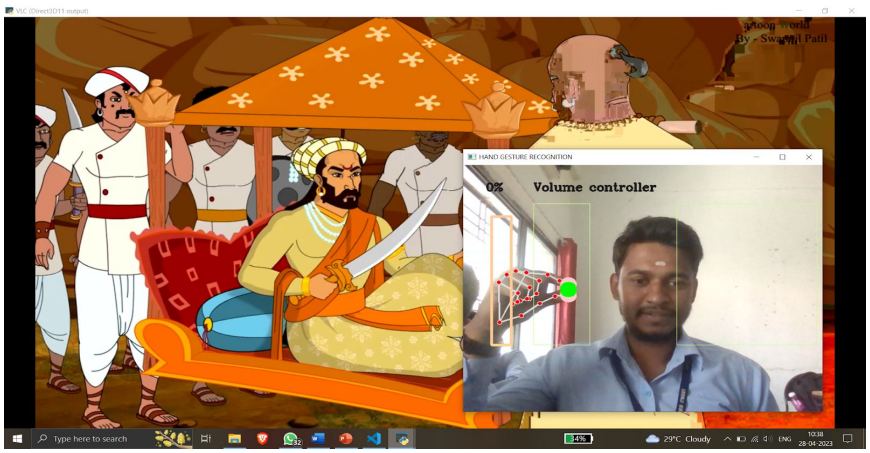


Fig. 3 Decreasing Volume

In this image you can see as the tips of the index finger and thumb get close to each other, the volume decreases. This process happens simultaneously, decrease in length will result in decrease in volume.



Fig. 4 Increasing Volume

In this image you can see as the tips of the index finger and thumb get far from each other, the volume increases. This process happens simultaneously, increase in length will result in increase in volume.



Fig. 5 Pausing the video

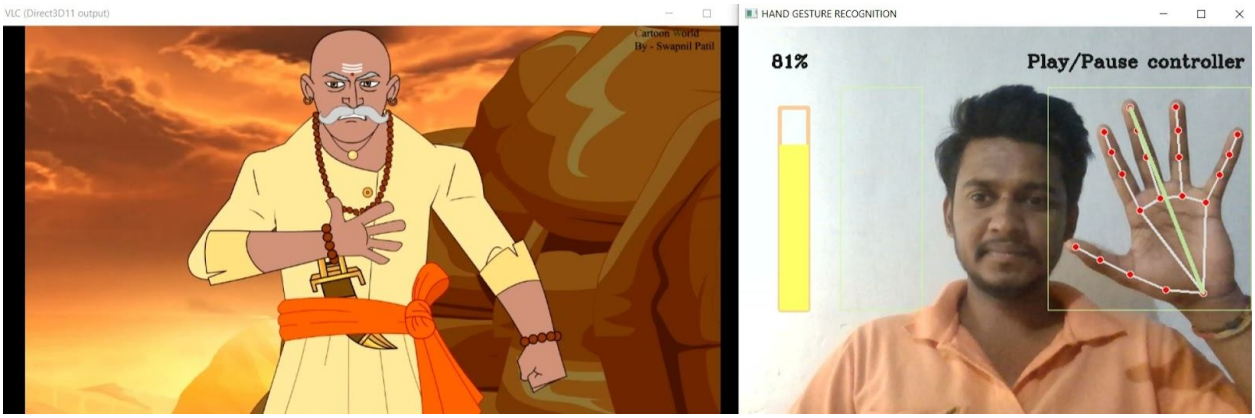


Fig. 6 Playing the video

In the above two images, you can see that when the fist of hand is closed the video is getting paused. And when the fist of hand is opened the video resumes playing.

VI. ALGORITHM OF WORKFLOW

In this part where the user is interacting with the application and controlling the media. It is a far more interesting and fascinating part of the application. In this part the users can control different features of the media player using different gestures.

A. Fingers Recognition

1) In the segmentation image of fingers, the algorithm which is applied is called the labeling algorithm which marks the regions of the fingers. And now in the segmented image or result picture on which labeling method is applied, if the detected regions in which the number of pixels is too small are regarded as noisy regions and discarded. On the other side the regions which are detected and number of pixels are good enough are regarded as fingers and remain.

2) We can check it out using a palm mask, as fingers and the palm can be segmented easily. The part of the hand's part covered with a palm mask is the palm, and the remaining parts are fingers. The features of media are controlled on the basis of length for eg.:

a) The media will be paused, If the length is in the range of 75-90.

b) The media will resume playing, If the length is greater than 150.

3) Same way, using the hypotenuse, we find the distance between the tip of the thumb and the index finger. There is a function called the math hypot function which can be taken in consideration, then we take the value of the difference between x2 and x1 and y2 and y1. And then calculating the length between the index finger and thumb finger the volume is manipulated and changed or controlled dynamically. The volume increases as the length between fingers is increased and vice versa.

VII. ACKNOWLEDGMENT

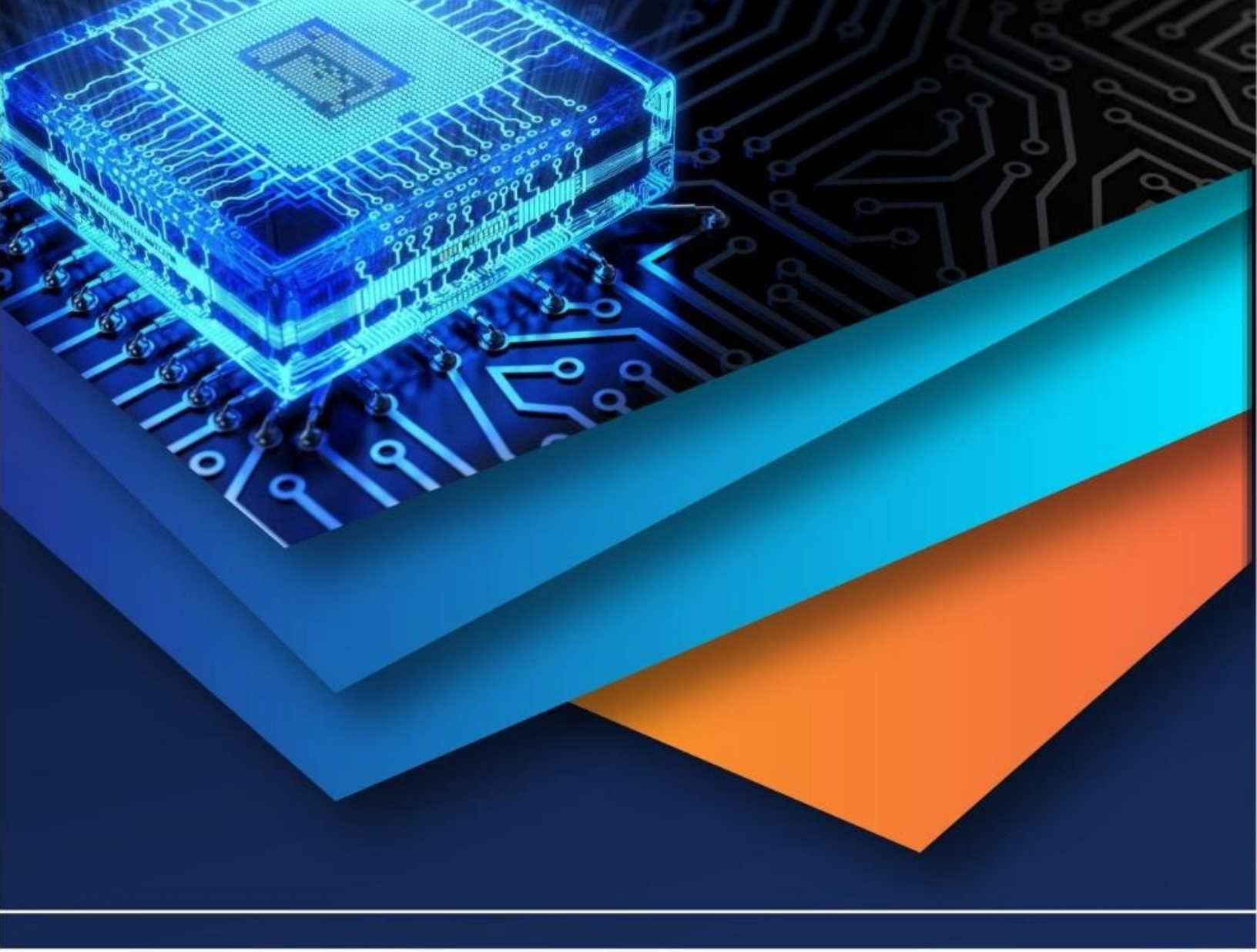
The satisfaction that accompanies the successful completion of any task will be incomplete without the mention of the people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crowns all the efforts with success.
We would also like to give our special thanks to Prof. A D Gujar who monitored our progress and gave his valuable suggestions. We choose this moment to acknowledge his contribution greatly.

VIII. CONCLUSION

In conclusion, hand gestures have become a convenient and intuitive way to control media devices. With advancements in technology, gesture recognition systems have enabled users to interact with their devices in a more seamless and immersive manner. By simply using hand movements and gestures, individuals can navigate through media content, adjust volume, play or pause videos, and perform other control functions without the need for physical buttons or remote controls. Hand gesture-based media control offers several advantages. Firstly, it provides a more natural and intuitive user experience, as humans are accustomed to using their hands for communication and interaction. This eliminates the learning curve associated with traditional input methods and makes media control more accessible to a wider range of users, including those with limited mobility or physical disabilities. Overall, hand gestures provide an exciting avenue for media control, offering a more natural and immersive interaction paradigm. With ongoing advancements in technology and research, we can expect further improvements in gesture recognition systems, leading to more seamless and intuitive media control experiences in the future.

IX. COPYRIGHT

Copyright © 2023 Sangamesh Narayanpethkar, Yashraj Khose, Samruddh S, Abhishek Makude. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)