IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, April 2023

Survey on Vision Based Hand Gesture Interface for Controlling Multimedia Player

Shashank Waghale[1], Sarthak Bankhele[2], Om Badade[3], Prof. Priyanka Raikar[4]
Students, Department Of Computer Engineering[1,2,3]
Professor, Department Of Computer Engineering[4]
Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: *As computers become additional pervasive in society, facilitating natural human–computer interaction (HCI) can have a positive impact on their use. Hence, there has been growing interest within the development of recent approaches and technologies for bridging the human-computer barrier. the last word aim is to bring HCI to a regime wherever interactions with computers are going to be as natural as associate degree interaction between humans, and to the current finish, incorporating gestures in HCI is a crucial analysis space. Gestures have long been thought-about as associate degree interaction technique that may doubtless deliver additional natural, creative, and intuitive strategies for human activity with our computers. Hand gesture recognition is one amongll|one amongst|one in every of} the systemsthat may notice the gesture of the hand in a period of time video. The gesture of hand is classed inside a definite space of interest. during this study, planning hand gesture recognition is one among the difficult jobs that involves 2 major issues. foremost is that the detection of the hand. Another drawback is to form an indication that's appropriate to be used one hand at a time. This project concentrates on however a system might notice, acknowledge and interpret hand gesture recognition through computer vision with the difficult factors that variability within the create, orientation, location, and scale. To perform well for developing this project, differing kinds of gestures like numbers and sign languages got to be created during this system. The image taken from the period of time video is analyzed via Haar-cascade Classifier to notice the gesture of hand before the image process is finished or in different words to notice the looks of hand in a very frame. during this project, the detection of hand are going to be done mistreatment the theories of Region of Interest (ROI) via Python programming. the reason of the results are going to be targeted on the simulation half since the distinction for the hardware implementation is that the ASCII text file to scan the period of time input video. the event of hand gesture recognition mistreatment Python, OpenCV, and YOLO V3 will be enforced by applying the theories of hand segmentation and also the hand detection system that uses the Haar-cascade classifier.*

Keywords: Hand Gesture, OpenCv, Python, Machine Learning

I. INTRODUCTION

With the development in Computer Vision and Human-Machine Interaction, the Computer holds the most important role in our daily life. Human-Computer Interaction can provide several advantages with introducing the different natural forms of device-free communication. Gesture recognition is one of the several types of them to interact with humans Gestures are the natural form of action which we often used in our day-to-day life. But for computer applications to interact humans with the machine, interaction with devices like keyboard, mouse, etc. must be required. As the various hand gestures are frequently used by humans, this project aims to reduce external hardware interaction which is required for computer application, and hence this causes the system more reliable for use with ease. The task of recognising hand gestures is one of the main and important issues in computer vision. With the latest advances in information and media technology, human computer interaction (HCI) systems that involve hand processing tasks such as hand detection and hand gesture recognition. The first step in any hand processing system is to detect and locate the hand in the real-time video from the webcam. The detection of hand is challenging because of variation in pose, orientation, location and scale. Image acquisition involve capturing image in the video frame by frame using a webcam.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, April 2023

The captured images go through the image pre-processing process which involves colour filtering, smoothing and thresholding. After successfully detecting the gesture it is recognised and is used to give commands to the multimedia player. Through this project, we aim to provide a natural device-free interface for controlling a multimedia player using Vision-based hand gestures recognition interface for controlling a multimedia player.

II. LITERATURE SURVEY

Several papers and projects have targeted the issue of hand gesture recognition. Francis et al compared methods for gesture recognition in cars, evaluating accelerometers-based, glove-based, and Kinect-based approaches. Mitra et al analyzed more computationally heavy methods using hidden Markov models and finite state machines. Ghotkar et al presented a novel approach to hand segmentation and gesture recognition using different color spaces. The methods proposed by Francis et al required additional hardware, while those proposed by Mitra et al were computationally heavy, requiring classification and processing time. Our goal was hence to follow the example of Ghotkar et al, and explore the more basic methods of hand segmentation and gesture recognition available, applying them to execute simple controls on a media player.In the Hand gesture recognition paper, simple scheme is presented. Usage of the skin color and labeling algorithm is introduced for hand recognition, it will do segmentation on the hand area from the background. Then the center of mass& the palm point is originated used for the production of the baseline . By which shapes of hand gesture the signature will constructed. As this can be forecast thelabels of class &will be classify in last phase. Using YOLO to recognise hand gesture, A new robust approach in gesture recognition based on skin color detection is used. Use of three different color correction algorithms is done before skin detection and then they are classified into gestures. It tracks the hand(s) location in real- time and recognizes several gestures. Only a webcam is required. It can differentiate if the user is performing one or two hand gestures. Cannot detect grab or swipe gestures. Problems when there is motion blur or hand trembling. These issues are solved by making a model and training it upto a certainaccuracy. To recognise dynamic hand gesture[6], Real time analysis of image streams generated using a camera deployed on a quadcopter or operated out of a ground control station. If camera is deployed on a quadcopter, account for its vibratory and oscillatory motion and then actively analyze the image stream for hand gestures. Send control commands to a quadcopter based on gestures recognized using machine learning with adaptive error control so that the algorithm understands when to execute which control command

III. PROPOSED SYSTEM

Gesture recognition the term collectively refers to the whole process of tracking human gestures to their representation and conversion to semantically meaningful commands. Research in hand gesture recognition aims to design and development of such systems than can identify explicit human gestures as input and process these gesture representations for device control through mapping of commands as output. The exact meaning of object detection, tracking and identification and point out the general problems regarding the object detecting, tracking and identification. To track multiple objects, different methods are in use such methods have shown very good performance, considering more frames before making association decisions should generally help better overcome ambiguities caused by longer-term occlusions and false or missed detections. Many global approaches that use more information have been explored to overcome errors of detections. Most proposed hand gesture systems can be divided into two categories of computer vision techniques. First, a simple approach is to use image processing techniques via OpenCV library and possibly other tools to provide interaction in real time, which considers time consumption because of real-time processing.

IV. SYSTEM ARCHITECTURE

The datasets are in the form of static or dynamic format. The usage of hand gloves also results in calculating the corner tip of gestures wrongly. The system gives better accuracy only when the images are static compared to the motion gestures. You Only Look Once (YOLO), which breaks through the CNN family's tradition and innovates a complete new way of solving the object detection with the most simple and highly efficient way. YOLOv2 has achieved 76.8 mAP at 67 FPS and 78.6 mAP at 40 FPS. It imposes strong spatial constraints on bounding box predictions such as recognising small objects in groups. It still struggles to generalise to objects in new or unusual aspect ratios or

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, April 2023

configurations. It's also not perfect that YOLO's loss function treats errors the same in small bounding boxes vs large boxes. It tracks the hands location in real- time and recognises several gestures. Only a webcam is required. It can differentiate if the user is performing one or two hand gestures. Cannot detect grab or swipe gestures. Problems when there is motion blur or hand trembling

Figure: Existing System Architecture

V. GOALS AND OBJECTIVES

5.1 Goal

The goal of a vision-based hand gesture interface for controlling a multimedia player is to provide a natural and intuitive way for users to interact with multimedia content. Instead of using a mouse, keyboard, or remote control, users can simply use their hands to control the playback, volume, and other features of the multimedia player. This type of interface can be especially useful for people with disabilities or for situations where a traditional input device is not available or convenient to use.

5.2 Objectives

- Accurate and reliable gesture recognition: The system should accurately recognize the user's hand gestures and interpret them as commands for the multimedia player. The recognition should be reliable and consistent across different lighting conditions, hand positions, and users.
- User-friendly interface: The interface should be intuitive and easy to use, even for users who are not familiar with the technology. The system should provide feedback to the user, such as visual or audio cues, to indicate that their gesture has been recognized and the command has been executed.
- Customizability: The system should allow users to customize the gestures and map them to specific commands in the multimedia player. This can help users who have specific preferences or disabilities, and can also improve the overall user experience.
- Compatibility: The system should be compatible with a range of multimedia players and platforms, such as video players, music players, and streaming services. It should also work with different devices, such as laptops, tablets, and smartphones.
- Performance: The system should have low latency and high accuracy to provide a seamless and responsive user experience. It should also be able to handle multiple users and gestures simultaneously, without any lag or delays.

5.3 List of Functionalities

- Play and pause: The user can make a specific hand gesture to play or pause the multimedia content.
- Volume control: The user can raise or lower their hand to increase or decrease the volume of the multimedia content.
- Track selection: The user can swipe their hand left or right to select the previous or next track in a playlist.
- Fast forward and rewind: The user can make a circular motion with their hand to fast forward or rewind the multimedia content.
- Zoom and rotate: For videos or images, the user can make a zoom-in or zoom-out gesture to enlarge or shrink the content. Additionally, they can rotate their hand to rotate the content.
- Mute and unmute: The user can make a specific gesture to mute or unmute the audio of the multimedia content.
- Full-screen mode: The user can make a gesture to enter or exit full-screen mode for videos or images.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, April 2023

- Gesture customization: The user can customize the gestures and map them to specific commands, such as assigning a fist gesture to stop the playback.
- Gesture recognition feedback: The system can provide visual or audio feedback to indicate that the gesture has been recognized and the command has been executed.
- Compatibility with different multimedia players: The system can work with different multimedia players and platforms, such as video players, music players, and streaming services.

VI. METHODOLOGY

- Webcam: When the media player opens, the webcam automatically starts to function and captures live video which then is read by OpenCV frame by frame and fed into the YOLO neural network for object detection.
- Image preprocessing: The proposed architecture begins with the user performing some hand gestures facing the camera. The camera captures live video which is then divided into multiple image frames
- Gesture Recognition System using yolo V3 algorithm: After image pre-processing, the output is fed into a gesture recognizing system that uses the YOLO algorithm. The YOLO algorithm recognizes the gesture in the following phases:
- Bounding Box Prediction: In this Each grid cell predicts a bounding box involving the x, y coordinate and the width and height and the confidence.
- Non-Max Suppression: In this it select one entity (e.g., bounding boxes) out of many overlapping entities. It is used to choose specific selection criteria to arrive at desired result.

After following the given steps the gesture is recognised and count is generated. This count (eg. o,1,2)is assigned to specific activity. When then the user shows open palm than the specific count is generated. Once the count is generated the task assigned with that task is performed. When a user perform specific gesture the unique count is generated, then this count is sent to database to understand what activity is assigned to that count and thus commands media player to perform the desired activity.

The methodology of a vision-based hand gesture interface for controlling a multimedia player involves several steps:

1. Hand gesture recognition: The system uses a camera or sensor to capture images or data of the user's hand gestures. The images or data are then processed using computer vision algorithms to recognize the specific gestures.
2. Gesture mapping: The recognized gestures are then mapped to specific commands for the multimedia player. For example, a wave gesture may be mapped to play/pause, while a circular motion may be mapped to fast-forward.
3. Command execution: The mapped commands are then executed on the multimedia player, either directly through an API or by simulating key presses on a keyboard.
4. Feedback and validation: The system provides feedback to the user to indicate that their gesture has been recognized and the command has been executed. The system also validates the gesture recognition to ensure accuracy and reliability.
5. Customization: The system allows users to customize the gestures and map them to specific commands according to their preferences or needs.
6. Compatibility: The system is designed to work with a range of multimedia players and platforms, and can be adapted to different devices and operating systems.

VII. CONCLUSION

At last, we conclude that A vision-based Hand body language gesture recognition using YOLO V3 generally refers to the process of recognising hand gestures using various processes. Hand gesture recognition using YOLO V3 is a young interdisciplinary field that draws using yolov3 algorithm, in this the gesture is recognised in three phases i.e using yolov3 algorithm, in this the gesture is recognised in three phases i.e bounding box prediction, running conventional network, non maximum suppression. There are different kinds of techniques used for hand gesture recognition.The effectiveness of the proposed method is verified by the recognition confusion matrix. And the proposed method has an

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, April 2023

High recognition accuracy for four custom dynamic hand gestures. This study has found that generally hand gesture recognition is made easy using YOLO V3.This work would be implemented to detect hand gestures and then further recognise and implement to control multimedia player.

REFERENCES

[1]. J. Francis and A. B K, "Significance of Hand Gesture Recognition Systems in Vehicular Automation-A Survey", International Journal of Computer Applications, vol. 99, no. 7, pp.50-55, 2014.

[2]. A. S.Ghotkar and G. K. Kharate, " Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", International Journal of Human Computer Interaction, vol. 3, no. 1, pp. 15-25, 2018.

[3]. Hand Recognition and Gesture Control Using a Laptop Web-camera Zi Xian Justin Yeo Exchange Student with Stanford University 450 Serra Mall, Stanford,CA94305.

[4]. Hand Gesture Recognition: A Review ,Sundus Munir, Shiza Mushtaq, Afrozah Nadeem,SyedaBinishZahra,5 may 2021.

[5]. Hand Gesture Based Quadcopter Control Using Image Processing And Adaptive Machine Learning: Krishna R Dixit, Tarun Verma, Subramanya U S , Dr. Umadevi V, IEEE 2018