SYNOPSIS

INTERACTIVE AI VIRTUAL MOUSE CONTROL SYSTEM

CONTENTS:

- ABSTRACT
- INTRODUCTION
- LITERATURE SURVEY
- EXISTING SYSTEM
- DISADVANTAGES OF EXISTING SYSTEM
- PROBLEM STATEMENT
- PROPOSED SYSTEM
- OBJECTIVES OF PROPOSED SYSTEM
- ADVANTAGES OF PROPOSED SYSTEM
- SOFTWARE AND HARDWARE REQUIREMENTS

ABSTRACT

A hand gesture recognition system provides a natural, innovative and modern way of non-verbal communication. It has a wide area of application in human computer interaction and sign language. The intention of this implementation is to discuss a noval approach of hand guesture recognition based on detection of some shape based features. The setup consists of a single camera to capture the gesture formed by the user and take this as input to the system. A primary goal of gesture recongition is to create a system which can identify specific human gestures and use them to convey information for device control and by implementing real time gestures recognition a user can control a computer by doing a specific gesture in front of a video camera linked to a computer. In this project we will develop a hand gesture volume control system with the help of OpenCV module. Here the system can be operated using hand gestures without using keyboard and mouse. A real-time fingertip-gesture-based interface is still challenging for human–computer inter actions, due to sensor noise, changing light levels, and the complexity of tracking a fingertip across a variety of subjects. Using fingertip tracking as a virtual mouse is a popular method of interacting with computers without a mouse device.

INTRODUCTION

With the development of augmented-reality technology, researchers are working to reduce people's workload while increasing their productivity by studying human–computer interactions (HCI). The Natural User Interface (NUI) of hand-gesture recognition is an important topic in HCI. Hand-gesture-based interfaces allow humans to interact with a computer in the most natural way, typically, by using fingertip movements. Fingertip detection is broadly applied in practical applications, e.g., virtual mice, remote controls, sign-language recognition, or immersive gaming technology.

The most efficient and expressive way of human communication is through hand gesture, which is a universally accepted language. It is pretty much expressive such that the dumb and deaf people could understand it. In this work, real-time hand gesture system is proposed. Experimental setup of the system uses fixed position lowcost web camera high-definition recording feature mounted on the top of monitor of computer or a fixed camera on a laptop, which captures snapshot using Red Green Blue [RGB] colour space from fixed distance. This work is divided into four stages such as image preprocessing, region extraction, feature extraction, feature matching. Recognition and the interpretation of sign language is one of the major issues for the communication with dump and deaf people. In this project an effective hand gesture segmentation technique has been proposed based on the preprocessing, background subtraction and edge detection techniques. Pre-processing is defined as procedure of formulating data for another process.

The main objective of the preprocessing process is to transform the data into a form that can be more effectively and effortlessly processed. In the proposed work, the pre-processing techniques are created on the basis of different types of combinations from the subsequent hand gesture image processing operations such as capturing image, removing noise, background subtraction, and edge detection and these image processing methods are discussed is as follows.

LITERATURE SURVEY

Sixth sense technology is a gesture based wearable interface that links the digital information around us with the physical world and it allows us to use our natural hand gestures to communicate or interact with the digital information. Several approaches have been presented on the concept of virtual mouse with different ideas. The approaches were done in which involved the concept of Image Processing and Image Acquisition. According to the study, the motto is to make a virtual mouse which is mainly useful for saving manual work. The future modification can use complex mouse workings using this simple image processing technique. By the is concept real world is interacting and getting well with the digital world using the concept of this technology known to be as Sixth Sense. Many works are done using Sixth sense technology some even uses JOT interaction with it as the use of RFID tags and image processing for potholes detection to overcome accidents that's a main problem in many parts of the world. Another work that is done in the similar domain is that train autonomous cars using block chain methods for faster and safer experience, the autonomous cars can use a review or ratting system which can help them to stack up which road is safe and shortest, this way a healthy route can be created for the autonomous industry be it cars or other autonomous vehicles.

There are some related works carried out on virtual mouse using hand gesture detection by wearing a glove in the hand and also using colour tips in the hands for gesture recognition, but they are no more accurate in mouse functions. The recognition is not so accurate because of wearing gloves; also, the gloves are also not suited for some users, and in some cases, the recognition is not so accurate because of the failure of detection of colour tips. Some efforts have been made for camera-based detection of the hand gesture interface.

The hand-mouse interface obtains high accuracy using a Kinect sensor; however, the gesture implementation is inconvenient because the user must control the mouse with both hands. Moreover, the work is limited by the resolution of the virtual monitor. This means that the width and height of the virtual screen depend on the skeleton joints provided by Kinect, e.g., the shoulder width and spine position. The hand-motion area is quite narrow for natural gestures.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)

IEEE Xplore Part Number: CFP20J32-ART; ISBN:978-1-7281-5518-0

EXISTING SYSTEM

- **ANN for Gesture Recognition using Accelerometer Data:**

  The authors introduced an Artificial Neural network application used for the classification and gesture recognition. The gesture recognition is done through the Wii remote, this remote will rotate in X, Y, Z directions. To reduce the computational cost and memory consumption the gesture recognition is processed in two levels. In first level User Authentication is done for gesture recognition. Accelerometer- Based gesture recognition method is used. In second level without any kind of signal processing for gesture recognition Fuzzy automata algorithm has been proposed. After recognizing the data of the gestures, the data was normalized and filtered by k-means and Fast Fourier transform algorithm. Using this Dynamic Bayesian Network. The recognition accuracy has increased up to 95%.

- **Combining multiple depth-based descriptors for hand gesture recognition Keyboard:**
  Based on the depth information of the image taken by the depth cameras the authors have introduced a scheme known as novel hand gesture recognition scheme. To properly recognize complex gestures by using 3-D information they used a set of 3-Dimensional features. The proposed hand gesture recognition system consists of three main steps. The first step based on colour and depth information the hand samples are segmented from the background. Wrist samples, palm and the fingers are subparts of the segmented hand samples. The proposed hand gesture recognition consists of four types of features. The second step is to extract these features for the segmentation.

DISADVANTAGE OF EXISTING SYSTEM

- Special hardware required.
- Has Specific hardware’s like mouse & keyboards for control of system is required.
- Need more desk space compare with virtual mouse.
- Using of gloves & hand tips in the hand for gesture recognition, mouse function is not accurate.
- Tip gesture of the hand is not efficient for detecting the gestures for the moment of mouse.

PROBLEM STATEMENT

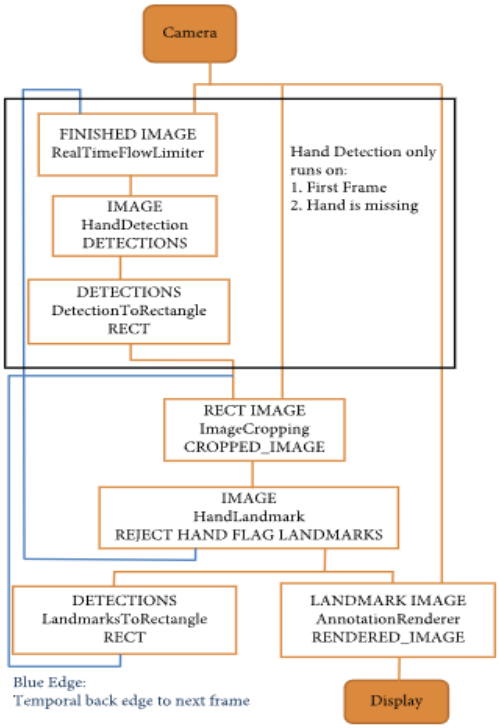
The proposed AI virtual mouse system can be used to overcome problems in the real world such as situations where there is no space to use a physical mouse and also for the persons who have problems in their hands and are not able to control a physical mouse. Also, amidst of the COVID-19 situation, it is not safe to use the devices by touching them because it may result in a possible situation of spread of the virus by touching the devices, so the proposed AI virtual mouse can be used to overcome these problems since hand gesture and hand Tip detection is used to control the PC mouse functions by using a webcam or a built-in camera.

Probably the most well-known problem in computer vision it consists of classifying an image into one of many different categories one of the most popular data set academia is ImageNet. Recent years classification models have surpassed human performance & it has been considered practically solved in contrast with problem like classification the output object detecting in variable in length since the number of objects detecting may change from image to image.

PROPOSED SYSTEM

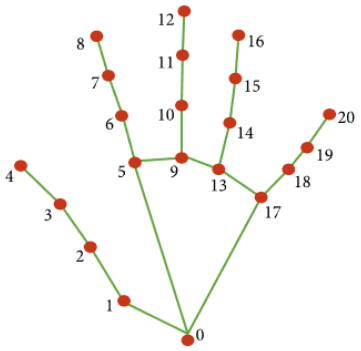
For the purpose of detection of hand gestures and hand tracking, the MediaPipe framework is used, and OpenCV library is used for computer vision. The algorithm makes use of the machine learning concepts to track and recognize the hand gestures and hand tip.

MediaPipe is a framework which is used for applying in a machine learning pipeline, and it is an opensource framework of Google. The MediaPipe framework is useful for cross platform development since the framework is built using the time series data. The MediaPipe framework is multimodal, where this framework can be applied to various audios and videos. The MediaPipe framework is used by the developer for building and analyzing the systems through graphs, and it also been used for developing the systems for the application purpose. The steps involved in the system that uses MediaPipe are carried out in the pipeline configuration. The pipeline created can run in various platforms allowing scalability in mobile and desktops. The MediaPipe framework is based on three fundamental parts; they are



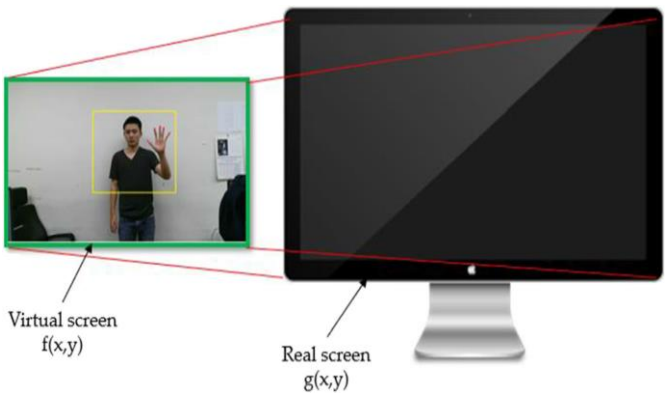
performance evaluation, framework for retrieving sensor data, and a collection of components which are called calculators, and they are reusable. A pipeline is a graph which consists of components called calculators, where each calculator is connected by streams in which the packets of data flow through. Developers are able to replace or define custom calculators anywhere in the graph creating their own application.

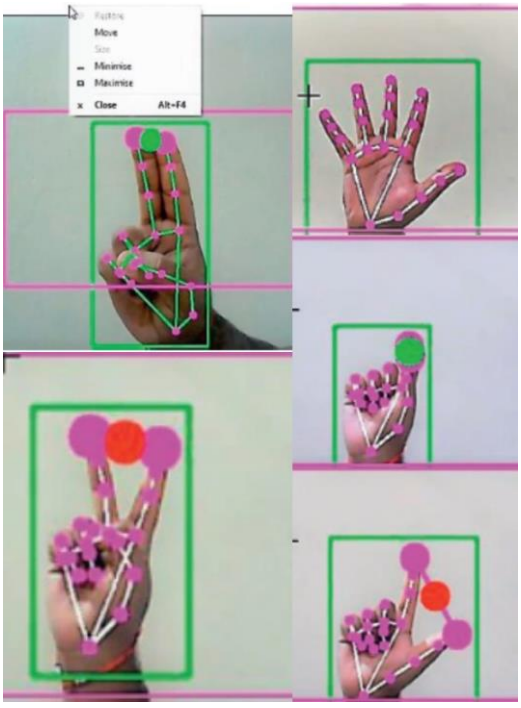
The calculators and streams combined create a data-flow diagram; the graph is created with Single-shot detector model is used for detecting and recognizing a hand or palm in real time. The single-shot detector model is used by the MediaPipe. First, in the hand detection module, it is first trained for a palm detection model because it is easier to train palms. Furthermore, the no maximum suppression works significantly better on small objects such as palms or fists. A model of hand landmark consists of locating 21 joint or knuckle co-ordinates in the hand region.



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

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. $e OpenCV library is used in image and video processing and also analysis such as face detection and object detection.





ADVANTAGE OF PROPOSED SYSTEM

- The proposed model has a greater accuracy of 99% which is far greater than the that of other proposed models for virtual mouse, and it has many applications.
- Amidst the COVID-19 situation, it is not safe to use the devices by touching them because it may result in a possible situation of spread of the virus by touching the devices, so the proposed AI virtual mouse can be used to control the PC mouse functions without using the physical mouse.
- The system can be used to control robots and automation systems without the usage of devices(iv)2D and 3D images can be drawn using the AI virtual system using the hand gestures.
- Persons with problems in their hands can use this system to control the mouse functions in the computer.
- In the field of robotics, the proposed system like HCI can be used for controlling robots.

SOFTWARE AND HARDWARE REQUIREMENT SPECIFICATION

- **HARDWARE REQUIREMENT:**

- Intel Pentium D processor 1.8 GHz or AMD Athlon X2 processor 1.8GHz or higher
- 3 GB RAM
- Peripheral webcam at least 30 frames/second, 640x480 resolution

- **SOFTWARE REQUIREMENT:**

- Windows 7 or higher
- .NET framework 3.5 or higher
- PYCHARM using Anaconda Navigator
- Open CV (Computer Vision)
- NumPy Lib

PROPOSED SYSTEM DESIGN

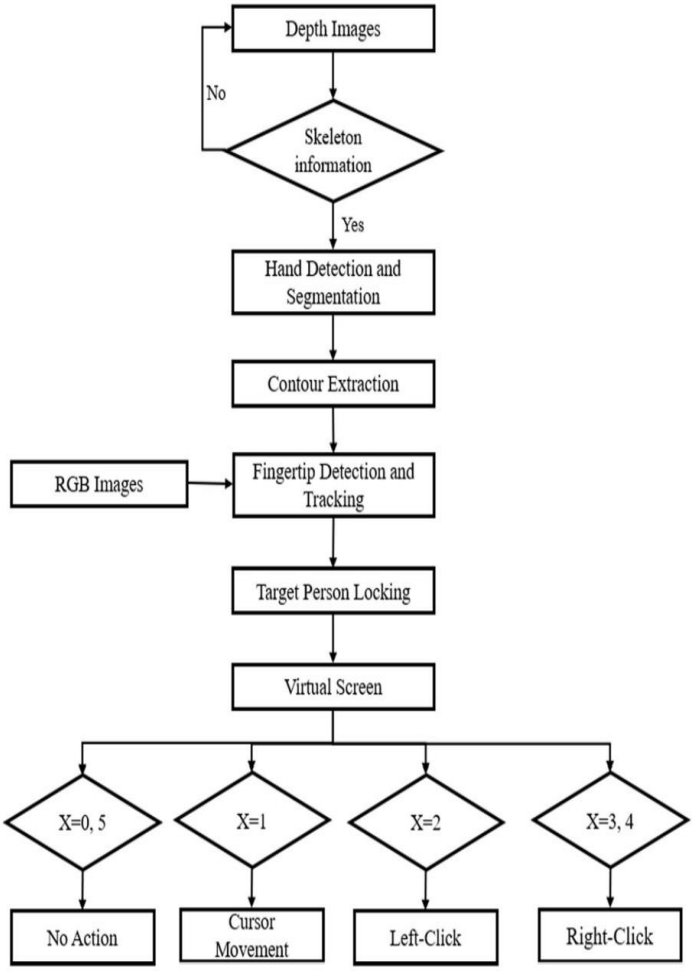


Fig. 1 Flowchart of the proposed method

Hand detection and segmentation:

The depth images used to detect the hand. These images were captured by a Microsoft Kinect V2 sensor, which estimates ach of the user's body parts through input depth images, and maps the learned body parts to the depth images through various user actions. In this manner, the camera can obtain the skeleton-joint information for 25 joints, e.g., hip, spine, head, shoulder, hand, foot, and thumb. Using the depth image of a Kinect skeletal tracker, the hand region of interest (HRI) and the centre of the palm are easily and effectively extracted.

Hand-contour extraction

Hand contours are the curve of the outmost points extracted from the hand-segmentation image. In the fingertip-detection process, contour extraction is a very important step to define the fingertip locations. In this step, the hand contours are detected using the Moore-Neighbour algorithm. This method is one of the most common algorithms used to extract the contours of objects (regions) from an image. After the binary images of the hand regions are detected, the algorithm can find the regional borders by scanning all of the pixels in the images.

Fingertip detection and tracking

After extracting the hand contour, the K-cosine Corner Detection algorithm computes the fingertip points using the coordinates of the detected hand contour. This is a well-known algorithm used for detecting the shapes of certain objects and also in fingertip detection.

Virtual screen matching

The virtual monitor is defined as a virtual space between a Kinect device and a user where a mouse cursor can be controlled by the hands. The advantage of this idea is that it can be implemented on different screen sizes and resolutions. The users only need to watch the virtual screen to control the gestures.

Virtual mouse

A computer mouse is a hand-held pointing device that is most often used to manipulate objects on a computer screen. This paper presents a method that allows the user to control the mouse using their fingertip without a mouse device.

CONCLUSION

The main objective of the AI virtual mouse system is to control the mouse cursor functions by using the hand gestures instead of using a physical mouse. The proposed system can be achieved by using a webcam or a built-in camera which detects the hand gestures and hand tip and processes these frames to perform the particular mouse functions. From the results of the model, we can come to a conclusion that the proposed AI virtual mouse system has performed very well and has a greater accuracy compared to the existing models and also the model overcomes most of the limitations of the existing systems. Since the proposed model has greater accuracy, the AI virtual mouse can be used for real-world applications, and also, it can be used to reduce the spread of COVID-19, since the proposed mouse system.

The project presented a program that allowed user to perform hand gestures for easy software control. A vision-based hand Gesture system that does not require any special markers or gloves and can operate in real-time on a commodity PC with low-cost cameras. Specifically, the system can track the tip positions of the counters and index finger for each hand. The motivation for this hand Gesture was a desktop-based volume control. system in which a user can control volume and cursor navigation in realtime using natural hand motions. Besides, we propose to employ the motion of the mouse cursor controlled by the hand, and give a suggestion about how to, on the bare hand, position a point through which to control the movement of the mouse cursor. For the sake of reliability, we, furthermore, propose a simple probabilistic model to effectively prevent the developed system from responding to invalid gestures.

REFERENCES

1. Akira Utsumi, TsutoniuMiyasato, Fumio KishinoandRyoheiNakatsu, “Real-time Hand Gesture RecognitionSystem,” Proc. of ACCV ’95, vol. 11, pp. 249-253, Singapore,1995

2. Attila Licsár, TamásSzirányi University of Veszprem, “Dynamic Training of Hand Gesture Recognition System” Department of Image Processing and Neurocomputing, H8200 Veszpré,23-26 Aug. 2004

3. L. Bretzner and T. Lindeberg, “Relative orientation from extended sequences of sparse point and line correspondences using the affine trifocal tensor,” in Proc. 5th Eur. Conf. Computer Vision, Berlin, Germany, June 1998, vol. 1406, Lecture Notes in Computer Science, pp.141–157, Springer Verlag

4. Intel Corp, “OpenCV Wiki,” OpenCV Library [Online], Available: http://opencv.willowgarage.com/wiki/

5. Z. Zhang, Y. Wu, Y. Shan, S. Shafer. Visual panel: Virtual mouse keyboard and 3d controller with an ordinary piece of paper. In Proceedings of Perceptual User Interfaces, 2001

6. W. T. Freeman and M. Roth, Orientation histograms for hand gesture recognition. International workshop on automatic face and gesture recognition. 1995, 12: 296-301.

7. G. R. S. Murthy, R. S. Jadon. (2009). “A Review of Vision Based Hand Gestures Recognition,” International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405410.8. Mokhtar M. Hasan, Pramoud K. Misra, (2011). “Brightness Factor Matching for Gesture Recognition System Using Scaled Normalization”, International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2)

8. Fossati A, Gall J, Grabner H, et al (2012) Consumer depth cameras for computer vision: research topics and applications. Springer Science & Business Media

9. Ge L, Liang H, Yuan J, Thalmann D (2018) Robust 3D hand pose estimation from single depth images using multi-view CNNs. IEEE Trans Image Process 27:4422–4436

10. Ge L, Liang H, Yuan J, Thalmann D (2018) Real-time 3D hand pose estimation with 3D convolutional neural networks. IEEE Trans Pattern Anal Mach Intell