Survey on Discrete Gesture Recognition Techniques

Akanksha Raut
akanksharaut62@gmail.com
St. Vincent Palloti College of Engineering and Technology, Nagpur, Maharashtra

Pallavi Wankhede
pwankhede@stvincentngp.edu.in
St. Vincent Palloti College of Engineering and Technology, Nagpur, Maharashtra

Hrishikesh Dhole
hrishidh@gmail.com
St. Vincent Palloti College of Engineering and Technology, Nagpur, Maharashtra

Nidhi Dontulwar
dontulwarn@gmail.com
St. Vincent Palloti College of Engineering and Technology, Nagpur, Maharashtra

Himanshu Bhole
himanshubhole.30@gmail.com
St. Vincent Palloti College of Engineering and Technology, Nagpur, Maharashtra

ABSTRACT

In present circumstances, for Human- Computer Interaction the hand gesture technology is a more habitual and usable system. Programmed hand gesture recognition framework gives us another strategy to be interconnected with the virtual environment. Humans use genuine physical controllers like mouse, keyboard to interlink with the computer which forestalls regular interface due to solid hindrance. To perceive the hand gesture and act according to the user commands to control the application is the sole purpose of hand gestures. This paper explores late strategies utilized in vision based Human Computer Interaction. All the approaches discussed in this paper are calculated and classified based on the type of work done, applied algorithms and proofs and also resources utilized.

Keywords: Human Computer Interaction, hand gestures, Face detection, Image Processing, Feature Extraction.

1. INTRODUCTION

Computers are basic for correspondence and are the highlight of data innovation. Human-beings cooperate with computers from numerous points of view which implies that having a decent interface that works with that communication is critical for our everyday exercises. Our intercommunication with intelligent devices today is not simply circumscribed to console and mouse. The way we communicate with each other comes from distinct ways of interaction like giving signals from face, body and speech. Having the option to reach out to the framework normally is turning out to be ever more convincing in numerous fields of Human Computer Interaction.

Inadequately planned human-machine interfaces can prompt numerous unforeseen issues. Therefore, it is very important to have an efficient, user friendly yet natural mode of interaction to create an authentic environment. At the point when computers were previously introduced into our lives, we figured out how to connect by utilizing the console, mouse and work area. Microsoft Excel for calculations, word for document, presentation and virtual desktop through Graphic user interface. However, GUIs have changed the method of communication with the users adding to the benefit that it makes our experience delightful.

Contrasted with many existing interfaces, hand gestures enjoy the benefits of being not difficult to utilize, characteristic, and instinctive. Fruitful utilizations of hand gestures acknowledgment incorporate laptop games control, Media players, Power-point presentation, human-robot cooperation, and gesture-based communication acknowledgment, to give some examples.

The paper is organized as follows: Section II focuses on different methods which are researched for the gesture recognition system. Section III introduces the literature survey table which gives a comparison of all the approaches used and also mentioning the advantages and disadvantages of each of them. Lastly, Section IV is the conclusion and also throws light on the future scope and further studies which can be carried out for the mentioned approaches.

2. DIFFERENT APPROACHES

2.1 Voila jones for Face Detection:

The Viola Jones algorithm is an object detection algorithm. It was impelled first and foremost as a result of the issues looked at during the face discovery measure. Face identification is the initial step of the face acknowledgment measure. Disregarding being an old-fashioned system, it is strong, and its utilization has demonstrated to be entirely prominent and continuously face acknowledgment[2]. The algorithm looks at many sub-locales and afterward attempts to discover the face by seeing explicit highlights around there. It requires taking a look at a wide range of positions and scales on the grounds that a picture can contain faces of different sizes. This can be executed utilizing OpenCV in Python[3]. The main steps in this algorithm are:

1. Selecting Haar-like features: Human faces possess some identical properties. These identical properties can be coordinated using Haar features. Examples of similar properties are:
• Nose region is brighter than eyes
• Eye region is darker than cheeks
There are three varieties of Haar features. They are edge , line , and four rectangle features. The Haar feature selection takes place accordingly.

2. Creating an integral image: An integral image analyses rectangular features in a given time which gives a good speed advantage. It calculates the sum of pixel values of an image or some part of an image. It's a very effective way of doing calculations and it also reduces the time of computation.

3. Running AdaBoost training: It is also known as Adaptive Boosting. It is a kind of algorithm which is used for boosting up the overall performance. A classifier is created and also a set of images which contains both positive and negative images is prepared. Then, the classifier with minimum error is finalized using mathematical calculations.

4. Creating classifier cascades: A cascade classifier is a classifier with a number of stages which can perform detection more accurately. This process is performed with a view that during the runtime, more and more negative images are rejected and the efficiency can be increased during the face detection.

2.2 Vision Based Approach

Human motion received by one or more cameras and vision devices can process many attributes to interpret gestures (such as color and texture), while sensors cannot. Although these methods are simple, many problems may arise, such as: Various lights, complex backgrounds and existing objects, the skin color is similar to the skin color of the hand (interference), and the system also needs some standards, Such as recognition time, speed, durability and computational efficiency[5].

Figure 1. Vision based approach

The Vision Based Approach consists of 3 modules for the feature extraction:

Camera Module:
This module is liable for associating and catching contribution through the various kinds of picture indicators furthermore, sends this picture to the identification module for preparing as edges. The generally utilized techniques for catching information are information gloves, hand belts and cameras. In our framework, to make the project more cost efficient, an inbuilt webcam was used. The framework has reasonable arrangement to permit contribution from a USB based webcam also yet this would require some consumption from the client. The picture outlines are in the type of a video[1].

Detection module:
This module is liable for the picture handling. The yield from the camera module is exposed to various picture handling procedures, for example, shading change, commotion evacuation, thresholding following which the picture goes through form extraction. In the event that the picture contains absconds, convexity surrenders are found by which the signal is identified. Assuming there are no imperfections, the picture is ordered utilizing Haar cascade to distinguish the signal. On account of dynamic signals, the location module does the accompanying; If Microsoft PowerPoint has been dispatched with a slideshow being empowered and the webcam recognizes palm in development, for 5 consistent edges at that point the unique signal swipe is distinguished[1].

Interface module:
The work of the interface module is to outline the gestures to their activities respectively. These activities are then passed to the fitting application. The front end comprises three windows. The primary window comprises the video input that is caught from the camera with the comparing name of the motion identified[1]. The second window shows the forms found inside the info pictures.

The third window shows the smooth threshold adaptation of the picture. The upside of adding the edge and shape window as a piece of the Graphical User Interface is to make the client mindful of the foundation irregularities that would influence the contribution to the framework and in this manner, they can change their PC or work area web camera to keep away from them. This would result in better execution.

2.3 Data Glove Approach

A variety of sensors are used in this approach to seize the arrangement and movements of the user's hand. The coordinates of the fingers and hand's composition will be easily computed. The sensors don't succeed in a simple reference to the PC, they hinder the actions of the user's hand as it has to be the end user in contact with the computer physically. They are not pocket friendly, and unsuitable to control in a setting of virtual reality. The sensors will become more economical and minor in few years, if Moore's law adhere accurate [5].

Figure 2. Data glove approach

2.4 Coloured Marker Approach

In this the user is supposed to wear the marked gloves and these gloves are dyed to benefit in course to single out fingers and palms. The geometric features of the hand will be derived by these gloves. The diagram shows 3 different dyed gloves for serving palms and different fingers. If we compare to data gloves this approach is quiet pocket friendly, however the innate Human Computer Interaction was still limited[5].
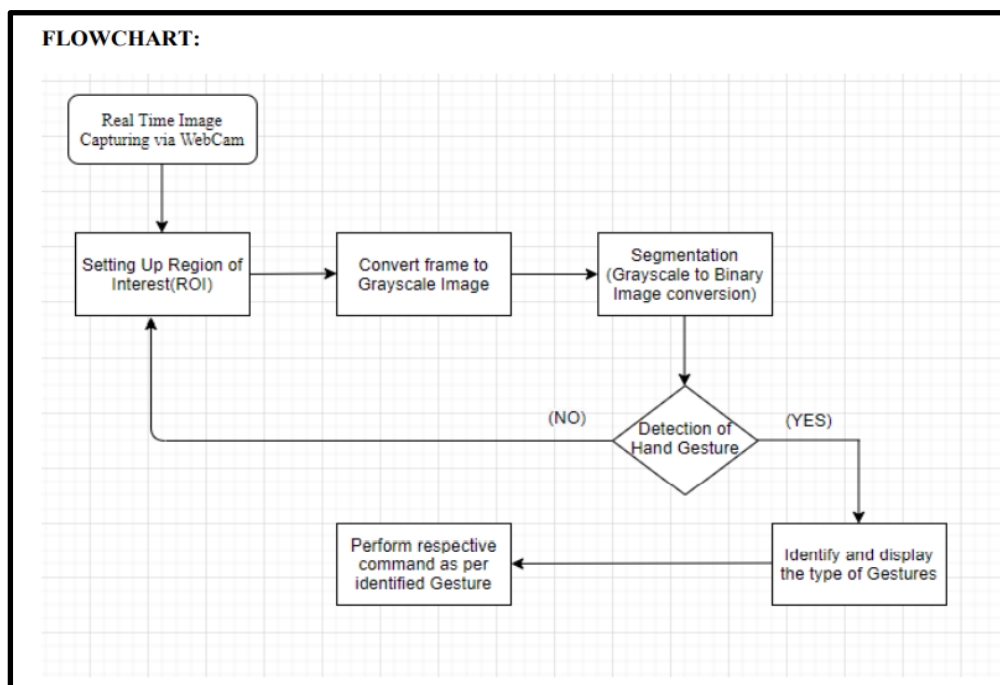
Figure 3. Coloured Marker Approach

FLOWCHART:

Figure 4. Illustration for the overall process.

Following are steps for any system as defined above to recognize the Hand Gesture:

1. Real time image capturing
To capture the real time image of a hand or any object external or internal webcam is used to capture the RGB image of the object or hand.

2. Image Pre-processing
Considering the reduction of computing time, only an area selected from the entire video sequence is used in this step. This is called the Region of Interest (ROI). The preferred method for image processing is to convert the images to color. If you want to increase the rendering in grayscale images, you can restore the images to their original color space after processing is complete. Then we convert the area of interest to a grayscale image. Then we blur the region of interest using Gaussian blur, this will cut down the high frequency object except the object.

3. Feature Extraction
The feature extraction takes place after the pre-processing phase in the character recognition system. The main task of pattern recognition is to take an input pattern and correctly assign it as one of the possible output classes. This process can be broken down into two general phases: feature selection and classification.

4. Segmentation
Segmentation is used to convert gray scale image into binary image so solely 2 objects in image that one is hand And second is background. This formula is employed for segmentation purpose and gray scale pictures are reborn into binary image consisting hand or background. An awfully smart segmentation is required to pick an adequate threshold of gray level to extract the hand from the background. Let's say there's no part of the hand that ought to have a background that should be clear.

5. Convex and Contour Hull Extraction
The contour is nothing but a boundary or contour of the object or hand that is in the image, the same coloured values that are connected points are known as the contour and are very important in the analysis of shapes, objects. Recognition.The set of leads around the target area is the convex hull; we must clarify the convex principle of the set in the envelope is absolutely under cover.

6. Haar Cascade Classifier
With Haar Cascades you can recognize any type of object as long as you have the correct XML file for it. You can even create your own XML files from scratch to recognize any type of object you want.

7. Features Extraction and Recognition
In this step we decide which is the way we decide the quantity of fingers. Next we perceive the hand gesture through the number of fingers we know for which we use Convex Hull to find the limit focus.The Convex Hull is the line totally encasing a bunch of focuses in a plane so that there are no concavities in the line. The principle of Convex Set is concentrated on the fact that lines amidst any two focuses inside the body are altogether inside it. The center of the palm and the radius is calculated using the Euclidean Distance. Using the threshold image and the bitwise AND operator to enumerate the estimated count of fingers. Subsequent to playing out the necessary guidance from the gesture, get back to the initial step to take another picture to be handled, etc.

2.5 The Arc Of Reference Approach:
The first step to the arc of reference approach is to take input for the system in the form of video. After scanning, background subtraction is to be performed. Develop an arc of reference and fingers which cover the arc will be known. So now we can easily straighten out the authorize the static gesture[6]. Main gesture recognition process: Look for authentic gestures. If the gesture is authentic, execute the accredited gesture else reset.

1. Background subtraction:
Background subtraction is an extensive preparatory step in computer vision operations[6]. Background subtraction figures the forefront cover playing out a deduction and keeping only the static parts of the image all that can be considered as foundation given the attributes of the noticed scene.
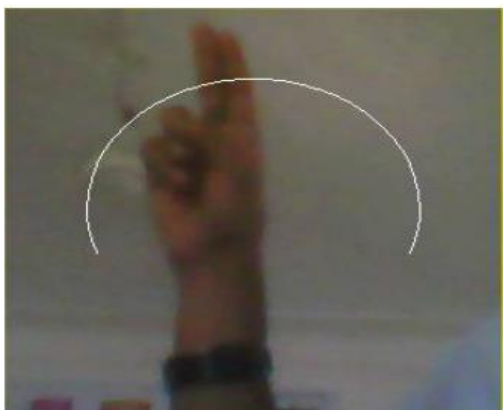

Figure 5. Input Image[5]   **Figure 6. Image after Background subtraction[5]**

2. Arc of reference[5]:

A white coloured arc is used to explore the number of fingers. This turns out to be the most smooth and agile technique to find out the number of fingers and to perform tasks accordingly. In this project to control the VLC media player we assign a single finger for the Play or Pause video stream. When fingers intersect the arc then a comparison is done between the pixel colours of arc and skin. Because of this the work to scan and compute all the pixels in the image is reduced. Gradually resulting in increased processing speed. Simply by scanning pixels in arc we can identify the number of fingers[6].

Table 1: Literature Survey Table

Sr. No.	Paper Name	Year of Public	Proposed Method	Advantages	Disadvantages
1.	Hand Gesture Recognition for Human Computer Interaction[1]	2017	● Camera Module ● Detection Module ● Interface Module	● Satisfactory recognition percentage of the gestures. ● Real time Image processing	Accuracy decreases with non-plain background
2.	Survey on Vision based Hand Gesture Recognition[2]	2019	● Voila Jones algorithm to detect Hand and Face	● Precise and rapid ● Well organized feature selection ● Detection rate is high	Not very effective in detecting inclined or bend face
3.	Real-time human-computer interaction based on face and hand gesture recognition [3]	2014	● Voila and Jones for Face detection. ● Tracking and measuring Threshold	● The gestures are more intuitive and user friendly.	May cause Image Processing errors sometimes.
4.	Hand gesture recognition on python and opencv [4]	2020	● Hand Segmentation method ● Haar cascade Classifier	● Quick ● vigorous and precise strategy for hand signals acknowledgment under unconstrained scenes	Lighting environment changes are still influential in the segmentation process, in particular the process of background removal.
5.	A new hand gestures recognition system[5]	2020	● Vision Based ● Coloured Marker ● Data glove	● Strong method for Gesture recognition ● Process faster and easier ● Reliability	Easily Affected by complicated Background
6	Operating an Application Using Hand Gesture Recognition System[6]	2016	● Background Subtraction ● Arc of reference	● Processing all pixel need is reduced ● Quick calculation to recognize number of fingers crossing	Sometimes detection may require more time.

3. CONCLUSION AND FUTURE SCOPE

In this paper work we conducted the study on different approaches where we can deduce some conclusions like the gesture recognition system for the camera, detection and the interface module did not use any markers making the system user friendly and cheap but accuracy of the system needs to be worked on also to benefit more gestures can be added. The Data glove approach is expensive but a very strong method for gesture recognition. The coloured marker approach is definitely a cheaper option but the drawback is that it requires you to constantly wear the coloured gloves for detection. Further research studies are going on this topic which will focus on using both the hands and increasing number of gestures. The other methods also give promising results to some extent but are affected by the complicated background, which is an advantage in the Arc of reference method as it includes background subtraction. All the approaches and the applications can be improved to a work with immense detachment so it uses data portrayal in a helpful and coherent field.

4. REFERENCES

[1] Aashni Hariaa , Archanasri Subramaniana , Nivedhitha Asokkumara , Shristi Poddara , Jyothi S Nayaka,. "Hand Gesture Recognition for Human Computer Interaction". 7th International Conference on Advances in Computing & Communications, ICACC-2017, 22- 24 August 2017, Cochin, India.

[2] Pranit Shah, Krishna Pandya, Jay Gandhi. "Survey On Vision Based Hand Gesture Recognition". INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING · May 2019

[3] Reza Azad, Babak Azad, Nabil Belhaj Khalifa, Shahram Jamali. "REAL-TIME HUMAN-COMPUTER INTERACTION BASED ON FACE AND HAND GESTURE RECOGNITION". International Journal in Foundations of Computer Science & Technology (IJFCST), Vol.4, No.4, July 2014

[4] Ahmad Puad Ismail , Farah Athirah Abd Aziz , Nazirah Mohamat Kasim and Kamarulazhar Daud. "Hand gesture recognition on python and opencv". IOP Conference Series: Materials Science and Engineering, Volume 1045, First International Conference on Electrical Energy and Power Engineering (ICEEPE 2020) 25th-26th August 2020, Penang, Malaysia

[5] Ahmed Kadem Hamed AlSaedi , Abbas H. Hassin AlAsadi. "A new hand gestures recognition system". Indonesian Journal of Electrical Engineering and Computer Science Vol. 18, No. 1, April 2020, pp. 49~55 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v18.i1.pp49-55

[6] Mr. Sudarshan G. Ghuge, Mr.Santosh G. Karkhil. "Operating an Application Using Hand Gesture Recognition System". International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 2, February 2016