Scientific Journal of Impact Factor (SJIF) : 3.134

ISSN (Print) : 2348-6406
ISSN (Online): 2348-4470

International Journal of Advance Engineering and Research Development

Control Windows Photo Viewer by Hand Gesture Recognition using MATLAB

Manoj Makhija[1],Kapil Goyal[2]

[1].Electronics and Communication Department,Jaypee Institute of Information and Technology,Noida,UP
manojmakhija9@gmail.com
[2].Electronics and Communication Department,Jaypee Institute of Information and Technology,Noida,UP
kapilgoyal150@gmail.com

Abstract*: First there was input of information or commands through keyboard, then touch input was introduced. Now the big question is, What Next? The answer is Gesture Recognition. The best method for gesture recognition is through image processing techniques. Image viewing is one of the most used applications of a personal computer. This paper works on controlling different functions of the image viewer through hand gesture recognition. In this paper a simple yet effective approach has been used to zoom in, zoom out and to view different images in an image viewer. Thumb plays an important role in the algorithm and the gestures used makes sense with the functions that these gestures are doing in the image viewer. This paper used a 30 fps with up to 30 mega pixel camera. Furthermore, pre-processing is done to segment the hand region from its background. The proposed algorithm can be used for controlling different applications as well.*

Key words*— Image processing, Hand gesture recognition, Human computer interaction, controlling applications, computer vision, Image viewer.*

I. INTRODUCTION

1.1 Motivation

Computer applications are controlled with the help of keyboard, mouse, joystick etc. and nowadays it is controlled our with a touch. Hand gestures are the next logical step towards making the control easier. Gestures are motion of the body/face/hand or any physical action form by the user in order to convey some information to someone else. Gesture recognition is the process by which gesture made by the user is made known to the system, through the use of computer vision via webcam connected to the computer. There is great emphasis on using hand gesture as a substitute of new input modality in broad range applications. Hand gestures have the natural ability to represent ideas and actions very easily, thus using these different hand shapes, being identified by gesture recognition system and interpreted to generate corresponding event, has the potential to provide a more natural interface to the computer system. This type of natural interaction is the core of immersive virtual environments. It can be simply realize that a wide range of gestures are utilized in our daily personal communication. People gesticulate more when they are talking on their cell phones and are not able to see each other as in face to face communication. In fact, we gesticulate twice more than we speak. The gestures vary greatly among cultures and context still are intimately used in communication. The significant use of gestures in our daily life as a mode of interaction motivates the use of gestural interface and employs them in wide range of application through computer vision.

1.2 Related Work

There are many gesture recognition techniques developed for tracking and recognizing various hand gestures. Each one of them having their own advantages and disadvantages. First of them is wired technology, in which users need to tie up themselves with the help of wire in order to connect or interface with the computer system. In wired technology user isn't able to move freely as they are connected with the computer system via wire. The issue was the length of the wire. Like a telephone it does not give us the freedom and therefore have limited movement. Then there are Instrumented gloves or electronics gloves or data gloves. These instrumented gloves are made up of some sensors, which provide the information related to hand location, finger position etc. through the use of sensors. The advantage of using sensors is that it almost never goes wrong. These data gloves provide good results but they are extremely expensive to utilize in wide range of common application. Data gloves are then replaced by optical markers. These optical markers project Infra-Red light and reflect this light on screen to provide the information about the location of hand or tips of fingers wherever the markers are wear on hand, the corresponding portion will display on the screen. These systems also provide the good result but require very complex configuration. Later on some advanced techniques have been introduced like Image based techniques which requires processing of image features like texture, color etc. If we work with these features of the image for hand gesture recognition the result may vary and could be different as skin tones and texture changes very rapidly from person to person from one place to another in the world. And also under different illumination condition, color texture gets modified which leads to changes in observed results. For utilizing various hand gesture to promote real time application this paper works on vision based hand gesture recognition system that work on shape based features for hand gesture recognition. Every person poses almost similar hand shape with one thumb and four fingers under. The proposed approach is designed and implemented for working on hand gestures.

II. THE PROPOSED ALGORITHM

2.1 Pre Processing

Image Pre-processing is necessary for image enhancement. Our cameras take images in RGB format. In this algorithm, the input RGB image gets converted in to YCbCr images as RGB color space is more sensitive to different light conditions YCbCr is not an absolute color space; rather, it is a way of encoding RGB information. The actual color displayed depends on the actual RGB primaries used to display the signal. Therefore a value expressed as YCbCr is predictable only if standard RGB primary chromaticity is used. Cathode ray tube displays are driven by red, green, and blue voltage signals, but these RGB signals are not efficient as a representation for storage and transmission, since they have a lot of redundancy. Redundancy is the number of bits used to transmit a message minus the number of bits of actual information in the message. It is the amount of wasted space used to transmit certain data [2].

$$\begin{aligned}
Y' &= 0 + (0.299 \cdot R'_D) + (0.587 \cdot G'_D) + (0.114 \cdot B'_D)\\
C_B &= 128 - (0.168736 \cdot R'_D) - (0.331264 \cdot G'_D) + (0.5 \cdot B'_D)\\
C_R &= 128 + (0.5 \cdot R'_D) - (0.418688 \cdot G'_D) - (0.081312 \cdot B'_D)
\end{aligned}$$

Where Y' is the luma component and C_B and C_R are Chroma components. Here, the prime ′ symbols mean gamma correction is being used; thus R′, G′ and B′ nominally range from 0 to 1, with 0 representing the minimum intensity (for black) and 1 the maximum (for white). The following figure shows RGB and YCbCr images.

RGB Image YCbCr Image

Fig.1 RGB and YCbCr images

Noise Removal: We fill holes in the image by a structuring element B.

$X_k = (X_{k-1} \oplus B) \cap A_c$

B is the symmetric structuring element. The algorithm terminates at the iteration step k if Xk= X_{k-1}. The set X_k then contains all the filled holes; the union of X_k and A contains all the filled holes and their boundaries contains all the filled holes. After removing noise it uses a fuzzy filter. Fuzzy filters provide promising result in image-processing tasks that copewith some drawbacks of classical filters. Fuzzy filter is capable of dealing with vague and uncertain information.

Image takenImage after noise removal

Fig 2. Noise removal

Then convert the noiseless image to binary image. In binary images, each pixel is stored as a single bit 0 or 1. Fig 3 shows the binary images.

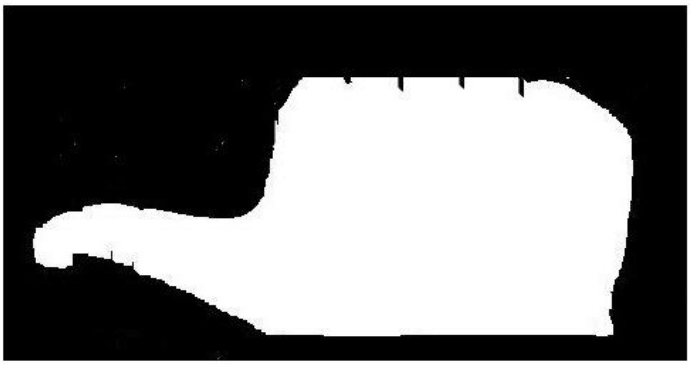


Fig 3.Binary Image

Dilation and Erosion: Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbours in the input image.

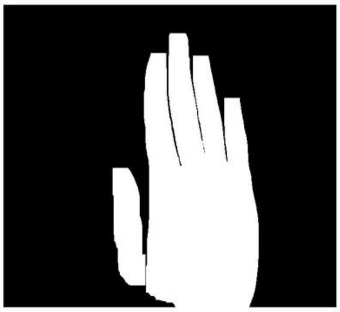


Fig 4. Image after performing Dilation and Erosion

2.2 Segmentation

Find boundary edges to locate the hand in the image. We localize the hand in the image. This is done by scanning the converted binary image from left to right, right to left and top to bottom. The first white pixel (or 1) on the left gives us the leftmost boundary of the hand, the first white pixel on the right gives us rightmost boundary of the hand and the first white pixel from the top of the image gives us the topmost boundary of the hand. We do not repeat the same process for the bottommost white pixel as it will always give the hand [1].

2.3 Orientation Detection

We calculate the coordinates of the different boundaries of the localized hand object. In the binary matrix (of the localized binary image) we calculate the number of rows as well as the number of columns.
(row2-row1) gives us the number of rows.
Where row2 is the last row and row1 is the first row of the binary matrix.
Similarly (column2-column1) gives us the number of columns.
Where column2 is the last column and column1 is the first column of the binary matrix.
We take the ratio of these two quantities and if this ratio is greater than 1 then the hand is vertical and if it is less than one then the hand is horizontal.

$$Y = \frac{(row2 - row1)}{(column2 - column1)}$$

Y=ratio>1, vertical
Y=ratio<1, horizontal

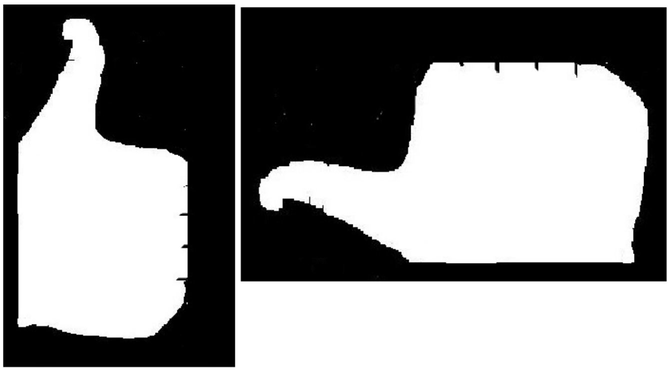


Fig 5. Vertical and Horizontal images

2.4 Centroid

Centroid is defined as the Centre of mass of an object of uniform density. Centroid of the hand object is calculated for further finding the number of fingers raised [3].

To find the centroid:
Product of X coordinates of all the white pixels and their intensity are added up.

$$\text{Sum x}=\sum\sum xf(x,y)$$

Similarly, product of Y coordinates of all white pixels and their intensity are added up
Sum y=$\sum\sum yf(x,y)$

To get the average we divide these values by the number of pixels or the area of the image.

M_{00} gives the area of binary image, where M is the Image moment. Image moment is the weighted average or moment of the image's pixel's intensities. For scalar or grayscale images, image moment [4][5] is given by

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y)$$

Where Mijis the image moment x and y are the coordinates of the pixels I(x,y) is the intensity of image's pixel at (x,y) order of the moment is i+j Centroid, C(a,b) = (M_{10}/M_{00},M_{01}/M_{00})

Where M_{10}=Sum x M_{01}=Sum y M_{00}=area of the binary image (a,b) are the coordinates of the centroid

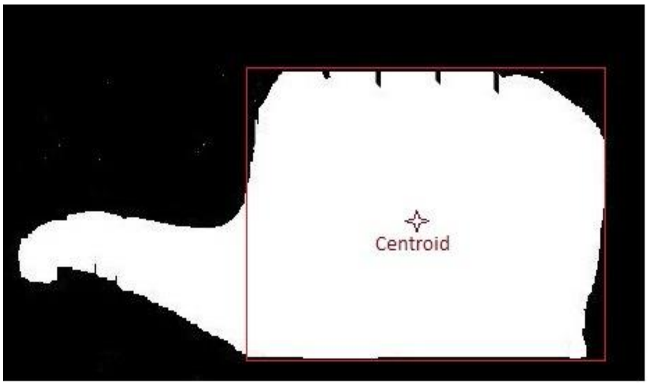


Fig 6. Centroid

2.5 Gesture Recognition

(a) For Vertical Image,

If the value of y is greater than '1' i.e. it is a vertical image,
then the gesture is found using y co-ordinates of the image.
First the hand region without the thumb is localized in a box, then find the bisector of the box using y co-ordinates. Now it is checked if the centroid of the full hand image lies on the bisector then the thumb is not present. if it lies above the bisector then the thumb is present on the upper side and if centroid lies below the bisector then thumb is present on the lower side of the hand.

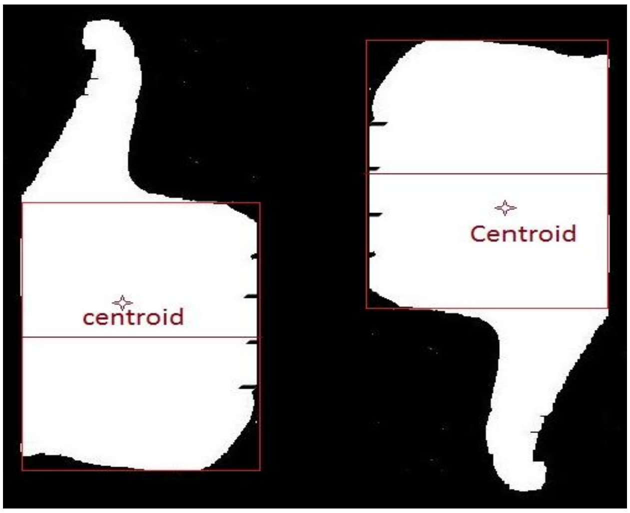


Fig 7. Vertical Gestures

(b) For Horizontal Image,
If the value of y is less than '1' i.e. it is a Horizontal image,then the gesture is found using x co-ordinates of the image.
First the hand region without the thumb is localized in a box, then find the bisector of the box using x co-ordinates. Now it is checked if the centroid of the full hand image lies on the bisector then the thumb is not present. if it lies on the left of the bisector then the thumb is present on the left side and if centroid lies on the right of the bisector then thumb is present on the right side of the hand.

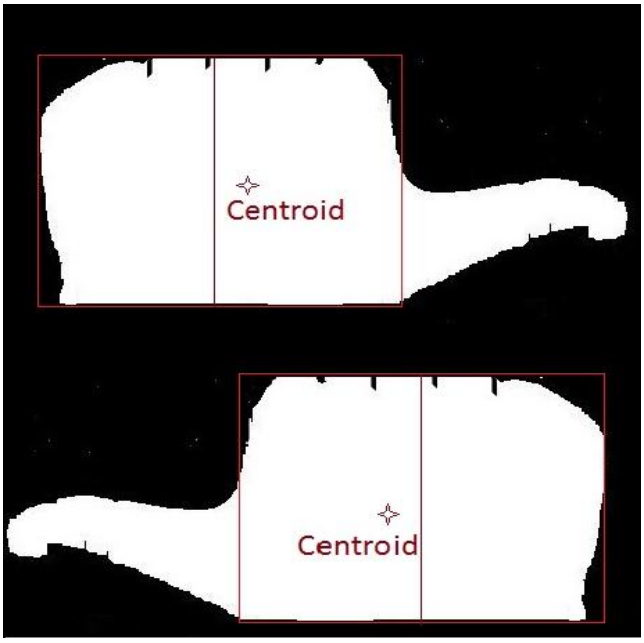


Fig 8. Horizontal Gestures

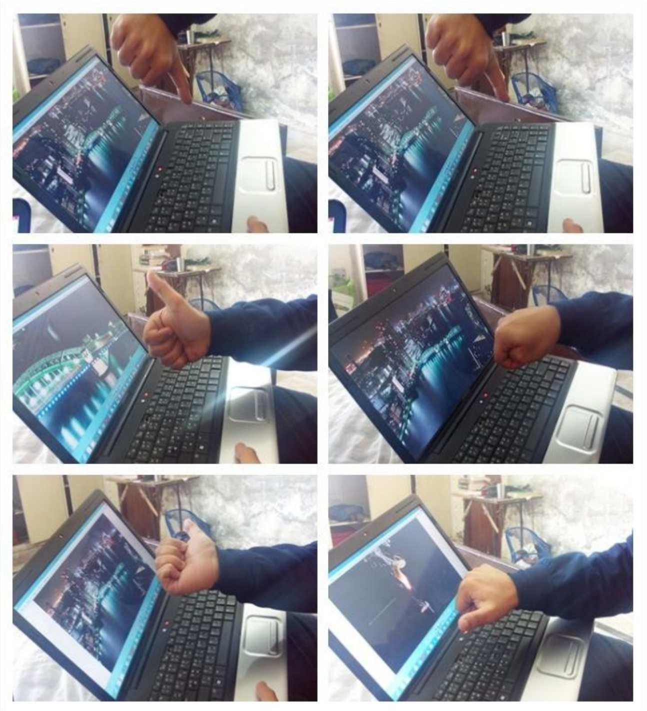
III. RESULT

This paper worked on controlling various functions of an image viewer. The algorithm was tested with over 60 plus hand images and the efficiency is over 95% for all the cases.

Gesture	Recognition Rate(%)
1.) Horizontal Hand Thumb on Left	97.2
2.) Horizontal Hand Thumb on Right	96
3.) Vertical Hand Thumb Down	97.4
4.) Vertical Hand Thumb Up	95.9
5.) No Thumb	97

Gesture	Functions
1.) Horizontal Hand Thumb on Left	Previous Image
2.) Horizontal Hand Thumb on Right	Next Image
3.) Vertical Hand Thumb Down	Zoom Out
4.) Vertical Hand Thumb Up	Zoom In
5.) No Thumb	Full Screen

Different gestures are assigned to different functions as follows



We proposed the shape based approach for hand gesture recognition to run multiple applications with just movement of our hand. This paper first pre-processed the image and then find centroid and check where the centroid lies. On the basis of the centroid gesture is recognized and according to the gesture function of the photo viewer application are to be controlled by gesture recognition. This can save a lot of time and works perfectly for people as time is of the essence. We can increase the number of application that we can use using this algorithm The proposed algorithm is simple and independent of user characteristics. In future, work on both the hands can be done.

REFERENCES

[1]Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction published in International Journal of Human Computer Interaction(IJHCI), Volume (3) : Issue (1) : 2012

[2] Hand Gesture Recognition based on Shape Parameters Meenakshi Panwar Centre for Development of Advanced Computing Noida, Uttar Pradesh , India

[3] Meenakshi Panwar and Pawan Singh Mehra , “Hand Gesture Recognition for Human Computer Interaction”, in Proceedings of IEEE International Conference on Image Information Processing(ICIIP 2011), Waknaghat, India, November 2011.

[4] Rajeshree Rokade , Dharmpal Doye, Manesh Kokare, “Hand Gesture Recognition by Thinning Method”, in Proceedings of IEEE International Conference on Digital Image Processing (ICDIP), Nanded India, pages 284 – 287, March 2009.

[5] Computer Music Controller Based on Hand Gestures Recognition Through Web-cam Junhao Jiang, Junji Ma, Yiye Jin Department of Electrical Engineering, Stanford University