Hand Gesture Recognition System

Swapnil D. Badgujar[1], Gourab Talukdar[2], Omkar Gondhalekar[3], Mrs. S.Y. Kulkarni[4]

[1,2,3] Department of Computer, Engineering, Pune University, India
[4] Assistant Prof., Department of Computer, Engineering, Pune University, India

Abstract- Gestures are a major form of human communication. Hence gestures can be found to be an appealing way to interact with computers, since they are already a natural part of how people communicate. A primary goal of gesture recognition is to create a system which can identify specific human gestures and use them to convey information for controlling device and by implementing real time gesture recognition a user can control a computer by doing a specific gesture in front of a video camera which is linked to the computer.

A primary goal of this gesture recognition research is to create a system which can identify specific human gestures for the control the traffic signals and mouse.

This project also covers various issues like what are gesture, their classification, their role in implementing a gesture recognition system for traffic and mouse control, system architecture concepts for implementing a gesture recognition system, major issues involved in implementing gesture recognition system, and future scope of gesture recognition system. For implementation of this system real-time hand tracking and extraction algorithm, and feature extraction are used.

Index Terms- Gesture Recognition, Hidden Markov Model, Hand Tracking, Feature Extraction, Background Subtraction.

I. INTRODUCTION

A primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control. To understand what gestures are, an examination is required of how other researchers view gestures. How do sociologists and biologists define and view "gesture"? How is information encoded in gestures? Also how humans use gestures to communicate with and command other people are explored. Furthermore, engineering researchers have designed a variety of "gesture" recognition systems.

People frequently use gestures to communicate. They are used for pointing to a person, to get his attention & convey information about spatial and temporal characteristics. Gesturing does not simply embellish spoken language, but it is part of the language generation process.

II. NATURE OF THE PROBLEM

When traffic is very heavy, an automated traffic light system is not efficient to control traffic causing traffic jam. Also sometimes at the peak hours traffic at some roads is very heavy while at the opposite end its less. But the time for each signal is same and that causes some of the crossroads to be jammed while some of them are empty.

Another problem is about mouse. Physically handicapped people can find it difficult to handle mouse. Also while travelling it might be inconvenient to use the mouse.

2.1 Previous Systems

There are many systems for regulating the traffic given as following:

The author Zhang Yuye et.al. [1] System use AT89C51 and CAN BUS controller which leads to complicated design and cost of the system more because of CAN BUS controller. Also in this case power requirement will be more in case of AT89C51.

The author Manoj Kanta Mainali et.al.[2] proposed a genetic algorithm approach to estimate the traffic volume in road sections without the traffic information of road sections. This method estimates the unknown traffic volume using only the known traffic volumes.

The author Cai Bai-gen et.al.[3] designed a vehicle detection system based on magneto-resistive sensor is composed by wireless traffic information collection nodes which are set on two sides of road to detect vehicle signal. The magneto-resistive sensor is expensive and maintenance cost of the system will be more if the system fails.

2.2 Proposed System and purpose of the system

To solve the problems, a system can be used called Gesture Recognition System. A primary goal of this gesture recognition is to create a system which can identify specific human gestures and use them to convey information to control traffic signals as per traffic controller's wish and also for controlling the mouse.

Two methods are considered suitable for gesture recognition. The first one is to use vision sensors like cameras to acquire images, which are analyzed to recognize the gestures. The second one is to place inertial sensor on the traffic police hand and extract the motion characters. The most advantage of the vision method is that it can recognize gestures without adding any extra hindrance to the traffic controller. Also sensors are inconvenient for use while travelling.

Also non-gesture recognition systems increase the cost and unnecessary hardware's while this sensorless system is less costly and also efficient to use.

So the purpose of this system is to control the traffic signals and mouse using hand gestures without using sensors at lower cost and with ease.

III. WORKING

During implementation one thing was clear that a system is going to be developed which can capture a hand gesture

performed by the user in front of web Cam, this captured image is then processed to identify the valid gesture through specific algorithm & execute the corresponding operation. The overall implementation of process is described as follows:

3.1 Human Generated Gesture

In the first step of implementation user will show one gesture. The gesture should be constant for some period of time, which is necessary for dynamic processing. These gestures should be already defined as valid gesture for processing.

3.2 Web Camera

The purpose of Web camera is to capture the human generated hand gesture and store its image in memory. The package called Java Media Framework is used for storing image in memory and again calling the same program after particular interval.

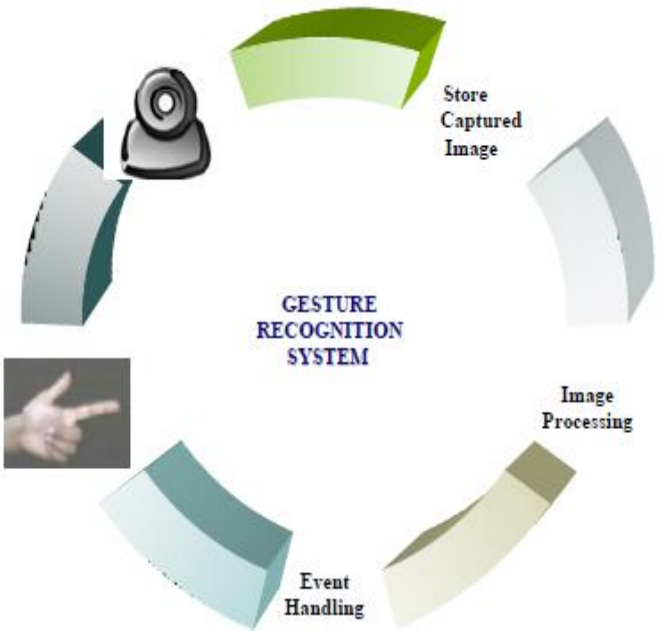

Fig 1 Interaction among the components

3.3 Image Processing Algorithm

This carries the major portion of implementation. First the captured image is preprocessed by techniques like real-time hand tracking and extraction algorithm, feature extraction, hidden markov model(hmm) training and gesture recognition.

3.4 Event Handling

Once the gesture is identified the appropriate command for it will be executed. This includes controlling mouse, performing its various applications like selecting, dragging and pasting any folder from one place to another, both left and right clicks and scrolling. This also includes controlling all traffic signals through different gestures for each signal.

IV. IMPLEMENTATION AND SYSTEM DESCRIPTION

4.1 Hand tracking and handshape extraction

Here, a real-time hand tracking method is developed. This method is robust and reliable in complex background. For tracking the moving hand and then for extracting the hand shape fast and accurately, the trade-off between the computation complexity and robustness need to be considered.

4.1.1 Feature extraction

To find the movement information, the input gesture is assumed to be non-stationary or moving. When objects move in the spatial-time space, an image sequence is generated, motion detector is able to track the moving objects by examining the local gray-level changes. Let $F_i(x,y)$ be the ith frame of the sequence and Di(x,y) be the difference image between the ith and the (i+1)th frame defined as:

$$D_i(x,y) = T_i\{|F_i(x,y)-F_{i+1}(x,y)|\} \quad (1)$$

where Ti is a thresholding function, $F_i(x,y)$ and $D_i(x,y)$ are all 160 X 120 images, and $D_i(x,y)$ is binary image defined as follows:

$$D_i(x,y) = \begin{cases} 1, & |F_i(x,y) - F_{i+1}(x,y)| \geq \text{threshold} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

(1)Thresholding- Having extracted the moving object region, the thresholding on the frame difference (i.e. Eq. (2)) can be aaplied for the extraction of the possible moving region in complex background. Conventional thresholding methods, such as Ostu thresholding are not suitable for the case of detecting motion difference. Instead, a simple thresholding technique is used to extract moving regions. The threshold for motion detection is determined as $t_M=0.2\mu$; where μ is the average luminous of captured image $F_i(x,y)$:

(2) Skin color detection- Skin can be easily detected by using the color information. First, we use the constraint, i.e. R < G < B, to find the skin color regions which may include a wide range of colors, such as pink,red, orange, and brown colors. So, we will find many regions other than the skin regions. However, those non-skin regions which satisfy our constraint will be excluded due to there is no motion information, e.g. orange color region will not be misidentified as the hand region. Second, some sample colors from the hand region may be obtained. To find the skin regions, we compare the colors in the regions with the prestored sample colors. If they are similar, then the region must be skin region..

(3) Edge detection- Edge detection is applied to separate the arm region from the hand region. There are fewer edges on the arm region than on the palm region. A simple edge detection technique (e.g. Kirsch edge operator) to obtain different direction edges is used, and then the absolute maximum value of each pixel is chosen to form the edge image of ith frame as $E_i(x,y)$:

(4) Combination of motion, skin color, and edge- The hand gestures information consists of skin color,movement and edge feature. We use the logic 'AND' to combine these three types of information, that is:

$$C_i(x, y) = D_i(x, y) \wedge S_i(x, y) \wedge E_i(x, y) \quad (3)$$

where $D_i(x,y)$, $S_i(x,y)$ and $E_i(x,y)$ indicate the movement, skin color and edge images. The combined region is shown in Fig 2. Fig. 3 shows the combined region $C_i(x,y)$: The combined image consists of a large region in the palm area and some small regions in the arm area.

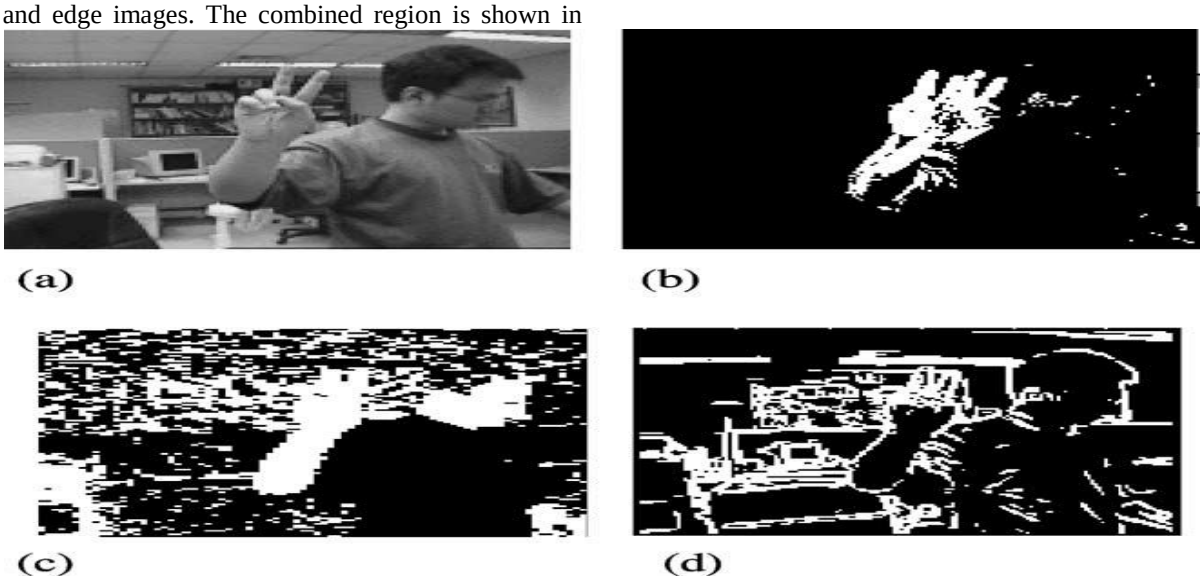

Fig. 2. The hand gesture information. (a) Original image $F_i(x, y)$, (b) motion region $D_iS(x, y)$, (c) skin color region $S_i(x, y)$, (d) edge region $E_i(x, y)$:

Fig. 3. The combined region $C_i(x, y)$:

(5) Region identification.- A simple method for region identification is to label each region with a unique integer number which is called the labeling process. After labeling, the largest integer label indicates the number of regions in the image. After the labeling process, the small regions can be treated as noise and then be removed. Fig. 4(a) shows that the labeling results and Fig. 4(b) shows the center position $p_c(i)$ of the hand region.

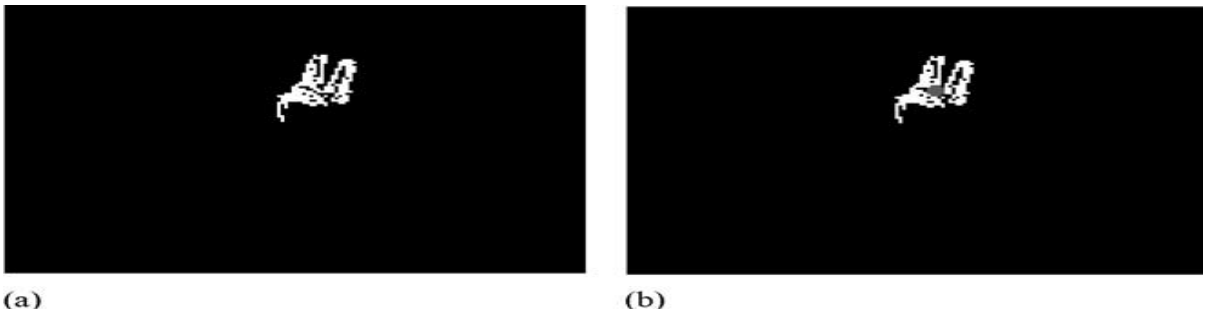

Fig. 4. (a) The labeling results $L_i(x, y)$; (b) the correct center position.

4.1.2 Robustness and low complexity

Only motion and color information is not sufficient, and also hand-shape is not always the largest labeled region. If there are other skin-color objects moving rapidly, the tracking process may fail. We need to take advantage of the motion smoothness constraint for trajectory justification, and use background subtraction to find the foreground object, and finally we can identify the hand region.

(1). Background subtraction- A simple background subtraction technique is used to obtain the hand gesture shape. We Fig. 5 shows the foreground region.

To update our background model, we adapt our background model by using current frame F_i and foreground region FG_i: We have generated two different types of foreground regions, one is $FG1_i = FG_i$; which is used to obtain the hand gesture region; and the other is $FG2_i$, which is applied for background updating process. $FG1_i$ has a compact shape while FG2i is generated for background updating.. The background update equation is:

$$BG_{i+1} = (1-w)BG_i + wF_i \quad (4)$$

Fig. 5 shows the flow diagram of our hand gesture tracking system and Fig. 6 shows the results of hand gesture region extraction process.

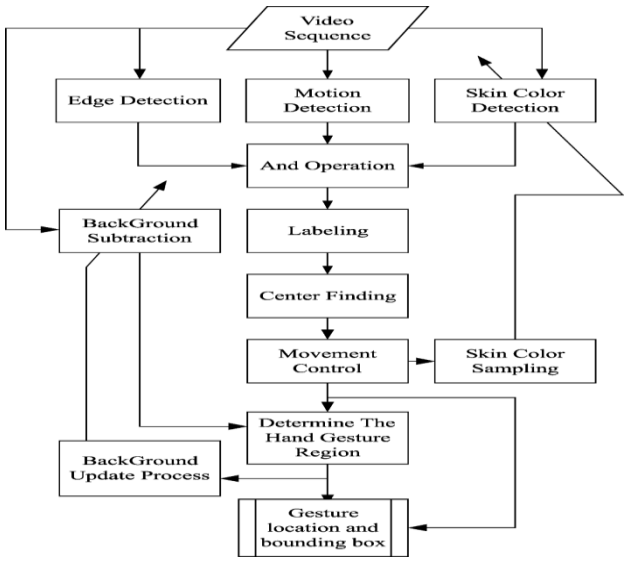

Fig. 5. The flow diagram of hand gesture tracking system.

(a) current frame. (b) foreground. (c) hand region

Fig. 6. Foreground region combining skin color and hand gesture position.

V. APPLICATIONS AND FUTURE ENHANCEMENT

5.1 Applications

This hand gesture recognition system is very vast in day-to-day technical solutions. Its main applications are

1. For controlling the traffic signals as per the wish of traffic controller to reduce the problem of traffic jam at peak hours.
2. To control the mouse so that physically handicapped people can use it and also it mouse will not be inefficient to use while travelling.

5.2 Future Enhancement

1. By integrating our system with voice recognition system we can embed it in ROBOTS.
2. We are also able to handle dynamic image processing and event handling accordingly.

VI. CONCLUSION

In this paper we have presented a method to recognize the unknown input gestures by using hand tracking and extraction method. We apply this system to recognize the single gesture. In the experiments, we assume stationary background so that our system will have smaller search region for tracking.

Using this model we have developed an application where we can control mouse with the finger using it on web cam. Also we have developed an application controlling traffic signals using hand gestures.

ACKNOWLEDGEMENT

We would like to thank our guide Prof. Mrs. S.Y Kulkarni Asst. Prof., Computer Department, Smt. Kashibai Navale College of Engineering, University of Pune, for her help and guidance to enable us to propose this system.

REFERENCES

[1] Zhang Yuye & Yan Weisheng;"Research of Traffic Signal Light Intelligent Control System Based On Microcontroller", First International Workshop on Education Technology and Computer Science,pp301- 303. 2009.

[2] Manoj Kanta Mainali & Shingo Mabu ;"Evolutionary Approach for the Traffic Volume Estimation of Road Sections", pp100- 105, IEEE. 2010.

[3] Ms promila sinhmar; " Intelligent Traffic Light and Density Control using Ir Sensors and Microcontroller"; International Journal of Advanced Technology & Engineering Research; ISSN NO: 2250-3536; VOLUME 2, ISSUE 2, MARCH 2012.

[4] Swapnali Shirke & P G Chilaveri;" Traffic Light Control Using Accelerometer Sensor on ARM Platform" International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 9, September - 2013

[5] Hand gesture recognition using a real-time tracking method and hidden Markov model sq Feng-Sheng Chen, Chih-Ming Fu, National Tsing Hua University, ROC Received 15 January 2001; received in revised form 2 January 2003; accepted 20 March 2003

AUTHORS

First Author – Swapnil D. Badgujar, Department of Computer, Engineering, Pune University, India
Second Author – Gourab Talukdar, Department of Computer, Engineering, Pune University, India
Third Author – Omkar Gondhalekar, Department of Computer, Engineering, Pune University, India
Fourth Author – Mrs. S.Y. Kulkarni, Assistant Prof., Department of Computer, Engineering, Pune University, India