Mouse Control using Color Bands

Prof. Bhakti Sonawane[1] Ashish Haram[2] Kavita Kamat[3] Nayan Mehta[4] Rachit Nandu[5]
[1,2,3,4,5]Department of Computer Engineering
[1,2,3,4,5]Shah and Anchor Kutchhi Engineering College, Chembur, Mumbai, India

Abstract— In the present era, although many technologies are developed which enable the interaction of the physical world with the digital world, very few are based on gesture recognition which is independent of any hardware. Basically in our project we are using 3 blobs which would replicate the three buttons of the mouse - left, right and middle button. By keeping the finger in a particular position for a specified amount of time, the user will be able to perform a mouse event. The goal of our project is to operate the mouse with the use of the fingers. It will thereby eliminate the gap between the physical and digital world.

Key words: Colour Recognition, Open CV, Image Processing, Blob

I. INTRODUCTION

Our project focuses on using Sixth sense technology which consists of using wearable devices to interact with the digital world with the help of gestures. In the present era, computers play an important role in our day-to-day life. Consequently, new applications and hardware are introduced for better interaction with the computer. The uses of existing computer devices are limited to hardware, such as mouse, trackball, keyboard etc. However, these devices are almost similar in functionality and features with regard to the speed with which we interact with the computer. Instead using hand gestures for interaction with the computer would prove to be a better solution and hence would avoid the use of any sort of hardware which limits the speed. Cursor movement using hand gestures will in-turn provides a natural-computer interface by allowing us to move the cursor on the screen with the help of gestures. Hand gestures are mainly of two main types: static hand gesture and dynamic hand gesture. A static hand gesture is only the position of hand using only a single image whereas a dynamic hand gesture records the movements of hand using sequence of images.This technique records the hand gestures of the user and takes it as an input to interact with the computer. The process first records the images continuously with the help of the web camera. The user must be wearing color bands on his hands during the entire process. The web camera captures only those images which consist of color bands and the rest of the images are discarded. The images consisting of those colors mentioned in the code are mapped with the image pixel in OpenCV are only detected and the mapping of the pixel position with the mouse input takes place. Thus the mouse cursor on the screen moves in synchronization with the user's hand which is taken as input. Different scaling techniques are used to cope up with size of the image such as image segmentation, image analysis and image processing.Using the concept of sixth sense technology, we can also implement different applications such as drawing of scenery on the screen with the help of hand gestures, watching the videos, taking live pictures and saving them and many more.

I-Mouse implementation is an application which makes use of computer system using finger movements. The reason behind using finger movements is that they are different from conventional computer inputs such as mouse, keyboard or any other hardware. The overall approach in designing interaction techniques is, wherever possible, obtain information of the user's finger movements while viewing the screen and match it to the mouse cursor.

The problem that the system tries to solve is particularly straightforward to understand.

The User Requirements can be summarized into: 'the software exploits the finger movements of a human in order to provide a Human Computer Interface usable without hands'. System Requirements are proposed considering various aspects:

1) Range of Users: the user must be close to the system.
2) Usability: the system must be transparent to the user. [1]

II. METHODS AND ANALYSIS

A. Sixth Sense Technology

Sixth sense technology [7] is a device which helps us to use hand gestures to interact with the Computer system without the need of any hardware. In scientific terms it is a wearable gestural interface. [2] It is based on the concepts of augmented reality and has well implemented the perceptions of it. Sixth sense technology has integrated the real world objects with digital world. . It associates technologies like recognizance of gesture using hand, recording video, manipulating image etc. It superimposes the digital world on the real world. Sixth sense technology is a perception of augmented reality concept. Like senses enable us to perceive information about the environment in different ways it also aims at perceiving information. Sixth sense is in fact, about comprehending information more than our available senses. The system recognizes the user's free hand gestures as well as icons/symbols drawn in the air with the index finger instead of a mouse.

The components needed for our system are:

1) Web Camera

The webcam records the free hand gestures of the user by capturing the video. The video is then sending to the code in OpenCV [4] and C# for processing. In image processing the video is sliced into images and only images which contain the color mentioned in the code are recognized. In short a camera is used to keep a track of the finger movements of the user.

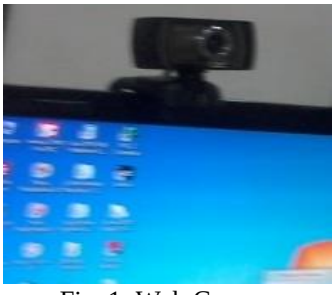

Fig. 1: Web Camera

2) Color Caps

Color caps are needed in order to perform the necessary mapping of the movements of the user's fingers with the mouse clicks. Our system uses same color of color caps, which are to be worn on the thumb, the index finger and the little finger. These are associated for performing left click cursor movement and right click respectively. It is also known as color bands or color marker.

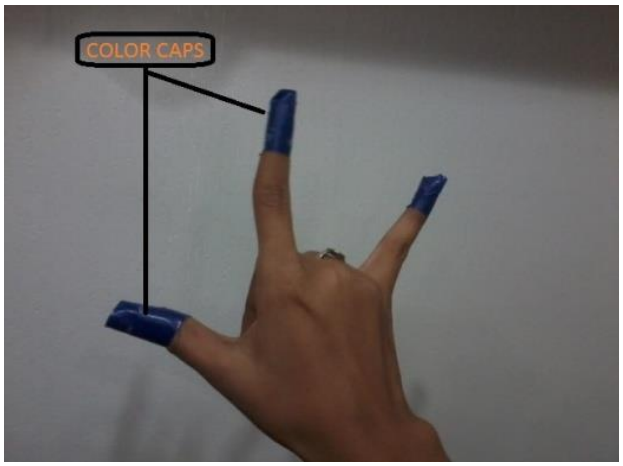

Fig. 2: Color Caps

1) Laptop/computer with the software Open CV and C# installed:

Open CV has many built-in functions which make us compatible to use it. The language used for our code is C#; therefore the need of a laptop with this software is essential for our system. [6]

Fig. 3: Laptop

B. Gesture Recognition

Gesture recognition is defined as any bodily movements that help us to communicate. Gesture recognition allows us to communicate with the machine and interact naturally without any mechanical devices and only with the help of gesture. The concept of gesture recognition allows us to move the mouse cursor on the screen with the help of gestures made through fingers.[3] In turn this will potentially make conventional input devices such as mouse, keyboards ,joysticks and even touch screen redundant and allowing the unencumbered body to give signals to the computer through gestures using fingers.

Recognizing gestures as inputs provides a better interaction for the physically impaired people to interact with the computer system and also provides a more natural interaction without relying on any hardware.

In gesture recognition process, the system is outfitted with the camera to take the gestures as input for the computer and accordingly move the cursor to perform different operations. Many different approaches have been developed to make use of computer vision algorithms that contribute to the development of sign languages. However the recognition and identification of human gestures proved to be useful. The concept of gesture recognition provides assistance and benefit to the physically impaired people who find difficulty in communicating, can use human gestures to interact with the help of sign languages. In gesture recognition the spoken words are interpreted into signs and symbols for better computer interaction. This technology is extremely useful to such kind of people.

In gesture recognition technology, the camera takes the movements of the human body as inputs and communicates the data to a computer. This data that is then used by the computer as input to control various devices and applications. The technology has the potential to change the way users interact with computers by eliminating input devices and make use of bodily movements. It provides a much better alternative to text user interfaces and graphical user interface to make use of fingers instead of a mouse or a keyboard to interact with the computer system. Thus with the help of this technique, the user does not need to solely depend on the hardware devices such as the mouse or a keyboard but in turn interact with the computer only with the help of gestures made by finger movements. The gestures are captured with the help of a web camera; the only need in our system is that the user needs to wear colour bands on his fingers as the system is designed accordingly.

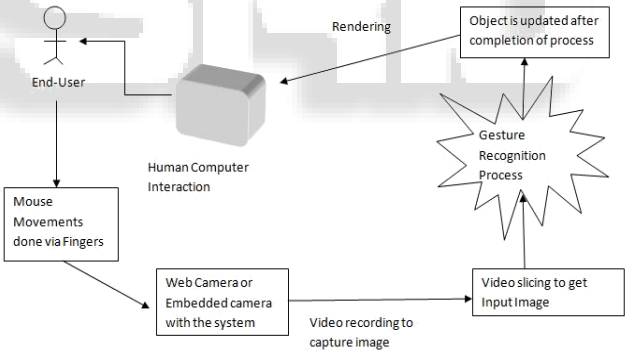

Fig. 4: System Architecture

In Open CV, code is already prepared to convert the live video which is captured by camera into a sequence of images; this process is called as Video Slicing. These sliced images are then processed for colour recognition process. The outcomes of the colour recognition process are the images which get detected by the colour caps are present at the fingertips of the user.

1) Colour Recognition

Initially, when camera will capture an image which is in the RGB format. RGB colour format is the standard colour format uses a combination of Red, Blue and Green components. Nowadays, this technique is most probably used in computer software. But in computer vision, RGB colour format shows some drawbacks such as strong or dim lighting conditions and shadows, etc. While HSV is better to handle these illumination problem.

– Hue: Hue means the basic characteristics of colour. There are infinite numbers of hues present between

any two colours. For example, there will be infinite number or orange hues present between Red and Yellow.

- Saturation: Saturation means the purity/clarity of colour. The colour with the high saturation appears rich and fresh, whereas the colour with low saturation appears to be dull and greyish.
- Value: Value means the Brightness of the colour. Colours with high saturation value have medium values (Because by mixing white or black colour, light or dark colours can be achieved).

2) Colour Thresholding

It is the simplest method to segment or divide an image. It is used to divide an image corresponding to blobs which the user wants to analyse. An image gets separated based on the variation of intensity between the object pixels and the background pixels. It used to separate properly the important pixels, by some thresholding value. [4] (I.e. one can assign them a value of 0 for black, 255 for white or any value that suits).

Thresholding values for the cursor operations:

Left clicking = left side colour cap

Right clicking = Right side colour cap

Move the cursor = left side colour cap and move the cursor accordingly

Dragging = left side colour cap hold for a certain time interval and move hand.

3) Binarization

Binarization is used to convert the image into binary image which have only two values for each pixel i.e. Black and white.

4) Blob Detection

While video slicing, number of blobs get detected. In OpenCV, blob detection refers detecting regions of an image that differ in properties, such as brightness, sharpness or colour, the distance from the camera to the captured blobs. A blob is an object of an image in which some properties are constant or vary within a prescribed range of values as compared to the other objects. These Blobs are renamed and labelled according to the decreasing order of their size.

III. PROPOSED METHODOLOGY

Here, we provide computer with a vision to make itself more intelligent to do activities that input device performs (i.e. mouse). An inter-communication is performed between a user & camera. Camera can be an embedded camera on laptop or a movable web camera. As result of this, a cascading process begins. A cascading is referred to the process where video is recorded during an intercommunication between an user & camera whose output will be a recorded video & which act as an input to video slicing process (conversion of video to image frame set) ; whose output will be image frame set & which act as an input to image manipulation process . Recognizance of colour & image manipulation as well as refining of an image is done, at this instance.

The sequence that we adopt, while implementing our project is as follows:

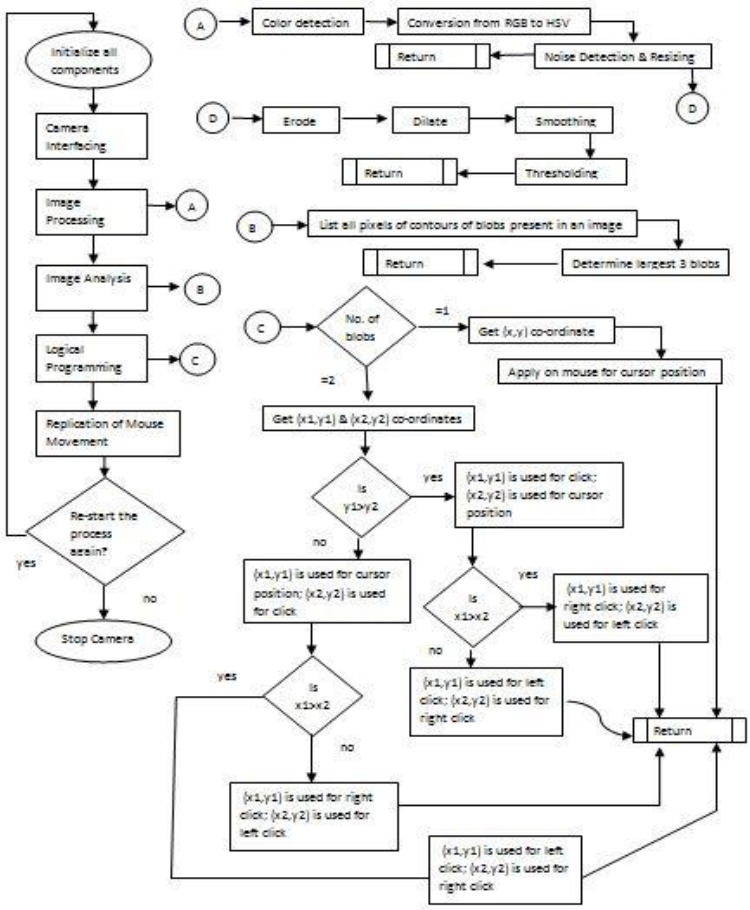

Fig. 5: Flow Chart

The interpretation of above diagram can be given as:

First, the initialization of all required components and process. Then, it follows 5 stages:

A. Camera Interfacing

It involves nothing but inter-communication between a user & camera. The gestures relating to moving fingers is taken as input & sent to back-end for processing in order to get proper output. So, it considered as important process.

B. Image Processing

The ultimate motive of this stage is to distinguish between color caps & fingers. For this purpose, recognizance of color is done as first step of this stage. Conversion from RGB format to HSV format is done, so that object with more intensity is detected. No image is completely noise free. To eliminate this noise, noise detection & resizing process is performed. Noise detection & resizing process comprises of 4 steps as follows:

1) Erode: It means reducing the noise from image as much as possible. For this, we use iterations. It is iterated as long as noise is not removed. As a result, it crosses an object's boundary line thereby shrinking the size of desired object.
2) Dilate: This step is used to get original size of desired object by extending the desired object. For this, we use iterations. It is iterated as much as iterations that have taken place in erode step. As a result, the boundary line of desired object is blurred.
3) Smoothing: After erode & dilate process, boundary line of desired object is blurred. To smoothen the object's boundary line, this step is essential. The boundary line becomes grayish in shade.
4) Thresholding: It is performed to get proper shape of the desired object (in our case, circular). [4]

After this, image is resized & forwarded to next stage.

C. Image Analysis

The goal of this stage is to analyze 3 blobs required to perform basic mouse events. To do this, we assemble all pixels of contours of blobs present in an image using "Linked List" data structure. The number of elements present in linked list is equal to the number of blobs detected in an image. To find largest 3 blobs, we need to calculate area of each blob. For this purpose, we use function contour. Area - an Open CV built-in function.

D. Logical Programming

We have developed code to interpret how a user is trying to fingers as a mouse. For this, we use time delineation by which we can code for various mouse clicks & mouse events accessed by the user. Also, locating proper co-ordinates of blobs are essential to differentiate between blobs used for right click & left click.

E. Replication of Mouse Movement

Scaling is needed as imaginary frame covers only certain region of the desktop. This imaginary frame will be mapped with desktop resolution. To perform fast motion, it is mandatory to use calibration factor. . To perform slow motion, it is mandatory to apply one-to-one position mapping.

To restart application process & run as many time as user wants, the above mentioned steps should be repeated. When user is done with his/her work, user can turn it off by stopping camera interfacing.

IV. OBSERVATION

While implementing the proposed project, we have noticed some positive aspects as well negative aspects of our project. I would like to throw light on positive aspects of our project & discuss on negative aspects of our project later.

First positive aspect of our project is user friendliness. It means, if the user can't perform single click operation in 10 seconds, he can change single click trigger value on user interface, according to his capability & need. This arises second positive aspect of the application, by making it dynamic to user. The third positive aspect is, user can control HSV values of color so as get his/her desired color detected by the system. For our project, we have used blue color.

Let's discuss about few negative aspect of our designed application. Firstly, our project works well if the place well illuminated. Well illuminated, in our project means light beam is coming in same direction in which web camera is placed or say in opposite of direction of user working on application. Thus, it can be interpreted as, our project possess illumination problem. Secondly, if all 3 fingers, which act as mouse, is shown at same instance, no operation will be performed & system messge will notify us about continous increment of count. That implies, mouse is in idle position. If we continuously move our fingers across the screen with fast speed, there is probability that the mouse gets disappear & application may get hang. This constitutes third negative aspect.These were the some of observations that we made during implementation of our project.

V. CONCLUSION

Our proposed methodology will completely eliminate the use of mouse, webcam and microphone, which are considered as an essential part of the computer system. The technology that we use will completely eliminate the use of mouse, with the need of only a webcam for interaction with the computer. Our system will develop an innovative and an easier way of interacting with the computer system, independent of external hardware. This way of human computer interaction will prove to be better where there is no need of any physical contact with the device. Image processing and recognizance of colour techniques in OpenCV and C# proved successful in implementing mouse cursor movement using sixth sense technology.

VI. FUTURE WORKS

We would like to enhance the proposed project more by eliminating illumination properly so that it works well in all possible real time environments. Also, we would like to design project such that it interacts with user efficiently irrespective of the users wearing color caps. We would like to overcome all negative aspects of our system, which were discussed earlier in Observation.

Our paper emphasized on the technique of mouse movement using fingers, the methods involved in achieving

the output, applications of our product in all possible sections. The use of our system is very high and also exhibits efficiency and speed in many desktop and real time applications. In future we would rather work on replacing our finger movements by eye movements and adding more features such as enlarging or closing the window and many more with the use of our palm and multiple fingers.

REFERENCES

[1] https://students.cs.unipi.gr/pub/docs/sysadmin-1992-1998/html/v05/i01/a4.htm
[2] http://www.quora.com/Sixth-Sense-technology
[3] http://en.wikiversity.org/wiki/Introduction_to_Computers/Future_Peripherals
[4] "Learning OpenCV" by Gary Bradski and Adrian Kaehler Published by O'Reilly Media, Inc.
[5] "The OpenCV Tutorials Release 2.4.2" Pg 13 –Pg 14
[6] " Beginning C# 2008 From Novice to Professional " Second Edition by Christian Gross Printed and bound in the United States of America Pg.1 to Pg.26
[7] http://www.slideshare.net/renjithravi/sixth-sense-technology-5706233.

