ISSN: 2321-7782 (Online)
Volume 2, Issue 3, March 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Paper / Case Study
Available online at: www.ijarcsms.com

Improved HCI using face detection and speech recognition

Sarnaik Supriya Vasantrao[1]
B.E. (final year)
Dept of Computer Engg
Dr. D.Y.Patil IET
University of Pune
Pune – India

Chalke Minal Prakash[2]
B.E. (final year)
Dept of Computer Engg
Dr. D.Y.Patil IET
University of Pune
Pune – India

Bombe Akshata Prakash[3]
B.E. (final year)
Dept of Computer Engg
Dr. D.Y.Patil IET
University of Pune
Pune – India

Ghadge Anupama Anant[4]
B.E. (final year)
Dept of Computer Engg
Dr. D.Y.Patil IET
University of Pune
Pune – India

Abstract: ***Computer plays an important role to perform various tasks with high speed. our goal is to make the system easy to use so that human computer interaction is made easy. In this paper we describe a graphical user interface navigation utility, same as the traditional mouse pointing device. In this movement of the pointer is achieved by tracking the motion of the head for this image processing technology is used, while button actions can be operated by voice command. Computerized face recognition is one of the field of biometrics which is commonly used because of the individuality of one human face to other. It has basically two parts one is face detection and other one is recognition of detected faces. for face detection thresholding and HSV color model is used.***

Keywords: Human computer interaction, image processing technology, thresholding, HSV component, color space model.

I. INTRODUCTION

In this project, we will use image-processing technology to detect face movement. Traditional mouse and keyboards required users to have fine motor skills in order to control them. This means that specific disabled people would not be able to utilize a computer, because they could not handle the mouse. The difficulty can be solved with the introduction of face detection technology. However, products of this kind that are in the market are normally costly. Thus, the purpose of our project is to build a solution that will be useable for people with disabilities, with little to no cost involved. The user could control the movement of the mouse and eventually operating system using his or her face movement (as captured by a webcam in real time).

Ordinary human at typing uses both hands and during this typing he cannot move the mouse similarly, many people are unable to operate a standard computer mouse or keyboard because of disabilities of their hands or arms. One likely substitute for these persons is our system, which allows controlling a computer without using standard mouse and keyboard for example (a) Using head movements to control the cursor across the computer screen; (b) using the speech for giving the control commands. Automatic speech recognition and face detection in joint multimodal system are combined in the system. In this paper, we offered complete head movement actions via webcam and button operations via microphone, using image processing technology to recognize the movements of head and command control .Users who has difficulties using the standard PC control devices could manipulate cursor simply by moving their heads and giving the speech command instead of clicking the buttons. Our project retrieves real-time images from a webcam and converts them to grayscale images. Then, it extracts already defined

feature vectors from the images and sends them to Support Vector Machine (SVM) to get the classification. Using the outcome our system will be able to control the mouse cursor in real-time.

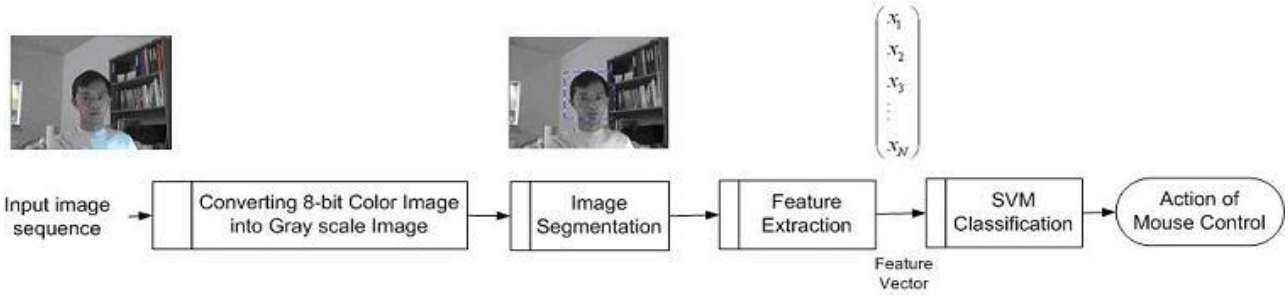

Fig.1. WebCam Face Tracking System Structure

II. System Objectives

The main objective of the project is to design, develop and validate the concept of image processing and face tracking for efficient application/ service/ content provisioning to the disabled people for communication. In order to achieve this main objective, the project will conduct work addressing the various technical challenges, which closely relate to the detailed objectives of the project:

The project replaces the standard computer mouse for people who cannot use or have limited use of their hands when controlling the computer or Augmentative communication device .it translates natural movements of the users head into directly proportional movements of the computer mouse pointer. Idea is to retrieve real time images from webcam and extract features from them for head tracking .using the result our program will be able to control the system in real time.

To develop the appropriate algorithms for enabling the system to function proper. This means the System using HSV, blob detection, skin thresholding algorithm, in order to convert the head movement to control mouse to run any application on computer.

III. Literature Survey

Existing System

The existing development of computer technologies has improved various applications in human computer interface that is related to the facial features that are extracted includes the eyes, mouth, nostrils and eyebrows. Face and gesture recognition is a part of this field, which can be useful in different applications such as in robotic, safety system, drivers, monitor, and record(video) coding system.

Since human face is a dynamic object and has a high degree of inconsistency. Various techniques have been projected previously. Face detection techniques can be classify into two methods: (a) feature based approach and (b) image-based approach. The techniques in the first kind makes used of apparent properties of face such as face geometry, skin color, and motion. Even feature-based technique can achieve high speed in face detection, but it also has problem in poor consistency under lighting condition. For second category, the image based approach takes advantage of current advance in model detection theory. Most of the image based approach applies a window scanning technique for detecting face, which requires huge computation. Therefore, by using only image based approach is not suitable enough in real-time application.

One of the present system is [1]combining head pose and eye location information for gaze estimation.[2] developed a visual based HMI for controlling a wheelchair by head gesture which were recognized by detecting position of the nose on user's face.[3]developed a user friendly human machine interface(HMI) for hands free control of an electric powered wheel chair(EPW).F Carrino *et al.* [4] developed a system, namely "Virtual Move", which allows users to navigate through Google Street View (GSV) using head movements, facial terminology, thoughts and emotional states. Gajwani and Chhabria [5] used

eye tracking and eye blinking obtained by a camera mounted on a cap to control a wheelchair. Huo and Ghovanloo [6] operated a wheel chair by using tongue movements, in which the movement data was obtained from a magnetic tracer on the tongue.

IV. Design and Implementation

Figure 2 shows the overall working of the system. The first step is taking the video feed from the web camera and grabbing the recent frame. Then on this grabbed image we apply the blur algorithm to soften the image means to reduce the noise from the image .then RGB to HSV conversion is done. In situations where color description plays an integral role, the HSV color model is often preferred over the RGB model. The HSV model describes colors similarly to how the human eye tends to identify color. RGB defines color in terms of a combination of primary colors, whereas, HSV describes color using more familiar comparisons such as color, vibrancy and brightness. The color camera, on the robot, uses the RGB model to determine color. Once the camera has read these values, they are converted to HSV values. The HSV values are then used in the code to determine the location of a specific object/color for which the robot is searching. The pixels are individually checked to determine if they match a predetermined color threshold. [7]

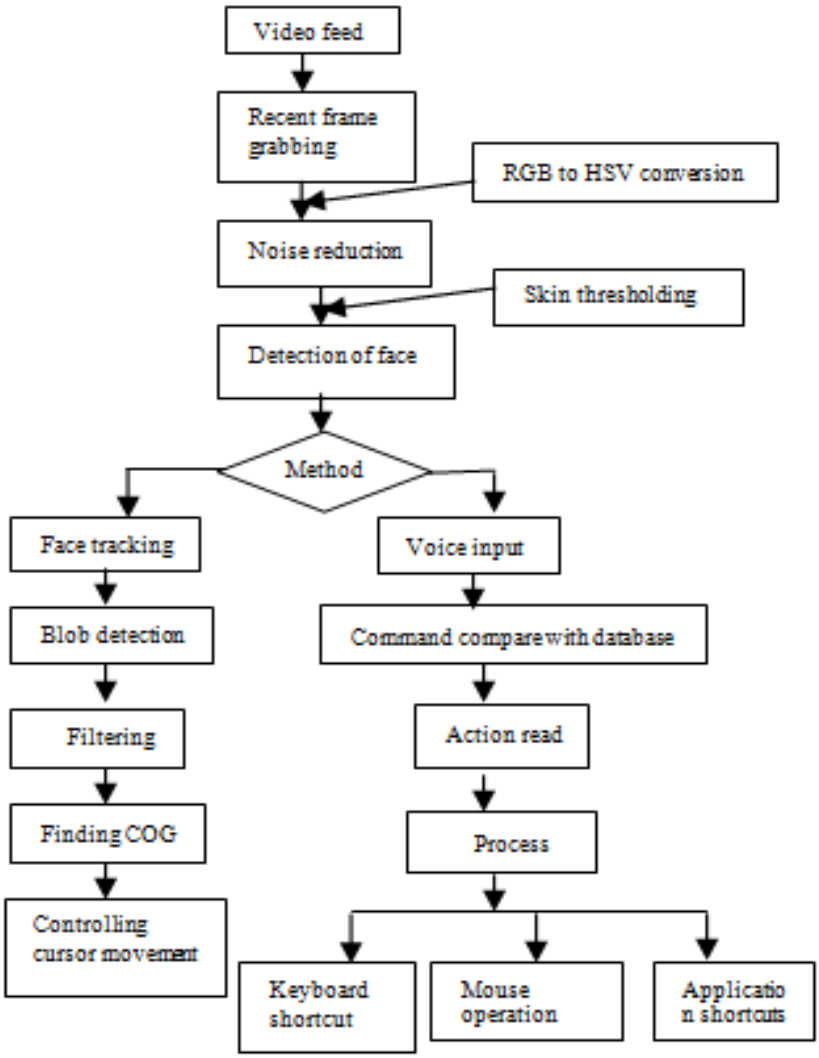

Fig.2: workflow of the system.

Then after this skin thresholding algorithm is applied on the image. Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images i.e. image with only black or white colors. It is usually used for feature extraction where required features of image are converted to white and everything else to black. (or vice-versa).The simplest property that pixels in a region can share is intensity. So, a natural way to segment such regions is through thresholding, the partition of light and dark regions .simply thresholding means differentiating between fore color and background color. Now due to this algorithm face detection is done. On this threshold image blob detection algorithm

is applied so as to filter the image. Blob detection refers to visual modules that are aimed at detecting points and/or regions in the image that differ in properties like brightness or color compared to the surrounding. after filtering the image ,the centre of gravity is calculated using Euclidian's distance formula through which mouse cursor movement is handled and then mouse actions i.e. click, double click etc are handled by voice input. In this process the commands are stored in database using java serialization then via microphone the voice input is given that input is then compared with the commands stored in the database. After comparing action is read and process is completed. The processes that are handled are mouse operations, keyboard shortcuts and application shortcuts .The main algorithms used are (a) Thresholding algorithm (b) Blob detection and (c) the Euclidian's formula is calculated to find the centre of gravity to operate the mouse cursor.

(a) Thresholding algorithm

Thresholding is the simplest method of image segmentation. From a gray scale image, thresholding can be used to create binary images i.e. image with only black or white colours. It is usually used for feature extraction where required features of image are converted to white and everything else to black (or vice versa). During the thresholding process, individual pixels in an image are marked as "object" pixels if their value is greater than some threshold value (assuming an object to be brighter than the background) and as "background" pixels otherwise. This standard is known as threshold over. Variants include threshold beneath, which is opposite of threshold over; threshold within, where a pixel is labelled "object" if its value is between two thresholds; and threshold outside, which is the opposite of threshold within. usually, an object pixel is given a value of "1" while a background pixel is given a value of "0." Finally, a binary image is created by colouring each pixel white or black, depending on a pixel's labels.

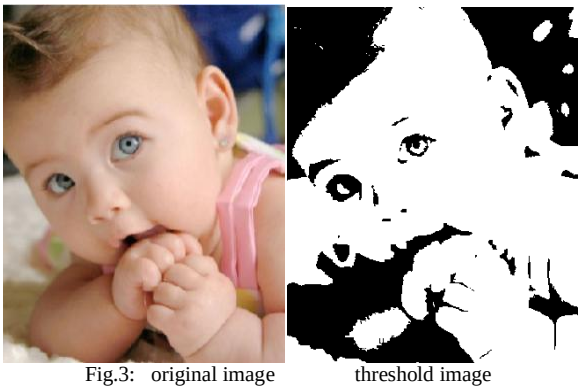

Fig.3: original image threshold image

- **Steps / Algorithm**

  – Traverse through complete input image array.

  – Read individual pixel colour value (24-bit) and convert it into gray scale.

  – Calculate the binary output pixel value (black or white) based on current threshold.

  – Store the recent value at same location in output image.

(b) Blob detection algorithm

Fig.4 Blob detection

In the area of computer vision, blob detection refers to visual modules that are aimed at detecting points and/or regions in the image that differ in properties like brightness or color compared to the surrounding.

There are two major classes of blob detectors

(i) Differential methods stand on derivative expressions and

(ii) Methods based on local extrema in the intensity landscape.

With the latest terminology used in the field, these operators can also be referred to as interest point operators, or alternatively interest region operators. There are several motivations for studying and developing blob detectors. One key reason is to provide corresponding information about regions, which is not gained from edge detectors or corner detectors. In previous work in the area, blob detection was used to obtain areas of interest for additional processing. These areas could signal the existence of objects or parts of objects in the image field with application to object recognition and/or object tracking. In other areas, such as histogram analysis, blob descriptors can also be exercised for peak detection with application to segmentation. a further general use of blob descriptors is as main primitives for texture analysis and texture identification. In latest work, blob descriptors have found increasingly popular use as interest points for wide baseline stereo matching and to signal the presence of informative image features for appearance-based object recognition based on local image statistics. There is also the related idea of point detection to signal the presence of extended objects.

Basic Algorithm

1) Filter with Gaussian at different scales

– This is done by just frequently filtering with the same Gaussian.

2) Subtract image filtered at one scale with image filtered at previous scale.

3) Look for local extrema.

– A pixel is bigger (smaller) than all eight neighbors, and all nine neighboring pixels neighboring scales.

4) More scales can produce greater accuracy and also more expense.

5) We are taking a derivative, so need to be careful about denominator.

– It turns out that we should increase scale multiplicative sigma, k*sigma, k*k*sigma&.

– Sigma = 1.6 produces reasonable results.

– k = cube root (2).

(c) Euclidian's distance formula

We are calculating centre of gravity and set its value. the center of gravity is obtained using Euclidian Distance.

The formula for finding Euclidian distance is as follows.

D = [(l2-l1) ^2+(m2-m1)^2]^(1/2)

And for this we have to calculate the displacement with respect to x and y axis.

The formula for displacement with respect to X and Y axis is as follows:

m=m2-m1 and l=l2-l1

Where,

D: Euclidian distance

m: displacement with respect to x axis.

L : displacement with respect to y axis.

V. CONCLUSION

In this paper, we proposed that this system was capable of performing the majority of an ordinary mouse's operations. With this system, users can operate computers by their head movements in front of web cameras and controls mouse actions by voice input via microphone. We see potential for our approach being integrated in interfaces intended for normal users and users with some disabilities. Particularly attractive should be the fact that it provides a low-cost means of human-computer interaction requiring the most basic of computer hardware. In the developed system the interaction between a user and a computer is performed by voice and head a movement through which computer is operated.

References

1. R.Valenti ,N.Sebe and T.Gevers, "combining head pose and eye location information for gaze estimation",IEEE transaction. On image processing vol.1, no.2, February 2012
2. P. Jia, H.H. Hu, T. Lu, and K. Yuan. Head gesture recognition for hands-free control of an intelligent wheelchair. Industrial Robot: An International journal, 34(1):60-68, 2007
3. E.J.Rechy-Ramirez, H.Hu and K.McDonald-Maier . Head movements based control of an intelligent wheelchair in an indoor environment .in proc. of the 2012 IEEE international conference on robotics and biomimetics, pages 1464-1469 IEEE 2012
4. F. Carrino, J .Tscherrig, E .Mugellini, O .Abou Khaled, and R. Ingold. Head-computer interface: a multimodal approach to navigate through real and virtual worlds. Human-Computer Interaction. Interaction techniques and environments, pages 222–230, 2011
5. P.S. Gajwani and S.A. Chhabria," Eye motion tracking for wheelchair control. International Journal of Information Technology, 2(2):185–187, 2010.
6. X. Huo and M. Ghovanloo," Using unconstrained tongue motion as an alternative control mechanism for wheeled mobility". IEEE Transactions on Biomedical Engineering, 56(6):1719–1726, 2009.
7. www.wikipedia.org
8. www.ieeexplorer.ieee.org

