A Dynamic Hand Gesture Recognition System for Touchless Writer and Accessing Media Player.

Ashwin Raje
Department of Computer Engineering
B. S. C. O. E. R.
Pune, Maharashtra, India

Ankit Pathak
Department of Computer Engineering
B.S.C.O.E.R.
Pune, Maharashtra, India

Ruturaj Jadhav
Department of Computer Engineering
B. S. C. O. E. R.
Pune, Maharashtra, India

Abstract— **In this paper a dynamic hand gesture recognition system for a touchless writer and accessing the media player is presented. Human machine interaction is an evolving branch of computer science. Image processing involves many mathematical algorithms. Gesture recognition system is an innovative way to interact with computers. Conventional systems used for interaction with computers are hardware dependent and costly. We present a system which does not require expensive hardware. Hence it is a cost effective system. This system is very simple to use as naive users can use it easily with little practice. We are using various algorithms like gaussian blurring, RGB to HSV conversion, thresholding etc. This is a real-time system which enables users to perform gestures dynamically rather than only using static gestures.**

Keywords — Gesture recognition, Touchless writer, RGB, HSV, COG, blob, threshold.

I. INTRODUCTION

Computers play a major role in our day to day life. Enormous amount of work is done on computers all over the world. Communication with the computers is an important aspect while dealing with various applications on the computer. Although computer technology has seen immense development in software as well as hardware, the way to communicate with computers has not changed a lot. Generally we use mouse and keyboard for interaction with computers. But these devices can not be the most convenient and natural input devices for most users. This gave rise to a new concept of gesture recognition technology.

Gesture is a form of non-verbal communication which includes movement of hands, face or other body parts. Since ancient times, gestures are a major way for communication and interaction between people. People could easily express the idea by gestures before the invention of language. Nowadays, gestures still are naturally used by many people and especially are the most major interaction way for deaf people [2].Gesture recognition technology in computer science is basically interpreting human gestures by mathematical

algorithms. The applications designed for gesture recognition generally require restricted background, set of gesture command and a camera for capturing images[1]. In this paper we propose a new system using gesture recognition technology, a real-time system which could be very useful and simple to operate. We are going to present the applications which use dynamic hand gestures as input for a media player and for writing application. The idea behind developing this application is to provide an innovative way for users to interact with computers without using expensive hardware. Already existing systems have some problems related with them and we try to overcome these problems.

In gesture recognition systems, the most important part is image processing. We use various image processing algorithms for example, RGB to HSV conversion, image blurring algorithm. In the applications of touchless writer and accessing media player image acquisition is done using a Webcam. Some functions in media player are used more frequently than others and thus we need to use media player for those functions using predefined gestures [1]. Webcam continuously captures images processed by a color matching algorithm, thus the movement of color marked object is analyzed and tracked [5].

II. LITERATURE SURVEY

Previous systems are largely dependent on hardware. As a result the cost increases. Also the cost of maintenance can be more. The traditional input devices such as mouse and keyboard require more attention in case of failure.

In the previous application of accessing media player, we observed that there are some restrictions such as requirement of appropriate background, etc. we overcame these problems by using different coloring algorithms. We are using HSV color system so that it will lead to perfection where previously RGB color system was used for color extraction.

Many different algorithms are used due to which it has become a time consuming task. Touchless cannot provide smooth movement and accurate positioning as other pointing devices [5]. Also due to insufficient light the application may lead to inaccuracy.

With the help of different algorithms we have moved one step further in accuracy as compared to the previous system. This would help in low maintenance cost as compared to hardware and can be easily accessible. Every algorithm helps in increasing accuracy while developing software.

III. PROPOSED SYSTEM

There are different techniques, algorithms that are used to develop the gesture recognition system. As we have seen that, in previous systems there different issues in different areas. To solve some of these issues, we have designed a strong robust, efficient gesture system for two applications, Touchless Writer and media player control.

Fig. 3(a): Example of character or number writing

Basically for any gesture system, video capturing of user action and decomposition of that video into frames or images is basic task. In our proposed system we are also going to grab images from the video, so that further processing is to be done.

Various problems related to images are needed to be solved. Common image problems are unstable brightness, poor resolution, brightness [2].

Various image processing techniques as shown in fig. 3(b) for the implementation touchless writer and accessing media player are used. There are different modules and mechanism that will carry out this image processing activities. One of them is blurring mechanism [7], which makes each and every frame smooth and soft. This mechanism is used to reduce sharp points, edges and corners.
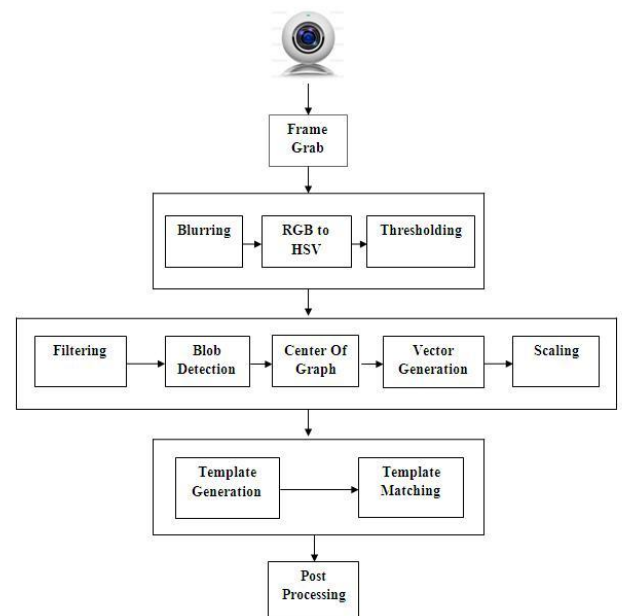

Fig. 3(b): System Design.

HSV color system must be used in this system. Therefore we need to convert RGB to HSV. This conversion will enable us to change the color space of images from RGB to HSV. It is necessary because HSV color space separates image intensity from color information. Here in this system we require only color information not the intensity of that color.

Another important concept is thresholding [5] which generates binary image (B/W) and it is also known as grayscale image.
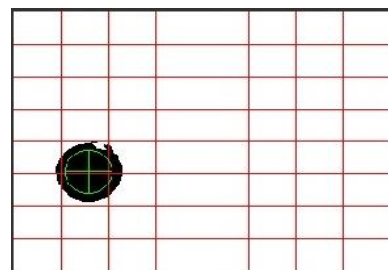

Fig.3(c): Thresholded Feed.

After this preprocessing and filtering (removing noise) of images, actual gesture generation and recognition process starts. After filtering, images are passed for Blob detection [4], in which consists of extracting only necessary part from the image. One main reason of this is to provide complementary information about regions, which is not obtained from edge detectors or corner detectors. In early work in this area, blob detection was used to obtain regions of interest for further processing. These regions could signal the presence of objects

or parts of objects in the image domain with application to object recognition or object tracking.

After extraction, blob is used as a main object for gesture generation. By deriving COG(centre of graph) [4] of that blob, vector generation activity is carried out according to user gestures. Vector generation is nothing but identification of gesture path that has been implemented by user. This generated vector will used by the system for template generation process. After this stage scaling is performed Scaling is a non-trivial process that involves a trade-off between efficiency, smoothness and sharpness.[6]

After all the processing, template of gesture is generated by the system. When that template will be matched with any template in the database result will be displayed.
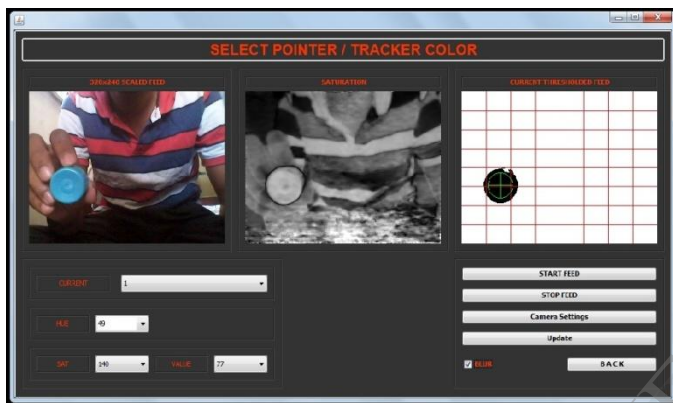

Fig 3(d): System Features

IV. MATHEMATICAL MODEL

Let S= {I, f}

Where I = Set of input Images
I= {i1,i2,i3…..in}

Input is taken in the video format, and it is broken into images so that the required processing will become easier

And F= Set of Filters
F= {fBlur, fHsv, fThresh, fBlob, fCog, fTempGen, fTempMatch}

Filters are nothing but the various processing techniques that are going to be applied on the SET OF IMAGES obtained from the video input.

fBlur =

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad [1]$$

(1)

where x is the distance from the origin in the horizontal axis

y is the distance from the origin in the vertical axis

and σ is the standard deviation of the Gaussian distribution.

Each and every image is scanned form left side top corner to right side bottom corner like RASTER SCAN method.

$$N \begin{cases} R= R/25 \\ G= G/25 \\ B= B/25 \end{cases}$$

(2)

25 is used to denote group of pixels on which blurring is applied. This size is not fixed

rgbmax= max(R', G', B')
rgbmin= min(R', G', B')

(3)

$$\{iH, iS, iV\} \longleftarrow fHsv(iBlur)$$

$$H = \begin{cases} 0 + 43*(g-b)/(rgbMax-rgbMin), & Cmax = R' \\ 85 + 43*(b-r)/(rgbMax-rgbMin), & Cmax = G' \\ 171 + 43*(r-g)/(rgbMax-rgbMin), & Cmax = B' \end{cases}$$

(4)

$$S = \begin{cases} 0 , & v = 0 \\ 255 * (rgbMax-rgbMin)/v, & v<>0 \end{cases}$$

(5)

V = rgbmax

(6)

$$iThresh \longleftarrow fThresh\{iH, iS, iV\}$$

Where T is a specific value known as THRESHOLD value, indicating specific intensity. If the intensity of a single pixel is above threshold value then the pixel value is set to 1otherwise it is set to 0.

$$iBlob \longleftarrow fBolb (iThresh)$$

L(x,y;t) = g(x, y, t) * f(x, y). To automatically capture blobs of different (unknown) size in the image domain

$$iTempGen \longleftarrow fTempGen\{ (x1, y1),(x2, y2)…………………(xi, yi) \}$$

Template is generated using the pixels which are visited by blob centre.

Template Matching is done for recognition of gesture.

As we are applying these set of functions on set of 'n' images obtained from video capture, time complexity of each function is O(n).

These are tractable problems and used in real time. Hence they are p-complete problems and time complexity of the system is also O(n).

V.   IMPLEMENTATION METHODOLOGY

The implementation of the system is as follows:

Step 1: Webcam Test

This is an initial module which checks if webcam is working properly.

Step 2: Pointer Selection

This is a mandatory step in order to successfully access both the applications of touchless writer and accessing media player. Here, users must select an object having a color so that it enables users to draw gestures in other modules. This module is shown in fig. 3(c).

Step 3: Letter Training

This module is a part of touchless writer application. Users can train the system any gesture pattern which they like for particular letter and add to database as shown in the fig. 5(a)
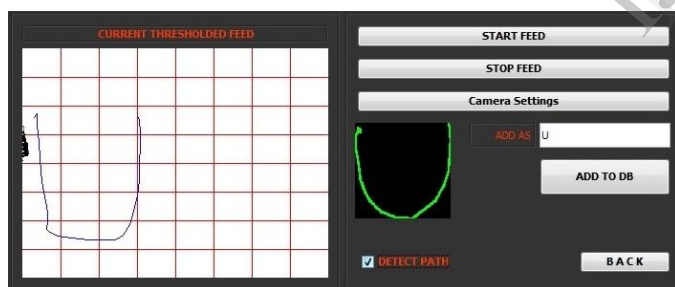

Fig. 5(a): Letter Training Module.

Step 4: Manage patterns

If user finds out that any trained letter is wrong or needs to be changed then manage patterns module will allow users to do the same.

Step 5: Touchless writer application

Users will draw a gesture and if the pattern is matched with trained pattern then that letter will be displayed in notepad as shown in fig. 5(b).
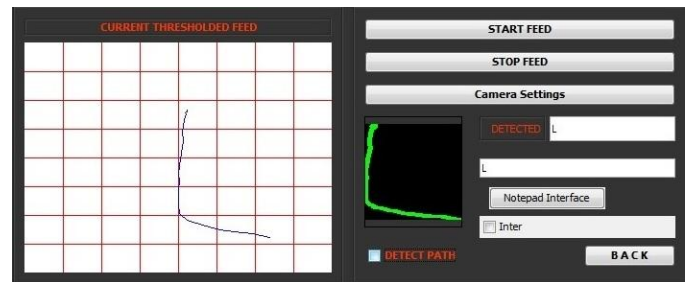

Fig. 5(b): Touchless Writer Application.

Step 6: Media player application

By performing predefined gestures in this module users can make use of various media player functions. In this system various gestures are assigned for increasing and decreasing volume, playing or pausing a video/audio, playing next or previous track as shown in fig. 5(c)
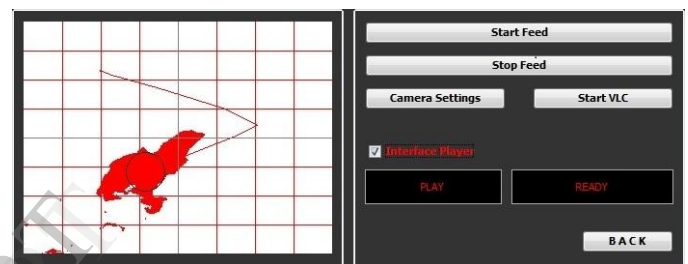

Fig. 5(c): Media Interface Module.

VI.   CONCLUSION

In this paper we have discussed importance of human computer interaction. We provide solution to the disadvantages from previous systems. We also presented implementation of Touchless writer and accessing media player. In future gesture recgnition system can be used to generate touchless fingerprints. More accuracy can be obtained in gesture recognition system.

VII.   FUTURE WORK

We are developing a system which is strongly related and dependent on image processing. For this system we are going to use color markers so that the modules in our proposed system will be implemented. There are several studies that should be carried on in future work. Skin detection model: There is strong need of skin detection model so that user will not have to depend on any type of color markers. This system can be implemented in various technologies e.g. devices and applications that do not have traditional mouse and keyboard interface, such as TabletPCs, media-center PCs hand-held computers, home appliances, and video games. In these devices, mouse and keyboard may seem to be clumsy and may require more effort.

VIII.   ACKNOWLEDGMENTS

IX.   REFERENCES

[1]   Astitwa Rathore, Manuj Paliwal, Gaurav Sharma, Dina Nath, Himanshu Mishra, Soumik Mondal, "A Dynamic hand gesture recognition system for controlling VLC media player" Advances in technology and engineering (ICATE)2013 International conference IEEE Jan 2013.

[2]   Ching-HaoLai,"A Fast Gesture Recognition Scheme for Real-Time Human-Machine Interaction Systems", IEEE, Smart Network System Institute, institute for Information Industry, Taipei City, Taiwan, and sep-2011

[3]   Yang Wu & Lu Yu," Touchless Writer: Object Tracking & Neural Network Recognition" The Milton W. Holcombe Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29631, 2012

[4]   Uma Sahu, Ditty Varghese, Gayatri Gole, Melanie Fernandes, Pratiksha Mishra, "Hand Cursor Implementation Using Image Processing & Sixth Sense Technology", International Journal of Infinite Innovations in Technology|ISSN:2278-9057   IJIIT|Volume-I|Issue-III|2012-2013 January.

[5]   Anoko moonlight, Braksus, BrotherE, Darth Sidious, Dawood majoka, "Thresholding   (image processing)",Source:http://en.wikipedia.org/w/index.php?oldid=451470 153,Nov 2007.

[6]   Andreas Kaufmann, ArroLu, Berland, Chrisballinger, Culix, Damian Yerrick,"   Image scaling",source:http://en.wikipedia.org/w/index.php?oldid=502249939, Feb 2011.

[7]   Axel Boldt, BenFrantz Dale, Darel Rex, "Gaussian blur", *Source*: http://en.wikipedia.org/w/index.php?oldid=436139840,June 2009.