ISSN: 2277-9655
Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# Cursor Movement in Real Time using Hand Gesture Recognition

**T. Suganya Thevi[*1], C. Chitra[2]**

[*1] Department of ECE, PSNA College of Engineering and Technology, Dindigul, India

[2] Assistant Professor, Department of ECE, PSNA College of Engineering and Technology, Dindigul, India

suganyaathevi@gmail.com

## Abstract

This paper proposes a method for cursor movements by using hand gestures recognized by Cascade Haar Classifiers. Purpose of this paper is to reduce the external hardware and to interact with the computer. The system consists of two modules namely, gesture recognition and control of cursor movements. Gesture recognition is attained using neural net training to the cascade Haar classifiers. Supervised neural net training helps in removing non-gesture pattern that helps to classify a frame as a gesture. The frame from the video recorded using web cam is given as input. Based on the extracted features, from the frame the cursor movements are controlled. This method improves the computational speed and also the effectiveness of communication. The tool used here is OpenCV 1.0 under Visual Studio 2008.

**Keywords**: Control of Cursor Movements, Haar Classifiers, Hand Gestures, Supervised neural net training.

## Introduction

Human Computer Interaction (HCI) is a domain where the interaction between human and computer is designed in an efficient manner. The conventional HCI devices such as keyboard and mouse are not enough for interactive applications since they essentially affect the effectiveness of communication with computers [1].

To achieve more efficient communication, human hand could be considered as an input device. The use of hand gestures provides an attractive alternative to cumbersome interface devices for human computer interaction. Hand gestures are defined to be a nonverbal form of communication in which visible bodily actions are used to communicate particular messages as in Fig.1.

**[a] [b]**

**Fig.1.Various Gestures for Control of cursor movements.**

This proposed system recognizes the hand gestures by analyzing the group of pixels in each frame that reduces time delay to a greater extent whereas the existing systems deal with gesture recognition using pixel wise analysis which results in a great delay.

Hand gesture recognition techniques are divided into two categories: appearance based approaches and 3D- hand model based approaches [2]. To get the region of interest, appearance based approaches use image features whereas 3D hand model approaches use 3D kinematic hand model. The region of interest is detected based on 2D appearance projected by the 3D hand model. Considering the real time performance as one of the important requirements, appearance based approach is quite easier. This is because, 3D hand models necessitates for the creation of a large database as a single 3D model can project many number of 2D images which is computationally expensive. Thus for gesture recognition, appearance based approach suits better.

## Related Works

The idea of using gestures for computer interactions was first applied by Ivan Sutherland [3], who demonstrated Sketchpad, an early form of stroke based gestures using a light pen to control graphical objects on a tablet display. The application of gestures in this way was recognized widely. Next to this, loads of researches and findings were done by many scientists.

ISSN: 2277-9655
Impact Factor: 1.852

C.W. Ng et al. in [4] proposed a system based on vision to detect 14 gestures to control windows and objects within a graphical interface.

A system based on the analysis of image property was proposed by Abe et al. [5] which recognizes hand gestures through the detection of the bending of the hand's five fingers.

One of the important findings is the use of glove based devices and markers instead of mouse. However, the gloves have wires attached to them which makes the users to feel uncomfortable and is also expensive. As computer systems have various size of monitors, touch screens cannot be opted for computer systems as they are very expensive (20" screen costs around 600$) and needs high maintenance. Hand gesture classification has become realistic due to the various advances in technical society. The requirements of the proposed system are cost effectiveness, real time performance, accuracy and robustness.

## Gesture Recognition

The proposed system considers only single handed gestures. A gesture is a composite action constructed by series of hand postures that act as transition states.

**Cascade Haar Classifiers**

Cascade Haar Classifiers are used for Gesture Recognition. Paul Viola and Michael Jones gave simple rectangular features called Haar-like features[6] as shown in Fig.2. They are based on Haar Wavelets. Haar classifiers consider adjacent rectangular regions of the image sub region. Presence of a Haar Feature is determined by subtracting the average dark region pixel value from average light region pixel value. If this difference is above the threshold (set during training), the desired feature is said to be present. Haar Classifiers are more advantageous than pixel wise analysis because it uses the adhoc knowledge of the input frame. Papageorgiou et al gave simple Haar basis functions. There are three most commonly used features namely, two-rectangle features, three-rectangle features and four-rectangle features.

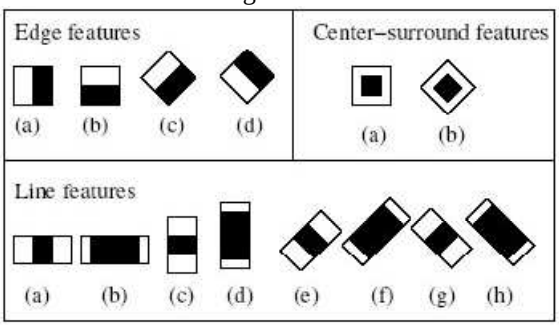



**Fig.2. A set of Haar-like features**

The accuracy of gesture recognition is poor when a single Haar feature is used. Cascade of classifiers could be used to improve the accuracy which is shown in Fig.3. High detection rates and radical reduction in the computation time are key features of cascade classifiers. If the first classifier gives a positive result then the particular sub-window is passed for the evaluation of a second classifier. The second classifier has to be tuned to get very high detection rates. A positive result from the second classifier causes the third classifier to do the evaluation, and so on. If a sub-window is determined to be a negative window at a particular stage, then the window is directed for rejection and there is no further processing in that sub window.

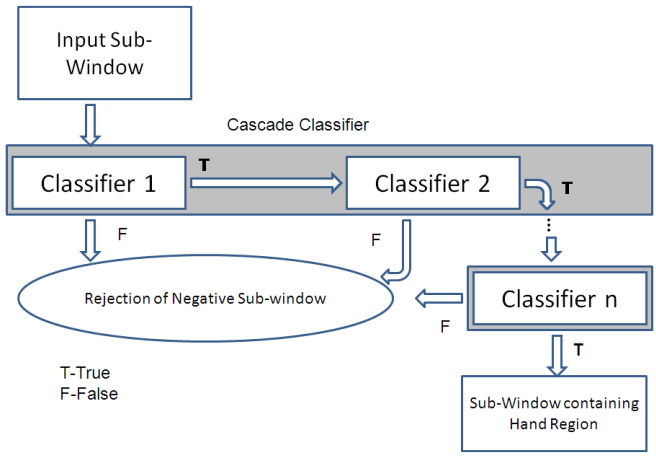



**Fig.3. Cascade Haar Classifiers**

**Haar Classifier Training**

The different stages in the cascade are constructed by using supervised learning technique. The supervised learning method (see Fig.4.) is used to improve the accuracy and achieve the real time performance. Supervised learning can be implemented using many approaches. The one used here is the Adaboost algorithm [7]. Adaboost algorithm constructs a strong classifier by selecting the best features in each step. High detection rates could be obtained, if the classifiers are with more features but the required computation time is high. Thus, the number of classifier stages, the number of features in each stage and the threshold used should be decided carefully for getting the desired detection rate. The samples that had passed in previous stages are used to train the classifier at subsequent stages. As many sub-window regions of the frame are negative, they are eliminated rapidly by cascade classifier and accurate detection happens. The proposed system of using Haar- like features along with nueral net training is approximately 15 times faster than any previous approaches[8].

ISSN: 2277-9655
Impact Factor: 1.852

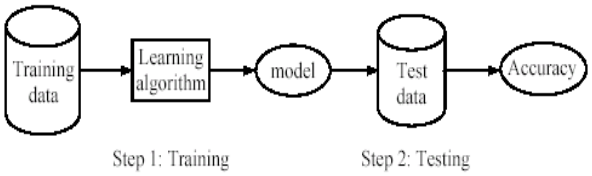



**Fig.4.Supervised learning method**

**Integral Image**

Each frame has hundreds and thousands of features in it. A technique to estimate the image features present in each frame is the Integral image (Fig 5.). A rich set of Haar-like features are computed using the integral image. This reduces the processing time significantly.

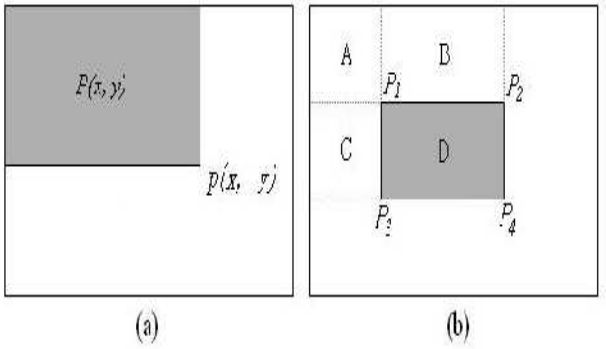



**Fig.5. Integral Image**

Let us consider an image sub-window as shown in Figure 5 (a). In this sub-window at any location (x, y), the value of the pixel contains the sum of the pixel values above and left of this pixel inclusive. This could be written as:

$$P(x,y) = \sum p(x', y')$$
$$x' \leq x, y' \leq y$$

For instance, the sum of the pixel value within the area *D* in Figure 5 (b) can be figured out by:

$$P_1+P_4-P_2-P_3$$

Where P1 = A, P2 = A+B, P3 = A+C, and
P4 = A+B+C+D.

## Control of Cursor Movements

Once the hand region in a frame is detected by using the cascade classifier the cursor movement is controlled by recognizing the desired gestures. The methodology used in this paper is as follows.

*A. Methodology:*

The proposed method involves the following steps:

*Step 1*: Negative samples, which do not contain the region of interest (Hand region) are collected. Sets of videos of 24 frames each from different users with the required gestures are also collected. These samples are called as positive samples.

*Step 2:* The frames are extracted from the videos. The video should be captured in uniform background. A new image is built, that is obtained by finding the difference between the image of the hand and the uniform background. The absolute difference value of image is calculated. The resultant image then consist of only the hand of user.

*Step 3:* From video, the images are in the form of uncompressed RGB image. A binary image is formed by converting the RGB image to grayscale image which eliminates the hue and saturation information while retaining the luminance factor. The noise present in the binary image is removed.

*Step 4*: As the region of interest is hand region, the edge pixels of the hand is calculated using canny edge detection algorithm. As the user may place the hand at a different distances every time, the images are to be resized to a size of 30 by 30.

*Step 5*: The Haar cascade classifiers are trained by using the obtained images for gesture classification. When the training is completed, testing is done by giving the input images. The trained images are compared to the input images to detect the region of interest and the cursor pointer is fixed to the region of interest to perform the assigned function.

The flowchart for the proposed system is shown in the Fig.6.

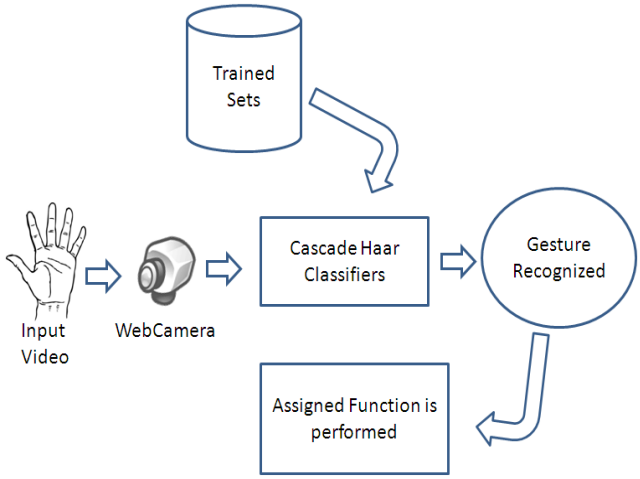



**Fig.6. System Overview**

## Experiments and Discussions

The software used in our paper is Visual Studio 2008 and the tool used is OpenCV 1.0. The basic simulations carried so far are -

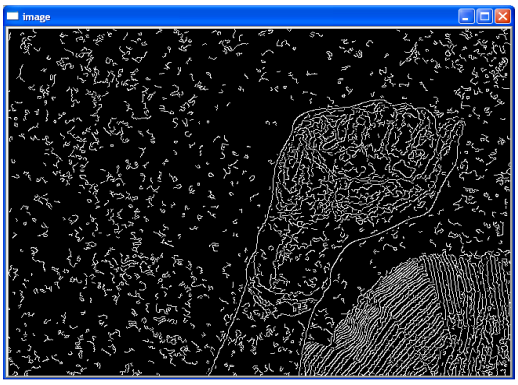
1. Converting a RGB image into gray scale image.

ISSN: 2277-9655
Impact Factor: 1.852

2. Smoothing an image using Gaussian and Median filter.
3. Edge detection of image using canny algorithm.
4. Capturing video using web cam (24 frames per second) in AVI format.
5. Recognition of Hand Gestures using Haar Cascade Clasifiers.

The outputs simulated so far are shown in the following figure:


**(a) Input Image**




**(b) Edge Detected Output**


**(c) Hand Region Recognized**

**Fig.7. Output Screen Shots**

*A. Future Work:*

All videos using web camera were taken under uniform background and normal illumination. Hand is the target region to be recognized. The positive samples are collected from different users. The Haar Classifiers are to be trained using these positive samples. The trained classifiers are then used for controlling the cursor. The simulations for training and testing are ongoing.

## Conclusion

The proposed system recognizes hand gestures with a single web camera as the input device. The first module is focused on the gesture recognition with cascade Haar classifiers and the Adaboost algorithm. The Haar- like features can effectively describe the hand gesture pattern with the computation of "integral image". The Adaboost algorithm speeds up the performance. The recognized gesture is then mapped to the trained image to perform the assigned function in the second module. This proposed system greatly reduces the delay for hand detection and thus improves Human Computer Interaction.

## *References*

*[1] Qing Chen, Nicolas D. Georganas, Emil M. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," Instrumentation and Measurement Technology Conference – IMTC 2007 Warsaw, Poland, May 1-3, 2007*

*[2] H. Zhou, T. S. Huang, "Tracking articulated hand motion with Eigen dynamics analysis," Proc. of International Conference on Computer Vision, Vol. 2, pp. 1102-1109, 2003.*

*[3] Sutherland, I. E. Sketchpad: A man-machine graphical communication system. In: Proceedings of the AFIPS Spring Joint Computer Conference 23. pp. 329–346,1963*

*[4] C. W. Ng and S. Ranganath,: Real-time gesture recognition system and application. Image Vis. Comput., vol. 20, no. 13–14, pp. 993–1007,2002*

*[5] K. Abe, H. Saito, and S. Ozawa,: Virtual 3-D interface system via hand motion recognition from two cameras. IEEE Trans. Syst., Man, Cybern. A, vol. 32, no. 4, pp. 536–540,2002*

*[6] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features," IEEE CVPR01, pp. 511-518, 2001.*

*[7] G. Bradski, A. Kaehler, V. Pisarersky, "Learning-based computer vision with Intels*

ISSN: 2277-9655
Impact Factor: 1.852

*open source computer vision library," Intel Technology Journal, Vol. 9, No. 2, pp. 119-130, 2005*

*[8] P. Viola, M. Jones, "Robust real-time object detection," Cambridge Research Laboratory Technical Report Series CRL2001 /01, pp. 1-24, 2001.*