2015 International Conference on Industrial Instrumentation and Control (ICIC)
College of Engineering Pune, India. May 28-30, 2015

Real time Finger Tracking and Contour Detection for Gesture Recognition using OpenCV

Ruchi Manish Gurav
P.G. Scholar, Dept. of Electronics and Telecommunication Engineering
VPCOE, Baramati, India
30ruchigurav@gmail.com

Premanand K. Kadbe
Assistant professor, Dept. of Electronics and Telecommunication Engineering
VPCOE, Baramati, India
be_prem@rediffmail.com

***Abstract*—Gestures are important for communicating information among the human. Nowadays new technologies of Human Computer Interaction (HCI) are being developed to deliver user's command to the robots. Users can interact with machines through hand, head, facial expressions, voice and touch. The objective of this paper is to use one of the important modes of interaction i.e. hand gestures to control the robot or for offices and household applications. Hand gesture detection algorithms are based on various machine learning methods such as neural networks, support vector machine, and Adaptive Boosting (AdaBoost). Among these methods, AdaBoost based hand-pose detectors are trained with a reduced Haar-like feature set to make the detector robust. The corresponding context-free grammar based proposed method gives effective real time performance with great accuracy and robustness for more than four hand gestures. Rectangles are creating some problem due to that we have also implement the alternate representation method for same gestures i.e. fingertip detection using convex hull algorithm.**

Keywords—Hand Gesture, Neural networks, Support vector machine, and Adaptive Boosting (AdaBoost), Reduced Haar-like feature set, Fingertip as contour detection, Convex Hull Algorithm, Human Computer Interaction (HCI), HSV (Hue, Saturation & Intensity Value)

I. INTRODUCTION

In recent years, computer vision development has great advancements and our day to day life tasks are incomplete without using computers. The major input devices like Keyboard and mouse are used to interacting with computers. Among the various interaction techniques use of hands as an input is an attractive method for establishing natural Human Computer Interaction. By using Hand gestures user can communicate more information in less time period. So for improving the interface between users and computers human computers interaction (HCI) technology has great utilization [1].

The primary goal of proposed system is to identify specific human gestures and we can use it to convey information or we can control any device or robot for offices and household application. The static pose of hand configuration may be defined as a posture. And a dynamic pose may be defined as a gesture that means physical movement of body organs such as hands, arms, face to convey meaningful information e.g. bye. For a successful communication, a sender and a receiver must have the similar information for a particular gesture.

There are two approaches which are commonly used to interpret gestures for Human Computer Interaction, which are specified as below:

A. Data Gloves based Method :

In this method user had to wear gloves, helmet & other heavy apparatus. For detecting hand gesture some optical or mechanical sensors, actuator & accelerometer are attached with the glove [2]. That equipment converts finger flexions into electrical signals for determining the hand posture. In this approach user had to carry a load of cables which were difficult to manage in real time environment. This method demands more maintenance due to the complex wired structures which is shown in figure 1.

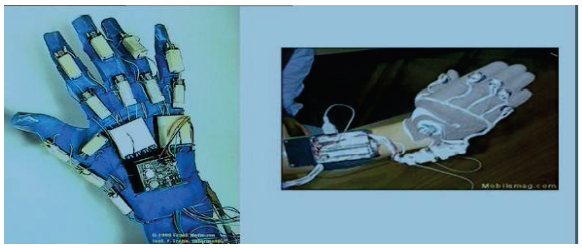


Fig 1: Data gloves based hand gesture recognition

B. Vision Based Method :

Recent trends of Computer vision techniques are easy, natural and less cost comparing[2]. Proposed method extracts the feature from the video frame. Today most of the laptops has an integrated webcam along with it so it is an easily available device. In our work, we are implementing a hand gesture recognizer capable of detecting a moving hand and recognize with its gesture. Here we are using integrated web camera of laptop to capturing image frame which is shown in figure 2.

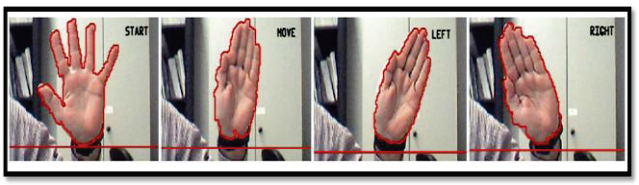


Fig 2: Hand postures captured by web camera

The paper organization is as follows: Section II highlights the various aspects of algorithms used to detect the hand gestures; Section III introduces the training method of individual hand-pose detectors using the reduced Haar-like feature set & AdaBoost and Convex Hull algorithm for finger tip & contour detection; Section IV discusses the introduction & application of software to be used and how we utilize that for our project; Section V provides the simulation result of proposed algorithms; Section VI presents gesture challenges of recent hand gesture recognition systems, Summary and conclusion.

II. LITERATURE SURVEY

First, In vision based approach, there are various techniques used for hand detection, training the gestures, background subtraction and finger tip detection which are reviewed as below :

The feature based hand detection techniques used by Viola and Jones detector and scale invariant feature transform [3][7] based hand detection have been implemented. These algorithms provide result with high accuracy but these are more sensitive to background. The second approach is image segmentation which uses HSV color space model rather than RGB color space to determine the color of human skin. This algorithm gives batter result for background separation and region boundary but it can't detect the object of skin color with similar color background [3]. The third approach is learning based gesture recognition in Adaptive Boosting algorithm that can integrate the information of same category of objects. It trains the network by combining all weak classifiers into one strong classifier. The AdaBoost learning algorithm [6] selects the best weak classifier from a set of positive and negative image samples. This algorithm provides result with better accuracy and fast speed but sometimes training period is more to train the network.

Another approach is for finding convex hulls. There are so many algorithms available for palm detection. In this section some of existing algorithms will be discussed which are used in our proposed technique. Graham's Scan Algorithm, Divide and Conquer algorithm, Jarvis's March or Gift wrapping Algorithm, Quick hull algorithm and Chan's algorithm. Graham Scan computes the convex hull of any given set of points. To implement the system for hand tracking and simple gesture recognition in real time, there is no need to touch or carry a peripheral device by user. By comparative analysis, we can conclude that only one detection technique not enough because different kind of methods can deal with different problem during detection & recognition. There are various available machine learning algorithms that are AdaBoost, support vector machine technique, hidden markov model, & principle component analysis for training classifiers [8]. There may also have different convex hull and contour detection of boundary of hand region.

Based on all these methods, we are going to implement the system by using the Adaptive boosting for hand detection and haar classifier algorithm to train the classifier. Here we also use HSV color model for background subtraction & noise removal, convex hull algorithm for drawing contour around palm and finger tip detection. There are multiple soft ware's for image processing application but among them OpenCV (Open Source for Computer Vision) software is very popular for Real time image processing applications such as object detection & gesture recognition. The major advantage is that one can easily integrate the code with hardware. We implement the proposed system on OpenCV library based on Linux environment.

III. METHODOLOGY

In this paper we started to implement a web camera initialization for live frame streaming. We also explain our three algorithms for unique purpose.

The simple Haar-like features (which are computed similarly to the coefficients in the Haar wavelet transform) are used in the Viola and Jones algorithm. The Haar-like features are robust to noise and various lighting condition because they compute the gray-level difference between the white and black area of rectangles. The noise and lighting variations affect the pixel values on the whole feature area. The integral image at the location of pixel [x, y] contains the sum of the pixel intensity values located directly above the pixel location [x,y] and at the left side of this pixel. So A[x,y] is the original image and AI[x,y] is the integral image that is calculated by below equation 1:

$$AI[x, y] = \sum_{x' \le x,\ y' \le y} A[x', y'] \qquad (1)$$

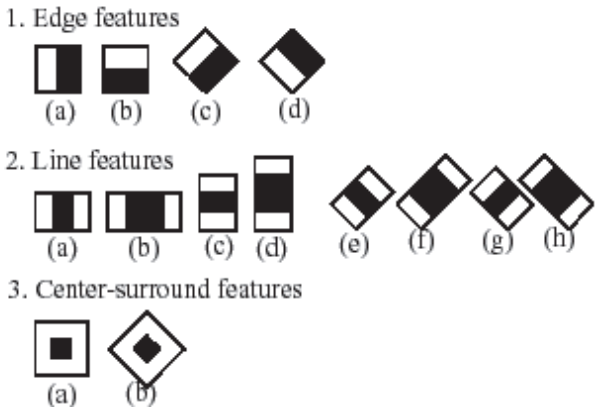


Fig 3: Common haar like features for integral images[5]

Practically no single Haar-like feature can identify the object with high accuracy. However, it is not difficult to find out one Haar-like feature-based classifier that has better accuracy than the random guessing.

The AdaBoost based learning algorithm improves stage by stage overall accuracy, by using a linear combination of these individually weak classifiers [4]. The AdaBoost learning algorithm initially assigns an equal weight to each training sample. We start with the selection of a Haar-like feature-based classifier for the first stage and got better than 50% classification accuracy. In next step this classifier is added to

the linear combination with the strength that is proportional to the resulting accuracy. So the training sample weights are updated i.e. training samples that are missed by the previous classifier are boosted in accordance. The next classification stage must achieve better accuracy for these misclassified training samples so that the error can be reduced. By this procedure we can improve the overall classification accuracy at further stage. The iteration goes on by adding new classifiers to the linear combination until the overall accuracy meets to the required level. At the final level the result is a strong classifier composed of a cascade of the selected week classifiers.

In Convex Hull algorithm first step is segmentation of the hand image that contains the hand to be located. In order to make this process it is possible to use shapes, but they can be changed greatly in interval that hand moves naturally. So, we select skin-color to get characteristic of hand. The skin-color is a distinctive cue of hands and it is invariant to scale and rotation [8]. In the next step we use the estimated hand state to extract several hand features to define a deterministic process of finger recognition. After the hand is segmented from the background, a counter is extracted. The counter vector contains the series of coordinates of edges of hand. Then the processing of counter vector gives the location of the finger tip.

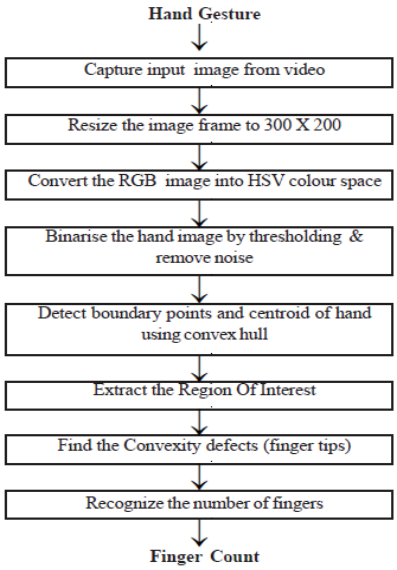


Fig 4: Flow of Methodology for finger counter

In convex hull implementation firstly calculate the points with minimum and maximum x and y-coordinates and by joining these points a bounding rectangle is defined, within which the hull is contained. There will be other points of convexity too; we find the convex defects i.e. between each arm of hull. The defect points are most likely to be the center of the finger valleys [8]. Then find out the average of all these defects which is definitely bound to be in the center of palm, but it is very rough estimate so average out and finds this rough palm center. Thus the radius of palm is an indication of depth of the palm using radius. The ratio of palm radius to the length of the finger triangle should be more or less same, thus we can find out number of location of tip of finger. Here AdaBoost algorithm and Haar like feature set algorithm are adopted for hand detection and recognition. But when we combine every gesture in same program, it get mixed and create confusion for real world hardware. So we have changed the representation way of rectangle & do it with fingertip by connecting lines for clear visibility. This experiment we developed on open source library for computer vision application called Open Computer Vision Library (OpenCV).

IV. Software Implementation

The OpenCV library is basically used in HCI, robotics, biometrics, image processing and other areas where visualization is important and includes an implementation of Haar classifier detection and training [4][5]. To train our own classifiers, two set of images are required. One set which referred as the negative images that contains an image or scene which does not contain the interested object. In this case the hand gesture feature that look exactly like the positive one, except those containing same hand gesture. The other set of images are positive images where they look exactly like the object which we want to detect. It is also important that they should be different in lighting and background. In our implementation, two hand postures are tested the "palm" posture and the "fist" posture. The camera that was used for the video input is a low-cost Web camera of laptop. This Web camera provides video capture with a maximum resolution of 640 x 480 up to 15 frames/s. The experiments are implemented with natural fluorescent lighting conditions. We collected 480 and 420 positive samples with different scales for the palm posture and the fist posture respectively and 500 random images for negative samples. In order to increase the robustness of the final classifier, the original positive set of images needs to be represented by different human hand in different color and size. After all of the positive and negative samples are ready, we set the required false alarm rate at 1 x 10^{-6} to terminate the training process which means that the accuracy of the classifier will meet to the requirement.

V. Result & Conclusion

As a first step we try the face detection based on available database of OpenCV. Then for capturing live streaming of camera the initialization has been done. The two gesture detection like palm and fist by green rectangle which is trained by integral images. The second step is the extracted image gestures which are compared with stored positive-negative integral image dataset and perform finger tip tracking by contour detection. All this requires analyzing the entire image with all present grammar. Using a 2.40 GHz intel® core™ processor Linux based OpenCV image processing software & Qt Creator IDE is used to analyze a 640 x 480 image size, a frame rate of 30 frames per second has achieved.

Fig.5 shows the detection of Fist and Palm gesture respectively by using C++ programming language. Fig.6 shows various finger gestures detected by proposed algorithm

and also contain segmented image and Identified gestures. Fig.7 shows the Finger tracking evaluation by multiple experiment.

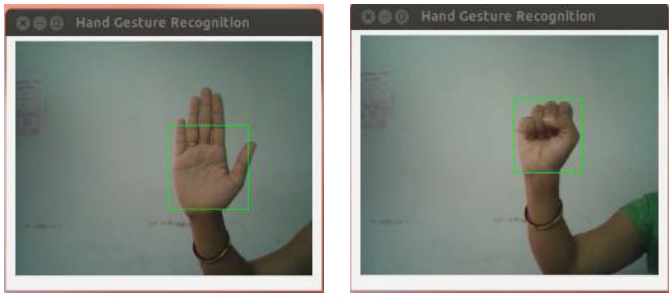
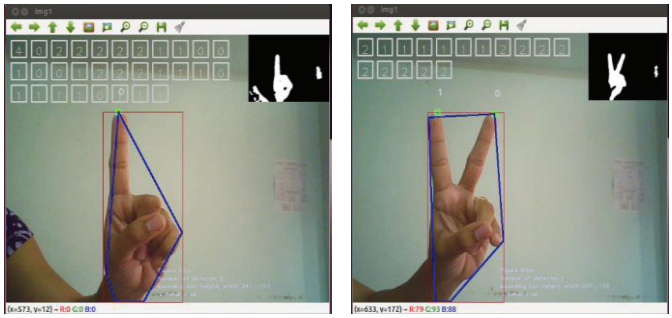
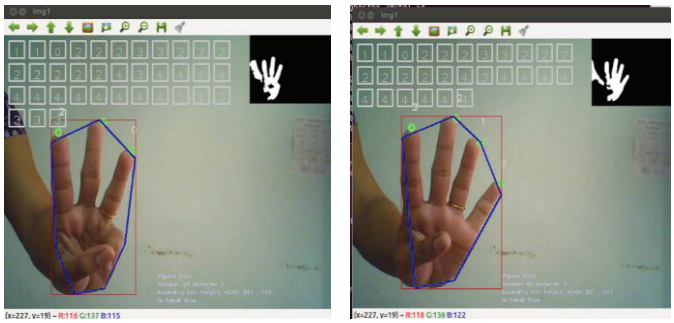


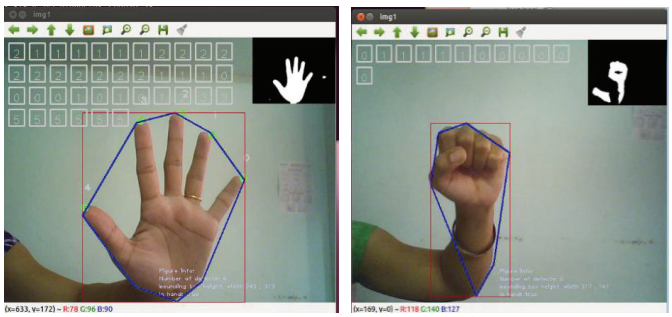
Fig 5: (a) "Palm" Detected (b)"Fist" Detected Gesture



(a) pointing finger (b) Two finger



(c) Three finger *(d) Four finger*



(e) Plam (f) Fist

Fig 6: The fingertip & contour detction using Convexhull algorithm

The aim of this project is to develop a real time Gesture recognition system. It is determined that contour is very important feature and can be used to discriminate two different gesture.

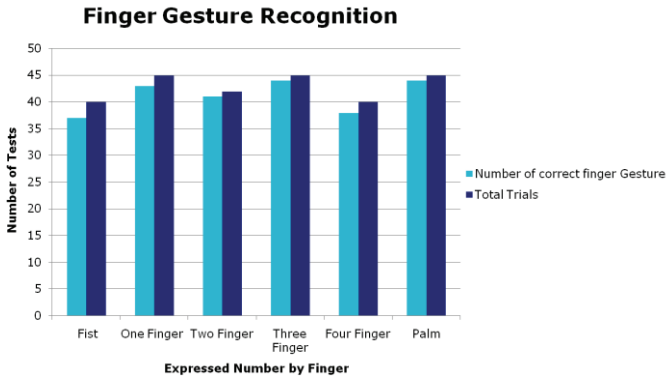


Fig 7: System's performance evaluation results

The processing steps to classify a gesture included gesture acquisition, segmentation, filtering, contour representation and classification using different techniques. The work was accomplished by training a set of feature set which is local contour sequence. The main advantage of Local contour sequence is that it is invariant to rotation, translation and scaling so it is a good feature to train the learning machine. We have achieved 92% accuracy with Convex Hull and 70% accuracy with AdaBoost.

References

[1] Radhika Bhatt, Nikita Fernandes, Archana Dhage "Vision Based Hand Gesture Recognition for Human Computer Interaction" University Of Mumbai (IJESIT) Volume 2, Issue 3, May 2013

[2] G. R. S. Murthy & R. S. Jadon "A Review Of Vision Based Hand Gestures Recognition" International Journal of Information Technology and Knowledge Management, July-December 2009, Volume 2, No. 2, pp. 405-410

[3] Qing Chen Nicolas, D. Georganas, and Emil M. Petriu "Hand Gesture Recognition Using Haar-Like Features And A Stochastic Context-Free Grammar" IEEE ,Vol. 57, No. 8, August 2008

[4] Soowoong Kim, Jae-Young Sim, And Seungjoon Yang "Vision-Based Cleaning Area Control For Cleaning Robots", IEEE Vol. 58, No. 2, May 2012

[5] Nasser H. Dardas And Nicolas D. Georganas "Real-Time Hand Gesture Detection And Recognition Using Bag-Of-Features And Support Vector Machine Techniques", IEEE Vol. 60, No. 11, November 2011

[6] Anupam Agrawal, Rohit Raj and Shubha Porwal "Vision-based Multimodal Human-Computer Interaction using Hand and Head Gestures" IEEE Conference on Information and Communication Technologies ICT 2013

[7] Kenji Oka and Yoichi Sato "Real-Time Fingertip Tracking and Gesture Recognition" IEEE proceeding on Computer Graphics and Applications Nov/Dec 2002