Vision Based Hand Gesture Recognition with Haar Classifier and AdaBoost Algorithm

Ruchi. M. Gurav
P.G. Scholar, Dept. of Electronics and Telecommunication *Engineering, VPCOE,Baramati, India*

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

Abstract- Gestures play an important role for communication among the human. In the recent years new technologies of Human Computer Interaction (HCI) are being developed to deliver user's command to the robots. Some of them are based on interaction 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 robot by using Computer Vision. Detection algorithms are based on the 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 two hand gestures.

Keywords – Hand Gesture, Gesture Recognition, reduced Haar-like feature set, AdaBoost, Human Computer Interaction (HCI).

I. INTRODUCTION

In recent years, computer vision development has great advancements so the essential aim of building hand gesture recognition system is to create a natural interaction between human and computer. Gestures are expressive, meaningful body motions involving the physical movements of the fingers, hands, arms, head, face or body with the intent of Conveying meaningful information or Interacting with the environment [2]. Gestures can be static or dynamic. A dynamic gesture is intended to change with respect to time whereas a static gesture is observed at fixed time. A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture. Gesture recognition is the process of recognizing and interpreting a continuous sequential stream gesture from the given set of input data but it's a complex process.

Human computer interaction (HCI) is the relation between the human and the computer. It is used for interacting with machine via various methods of interaction such as hand, head, facial expression, voice & touch [3]. For this project we are using hand gesture because it is very simple and natural body language. The key problem in gesture interaction is how to make hand gestures understood by computers. The approaches are mainly divided into Data-Glove based and Vision Based approaches. The Data-Glove based method sensor devices for digitizing hand, collect hand configuration and movement. However, these devices are quite expensive and difficult to manage in real time environment due to its complex wired structure.

Alternate Method is Vision Based, which is based on recent technology of Computer vision. It doesn't requires any gloves or other heavy apparatus. This method requires only web camera to extracts the video frame. Today most of the laptops have an integrated webcam along with it so it is an easily available device. This paper is organized as follows: In section II highlights a survey on vision based hand gesture recognition; Section III introduces the algorithms to be used for hand detection & gesture recognition; Section IV provides Software introduction, Implementation and results which are simulated and section V presents the conclusion and future work.

II. RELATED WORK

Most of the researchers classified gesture recognition system into mainly three steps after acquiring the input image from camera. These are: Extraction Method, features estimation and classification or recognition as shown in Figure 1.

Figure 1. Gesture recognition system steps.

A. *Extraction method and image pre-processing*

Segmentation is the first process for recognizing hand gestures. It is the process of dividing the input image into regions separated by boundaries. The segmentation process is depends on the type of gesture, if it is dynamic gesture then the hand gesture need to be located and tracked but if it is static gesture then input image have to be segmented only. The hand should be located firstly so a bounding box is used to specify the skin color, since it is easy and invariant to scale, translation, and rotation changes. In segmentation process the color space is used but color spaces are sensitive to lighting changes for these reason HSV color models are used [7]. This technique concentrates on the pigments of the pixel used to normalized R-G color space. Some preprocessing operations are applied such as background subtraction, edge detection and normalization to enhance the segmented hand image.

B. *Features Extraction*

Good segmentation process leads to perfect features extraction process and it play an important role in a successful recognition process. Features vector of the segmented image segmented image can be extracted in different ways according to particular application [4]. Various methods have been used for representing the features extraction. Some methods used the shape of the hand such as hand contour detection while others detect fingertips position or palm center.

C. *Gestures Classification*

After modeling and analysis of the input image, gesture classification method is used to recognize the gesture. Recognition process is affected by proper selection of features parameters and suitable classification algorithm [6]. For example edge detection or contour operators cannot be used for gesture recognition since many hand postures are generated and could produce misclassification. The statistical tools used for gesture classification are Hidden Markov Model (HMM) [4], Finite State Machine (FSM), Principal Component Analysis (PCA) [8] and Neural network has been widely applied for extraction the hand shape. Other soft computing tools are Fuzzy C Means clustering (FCM), Genetic Algorithms GAs. In vision based approach, we survey of all available techniques used for hand detection, gesture recognition. From that we selected HSV for Preprocessing, AdaBoost for hand detection and Haar classifier for training and reorganization purpose which has fast detection with good accuracy.

III. PROPOSED ALGORITHM

The methodological analysis has been presented in Figure 2. The work commence with capturing Hand gesture images using cameras. These images are feed to pre-processing steps like filtering,background subtraction and segmentation. Then different colour features & texture are extracted from extracted input image. Finally, the feature values are fed as input to the classifier to classify the given image.

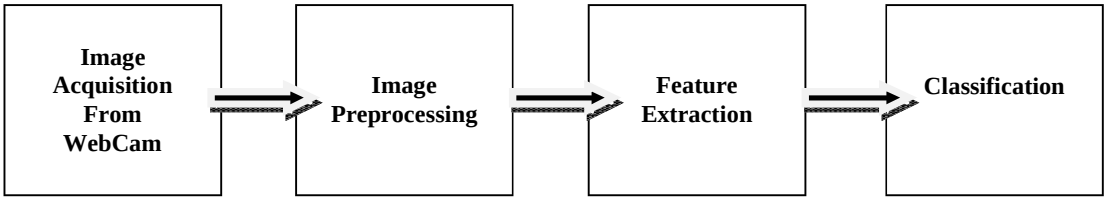

Figure 2 Block Diagram of hand gesture recognition system

A) Haar like feature algorithm –

Viola and Jones [5] proposed a statistical approach to manage the variety of human faces. This famous face detector architecture called Haar like feature to describe hand. This is very simple algorithm because rather than using the intensity values of a pixel, they use the change in contrast values between adjacent rectangular groups of pixels [3]. There are four different haar like feature such as Edge feature, Line feature, Center surround feature and Diagonal feature.

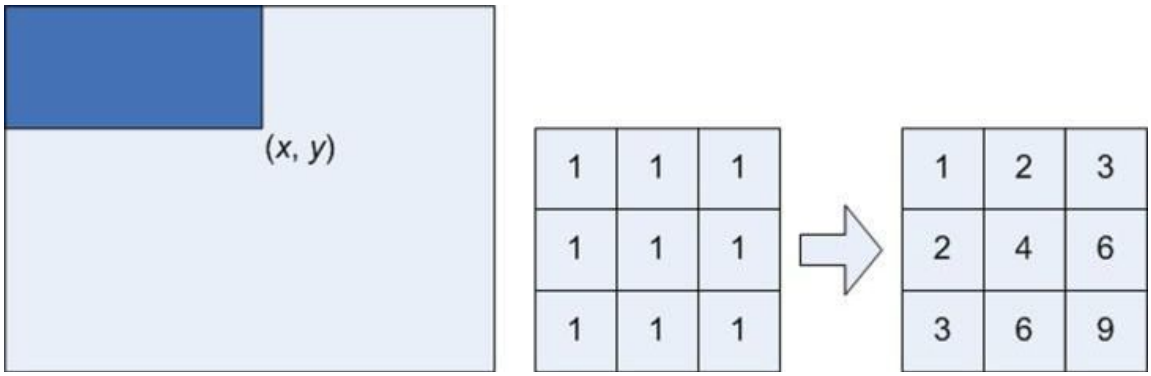

Figure 3. Integral image generation. The shaded region represents the sum of the pixels up to position (*x*, *y*) of the image. It shows a 3×3 image and its integral image representation.

There are two motivations for the employment of the Haar-like features rather than raw pixel values. The first is that the Haar-like features can encode ad hoc domain knowledge, which is difficult to describe finite quantity of training data. The Haar-like features describe the ratio between the dark and bright areas within a kernel [7]. One typical example is that the eye region on the human face is darker than the cheek region, and one Haar-like feature can efficiently catch that characteristic. The second motivation is that a Haar-like feature-based system can operate much faster than a pixel based system. Besides the above advantages, the Haar-like features are also relatively robust to noise and various lighting condition because they compute the gray-level difference between the white and black rectangles. The value of a Haar-like feature is the difference between the sums of the pixel values in the black and white rectangles i.e.,

$$F(x) = \sum_{\text{Black}} (\text{pixel value}) - \sum_{\text{white}} (\text{pixel value})$$

B) Adaptive Boosting algorithm-

The AdaBoost based learning algorithm improves stage by stage overall accuracy, by using a linear combination of these individually weak classifiers [5]. The AdaBoost learning algorithm initially assigns an equal weight to each training sample.

The concepts of weak learner and strong learner derived from PAC model, a weak learner is a classifier and slightly correlated with the labels. A strong learner is a well correlated with true labels. Random guessing would be like tossing a coin to get a probability accuracy of 50%. If there is a condition that can help the estimate to slightly improve the accuracy, we say this classifier is a weak classifier.

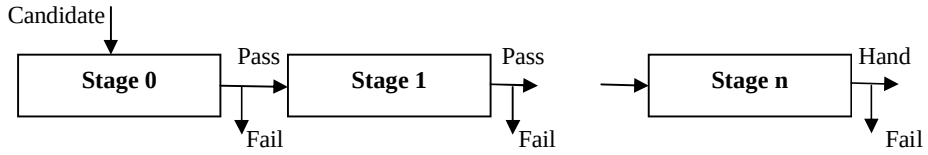

Figure 4. Cascade of stages. Candidate must pass all stages in the cascade to be concluded as a Hand.

At the same time, if there is another condition that can improve the accuracy greatly, it is a strong classifier, the procedure that gets this condition is considered a strong learner. The boosting algorithm is a machine learning algorithm and it can upgrade the weak learner to the strong learner, it brings a big benefit for machine learning. Boosting algorithm is useful for classification such as creation of models, image segmentation and data mining. AdaBoost's full name is Adaptive Boosting we can understand it is a improvement from Boosting algorithm. The difference with Boosting algorithms is that boosting algorithms needs to know the error rates lower limit. AdaBoost algorithm automatically adjusts the lower rate of error rate according to the weak learning's feedback. The AdaBoost algorithm does not need to know the details of weak learning and it will achieve the same efficiency as Boosting [6]. After the AdaBoost algorithm was proposed, it obtained lots of attention in machine learning field and it has been applied in many areas. No matter what kind of data is, AdaBoost algorithm is able to enhance the learning accuracy and very easily apply it in related areas in the real world. It effectively improves the accuracy of a given learning algorithm. The Adaboost learning algorithm is a variation of the regular boosting algorithm, and can adaptively select the best features at each step and combine a series of weak classifiers into a strong classifier.

IV. Experiment and Result

A. *Software Introduction*

Here this experiment is developed on open source library for computer vision application called Open Computer Vision Library (OpenCV). The OpenCV library is basically used in HCI, robotics, image processing, biometrics and other areas where visualization is important. This software is very popular for Real time image processing applications such as object detection & gesture recognition. Here OpenCV is adopted because there are several advantages and major advantage is that one can easily integrate the code with hardware. We implement the proposed system on OpenCV library based on Linux environment.

B. *Experiment*

One set of images referred as the negative images that contains an image or scene that does not contain the interested object, in this case of hand gesture feature, that look exactly like the positive ones, except that they don't contain the same hand gesture. The other set of images are positive images that are look exactly like the object which we want to detect. It is also important that they should be different in lighting and background condition. 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. 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 1×10^{-6} to terminate the training process, which means that the accuracy of the classifier will meet to the requirement.

C. *Result*

The first step of hand feature detection is detecting the shape and size of hand gestures which are trained by integral images. The second step is the extracted image gestures which are compared with stored positive-negative integral image dataset.

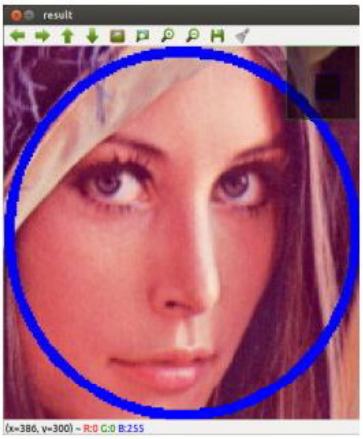

Figure 5: Face detection on available database of OpenCV

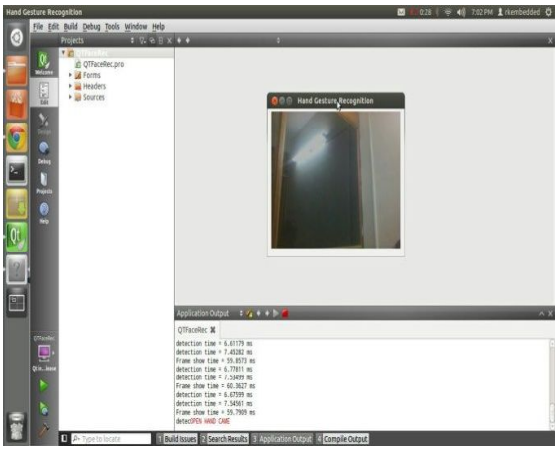

Figure 6: Initialization of web camera

This requires analyzing the entire image with all present grammar. Using a 2.40 GHz intel® core™ processor to analyze a 640 x 480 image, a frame rate of 3 frames per second has achieved.
Fig.5 shows detection of facial feature by available database of OpenCV, Fig.6 Shows the web camera initialization, Fig. 7 & 8 shows the detection of Fist and Palm gesture respectively by using C++ programming language.

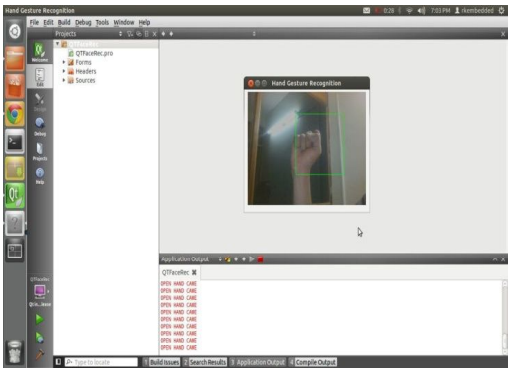

Figure 7: Detected "Fist" Gesture

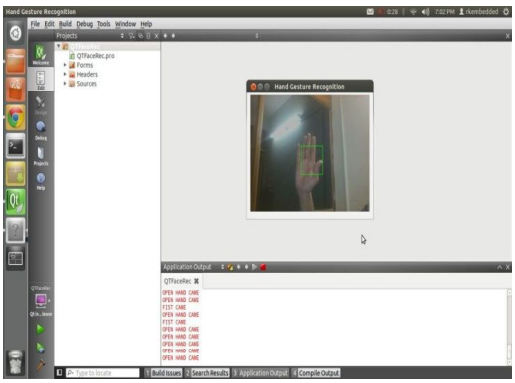

Figure 8: Detected "Palm" Gesture

V. CONCLUSION

We compared various algorithms to train the gestures of hand and selected haar like feature set with AdaBoost algorithm. By training the system, we can detect the gesture of live images. We have trained the system for two gestures and in future we will be training more gestures.We will also impliment the contour detection algorithm for fingertip detection for all gestures.
We can demonstrate that this Hand gesture detection and recognition, combined with other technologies, can produce effective and powerful applications.

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] Hardy Francke, Javier Ruiz-del-Solar and Rodrigo Verschae "Real-Time Hand Gesture Detection And Recognition Using Boosted Classifiers And Active Learning" Department Of Electrical Engineering, Universidad De Chile

[4] 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

[5] Phillip Ian Wilso, Dr. John Fernandez, "Facial Feature Detection Using Haar Classifiers" South Central Conference

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

[7] 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

[8] 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