Real Time Hand Gesture Recognition System

Shival Abhilasha Shamsundar, Suryawanshi Rupali Vitthalrao

Abstract- Hand gestures can be used for natural and intuitive human-computer interaction. Our new method combines existing techniques of skin color based ROI segmentation and Viola-Jones Haar-like feature based object detection, to optimize hand gesture recognition for mouse operation. A mouse operation has two parts, movement of cursor and clicking using the right or left mouse button. In this paper, color is used as a robust feature to first define a Region of Interest (ROI). Then within this ROI, hand postures are detected by using Haar-like features and AdaBoost learning algorithm. The Ada Boost learning algorithm significantly speeds up the performance and constructs an accurate cascaded classifier by combining a sequence of weak classifiers.

Index Terms— Human Computer Interaction, Hand Detection, Segmentation, Hand Tracking and Gesture Recognition..

I. INTRODUCTION

This since in recent years applications like Human-computer interaction (HCI) and robot vision have been active research areas. The conventional Human Computer Interaction devices such as keyboard, mouse, joysticks, roller-balls, touch screens, and electronic pens, are inadequate for latest Virtual Environment (VE) applications. These devices restrict the complete utilization of high performance hardware due to their limited input characteristics. The VE applications offer the opportunity to integrate various latest technologies to provide a more immersive user experience [1]. The main purpose of our implementation is to operate computer mouse actions by hand gestures using a low cost USB web camera. The computer mouse operations involve the motion of the cursor plus clicking operation. The proposed technique first uses robust skin color features for defining the ROL Then within this, ROI hand gestures are then recognized by using Viola Jones algorithm. Defining ROIs and searching within them significantly reduces computational time. Viola-Jones algorithm uses a set of Haar-like features, which processes a rectangular area of the image instead of a single pixel. To achieve real time performance, AdaBoost algorithm is used to automatically select the best features and build a cascaded classifier. Once gestures are recognized, and then they are assigned to different mouse events, such as right click, left clicks, undo [4], etc.

II. FLOW OF HAND GESTURE RECOGNITION SYSTEM

In this section, the flow of hand gesture recognition system algorithm is presented as shown in Fig.1.An effective vision based gesture recognition system for human computer interaction must accomplish two main tasks. First the position and orientation of the hand must be determined in each frame. Second, the hand and its pose must be recognized and classified to provide the interface with information on actions required i.e. the hand must be tracked within the work volume to give positioning information to the interface, and gestures must be recognized to present the meaning behind the movements to the interface. Due to the nature of the hand and its many degrees of freedom, these are not insignificant tasks. In order to accomplish these tasks, our gesture recognition system follows the following architecture as shown: The system architecture as shown in figure 1 uses an integrated approach for hand gesture recognition system. It recognizes static and dynamic hand gestures [1].

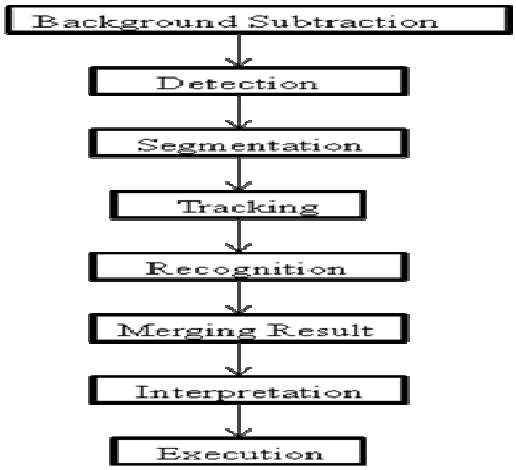

Fig.1: Flow of Hand Gesture Recognition System

III. HAND DETECTION USING HAAR LIKE FEATURE

During the image acquisition phase system extracts the static background and sits idle till user puts his hand in front of the corresponding camera. Once the hand is placed in front of camera it detects the hand using haar like features. The noise and lighting variations also strike the pixel measures on the entire characteristic region, which could be counteracted[3].The simple Haar-like features (so called because they are computed similarly to the coefficients in the Haar wavelet transform) are used in the Viola and Jones algorithm. The Haar-like features describe the ratio between the dark and bright areas within a kernel. 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. Each Haar-like feature consists of two or three connected “black” and “white” rectangles. Fig. shows the extended Haar-like features set that was proposed by Lienhart and Maydt [2]. The value of a Haar-like feature is the difference between the sums of the pixel values in the black and white rectangles [3].

Manuscript Received on May 15, 2015.

Prof. Shival Abhilasha Shamsundar, Electronics & Telecommunication Engg., Solapur University/ Brahmdevdada Mane Institute Of Technology, Solapur, India.

Prof. Suryawanshi Rupali Vitthalrao, Electronics & Telecommunication Engg., Solapur University/ SPM Poplytechnic, Solapur.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

$$f(x) = \sum_{\text{black}} (\text{pixel value}) - \sum_{\text{white}} (\text{pixel value})$$

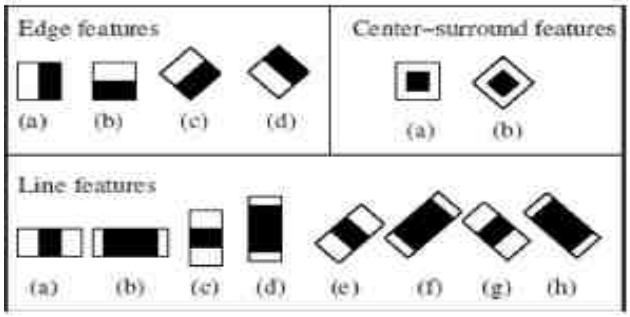

Fig.2: Extended set of Haar-like Feature

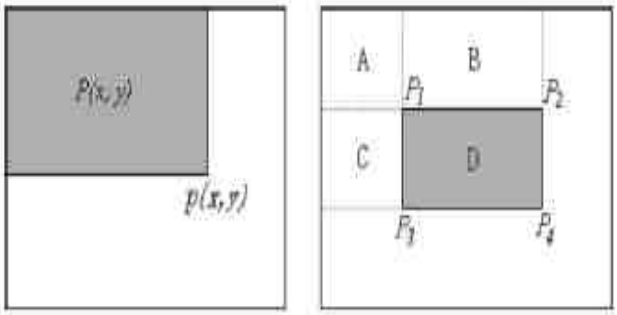

Fig.3: Integral Image

The "integral image" at the location of pixel(x, y) contains the sum of the pixel values above and left of this pixel, which is inclusive

$$P(x,y) = \sum_{x' \le x,\, y' \le y} p(x', y').$$

To detect an object of interest, the image is scanned by a subwindow containing a specific Haar-like feature based on each Haar-like feature *fj*, a correspondent weak classifier *hj*(*x*) is defined by

$$h_j(x) = \begin{cases} 1, & \text{if } p_j f_j(x) < p_j \theta_j \\ 0, & \text{otherwise} \end{cases}$$

Where x is a sub window, and θ is a threshold. *Pj* indicates the direction of the inequality sign.

IV. SEGMENTATION TECHNIQUES

IV.I Background Subtraction

Hand segmentation can be done with the help of background subtraction. In this method first, the image of our working background (without gesturer) is stored. Now the image frame (with gesturer) is subtracted from the previously stored background image plane. This gives the image of gesturer's body parts. Now we required to perform the operation for detecting face and segmented the moving hand. Disadvantage of this system is that if the lighting conditions change abruptly then there is a change in pixel value where the light intensity changed and additive noise contributes to the output [1]

The core of each motion detection system is the part of background subtraction that effectively extracts the correct shape of moving objects. In this paper, we have used the image of moving points detected by the temporal image analysis and the reference background image Bt evaluated at the time t. Radiometric similarity is used again for selecting from the moving points in those that are different from the model background image Bt. The resulting binary foreground image is obtained as follows [1]:

$$F^t(x,y) = \begin{cases} 1 & \text{if } R(I_M(x,y), B^t(x,y)) < \sigma_s \\ 0 & \text{otherwise} \end{cases}$$

IV.II Morphological Operation:

We cannot get a good estimate of the hand image because of background noise. To get a better estimate of the hand, we need to delete noisy pixels from the image. We use an image morphology algorithm that performs image erosion and image dilation to eliminate noise [1]. Erosion trims down the image area where the hand is not present and Dilation expands the area of the Image pixels which are not eroded. Mathematically, Erosion is given by,

$$A \ominus B = \{x \mid (B)_x \cap A^c = \phi\}$$

Where A denotes input image and B denotes Structure elements. The Structure element is operated on the Image using a Sliding window and exact matches are marked Dilation is defined by,

$$A \oplus B = \{x \mid (\hat{B}) \cap A \neq \phi\}$$

$$= \{x \mid [(\hat{B})_x \cap A] \subseteq A\}$$

Where A denotes the input image and B denotes the structure element. The same structure element is operated on the image and if the center pixel is matched, the whole area around that pixel is marked [5].

V. HAND TRACKING

V.I Camshift Algorithm:

The segmented hand from the input sequence is tracked in the subsequent phase. This is done by using modified CAMSHIFT based on the CAMSHIFT algorithm for hand tracking. CAMSHIFT is designed for dynamically changing distributions [3]. This unique feature makes the system robust to track moving objects in video sequences, where size of the object and its location may change over time. In this way, dynamic adjustment of search window size is possible. CAMSHIFT is based on colors, thus it requires the availability of color histogram of the objects to be tracked the desired objects within the video sequences. As mentioned earlier, the color model was built in the HSV domain on the basis of the hue component. The spatial mean value is computed by selecting the size and initial position of the search window. The search window is moved towards center of the image in subsequent steps. Once the search window is centered computing is done through the centroid with first-order instant for x, y. The process is continued till it arrived at the point of convergence. In the implemented tracking process the image is divided into four regions R1, R2, R3, and R4. Then the number of white pixels in each of the four divided regions is calculated based on the total number of pixels. In order to find the position of hand in the divided regions we calculate p1, p2, p3, p4 the position values of hand tracked in the corresponding region. The position values are calculated using the Camshift algorithm is based on the Mean Shift algorithm[1].

VI. GESTURE RECOGNITION

In the recognition phase as shown in figure the system extracts the features of the hand for gesture classification. In the

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

recognition phase the system extracts the features of the hand for gesture classification. The different gestures are differentiated by extracting features from the contours around hand and getting the convex hull of the given contour. One can identify different gestures by counting the numbers of defects in the convex hull and it is relative orientation in its bounding rectangle. The features are classified into 8 classes, which are classified by a widely used algorithm for recognition called NCN (nearest conflicting neighbors).The recognition phase of the proposed two hand gesture recognition system takes the segmented image as input. The region of interest in the segmented image is identified and set first. Then the system finds the contour around the image and biggest area around the contour of the image is extracted to find the convex hull of this contour. Once the convex hull s found the number of defects in the image is calculated and orientation of the bonded region of image is found. Next the direction of the image is taken from the movement of the tracking region [1].

VII. CONCLUSION

Gesture based interfaces allow human computer interaction to be in a natural as well as intuitive manner. It makes the interaction device free which makes it useful for dynamic environment It is though unfortunate that with the ever increasing interaction in dynamic environments and corresponding input technologies still not many applications are available which are controlled using current and smart facility of providing input which is by hand gesture. The most important advantage of the usage of hand gesture based input modes is that using this method the users get that ability to interact with the application from a distance without any physical interaction with the keyboard or mouse [1]

REFERENCES

1. Conic, N., Cerseato, P., De Natale, F. G. B.,: Natural Human- Machine Interface using an Interactive Virtual Blackboard, In Proceeding of ICIP 2007, pp.181-184, (2007) pdf.
2. R. Lienhart and J. Maydt, “An extended set of Haar-like features for rapid object detection,” In Proceedings of ICIP02, pp. 900-903, 2002.
3. Nguyen Dang Binh, Enokida Shuichi, Toshiaki Ejima, “Real-Time Hand Tracking and Gesture Recognition System”, 2005.
4. “New Hand Gesture Recognition Method for Mouse Operations”, IEEE 2011
5. R.C.Gonzales, R.E.Woods, Digital Image Processing. 2-nd Edition, Prentice Hall, 2002.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.