International Journal of Innovations in Engineering and Technology (IJIET)
http://dx.doi.org/10.21172/ijiet.112.09

Face Tracking using Camshift in Head Gesture Recognition System

Er. Rushikesh T. Bankar[1], Dr. Suresh S. Salankar[2]

[1]*Department of Electronics Engineering, G H Raisoni College of Engineering, Nagpur, Maharashtra, India*
[2]*Department of Electronics & Telecommunication Engineering, G H Raisoni College of Engineering, Nagpur, Maharashtra, India*

Abstract - The gesture recognition has wide range of applications such as, developing aids for the hearing impaired, designing techniques for forensic identification, recognizing sign language, medically monitoring patient's emotional states or stress levels and monitoring automobile driver's alertness / drowsiness levels etc. In head gesture recognition based interface (HGI) process, the face detection, as well as face tracking is important to track under different environmental conditions. Some difficulties with head gesture based control are, the head of the user may be out of the image view, the face color of the user may change in the varying illumination conditions and the user may have different facial appearances like mustache and glasses. In this paper we proposed a novel head gesture recognition approach with traditional face detection algorithm and object tracking algorithm are combined in the system to achieve accurate face detection, tracking and gesture recognition. The system consists of face detection, face tracking, and eye template matching method and also has obstacle detection system comprising of the four ultrasonic sensors.
Keywords - Face detection, Face tracking, Gesture recognition, Haar like features.

I. INTRODUCTION

In our daily lives, today's framework of interactive, intelligent computing, an efficient human computer interaction is assuming utmost importance. The gesture recognition can be termed as an approach in this direction. It is the process by which the gestures, hand / head made by the user are recognized by the receiver. The gestures are expressive, having meaningful body motions involving the physical movements of the hands, arms, fingers, face, head or body with the intent of conveying the meaningful information or doing some needful actions. In any gesture recognition, the environmental conditions may play an important role. The gesture recognition has wide range of applications such as, developing aids for the hearing impaired, designing techniques for forensic identification, recognizing sign language (Hand Gesture Recognition), medically monitoring patients emotional states or stress levels and monitoring automobile driver's alertness / drowsiness levels etc. In head gesture recognition based interface (HGI) process, the face detection, in the particular environmental condition or the user using the system or wheelchair in indoor or outdoor environmental conditions, is very useful. The face tracking is also important to track under different environmental conditions. Some difficulties deal with the new generation of the head gesture based control of wheelchairs. There difficulties are, the head of the user may be out of the image view, the face color of the user may change in the varying illumination conditions and the user may have different facial appearances like mustache and glasses. These limitations can be solved in the paper [1] P. Jia, H. Hu, T. Lu and K. Yuan, "Head Gesture Recognition for Hands Free Control of an Intelligent Wheelchair". The paper presents a novel hands free control system for an intelligent wheelchairs based on the visual recognition of the head gestures. The traditional face detection algorithm and object tracking algorithm are combined in the system to achieve accurate face detection, tracking and gesture recognition in the real time. This system is used as the human friendly interface for the elderly and disabled people to operate an intelligent wheelchair using their head gestures rather than their hands. The system uses camshift algorithm for the face tracking of the user. But the camshift has some limitations. These are; it cannot accurately track the face when the illumination condition changes, it cannot work well under the cluttered background.

Rathore et al. [2] proposed an intelligent system which can assist physically handicapped, visually impaired as well as elderly people. It consists of a navigation system. The system uses the accelerometer and magnetometer. The system also consists of a navigation pad which can be tied to the head foe navigating the intelligent wheelchair. The obstacle detection / avoidance are one of the most important parameter for the intelligent wheelchair used by the user. For the obstacle avoidance system, the four ultrasonic sensors are used.

Manju Davy and R. Deepa [3] proposed an intelligent wheelchair based on the accelerometer sensor to recognize the head movements of the user. The project works for the handicapped peoples who cannot perform their controlled movement. The system or an intelligent wheelchair will be used for the patients who are particularly suffering from the diseases. To generate the motion control commands for the controller, the recognized gestures of the user are used. So that wheelchair can control the motion of intelligent wheelchair according to the interaction of the user.

The position controlled by the head is the gesture which can be performed by the patients having quadriplegic. So the movement of the user's head is the gesture for such patients.
The paper [4] presents about nose template matching for the recognition of the frontal face posture. Unlike the profile left and profile right head gestures, the further more investigation is conducted to recognize the frontal head gestures.

Figure 1: Nose Template Matching Method

The figure [1] above shows clearly that the head gesture recognition method is fairly feasible. There are five frontal face gestures to be recognized namely frontal center; frontal up; frontal down; frontal left; frontal right.

II. THE EXISTING GESTURE RECOGNITION METHODS

The author [5] Jiao Zhang et al. presents a hand gesture recognition system. The system is a novel intelligent wheelchair control for those with physical accessibility problem. In this system, the haar like features and the adaboost learning algorithm are used for the hand gesture detection. The hand gesture commands are determined by comparing the center of the minimum rectangle which contains the hand gesture with a fixed area. The real time performance and the high recognition accuracy can be obtained with Adaboost algorithm.
To detect and track hand gesture, each frame of input image series is swept from the top left corner to the bottom right corner by a stretchable sub window so that the different hand sizes can be detected. The sub window of a specific size scans the image pixel by pixel.
Cuong Tran et al. [6] proposed an optical flow based head movement and gesture analyzer system. In this system, the head movements can be broken into its elemental forms that are moving and fixation states and combinations of these elemental forms give rise to various head gestures. The optical flow based head movement and gesture analyzer segments head gestures into moving and fixation states using optical flow tracking and intermittent head pose estimation. The segmenting head gestures into its elemental forms gives access to the higher level semantic information such as fixation time and rate of head motion.
To estimate the head motion used for the state transitions, the optical flow vectors are computed over sparse interest points in the image using the lucas kanade algorithm. The experimental analysis shows promising results of an average of 88% accuracy in the fixation state and an average of 87% accuracy in move state level classification.
Parimita Saikia and Karen Das [7] presents head gesture recognition system using optical flow based classification with reinforcement of gaussian mixture model based background subtraction. The gaussian mixture model is used to get the required information regarding the head movement of the user. The system model can be implemented in the various control system. A Gaussian mixture model is a parametric probability density function represented as a weighted sum of Gaussian component densities.
The shadows which are detected as foreground can cause several problems when extracting the objects. The two examples are object shape distortion and several objects merging together. The system is especially crucial and these problems should be avoided.

III. ADABOOST ALGORITHM & HAAR LIKE FEATURE

The adaboost algorithm is built on the attential cascade structure. The basic principle is, first, scaling the original image through the multi scaling and then using a fixed size sub window scans all the zoom of the scaled image to obtain a large number of sub window images. Each sub window image needs to go through a multi layer structural classifier to determine whether these sub window images are the target windows, and then obtains the final detected result. The Adaboost algorithm of viola and jones provides good performance for the face detection, license plate detection and pedestrian and exceptional performance in the binary classification problems. In this paper it was used for head gesture detection. The Adaboost algorithm is one kind of self adaptation boosting algorithms.

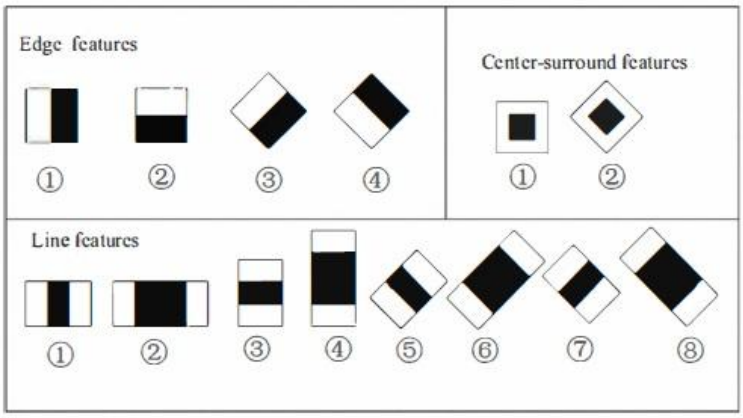

Figure 2: Extended set of haar like features

The haar like features focus more on the information within a certain area of the image rather than each single pixel. The haar like feature based system can operate much faster. Each haar like feature consists of two or three connected black and white rectangles. The haar like features describe the ratio between the dark and bright areas within a kernel. Therefore they are relatively robust to noise and lighting changes. For the face detection, firstly, the haar like features was used because the eye region on human face is darker than the cheek region and one haar like feature can efficiently catch that characteristic.
The value of a haar like feature is the difference between the sums of the pixel values in the black and white rectangles. To detect an object, the image is scanned by a sub window containing a specific haar like feature.

IV. HEAD GESTURE RECOGNITION

The existing head gesture recognition method used nose template matching method for the recognition of the head gestures. The profile left, profile right, profile up and profile down head gestures are recognized using the nose template matching method. As per the nose template directions of the user, the movement of the head is to be decided. Accordingly that the control signal goes to the wheelchair system and the wheelchair moves as per the directions.
But the system deals with some uncertainty in the practical application of the electric wheelchair. If any person calls the user and the user seated on the wheelchair is in the opposite direction the person, then in this situation we normally moves the head. So the user moves his head and accordingly the wheelchair moves. But the users don't want to move the wheelchair. So such misoperation occurred while the system uses nose template matching method with head movement.
To overcome this limitation we have to propose a system with head titling. For the head tilting of the user, system requires eye template matching method. For the movement of the wheelchair, the user uses the head tilting method. In this situation, if anybody calls the user then the user move his head and not tilt the head. So the misoperation of the wheelchair is not performed by the user. For the movement of the electric wheelchair, the head tilting method is the best method. For performing the head tilting method we require the method of template matching. So the eye template matching method is the best method to recognize the head gestures of the user.

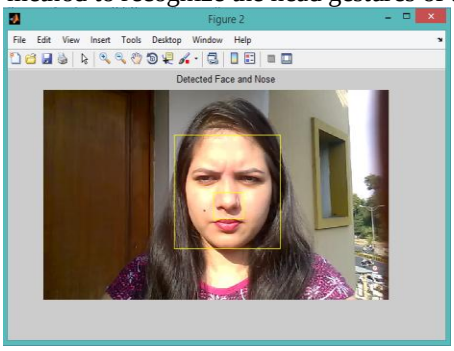

Figure 3: Face detection

Figure 3 shows the performance of the viola jones algorithm. The face of the user is detected using viola jones algorithm.

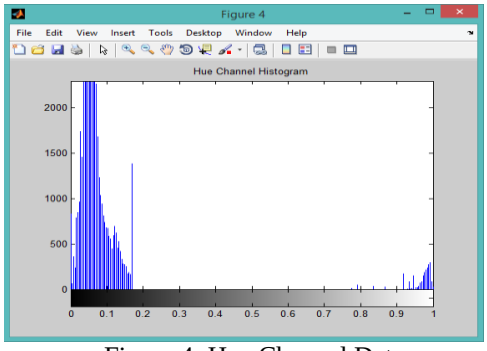

Figure 4: Hue Channel Data

Figure 4 shows the hue channel data for the face detection of the user.

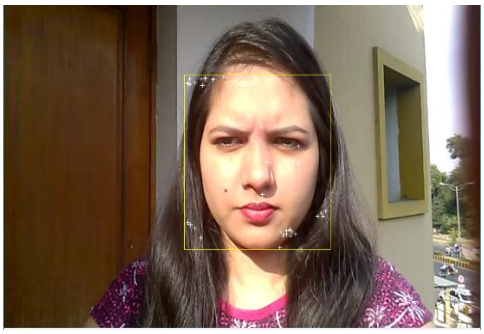

Figure 4: Hue Channel Histogram

Figure 4 shows the hue channel histogram for the detected face of the user which is shown in the figure 3. Basically, a color model is a method for specifying colors in some standard way. It generally consists of a 3 D coordinate system and a subspace of that system in which each color is represented by a single point. In RGB model, each color is represented as 3 values R, G and B, indicating the amounts of red, green and blue which make up the color. The HSV Model includes hue, saturation and value properties. The hue property states that the true color attribute that is red, green, blue, orange, yellow and so on. The saturation property states that the amount by which the color as been diluted with white. The whiter in the color, the lower the saturation. The value property states that the degree of the brightness a well lit color has high intensity; a dark color has low intensity.

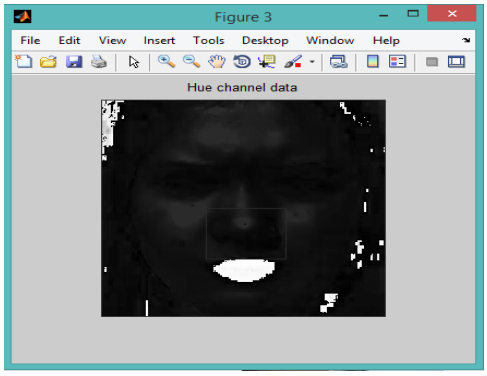

Figure 5: Detected Key points

Figure 5 shows the key points detection in user's face using the scale invariant feature transform method. The transform method generates the image feature that is the key points. These key points are invariant to the image

scaling and rotation. This method is the partially invariant to the change in illumination. The key point localization rejects the low contrast key points. With this information, points are rejected that have low contrast because they are sensitive to the noise or are poorly localized along an edge.

V. FACE TRACKING ALGORITHM

The head gesture based interface consists of real time face detection, tracking and gesture recognition. The Adaboost is the most recent face detection method with both high accuracy and fast speed. The Adaboost extracts the haar like features of the image which contain the image frequency information only by the integer calculation that is fast. Then a set of key features are selected from all the extracted features. Instead, we compute likelihood maps where the value at a pixel is proportional to the likelihood that the pixel comes from the object we are tracking. The computation of the likelihood can be based on the color; texture; shape (boundary) and predicted location.

The mean shift algorithm is an efficient approach to tracking the objects whose appearance is defined by the histograms since not limited to only color. The mean shift represent the color distribution with a histogram. The mean shift finds the region that has most similar distribution of colors. For adopting the window size with the size and rotation of the targeted head movement of the user, it goes to the camshift that is continuously adaptive meanshift. Again it applies the meanshift with new scaled search window and previous window location. The process continued until the required accuracy will meet.

The camshift is a very efficient color tracking method. It is a fast object tracking method based on image hue. The camshift is the continuously adaptive mean shift algorithm. The camshift sets the region of interest in the image. It selects the initial location of the object to be tracked. The window goes from minimum pixel to the maximum pixel which finds the centre of the image then tracked the object.

Table 1: Speed and Minimum Face Size of Camshift Algorithm

Environmental Conditions	Directions	Minimum Face Size	Time cost / Frame
Sunshine	Frontal	123 * 123	0.0370
	Left	123 * 123	0.0026
	Right	123 * 123	0.0026
	Up	123 * 123	0.0030
	Down	123 * 123	0.0043
Cluttered	Frontal	100 * 100	0.0333
	Left	100 * 100	0.0047
	Right	100 * 100	0.0039
	Up	100 * 100	0.0026
	Down	100 * 100	0.0043
Shadow	Frontal	143 * 143	0.0406
	Left	143 * 143	0.0003
	Right	143 * 143	0.0040
	Up	143 * 143	0.0007
	Down	143 * 143	0.0032

Table 1 above shows the speed and minimum face size of face tracking algorithm. The different environmental conditions are sunshine, cluttered and shadow. The proposed system tracks the face under different head gestures of the user is shown in the above table with speed and face size.

VI. CONCLUSION

The head gesture based interface consists of real time face detection, tracking and gesture recognition. The Adaboost algorithm is used for face detection shows high accuracy and fast speed. The Adaboost extracts the haar like features of the image which contain the image frequency information only by the integer calculation that is fast. Then a set of

key features are selected from all the extracted features. The camshift algorithm is used for color tracking concluded to be fast object tracking method based on image hue.
The head gestures are determined by eye template matching method. Then the movement of the system is controlled by the head gestures of the user. Experimental results are tested on wheel chair system. The system is equipped with accelerometer as well as eye template matching method for the movement of the chair and the ultrasonic sensors are used for the obstacle detection.

VII. REFERENCES

[1] P. Jia, H. Hu, T. Lu and K. Yuan, "Head Gesture Recognition for Hands Free Control of an Intelligent Wheelchair", 2010 IEEE International Conference on Computer Application and System Modeling (ICCASM 2010).
[2] Deepesh K. Rathore, Pulkit Srivastava, Sankalp Pandey and Sudhanshu Jaiswal, "A Novel Multipurpose Smart Wheelchair", 2014 IEEE Student's Conference on Electrical, Electronics and Computer Science.
[3] Manju Davy and R. Deepa, "Hardware Implementation Based on Head Movement using Accelerometer Sensor", International Journal of Applied Sciences and Engineering Research, Volume 03, Issue 01, 2014.
[4] P. Jia, H. Hu, T. Lu and K. Yuan., et al., "Head Gesture Recognition for Hands Free Control of an Intelligent Wheelchair", 2007: 1 - 10.
[5] Yi Zhang, Jiao Zhang and Yuan Luo, "A Novel Intelligent Wheelchair Control System Based on Hand Gesture Recognition", Proceedings of the 2011 IEEE / ICME International Conference on Complex Medical Engineering May 22 - 25, Harbin, China 334 - 339.
[6] Sujitha Martin, Cuong Tran and Mohan Trivedi, "Optical Flow Based Head Movement and Gesture Analyzer", 21st International Conference on Pattern Recognition (ICPR 2012), November, 11 - 15, 2012, Tsukuba, Japan.
[7] Parimita Saikia and Karen Das, "Head Gesture Recognition using Optical Flow based Classification with Reinforcement of GMM based Background Subtraction", International Journal of Computer Applications (0975 - 8887), Volume 65, No. 25, March 2013, 05 - 11.
[8] Zhang Fang Hu & Lin Li, "A Novel Intelligent Wheelchair Control Approach Based on Head Gesture Recognition, " 2010 IEEE International Conference on Computer Application & System Modeling (ICCASM 2010), PP. 159 - 163.
[9] Jingbo Zhao & Robert Allison, "Real Time Head Gesture Recognition on Head Mounted Displays using Cascaded Hidden Markov Models, " IEEE SMC 2017, Banff, Canada, PP. 159 - 163.
[10] Hyunduk Kim, Sang Heon Lee, Myoung Kyu Sohn and Dong Ju Kim, "Illumination Invariant Head Pose Estimation using Random Forests Classifier & Binary Pattern Run Length Matrix", Human Centric Computing and Information Sciences 2014, a Springer Open Journal, 2014.
[11] Dajie Cong, Ping Shi, and Daxi Zhou, "An Improved Camshift Algorithm Based on RGB Histogram Equalization", 2014 7th IEEE International Conference on Image & Signal Processing .
[12] Eshed Ohn Bar & Mohan Manubhai Trivedi, "Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision based Approach & Evaluations", IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 6, PP. 2368 - 2377.
[13] Cloud Yanwen Chong, Jianfeng Huang, Shaoming Pan, "Hand Gesture Recognition Using Appearance Features Based on 3D Point", Journal of Software Engineering and Applications, 2016, 9, 103 - 111.
[14] Srishti, Prateeksha Jain, Shalu, Swati Singh, "Design & Development of Smart Wheelchair using Voice Recognition & Head Gesture Control System", International Journal of Advanced Research in Electrical, Electronics & Instrumentation Engineering, 2015, 04, 4790 - 4798.