

Mouse Handling Using Facial Features

Neha Nikhade¹, Maninee Chaki², Nazneen Shaikh³, A.B.Lamgunde⁴

Department of Computer Engineering, Sinhgad Institute of Technology, Lonavala. Maharashtra, India^{1,2,3,4}

Abstract: Man interacts with computer in a number of ways, the interface between humans and computers they use is important in facilitating this interaction. New methods of Human Computer Interaction (HCI) has been developed in recent years, such as interaction with machine using hand movements, head movements, facial expressions, voice, touch etc. HCI aims to improve the interactions between users and computers by making computers more user friendly. The paper presents a system for eye blink detection and head movements for communication between man and machine. Facial features (eyes and nose) are detected based on which actions such as head movements and eye blinks are performed so that cursor can carry out appropriate functions on screen. Our system provides a comprehensive solution to control computer mouse cursor movement with eye-blinks and head movement for people with special needs. The user may perform the clicking of mouse through eye-blinks. For detection of face candidates our basic strategy is use of Six-Segmented Rectangular Filter (SSR) along with a support vector machine (SVM) for face verification. The patterns of between-the-eyes are tracked with update template matching in face tracking

Keywords: HCI, SSR, SVM, VOG

I. INTRODUCTION

In 21st century almost everything is computerize. Human nature demands independence, similar is the case with physically handicapped people. Computer has affected people, society and surrounding in many ways. With the availability of low cost webcams and high speed processors, many people are taking keen interest in image processing. One of the important fields in artificial intelligence is human computer interface which uses human gestures and features to interact with the computer.

“The planning, study and design of the interaction between computer and users” is known as Human Computer Interaction (HCI). The user interface (hardware and software) represents the point of communication between a computer system and people and this is where the interaction begins. In other words input to the system and output of the system (i.e. interaction) is studied along with emotions, intend etc.

Head and face detection is the first step to be taken for facial expression recognition and head gesture recognition. In this paper we propose a real time face detection and tracking algorithm for video sequences. In recent years many methods, algorithms have been proposed for detection of face. Hjelmas [1] did a comprehensive study on this subject, listing around 200 references. According to him face detection techniques were classified into two broad categories: the feature based approach and the image based approach. A window scanning technique [4] is mainly applied to most of the image based approaches. Many approaches were built for iris tracking as well. Video - oculography (VOG) [6] is one such traditional approach to provide the gaze position with high precision. VOG uses illumination and data acquisition techniques in order to determine the point the subject is looking at using captured images. Biggest advantage of VOG is providing a solution for building non-intrusive systems which avoids use of helmets and special glasses holding small cameras. The most popular

face candidate extraction method is to extract a skin-tone region in color images [2] [5] [17] [19] [20]. The problem with color information is that it is extremely sensitive to lighting conditions, and to adapt the skin-tone model to the lighting environment in real-time increases complexity. From stereo systems detailed information is also used to extract foreground objects (i.e. heads) from the background [10][3][12]. However, this detailed information involves complex computation. The background subtraction technique is another way to extract face candidate regions [11].

Another problem we want to fix here is the fact that most feature-based approaches usually require features related not only to the eyes but also to mouth, or skin-colors and/or face contours, which are likely to be affected by beards, mustaches or hair styles.

Another step that we take to reduce complexity is conversion of color image to gray scale. Apart from this our main focus lies on reducing the hardware so that it becomes readily available and affordable to people belonging to all classes.

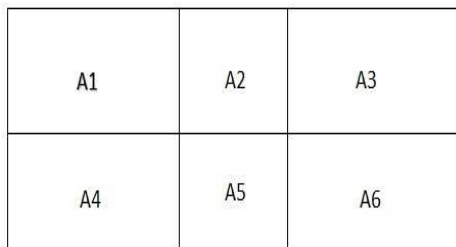
Here we proposed a filtering technique to extract between the-Eyes in real time in [9]. But this filtering approach may fail to detect Between-the-Eyes when hair covered the forehead, which can happen in some cases. Thus in orders to reduce unwanted complexity we first extract face candidates and apply the filter to only those candidates. Due to location and shape of nose, the nose tip is selected as pointing device. Also because nose is located at the middle of the face, it makes it easier to track the face.

Eye blinks is used by user to perform mouse events on screen. We are aimed at designing a system that uses head movements and eye blinks to interact with computer. The left/right eye blinks perform left/right mouse click events on screen. The only external device that the user needs is a webcam.

As almost everything today is being computerized it is important to increase the human-computer interaction (HCI) even for people with special needs. Our software system primarily focuses on this aspect of the computer system, that is, improved HCI by providing alternative mouse controls for physically challenged people such as people suffering from Quadriplegia, Cerebral Palsy, and Multiple Sclerosis

II. SSR

The extraction of face candidates is done using the six segment rectangular filter or the SSR filter. The rectangular region is divided into six segments A1-A6. Initially, a rectangle is scanned throughout the input image. Real time tracking of eyes, many a times, relies on the location of the region between-the-eyes (BTE). By placing the SSR filter, the eyes and between-the-eyes region is found out by calculating the average pixel value within the segment.



We denote the average pixel value within a segment A_i as A_i'' . Then when one eye and eye brow is within A1 and the other within A3, we can expect

$$A1'' < A2'' \quad \text{and} \quad A1'' < A4'' \quad (1)$$

$$A3'' < A2'' \quad \text{and} \quad A3'' < A6'' \quad (2)$$

A point where (1) and (2) can be satisfied can be considered to be face candidate. Once the face candidates are identified, the remaining area of the face that will not be used is discarded.

III. SVM

The face candidates that are extracted using the SSR filter are verified using the Support Vector Machine (SVM). The SVM are a new type of maximum margin classifiers. In Learning theory a particular theorem states that to get minimum classification error the hyper plane that separates positive samples from the negative ones should be with the maximum margin of the training samples. An input training data sample is given to the SVM, where each sample consists of attributes and a class label. Support vectors are the data samples that are nearest to the hyper plane. In order to get the maximum margin of the training data set, the hyper plane is defined by balancing its distance between the positive and the negative support vectors. Thus the verification of the BTE, extracted within the face candidates, is done using the SVM.

IV. JMF

The Java Media Framework (JMF) is a Java library that enables audio, video and other time-based media to be added to Java applications and applets. This optional package, which can capture, play, stream, and transcode multiple media formats, extends the Java Platform, Standard Edition (Java SE) and allows development of cross-platform multimedia applications. Currently used API for Java, the Java Media Framework (JMF) is dealing with real-time multimedia presentation and effects processing.

JMF handles time-based media, media which changes with respect to time. Examples of this are video from a television source, audio from a raw-audio format file and animations. For our project we will be using JMF version 2.1 and above.

V. PROCESS EXECUTION

1]Nose Tip Tracking

Once the face candidates are extracted, the nose tip is tracked for moving the cursor on the screen.

2]BTE Tracking

BTE is the only region perpendicular to the eyes and nose and hence it is tracked in the refinement process. The BTE is tracked just like the nose tip.

3]Detection of Motion

The pixels in a certain region are subtracted from the same pixels of the previous frame and motion is detected. The motion at a pixel is detected if the absolute value of the subtraction is larger than a certain threshold.

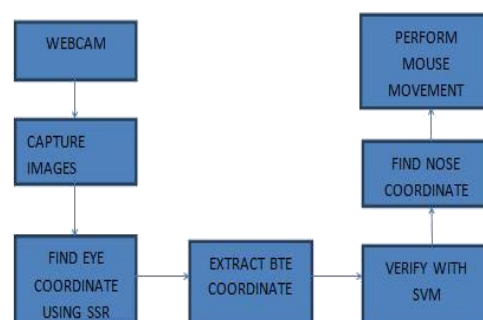
4]Detection of Blink

Only if the eye is not moving can the blink detection process be run. The face will be used instead of the mouse to move the cursor on the screen and then the blink is performed to click on eyes.

5]Tracking of Eyes

If a left/right blink was detected, the tracking process of the left/right eye will be skipped and its location will be considered as the same one from the previous frame (because blink detection is applied only when the eye is still). Eyes are tracked differently from tracking the nose tip and the BTE, because these features have a steady state while the eyes are not opening, closing, and blinking.

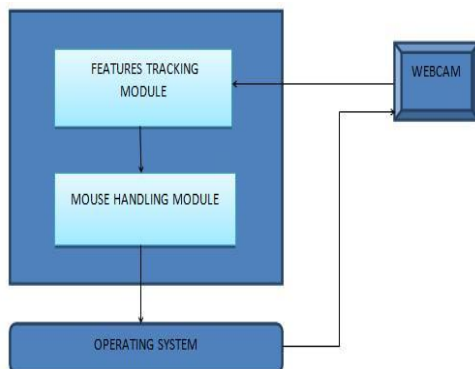
VI. SYSTEM DESIGN



Person sitting in front of computer having external camera or inbuilt camera. Adjust the camera such that person face should be in the center of camera's view. We need to extract facial features using SSR and SVM. Track those features in the video stream using integral images method. Tracking involves skin pixels like eyes, eyebrows, between the eyes (BTE), nose tip. Mouse movement is performed by the nose tip. Blinks performed by the eyes are tracked so as to perform click operations. If left blink detected, mouse will perform left click operation whereas if right blink detected, mouse will perform right click operation..

VII. SYSTEM FRAMEWORK

Webcam is connected to the machine, it may be either inbuilt or externally. As we know the role of webcam is to capture images whether in a form of single shot or multiple shot. Multiple shot is nothing but the images in the form of video stream. But we need to extract some features rather than total image. Features like eyes, eyebrows, between the eyes and nose tip. Using those features, we can perform mouse movement and some basic operations like left click, right click, and double click in operating system.



VIII. CONCLUSION

Entire image is scanned by SSR filter. SSR filter is similar to the window scanning technique. Important advantage of this is we do not work on skin pixels only we need is location of face. One more advantage is it does not need to work on several colors. Moreover, it works on only two colors black and white. Our system only uses webcam no other external hardware rather than it.

REFERENCES

[1] ATT Laboratories, Cambridge, UK. "The ORL Database of Faces". <http://www.uk.research.att.com/facedatabase.html>.
 [2] D. Chai and K. N. Ngan. Face segmentation using skin-color map in videophone applications. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(4):551–564, 1999.
 [3] T. Darrell, G. Gordon, M. Harville, and J. Woodfill. Integrated person tracking using stereo, color, and pattern detection. *Proc. CVPR 1998*, pages 601–608, 1998.
 [4] E. Hjelmas and B. K. Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, 2001.
 [5] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face detection in color images. *IEEE Trans. on PAMI*, 24(5):696–706, 2002.
 [6] T. Joachims. SVM. <http://svmlight.joachims.org/>.

[7] T. Joachims. Making large-scale svm learning practical. *Advances in Kernel Methods — Support Vector Learning*, B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.
 [8] S. Kawato and J. Ohya. Two-step approach for real-time eye tracking with a new filtering technique. *Proc. Int. Conf. on System, Man & Cybernetics*, pages 1366–1371, 2000.
 [9] S. Kawato and N. Tetsutani. Real-time detection of between-the-eyes with a circle frequency filter. *Proc. ACCV 2002*, II:442–447, 2002.
 [10] E. Hjelmas and B. K. Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, 2001.
 [11] D. Chai and K. N. Ngan. Face segmentation using skin-color map in videophone applications. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(4):551–564, 1999.
 [12] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face detection in color images. *IEEE Trans. on PAMI*, 24(5):696–706, 2002.
 [13] J. C. Terrillon and S. Akamatsu. Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images. *Proc. 12th Conf. on Vision Interface (VI '99)*, 2:180–187, 1999.
 [14] H. Wu, Q. Chen, and M. Yachida. Face detection from color images using a fuzzy pattern matching method. *IEEE Trans. on PAMI*, 21(6):557–563, 1999.
 [15] J. Yang and A. Waibel. A real-time face tracker. *Proc. 3rd IEEE Workshop on Application of Computer Vision*, pages 142–147, 1996.
 [16] Mehrube Mehrubeoglu*, Linh Manh Pham, Hung Thieu Le, Ramchander Muddu, and Dongseok Ryu "Real-Time Eye Tracking Using a Smart Camera"
 [17] S.-H. Kim, N.-K. Kim, S. C. Ahn, and H.-G. Kim. Object oriented face detection using range and color information. *Proc. IEEE 3rd Int. Conf. on Automatic Face and Gesture Recognition*