

Visual Based Control Using Human Computer Interface

Varsha N
Dept of CSE
PES College of Engineering,
Mandya

Varun N
Dept of ECE
Vidyavardhaka College of Engineering,
Mysuru

Abstract- A Human Computer Interface (HCI) System for playing games is designed here for more natural communication with the machines. The system presented here is a vision-based system for detection of long voluntary eye blinks and interpretation of blink patterns for communication between man and machine. This system replaces the mouse with the human face as a new way to interact with the computer. Facial features (nose tip and eyes) are detected and tracked in real-time to use their actions as mouse events. The coordinates and movement of the nose tip in the live video feed are translated to become the coordinates and movement of the mouse pointer on the application. The left/right eye blinks fire left/right mouse click events. The system works with inexpensive USB cameras and runs at a frame rate of 30 frames per second.

Keywords- Human Computer Interface (HCI), SSR Filter.

I. INTRODUCTION

This project aims to present an application that is capable of replacing the traditional mouse with the human face as a new way to interact with the computer. Facial features (nose tip and eyes) are detected and tracked in real-time to use their actions as mouse events. The coordinates and movement of the nose tip in the live video feed are translated to become the coordinates and movement of the mouse pointer on the user's screen. The left/right eye blinks fire left/right mouse click events. The only external device that the user needs is a webcam that feeds the program with the video stream. One way to achieve that is to capture the desired feature with a webcam and monitor its action in order to translate it to some events that communicate with the computer. The nose tip was selected as the pointing device as it is more comfortable to use it as the feature that moves the mouse pointer and defines its coordinates. Eyes were used to simulate mouse clicks, so the user can fire their events as he blinks.

An Algorithm is used that detects and tracks the desired facial features precisely, and fast enough to be applied in real-time.

II. FACE DETECTION AND TRACKING

The system uses color, motion, and correlation-based template matching to detect and track faces. It can detect and track different size faces at various distances to the camera. This allows users to move closer or farther from the camera and still be detected automatically. To achieve this, the system uses image pyramids [60]. Each pyramid consists of eight levels. The highest resolution image in the pyramid is

the 640×480 video frame, and the lowest resolution image contains 32×24 pixels. In each level of the pyramid, the system searches for a face of size 12×16 . This approach allows the system to operate in real time. The level of the pyramid at which the face is detected can then be used to infer the "true" size and location of the face, which is the size and location of the face in the original video frame. In the case where the person is far from the camera and appears relatively small in the video frame, the system can efficiently detect the face in a high-resolution level of the pyramid. In the case where the person is close to the camera and therefore appears relatively large in the original video frame, the face is detected efficiently in a lower resolution level of the pyramid. The face tracking algorithm computes pyramids PColor and PMotion from the result of a color histogram lookup C and a motion image M, as described in Sections I and II hereafter. These pyramids are used to mask the pyramid PInput of the input video frame to yield a pyramid PMaskedInput, which is then used to compute a pyramid PCorrelation of template correlation values.

III. FACE RECONGNITION PRE-PROCESSING

An image consists of a two-dimensional array of numbers. The color or gray shade displayed for a given picture element (pixel) depends on the number stored in the array for that pixel. The simplest type of image data is black and white. It is a binary image since each pixel is either 0 or 1. The next, more complex type of image data is gray scale, where each pixel takes on a value between zero and the number of gray scales or gray levels that the scanner can record. These images appear like common black-and white photographs | they are black, white, and shades of gray. Most gray scale images today have 256 shades of gray. People can distinguish about 40 shades of gray, so a 256-shade image looks like a photograph." This concentrates on gray scale images. The most complex type of image is color. Color images are similar to gray scale except that there are three bands, or channels, corresponding to the colors red, green, and blue. Thus, each pixel has three values associated with it. A color scanner uses red, green, and blue letters to produce those values. Images are available via the Internet, scanners, and digital cameras. Any picture shown on the Internet can be downloaded by pressing the right mouse button when the pointer is on the image. This brings the image to your PC usually in a JPEG format. Your Internet access software and other software packages can convert that to a TIFF or BMP.

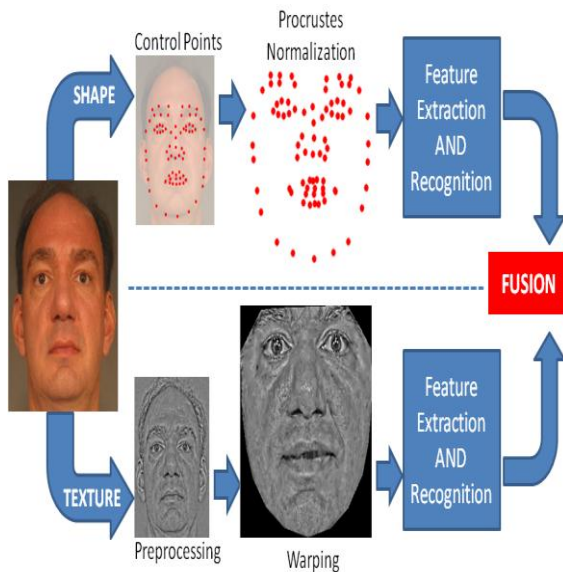


Figure-3.1 face recognition pre-processing

IV. MOTION DETECTION

As in the skin color analysis, a preprocessing mask—here, the motion mask is created to aid face detection. The analysis is based on the assumption that, if the user moves, frame differencing should find the pixels where motion occurs in the image. The frame differencing creates the motion image .

$M(x, y) = |I_t(x, y) - I_{t-1}(x, y)|$ (1) where the pixel values $M(x, y)$ are the absolute difference between the grayscale values of the current image frame I_t and the previous image frame I_{t-1} (grayscale values are provided by the Y channel of the YUV color image). Since the frame rate of the camera used is 15 frames per second, those frames represent images taken approximately 67 ms apart. A higher frame rate may require the algorithm to be modified to maintain a similar temporal separation between frames.

The motion image is decimated into a pyramid. As in the color analysis, the goal is to find a face of size 12×16 . A low-pass averaging filter of support 12×16 is therefore applied to remove motion that cannot be due to the motion of the face. Subsequent thresholding with a threshold value of ten gray levels then results in the binary motion mask shows a motion mask pyramid with the low-pass filter disabled to better highlight pixels that contribute to the motion image. In the case when there is little or no motion, the system tries to find the face near the same location and scale as found in a previous frame. It sets the locations within five pixels of the previously found face location to one in the binary motion image and thus prevents the motion mask from excluding the previous face location from subsequent processing.

The two adjacent motion pyramid levels are also modified in this way to account for small movements toward or away from the camera, which are not detected by the motion analysis. When the system initializes and there is no

previously detected face location, the center of the image is set as the default prior location for all pyramid levels.

V. EYE DETECTION

EyeKeys was tested with the camera mounted on the end of an articulated arm, which allowed the camera to be optimally positioned in front of the computer.

VI. CONCLUSION

Assistive technology enables people with severe paralysis to communicate their thoughts and emotions. It also allows them to exhibit their intellectual potential—sometimes disproving a previous diagnosis of a mental disability. To provide such communication technology, we have created the camera-based human-computer interface EyeKeys, which is a new tool to use gaze direction to control the computer.

The EyeKeys' face tracker combines existing techniques in a new way that allows the face to be tracked quickly as a means to locate the eyes. The method of mirroring and projecting the difference between the eyes is a novel approach to detecting to which side the eyes look. Experiments with EyeKeys have shown that it is an easily used computer input and control device for able bodied people and has the potential to become a practical tool for people with severe paralysis.

REFERENCES

- [1] J.-D. Bauby, *The Diving Bell and the Butterfly*. New York: Vintage Books, 1997.
- [2] *The Diving Bell and the Butterfly*, 2007, France: PathéRenn Productions. Film, Directed by Julian Schnabel.
- [3] [Online] Available: <http://www.nationalmssociety.org/>
- [4] ALS Association. [Online]. Available: <http://www.alsa.org/>
- [5] Bedford, MA: Appl. Sci. Lab. [Online]. Available: <http://www.a-s-l.com>
- [6] Y. L. Chen, F. T. Tang, W. H. Chang, M. K. Wong, Y. Y. Shih, and T. S. Kuo, "The new design of an infrared-controlled human-computer interface for the disabled," *IEEE Trans. Rehabil. Eng.*, vol. 7, no. 4, pp. 474–481, Dec. 1999.
- [7] Don Johnston, Inc., *Infrared head-mounted mouse alternative*. Penny & Giles HeadWay. [Online]. Available: <http://www.donjohnston.com>
- [8] D. G. Evans, R. Drew, and P. Blenkhorn, "Controlling mouse pointerposition using an infrared head-operated joystick," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 1, pp. 107–117, Mar. 2000.
- [9] LC Technologies, *Eyegaze system*. [Online]. Available: <http://www.lctinc.com>
- [10] Madentec. [Online]. Available: <http://www.madentec.com>
- [11] Tash solutions. [Online]. Available: <http://www.tashinc.com>
- [12] R. Vaidyanathan, B. Chung, L. Gupta, H. Kook, S. Kota, and J. D. West, "Tongue-movement communication and control concept for hands-free human-machine interfaces," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 4, pp. 533–546, Jul. 2007.
- [13] L. Young and D. Sheena, "Survey of eye movement recording methods," *Behav. Res. Meth. Instrum.*, vol. 7, no. 5, pp. 397–429, 1975.
- [14] R. Barea, L. Boquete, M. Mazo, and E. López, "System for assisted mobility using eye movements based on electrooculography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 4, pp. 209–218, Dec. 2002.

- [15] P. DiMattia, F. X. Curran, and J. Gips, An Eye Control Teaching Device for Students Without Language Expressive Capacity: EagleEyes. Lewiston, NY: Edwin Mellen Press, 2001. [Online]. Available: <http://www.bc.edu/eagleeyes>
- [16] Y. Ebisawa, "Improved video-based eye-gaze detection method," *IEEE Trans. Instrum. Meas.*, vol. 47, no. 4, pp. 948–955, Aug. 1998.
- [17] A. Gee and R. Cipolla, "Determining the gaze of faces in images," *Image Vis. Comput.*, vol. 12, no. 18, pp. 639–647, Dec. 1994.
- [18] T. E. Hutchinson, K. P. White, Jr., W. N. Martin, K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, no. 6, pp. 1527–1533, Nov./Dec. 1989.
- [19] Q. Ji and Z. Zhu, "Eye and gaze tracking for interactive graphic display," *Mach. Vis. Appl.*, vol. 15, no. 3, pp. 139–148, Jul. 2004.
- [20] A. Kapoor and R.W. Picard, "Real-time, fully automatic upper facial feature tracking," in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recog.*, Washington, DC, May 2002, pp. 10–15.