Facial Expression Based Computer Cursor Control System for Assisting Physically Disabled Person

M. Vasanthan
School of Mechatornic Engineering,
University Malaysia Perlis (UniMAP)
Perlis, Malaysia
vasan6647@yahoo.com

M. Murugappan
School of Mechatornic Engineering,
University Malaysia Perlis (UniMAP)
Perlis, Malaysia
murugappan@unimap.edu.my

R. Nagarajan
School of Mechatornic Engineering,
University Malaysia Perlis (UniMAP)
Perlis, Malaysia
nagarajan@unimap.edu.my

Mr. Bukhari Ilias
School of Mechatornic Engineering,
University Malaysia Perlis (UniMAP)
Perlis, Malaysia
bukhari@unimap.edu.my

J. Letchumikanth
School of Mechatornic Engineering,
University Malaysia Perlis (UniMAP)
Perlis, Malaysia
letchu8787@yahoo.com

Abstract – Recent years, several researchers are developing different kinds of assistive devices for physically disabled peoples. In this work, movement of illuminant markers through facial expressions is used to control the cursor movement in computer applications. A set of five facial expressions namely left and right cheek movement, eye brow rise and down and mouth open are used for controlling cursor movement in left and right direction, up and down and click, respectively. Four very small luminous stickers are fixed on subject's face and the subject is instructed to perform the above said facial expressions. Conventional web-camera is used for capturing the facial expression and sends the data into BASIC STAMP microcontroller through serial port interfacing. Movements of markers are detected through its x-y coordinate's changes on the video image and each facial expression is uniquely represented by a binary number. As a result of change of x-y co-ordinates, the BASIC STAMP microcontroller sends the binary code to the computer for controlling the mouse actions.

Keywords - Facial features, facial expressions, cursor control, BASIC STAMP microcontroller.

I. INTRODUCTION

Application of personal computer (PC) is widely spread over all the areas such as communication, environmental control, education, research and entertainment and used by several peoples. However, physically disabled peoples such as bed-ridden, paralyzed, etc are facing a huge challenge on using the computers for their day to day activities. Several research works have been reported on assistive device development for physically challenged peoples for their easy living. Most of the researchers have developed a head-operated mice or joysticks in order to interact with a computer and to type their required aid on the screen through keyboard [3]. However, these devices are higher in cost and required pre-learning before going to use the systems. Some of the researchers in United Kingdom have developed assistive devices based on movement of infrared transmitters on their head. To name few devices based on this working principle are, Prentke Romich's HeadMaster Plus [1], Orin Instrument's HeadMouse [2], etc. However, some of the assistive devices are developed at lower cost such as Granada Learning's Head Mouse [3] and Penny and Gilles' device (£400, RM2400). In general, most of the assistive devices are based on the subject head movements and required to wear a relatively complex piece of equipment either on their head, face, etc.

In this work, we aim to develop a simple, portable and low cost assistive device through non-contact interface technology. Through our system, the subject is free from wearing any kind of controller to control the cursor movement in PC. The proposed system is only requires a microcontroller, few number of illuminant stickers and conventional web camera for easier cursor control. The web camera is used to capture the facial expression of the subject and interfaced with the PC through serial port interfacing. The PC is continuously track the movement of markers and sends the information about the X-Y co-ordinates of the cursor position to the microcontroller. The microcontroller will send binary information related to different facial expression to the cursor for controlling its movement.

In this work, the position of the face fixed in the range of rectangular to avoid the effects from any kind of background noises and movements. A set of four luminous stickers are placed at left cheek, right cheek, mouth, and centre of forehead on the face. Within the rectangular areas, the system will detect the four circles corresponding to the circular shape markers and its position. Movement of markers in X-Y co-ordinates of the circle is detected and sent it to the PC for moving the cursor. If the right cheek marker is moved on right side, the system will increase the value of X coordinate from the current position of X-axis. Hence, the cursor will be moved to right side continuously until the facial expression is stopped. In a similar manner, if the left cheek is moved, then the system will decrease the X coordinate from the current position of X-axis. When mouth opens, the system will decrease the value of Y coordinate from the current Y-axis. During eye brow rises, the value of Y coordinate will be increased from the current Y-axis value. However, if both left and right cheeks are moved (considered here as smiling) simultaneously, the system will send a command to "click" the cursor. This operation will continue until user stop the system. This project has been implemented using Labview software for image processing operations and programming in microcontroller to control the cursor.

II. LITERATURE REVIEW

Development of assistive devices based on computer control is increasing day by day. These assistive device can be used for the physically challenged people to interact with many applications such as communication (mobile phone, tablet PC), living environmental control (intelligence home appliance control), education and entertainment (video games). In order to design a human assistive device based on facial features, it is essential to review the research works related to facial features recognition methods. Recent years, vision based systems are more popular on designing and developing different types of application oriented systems such as biometric, security systems, etc. Face recognition systems are being tested and installed in the airports to provide new level of security [4]. Indeed, human-computer interfaces (HCI) based on facial expression and body gestures are being exploited as a way to replace the traditional interfaces such as the mouse and keyboard [4]. And actions of face could play an important complementary or supplementary role to that played by the hands in machine-interaction (HMI) [5].

Ideally, acquisition of facial features is uses as an automatic face detector and it allows locate the faces in complex scenes with cluttered backgrounds [6]. Here, the limitations on out-of-plane rotations can be addressed by using different type of warping techniques. In this approach, the center positions of distinctive facial features such as the eyes, nose and mouth are considered as a reference points in order to normalize the test faces according to some generic or reference face models [7]. Scale changes of faces may be tackled by scanning images at several resolutions in order to determine the size of present faces, which can then be normalized accordingly [7, 8].

Several face detection algorithms have been developed by different researchers. In general, face detection techniques are categorized into two main categories, the feature-based approach and the image-based approach [9]. Most of the vision based approaches are exhaustively scan the facial images at different scales for face detection. However, this approach requires more time for facial recognition and window scanning technique is more complex. In contrast to this approach, the feature-based approach eventually leads to the localization of facial features for face recognition. Hence, most of the real time applications are adopting with feature based approach for face recognition [10].

The most well-known method that is applied in the feature based approach is skin model which is effective in image segmentation and face extraction. The inspiration to use skin color analysis for initial classification of an image into probable face and non-face region stems from a number of simple but powerful characteristics of skin color [11]. Firstly, processing skin color is simpler than processing any other facial feature. Secondly, under certain lighting conditions, the color is orientation invariant. The major difference between skin tones is intensity, e.g. due to varying lighting conditions and different from the color of most other natural objects in the world [12]. Yang [13] presented a review of face detection methods, which fell into a small number of categories, the most important of which were methods depending of the color and physiognomy of the face. Störring[14] suggested that a face's apparent color was due to two factors: the amount of melanin in the skin and the ambient illumination. Of the two, ambient illumination caused the greater variation to the perceived color. They concluded that if normalized color values were used (i.e. the effect of illumination were removed) then skin colors were consistently with a fixed and quite narrow set of limiting values, independent of the subject's natural skin coloration. This approach, and slight modifications, has become very popular due to its simplicity [15,16,17,18,19].

For mouse control module, the conversion from human motion parameters to mouse cursor navigation can be categorized into direct mode, joystick mode, and differential mode. For direct mode, a one-to-one mapping from the motion parameter domain to screen coordinates is established by off-line calibration or by design based on the prior knowledge about the human-monitor setting [15]. Joystick mode navigate mouse cursor by the direction (or the sign) of the motion parameters. And the speed of the cursor motion is determined by the magnitude of the motion parameters [20]. In differential mode, the cumulating of displacement of the motion parameters drives the navigation of the mouse cursor, and some extra motion parameter switches on/off the cumulating mechanism so that the motion parameter can be reset without influencing the mouse cursor. Therefore this mode is very similar to hand-held mouse mode: user can lift mouse and move back to the origin on mouse pad after performing a mouse dragging operation [20]. After mouse cursor is navigated to desired location, the execution of mouse operations, such as mouse button clicks are triggered by detection of specific motion patterns. The most straightforward trigger is that some motion parameters exceed specified thresholds. In [21], a mouse-click event is triggered by "dwell time", e.g. a mouse click is triggered if the user keeps the mouse cursor still for 0.5s. In [22], the confirmation and cancelation of mouse operations is conveyed by head nodding and head shaking. Timed finite state machine is designed to detect the nodding and shaking in the raw motion parameters. The other method used to control cursor is by moving nostril [23]. The distances from the nostril point to the boundaries of the face region in the vertical and horizontal directions is computed. This is translated directly into cursor position coordinates by linear scaling. The scaling coefficients are defined by a user specific calibration stage, we require that the user's maximum nostril movements in the horizontal and vertical directions will map into movements of the cursor that completely cover the monitor's screen. In a future version of the software we will replace the absolute cursor positioning with a joystick-like control method, whereby the cursor's velocity is controlled by the nostrils' positions as this offers improved cursor positioning.

Brunelli [24] used template matching for face recognition. The algorithm prepares a set of four masks representing the eyes, nose, mouth and face for each register person. To identify the unknown person in the image, the algorithm first detects the eyes via template matching and then normalized position, scale and rotation of the face in the image plane using the detected eye position. Next, for each person in the database, the algorithm places four masks on their positions relative to the eye position and computes the cross-correlation between the four masks and the image. The unknown person in the image is classified as the person giving the greatest sum of the cross-correlation values to the four masks. This method can be used for detecting facial expression also. This method can identify the movement of the eyes, nose, and mouth within the masks. From there it can easily identify the facial expression.

The goal of this project is to use face expression as an alternative way to control the cursor movement. It's very useful project for the physical disabilities people. The most obvious way of doing this is to use a webcam, interfaced to the laptop. The system automatically detects the user's face by means of the algorithm of Viola and Jones[25]. The user should just stay still for a second to get the initial position of the face and stickers. The dynamic movement will track by the system and give an output to the cursor operation system. The cursor will move according to the output data given by the system. This system is less cost easy to handle.

III. METHODOLOGY

The methodology of this project is divided into two main parts, namely Detection of Facial Expression and Data Acquisition and Cursor Control.

A. *Data Acquisition*:

In this work, four numbers of Luminous Stickers are cut into 1cm diameter and pasted on the specific locations on the user face (forehead, right cheek, left cheek and mouth). Conventional web camera is used for capturing the facial expressions of the subject. During the system is initialization, the webcam will be ON and capture the facial image data for processing. A set of five facial expressions namely left and right cheek movement, eye brow rise and down and mouth open are used for controlling cursor movement in left/right direction, up/down and click, respectively.

B. *Facial Expression Detection*

After the system start initializing, the webcam will be ON and it starts acquiring the image (Figure 1). Acquired images of different resolutions are converted into 8 bit per pixel. This will highly reduce the computational time and memory requirement of the processor to perform the cursor control. Acquired images are first converted into gray scale images and then preprocessed using median filter to reduce the effect of background noises. In this work, edge detection method used to detect the shape of face and luminous sticker circles (Figure 2). Thresholding method is used to find the perfect circles of luminous stickers on the face (Figure 3). The initial position of the luminous markers in X-Y coordinates are given in Table 1. These X-Y coordinates position are passed to the central processing unit (CPU) with Region of Interest (ROI) values. In the case of detecting the circles in the ROI region, the system can send the types of facial expression to the microcontroller for sending the control command to the cursor.

C. *Cursor Control*

Serial communication is used for interfacing the CPU and microcontroller. Based on the changes in X-Y coordinates of the luminous stickers, the output in the form of American Standard Code for Information Interchange (ASCII) code is send to the microcontroller through serial port at the baud rate of 9600. In this work, we have used five ASCII characters for five different actions and are represented by a binary numbers. Thereby, the cursor is moved according to the input data derived from the microcontroller. Complete algorithm has been implemented by using LabView Software.

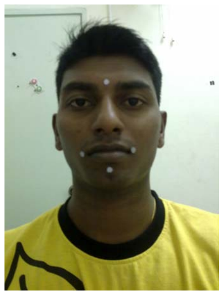

Figure 1 – Webcam image

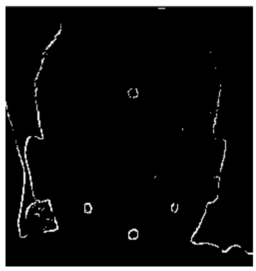

Figure 2 – Gray scale *Figure 3 – After image processing*

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this work, we have tested this cursor control algorithm with 15 subjects of 8 males and 7 females in the mean age of 24.5 years. Region of Interest (ROI) of rectangle form is fixed for all four stickers in the video stream. The circular form of sticker is in the center of the rectangle ROI. The value of ROI coordinate is changed according to the user face. Using the image processing techniques, edge detection method is used to trace the sticker movement in the ROI. The area of the circle detection is limited to 50 to 150 pixels. Facial test image after setting the ROI and the detection of the circle within ROI is shown in Fig 4. If the movement of sticker is beyond this pixel value, the system does not send any comment to the microcontroller for cursor action and it is still in same position of last sticker movement. In the case of user to control the cursor movement, the user face should be brought to the initial position.

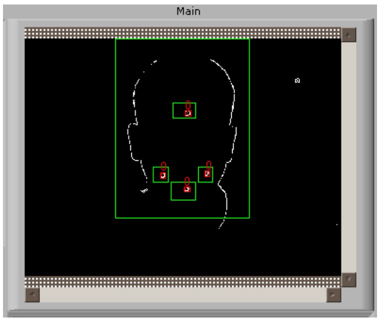

Figure 4 – Position ROI and Sicker Detection

An algorithm which uses area of the circle is developed to identify the center of the circle. The system then trace the movement of the center coordinates and compute in X-Y coordinates form. All the information about the coordinates of the stickers is collected in cluster format. From this data, we separate the important data such as center coordinate area of the sicker and its diameter by using array format. The center coordinates of X axis and Y axis is used as a reference point for the static position. The movement of X-Y coordinated is used as input for the system to detect the up, down, left and right values of X-Y coordinates for each sticker. These X-Y coordinates are then used for finding the suitable action to create an output data. In this work, there are five output cursor movements such as up, down, left, right and click for each expression namely eye brow rise up, open mouth, left cheek move, right cheek move and smile respectively. In the case of no user face is detected by the system then the algorithm will not send any comment to the microcontroller for cursor movement. The algorithm will send an output when any one of the ROI is not filling with an input user face (object). Initially all outputs are initialized to '0' and if there is a movement, the output will be '1' for the relevant sticker motion. This output will be automatically sent to the microcontroller through to serial port.

TABLE I. DATA SENT BY THE SYSTEM AND RECEIVED BY THE MICROCONTROLLER.

	Send		**Receive**	
Expression	**Binary**	**Sending Data**	**Receiving data**	**Action**
Rise Brow	10000	a	a	+Y-direction circuit activated (move up)
Left Cheek	01000	b	b	-X-direction circuit activated (move left)
Right Cheek	00100	c	c	+X-direction circuit activated (move right)
Open Mouth	00010	d	d	-Y-direction circuit activated (move down)
Smile	00001	e	e	Active Click

Table 1 shows the data that send by the system and received by the microcontroller basic stamp. In the LabView software, "CASE" method was used to send the data serially to the output to microcontroller. In this "CASE" method, a set of ASCII letter with respect to the binary value is created to differentiate the user facial expression and it should be defined before the algorithm start. This will send data to the microcontroller only if the input is same as the case structure in the box. For example, the expression rise eye brows have the binary value '10000'. When this action was done, the system will send '10000' to the case box. As defined earlier, the box will create output as 'a' and this value 'a' will be send to microcontroller. After microcontroller receives the data, the processer will process it and move the cursor in the Y direction circuit will activate to move the cursor upward. Similarly, all other facial actions are used for controlling the X-Y coordinates of the cursor to control its movement.

V. CONCLUSION

In this work, we have successfully utilized the four stickers to control the movement of cursor in computer applications. The proposed methodology is very simple and efficient on controlling the cursor movement in an efficient manner. The system also have enough intelligent in acquiring the users image from web camera. This system will replace the use of conventional head movement based PC control by physically challenged people. However, this system has the following disadvantages: (i) the output of cursor control is very sensitive on lighting condition (ii) wearing of four markers permanently on the user face is highly intrusive for the subjects (iii) the response time on cursor control after face detection is slow due to the lower clock frequency of the microcontroller. Hence, the current focus on this work is relies on expanding the design using high speed microcontroller with different light settings.

VI. ACKNOWLEDGEMENT

This work is financially supported by Ministry of Higher Education (MOHE), Malaysia through Fundamental Research Grant Scheme (FRGS). Grant Number: 9003-00254.

6. References

[1] Prentrom 2001, http://store.prentrom.com/. retrieved on August 2010

[2] Orin Instruments, 2001 http://www.orin.com/access/: retrieved on August 2010

[3] R. Drew, S. Pettitt, P. Blenkhorn and D.G. Evans (1998), A head operated 'joystick' using infrared. In: A.D.N. Edwards, A. Arato and W.L. Zagler, Editors, *Computers and Assistive Technology ICCHP '98, Proc XV IFIP World Computer Congress*, *Österreichische Computer Gesellschaft*.

[4] Christopher, A. W. and Xiuwen, L.(2002) Face Detection Using Spectral Histograms and SVMs[Online]. The Florida State University. www.nd.edu/~kwb/nsfufe/Security Privacy.pdf

[5] Michael J. Lyons,(*2004),* Journal of Facial Gesture Interfaces for Expression and Communication, *IEEE International Conference on Systems,Man and Cybemetics*.

[6] B. Fasel, and Juergen Luettin, Ascom Systec AG, (2006), Applicable Research and Technology, Gewerbepark CH-5506, Maegenwil, Switzerland.

[7] I. Essa and A. Pentland, (1997), Coding, analysis, interpretation and recognition of facial expressions. IEEE Trans. Pattern Anal. Mach. Intell.

[8] H. Rowley, S. Baluja and T. Kanade, (1998), Neural network-based face detection. IEEE Trans. Pattern Anal. Mach. Intell. 2001.

[9] Erik, H. (2001). Face Detection: A Survey. [Online]. Department of Informatics, University of Oslo.http://www.tsi.enst.fr/~chollet/Biblio/Articles/Domaines/BIOMET/Face/sdarticle.pdf

[10] Shinjiro, K. and Jun, O. (2000, Aug.) Automatic Skin-color Distribution Extraction for Face Detection and Tracking. [Online]. ATR Media Integration and Communications Research Laboratories. www.cipprs.org/vi2002/pdf/s7-1.pdf

[11] T. Ezgi, (2002) Face Detection Using a Mixture of Subspaces. [Online]. The Middle East Technical University.www.ece.arizona.edu/~pgsangam/ietepaper.pdf.

[12] M. Artur, W. Stefan, and R. Martin Line [1996] Based Robot Localization under Natural Light Conditions[Online].www.ni.uos.de/fileadmin/user_upload/publications/merke_welker_riedml.ecai04ws.pdf

[13] M. -H. Yang, N. Ahuja, D. Kriegman and A. survey, (1999), A survey on face detection methods, *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[14] M. Störring, H.J. Andersen and E. Granum, (1999), Skin colour detection under changing lighting conditions. In: H. Araujo and J. Dias, Editors, *Seventh International Symposium on Intelligent Robotic Systems, Coimbra, Portugal* , pp. 187–195.

[15] K. Toyama, Look, (1998) Ma—No Hands! hands free cursor control with real-time 3D face tracking, *Proceedings of Workshop on Perceptual User Interface (PUI'98)*.

[16] J. Sobottka and I. Pittas, (1996.) Segmentation and tracking of faces in color images, *In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition* p. 236–241

[17] J. Yang and A. Waibel, (1996), A real-time face tracker, *Proceedings of third IEEE workshop on application of computer vision*, pp. 142–147. View Record in Scopus | Cited By in Scopus (140)

[18] B. Schiele and A. Waibel, (1995.) Gaze tracking based on face-color, *Proceedings of the international workshop on auto. Face and gesture recognition, Zurich* p.p. 344–349.

[19] C.L. Zitnick, J. Gemmell and K. Toyama, (1999) Manipulation of video eye gaze and head orientation for video teleconferencing, *Technical Report MSR-TR-99-46*, Microsoft Research, Redmond, WA, p.p. 12.

[20] D. Gorodnichy, S. Malik, G. Roth, Nouse (2002), 'use your nose as a mouse' - a new technology for hands-free games and interfaces, in: Intern. Conf. on Vision Interface, Calgary, pp. 354–361.

[21] B. M., G. J., F. P., (2002), The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities, IEEE Trans Neural Syst Rehabil Eng. 10 (1) pp1–10.

[22] J. W. Davis, S. Vaks, (2001), A perceptual user interface for recognizing head gesture acknowledgements, in: Workshop for Perceptive User Interface, Orlando, FL, pp. 1–7.

[23] Evans DG, Drew R, Blenkhorn P. (2000), Controlling mouse pointer position using an infrared head operated joystick. IEEE Transactions on Rehabilitation Engineering pp107–17.

[24] R.Brunelli, T.Poggio,(1993) "Face recognition: features versus templates", IEEE Trans. Pattern Anal. Mach. Intell., vol. 15, no.10, pp.1042-1052.

[25] Viola and Jones, (2004) P. Viola and M. Jones, Robust real-time face detection, Int J Comput Vis 57