Design and Implementation of a Mouth-Controlled Mouse

Chu-Hsing Lin
Department of Computer Science
Tunghai University
Taichung, Taiwan
chlin@thu.edu.tw

Shang-Sheng Yang
Department of Computer Science
Tunghai University
Taichung, Taiwan
s942921@thu.edu.tw

Jeng-Heng Chen
Department of Computer Science
Tunghai University
Taichung, Taiwan
s942941@thu.edu.tw

Yi-Zhu Cao
Department of Computer Science
Tunghai University
Taichung, Taiwan
g98350017@thu.edu.tw

***Abstract*—Nowadays, people try to make more convenient assistive device for special needs persons. In this modern society, how to combine technology with assistive device has become a popular issue. However, assistive devices of hand impaired person for using computer are lack of demand. Therefore, to make the assistive device for whom are unable to use mouse is the object of this paper. In this paper, we implement mouth-controlled mouse using face detection and tracking system. The images of mouth moving are captured into computer by webcam, and then we can control cursor and click by moving mouth. With the technical equipment, we can surely improve living quality of disabilities.**

Keywords-mouth-controlled; mouse; assistive device

I. INTRODUCTION

In recent years the development of technology has been greatly assisted by the application and advancement of computers. Take biological technology and astronomy for example. Computers' powerful calculating abilities have extended the range of human's observation to the extent that we have been provided with endless opportunities to increase our awareness and knowledge of many aspects in our life [1].

As a result, we have chosen to apply a simple facility of computing to create this experiment, with the prediction that we may use other parts of our body to control computers rather limiting usage to only hands. Windows are installed in most of the computers, and Mouse plays an important role in using this software. Therefore, our plan is to use the human's head to control the mouse. The mouse mostly controlled by hands, is not really convenient to many disabled people. As a result, we have envisioned using other parts of our body to substitute for hands or to help hands to control computers. We have considered several factors, and most importantly, hands and feet are not available for most of disabled people. We have decided to control mouse by mouth [2].

In this research, the most important task is the shape of the mouth and how to judge when the mouth is shut. To test this we created a systematic process. Firstly we captured the webcam images and then flip them, and then we removed color noise, then applied detection and treatment, identification. Finally with a number of improvements for the algorithm, with a view to the user for easier use we were able to move the mouse. A brief description of the system is as follows: At the beginning, the webcam transports the image to the program in the computer dealing with replacement processing. Then a comparison of the images from the identification system to that of the real person is made. The original RGB image is converted into HSV, and the default threshold of HSV is used to identify the mouth, erosion and dilation are used to remove noise. Finally using their mouth the user finds a way to connect to and position the location of the cursor, a finally analysis of the mouth-type will determine when the 'click' function is applied.

II. RELATED WORK

Recently, the research in the area of assistive device which integrated computer science is becoming more and more. In a conventional keyboard, like a QWERTY keyboard, there are too many keys such that the spaces between neighboring keys are too small for physical disabled, Sheng-Wen Shih and Chan-Hao Huang proposed a novel prosthetic keyboard with reduced number of keys such that the space between neighboring keys is sufficient for physically disabled [3]. Given only 12 keys in the designed keyboard, and each character encoded by using radix-12 Huffman algorithm. They compare the performance of three different input methods, the proposed Huffman method can help the hand impaired person to type more words per minutes than other two methods.

III. PROPOSED MECHANISM

The propose of this paper is to find the position of mouth and to identify the shape of an opened and closed mouth. We obtain the image which was captured by webcam, and identify the shape of the mouth. We hope that disabilities will control the cursor easily with our system. The following are introductions for this system: After the system begins, the system changes the positions of images from left to right; because the image taken by a webcam is opposite to the real person. Next, the system changes the RGB image into HSV [4][5], with the default setting of HSV to find the position of mouth. The system uses erosion and dilation to erase the stains, and then connects with the webcam to locate the user's mouth. Finally, the system can locate the position of cursor, and then analyze the shape of mouth to identify when the users 'click' the Mouse.

A. *Flip Loaded Image*

The image captured from webcam will be inversed from user's view. So we need to flip the loaded image to make it appear like a mirror of the users on the screen. To flip an image, we need to know the width of the image, and then make a flipped image with a function shown in (1).

$$Pfilpped(x, y) = Porig(Width - x, y) \quad (1)$$

B. *HSV Color System*

First of all, we use the RGB pixel number to find the range of RGB pixel number for the mouth. However, the RGB color system is easily intermittent with a great influence of light source. As a result, two similar colors could have different RGB numbers, and the difference in lighting can greatly influence the color of the lips.

So we locate the mouth on HSV color space, which H stands for Hue, S stands for Saturation, and V stands for Value. The conversion is from RGB to HSV. The RGB value is between decimal 0 to 1, and conversion from 0 to 255. Information from the HSV color map shows HSV color is more continuous and less susceptible to the effects of brightness. It's easier to set a specified set of specific HSV, and the results have less noise.

$$\mathrm{H} = (0 + (G - B)/(MAX - MIN)) \times 60, \text{ if } \mathrm{R} = \mathrm{MAX} \quad (1)$$

$$\mathrm{H} = (2 + (B - R)/(MAX - MIN)) \times 60, \text{ if } \mathrm{G} = \mathrm{MAX} \quad (2)$$

$$\mathrm{H} = (4 + (R - G)/(MAX - MIN)) \times 60, \text{ if } \mathrm{B} = \mathrm{MAX} \quad (3)$$

$$S = MAX - MIN \quad (4)$$

$$V = MAX \quad (5)$$

C. *Erosion and Dilation*

Although using the results of HSV better than the use of RGB, there still is some noise. The proposed method, erosion and dilation, is the method we have applied to achieve the removal of the effect of noise. First, we use Erosion to filter out the smaller noise, but the scope of mouth will also be affected by deformation, so it's necessary to use Dilation to recover the mouth. In addition, usage of Dilation and Erosion once is the most efficient. A comparison for the steps is then used to find the largest set, the part we want, as shown in Fig. 1 [6].

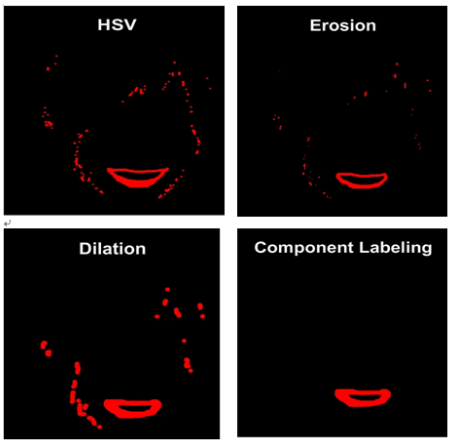

Figure 1. The results of various steps

D. *Component Labeling*

First step, we begin searching from the point of upper left side, and employ Breadth-First-Search (BFS). If we finish checking all points, then we can find the largest set. The detailed steps are as follows:

1) Fig. 2: Search point (1, 1) is 1 from left to right, then push it in the queue. Take out (1, 1) from the queue and search 1 from top to down.

2) Fig. 3: Modify point (1, 1) to 1 of the another array, and modify point (1, 1) to 0 of the original array.

3) Fig. 4: Search point (1, 4) from left to right then push it to the queue.

4) Fig. 5: Take out (1, 4) from the queue and search the neighbor block to find (1, 5) and (2, 4) then push them to the queue. And modify point (1, 4) to 2 of the another array and point (1, 4) to 0 of the original array.

5) Fig. 6: Take out (1, 5) from the queue and search the neighbor block to find (2, 5) is 1 then push it to the queue, and modify point (1, 5) to 2 of the another array and point (1, 5) to 0 of the original array.

6) Fig. 7: Take out (2, 4) from the queue and search the neighbor block to find (2, 3) is 1 then push it to the queue, and modify point (2, 4) to 2 of the another array and point (2, 4) to 0 of the original array.

7) Fig. 8: We can get the largest group of the final.

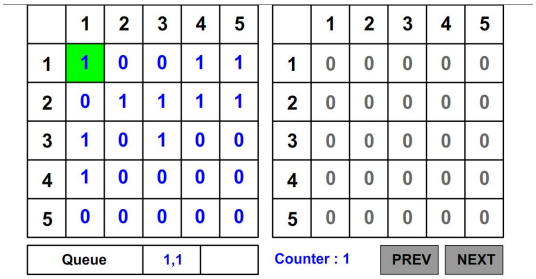

Figure 2.

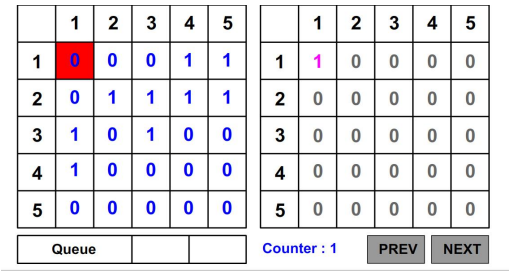

Figure 3.

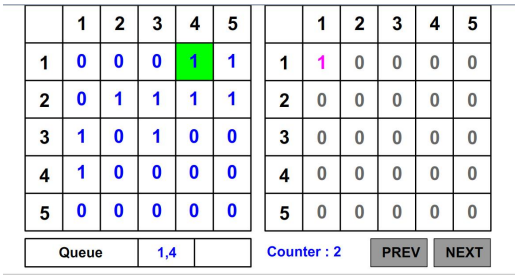

Figure 4.

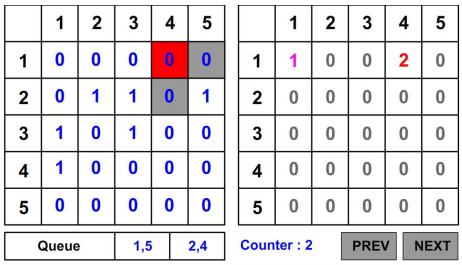

Figure 5.

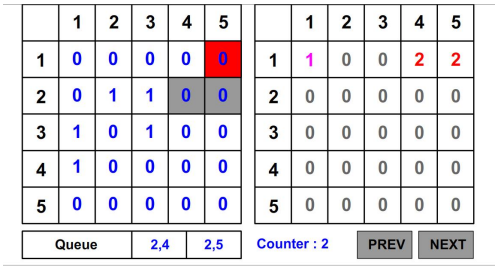

Figure 6.

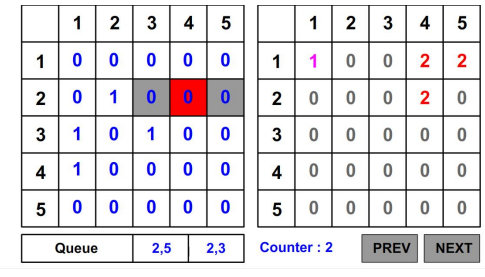

Figure 7.

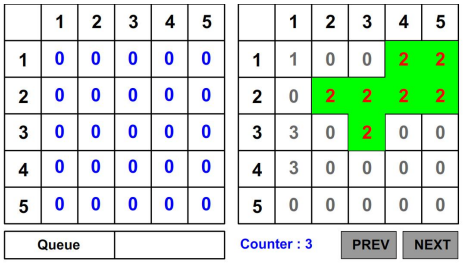

Figure 8.

E. Mouse Moving Control

We can get the largest group from the captured image through several steps. From this group, two points (MIN_x, MIN_y) and (MAX_x, MAX_y) could be found, and then we figure out the center point of mouth sketch which is painted by these two points. The shift of center point would show us if the cursor moves or not. We define two kinds case of cursor movements; We average the five times of center point. If the value under threshold we set initially, then we determine to control the cursor stay still; but when the value is beyond the threshold, we make cursor move.

*1) CASE 1:*We draw five points into array in order, then we had the fifth coordinate value subtract the average of first four coordinate values to reducing the effect of noise and avoiding mouse irregularly moving.

2) CASE 2: From the average obtained by CASE1, we add pre-five points and the new point together while the new point is in a higher proportion. Such as: 0.4*(the average of pre-five times) + 0.6*(the value of the new center), this approach allows the cursor to move instantaneously.

F. Mouse Click

In order to decide whether click mouse or not, the mouth shape must be detected. We get the mouth sketch through several steps as mentioned above, and infer that the mouth shape from the aspect ratio of the mouth sketch. It was observed that the aspect ratio of mouth sketch approximate 1 when opening mouth. In other words, the aspect ratio of mouth sketch between 2 and 4 when closing mouth, our research identify mouth opening with left key pressing of mouse, and vice versa.

In addition, we set the threshold value for noise interference. First of all, using a one-dimensional array to record the most recent 5-times the mouth size of pixel group,

and removed the value when not detecting the mouth, that is, when group size=0, we record the average value of the 5-times and compared to the recent group size. Given the scope of error, this can be removed with the significant difference of minimum and maximum values; thus effectively reducing the noise caused by interference.

IV. Experiments & Discussions

Fig. 9 shows the program interface, there are six sub-windows and adjustment buttons in the program. The upper sub-windows are mirror, HSV, erosion, dilation and mouth image respectively, the lower sub-window is the image of webcam. The adjustment button adjusts appropriate value of HSV threshold.

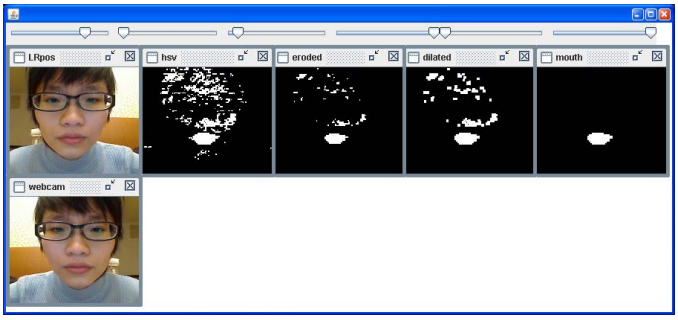

Figure 9. The program interface

We evaluate the performance of mouth-controlled mouse through two experiments. One is to draw a square (6cm*6cm), and we record the spending time. In Fig. 10, we can see that, the spending time of mouth-controlled mouse is 5.77 times of the hand-controlled mouse. The result shows that the cursor moves faster by hand.

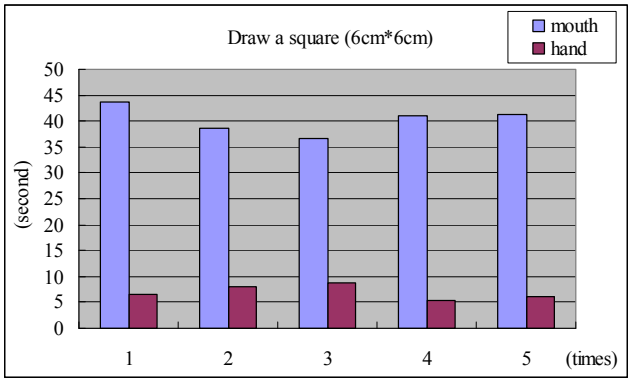

Figure 10. Spending time of drawing a square

Although we can detect if the mouth open or not, there are several problems need to be improved. First of all, the system may identify the whole mouth as upper lip or underlip. Secondly, the system may also fail to identify the mouth shape when unpredictable objects appear such as another mouth, fringe, and the reflect light of glasses. These may affects the detection result, even the clothing that user wears may cause an error. Table I shows the accuracy of detection, it indicates that the accuracy of mouth opening (aspect ratio is 1) is higher than mouth closing (aspect ratio is larger than 2); The accuracy are higher when mouth stay still whether users opens mouth or not, thus we see the detection result is susceptible to mouth moving.

TABLE I. The Accuracy of Mouth Shape Detection

Moving Mouth or Not	Accuracy	
	Opening mouth	***Closing mouth***
stay still	100%	93%
keep moving	91%	77.5%

V. Conclusions

In this research, we have controlled the cursor and click functions, which are the two main functions of mouse. However, we still have several problems need to be solved. For example, the webcam is fixed, and it's difficult to identify the position of the mouth especially when user nod or raise his/her head. That is a great restriction when moving the cursor. The light of the environment would have a huge influence on processing the image captured by webcam. In this research, we have chosen the HSV mode to decrease the influence of lights which mentioned above. However, the function of our system is still insufficient, for example, double click of mouse and right key of mouse are not allowed in the research, it is definitely need to be improved.

In the future, we hope that this technique can be developed to be widely used in any public computer and help disable people use the computer conveniently.

VI. Acknowledgment

This work was supported in part by the National Science Council under the grant NSC99-2221-E029-034-MY3.

References

[1] W. J. Cao, "Finger-Guessing Game," National Central University Library Electronic Theses & Dissertations System, vol. 945201096, pp. 1-52, 2006.

[2] J. Kuch, "Vision-based hand modeling and gesture recognition for human computer interaction," M.S. Thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1994.

[3] Sheng-Wen Shih, and Chan-Hao Huang, "Improving the input speed of a multi-key prosthetic keyboard based on Huffman coding," National Digital Libraty of Theses and Dissertations in Taiwan, pp. 1-93, 2011

[4] O. Ikeda, "Segmentation of faces in video footage using HSV color for face detection and image retrieval," Proc. International Conference on Image Processing (ICIP 2003), Sep. 2003, vol. 2, pp. 913-916.

[5] C. M. Pun, and C. F. Wong, "Image retrieval using a novel color quantization approach," Proc. International Conference on Signal Processing (ICSP 2008), Oct. 2008, pp. 773-776.

[6] M. Reiter, R. Dormer, G. Langs, and K. Bischof, "3D and Infrared Face Reconstruction from RGB data using Canonical Correlation Analysis," Proc. International Conference on Pattern Recognition (ICPR 2006), Aug. 2006, vol. 1, pp. 425-428.