A Robust Low Cost Virtual Mouse Based on Face Tracking

Guoqing Xu[1,2], Yangsheng Wang[1], Xuetao Feng[1]

[1] Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China
E-mail:guoqing.xu@ia.ac.cn

[2] School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430074, China

Abstract: This paper introduces a non-contact virtual mouse controlled by tracking the face movement. It is specially designed for the disabled with mobility impairments in the upper extremities. The whole work is based on the common web camera. Improved optical flow algorithm is proposed through locating the facial features defined by active appearance model. Experiments show it is robust to the various background and environment illumination. It also can effectively track the features when the face moves fast in a large scope.

Key Words: Pattern recognition, face detection, human computer interaction

1. INTRODUCTION

Gaze tracking is one of the key researches of convenient and efficient non-contact natural Human Computer Interaction(HCI) design. It can help the people with disabled limb movement to operate computer with eye gaze [1, 2, 3, 4, 5]. It is usually based on infrared lights and high-resolution image acquisition equipments. In addition, the calibration process is complicated. The system is not easy to carry. In recent years, some researchers have used low cost common web camera to acquire face images. Many effective image analysis algorithms have been proposed to detect the face and track the facial features in the video sequences. These methods do not need infrared lights and special video equipments. So they have fewer restrictions on the calibration procedure, and have become feasible research directions.

This paper is organized as follows. Section 2 presents the related work. In section 3, the face features location algorithm based on AAM is proposed. Considering the diversity of the training sample, the algorithm can adapt various pose, head tilting and even face being partially occluded. Section 4 describes the improved optical flow blink detection based on the local eye area. Section 5 gives the experiment result of a real application.

2. RELATED WORK

Virtual mouse can help the user to control the cursor in the screen through tracking the face or head movement. It can be realized using a common camera in front of the user. The head and face trajectory of the user can be converted to the movement of the screen cursor. Various image vision and pattern recognition algorithms have been proposed, e.g. Betke[6] presented a system initialized by manual selection of the face area. Grauman [7] used the difference images between two consecutive camera frames to detect the eye blink and eyebrow raise. Paola Campadelli[8] proposed an algorithm for the automatic detection and description of facial features in 2D color images of either frontal or rotated human faces in color images with homogeneous light-colored background and frontal diffuse illumination. Michael Chau[9] presented a real-time vision system that automatically detected a user's eye blinks and accurately measured the blink durations. It enabled the disabled to communicate using blink pattern.

T. Morris[10] presented a non-contact computer system. It was controlled by head and eye for people with motor difficulties, it used spatial temporal filtering and variance maps to locate the head and find the eye feature points respectively. Tu and Thomas[11] proposed a camera mouse driven by 3D model based visual face tracking. It retrived the head orientation and translation to navigate the mouse cursor. The mouse events were utilized through the detection of the mouth movement. Kim and Ryu[12] presented a system for the hands-free control of the computer. It used the web camera to track the movement of the user's head and translated it into computer cursor movement. Iain Matthews and Simon Baker[13] proposed Active Appearance Models (AAM), which is powerful for modeling and registering deformable objects. In the recent decade, AAM has deserved further research and extensive application in the facial feature location, especially in the eye and mouth region location. It was robust to the various environment illumination and background motion. Tomàs Pallejà[14] created a relative virtual mouse based on the interpretation of head movements and face gestures. It realized motion detection through the optical flow algorithm. It attached mouse positioning accuracy close to the joysticks, but the speed was lower than a real mouse. Optical flow is sensitive to the illumination, so the accuracy will be weakened in a changing light environment. Moreover, this method was based on the assumption that the initial region of interest(ROI) was a front face, therefore, when the non-face object in front of the camera was active, misjudgment will occur. And when the face was far away from the camera, it also failed.

978-1-4244-4199-0/09/$25.00 ©2009 IEEE

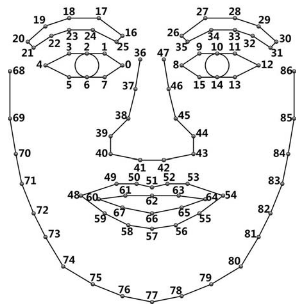

Fig. 1: AAM facial features

This paper proposes a novel local area optical flow blink detection algorithm based on Tomàs Pallejà's work. Experiments indicate a precise result.

3. FACE AND FACIAL FEATURE DETECTION

According to the AAM model in Fig. 1, eight points in the area around eyes are defined. Among them, point 0, 8 and 4, 12 are inside and outside corners of the eyes, point 2, 10 are the right and left upper eyelid center respectively. With AAM model, these feature points can be quickly detected. Moreover, local feature regions used for blink detection can be achieved. In the beginning, AdaBoost face detection algorithm is used to quickly detect the face in front of the camera, then AAM is used to localize face features. When the face moves in the XY plane, those feature points can be tracked, and the trajectory is mapped to the movement of the mouse cursor position. However, some feature point positions will change when blinking, opening mouth even in a static face image. Therefore, a special triangle is defined. It is composed by the points including the lower eyelid centers of the eyes and the center of nostrils. That is, point 6, 14 and the center of line connecting point 40 to point 43 in Fig. 1. Fig. 2 shows the movement tracking results of the triangle under the conditions of various face gestures. In order to reduce the tracking error, the gravity of the feature triangle is calculated as the reference point of the virtual mouse mapping, as marked in Fig. 2.

4. BLINK DETECTION

In the virtual mouse, the click event is trigged by double quick blink when the mouse stops i.e. the face stops moving. In [14] the optical flow is calculated in the whole image without considering the real eye position. Using AAM, precise eye features can be tracked, so the optical flow can be limited within the local AAM feature area to improve the tracking accuracy. The optical flow between two consecutive images F and $F-1$ is defined as:

$$I_{i,j}^{F}(x,y) = I^{F}(i \times m + x, j \times n + y) \tag{1}$$

Fig. 2: AAM facial features tracking

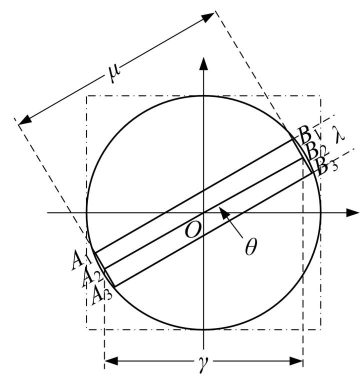

Fig. 3: Facial rotation independent eye model

where (x,y) is the relative position of the pixel in the block, (i,j) is the column and row index of the block. Considering the calculating effective, the whole image is divided in blocks of m by n pixels, dx and dy are the pixel offsets within a block. Therefore, the whole optical flow is the cumulation of these blocks:

$$S_{i,j}(d_x,d_y) = \sum_{x,y} [I_{i,j}^{F}(x,y) - I_{i,j}^{F-1}(x+d_x, y+d_y)]^2 \tag{2}$$

The moving direction of mouse can be decomposed by horizontal and vertical orientation respectively. For this reason, the optical flow can be simplified to these two axis:

$$H_{i,j}(d_x,d_y) = \sum_{x,y} \left| I_{i,j}^{F}(x,y) - I_{i,j}^{F-1}(x+d_x, y) \right| \tag{3}$$

and

$$V_{i,j}(d_x,d_y) = \sum_{x,y} \left| I_{i,j}^{F}(x,y) - I_{i,j}^{F-1}(x, y+d_y) \right| \tag{4}$$

Based on the optical flow, the orientation of the movement can be calculated. On the premise of constant illumination, the same object will have least optical flow value between consecutive frames:

$$\begin{cases} x_{\min} = \arg\min\limits_{d_x} H_{i,j}(d_x,d_y) \\ y_{\min} = \arg\min\limits_{d_y} V_{i,j}(d_x,d_y) \end{cases} \tag{5}$$

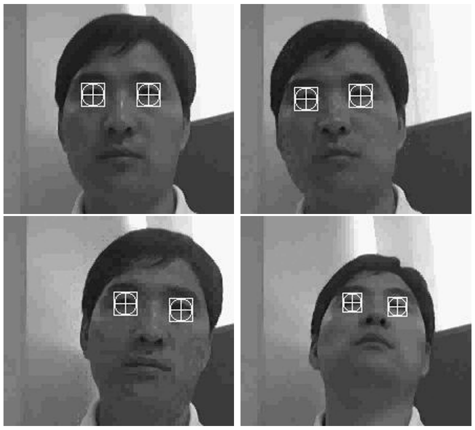

Fig. 4: Face movement independent eye tracking

Therefore, the offset of the block with the least optical flow is determined as the moving direction. Consequently, the horizontal movement right(R) and left(L) are defined by[14]:

$$\begin{cases} R_{i,j} = |x_{\min}|, & \text{if } -p \leq x_{\min} < 0 \\ L_{i,j} = x_{\min}, & \text{if } \quad 0 < x_{\min} \leq p \end{cases} \tag{6}$$

accordingly, the vertical movement up(U) and down(D) are determined by:

$$\begin{cases} U_{i,j} = |y_{\min}|, & \text{if } -q \leq y_{\min} < 0 \\ D_{i,j} = y_{\min}, & \text{if } \quad 0 < y_{\min} \leq q \end{cases} \tag{7}$$

where $x_{\min}$ is the position of the minimum value in the function H_{ij} and V_{ij}. p and q are the maximum horizontal and vertical offset of the blocks.

Usually the movement of the user's head or face is in the XY plane, and the Z coordination remains unchanged. When face rotates, the two eyes are no longer in the same horizontal line, as shown in Fig. 3. The angle of the eye corners is in the scope of -30 to 30 degree. However, the Euclidean distance μ between the inner and outer corners of the eye will remain unchanged. Therefore, the circle with diameter μ also remains the same area although the circle center varies when the face rotates. Fig. 4 shows the face movement independent eye tracking. The eye area is registered as an invariable square with side length μ. For the convenience of calculation, the optical flow is calculated in the limited square. If the user blinks while the face remains still, the optical flow of the consecutive images will give obvious local characteristics. From the binary image, the cumulative gray value of columns has two peaks while the cumulative gray value of the rows has one peak in area of the eyes. The curves are illustrated in Fig.5. The blink can be accurately determined by these peaks with proper thresholds.

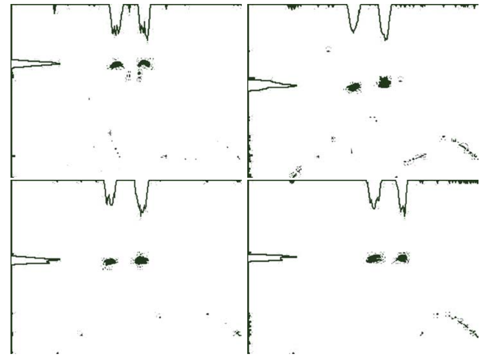

Fig. 5: Blink detection of the local eye area

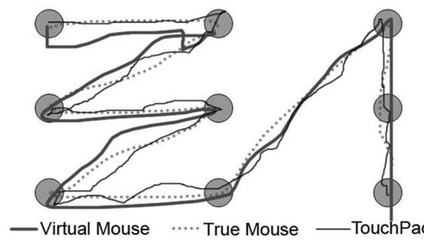

Fig. 6: Trajectory of different devices

5. EXPERIMENT AND CONCLUSION

In the experiment, a virtual mouse is designed. The validation is comparing the trajectory and speed among several input devices. They are the touch pad, the true mouse and the virtual mouse. The test computer is configured with Pentium4 2.4GHz CPU and 512Mbit memory. The screen resolution is 800×600. Fig. 6 shows the trajectories produced by these input devices. The target is moving among the 9 points through a zigzag path. The smoothest movement occurs in the true mouse while the virtual mouse performs mediately.

Fig. 7 shows the control speed contrast among these devices. Four kinds of tests are presented, they are running one circle outside the 9 points, walking along the zigzag like Fig. 6 one time, running two circles outside the 9 points, running four circles outside the 9 points respectively. The virtual mouse is slower than the other two devices because face movement is not as convenient as others. Experiments show that this method has strong robustness. It can stably track the face features and detect the blink in various conditions such as backgrounds movement and environment illumination changing. It can be used for the design of high available virtual mouse. In the next work, the other mouse event including double click, dragging and right button click will be considered.

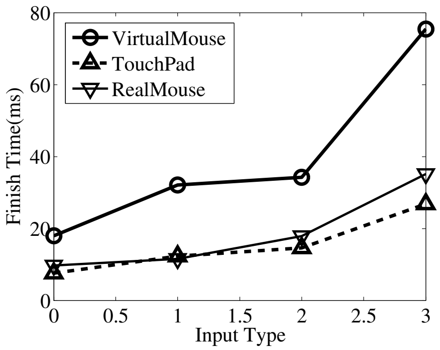

Fig. 7: Speed of different devices

6. ACKNOWLEDGEMENTS

This work has been funded by National Key Technology R&D Program in the 11th Five Year Plan of China under the program of Research on Disabled People Function Rehabilitation Aids. This program aims to foster innovation to make technology accessible for the disabled.

References

[1] Thomas E. Hutchinson, K. Preston White, JR. et al. Human-Computer Interaction Using Eye-Gaze.*IEEE Transactions on Systems, Man, and Cybernetics.*, 19(6):1527-1534,1989.

[2] Shen-Wen Shin and Jin liu. A Novel Approach to 3-DGazeTrackingUsing Stereo Cameras.*IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics.*, 34(1):234-245,2004.

[3] Dongheng Li, David Winfield, Derrick J. Parkhurst. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches.*Proc. CVPR05*, San Diego, CA, 2005, Vol.1, pp.75-82.

[4] Kang Ryoung Park. A Real-Time Gaze Position Estimation Method Based on a 3-D Eye Model.*IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics.*, 37(1):119-212,2007.

[5] Arantxa Villanueva and Rafael Cabeza. A Novel Gaze Estimation System With One Calibration Point.*IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics.*, 38(4):1123-1138,2008.

[6] M. Betke, J. Gips and P. Fleming. The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People With Severe Disabilities.*IEEE Trans. On Neuronal Systems and Rehabilitation Engineering.*, 10(1):1-10,2002.

[7] K. Grauman, M. Betke, J. Lombardi, J. Gips and G.R. Bradski. Communication via eye blinks and eyebrow raises: video-based human-computer interfaces.*Universal Access in the Information Society.*, 2:359-373,2004.

[8] P. Campadelli and R. Lanzarotti. Localization of Facial Features and Fiducial Points.*Proc. VIIP02*, Marbella, Spain, 2002, pp.491-495.

[9] Michael Chau and Margrit Betke. Real time eye tracking and blink detection with USB cameras.*Boston University Computer Science Technical Report.* no.2005-12.

[10] T. Morris, P. Blenkhom, Farham Zaidi. Blink Detection for Real-Time Eye Tracking.*Journal of Network and Computer Applications.*, 25(2):129-143,2002.

[11] Jilin Tu, Thomas Huang and Hai Tao. Face as Mouse Through Visual Face Tracking.*Proc. CRV'05.*, Victoria, BC, Canada, 9-11 May 2005, pp.339-346.

[12] HyungJun Kim and DaeHyun Ryu. Computer Control by Tracking Head Movements for the Disabled.*Proc. ICCHP'06.*, Linz, Austria, July 12-14, 2006, pp.709-715.

[13] Iain Matthews and Simon Baker. Active Appearance Models Revisited.*Journal of Computer Vision.*, 60(2):135-164,2004.

[14] Tomàs Pallejà, Edgar Rubión, Mercè Teixidó, et al. Using the Optical Flow to Implement a Relative Virtual mouse controlled by head movements.*Journal of Universal Computer Science.*, 14(19):3127-3141,2008.