FaceController: A Facial Feature Based Hands-free Interface for Computer Accessibility

Yuwei Jiao
University of Waterloo
Waterloo, On, Canada
jyuwei@uwaterloo.ca

ABSTRACT

FaceController is a face tracking based system that enables computer manipulations with simple facial expressions and enhances the computer and software accessibility, especially for the disabled. Beyond a camera on the top of a computer monitor, there is no other hardware requirement of the system. Supplemented by the power of face detection in OpenCV, FaceController provides a complete solution for building hands-free input system. Foremost, FaceController allows people to move the cursor, click on buttons, control scroll bar, and type text via virtual keyboard in the screen with facial movements. Users can also set shortcuts for commands like opening/closing file or switching between panels with expressions like smiling or sad face. FaceController has an innovative capability of sending system-wide click and keystroke signals. Thus, it is conceivable that FaceController would provide another solution for the disabled to manipulate any software, and it could become the next generation input device.

Author Keywords

face tracking, hands-free control, input device

INTRODUCTION

The mouse and keyboard have been principal computer input devices for several years. It is undeniable that they provide easy, accurate and sensible controls over computers. However, input accessibility is far from solved. Firstly, it is rather frustrating and exhausting if users need to switch between mouse and keyboard constantly for a period of time. Secondly, mouse and keyboard are not elegant solutions for direct manipulation as there remains a distance between hands and eyes. Therefore users have to handle it carefully on what they do and what they see. Thirdly, as input devices, mouse and keyboard are not accessible to the disabled. As a result, traditional input devices somehow impair the accessibility and user experience, and also prevent people from realizing the full potential of computers.

With the increase of machine power and decrease of camera cost, vision technology provides a way for a new generation input device. To be operational, vision based input interface requires to be "affordable, flexible, precise and robust"[7]. A few researchers have explored the possibility of using facial features to control computers and developed hands-free mouse replacements. However, most prototypes use the movement of eyes and nose, or markings on user's face to navigate the cursor. These solutions do not take fully advantage of the information delivered by human's facial expressions. There is still a large possible space unexplored. Furthermore, the functionality of conventional input devices is weakened, for example, the ease of use and rapid response.

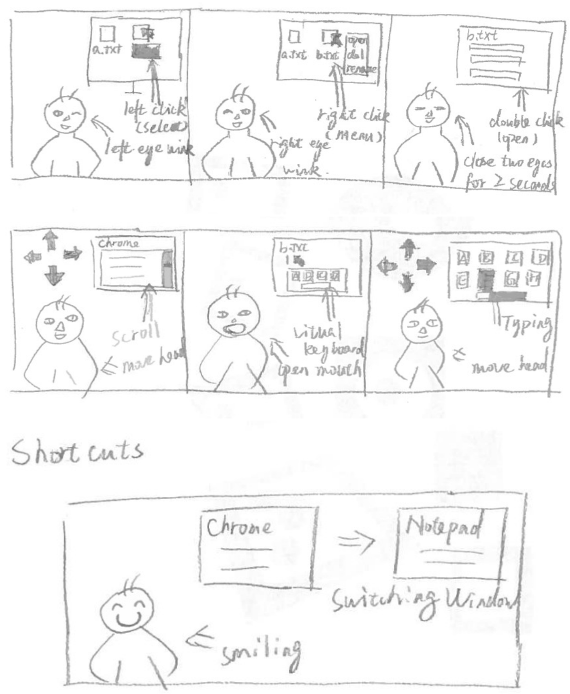

Figure 1. Sketch to illustrate FaceController

In this paper, I present FaceController, a novel facial tracking based system that enables most software manipulations with basic human expressions (see Figure 1). The system-wide implementation allows users to move the cursor, click on buttons, control scroll bar and type characters directly via various facial expressions. Each facial movement is bound with an input command in the system. Currently FaceController is able to detect when users wink, raise either eyebrow, open their mouths and move heads around. And it can also recognize combinations of facial movements, which enlarges expressing space and improves the flexibility.

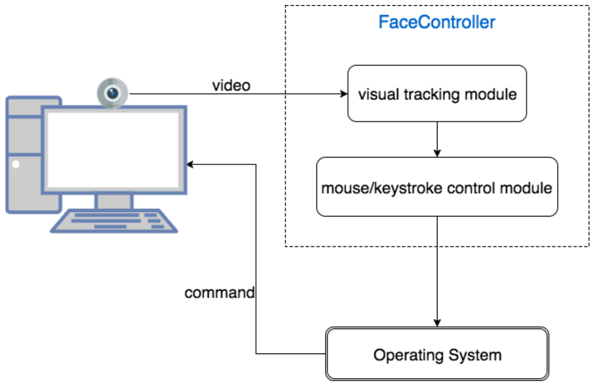

Figure 2. The framework of FaceController

Movement	
left / right eye wink	left / right button click
close two eyes (> 2s)	double click
open month	activate / deactivate virtual keyboard
move closer to screen	activate scroll mode
move head up / down / left / right	default mode: move the cursor
	scroll mode: scroll panel
	typing mode: typing alphabet
raise left / right eyebrow	switch left / right windows
Expression	
smiling face	zoom in
sad face	zoom out
angry face	return to desktop

Table 1. Facial movements and expressions commands

In my implementation, I first applied a variety of face tracking algorithms in an attempt to find the optimal method in regard to accuracy and complexity. Since the appearance of facial features differs among people, tracking performance varies among individuals. Besides, enviromental factors, for example, lighting, might affect the performance as well. Therefore it's necessary to take those factors into account. Next step was to project from 3D facial movement to 2D cursor position. FaceController provides three mouse cursor control modes: direct mode, joystick mode and differential mode. And I will explore more on hands-free typing in the future. Finally I present the discussions, as well as strengths and limitations of my prototype.

RELATED WORK

Assistive computer input systems that provide alternative methods to get controls of computers have been used and explored for a considerable period. Besides using facial movements, people have tried to take advantages of other physical signals to communicate with computers effectively, for example, Keirn and Aunon suggested to control using brain-wave [11]. Based on those ideas, people have tried to improve the recognition accuracy to communicate better [17] [19]. In the field of facial expression recognition, several effective algorithms and methods have been published and applied for years. And with the development of hardware and technology, computing efficiency has been improved greatly, which enlightens more complicated while accurate algorithms [14] [2] [15] [22].

Methods

Electrooculography

Electrooculography (EOG) is a technique for measuring the state of retina with pairs of electrodes placed near the eyes. When eyes moving from one position to another, a possible difference occurs between the two electrodes and it becomes possible to estimate the potential eyes position and movement. Kaufman et al developed an eye control system using EOG for the disabled [10] and indicated it was very inexpensive but applicable for most basic interaction devices. The headphone prototype by Manabe et al also applied EOG and used Kalman filter to analyse the signals captured [13]. However, researchers indicate that EOG method is not accurate enough because of noisy signals. People partly solved the problem using multiple pairs of electrodes. After a training session of a few minutes, they indicated users could control several window functions naturally. However, the hardware requirements built a barrier for users. Some users refused to take electrodes on face, and in some cases, the electrodes fall off, leading unpleasant user experiences.

In general, the two disadvantages of EOG method are extra hardware equipments and low manipulation accuracy in practice.

Vision-based Methods

Face detection using computer vision techniques has been explored thoroughly and applied to control mouse for years [23]. Pantic and Rothkrantz gave an ideal system for facial expression detection, extraction, classification and analysis [17], including stereo algorithm, color detector, eigenfaces and eigenfeatures. This could be used as baseline in future studies.

Stereo face tracking provides simple yet effective approach which is suitable for real-time processing [14] [19] [21] [7]. Matsumoto and Zelinsky's work included two cameras and re-constructed 3D facial feature model if any face found in both image streams [14]. Saragih et al proposed a method to improve the performance of stereo using deformable model fitting method [19]. Driven by visual face tracking and 3D model, Tu et al developed a camera mouse system using only one camera as video input but able to retrieve 3D head motion parameters [21]. They designed three mouse control modes and compared the controllability under Windows XP.

PCA [1] [19] and Haar Cascades [18] [7] are also commonly used face recognition methods.

Prototypes

Man to machine communication through brain-wave processing is an interesting research topic for a long period of time. Keirn and Aunon was interested in determine if it's possible to capture electrical activities in brain and monitor in the electroencephalography(EEG) [11]. EEG measures voltage fluctuations with two electrodes place along user's scalp. It's a typically noninvasive method. Their work presented that it possible to capture and distinguish among various mental tasks with high degree of accuracy using EEG. Therefore, translating these mental signals into a set of commands to external devices could be feasible. One of the significant extensions to this work is the real-time music notation developed by Eaton and Miranda [5].

Since last century, people start to focus on eye movements as a new method for computer input based on the observations that human direct visual attention by the movements of eyes. Colin [23] first carried out experiments to prove the feasibility of the eye tracker as a fast selection device for computers. Takami et al developed a eye-control system for the disabled in 1995 [20], which could be recognized as the first stage of smart home because users could control objects including TV, telephone and camera, etc. Chen et al shared the similar idea of interaction with computers with special designed headphone and desk equipments [3], but the system was infrared-controlled. A wearable headphone-type gaze detector was proposed by Manabe et al in 2006 [13]. With four EOG electrodes attached on the headphone near the eyes, the prototype could estimate gaze direction easily. However, such headphones are not convenient and users would feel overwhelmed to wear then continuously.

Researchers are engaged in finding more natural methods to assist computer communication. Combined head pose and gaze direction, several systems provide robust and accurate solutions to some degree [14] [1] [20] [12]. Hornof and Caveder developed the system "EyeDraw" enabling children to draw with eyes [8]. And a system based on gestures and movements helped doctors to control the position of laparoscope [16] Jota and Wigdor explored the design space of eyelid gestures systematically [9] and they also developed prototypes for both computers and smart phones. However, one problem for these systems is that, they only take the advantage of part of human facial movements (eyes). Therefore the application fields are limited and the systems are not feasible in some cases.

In recent years, more facial movements and expressions have been take into considerations. Gorodnichy and Roth proposed to use user's nose tip as mouse [7]. Camera Mouse has a developing history of more than ten years and keeps updating online [2] [6]. It's a truly powerful and wonderful program. Besides controlling and communicating through eyes, nose and head, it also allows user defined position on face to act as cursor. And users can set parameters in GUI to adjust respond speed and smoothness. More experiments and evaluations were provided by Cloud et al [4].

IMPLEMENTATION

Framework

The framework of FaceController is shown in Figure 2. No other hardware requirement is needed beyond a commodity computer with camera on the top. There are at least three components in my system: the operating system to handle commands, the visual tracking module and the mouse/keystroke control module. At present, FaceController is built on Mac OS Sierra 10.12.3 and uses the built-in webcam (720p FaceTime HD camera).

Language, Libraries, Tookits

FaceController is developed using Python 3.5.1 and takes the advantage of OpenCV (3.2.0) and Numpy (1.11.2) to handle image signals. To control the mouse and keyboard programmatically, I adopt PyAutoGui (3.2.1) in FaceController. PyAutoGui is a cross-platform Python module which can simulate moving, clicking, dragging the mouse and pressing, releasing and holding keys. It even supports pressing keyboard hotkey combinations, which enables the functions of shortcuts in my implementation. See `http://pyautogui.readthedocs.io/en/latest/index.html` for more information about PyAutoGui.

Face Detection Algorithms

Face Detection is the first step in the system (see Figure 4). There are several face detection algorithms to locate a human face in the screen. Check out `http://www.face-rec.org/algorithms/`. In my implementation, I applied the following techiniques:

- Principal Component Analysis (PCA)
- Haar Feature-based Cascade Classifiers
- ...

To be finished

DISCUSSION

To be finished

FUTURE WORK

To be finished

REFERENCES

1. Kohei Arai and Ronny Mardiyanto. 2010. Camera as mouse and keyboard for handicap person with troubleshooting ability, recovery, and complete mouse events. *International Journal of Human Computer Interaction (IJHCI)* 1, 3 (2010), 46.

2. Margrit Betke, James Gips, and Peter Fleming. 2002. The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on neural systems and Rehabilitation Engineering* 10, 1 (2002), 1–10.

3. Yu-Luen Chen, Fuk-Tan Tang, Walter H Chang, May-Keun Wong, Ying-Ying Shih, and Te-Son Kuo. 1999. The new design of an infrared-controlled human-computer interface for the disabled. *IEEE Transactions on Rehabilitation Engineering* 7, 4 (1999), 474–481.

4. RL Cloud, Margrit Betke, and James Gips. 2002. Experiments with a camera-based human-computer interface system. In *Proceedings of the 7th ERCIM Workshop" User Interfaces for All," UI4ALL*. 103–110.

5. Joel Eaton and Eduardo Miranda. 2013. Real-time notation using brainwave control. In *Sound and Music Computing Conference*.

6. James Gips, Margrit Betke, and Peter Fleming. 2000. The Camera Mouse: Preliminary investigation of automated

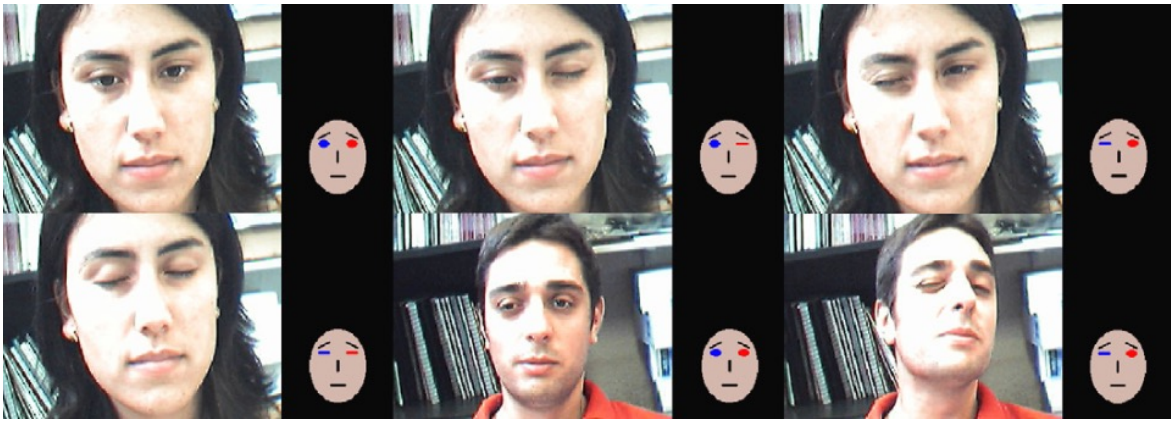

Figure 3. Wink detection [22]

visual tracking for computer access. In *In Proc. Conf. on Rehabilitation Engineering and Assistive Technology Society of North America*. 98–100.

7. Dmitry O Gorodnichy and Gerhard Roth. 2004. Nouse 'use your nose as a mouse' perceptual vision technology for hands-free games and interfaces. *Image and Vision Computing* 22, 12 (2004), 931–942.
8. Anthony J Hornof and Anna Cavender. 2005. EyeDraw: enabling children with severe motor impairments to draw with their eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 161–170.
9. Ricardo Jota and Daniel Wigdor. 2015. Palpebrae superioris: exploring the design space of eyelid gestures. In *Proceedings of the 41st Graphics Interface Conference*. Canadian Information Processing Society, 273–280.
10. Arie E Kaufman, Amit Bandopadhay, and Bernard D Shaviv. 1993. An eye tracking computer user interface. In *Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in*. IEEE, 120–121.
11. Zachary A Keirn and Jorge I Aunon. 1990. Man-machine communications through brain-wave processing. *IEEE Engineering in Medicine and biology magazine* 9, 1 (1990), 55–57.
12. John J Magee, Matthew R Scott, Benjamin N Waber, and Margrit Betke. 2004. Eyekeys: A real-time vision interface based on gaze detection from a low-grade video camera. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*. IEEE, 159–159.
13. Hiroyuki Manabe and Masaaki Fukumoto. 2006. Full-time wearable headphone-type gaze detector. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1073–1078.
14. Yoshio Matsumoto and Alexander Zelinsky. 2000. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*. IEEE, 499–504.
15. Tim Morris and V Chauhan. 2006. Facial feature tracking for cursor control. *Journal of Network and Computer Applications* 29, 1 (2006), 62–80.
16. Atsushi Nishikawa, Toshinori Hosoi, Kengo Koara, Daiji Negoro, Ayae Hikita, Shuichi Asano, Haruhiko Kakutani, Fumio Miyazaki, Mitsugu Sekimoto, Masayoshi Yasui, and others. 2003. FAce MOUSe: A novel human-machine interface for controlling the position of a laparoscope. *IEEE Transactions on Robotics and Automation* 19, 5 (2003), 825–841.
17. Maja Pantic and Leon J. M. Rothkrantz. 2000. Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on pattern analysis and machine intelligence* 22, 12 (2000), 1424–1445.
18. Kaushik Parmar, Bhavin Mehta, and Rupali Sawant. 2012. Facial-feature based Human-Computer Interface for disabled people. In *Communication, Information & Computing Technology (ICCICT), 2012 International Conference on*. IEEE, 1–5.

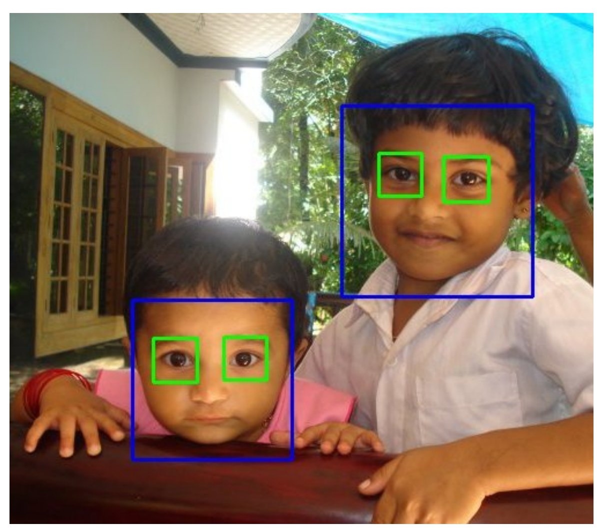

Figure 4. Face detection results using Haar Cascades in OpenCV

19. Jason M Saragih, Simon Lucey, and Jeffrey F Cohn. 2011. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision* 91, 2 (2011), 200–215.

20. Osamu Takami, Kazuaki Morimoto, Tsumoru Ochiai, and Takakazu Ishimatsu. 1995. Computer interface to use head and eyeball movement for handicapped people. In *Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on*, Vol. 2. IEEE, 1119–1123.

21. Jilin Tu, Hai Tao, and Thomas Huang. 2007. Face as mouse through visual face tracking. *Computer Vision and Image Understanding* 108, 1 (2007), 35–40.

22. Javier Varona, Cristina Manresa-Yee, and Francisco J Perales. 2008. Hands-free vision-based interface for computer accessibility. *Journal of Network and Computer Applications* 31, 4 (2008), 357–374.

23. Colin Ware and Harutune H Mikaelian. 1987. An evaluation of an eye tracker as a device for computer input2. In *ACM SIGCHI Bulletin*, Vol. 17. ACM, 183–188.