Volume 4, Issue 1, January 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Camera Based Mouse by Detecting Real-Time Face

Sonali Dohe, Tarun Yengantiwar

Department of CSE

T.G.P.C.E.T., Ngpur, India

Abstract: ***This paper introduces a very distinctive camera mouse driven visual face tracking technique based on three-dimensional model. whereas camera becomes customary configuration for personal computer(PC) and laptop computer speed becomes fast for achieving human machine interaction through visual face pursuit becomes a doable answer to hand-free management. The human facial movement is rotten into rigid movement, e.g. translation and rotation, and non-rigid movement, such as the eyes, open/close of mouth, and facial expressions, etc. We introduce our visual face pursuit system which will robustly and accurately retrieve these motion parameters from video at fundamental quantity. Once activity, the retrieved head orientation and translation is employed to navigate the mouse pointer, and additionally the detection of mouth movement is utilized to trigger mouse events. 3 mouse management modes unit of measurement investigated and compared. Experiments in Windows XP environment verifies the convenience of navigation and operations using our face mouse. this system are academic degree alternative knowledge data input device for people with hand and speech disability and for movement vision-based game and interface.***

Keywords: ***Visual face tracking, Human computer interaction(HCI), Three-D face model, Camera mouse***

I . INTRODUCTION

People have long been speculating the possibility of interacting with laptop and computer in associate natural suggests that, instead of mistreatment hand-controlled input devices, e.g. mouse and keyboard. However therefore on attempt to try this, the laptop/computer has to accommodate to human's natural sensing and activity behaviors which is usually not trivial. One potential resolution is to track frame movement in video input, from which the intention of human could also be inferred for laptop to retort to. This kind of technology is termed activity user interface [1][2].

By retrieving human motion parameters in video captured at real-time by a camera hooked to computer/laptop, mouse cursor could also be controlled by the motion parameters. This new mouse tool is termed camera mouse[3]. A camera mouse is usually composed of a visible following module and a mouse control module. Visual following module retrieves motion parameters from the video, and mouse management module specifies the rule of management. The framework are going to be illustrated by Fig1.Among frame components, face has been the most studied for visual human following and activity interface, because face look is additional statistically consistent in color, type and texture, and so allow laptop to sight and track with lustiness and accuracy. With different assumptions, of us have projected to navigate mouse with the movement of eyes[4][5][6][7], nose[8][9], and nostrils[10][11], etc. For eye following, of us generally employ infrared lighting cameras, and create the foremost of the fact that the iris of human eye has large infra-red light-weight reflection. For nose following, [8] claimed that method nose as an extremum of the 3D curvature of the nose surface makes nose the foremost robust feature for following with high accuracy. For orifice following, skin-color region is usually extracted first, and orifice could also be distinguished by its dark color and unique contour type. By following the X-Y facial feature coordinates, mouse pointer could also be navigated. However, we also notice the movement and website of facial feature in video generally doesn't coincide with people's focus of attention on screen. This therefore makes the navigation operations un-intuitive and inconvenient. therefore on avoid that problem, of us have projected to navigate mouse pointer by 3D head cause. The estimation of 3D head cause generally requires following of over one feature. Head cause can be inferred by stereo triangulation if over one camera are employed[7][12] or by abstract thought from science characteristics of face geometry[13][14]. supported the technical developments throughout this house, some industrial product have been developed in recent years[15] [16][17].

For mouse management module, the conversion from human motion parameters, i.e. position and/or rotation (orientation), etc., to mouse pointer navigation could also be categorized into direct mode, joystick mode, and differential mode. For direct mode, a matched mapping from the motion parameter domain to screen coordinates is established by activity off-line or designedly supported the a prior information concerning the human-monitor setting[13]. Joystick mode navigate mouse pointer by the direction(or the sign) of the motion parameters. and so the speed of the pointer motion is ready by the magnitude of the motion parameters[9]. Using the differential mode, the cumulation of displacement of the motion parameter displacements drives the navigation of the mouse pointer, and a number of further motion parameter switches on/off the cumulation mechanism therefore the motion parameter could also be shifted backwards whereas not influence the pointer position. therefore this mode is very similar to commonplace mouse mode as user can raise mouse and move back to the origin on mouse pad once acting a mouse dragging operation[9].

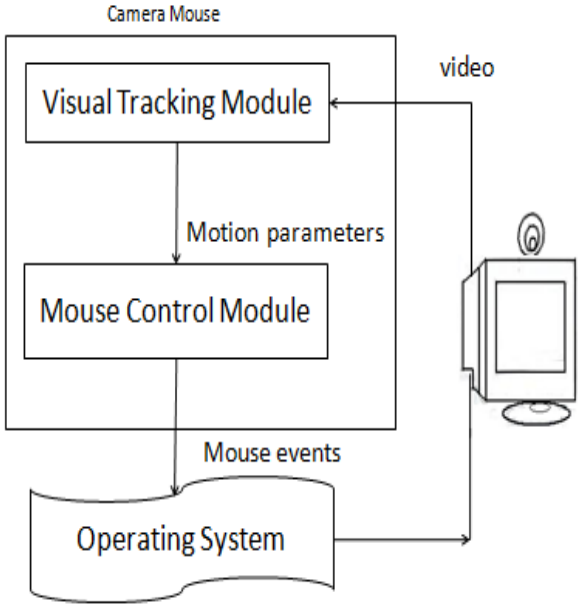

Figure 1: The framework of camera mouse

After mouse pointer is navigated to desired location, the execution of mouse operations, like electric switch clicks, is disbursed keep with additional interpretation of user's motion parameter changes. The foremost straightforward interpretation is to threshold some such motion parameters. In [3], the mouse clicks is generated supported"dwell time", e.g. a click is generated if the user keeps the mouse cursor still for zero.5s. In [6], the confirmation and cancelation of mouse operations is shipped by head cernuous and head shaking. Regular finite state machine is supposed to sight the cernuous and shaking from raw motion parameters.

In this paper, we have an inclination to projected to use 3D model-based visual face following approach [18] to retrieve facial motion parameters for mouse management. This approach utilizes just one camera as video input, but is prepared to retrieve 3D head motion parameters. it's in addition not merely a facial feature hunter, the non-rigid facial deformation is in addition developed as a linear model and may well be retrieved. Supported the motion parameters retrieved by our system, we have an inclination to designed 3 completely totally different mouse control modes. at intervals the experiments, the controllability of the 3 mouse management modes is compared. At last, we have an inclination to demonstrate how our system could also be utilized to play laptop card game in Windows XP setting. The rest of the paper is organized as follows: Section 2 summarizes our face model. Section 3 explains but our tracking system works. Section four describes the planning of our mouse management ways in which. Section 5 shows the experimental evaluation of the controllability of our system. And Section six summarizes the paper and provides some analysis of future directions.

II. FACE MODELING THREE-DIMENTIONAL

The 3D pure mathematics of human facial surface is pictured by a collection of vertices {(xi, yi, zi)|i = 1,….N} in space, wherever N is that the total range of vertices. so as to model facial articulations, a supposed Piecewise Bezier Volume Deformation Model is developed[18]. With this tool, some pre-specified 3D facial deformations is manually crafted. These crafted facial deformations ar known as Action Units(AU)[19]. For our trailing system, twelve action units ar crafted as shown in Fig. 1.

Table 1: Action Units

AU	Description
1	Vertical movement of the center of lower lip
2	Vertical movement of the center of upper lip
3	Vertical movement of left mouth corner
4	Horizontal movement of left mouth corner
5	Horizontal movement of right mouth corner
6	Vertical movement of right mouth corner
7	Vertical movement of left eyebrow
8	Vertical movement of right eyebrow
9	Lifting of left cheek
10	Lifting of right cheek
11	Blinking of left eye
12	Blinking of right eye

III. VISUAL FACE TRACKING BY CAMERA

A.Initializ the work of face tracking

The face modeling equation three defines a extremely nonlinear system. luckily visual face pursuit isn't a method of finding answer with random initial guess. A initial answer is usually supplied with high accuracy by manual labeling or by automatic detection of the face to be half-track within the first frame of the video. Our system provided associate degree automatic pursuit data format procedure. within the initial frame of the video, we have a tendency to do face detection using Ada boosting algorithm[20]. when face detection, the location of facial expression ar known by ASM techniques[21][22]. And eventually, the generic three-dimensional face model is adapted and malformed to suit to the detected two-dimensional facial expression. In the worst case, if the automated procedure goes awry, user is additionally supplied with GUI tool to fine tune the initialization result.

At the first frame of the video, the model parameter is initialized. From the second frame on, the optical flow at each vertex on face surface is computed, the displacement of the model parameters is estimated by solving the system of linear equations and the model parameters are updated accordingly. This procedure iterates for each frame, and for each frame, it also iterates in a coarse-to-fine fashion.

B. Fast computation of optical flow

One of the most long time-consuming operation of tracking procedure is that the computation of optical flow. The multidimensional language of tracking procedure optical flow by model matching through normalized correlation. Given the feel model centered at (X, Y) at frame t-1, the situation of the supreme normalized correlation of the templates in looking out window at frame t defines the optical flow at (X, Y). Denote the model centered at (X, Y) as { t(i , j)}. and therefore the model centered at (x, y) as {f(x + i , y + j)}. For reducing double decision computations and intermediate variable storage formula are used .So, for increasing the speed of computation and speed up the intermediate variables by using integral image techniques[20]

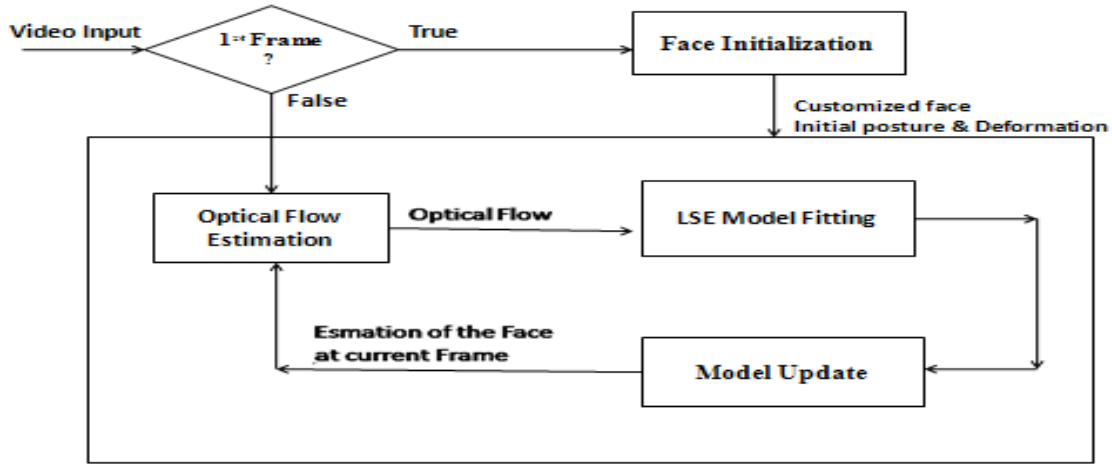

Figure 2: The flowchart of tracking procedure

IV. PROPOSED WORK

A.Performance of tracking system

The praposed system presently runs on a computer. The setup use the camera and computer system. The camera is mounted below the screen, and looks upward to the user's face. Some head motion parameters(Tx, Ty, Tz, Rx, Ry, Rz, and sequences(120 frames for concerning ten seconds) of the pinnacle cause parameters and AU coefficients associated with mouth and eyes) calculable by the huntsman area unit shown in Figure half dozen. The figure indicates the user enraptured his head horizontally (in x coordinate) and back-forth(in z coordinate), turned his head concerning y axis (yaw movement) within the 1st forty frames, and opened his mouth wide at concerning the 80-th frame, and blinked his eyes at concerning the 20-th, 40-th, and 85-th frame.

The frame rate of the huntsman achieves 22 fps when only camera mouse system rigid motion parameters area unit computed and 13fps once all motion parameters area unit computed. as a result of the optical flow is computed by normalized correlation, the huntsman is powerful to illumination changes throughout following. as a result of the LMSE resolution of the huntsman is globally optimum in nature, it will handle partial occlusions of the face. Besides, the LMSE model fitting error and therefore the confidence of the optical flow computation(derived from the total of the normalized correlation in any respect vertices) indicates whether or not the huntsman has failed or not, and might trigger automatic re-initialization of the huntsman. of these blessings build our following system a good candidate for the visual following module privately mouse.

B.Mouse cursor movement control

The direct mode, joystick mode and differential mode are implemented for the mouse control module. For the direct mode, the face orientation angle Rx, Ry (the rotation angle with respect to x and y coordinate)are mapped to the

mouse cursor coordinates (X, Y) on the screen. As the reliable tracking range of Rx and Ry is about 40 degree celciouse, and the resolution of the screen is 1600 x 1200, we therefore empirically let the mapping function to be

$$X = 40 (Ry – Roy)$$
$$Y = 30 (Rx – Rox)$$

where R0x and R0y are the initial face orientation angles. For joystick mouse control mode, the following control rule is employed.

$$Xt+1 = Xt + \Delta (Ry – R\ 0y)$$
$$Y\ t+1 = Yt + \Delta (Rx – R\ 0x)$$

The Δ (x) function is a step function in which the constants are specified empirically. We found it is easier for the user to learn to move the mouse cursor with desired speed and to keep cursor still at desired location by changing user head pose with such a step control function. For differential mouse control mode, we have the following control rules

$$Xt+1 = Xt + \alpha \Delta R\ ty\ bt$$
$$Y\ t+1 = Yt + \beta \Delta R\ tx\ bt$$

That why the mouse is navigate by using the cumulation of head orientation displacements Δ Rt x and Δ Rt y with head moving toward to the camera turns on the mouse dragging state, head moving away from the camera turns on the mouse lifting state.

The variations in nonrigid motion parameters trigger mouse button events. While there are 12 AUs for selection, not all of them are good for triggering mouse event. Ideally the detection of AU should be robust against head pose change and noisy outliers. AU7 and AU8(eyebrows raising) are not good because eyebrow movements are relative subtle for detection. AU9, AU10(cheek lifting) are not good because the lack of texture on cheek makes the estimation unreliable. And AU11 and AU12(eye blinking) are not good because user may blink his eye unintentionally and user may have difficulty to do click-and-drag operation with their eyes closed. We chose using , and the detection of mouth corner stretching to trigger right-buttonclick event, and the detection of mouth opening to trigger left-button-click event.

V. CONCLUSION

In this paper, we tend to investigated the state of the art of camera mouse techniques. specifically, we tend to projected to use a 3D model based mostly visual face huntsman to regulate the mouse 6 victimization camera mouse to play Windows game Solitaire. The user is dragging Spade-7 from right toward Heart-8 within the left by turning his head with mouth opened. and perform mouse operations. The implementation of our face trackers is careful. supported the calculable rigid and non-rigid facial motion parameters, three mouse management modes, direct mode, joystick mode and differential mode, ar enforced for the mouse management. The experiments verified the effectiveness of our camera mouse system. specifically, the accuracy of mouse navigation victimization the three mouse management modes is evaluated numerically, and also the execs and cons of each management mode is summarized. We believe associate best answer for mouse management could be a combination of direct mode and joystick mode. The direct mode will navigate the mouse indicator to the realm of interest on screen, then joystick mode is triggered to fine tune the indicator location per user's need. The two modes is switched on/off by the detection of some nonrigid facial deformation, i.e., the blinking of eye. An interesting future direction is the way to estimate human emotion supported the visual trailing of user facial motions. Some mouse events can be triggered by user emotions, and the application(such as pc tutoring software) will respond to user during a proactive manner per user's emotion (or non secular state). We additionally notice the visual face trailing system additionally play a very vital role privately mouse. Our current system is still not totally optimized in rule and in implementation. We still got to more improve our system so it can run quicker {and will|and may|and might} reach a lot of sturdy trailing performance in the context of camera mouse application.

REFERENCES

[1] Matthew Turk and George Robertson, "Perceptual user interfaces," in Communications of the ACM, 2000, vol. 43, pp. Figure 10: Using camera mouse to play Windows game Minesweeper. The user is sweeping the remaining squares at the cursor location(where the 2 mines have been marked) by simultaneous mouth opening and stretching which corresponds to clicking the left and right mouse button simultaneously. 32–34.

[2] M. Turk, C. Hu, R. Feris, F. Lashkari, and A. Beall, "Tla based face tracking," in Intern. Conf. on Vision Interfacec, Calgary, 2002, pp. 229–235.

[3] Betke M., Gips J., and Fleming P., "The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities," IEEE Trans Neural Syst Rehabil Eng., vol. 10, no. 1, pp. 1–10, 2002.

[4] Carlos Morimoto, Dave Koons, Arnon Amir, and Myron Flickner, "Real-time detection of eyes and faces," in Workshop on Perceptural User Interfacecs, San Fransisco, 1998.

[5] Robert J.K. Jacob, "What you look at is what you get: Eye movement-based interaction techniques," in Human Factors in Computing Systems, 1990, pp. 11 – 18.

[6] JamesW. Davis and Serge Vaks, "A perceptual user interface for recognizing head gesture acknowledgements," in Workshop for Perceptive User Interface, Orlando, FL, 2001, pp. 1–7.
[7] Ravikrishna Ruddarraju and et al, "Perceptual user interfaces using vision-based eye tracking," in the 5th international conference on Multimodal interfaces, Vancouver, British Columbia, Canada, 2003, pp. 227 – 233.
[8] D.O. Gorodnichy, "On importance of nose for face tracking," in 5th Intern. Conf. on Automatic Face and Gesture Recognition, Washington, DC, USA., 2002.
[9] D.O. Gorodnichy, S. Malik, and G. Roth, "Nouse 'use yournose as a mouse' - a new technology for hands-free games and interfaces," in Intern. Conf. on Vision Interface, Calgary, 2002, pp. 354–361.
[10] Layal El-Afifi, Mohamad Karaki, and Joelle Korban, "'handfree interfacee'- a fast and accurate tracking procecdure for real time human computer interaction," in FEA Student Conference, American University of Beirut, 2004.
[11] Vikash Chauhan and Tim Morris, "Face and feature tracking for cursor control," in 12th Scandinavian Conference on Image Analysis, Bergen, 2001.
[12] Trevor Darrel, Neal Checka, Alice Oh, and Louis-Philippe Morency, "Exploring vision-based interfaces: How to use your head in dual pointing tasks," Tech. Rep. AIM-2002- 001, Artificial Intelligence Laboratorty@MIT, 2002.
[13] K. Toyama, "Look, ma-no hands! hands-free cursor control with real-time 3d face tracking.," in Workshop on Perceptual User Interfaces, San Fransisco, 1998.
[14] Y.X. Hu, L.B. Chen, Y. Zhou, and H.J. Zhang, "Estimating face pose by facial asymmetry and geometry," in the 6th Intl. Conf. on Automatic Face and Gesture Recognition, 2004.
[15] Camera Mouse, "http://www.cameramouse.com/," .
[16] Mouse Vision Inc., "www.mousevision.com," .
[17] Smart Eye, "http://www.smarteye.se/," .
[18] Hai Tao and Thomas Huang, "Explanation-based facial motion tracking using a piecewise bezier volume deformation model," in Proc. IEEE Computer Vision and Patt. Recogn., 1999, vol. 1, pp. 611–617.
[19] P. Ekman and W.V. Friesen, Facial action coding system, Psychologist Press., Palo Alto, 1978.
[20] P. Viola and M. Jones, "Fast and robust classification using asymmetric adaboost and a detector cascade," in Advances in Neural Information Processing System. 2002, vol. 14, MIT Press, Cambridge, MA.
[21] T. F. Cootes and C. J. Taylor, "Active shape model search using local grey-level models: A quantitative evaluation," in 4th British Machine Vision Conference. 1993, p. 639648, BMVA Press.
[22] Jilin Tu, Zhenqiu Zhang, Zhihong Zeng, and Thomas Huang, "Face localization via hierarchical condensation with fisher boosting feature selection," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004, vol. 2, pp. 719–724.

