Intelligent Finger Movement Controlled Interface for Automotive Environment

Pradipta Biswas[1], Samuel Twist[2] and Simon Godsill[2]

[1]Centre for Product Design and Manufacturing
Indian Institute of Science
Bangalore 560012, India
Email: pradipta@cpdm.iisc.ernet.in

[2]Department of Engineering
University of Cambridge
Cambridge, UK
Email: sjg30@cam.ac.uk, sam.twist@cantab.net

ABSTRACT: This paper presents an intelligent finger tracking system to operate infotainment systems or instrument panels in automotive environment. We developed algorithms to control an on-screen pointer using off-the-shelf infra-red sensors and reported a user study on pointing and selection tasks undertaken using the finger tracking system. We proposed polynomial models to predict probable targets and a Bayesian Fusion Model to integrate eye gaze locations of users with their finger tracks. The predictive fusion model resulted in less than 2 seconds pointing and selection times on average inside a car running on a motorway.

Keywords: Finger Tracking, Eye gaze Tracking, Multimodal Fusion

1. INTRODUCTION

Recent research on intelligent dashboard design often investigates new modalities of interaction inside a vehicle such as head-up displays, gaze controlled interface [Biswas 2016], finger movement tracking systems [Ahmad 2017], haptic feedback [Chang 2011], hand gesture tracking based input [Ohn-Bar 2014] and so on. However, most research either collects data in driving simulators or analyses offline data [Ahmad 2016]. Driving simulators miss the effect of on-road vibration while analysis of offline data misses the presence of a feedback loop by an operator. This paper reports a user study in which a person used a finger tracking based dashboard, inside a car running on a highway, and a set of intelligent multimodal algorithms resulting in an average response time, for standard point and selection tasks, of less than 2 seconds. We have explored the use of an intelligent finger tracking system that can predict users' pointing targets before the users physically touch a touchscreen. If the driver need not physically touch the instrument panel it:

- can be placed even out of the reach of the driver.

- will be helpful for an elderly driver who has reduced range of motion at the shoulder due to age related physical impairments, such as arthritis.

- can reduce the pointing and selection time as targets will be selected even before users actually touch them.

We initially collected data on finger tracking systems in desktop computing and automotive environments and used the cursor trajectories to develop predictive models. Then we conducted a user study inside a car on the intelligent finger tracking system.

2. Predictive Fusion Model

2.1. Cursor Trajectory Analysis

We used data from our previous studies for developing a predictive model for the finger tracking system. Data were collected in desktop computing environment and inside a moving car, bus and train using finger and eye gaze tracking sensors [Biswas 2016]. We fit different polynomial models on variable numbers of previous data points. We considered 3, 4 and 5 previous data points and fit linear, quadratic, cubic and quartic polynomial equations, as appropriate. We then predicted the y-coordinate from a given x-coordinate after developing different polynomial models from previous data points and compared the average R^2 and error.

In the tables below, we have bold-faced the highest R^2 values and the lowest error terms. It may be noted that prediction using the last 3 points increased R^2 and reduced error for most models. The cubic and quartic models increased R^2 values over their linear and quadratic counterparts but also increased error. So we further analysed the linear and quadratic models in the following study described in section 3.

2.2. Fusion Model

Besides the predictive model, we also explored the possibility of fusing another input modality to increase the accuracy of the predictive model. We have used an eye gaze tracker to improve the accuracy of finger tracking. Eye gaze tracking is the process of measuring either the point of gaze (where one is looking) or the motion of an eye relative to the head while an eye tracker is a device for measuring eye positions and eye movement.

Table 2. Polynomial Model Fitting in Desktop Computing Environment

	Avg R²	Avg Error
5-points		
Linear	0.78	15.63
Quadratic	0.91	17.61
Cubic	0.92	19.47
Quartic	0.92	19.42
4-points		
Linear	0.83	13.65
Quadratic	0.94	15.85
Cubic	0.95	17.73
3-points		
Linear	0.89	**11.86**
Quadratic	**0.97**	14.34

Table 3. Polynomial Model Fitting in Vibrating Environment (inside moving vehicle)

	Avg R²	Avg Error
5-points		
Linear	0.64	7.69
Quadratic	0.74	8.09
Cubic	0.74	9.94
Quartic	0.76	9.90
4-points		
Linear	0.67	6.47
Quadratic	0.77	7.23
Cubic	**0.81**	15.07
3-points		
Linear	0.60	**5.44**
Quadratic	**0.81**	7.38

Our fusion technique considers eye gaze and finger tracks as two independent variables. The fusion model works at feature level [Sanderson 2002] and takes x and y coordinates as input which are calculated from raw eye gaze and finger tracking signals. The eye gaze tracker was calibrated using 9-point calibration on a 2-dimensional screen. We used a least square predictor with the finger tracker that can predict the next probable point of finger movement as described in the previous subsection. We constructed two 2-dimensional Gaussian distributions at the point of eye gaze and at the point of predicted finger track on the screen. The standard deviations of the Gaussian distributions were proportional to the published accuracy of the trackers. In this particular implementation, we considered all points on the screen are equally likely to be a probable target and so multiplied the likelihood probabilities of eye gaze

and finger tracks. The pointer is moved to the most probable point on the screen based on the Maximum Expected probability (MEP). The pointer trajectory is fed to the target prediction system, discussed in previous section, to predict a target. The following set of equations further explains the fusion strategy. In these equations P(x) stands for probability of variable x while P(y|x) stands for conditional probability of variable y given x.

P(target|eye-gaze, finger-track)

= P(eye-gaze, finger-track|target) × P(target)

= P(eye-gaze|target) × P(finger-track|target) × P(target)

Considering eye-gaze and finger track are independent to each other

∝ P(eye-gaze|target) × P(finger-track|target)

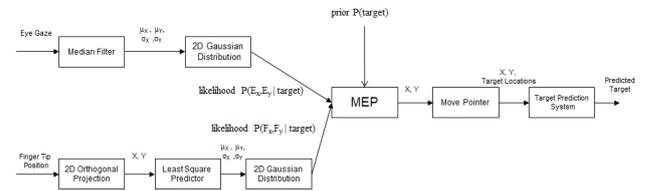Considering all targets are equally likely to be selected

Figure 2. Bayesian Fusion Strategy

3. Evaluation of the Fusion Model

We conducted the following study to evaluate the fusion model. We compared the linear and quadratic model with a naïve nearest neighbour predictor. The nearest neighbour predictor only predicts the nearest target from the latest cursor position. We have described the study in the following subsections.

Participants: We collected data from 9 participants (5 male, 4 female, average age 29.8 years).

Material: We used similar set of materials as we used in the study described in section 2.2. The car was driving along a motorway at approximately 60 MpH while participants undertook pointing and selection tasks. We used a Tobii EyeX gaze tracker with the Tobii SDK to record eye gaze locations and a Leap Motion motion tracker to record fingertip locations.

Design: The pointing and selection task was initiated through an auditory cue. It mimicked a car dashboard (figure 2) and participants were instructed to select a button on it after hearing the auditory cue. The auditory cue was set to appear between 5 and 7 second intervals. The target button was randomly selected on the dashboard. The pointing was undertaken through the intelligent finger tracking system as described in the previous section and selection was done through dwelling on target.

For the control condition, the dwell time was set to 500 msec while, for the predictive conditions, a target was automatically selected if it was predicted for seven consecutive prediction cycles.

Figure 2. Screenshot used in the study

Procedure: Initially, the participants were briefed about the study and its goal. We calibrated the eye gaze tracker while the car was not moving and demonstrated the pointing and selection tasks to the participants. They undertook the experimental task once to familiarize themselves while the car was not moving although we discarded the data from this familiarization session.

Once the car was driven to the motorway, we commenced the actual trial. Participants undertook pointing and selection tasks under the following three conditions in random order:

1. Nearest Neighbour Predictor (Control Condition)
2. Fusion Model with Linear Predictor
3. Fusion Model with Quadratic Predictor

Participants were allowed five minutes under each condition. They were instructed to take their hands out of range of the motion tracking sensor after each pointing task and start hand movement only after hearing the auditory cue. After each condition, they filled up the TLX sheet based on their average performance.

Results: We measured three dependent variables:

1. **Selection Time:** the time difference between the last auditory cue and timestamp recorded at the selection of a target.
2. **TLX scores:** which were given by participants after each individual session based on their average subjective feeling.
3. **Number of wrong selections:** which were automatically measured by the logging software.

As expected, the selection times and TLX are lower in the predicted condition while the number of wrong selection are lower in the control condition. The average selection times were less than 2 seconds in both linear and quadratic prediction models. However, none of the dependent variables show a statistically significant difference at $p < 0.05$ in repeated measure ANOVA.
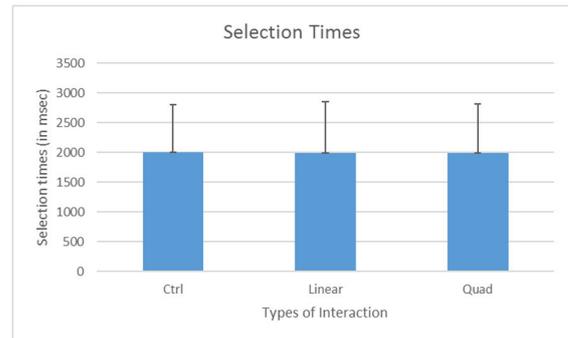

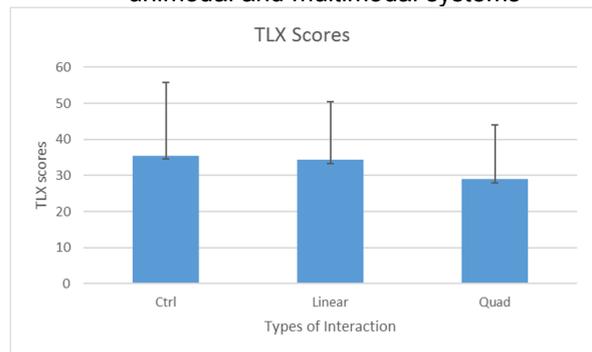**Figure 3.** Comparing pointing and selection times for unimodal and multimodal systems

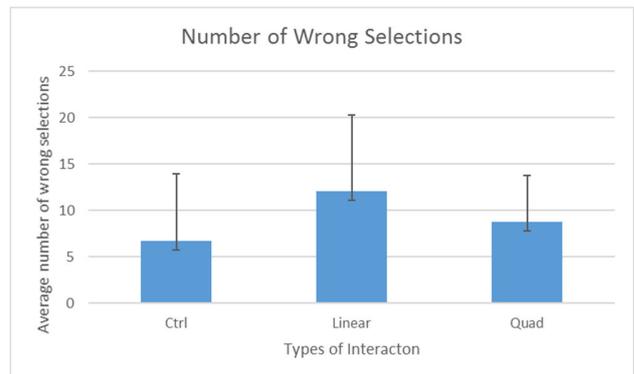**Figure 4.** Comparing cognitive load in terms of TLX scores for unimodal and multimodal systems

Figure 5. Comparing number of wrong selections for unimodal and multimodal systems

Discussion: Our study found that the fusion model reduced the average pointing and selection times below 2 seconds and the quadratic model also reduced cognitive load in terms of TLX scores from the control condition. However, we could not find statistically significant difference for the dependent variables. Due to safety issues, we could not let the driver undertake the pointing and selection tasks while driving and the study was conducted on the passengers. Although we instructed the participants to take their fingers away from the sensor after each pointing task, in effect many of them did not fully remove their hand from sensing range of the motion tracking sensor in-between the pointing tasks. As a result, the sensor could continuously track fingers and the quality of tracking was too

good to properly leverage the prediction mechanisms.

However, in a realistic situation, the hands of the driver will be engaged to the steering wheel and the sensor would only need to construct the hand model for each and every pointing task, while the driver reaches towards the dashboard (or instrument panel). Constructing the hand model for a moving hand may reduce the accuracy of sensing and result in the eye gaze tracker and predictive model becoming more useful. One way of simulating this situation would be by using a driving simulator but, in that case, we shall miss the vibration from the road and, as our study pointed out earlier, pointing and selection times were significantly affected due to vibration. Our present research is focusing on improving the accuracy of the prediction model as well as simulating a realistic driving situation to validate the model.

4. Conclusion

This paper proposed an algorithm to control an on-screen pointer using finger movement recorded by off-the-shelf infra-red trackers. Our algorithm was different from existing ones as we did not try to recognize a limited set of gestures but rather proposed an algorithm to control a graphical user interface by unconstrained finger movement. We analysed cursor trajectories of the finger tracking system and proposed a set of polynomial models to predict cursor trajectories a-priori. We have also proposed to use an eye gaze tracker, to increase the accuracy of prediction, and designed a Bayesian Fusion Model that combined signals from eye gaze and finger trackers. The fusion model can also be updated based on the history of interaction of users although, for the present analysis, we considered all pointing targets to be equally likely to be selected. Our study inside a car, involving 9 users undertaking pointing and selection tasks, reduced the target selection times and cognitive load using the Bayesian model compared to a naïve nearest neighbour predictor although the number of wrong selections increased using the fusion model. The average pointing and selection times were less than 2 seconds using the fusion model. Our future research is trying to further reduce the pointing and selection times by improving the fusion model.

REFERENCES

[Bashar 2015] Ahmad B. I., Langdon P. M. and Godsill S. J., Intelligent Intent-aware Touchscreen Systems Using Gesture Tracking with Endpoint Prediction, Proc. of the 17th International Conference on Human Computer Interaction (HCII '15) , 2015.

[Biswas 2016] Biswas P., Exploring the use of Eye Gaze Controlled Interfaces in Automotive Environments, Springer 2016, ISBN: 978-3-319-40708-1

[Chang 2011] Chang W., Hwang W. and Ji Y. G.(2011) Haptic Seat Interfaces for Driver Information and Warning Systems, International Journal of Human-Computer Interaction, 27:12, 1119-1132, DOI: 10.1080/10447318.2011.555321

[Duarte 2015] Duarte C., Costa D., Feiteira P. and Costa D., Building an Adaptive Multimodal Framework for Resource Constrained Systems, P. Biswas (Ed.): A Multimodal End-2-End Approach to Accessible Computing (2nd Edition), Springer 2015

[Feld 2013] Feld M., Meixner G., Mahr A., Seissler M. and Kalyanasundaram B.(2013), Generating a Personalized UI for the Car: A User-Adaptive Rendering Architecture, UMAP 2013, LNCS 7899, pp. 344–346

[Ford 2015] Ford Sync Voice Command System, Available at http://owner.ford.com/how-tos/sync-technology/myford-touch/get-started-with-sync/sync-voice-commands-by-category.html, Accessed on 2/11/2015

[Kern 2010] Kern D., Mahr A., Castronovo S., Schmidt A. and Müller C.(2010), Making Use of Drivers' Glances onto the Screen for Explicit Gaze-Based Interaction, Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications

[Kim 2015] Kim J., Lim J., Jo C. and Kim K. (2015), Utilization of Visual Information Perception Characteristics to Improve Classification Accuracy of Driver's Visual Search Intention for Intelligent Vehicle, International Journal of Human-Computer Interaction, DOI: 10.1080/10447318.2015.1070561

[Leap 2015] Leap Motion Controller, Available at https://www.leapmotion.com/, Accessed on 4/11/2015

[Lexus 2015] Lexus Voice Command System, Available at http://drivers.lexus.com/t3Portal/document/om-nav/OM48C25U/pdf/07.pdf, Accessed on 2/11/2015

[Normark 2015] Normark C. J. (2015): Design and Evaluation of a Touch-Based Personalizable In-vehicle User Interface, International Journal of Human-Computer Interaction, DOI: 10.1080/10447318.2015.1045240

[Ohn-Bar 2014] Ohn-Bar E. and Trivedi M.(2014), Hand Gesture Recognition in Real-Time for Automotive Interfaces: A Multimodal Vision-based Approach and Evaluations, IEEE Transactions on Intelligent Transportation Systems

[Poitschke 2011] Poitschke T., Laquai, F. , Stamboliev S. and Rigoll, G. (2011), Gaze-based interaction on multiple displays in an automotive environment, IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2011 Page(s):543 - 548, ISSN : 1062-922X

[Sanderson 2002] Sanderson C.and Paliwal K. K., Information Fusion and Person Verification Using Speech & Face Information, Research Paper IDIAP-RR 02-33 (2002)

[Schwaller 2013] Schwaller M. and Lalanne D., Pointing in the air: Measuring the effect of hand selection strategies on performance and effort, Proceedings of SouthCHI 2013.

[SUS 2014] System Usability Scale, Available at http://en.wikipedia.org/wiki/System_usability_scale, Accessed on 12/07/2014

[Tobii 2015] Tobii EyeX Eye Tracker, Available at: http://www.tobii.com/xperience/, Accessed on: 31st August, 2015