

Improving Eye Gaze Controlled Car Dashboard using Simulated Annealing

Pradipta Biswas¹, Gowdham Prabhakar¹, J. Rajesh¹, Kavesh Pandit² and Ankit Halder²

¹Centre for Product Design and Manufacturing
Indian Institute of Science
Bangalore 560012, India
{pradipta, gowdham, rajesh}@cpdm.iisc.ernet.in

²Manipal Institute of Technology
Manipal University
Manipal 576104, India
{kaveshpandit03, ankithalder1995}@gmail.com

This paper proposes and evaluates a new algorithm to improve performance of pointing and selection tasks in an eye gaze controlled graphical user interface in automotive environment. The algorithm aims to reduce number of wrong selection of targets due to inaccuracy of either the eye gaze tracker or the calibration routine. The proposed method does not require changing layout of the interface rather puts a set of hotspots on clickable targets using Simulated Annealing algorithm. A couple of user studies involving a driving simulator and ISO 26022 lane changing task found that the proposed method can reduce number of wrong selection and can also improve driving performance compared to an existing touchscreen system.

Eye Gaze Tracker, Automotive Environment, Gaze Controlled Head-Up Display, Simulated Annealing

1. INTRODUCTION

Modern infotainment systems in automobiles facilitate driving at the cost of adding secondary tasks in addition to the primary task of driving. These secondary tasks have considerable chance to distract a driver from his primary driving task, thereby reducing safety or increasing cognitive workload. Thus easing out human machine interaction (HMI) between drivers and infotainment systems can potentially raise safety and help to leverage the true potential of those systems.

Existing infotainment systems in modern automotive use either electronic screens or head up displays as output modalities while input is taken through touchscreen, physical buttons and voice recognition technologies. To facilitate HMI in automotive environment, researchers have already explored use of haptic feedback [Chang 2011], hand gesture tracking based input [Ohn-Bar 2014], personalizing instrument displays and predictive models to help drivers in regular driving tasks like parking [Feld 2013; Normark 2015]. Majority of existing cars use touchscreen for operating the dashboard. Operating a touchscreen involves visual search for intended target on the screen and then stretching arm to physically touch the screen with a finger. This paper investigates if we can use the eye gaze movement during visual search to automatically

activate intended target on the touchscreen. An eye gaze controlled instrument panel

- can be placed even out of the reach of the driver
- will reduce hand movement to physically touch the dashboard which may reduce response times and
- will be helpful for elderly driver who has reduced range of motion at shoulder due to age related physical impairment like arthritis.

Kern [2010] and Poitschke [2011] already reported user studies involving simulated driving tasks while comparing eye gaze controlled interface with traditional touchscreen system. In this paper, we have proposed and then evaluated an algorithm to improve driving performance and reduce errors in pointing and selection times for eye gaze controlled displays in automotive environment.

2. IMPROVING EYE GAZE CONTROLLED INTERFACE

Even though state-of-the-art eye gaze trackers have 0.4° visual angle of accuracy but while people operate a car or even a driving simulator they change their posture and head position reducing the accuracy of the eye gaze tracker. As a result, even though operators fixate attention to a button, the system often activates a neighbouring button requiring users to adjust

the offset. The offset is random and different for different people in different situations and difficult to be fixed by simple translation operation. In the following section, we have presented an algorithm to increase the accuracy of the gaze controlled interface in gaze controlled interface.

In user studies involving gaze controlled interface, we noted that as users stare at the middle of the target, due to the inaccuracy of the tracker or their head movement, the neighbouring button was occasionally selected. This probability of wrong selection increases if the buttons are closely spaced in the interface. Hence the probability of wrong selection will be reduced if we can increase the inter-button spacing. However, we cannot change the design of an interface just to make a particular interaction technique working better.

We have explored the option of introducing hotspots inside each button to facilitate eye gaze tracking interaction. If we can introduce a hotspot on each button and keep them well separated, we can instruct users such that the first saccade on a button would launch on these hotspots. We hypothesize that keeping these hotspots well-separated may reduce chances of wrong selection.

To find the best position of hotspots, we have represented an interface as a graph where each node corresponds to a target button (clickable objects) and neighbouring buttons are connected by an edge. For example, the interface in figure 1 corresponds to the following graph.

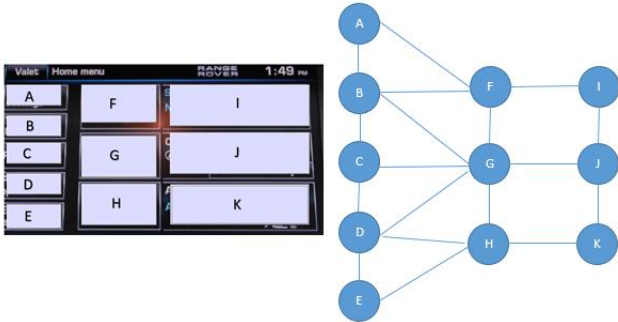


Figure 1. Graph structure of an interface

We assumed each button has a hotspot on it, which is initially located at the centre of the button. The weight of each edge is equal to the Euclidian distance between the hotspots of two neighbouring buttons. We explored two different algorithms to increase the distances between hotspots. We have defined the following cost function and tried to minimize it.

$$\text{Cost Function} = \sum_{i,j} d_{ij} \frac{1}{d_{ij}^2}, \text{ where } d_{ij} \text{ is the distance of}$$

the hotspots between buttons i and j and is equal to the weight of the edge between nodes i and j .

We have modelled the problem of finding optimum locations of hotspots as a state space search prob-

lem. Each state corresponds to a particular organization of hotspots. A state transition occurs when any hotspot changes its position. If we consider each button has k possible positions and if an interface has n buttons, then an exhaustive search algorithm need to evaluate k^n states. Even for a moderately complex interface an exhaustive search algorithm will be computationally intensive or almost impossible. Hence we used the following two algorithms.

2.1 Greedy Algorithm

This algorithm picks up the edge with minimum weight, which means the two most closely spaced buttons. It checks the degrees of the two nodes of the minimum-weight edge and updates the hotspot of the node with higher degree. The algorithm calculates the centroid of the hotspots of neighbouring nodes of the selected node and the new hotspot is calculated as the nearest point on the selected button (or node) to the centroid. While selecting the next node for updating hotspot, the algorithm checks whether the node is visited earlier and if so, it selects a different node. The algorithm is greedy in the sense that it only updates the hotspot if the overall value of the cost function is reduced from the previous value.

2.2 Simulated Annealing

This algorithm randomly selects a node and also randomly selects a point on the node as its new hotspot. If the new hotspot reduces the value of the cost function, then it is selected and updated. However even if the new hotspot increases the value of the cost function, it may still be selected based on the following condition

$$e^{\frac{\text{oldCostFn} - \text{newCostFn}}{T}} > \text{A random number between 0 and 1.}$$

In the above equation, the value of T runs from 5000 to 1 and reduced by 1 in each iteration.

Both algorithms were run for 5000 iterations and figure 2 below plots the values of the cost function for the minimum value of the cost function obtained in the algorithms. In this figure, 'Original' stands for the initial state of the algorithm when all the hotspots were at the middle of each button.

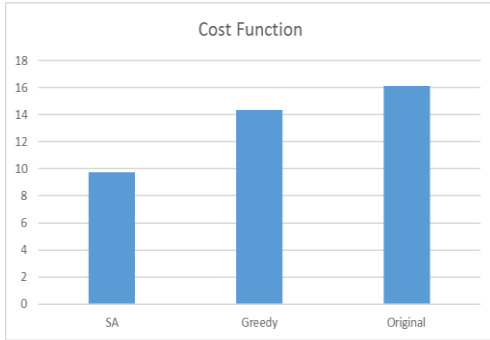


Figure 2. Comparing the cost function for different algorithms

We also plotted the sum the of the weight of edges of the graph for the minimum value of the cost function for each algorithm in figure 3 below.

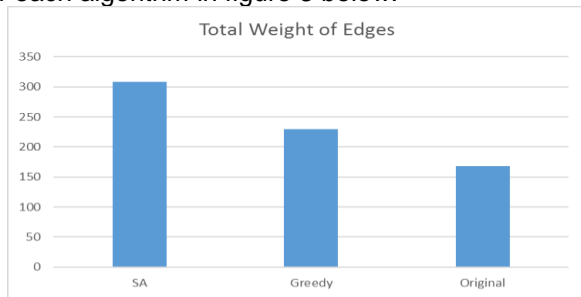


Figure 3. Comparing total distances among hotspots for different algorithms

It may be noted that both algorithms reduced the cost function and increased the overall weight of edges. The Simulated Annealing (SA) algorithm reduced the cost function further than the greedy algorithm. The greedy algorithm was stuck in cycle and a local optimum after visiting all nodes a couple of times. The SA algorithm never stuck in local optimum due to randomly choosing node and hotspots although we also could not conclude whether it reached the global optimum. It may be noted that the weights of edge in the final state of SA are significantly lower ($p < 0.01$) from the initial state in a paired t -test. Figure 4 below shows final positions of hotspots on the buttons as red dots.

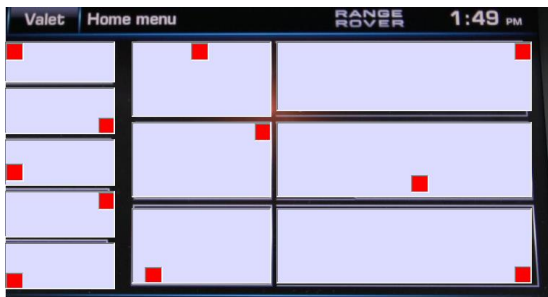


Figure 4. Final positions of hotspots after Simulated Annealing

3. USER STUDY TO EVALUATE HOT SPOTS

We hypothesized that introducing hotspots will reduce the number of wrong selection and will also reduce pointing and selection times. In the following paragraphs we presented a user study to compare a touchscreen system and eye gaze controlled systems with and without hotspots.

3.1 Participants

We collected data from 9 participants (7 male, 2 female, average age 29.4 years). All participants used to drive cars and had no problems to use the experimental set up.

3.2 Material

We used a Tobii EyeX eye gaze tracker and Lenovo Yoga Laptop for the secondary task. A Logitech G4 driving wheel and associated pedals were used for the primary driving task. We used the ISO 26022 lane changing task to measure participants driving performance and it was run on a 40" MicroMax TV screen. Selection was done by pressing a switch on steering wheel as a previous study [Biswas 2015] found hardware switch based selection is faster compared to dwell time or voice based selection for gaze controlled interface.

3.3 Design

In this dual task study, participants undertook ISO 26022 lane changing task as their primary task. In this task participants needed to drive along a 3-lane motorway. At a regular interval, drivers were shown a sign and instructed to change lane. The destination lane was randomly chosen and the driving path is automatically logged.

The secondary task involved pointing and clicking on a screen. We used the following three modalities for pointing

1. Eye gaze tracking with hot spots
2. Eye gaze tracking without hot spots
3. Touching

We have used an existing dashboard display from Jaguar Land Rover (figure 4). We used the same dimensions of the buttons as the original display but removed all button captions. During the study, one of the buttons of the display was randomly selected as target and a caption 'Target' appeared on that button. The primary and secondary tasks were linked through an auditory cue. While driving, participants were instructed to point and select the designated target on the secondary screen after hearing an auditory cue.

The auditory cue was set to appear between 5 and 7 seconds interval. The auditory cue was kept on repeating during the whole duration of driving.

3.4 Procedure

Initially we explained the aim of the study to participants. They were first trained with the gaze controlled interface and allowed to use the driving simulator to undertake a test run. After training, they undertook trials in different conditions in random order. Participants were instructed to drive realistically without veering off from the driving lane. While using the gaze controlled interface with hot spots, participants were instructed to focus on the hot spot of the designated target while no such instructions was given in other two conditions. After each condition, participants filled up the NASA TLX and SUS questionnaire.

3.5 Results

We have measured the following dependent variables:

1. **Driving performance** is measured as
 - a. **Mean deviation** from designated lane calculated according to Annex E of ISO 26022 standard.
 - b. **Average speed** of driving, in particular we investigated if the new modality significantly affected driving speed.
 - c. **Standard Deviation of Steering Angle**, a large standard deviation means drivers made sharp turns for changing lanes.
2. **Pointing and Clicking performance** is measured as
 - a. **Error in secondary task** as the number of wrong buttons selected compared to total number of selections.
 - b. **Response time** as the time difference between the auditory cue and the time instant of the selection of the target button. This time duration adds up time to react to auditory cue, switch from primary to secondary task and the pointing and selection time in the secondary task.
3. **Cognitive load** measured as the NASA Task Load Index (TLX) score.
4. **Subjective preference** as measured as the System Usability Score (SUS).

For each dependent variable, we initially compared the descriptive statistics and then undertook parametric or non-parametric statistical hypothesis test. If an ANOVA or H-test found a significant difference, we also used t-test or U-test for pairwise comparisons. Before undertaking the trial in dual task condition, participants used the driving simulator for undertaking the primary driving task only. The driving path of this sin-

gle task situation was used as a baseline for evaluating degradation in driving performance for the secondary task. Following the description of ISO 26022 Annex E, a reference path trajectory with a constant lane change start position and lane change length was calculated, which has a maximum correlation with the base line driving path. For example, in figure 5 below the green line showing the reference path while the red line is the driving path in dual task situation.

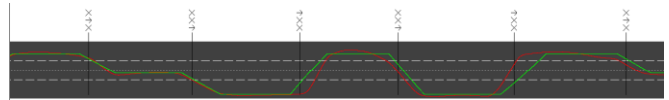


Figure 5. Illustrating calculation of mean deviation from lane to measure driving performance

We compared the arithmetic mean of the deviation from the reference path as a metric of driving performance. It may be noted from figure 6 below, the mean deviation was lowest for touchscreen based selection while highest for eye gaze tracking based selection. Using hotspots resulted in 22% reduction of deviation from reference path than without using hotspot in eye gaze controlled system.

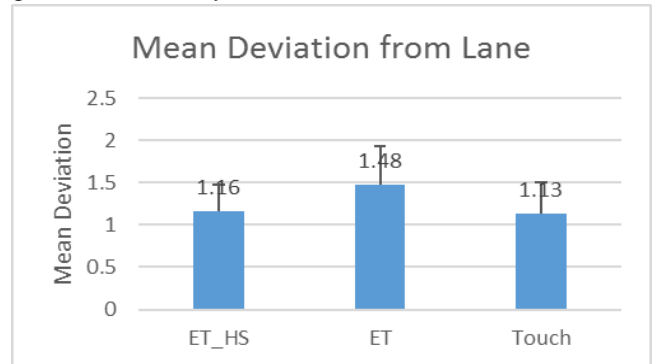


Figure 6. Comparing Driving Performance in terms of mean deviation

The average speed of driving was highest for eye gaze controlled interface while lowest for hotspot equipped gaze controlled system. However, the difference between highest to lowest average speed is less than 4km/hr and the differences were not statistically significant.

The standard deviation of steering angle was highest for touchscreen system and lowest for hotspot equipped gaze controlled system. It may indicate drivers drove slower and more cautiously for hotspot equipped gaze controlled system compared to other systems.

We hypothesized that using hotspot will reduce probability of wrong selection in gaze controlled interface and indeed the average number of wrong selection was reduced by 57% in gaze controlled interface. The average number of wrong selection is lowest and

equal for both touchscreen and hotspot equipped gaze controlled interface. Average response time was lowest for touchscreen system and highest for eye gaze controlled system. Hotspots decreased the average response times marginally (3%) than without using it in gaze controlled interface.

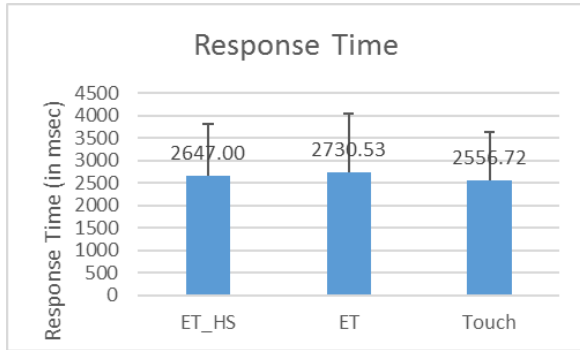


Figure 7. Comparing average response times

The TLX scores were highest for hotspot equipped gaze controlled system and lowest for touchscreen. An one-way ANOVA among the different components of TLX scores for hotspot equipped gaze controlled system found a statistically significant difference [$F(5,48) = 3.15, p < 0.05$] with highest score for temporal demand (mean 55.44, stdev 11.93) followed by mental demand (mean 45.56, stdev 20.29) and effort (mean 42.67, stdev 18.38).

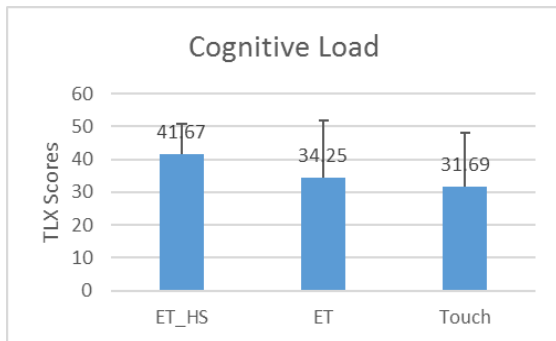


Figure 8. Comparing cognitive load in terms of average TLX scores

The SUS scores were greater than 68 for all cases and highest for hotspot equipped gaze controlled system, which means all systems were usable with hotspot equipped gaze controlled interfaces was most preferred by users. We did not find any significant differences among the different conditions for any dependent variable.

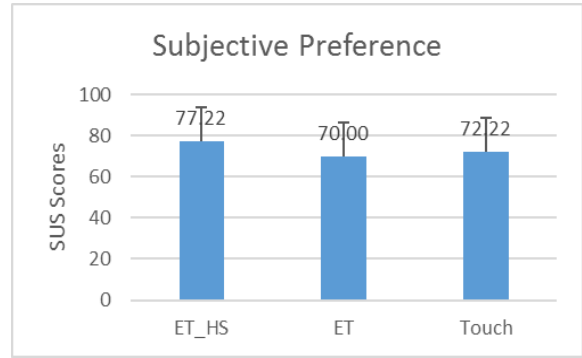


Figure 9. Comparing preference in terms of average SUS scores

3.6 Discussion

This study demonstrated that hotspots reduced the number of wrong selections by more than 50% and also marginally reduced the response time for eye gaze controlled interface. Users found the system easy to use and preferred to use the system frequently although at an expense of high cognitive load, in particular the temporal demand. Touchscreen system was still fastest to use and the driving performance was also best for touchscreen system than eye gaze controlled systems even though hotspots improved driving performance in gaze controlled interface.

One problem with gaze controlled interface is it still needs drivers to take their eyes off from road even the average duration of response times was 2.6 seconds. The next section presents a modification of the gaze controlled system that does not require drivers to take their eyes off road but even then allow to operate the dashboard using eye gaze.

4. PROJECTED GAZE CONTROLLED INTERFACE

In existing automobiles, drivers need to take their eyes off the road and physically touch the dashboard to operate the infotainment system. The eye gaze or finger movement controlled based system do not need drivers to physically touch the interface but still needs to take eyes off from road ahead. Nowadays drivers often put their iPad or smartphone on top of the dashboard to avoid looking down but the device itself occlude part of the windscreen (figure 10). The voice output system is not often useful for different language speakers and in noisy environment. In this section, we have proposed a system that projects the dashboard on a semi-transparent screen in front of the windscreen, so that even when drivers interacting with it they are not completely losing view of the road ahead. The following subsections present another study to

evaluate the utility of the hotspot for a projected gaze controlled interface.



Figure 10. Drivers often attach instruments on windscreen to avoid looking down

Initially we developed a projected display to use it with the driving simulator. The setup consists of the following three components, viz., Screen, Tilting mechanism and Frame.

The Screen consisted of layers of semi-transparent Polyethylene sheets, which were fixed to a rigid plastic frame on all four sides. The thickness of the plastic frame was more than the plastic screen (~8mm). In order to keep it upright, the screen was fixed into two slots present on the support structure made in the wooden frame.

A tilting mechanism was made to support the eye tracker and also to adjust the angle of the eye tracker. It consists of a magnetic holder, pivot, base and threaded block. The base component held the other components in place. The magnetic holder was attached to the base component with the pivot. The threaded block was attached to the base. The screw in the threaded block allowed to vary the angle of magnetic holder with respect to the base. The eye tracker was magnetically fastened to the magnetic holder.

The support frame held the screen and the tilting mechanism in place. The frame for automotive setup consisted two wooden blocks held in position with two sheet metal support held together by 4 numbers of M5 Screws. The setup was placed above the dashboard of the automobile. The screen is mounted in the slots of the frame. The eye tracker along with the tilting mechanism was placed on the sheet metal support. The overall structure is shown below in figure 11.

The following section describes a used study to measure both objective performance and subjective preference of users with the projected display.

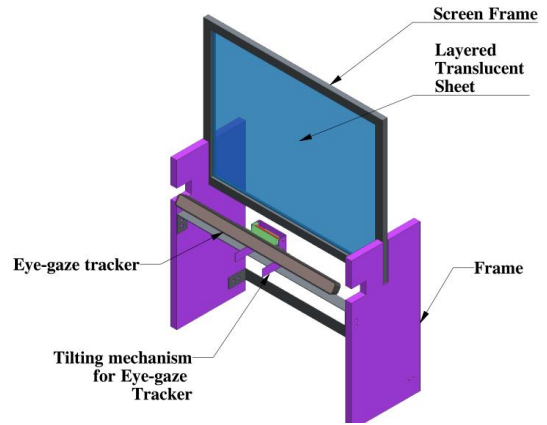


Figure 11. The projected screen set up

4.1 Participants

We collected data from 11 participants (8 male, 3 female, average age 29.2 years). All participants used to drive cars and had no problems to use the experimental set up.

4.2 Material, Design, Procedure

We used a Phillips PicoPix 56 Lumen projector to display the secondary task on a semi-transparent display (figure 12). The rest of the material, design and procedure was same as previous study. Seven out of 11 participants took part in the previous trial about 2 months ago but did not use the gaze controlled system afterwards, So, we assumed all participants were novice to gaze controlled interface.

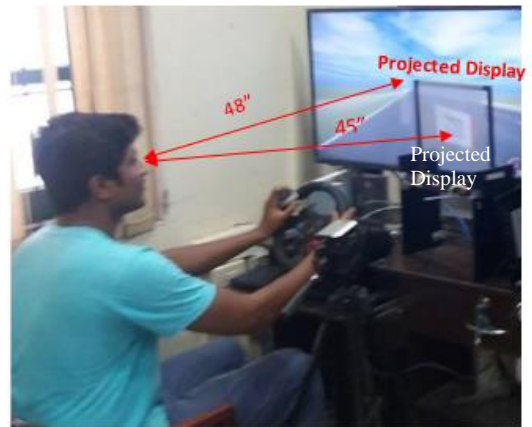


Figure 12. Set up of gaze controlled projects display

4.3 Results

We used similar analysis strategy as the previous section. The driving performance in terms of mean deviation from designated lane was significantly different in

a Kruskal-Wallis H-test [$\chi^2(2,28) = 10.56, p < 0.05$]. Pairwise signed rank test also found that driving performance was significantly different between projected gaze controlled systems for using hotspots. It may be noted that using hotspots, mean deviation from designated driving lane was reduced by 41% for projected gaze controlled interface and the mean deviation for hotspot equipped projected gaze controlled interface was even lower than the touchscreen based system.

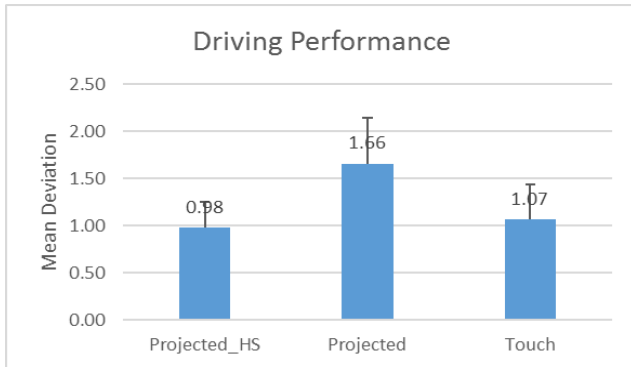


Figure 13. Comparing Driving Performance in terms of mean deviation

The differences among average speeds of driving were not significant. It was approximately 3kn/hr slower for hotspot equipped projected gaze controlled system than touchscreen system.

The standard deviation of steering angle was significantly lower for projected gaze controlled system than the touchscreen system in a signed rank test ($p < 0.05$). It may indicate drivers needed to drive more carefully with the projected gaze controlled system compared to touchscreen system.

The average number of wrong selections was higher in projected gaze controlled interface than touchscreen system although using hotspots reduced it by 15.6% than without it.

Average response time was still lowest for touchscreen based system but it is only 2% higher in hotspot equipped projected screen. A single way ANOVA found significant difference among the reaction times [$F(2,257) = 4.84, p < 0.05$]. A set of unequal variance t -tests found that touchscreen had significantly lower response times [$p < 0.05$] than projected screen without hotspot while the difference between touchscreen and hotspot equipped projected screen was not significant.

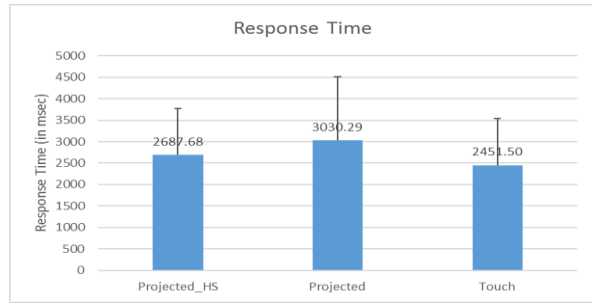


Figure 14. Comparing average response times

The TLX scores were highest for projected gaze controlled system and lowest for touchscreen. The hotspots reduced the average cognitive load by approximately 6% from the projected gaze controlled system without hotspots. However, we did not find any significant difference among different components of TLX scores for hotspot equipped gaze controlled projected system although the temporal demand (mean 52.89, stdev 19.38) was still higher than other components.

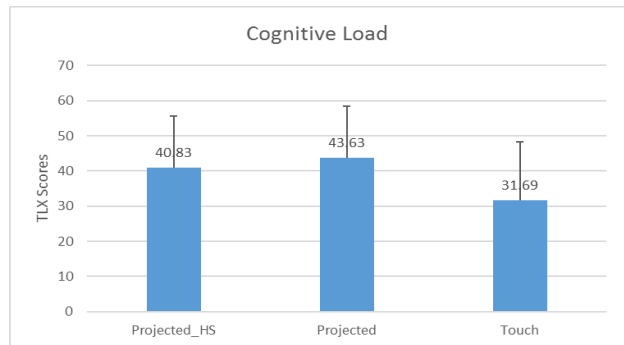


Figure 15. Comparing cognitive load in terms of average TLX scores

The SUS scores were greater than 68 for all cases and highest for hotspot equipped gaze controlled system, which means all systems were usable with hotspot equipped gaze controlled interfaces was most preferred by users.

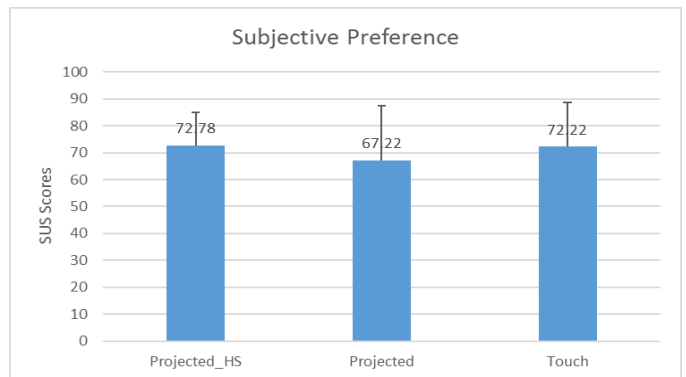


Figure 16. Comparing preference in terms of average SUS scores

4.4 Discussion

This study demonstrates that eye gaze tracking devices can be used for projected displays as well. While participants used the projected display, they need not to look down, stretch hand and could see through the road behind the semi-transparent display. We indeed found that driving performance was better than touchscreen for the hotspot equipped projected gaze control system. The average response time was only 2% higher than touchscreen system. The SUS scores also indicated users did not face any serious trouble in using this system. The cognitive load was still higher than touchscreens but it should also be noted that our participants use touchscreen enabled device everyday while they were using the eye gaze controlled interface for first time during the trials.

The main problem with the projected display was the number of wrong selections. Using hotspots significantly reduced it but it was still higher than both touchscreen and screen based gaze controlled system. Actually, we could not completely eliminate the skewness of the projected display and positioning the projector without hampering driving equipment was also a challenge, which all resulted in poor calibration of the system. Our future work is investigating appropriate ways of positioning the projector and eye gaze tracker.

Finally, these gaze tracking studies were conducted in a driving simulator where we missed out the effect of on-road vibration. Our future work will investigate effect of vibration in detail on gaze controlled system.

5. LITERATURE SURVEY AND GENERAL DISCUSSION

5.1 Alternative Input Modalities

A huge portion of research in HCI has been devoted to investigate and compare input and output modalities of interaction. Perhaps the most famous one of all is Stuart Card's initiative at Xerox Parc of using Fitts' Law [Fitts 1954] to compare mouse, joystick and trackball, which triggered a plethora of research during the last thirty years on inventing and comparing new modalities of interaction. Most studies on inventing and comparing new modalities of interaction conduct a pointing and clicking or dragging task in a screen and report the Index of Performance of the new input device according to Fitts' Law. For example, Mackenzie, Sellen and Buxton [1991] compared Mouse, Stylus and Trackball while Jagacinski and Monk [1985] compared Helmet-mounted sight and

joystick. In recent time, Vertegaal [2008] conducted similar studies with mouse, stylus and eye-gaze tracker. Ware and colleagues [1987] and Miniotas [2001] also conducted similar Fitts' law analysis for eye gaze tracker. Though these studies found mouse or stylus are faster than trackball, head or eye gaze tracker, but still there are certain cases of physical and situational impairment those require the use of these devices. Additionally, with advent of computer vision and infrared trackers, eye gaze and hand or finger movement trackers are turning cheaper and portable finding their use in everyday products like tablets and smart-phones.

5.2 Gaze Controlled Interface

Research on eye gaze tracking started even in late 18th century although controlling an electronic interface using eye gaze or finger movement is relatively new concept. For eye gaze controlled interface, Zhai [1991] pioneered in MAGIC system, which aimed to use eye gaze track directly to improve pointing, in particular homing on a target using mouse. Ashdown and colleagues [2005] addressed the issue of head movement while tracking eye-gaze in a multiple monitor scenario. They used head tracking to switch pointer across screens, which was preferred by participants, but in effect increased pointing time. Dostal and colleagues [2013] addressed similar issues by detecting which monitor the user is looking at through analyzing webcam video. The Sideways system [Zhang 2013] even eliminates personalized calibration and can scroll contents of a display screen by detecting eye gaze. The system identifies whether users are looking at the middle or sides of a display and if they are looking to the sides, the system scrolls content at the middle. Both Dostal's system and Sideways system does not use precise x and y coordinates to a move a mouse pointer. Fu and Huang [2013] proposed an input system hMouse, which moves a pointer based on head movement. They detected head movement by analyzing video input and their system is found to outperform another similar system called CameraMouse [2013]. Fejtova's [2009] Magic Key system also uses a webcam like CameraMouse but the pointer is moved in screen based on position of nose (nosetrills to be precise). Selection is done by eye blinks. Bates [1999] multimodal eye tracking system allows zooming portion of screen using a polhemus tracker. Zandera and colleagues [2010] combine a BCI system with eye-gaze tracking, where EEG generated by imagining a rinsing action is trained to make a selection. However, their system had limited success in reducing pointing times. Penkar [2012] and colleagues investigated different dwell time durations for selecting target for an eye gaze con-

trolled interface, although their study involved only selecting target at the middle of the screen. Later they extended the study [Lutteroth 2015] for selecting hyperlinks in a webpage but dwell time based selection would be inappropriate in automotive environment as it requires operators staring away from road for selecting target. Pfeuffer and Voelker explored fusing touchscreen interaction with gaze controlled system by using eye gaze for object selection and touch interaction for object manipulation. Pfeuffer [2016] explored desktop computing tasks like image searching and map navigation while Voelker [2015] investigated multi-screen display, which is more advanced in terms of coordinate mapping between horizontal and vertical displays compared to Dostal's [2013] system. However, our proposed work uses eye gaze for not only object selection but also for cursor movement.

5.3 Automotive Environment

It may be noted that, majority of studies comparing input modalities considered desktop computing environment. Although Grandt [2013] compared trackball, touchscreen, speech input and mouse for an Anti-Warfare system but the experimental set up was similar to desktop computing environment with pointing and selection be the only task to be undertaken. The automotive environment is different in two distinct ways

- The pointing and selection task is a secondary task while driving is the primary task.
- The whole environment of the task is vibrating, which may affect both the human performance as well as accuracy of the sensors used for pointing

Most cars nowadays use touchscreen system. If we investigate the intermediate steps involved in operating a touch screen following the concept of Model Human Processor and GOMS modelling [Card 1983] while driving, we may find the steps as

- (a) Context switching from primary driving to secondary task
- (b) Visually search for intended target on display
- (c) Mental preparation to select intended target
- (d) Stretching arm to select intended target
- (e) Touching the display at appropriate point
- (f) Context switching back to driving

Our research explores whether we can use step 2 (visual search) to select intended target directly. Farrell [2005] noted that "*humans use their eyes naturally as perceptive, not manipulative, body parts. Eye movement is often outside conscious thought, and it can be stressful to carefully guide eye movement as*

required to accurately use these target selection systems". However, automotive user interfaces do not require continuous manipulation of an on-screen pointer like graphical user interfaces in desktop computing. Unless a particular interaction is very familiar to the driver (like reaching for the gearbox while driving), he has to glance at the user interface [Duarte 2014]. An accurate gaze tracking with target prediction technology can leverage this glance for pointing. Researchers attempted to eliminate or reduce visual search using gesture recognition techniques but the systems either require to remember a set of gestures (AirGesture System [May 2014]) or relative positions of screen items (BullsEye system [Weinberg 2012]). Additionally, such systems worked inferior to a touchscreen system in terms of driving performance or secondary task.

Gaze controlled interfaces were already patented for automotive environment. For example, Mondragon and colleagues [2013] presented an eye gaze controlled smart display for passengers of vehicles. Users may point and select icons on the display by staring at appropriate portion of the screen. However, our proposed system is intended to be used by drivers and selection of target has to be faster than dwelling or staring away from road. Seder and colleagues [2012] presented a graphic projection display for drivers showing objects in road ahead. Users can select object on the projection display using different input modalities including eye gaze. However, the patent does not address a method of improving accuracy of the gaze tracking itself and it does not intend to operate the dashboard as the proposed system. Poitschke's [2011] study compared gaze controlled dashboard with touchscreen and reported higher reaction time for gaze controlled dashboard. They did not propose any intelligent algorithm to reduce pointing and selection times for gaze controlled interface.

Contribution of our studies: Our algorithm involving hotspots tried to reduce number of wrong selection due to inaccuracy in eye gaze tracking. It leverages the pop-out effect of visual search by placing a set of well separated spots on target buttons of a screen. We explored state space search algorithms to find the best positions of these hot spots on the interface. Our algorithm defined an energy parameter that is inversely proportional to distance between neighbouring hot spots and tried to minimize overall energy by increasing separation among hotspots. The Metropolis Keyboard [Zhai 2000] used similar concept of minimizing energy in optimizing positions of buttons in a virtual keyboard. Our first trial confirmed that hotspots indeed reduce the number of wrong selections in a gaze controlled interface. We also tried to project the content of dashboard on the windscreen so that drivers need not to look down and reported eye gaze can also be used

to control a projected display. Using hotspots with the projected display improved driving performance even compared to touchscreen.

Overall we have found that with projected gaze tracking system, drivers can undertake pointing and selection tasks at 2.6 seconds on average without hampering driving performance. Additionally, the eye gaze tracker attached to a gaze controlled display can be used to detect cognitive load or distraction from driving.

5.4 Cluttering of display due to hotspots

Putting hotspots on every clickable items on a screen may clutter the display. Large [2016] and colleagues classified existing automotive user interfaces into four types – list, menu, map and text entry. The clickable items for list and menu selection has less density (#items/screen area) and putting a small spot on them will not clutter the display. For Map manipulation, hotspot will not be sufficient as users may need to click and drag anywhere in the map, a different study [Biswas 2016] proposed a fixation duration based technique for map manipulation. For virtual keyboard, the minimum distance between two hotspots may not be greater than the accuracy of the gaze tracker due to high screen-density, making hotspots less worthy. A ‘Dasher’ [Ward 2000] type of virtual keyboard will be more appropriate for gaze controlled interface.

5.5 Interference with caption for hotspots

When the buttons will have captions on them, it may be possible that the optimum position of the hotspot will be on the caption itself. It may be noted that the aim of the hotspot is to leverage the pop out effect and it needs not to be a physical spot itself. Bold-facing or rendering in reverse-contrast a single letter of the caption or placing a high-contrast icon can also serve as a hotspot.

5.6 Vibrating environment

It may be noted that all our studies were conducted in driving and flying simulators inside laboratory and thus we missed to measure effect of vibration on the performance of eye gaze and finger tracking sensors. A study in automotive environment [Biswas 2016] reported variation in vibration along vertical axis of a screen is more detrimental to performance of eye gaze controlled system than along X or Z axes. In a more precise experiment, Adelstein [2008] and colleagues reported “significant degradations in both error rate and response time in a reading task at 0.5 and 0.7 g for 10-pt, and at 0.7 g for 14-pt font displays”. We undertook a study where we collected data from 9 participants on an eye gaze and finger controlled system inside a car running at approximately 60km/hr on

a highway. However, unlike our previous studies, the driver himself could not undertake trials due to risk of causing accident and a fellow passenger undertook the trials. The design of the study was otherwise similar to the studies described earlier in the paper. We found an average selection time of 1.99 seconds (std dev 0.86 seconds). We are now designing a more accurate study under variable vibrating condition.

6. CONCLUSION

This paper proposes an algorithm to facilitate human machine interaction in automotive environment by using a gaze controlled projected (head-up) display. Our proposed system does not require drivers to look down and taking eyes off from road while operating a dashboard. The performance of the system was evaluated using a dual task setting involving a driving simulator and it is found that the pointing and selection times are not significantly different between the proposed system and touchscreen display, rather driving performance was improved with the proposed system than existing touchscreen display.

7. REFERENCES

- [Ashdown 2005] Ashdown M., Oka K. and Sato Y.(2005), Combining Head Tracking and Mouse Input for a GUI on Multiple Monitors, CHI Late breaking Result
- [Bates 1999] Bates R., Multimodal Eye-Based Interaction for Zoomed Target Selection on a Standard Graphical User Interface. Proceedings of INTERACT 1999
- [Biswas 2015] Biswas P. and Langdon P., Multimodal Intelligent Eye-Gaze Tracking System, International Journal of Human-Computer Interaction 31 (4), Taylor & Francis, Print ISSN: 1044-7318
- [Biswas 2016] Biswas P., Exploring the use of Eye Gaze Controlled Interfaces in Automotive Environments, Springer 2016, ISBN: 978-3-319-40708-1
- [CameraMouse 2013] CameraMouse, Available at: <http://www.cameramouse.com>, Accessed on 22nd September 2013
- [Card 1983] Card S., Moran T. and Newell A. "The Psychology of Human-Computer Interaction." Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1983
- [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

- [Dostal 2013] Dostal J., Kristensson P. O. and Quigley A.(2013), Subtle Gaze-Dependent Techniques for Visualising Display Changes in Multi-Display Environments, ACM International Conference of Intelligent User Interfaces (IUI) 2013
- [Duarte 2014] Duarte R. S. and Oliveira B. F., Comparing The Glance Time Required To Operate Standart Interface Devices And Touchscreen Devices In Automobiles, Proceedings of ACM INTERACCIÓN 2014
- [Duchowski 2007] Duchowski A. T.(2007), Eye Tracking Methodology. Springer, 2007
- [Farrell 2005] Farrell S., Zhai S.(2005), System and method for selectively expanding or contracting a portion of a display using eye-gaze tracking, US Patent No.: 20050047629 A1
- [Fejtova 2009] Fejtova M. et. al. (2009) , Hands-free interaction with a computer and other technologies, Universal Access in the Information Society 8
- [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
- [Fitts 1954] Fitts P.M. (1954) The Information Capacity of The Human Motor System In Controlling The Amplitude of Movement, Journal of Experimental Psychology 47 pp 381-391.
- [Fu 2007] Fu Y and Huang T. S. (2007), hMouse: Head Tracking Driven Virtual Computer Mouse, IEEE Workshop on Applications of Computer Vision
- [Grandt 2013] Grandt M., Pfendler C., Mooshage O., Empirical Comparison of Five Input Devices for Anti-Air Warfare Operators, Available at: http://dodccrp.org/events/8th_ICCRTS/pdf/035.pdf, Accessed on 30/11/2013
- [Jagacinski 1985] Jagacinski R. J. and Monk D. L. (1985), Fitts' Law in Two Dimensions with Hand and Head Movements, Journal of Motor Behaviour, 17(1), 1985, pp 77-95
- [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
- [Large 2016] Large D., Burnett G., Crundall E., Lawson G. and Skrypchuk L., Twist It, Touch It, Push It, Swipe It: Evaluating Secondary Input Devices for Use with an Automotive Touchscreen HMI, Proceedings of ACM AutoUI 2016
- [Lutteroth 2015] Lutteroth C., Penkar M. and Weber G., Gaze vs. Mouse: A Fast and Accurate Gaze-Only Click Alternative, Proceedings of ACM UIST 2015
- [MacKenzie 1991] MacKenzie, I. S., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. Proceedings of the CHI '91 Conference on Human Factors in Computing Systems, pp. 161-166. New York: ACM.
- [May 2014] May K. R., Walker B. N., Gable T. M., A Multimodal Air Gesture Interface for In Vehicle Menu Navigation, Adjunct Proceedings of ACM Automotive UI 2014
- [Miniotas 2001] Miniotas D. (2001), Application of Fitts' Law to Eye Gaze Interaction, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) 2001, pp 339-340
- [Mondragon 2013] Mondragon C. K. and Bleacher B., Eye tracking control of vehicle entertainment systems, Patent No. WO2013036632
- [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
- [Penkar 2012] Penkar A.M., Lutteroth C. and Weber G., Designing for the Eye - Design Parameters for Dwell in Gaze Interaction, OZCHI 2012
- [Pfeuffer 2016] Pfeuffer, K., & Gellersen, H. (2016, October). Gaze and Touch Interaction on Tablets. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (pp. 301-311). ACM.
- [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
- [Seder 2012] Seder T. A., Szczerba J. F. and Cui D., Virtual cursor for road scene object selection on full windshield head-up display, Patent No.: US20120174004
- [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
- [Vertegaal 2008] Vertegaal R. (2008), A Fitts' Law Comparison of Eye Tracking and Manual Input in the Selection of Visual Targets, Proceedings of the International Conference of Multimodal Interaction 2008, pp 241- 248
- [Voelker 2015] Voelker, S., Matvienko, A., Schöning, J., & Borchers, J. (2015). Combining Direct and Indirect Touch Input for Interactive Desktop Workspaces using Gaze Input.
- [Ward 2000] Ward D. J., Blackwell A. F. and MacKay D. J. C., Dasher - a Data Entry Interface Using Continuous Gestures and Language Models, In proceedings UIST 2000
- [Ware 1987] Ware C. and Mikaelian H. M (1987), An Evaluation of an Eye Tracker as a Device for Computer Input, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) 1987, pp 183-187
- [Weinberg 2012] Weinberg G., Knowles A. and Langer P., BullsEye: An Automotive Touch Interface that's always on Target, Adjunct Proceedings of ACM Automotive UI 2014
- [Zandera 2010] Zandera T. O., Gaertner M., Kothea C. & Vilimek R.(2010), Combining Eye Gaze Input With a Brain-Computer Interface for Touchless Human-Computer Interaction, International Journal of Human-Computer Interaction, 27(1)
- [Zhai 199] Zhai S., Morimoto C. and Ihde S.(1999), Manual and Gaze Input Cascaded (MAGIC) Pointing. ACM SIGCHI Conference on Human Factors in Computing System (CHI) 1999.
- [Zhai 200] Zhai S., Hunter M. and Smith B. A., The Metropolis Keyboard – An Exploration of Quantitative Techniques for Virtual Keyboard Design, Proceedings of ACM Symposium on User Interface Software and Technology (UIST 2000), November 5-8, 2000
- [Zhang 2013] Zhang Y., Bulling A. and Gellersen H.(2013), SideWays: A Gaze Interface for Spontaneous Interaction with Situated Displays, ACM SIGCHI Conference on Human Factors in Computing System (CHI) 2013