Gaze-controlled Driving

Martin Tall

IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen, Denmark
info@martintall.com

Alexandre Alapetite

Technical University of Denmark
Dept. of Management Engineering
Produktionstorvet 426, DK-2800
Kongens. Lyngby, Denmark
alexandre@alapetite.net

Javier San Agustin

IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen, Denmark
javier@itu.dk

Henrik H.T Skovsgaard

IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen, Denmark
henrik.skovsgaard@itu.dk

John Paulin Hansen

IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen, Denmark
paulin@itu.dk

Dan Witzner Hansen

IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen, Denmark
witzner@itu.dk

Emilie Møllenbach

Loughborough University.
Applied Vision Research Centre
(AVRC), Loughborough,
Leicestershire, UK. LE11 TU
e.mollenbach@lboro.ac.uk

Abstract

We investigate if the gaze (point of regard) can control a remote vehicle driving on a racing track. Five different input devices (on-screen buttons, mouse-pointing low-cost webcam eye tracker and two commercial eye tracking systems) provide heading and speed control on the scene view transmitted from the moving robot. Gaze control was found to be similar to mouse control. This suggests that robots and wheelchairs may be controlled "hands-free" through gaze. Low precision gaze tracking and image transmission delays had noticeable effect on performance.

Keywords

Gaze, control, input, robot, mobile, wheelchair

ACM Classification Keywords

H5.2. User Interfaces — Input devices and strategies.

Introduction

Control of wheelchairs and remote vehicles could both benefit from effective hands-free input. Previous research has suggested voice control of wheelchairs [1], and electromyogram (EMG) signals, face gestures, etc. [2]. There are only a few reports on the use of gaze control [3, 4]. Matsumo et al. [3] achieved a precision of 3° with their wheelchair mounted tracking system. They were inspired by the fact that a person

often looks into the direction of the next move when walking or driving, so this may be a natural and intuitive way to control a wheelchair. However, they decided not to utilize gaze direction but to use the tracking of face direction (head movements) instead, because gaze cannot be reliably controlled by intention in a dynamic environment where, for instance, other people walk around. Issues with highly variable lighting conditions (e.g. sunlight and neon lights shining directly into the camera pointing upwards a user's face) have been reported [5]. According to their experience vibrations from the moving wheelchair also complicated tracking. This suggests that gaze tracking needs to be precise and robust to control a vehicle.

The "TeleGaze" interface [6] for gaze control of a remote robot is overlaid on top of the video stream from the robots camera. The user gives navigation commands by fixating on overlaid interface elements. A command will be given for as long as the user fixates on the region. Looking outside of the control regions will not issue any commands.

There are three different design approaches to controlling a vehicle by gaze: 1) Directly by just looking where you would like to drive. 2) Indirectly by gazing at on-screen buttons that executes commands like "forward", "stop" etc.; 3) By gazing at an image of the front-view. We decided to explore the last version since it provides a direct mapping between "target point" and gaze point that is the virtue of the first approach, but with a fast and obvious way of braking – namely to look away from the screen.

Eye gaze robot construction

The robot prototype was built around a plastic frame using some Lego® Mindstorms NXT components, which are fast, easy and affordable to produce. The robot platform has two independent drive wheels at the front, and two passive wheels at the back (figure 1). This setting is similar to some motorized wheelchairs, and allows driving forward, backward, left, right, including in-place rotations.
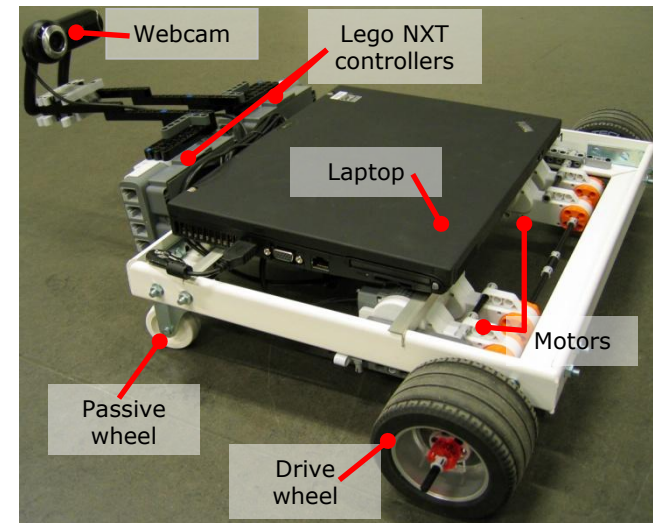

figure 1. Mobile robot carrying a laptop computer.

The robot alone weights 1.9 Kg, and it is able to carry up to 10 Kg. With a medium load such as a laptop computer, its speed is ~0.5 m/s (depending on battery level). The laptop computer controls the robot via Bluetooth, and communicates with the user's computer over Wi-Fi or 3G. The laptop computer can be avoided if the robot stays within Bluetooth class-2 range

(~10m) of the operator, and if a wireless camera is used (which was not the case for the experiments).

Our interface design provides a direct feedback loop with no visible interface components displayed. We utilize the point of regard on the screen directly as the user observes the streaming video to continuously adjust the locomotion of the robot. The direction and speed is the modulated linearly by the distance from the centre point of the monitor (figure 2). Two dimensions are combined, the X-axis modulates steering and the Y-axis modulates speed. Commands are issued every 100ms., continuously updating the navigation instructions.
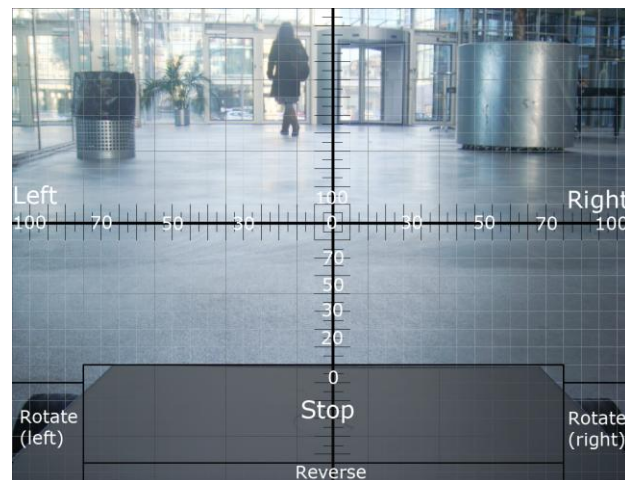

figure 2. An illustration of the invisible control functions put on top of the video scene image from the robot. The X-axis modulates steering (from -100% to +100% where 0% driving straight) and the Y-axis modulates speed (From -50% to +100% where 0% is stop).

Looking at an object right in front of the robot will reduce speed – a natural way to brake for an obstacle. Fixating one of the wheels will issue a rotation (sharp turn) in that direction.

Method

Participants
A total of 5 male volunteers, ranging from 27 to 49 years old, participated in this study. All of them had previous experience with gaze interaction, and three of them used contact lenses.

Apparatus
Five different input devices were used to control the robot: optic mouse, on-screen buttons, and three gaze trackers. Two of the three gaze trackers were commercial systems (SMI iViewX RED and Tobii 1750), while the last one was a low-cost, webcam-based system that we have developed.

figure 3. Experimental setup for the control interface.

Design and Procedure

The experiment required the participants to remotely drive a robot along a track built on the ground (figure 4). The experiment was conducted using a within-subjects factorial design, with *input device* (mouse pointing, on-screen buttons mouse clicking, Tobii, SMI and webcam) being the factor under study. The order of input device was counterbalanced across participants, i.e. each participant completed a lap with each input device. In every trial we measured *lap time*, *bin hits* (number of times the robot hit a bin), and *line crossings* (number of times one wheel was outside the track). A brief 2 minute introduction to the interface was given. No test runs were allowed.

figure 4. Experimental setup for the test circuit.

Results

Analysis of the robot-control task was performed using two ANOVAs, with *input device* as the independent variable. Lap time and total error rate were analyzed as the dependent variables. All data were included.

LAP TIME

The mean over-all time for a successful lap was 184s ($SD = 53$). There was a significant effect from input device, $F(4, 24) = 4.4$, $p < 0.05$ (figure 4). Lap completion time was fastest with the mouse ($M = 147s$, $SD = 6$) and the webcam gaze tracker was slowest ($M = 247s$, $SD = 34$). A LSD post-hoc test showed significant difference between the webcam gaze tracker and all other input devices.
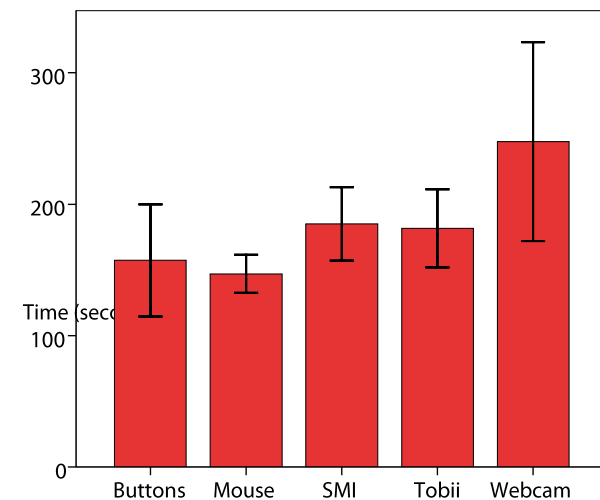

figure 5. Mean lap completion time for each input device. Error bars show ± 1 standard deviation.

TOTAL ERROR RATE

Line crossing and obstacle hits were combined in a measure of the overall error rate of 4.3 (SD = 2.6). There was a significant effect from input device, $F(4, 24)$ = 3.0, $p < 0.05$. Error rate was lowest with the Tobii tracker (M = 3.0, SD = 1.7) and the mouse (M = 3.0, SD = 2.0); the webcam gaze tracker produced the most errors (M = 7.2, SD = 2.3). The LSD post hoc analysis showed significant difference between the webcam gaze tracker and all other input devices. Figure 6 shows the mean error rate and standard deviation for each input device in the experiment. A Pearson's Correlation showed a significant relationship between lap time and error rate (r = 0.58, $p < 0.01$).
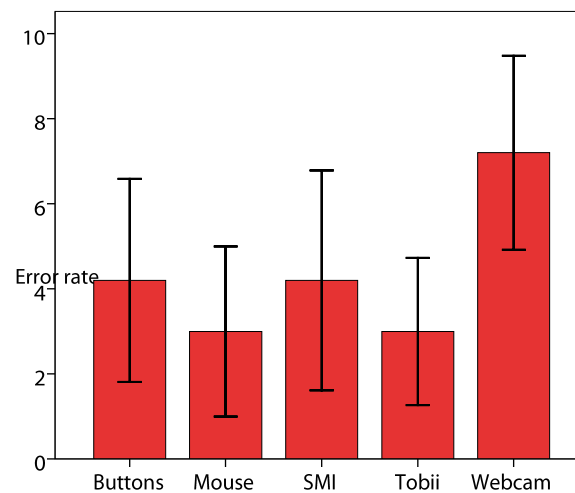

figure 6. Mean total error rate from the experiment. Error bars show ± 1 standard deviation.

Discussion

Our initial prototype was developed to investigate navigation by directly gaze input, without using traditional GUI components, to achieve hands-free control of a vehicle. In this first experiment our focus was on gaze as a unimodal input, leaving aside other considerations such as safety issues, sunlight disturbances and vibrations of cameras from driving on rough surfaces. Further on we intend to investigate possibilities for multimodal interaction e.g. combining gaze with voice or EMG inputs.

In our experiment all participants managed to complete the circuit on their first attempt, on all of the input devices. The low-cost eye tracker, using a head-mounted webcam, caused most errors and longest lap times. In the webcam setup, head movements will slightly off-set the gaze position. When users tried to reacquire correct gaze positioning erroneous navigation commands were sometimes issued. Hence, the stability of the eye tracking device is crucial. This is demonstrated by the results as there is a significant difference between the webcam eye tracker and all other devices. There was no significant difference between the other devices; this may suggest that the highly accurate eye trackers are able to provide control of vehicles that are as good – in terms of speed and errors – as mouse control. Future experiments will investigate how gaze control compares with more traditional modalities for vehicle control, such as joysticks and steering wheels. However, these devices are troublesome individuals with severe motor impairments such as Amyotrophic Lateral Sclerosis (ALS).

We acknowledge the differences between our prototype and controlling a wheelchair by gaze. The video images from a camera fixed to the chair will be different from our experimental setup, since the user will then move with the camera. We hope to acquire an electronic wheelchair to evaluate the performance of this method of interaction and navigation.

Furthermore, this type of navigation could be beneficial in a multimodal remote control scenario where the hands are required for other tasks. In this setting it is crucial with a high quality video link. We observed that lags in the image sequence might cause commands to be issued towards points that had already have been passed (this will not be an issue in the case of wheelchair control).

In conclusion, our initial experimental setup may serve as a simple, safe and affordable test bed for future design of gaze-controlled mobility, possibly supplemented with other modalities.

References
[1] Mazo, M., Rodríguez, F.J., Lázaro, J.L., Ureña, J., García, J.C., Santiso, E., Revenga, P., García, J.J. (1995). Wheelchair for physically disabled people with voice, ultrasonic and infrared sensor control. Autonomous Robots 2(3):203-224.

[2] Moon, I., Lee, M., Ryu, J., Mun, M. (2003). Intelligent robotic wheelchair with EMG-, gesture-, and voice-based interfaces. Proceedings of Intelligent Robots and Systems 2003, pp 3453-3458. DOI:10.1109/IROS.2003.1249690

[3] Matsumoto Y., Ino, T. & Ogasawara, T. (2001). Development of Intelligent Wheelchair System with Face and Gaze Based Interface, Proceedings of 10th IEEE Int. Workshop on Robot and Human Communication (ROMAN 2001), pp. 262-267.

[4] Roberts, A., Pruehsner, W. & Enderle, J.D. (1999). Vocal, motorized, and environmentally controlled chair. Proceedings of the IEEE 25th Annual Northeast Bioengineering Conference, 1999, p. 33-34.

[5] Canzler, U. & Kraiss, K.-F. (2004). Person-Adaptive Facial Feature Analysis for an Advanced Wheelchair User-Interface. In: Paul Drews (Eds.): Conference on Mechatronics & Robotics 2004 , Volume Part III, pp. 871-876, September 13 – 15, Aachen, Sascha Eysoldt Verlag, ISBN 3-938153-50-X

[6] Hemin, O. L., Nasser, S. & Ahmad L. (2008) "Remote Control of Mobile Robots through Human Eye Gaze: The Design and Evaluation of an Interface", SPIE Europe Security and Defence 2008, Cardiff, UK