INTERMODAL DIFFERENCES IN DISTRACTION EFFECTS WHILE CONTROLLING AUTOMOTIVE USER INTERFACES

Michael Geiger[*], Martin Zobl[*], Klaus Bengler[+], Manfred Lang[*]

[*] Institute for Human-Machine Communication
Technical University of Munich, D-80290 Munich, Germany
[+] BMW AG, dep. EV-22, D-80788 Munich, Germany

geiger@ei.tum.de, zobl@ei.tum.de, klaus.bengler@bmw.de, lang@ei.tum.de

ABSTRACT

In this study, the haptical and gestural user input modes are compared with regard to distraction from a controlling task similar to steering a car. The examination is carried out in a driving-simulation laboratory. While performing the controlling task permanently, the test subject has to execute a variety of user input with a given modality. Haptical input is done by using buttons or a rotary knob while gestural input consists of dynamic right hand movements in a designated space. During the experiment, the controlling error, the object recognition performance, and the required duration for each user input are measured. The results show a significant benefit for gestural input.

1. INTRODUCTION

Due to the multiplicity of coexisting electronic devices in modern luxury and upper class automobiles, the drivers' task to control this huge functional range while driving has become increasingly complex. Some examples of these devices are navigation systems, audio and video components like CD player, radio, and television, mobile phone, car computer, air condition, and any additional conceivable unit such as internet applications. For the purpose of enabling the user to handle these car devices, an optimized man machine interface is desirable. Nowadays, user input in cars is mainly done haptically via buttons or rotary knobs. Our basic approach is to use additional input channels besides the tactile one. Humans mainly pass on information using speech and gestures. This study examines the potential of visual gesture input to control in-car devices with regard to distraction. We understand gestures as dynamic right hand movements performed in the field of vision of a camera mounted at the car's ceiling. The gesture vocabulary used is the outcome of foregoing studies and is reduced to an intuitive set of hand-gestures. Our assumption that gestural input might cause less distraction effects than haptical input bases on recent psychological experiments on selective visual processing [Deu98] [Pap99]. These studies have shown that goal-directed deictic movements are generally coupled with substantial mental distraction effects, which are verified by an almost nonexistent capability of object recognition performance. They conclude that it is not possible to attend to one object and, at the same time, to point to another. Transferring these results to our problem, the haptical actuation of a button is similar to a goal directed movement. Accordingly, the use of dynamic hand gestures, which are less directional than haptical interactions, promises an efficient reduction in loss of attention.

2. METHODOLOGY

2.1 Test Environment and Conditions

To ensure an authentic environment, this study is carried out in a driving-simulation laboratory (cf. Fig. 1). In order to avoid undesired side effects on the data, caused by the characteristics of the driving simulation application (implemented car model, graphics etc.), we have developed an abstract steering task (cf. Fig. 2) displayed on a projection area in front of the car. Thereby, the steering nominal is given as a marker following horizontal movements similar to a road course. The controlling task consists of following the nominal marker with a second marker, which is controlled by the subject's steering wheel, whereby the angle of the steering wheel is translated into horizontal deflection of the actual marker.

While steering, the subject has to perform a variety of user input (cf. input tasks, Tab. 1) with a given modality. Haptical input is done using buttons or a rotary knob placed in the midconsole (cf. Fig. 3a) and according dynamic right hand gestures in a designated space above the midconsole (cf. Fig. 3b).

To get comparable data, the sequential control of the experiment is automated, which means that all events happen exactly at the same time. The input tasks are read out to the subject via pre-recorded speech. After the test, the subject is asked by which kind of user input (haptical or gestural) it felt more distracted from the steering task, and which one was more comfortable.

From: Proc. HCI 2001, Lawrence Erlbaum Ass., NJ

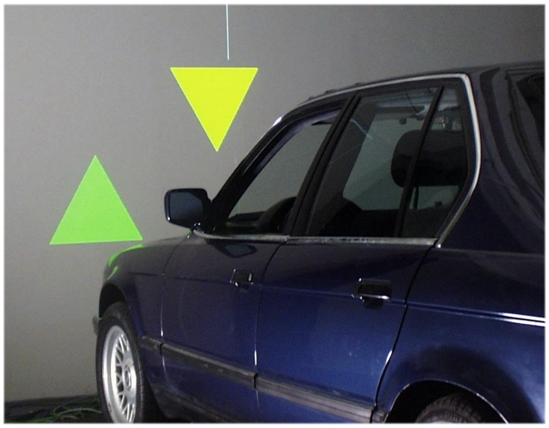

Figure 1: driving simulation laboratory

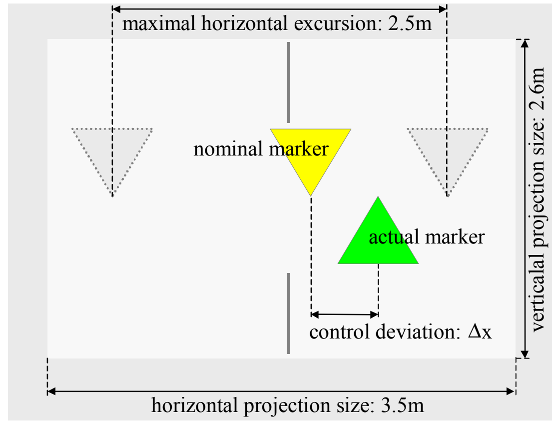

Figure 2: steering task

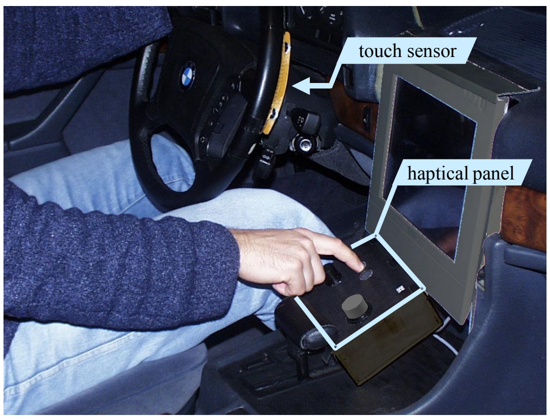

Figure 3a: haptical user input (select button)

Figure 3b: gestural user input (telephone gesture)

Input Tasks

Input Method		Select	Phone	Increase	Decrease
	Haptical	Press ‘select’ button	Press ‘phone’ button	Turn rotary knob clockwise	Turn rotary knob counter clockwise
	Gestural	Point to front	Lift virtual phone receiver	Up-movement	Down-movement

Table 1: input tasks

2.2 Data Logging

In order to estimate the amount of distraction while processing user input, the nominal and actual x-positions, the object recognition performance and the required duration are measured and stored to a log file.

2.2.1 Nominal x-Position and Actual x-Position

Both nominal x-position and actual x-position are recorded with a sampling rate of 50 Hz (cf. curves Fig. 5). During the test design, the motion of the nominal marker was set to an average speed that allows the controlling to be done with a very small deviation error if the subject exclusively concentrates on that main task. The x-position data is used to calculate the control error (cf. Results) which describes the subject’s capability of following the road.

2.2.2 Required Duration

The required duration is the time interval while the steering wheel is released, measured by a touch sensor (cf. Fig. 3a) mounted to the steering wheel. Before the experiment starts, the subject is instructed to keep this sensor pressed all the time unless releasing is necessary to do an input task. Additionally, all input tasks have to be done as fast as possible and without interruption. To stimulate the subject to perform as fast as possible, an acoustical feedback is presented at the end of every task in form of a sine tone with a frequency that increases (and becomes more unpleasant) with elapsed time. The required duration is one measure for the efficiency of the respective input method.

2.2.3 Object recognition performance

Object recognition performance is measured as recognition error rate with 2AFC, in which the subject's task is to recognize the identity of a visual object presented for a short time. The object (seven-segment display, cf. Fig. 4) is presented at the actual position of the nominal marker within a varying time delay after the release of the steering wheel.

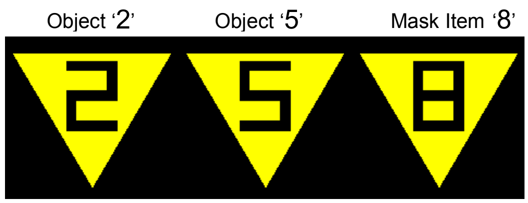

Figure 4: visual objects to identify and mask item

To assure the object's presentation coming along during the execution of the input task, this delay lasts at least 100 ms but does not exceed 500 ms. The object is then replaced by a mask item ('8') 70 ms later in order to avoid the imprint of the pattern to the retina. The object recognition performance indicates whether the subject's visual attention is directed to the front (traffic event) or not.

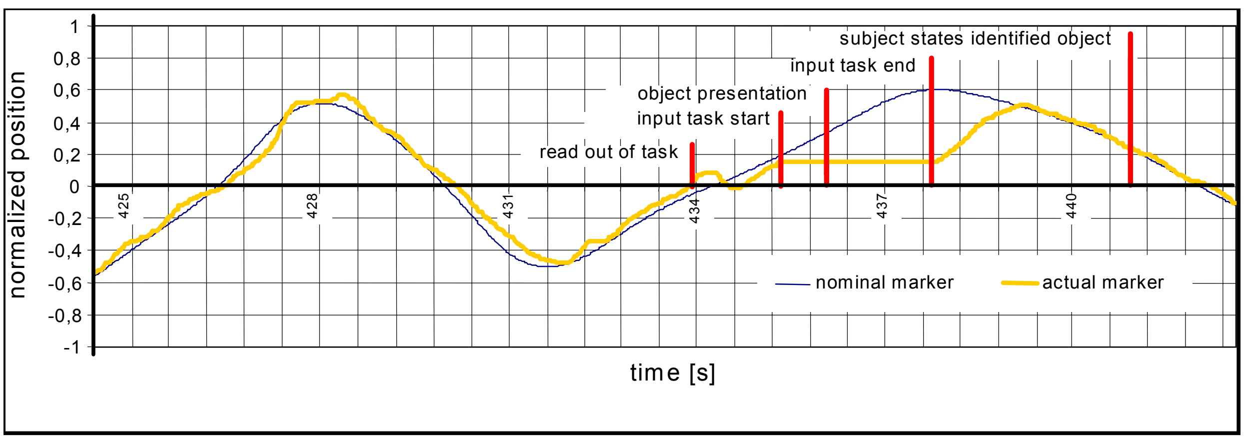

Figure 5: Timechart (excerpt): x-positions of nominal and actual marker as well as occurring events during input task; Remarkable: even the task's read out causes an attention loss; in this example, the subject stops the controlling task entirely while doing user input.

2.3 Problem of delayed input task performance

During the test design the following problem occurred. The subjects did not perform the input tasks - as demanded - as fast as possible, but waited for the object presentation before doing user input. The implementation of acoustic feedback (cf. Required Duration) could not solve this problem completely. Therefore, we invented additional events in order to prevent the subject attuning to the object presentation during an input task: On the one hand, now there are tasks without object presentation. On the other hand, random object occurrence outside an input task is implemented. Altogether, there are 32 input tasks: 16 haptical and 16 gestural user inputs, whereas only half of the tasks come along with a displayed object. Furthermore, there are 16 object presentations outside input tasks.

3. RESULTS

For data comparison, the differences of the matched observations $d_i=x_i^{hap}-x_i^{ges}$ for all measured factors are regarded. Since the differences are apparently normally distributed the t-test is used. The null hypothesis H_0: $\mu_d=0$ suggests that gestural and haptical values are identical and is tested against the alternative H_A: $\mu_d>0$ that haptical input produces higher values than gestural input. In our case, higher values stand for poorer suitability for in-car use concerning all measured data.

3.1 Controlling Error

The controlling error is calculated by the root mean square of the horizontal deviation Δx (cf. Fig. 2) between the nominal and actual value of the steering task [Joh93]. To obtain distance independent values, the pixel difference is converted to a deviation angle. The shape of the two distributions (cf. Fig. 6) is almost identical whereas the distribution for haptical input is shifted to higher deviation angles. The null hypothesis that the distribution of the differences has a mean of $\mu_d=0$ can be rejected on a significance level $\alpha<0.01$. The $1-\alpha$ interval for the true mean is $\mu_d\in[0.18\ 0.58]$, whereas a deviation angle of 0.58 deg corresponds to an offset of about 1m in a distance of 100m.

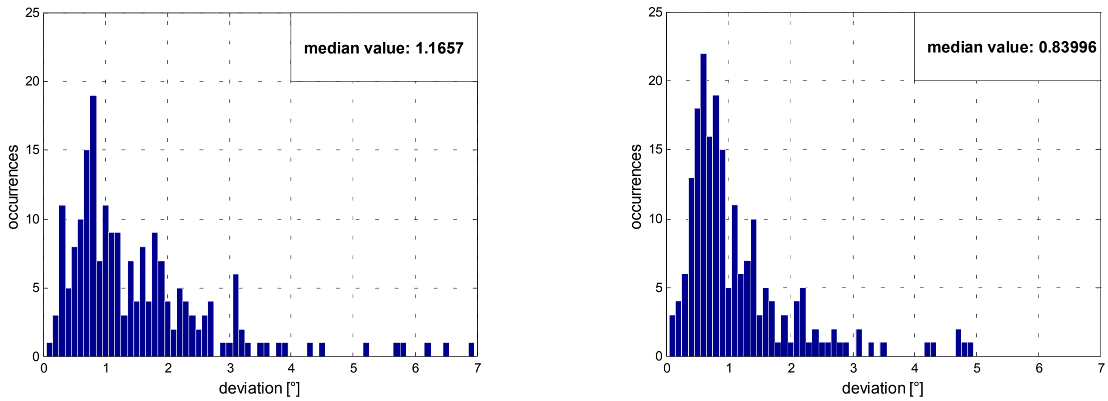

Figure 6: control error for haptical input (left) and for gestural input (right)

3.2 Required Duration

The data for the required duration (cf. Fig. 7) show the same results as for the controlling error, with the $1-\alpha$ interval for the true mean $\mu_d\in[0.36\ 0.57]$. This means that in average haptical input takes about *1.4* times longer than a gestural one.

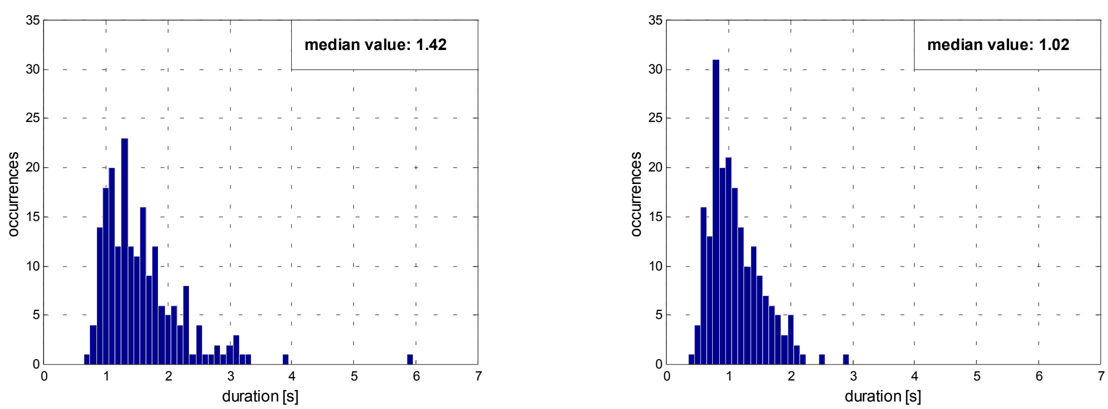

Figure 7: required duration for haptical input (left) and for gestural input (right)

3.3 Object Recognition Performance

The error rates (cf. Fig. 8) also show that gestural input is superior to haptical input regarding the recognition performance. The median recognition error rate for gestural input is about half the size of the haptical one. Using 2AFC, an error rate of *0.5* means pure guess. The null hypothesis can be rejected with $\alpha<0.01$, too. The $1\text{-}\alpha$ interval for the true mean is $\mu_d \in [0.08\ 0.32]$. The subject´s attention is obviously not directed to the front with haptical input.

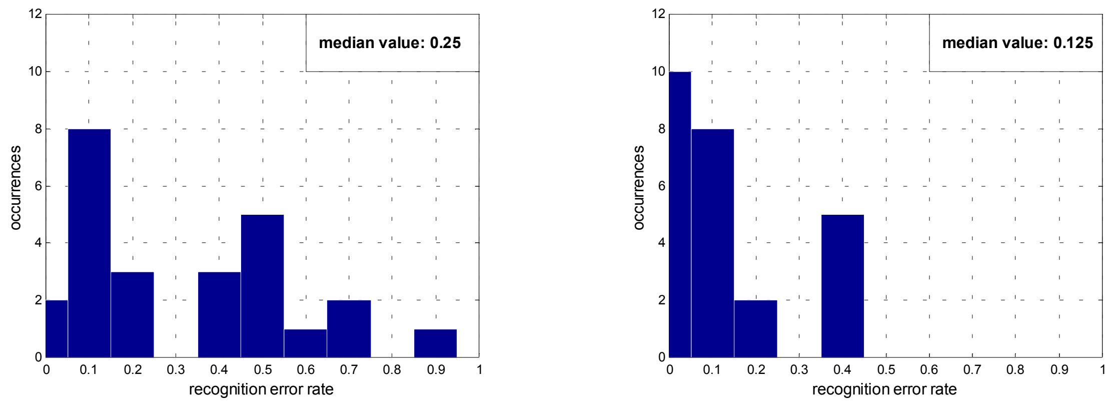

Figure 8: recognition error rate for haptical input (left) and for gestural input (right)

3.4 User Acceptance

After the experiment, the subjects complete a questionnaire. The questionnaire results are in correspondence to the objective outcomes. According to the subjects' statements, gestural user input distracts less (94% of the subjects) than haptical (6%) and is more pleasant (76% vs. 24%).

4. CONCLUSIONS

The findings argue for substantially reduced distractions when using gestural user input for controlling an automotive MMI. Therefore, we expect a significant gain in user acceptance by using gestural user input for certain tasks in an automotive environment. The next step in our work will be to implement the institutes gesture recognition system [Mor99] into the car environment.

REFERENCES

[Deu98] Deubel, H., Schneider, W. X., Paprotta, I. (1998). Selective Dorsal and Ventral Processing: Evidence for a Common Attentional Mechanism in Reaching and Perception. *Visual Cognition*, pp.81-107

[Joh93] Johannsen, G. (1993). *Mensch-Maschine-Systeme* Available: Springer Verlag, ISBN 3-540-56152-8, Berlin-Heidelberg-New York.

[Mor99] Morguet, P., Lang, M. (1999). Comparison of Approaches to Continous Hand Gesture Recognition for a Visual Dialog System. *Proceedings ICASSP 99 (Phoenix, Arizona, USA), IEEE, Vol. 6*, pp. 3549-3552

[Pap99] Paprotta, I. (1999). *Selektive Wahrnehmung und Handlungssteuerung, Die Kopplung von visueller Aufmerksamkeit und Bewegungszielselektion*. Available: Shaker Verlag Aachen, ISBN 3-8265-6588-6.