A Context-Adaptive Search Engine Concept and Multimodal Input Strategies for Automotive Environments

Markus Ablassmeier and Gerhard Rigoll
Institute for Human-Machine Communication
Munich University of Technology
Arcisstr. 16, 80290 Munich, Germany
phone: +49 89 289-28541
{ablassmeier | rigoll}@ei.tum.de

ABSTRACT

The constantly growing amount of information in cars implicates the development of new strategies to cope with this amount of information for drivers. An intelligent search engine reduces the problem of distracting menu navigation in deep hierarchies and consequently increases the concentration on the primary driving task. The evaluation shows that the new search engine concept reached a very high user acceptance and the objective data proved observably acceleration in handling. Also different input strategies for selection and alphanumeric input were evaluated for using while driving. The major problem of alphanumeric input can be resolved by multimodal input devices.

Keywords

search engine, automotive, GUI design, menu, driver information systems, input devices

INTRODUCTION

Nowadays, besides the function of merely a transportation system, automobiles take more and more the task of a multimodal information system. Due to the enormous increase of functions, especially over the last ten years, the tendency to integrated menu-driven displays and central handling concepts is unstoppable [1]. Therefore, the cabin is also called driver's working environment. In this case, a major problem is the fast and efficient retrieval of information and functions within these interaction concepts.

For this reason, a new intelligent search engine concept for cars has been developed to accelerate and simplify the interaction of the driver with the system, and has been implemented in a user-interface for cars. The goal of the intelligent search engine reduces the problem of menu navigation in deep hierarchies of driver-information-systems, and must not distract the driver from the main task, which is the vehicle guidance.

For multimodal interaction several input strategies for the text input and result selection have been evaluated for appropriateness while driving.

BACKGROUND

This chapter gives some background information about the driving task in general and context influence. Further, main aspects of graphical user interfaces (GUI), multimodal interfaces, clustering in menu structures and the function of search engines are explained. Afterwards relevant guidelines are presented.

Driving-Tasks

Taxonomy

Compared to the automobile domain, in front of a desktop PC, the user can predominantly execute her or his operations in a concentrated way, as there is no dual task competition. Especially in the car domain, often error-prone situations occur regarding human-machine interaction with different in-car applications, as the driver often has a certain mental workload. This basic stress level is due to the execution of so-called primary and secondary tasks, and may be increased by environmental impacts, like the conversation with a co-driver. If the driver interacts, i.e., with a communication and infotainment system in such a stress phase (tertiary task), inattention, distraction, and irritation occur as a consequence of the high workload resulting from a superposition of the tasks mentioned above, which will become manifest in an increased error potential and in erroneous operations of these tertiary systems. [2] introduce an in-depth classification of driving tasks.

Context Influences

During driving, a large set of influencing variables are effective on the interaction and, as a consequence, on the dialog between driver and the tertiary systems to be operated. These factors are summarized as context

IUI 06, Sydney, Australia
ACM Press

parameters and can be divided into three subgroups: environ-mental, user, and system context parameters (see Figure 1).

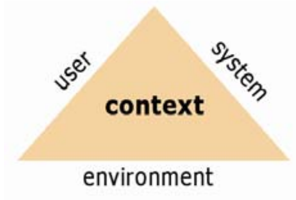

Figure 1. Contextual triangle

Graphical User Interfaces (GUI)

Multimodal Input

Multimodal interfaces combine natural input modes - such as speech, pen, touch, manual gestures, gaze and head and body movements - in a coordinated manner. Multimodal interfaces are largely inspired by the goal of supporting more transparent, flexible, effective, effcient and robust interaction. The flexible use of input modes is an important design issue. This includes the choice of the appropriate modality for different types of information, the use of combined input modes, or the alternate use between modes. Input modalities can be selected according to context and task by the user or system. Especially for complex tasks and environments, multimodal systems permit the user to interact more effectively. Because there are large individual differences in abilities and preferences, it is essential to support selection and control for diverse user groups. For this reason, multimodal interfaces are expected to be easier to learn and use. The continuously changing demands of mobile applications enables the user to shift these modalities, e.g. in-vehicle applications

Menu Structures

Generally, a menu is a list of selectable options, which after a choice of the user leads to a change of the system status [3]. If the number of options is too large to be presented efficiently in a single menu panel, it is often necessary to design a menu structure.

Hierarchical menus are a series of menus which are structured in a hierarchical or "tree-like" manner, where the selection of an initial option leads to another menu containing additional options, until the desired results are obtained.

Network menus are series of menus structured as a network providing redundant pathways to either all or some of the menus within the structure. A special application of network dialogs is the internet; the term Hypertext was created by Ted Nelson in 1965 to describe non-linear writing in which you follow associative paths through a world of textual documents [3]. The most common use of hypertext these days is found in the World Wide Web (WWW) pages and is the basis of the internet form today. The user of the WWW follows specially highlighted links on a web page to get from one node to another node within the network.

Search Engines

A search engine is a program that looks up documents for specified keywords, and returns a list of the documents, where the keywords were found. Although a search engine is a general class of programs, the term is often used to specify systems, like Alta Vista or Google that enable users to search for documents on the WWW. [5]

Search engines are the key to finding specific information on the vast expanse of the www. There are basically three types of search engines: those that are powered by crawlers, or spiders; those that are powered by human submissions; and those that are a combination of the two. Crawler-based engines send crawlers or spiders out into cyberspace. These crawlers visit a web site, read the information on the actual site, read the site's meta tags and also follow the links that the site connects to. The crawler returns all that information back to a central depository where the data is indexed. The crawler will periodically return to the sites to check for any information that has changed, and the frequency with which this happens is determined by the administrators of the search engine. Human-powered search engines rely on humans to submit information that is subsequently indexed and cataloged. Only information that is submitted is put into the index. In both cases, a search engine query to locate information is actually searching through the index that the search engine has created and is not searching the Web. These indices are giant databases of information that is collected, and stored, and subsequently searched.

The same search on different search engines produces different results, because not all indices are going to be exactly the same. It depends on what the spiders find or what the humans submitted. But not every search engine uses the same algorithm to search through the indices. The algorithm is what the search engines use to determine the relevance of the information in the index to what the user is searching for. One of the elements that a search engine algorithm scans for is the frequency and location of keywords on a Web page. Those with higher frequency are typically considered more relevant. Another common element that algorithms analyze is the way that pages link to other pages in the Web. By analyzing how pages link to each other, an engine can both determine what a page is about and whether that page is considered "important" and deserving of a boost in ranking.

Guidelines for Dialog Structures

To improve the usability of an application, it is essential to have a well designed interface. Shneiderman's "Eight Golden Rules of Interface Design" [6] are an important guide to good interaction design. These are: strive for consistency, enable frequent users to use shortcuts, offer informative feedback, design dialog to yield closure, offer simple error handling, permit easy reversal of actions, support internal locus of control, reduce short-term memory load.

Several other guidelines are very useful for designing and evaluating dialog structures:

The ISO 9241 part 10 [7] presents ergonomic dialog principles which apply to the design: suitability for the task, self-descriptiveness, controllability, conformity with user expectations, error tolerance, suitability for individualization, and suitability for learning.

The ISO/DIS15005 2000 [8] describes the ergonomic principles to be applied in formulation of dialogs between the driver of a road vehicle, and the information and control system, when on the move: The dialog between the system and the driver must therefore take account of the workload on the driver as a whole, including the cognitive, perceptive and physical functions associated with driving so that the dialog does not prevent the driver from conducting the vehicle from being driven properly and safely.

This background knowledge is the basis for the following concept and implementation of using a search engine as a graphical-user-interface in cars.

CONCEPT AND IMPLEMENTATION

In this new search engine concept for cars, the driver is supported by rapid access to the relevant information under consideration of actual context information. The information basis can be accessed via one graphical input mask on a central information display (CID). The entry of search items or words is supported with several input modalities. Especially the input of alphanumeric terms is facilitated via speech recognition, handwriting detection via touch pad, text over a numerical pad (T9), and a speller via a turning knob.

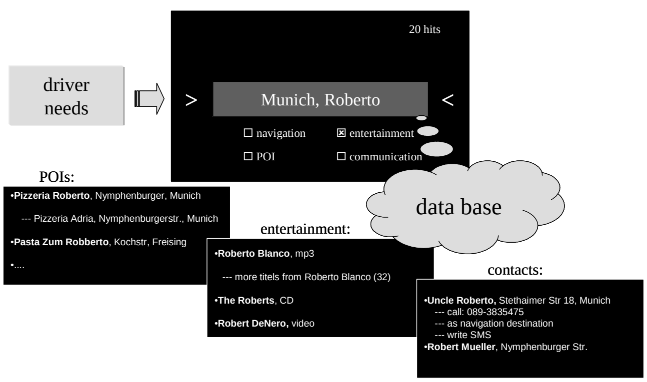

Figure 2. Search engine concept idea

For each hit, an individual and dynamic ranking for assignment of relevance is identified. After calculation of the ranking (a score-based algorithm) the search behavior of the driver (search history and frequency) and the distance to the actual location can be interpreted. Also a unit for error management is integrated in the search algorithm to increase the robustness of the system. Different spellings of search words or names (e.g. Meier or Mayer) and twisted letters (e.g. rsetaurant instead of restaurant) were also accepted and interpreted.

The hit list can be browsed or navigated with the central turning knob or the speech recognition to reach the relevant information or function. Two databases are available for retrieval of relevant hits: a local (local onboard car data-base) and a global database (centrally administrated). The global database can be accessed, e.g., via a wireless telecommunication system. The database consists of point of interests (POIs), music titles, contacts, menu entries, handbook, etc. Beyond this, the system has the possibility to send inquiries to real internet search engines. To guarantee an integrated information management, the search system is connected with the vehicle functions and the conventional hierarchical menu structure. Thus, for each hit, connected functionalities can be used. It is, for example, possible to send an address or GPS data directly to the navigation system, and let the system calculate a possible driving route. A phone number of an entry can be transmitted to the mobile phone system or shortcuts to menu items can be selected directly.

This concept idea provides parallel function access to the user [9]. The driver can access all the information directly from only one initial point, which is the entry mask. From the result list, the driver can select the desired information or functionality by browsing. No paths for example through a menu structure have to be memorized by the driver.

This search engine concept idea has been implemented on the basis of the FERMUS development framework [10] and a rapid-prototyping simulation software using an existent driver-information-system simulation.

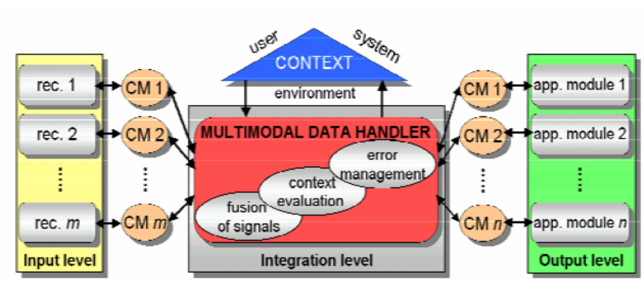

Figure 3. Multimodal framework based on a late semantic fusion of input

The system architecture basically consists of three main processing levels: the input level, the integration level, and the output level (see figure 3). The input level contains any kind of interface that is capable of recognizing user in-puts (e.g., turning knobs, buttons, speech recognizer, etc.). Dedicated command mappers (CMs) encode the information bits of the single independent modality recognizers and context sensors into a meta language based on a context-free grammar (CFG). In the integration level, the recognizer outputs and additional information of context sensors (e.g., information about application environment, user state, etc.) are combined in a late semantic fusion process that is extensively illustrated in [11].

EVALUATION

The developed search engine was analyzed in a usability test (Experiment 1). To verify, whether the concept idea has a real added value for the driver, a driving simulation task has been applied. In a second Experiment (Experiment 2) the focus was on input devices for alphanumeric input. Input devices play a major role for the use of this intelligent search engine.

Experiment 1 “Search Engine”

The automotive trial platform was a test car in a laboratory environment. Test persons had to follow a predefined street course in a laboratory driving simulation (see figure 4). This test was developed within the scope of the project FERMUS for assessing ease-of-use and distraction potential in a reliable, reproducible and economical manner [12].

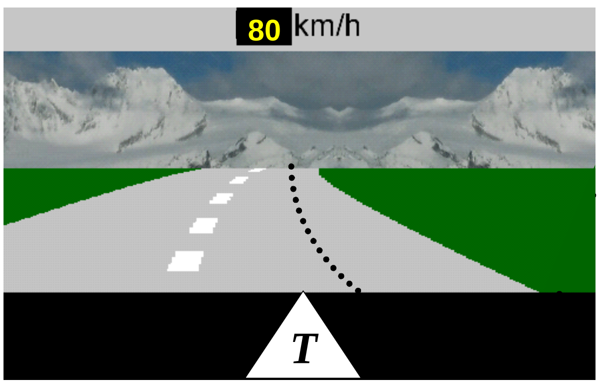

Figure 4. Screenshot of the 3D-driving simulation

In this study 14 test persons of higher qualifications participated with the age of 19 till 48 (average age was 28 years). The test persons had to perform diverse tasks regarding the operation of the multimodal in-car system. The target application was a navigation and entertainment system, consisting of an MP3-player, navigation, and a POI database simulation. As tactile modalities, the user was provided a central control unit integrated in the armrest and a 10” screen located at the center console (see figure 5). Input could also be delivered via command-based speech.

The subject was told to solve different tasks of interaction with an existing graphical menu structure and the new search engine. For evaluating the input modalities different instructions of usage were given. The test persons fulfilled the following tasks in several passes:

- entering a driving destination: e.g. using the turning knob as input modality for navigation in the menu structure or the search engine, and the speller for alphanumeric input of street and city
- entering a phone number: e.g. via numeric pad as input modality
- menu navigation to car functions: e.g. change of sound settings

For each task sequence, they had to proceed one ride by using the new search engine and one ride by using the ordinary hierarchical menu style. Afterwards, the persons were interviewed with a questionnaire about handling, acceptance, benefit, and distraction effects.

Figure 5. The cockpit interface

For evaluation of the deviation from the lane keeping an error score has been logged. Furthermore, interaction times for each tasks were analyzed. When the test series had been finished, the constellation has been transcribed and filed in terms of test data sets.

Results 1 “Search Engine”

Regarding time data, it shows that through the newly introduced search engine, significant time savings have been reached. In the case of the task “entering a driving destination” an average time saving of 45,2 % resulted. The task for “entering phone numbers” can be absolved with an average time saving of 8,6 % compared to the menu structure. The deeper the menu points are located in the hierarchy, the higher time savings have been reached with the search engine. With regard to the average track deviation, the new functions evoked no verifiable irregularity.

In the subjective rating, a high agreement of the test persons with the new search engine occurred: To the statement “This function makes sense”, 12 of 14 persons (85,7 %) agreed with this statement, so that an average rating of 1.14 has been reached (1- I agree; 4 - I disagree).

Evaluating the different modalities, the command-based speech entry got the best rating with 1.36, the numeric pad got 2.79, and the turning knob 3.71 (1- very good; 6 - unsatisfactory). But all test persons considered a turning knob appropriate for scrolling and selecting in the result list. Every test subject classified the speech entry suitable for entering search items and 9 of 14 the numeric pad. As an additional input device 5 persons wished a touch screen or touch pad.

According to the ranking of the POI search results the test persons (7 of 14) considered personal preferences as most important followed by distance to destination (4 of 14) and alphabetical order (4 of 14). Only two wished a consideration of the search behavior of other drivers. Four persons suggested to regard the time to destination instead of the distance.

Summarizing, the new search engine concept for cars has reached a very high subjective acceptance by the test subjects, and the objective data has proved significant acceleration in handling.

Experiment 2 “Input devices”

Various tactile input elements were implemented and tested in the static driving simulator located in the Navigation Laboratory of our Institute. Four of these input devices were further investigated within a usability test: a multifunctional turning knob (MTK), a touchscreen for handwriting recognition (HRD), a numerical keypad adapted for T9 text input (NUM) and finally a touchscreen adapted for on-screen use (OSK). 15 participants, ages ranging from 22 to 60, were asked to enter an address as a secondary task, while driving. The order in which the input devices were tested, varied between the test subjects. For simulation purposes, the program Lane Change was used. This test was developed within the scope of the project ADAM (Advanced Driver Attention Metrics) for assessing ease-of-use and distraction potential in a reliable, reproducible and economical manner. The setup consists of a PC monitor, a steering wheel, and a set of pedals. The probands see a three-lane road stretching out before them (see Figure 6). Signs on the left and right sides of the road indicate the lane they should be driving in.

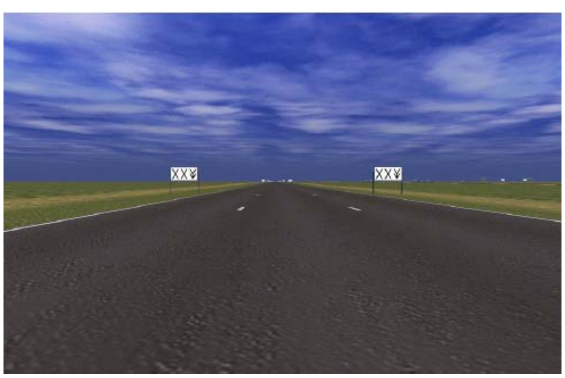

Figure 6 The Lane Change Task

With a special software tool the lane deviation from the ideal course as a measure for distraction and the duration of tasks can be calculated afterwards (see Figure 6).

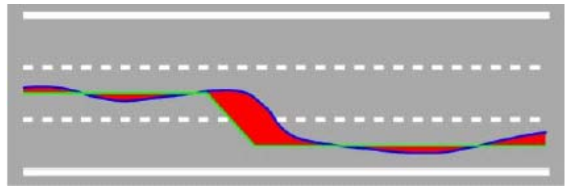

Figure 6 Lane deviation from the set course

In addition to the video material gathered during the test, the driving proficiency was also measured with the Lane Change Task program. Subjective criteria was compiled using a structured questionnaire.

Results 2 “Input devices”

Driving Performance

The driving performance was measured as explained in Figure 6 by comparing the mean lane-deviation. The following chart 1 summarizes the results. It represents the mean deviation while performing the complete input sequence. This sequence includes entering the city-name, the street-name and the house-number. The NUM shows the largest lane deviation and therefore a bad driving performance. This can be attributed to different aspects. One important fact is, that the user actually had to look down to the middle-console in order to manipulate the controller. This affected the driving performance. The OSK and the HRD performed well in this aspect, showing mean lane deviations of 1.45 m and 1.46 m respectively. Here the HRD benefited from the almost blind use. The visual distraction was greatly reduced with this device. Just one control gaze to the CID was necessary, instead of constant visual contact with the input device.

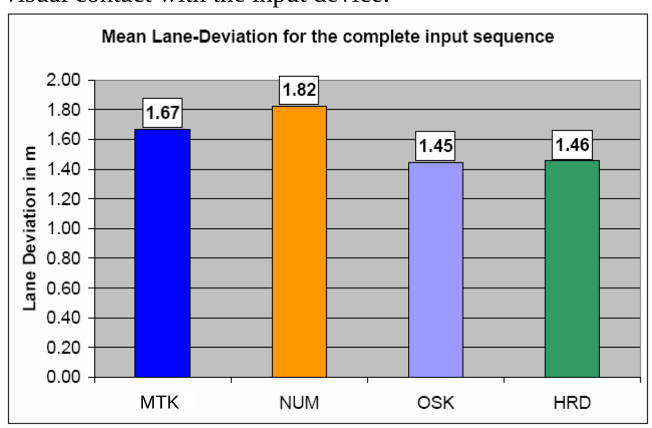

Chart 1. Mean Lane-Deviation

Input Time

Another aspect considered during the test was the actual time each user spent, while performing the input. The results can be seen on chart 2. The total input time can be considered as the time the user manipulated the controller in addition to the time he spent correcting his mistakes. Therefore different aspects have to be considered when interpreting the results. The manipulation time, that refers to the time it takes to perform the input of any character with a specific devices, greatly variates from device to device. Here, the OSK proved as a very fast input method, since just a touch on the specific letter was required for the entering it. On the other hand, the manipulation of the HRD was more time consuming, since the complete letter had to be "painted" onto the touchscreen. The use of the MTK-speller involves turning the knob to the desired character and then confirming the highlighted character by pressing the controller. This step require a not negligible amount of time. Finally the time required for manipulating the NUM depended on the experience the user had with the device. While some users had great difficulties getting used to typing the letters with the keypad, others did this very fastly. The difference is mainly based on the experience with the predictive text input system and the knowledge of the key-layout. While most of the users usually type their short messages on their cellphones using their thumbs, the NUM had to be manipulated with the other fingers, since the device was fixed into the middle-console. The experience using the numeric keypad on cellular phones could therefore not be directly transferred to this device.

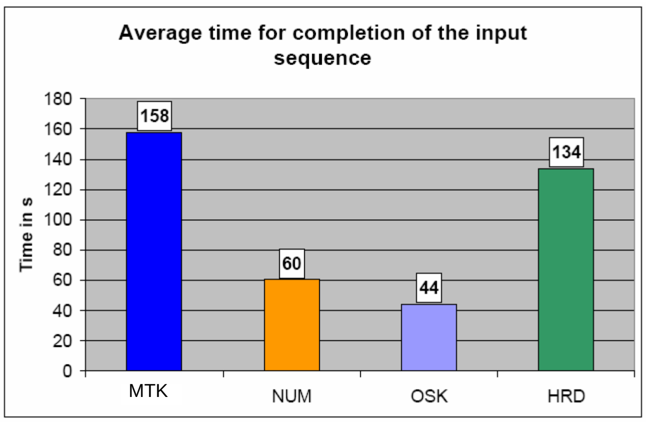

Chart 2. Average input time

Subjective Results

Concerning the safety, most of the users agreed that alphanumeric input is risky while driving. The device that was considered as the safest one, was the HRD, mainly because it doesn't require a control gaze and can be manipulated blindly. The rest of the devices were classified as risky, since all of them required visual confirmation of the entered data and therefore distracted the driver. As can be seen in chart 3, the MTK controller was considered as the most stressful input method. This is probably due to the difficult error correction and the long amount of time each input requires. None of the devices was actually considered relaxing but the OSK was evaluated as the most relaxing input method of this study. Concerning the ease of use, the OSK, the NUM and the HRD were judged as easy and intuitive devices as can be seen of Figure 5.6. On the other hand, the manipulation of the MTK controller was more confusing, specially when the users had to figure out if they had to move or turn the knob.

General Preference

The question "Which input method do you prefer for in-car use?" was answered as follows. Supporting the objective data, 47% of the users favored the OSK as an input device for in car use, followed by the HRD with 33%. Just one person actually prefered using the NUM. This person had previous experience with the NUM and therefore performed the input blindly and in a very short time. These results further support the interpretation of the objective data, where the OSK was clearly the device which performed best in a driving environment. However the impossibility to perform a blind input can be very distracting.

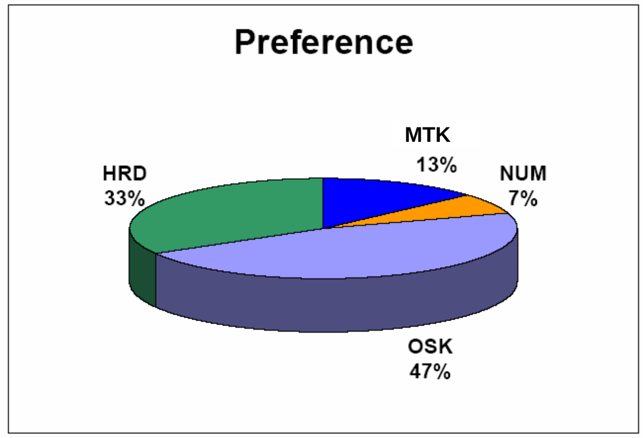

Chart 3. Input preference

Concerning the safety, most of the users agreed that alphanumeric input have a risky potential while driving. The device that was considered as the safest one, was the HRD, mainly because it doesn't require a control gaze and can be manipulated blindly. The rest of the devices were classified as risky, since all of them required visual confirmation of the entered data and therefore distracted the driver. The MTK controller was considered as the most stressful input method. This is probably due to the difficult error correction and the long amount of time each input requires. None of the devices was actually considered relaxing but the OSK was evaluated as the most relaxing input method of this study. Concerning the ease of use, the OSK, the NUM and the HRD were judged as easy and intuitive devices. On the other hand, the manipulation of the MTK controller was more confusing, especially when the users had to figure out if they had to move or turn the knob.

CONCLUSION AND OUTLOOK

One main aspect of these studies had been to primarily improve the traffic safety and to accelerate the intuitive retrieval of information and functions in vehicles. The results of the usability study have shown that through an intelligent search engine, time savings concerning the HMI-interaction of about 30 % can be reached. As there has been no evident influence on the drivers' attention, the main goal of more safety and user-friendliness for handling is achieved. The preferred input modality is speech. In respect to alphanumeric tactile interaction many persons asked for two parallel, multimodal input devices, so that they could choose according to the situation. In a standing car, the users preferred the use of the OSK, since it proved to be a fast and reliable input device. During the actual driving phase the persons favored the use of the touchpad with handwriting recognition or the MTK controller. To meet the drivers' intention, further research should be focused on cognitive stress while driving, intelligent information management for respective driving situation and transparent search criteria.

REFERENCES

1. Niedermaier, B.: Development and evaluation of a rapid-prototyping approach for multimodal human-machine-interaction in automobiles. Munich, Technical University of Munich, Institute for Human-Machine Communication, Dissertation, 2002.
2. Donges, E.: The principle foreseeableness as interpretation concept as measures for active security measures. In: Man-machine the system in traffic, VDI reports (1992), NR. 948
3. Norm ISO9241 part 14: Menu Dialogues. In: Ergonomic requirements for officework with visual display terminals (VDT-s), 1996.
4. Mulzer, W.; Zhang, S.: The Little History of The World Wide Web, FU Berlin, Department of Informatics, 2001. URL: http://www.inf.fuberlin.de/lehre/SS01/hc/www, 2003-08-12.
5. Searching & Finding in the Internet or: "The Needle in a Haystack", 6. edition, September 2001, Regionales Rechenzentrum für Niedersachsen, Universität Hannover
6. Shneiderman, Ben: Designing the user interface: strategies for effective human-computer interaction. 3. ed., Mass. [u.a.]: Addison-Wesley, 1998.
7. Norm ISO9241 part 10: Dialog Principles. In: Ergonomic requirements for officework with visual display terminals (VDT-s), 1996.
8. Norm ISO/DIS15005: Road vehicles In: Ergonomic aspects of transport information and control systems - Dia-logue management principles and compliance procedures, 2002.
9. Ablaßmeier, M.; Niedermaier, B.; Rigoll, G.: A New Approach of Using Network Dialog Structures in Cars. In: SCI 2004, 8th World Multiconference on Systemics, Cybernetics, and Informatics SCI 2004, 18.-21.07.2004, Orlando, Florida, USA. Publ.: N. Callaos; W. Lesso; B. Sanchez. IIIS Orlando 2004, Vol. XI, p. 227.
10. G. McGlaun, F. Althoff, M. Lang, and G. Rigoll: Development of a Generic Multimodal Framework for Handling Error Patterns during Human-Machine Interaction, In: SCI 2004, 8th World Multi-Conference on Systems, Cybernetics, and Informatics, Orlando, FL, USA
11. G. McGlaun, et al.: "A New Approach for Integrating Multimodal Input via Late Semantic Fusion," USE-WARE 2002, Darmstadt, Germany; in: VDI Bericht 1678; pp. 181-185, 2002
12. G. McGlaun, F. Althoff, B. Schuller, and M. Lang: A New Technique for Adjusting Distraction Moments in Multitasking Non-field Usability Tests, In: CHI 2002, Int. Conference on Human Factors in Computing Systems, Minneapolis, MN, USA