2nd Multimodal Interfaces for Automotive Applications (MIAA 2010)

Michael Feld, Christian Müller
German Research Center for Artificial Intelligence
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
{*michael.feld*, *christian.mueller*}*@dfki.de*

Tim Schwartz
Cluster of Excellence, Multimodal
Computing and Interaction
University of the Saarland
Germany
schwarz@cs.uni-sb.de

ABSTRACT
This paper summarizes the main objectives of the 2nd IUI workshop on multimodal interfaces for automotive applications (MIAA 2010).

Keywords: multimodal interfaces, human-machine-interaction, automotive applications

ACM Classification: H.5.1 Multimedia Information Systems. – Evaluation – Methodology; H.5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces, Natural Language, User-centered design, Voice I/O, Interaction styles, Haptic I/O, Ergononmics; J7 [Computers in other systems]: Command and Control, Consumer Products

General terms: Design, Reliability, Experimentation, Human Factors

INTRODUCTION

The technology we interact with is no longer static, but will react completely different depending on who is using it. Personal computers and mobile devices have already been affected by this development in recent years. As vehicles are accumulating more advanced technological components, it is clear that they are heading in the same direction. But not only because a certain degree of customizability is expected today: As cars host such a large range of accessories and devices on little space, a seamless interaction between user and system entails an enormous benefit with respect to usability.

The basis for a modern human-machine interaction concept is a user-centric design that focuses on the needs of the target users. An advanced user-adaptive system however goes further: There it is not sufficient to develop the interface on the drawing board with a specific target group in mind and then implement a set of fixed interactions for that group, but to also incorporate the actual state of the user and its interrelations with the context in the dialog system. This can be done for example by adjusting system output to the cognitive load of the user or enable different input methods depending on what interactions related to other tasks such as driving or communicating with passengers are being performed in parallel.

IUI'10, February 7–10, 2010, Hong Kong, China.
ACM 978-1-60558-515-4/10/02.

MAIN OBJECTIVES OF THE WORKSHOP

One research aspect of such user-adaptive systems is the acquisition of knowledge. Different functions will be able to provide different degrees of service and personalization depending on what is known about the subject. The highest level of service can be accomplished when we know exactly whom we are dealing with, i.e. if the user identifies himself by some credentials, voice identification or an ID token. There are several examples where such tokens are incorporated into car keys, ID cards or other common accessories. In that case, we can look up basic information in a local database. With more vehicles being equipped with Internet access, such information can also be looked up on web services, enabling an easier sharing of preferences when multiple cars are used. The greater challenge however is the situation when the user does not identify himself, either because he is in a hurry, because he is not a regular driver of the car (imagine a car rental service) or because he is simply joining in some other car and not driving at all. Explicit information input – while generally less desirable because of the overall discomfort generated by entering it – is not a realistic option for the driver. In this case, we can still fuel our systems with knowledge from probabilistic models, typically based on sensor information. The same applies to most of the rapidly changing aspects of user state. Such non-intrusive information acquisition methods have a strong potential for in-car applications. While many other scenarios suffer the problem of difficult instrumentation, in the automotive field, we are dealing with a rather small and well-defined interaction space (i.e. a single seat), where it is easier to place cameras and other sensors. In addition, there are already many existing sensors available to be exploited, for instance weight and temperature sensors. Also, while the speech of a user is typically used only as a primary input to control a system, there is much more information we can obtain about the speaker from his voice, like age, gender, emotion, cognitive load and even the level of alcohol in his blood.

The collection of data from external sources is not the only aspect related to user knowledge acquisition. When we have many pieces of information from many different sources, an intelligent fusion mechanism needs to be designed that also takes into account context parameters. This is of particular importance whenever uncertainty is involved. When we consider for example the attention state of driver, we may have some the visuals of the person's eyes as one cue, while other cues are derived from reaction times and driving style. Then again, the time of day and duration of driving play a role as context parameters. In the end, we need to come up with a single value – preferably annotated by a meaningful confidence value – that an application or driving assistance function can make use of to adapt itself to this circumstance. There are many different ways of how such fusion of information can be done. One option is logical inference systems, which work quite well for fact knowledge. Rule engines may be another choice, and they are preferred in cases where complex inferences based on expert knowledge is needed, or when are they serve to model rules set by laws, a case that is omnipresent in automotive system design.

A different area of research addresses the question of how user knowledge is represented. In order to provide an accessible platform for other services, a certain degree of content is needed with respect to how the data is retrieved and possibly updated. Semantic knowledge representation in the form of ontologies is certainly a state-of-the-art solution here, yet there are few examples of public ontologies that focus on the driver and passengers as a special case of a user and that also take incorporate context aspects like car features and the current driving situation. The fact that information sharing between cars like with the car2x paradigm is becoming increasingly important for new safety and convenience features too requires us to think about the privacy aspects of data sharing too. Essentially there need to be methods to ensure that no sensitive data leaves the car or is "desensitized" on its way out. Further, in car systems we have the contrast of lengthy driving on the one hand and highly fluctuating information on the other, which makes the dimension of time another essential issue for knowledge representation. Services have to know the validity period of a datum, and may need to access data from already passed points in time.

The final objective in the design process of personalized in-car systems is the realization of the actual adaptation concepts. Eventually every aspect of multimodal input and output is adaptable: The way gestures are recognized, the type of output produced, the speed of voice output, the temperature settings and window operation mode, and the elements presented on the screen. The motivation behind it may vary between settings. It can serve to ensure an intuitive usage by making a system react like the user expects it to. It can also be used to customize functions to the user's preferences. Adaptation may be employed to compensate a physical or mental handicap of the user, or simply to reduce his cognitive load while driving. Identifying places where adaptation can be useful is the first step, followed by the identification of those characteristics in the user model that it should be based on. Then an adaptation strategy has to be chosen, which will determine the result visible to the user. Strategies can range from very function-specific effects, like the choice of roads in a navigation component, to generic ones, such as the increasing of font size for people with a bad visual acuity. Generic adaptation components are a further pertinent topic for further research. The selection of strategies typically comes along with a conflict resolution mechanism in case multiple strategies are in conflict or compete for a single output modality. This can avoid for instance that a user is presented with a long sequence of spoken output while the screen remains unused. How far an adaptation can go largely depends on the confidence in the underlying information.

The car HMI is subject to much stricter rules and undergoes a more rigorous testing than interfaces in other areas. Therefore, a formal specification of the HMI on multiple layers of abstraction is necessary in the design phase. However, current methods often do not include adaptivity as part of such a specification, hence new procedures have to be developed and existing ones need to be extended to incorporate these aspects. In a similar manner, current industry standard evaluation methods may also need to be updated.

Adaptation does not end when the user is confronted with a personalized output. Most of the time, uncertainty is involved to some extent in the process, hence the result may not always be appropriate, or it may be difficult to understand – up to the case where the user mistakes adaptive behavior for a system malfunction. The user's reaction to this at least partially autonomous decision of the system to adapt the way it did may provide important cues as to whether the intended goal was reached. Moreover, it can shed some light on the user's model of the system and his learning process. Trying to find out about the user's view of an adaptive system and using this information to improve the learning and make the behavior more transparent is subject to ongoing research. Future systems are expected to be much more user-friendly when they approach adaptivity with this meta-level aspect in mind.

REFERENCES

[1] Brandherm, B. und Jameson, A. (2004). An Extension of the Differential Approach for Bayesian Network Inference to Dynamic Bayesian Networks. International Journal of Intelligent Systems, 19(8), 727–748.

[2] Heckmann, D. (2005). Ubiquitous User Modeling. Department of Computer Science, Saarland University.

[3] Müller, C. (2005). Zweistufige kontextsensitive Sprecherklassifikation am Beispiel von Alter und Geschlecht [Two-layered Context-Sensitive Speaker Classification on the Example of Age and Gender]. Ph.D. thesis, Computer Science Institute, Saarland University