Natural Visual User Interfaces – Beyond Input Modalities

Nora Broy
BMW Group Research and Technology
Hanauerstraße 46
80992 Munich, Germany
+49 89 382 19583
nora.nb.broy@bmw.de

Sonja Rümelin
BMW Group Research and Technology
Hanauerstraße 46
80992 Munich, Germany
+49 89 382 51985
sonja.ruemelin@bmw.de

ABSTRACT
The understanding of Natural User Interfaces (NUIs) commonly refers to user input which is often realized through speech or gesture interaction. However, car manufacturers have not yet extensively integrated natural input into their products. In our opinion, the feedback channel back to the driver as well as natural output modalities have been neglected in the discussions of NUIs. Visual presentations that integrate strategies of human visual perception can possibly serve as an appropriate output channel, enrich speech and gesture interaction with feedback and in turn increase user acceptance. We propose the integration of stereoscopic or large display solutions for the development of NUI concepts. Our approach allows for a seamless integration of displays into the vehicle's environment and interior so that the borders between virtual and physical space become indistinct.

In this paper, we discuss potentials and limitations of Natural Visual User Interfaces (NVUIs) in an automotive context. We provide first steps towards assessing the design space of visual output for NUIs to inspire future research.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User interfaces – *Graphical user interfaces (GUI), Interaction styles (e.g., commands, menus, forms, direct manipulation).*

Keywords
Automotive user interfaces; graphical user interfaces; large display spaces; natural user interfaces; stereoscopic displays.

1. INTRODUCTION
The aim of designing NUIs is to turn the human machine interface (HMI) invisible by providing natural and easy to learn interactions. Recent investigations in NUIs concentrate on input design and more concrete on speech and gestures [8]. In general, UI development affects not only the communication and translation of the user's intended goals to the system but a much richer scope of presenting information [4]. We claim that NUIs need appropriate user feedback and output of the system's current state. In particular, automotive user interfaces have to provide urgent information about the vehicle's state and the driving situation in a comprehensible way. This is the case for visualizing information related to primary (e.g. velocity) and secondary tasks (e.g. indicator). Tertiary tasks (e.g. initiating a phone call, browse music library) offer a design space that allows the application of natural input techniques as gesture and speech. This paper discusses the shortcomings of these input techniques and proposes natural output modalities which might foster a more natural interaction with the car's user interface. We focus on visual output and suggest the integration of modern display technologies as autostereoscopic and large displays. In this context, we elucidate risks and potentials for NUIs inside the car and highlight opportunities for future work.

2. COMMUNICATION CHANNELS
In general, NUIs are understood as synonym for speech and gesture interaction. In the following, we discuss these input channels. Furthermore, we consider different output channels and estimate their contribution to NUIs in the automotive context. Thereby, we highlight the special role of visual feedback.

2.1 Typical Input Modalities for NUIs
Speech is one of the basic communication channels between humans. However, current speech input in cars is less natural since they are command-based. Alvarez et al. [1] tackle this problem by applying natural conversations as input modality and auditory system responses for output. Nevertheless, we think that people have inhibitions to talk to machines [3]. Moreover, users can phrase their goals in various ways which complicates the unambiguous interpretation of speech input even more.

Human communication heavily builds on body language as a modality to transmit information. In human-machine communication, gesture input is commonly applied for touch displays and freehand interaction. Touch gestures are integrated in cars already. However, it is often not clear in which parts of the screen a touch interaction can take place and which parts are inaccessible. Beside touch gestures, there is current research about freehand interaction in the vehicle [5]. Car manufacturers have not yet integrated freehand gestures in their products but are working hard to make it a part of the HMI of future vehicles. One of the problems they face is ambiguity; starting a certain action would require an activation gesture to prevent non-intended input. Moreover, in most use cases gestures induce an abstract function matching so the user needs to learn how to perform a specific action to trigger a function.

2.2 Output Modalities for NUIs
Auditory output in form of speech can transmit rich information. However, processing time causes delays in perception. The temporal dependency on the moment of output limits the driver in focusing on the feedback. This can be useful for urgent information transmission, for example abstract alert sounds. However, those cannot communicate a lot of information. For any kind of auditory output, the driver cannot choose the moment of attention switch. If the driving situation does not allow the perception, the temporal-dependent presentation is missed.

Haptic output can be used as feedback to enhance gestural interaction. For example, touch displays do not provide perceivable shapes of the displayed controls and confirming feedback while manipulation. Tactile feedback in form of vibrations is one way to compensate the lack of haptic feedback. With freehand gestures, the driver does not have a physical connection to the controls, suggesting the application of remote tactile feedback [6]. However, tactile output can interfere with car vibrations, making it unperceivable. Furthermore, haptic output is used to encode spatially related information like indicating an unintended lane change to the left or right. Similar to auditory alerts, the information density is limited and not sufficient for presenting complex information.

In contrast to the previously discussed output channels, the visual channel is rich and can present a large amount of information in a short time frame. Since the visual channel provides most important feedback for the primary driving task, it is crucial to design visual output in a reasonable way. The most important issue is to control information overload to inhibit perceptual tunneling and cognitive capture on the driver's side [7]. The reasons for particularly engaging interfaces have to be identified and avoided. Keeping these criteria in mind, the next section discusses the potential of visual output for NUIs.

3. VISUAL OUTPUT FOR NUI

We propose visual output as a new output channel for NUIs. We want to translate existing human visual perception strategies into HMI concepts in an automotive context. We claim that a natural visual presentation requires a seamless integration of displays into their environment. In this case, seamlessness is not restricted to the physical construction into the cockpit's surface but also includes the integration of the displayed content into the real world. We consider that such a visual output can enhance existing natural input, as speech and gesture, by providing a natural visualization of feedback, virtual content, and the vehicle's state. Furthermore, representing the car in a natural way possibly fosters the communication between driver and vehicle.

In the following, we present concrete approaches by discussing stereoscopic and large displays for NUIs in the car. At this point, we do not give a definition of natural visual output. Instead, we aim at motivating starting points for the design of visualization concepts to integrate natural input and output.

3.1 Stereoscopic Displays

Humans have a three-dimensional vision through the perception of horizontally shifted images. The discrepancy between the two projections onto the retinas enables an internal spatial representation of the environment. Stereoscopic displays generate an illusion of depth on a flat surface by presenting two distinct images to the viewer's eyes. Since humans perceive their environment with two eyes, this technology possibly fosters the naturalness of information visualization. Realistic spatial presentations can encourage the user to explore natural input techniques as freehand gestures. Moreover, the use of 3D depth allows structuring the displayed content, for example by putting the interaction focus to the foreground [2] or, for touch interaction, pushing inaccessible objects to the back.

Stereoscopic technologies use several encoding and decoding strategies for a correct projection of the two images into the eyes. Glasses-based technologies are not sufficient for an in-car situation since it is not applicable to force the driver to wear glasses in order to perceive the user interface correctly. Autostereoscopic approaches do not require any headgear to create a spatial impression. However, these technologies are limited to defined spatial areas where the viewer's eyes perceive the respective images. Tracking the position of the user and adjusting the viewing zones in accordance provides wider degrees of freedom for head movements. This technological approach simultaneously offers the integration of motion parallax which makes the spatial impression even more natural. The peril of stereoscopic technologies is that they can cause visual fatigue. We consider that future technological developments, e.g. holographic and volumetric displays, and appropriate content design can compensate this risk.

3.2 Large Display Spaces

Integrating screens that offer a large surface for the presentation of information and that are distributed over the cockpit's area result in what we call large display spaces. Different concept cars have been presented that integrate large display spaces spanning over the whole cockpit, like DICE[1] or Fun Vii[2].

Currently integrated screens do not take advantage of the large human field of view but clutter the huge amount of information on a small spot of the available cockpit area. The size of large display spaces offers the possibility to declutter virtual presentations in the human field of view by logically structuring the information. A spacious presentation of information can support the perception of content by giving room for natural visualizations and by localizing them according to their semantics. For example, rarely used content can be placed further away. Large display spaces present an overview of available information and starting points for enabling different functions concurrently. In combination with freehand gestures and touch functionality distributed over the cockpit's display spaces, various functions can be directly accessed without transitions in the menu hierarchy.

4. CONCLUSION AND FUTURE WORK

This paper aims at integrating visual output to enhance NUIs. We argue that the output channel has been neglected in discussions about NUIs. Information can be presented in a way users know from real life. We proposed stereoscopic and large displays as a natural output channel. Both approaches provide space for structuring information by using 3D depth or large areas. Speech and gesture interaction can benefit though a consistent visual presentation. Visual output has the potential to transmit rich information at one glance. However, the visual channel is the main feedback channel for the driver's primary task driving. Therefore, we structure our research as follows. We will develop interaction concepts integrating stereoscopic and large displays and implement them in prototypical realizations. We evaluate the concepts in driving situations to assess, and iteratively exclude, negative effects on driving performance and workload. Moreover, we believe that the presented output approaches increase joy of use and attractiveness.

In summary, we want to enrich the discussion of NUIs by adding visual output as a further modality to existing input approaches, facilitating future driver-vehicle communication.

[1] http://www.youtube.com/watch?v=-BfWS83vPks
[2] http://www.youtube.com/watch?v=PWq541Q95so

5. REFERENCES

[1] Alvarez, I., Alnizami, H., Dunbar J., Johnson A., Jackson F. and Gilbert, J. E. Designing Driver-centric Natural Voice User Interfaces. In *Adj. Proc. of AutomotiveUI'11*. 156-159.

[2] Broy, N., Schmidt, A., André, E. Is Stereoscopic 3D a Better Choice for Information Representation in the Car? In *Proc. of AutomotiveUI '11*.

[3] Grasso, M. The long-term adoption of speech recognition in medical applications. In *Proc. Symposium on Computer-Based Medical Systems.* 2003, 257–262.

[4] Norman, D. *The Psychology Of Everyday Things*. Basic Books. 1988.

[5] Mahr, A., Endres, C., Schneeberger, T., Müller, C. Determining Human-Centered Parameters of Ergonomic Micro-Gesture Interaction for Drivers Using the Theatre Approach. In *Proc. of AutomotiveUI '11*.

[6] Richter, H., Wiethoff, A. Augmenting Future In-Vehicle Interactions With Remote Tactile Feedback. In *Adj. Proc. of AutomotiveUI'11*. 162-163.

[7] Tönnis, M., Broy, V., Klinker, G. A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers. In *IEEE Virtual Reality Conference 2006*. 127-134

[8] Wigdor, D. and Wixon, D. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann. 2011.