Workshop on In-Vehicle Auditory Interactions at the 21st International Conference on Auditory Display

ICAD in space —
Interactive spatial sonification
International Conference on Auditory Display
08.–10. 07. 2015 I Graz, Austria
http://icad15.iem.at
Table of Contents

**Message from the Workshop Organizers**

Introduction to the Workshop on In-vehicle Auditory Interactions

Objectives

Summary of Contributions

Conclusion

Peculiarities

**Workshop Papers**

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soundscapes for In-vehicle Technologies</td>
<td>KatieAnna E. Wolf, Michael A. Nees</td>
<td>WP1:1-4</td>
</tr>
<tr>
<td>Influence of Inattentively Perceived Auditory Stimulation on Driving Performance</td>
<td>Khashayar Aghdaei, Andreas Riener</td>
<td>WP2:1-4</td>
</tr>
<tr>
<td>Target Matching Auditory Displays To Support Improved Driver Skills And Vehicle Coordination</td>
<td>Nicholas Powell, Jo Lumsden</td>
<td>WP3:1</td>
</tr>
<tr>
<td>In-car Active Sound Generation for Enhanced Feedback in Vehicles with Combustion engines or Electric Engines</td>
<td>Alois Sontacchi, Matthias Frank, Robert Höldrich</td>
<td>WP5:1-3</td>
</tr>
<tr>
<td>Experiment on the Use of Auditory Interfaces for Take-over Requests in Highly Automated Driving</td>
<td>Pavlo Bazilinskyy</td>
<td>WP6:1-4</td>
</tr>
<tr>
<td>A Software Library for In-Vehicle Auditory Display</td>
<td>Yota Morimoto</td>
<td>WP7:1</td>
</tr>
<tr>
<td>User Experience Evaluation Framework for Auditory In-vehicle Systems</td>
<td>Alexander Meschtscherjakov, Nicole Perterer, Arno Laminger, David Wilfinger, Sandra Trösterer, Manfred Tscheligi</td>
<td>WP8:1-2</td>
</tr>
</tbody>
</table>
Message from the Workshop Organizers
Introduction to the Workshop on In-vehicle Auditory Interactions

The scope and prevalence of in-vehicle technologies has dramatically expanded, which increases drivers’ visual, physical, and mental workload. In an attempt to reduce this workload, but allow for necessary interactions, auditory displays have been used in vehicles over the years, but only for basic information transmission. The goal of our joint workshop is to discuss the use of auditory displays for interaction in the vehicle at a more advanced level in order to offer better driver experience in rapidly changing vehicle environments. This full day workshop is intended to secure sufficient time for intermingling participants, presenting conceptual sounds, discussing issues, and integrating ideas.

Objectives

We have five explicit goals in our workshop:

• Provide an organized discussion about the topic of how auditory interactions can be efficiently and effectively applied to in-vehicle contexts;
• Build and nurture a new community that bridges the auditory display community with the automotive user interface community;
• Discuss and exchange ideas interactively within and across sub-communities;
• Suggest promising directions for future trans-disciplinary work; and
• Yield both immediate and long-term community-, research-, and design-guidance products.

To this end, we invited researchers and practitioners from all backgrounds, who are interested in auditory display and automotive user interface fields. In this discussion and the formation of future directions, we particularly focus on “unobtrusive interfaces” and aspects of “gamification”. Unobtrusiveness will allow auditory displays to be informative, but stay at the periphery of the driver’s attention and be easily accepted by end users. To accomplish this, we discussed several design guidelines based on blended sonification. Gamification can promote users to keep their motivation for behavioral change for a longer period of time (e.g. in the context of eco-driving applications).

This is a focused, relevant, timely workshop topic for ICAD attendees. The information that can be sonified for the drivers and passengers, or ways in which they can interact with the vehicle through auditory modality are endless, and the topic requires a rigorous discussion with those highly knowledgeable in auditory interactions. It is also a timely topic, as the types of vehicle (e.g., electric and autonomous vehicles) and in-vehicle technologies (e.g., PND Personal Navigation Devices, ADAS: Advanced Driver Assistance Systems, ITS: Intelligent Transportation Systems, etc.) have rapidly been increasing. This increasing information still needs to be passed to the driver in some way and these interactions must be designed and studied in a highly scientific but creative manner. Through achieving the goals of this workshop we expect to provide an opportunity to move this integrated field forward and build a solid community that includes ICAD.

Summary of Contributions

The position papers submitted to the workshop, “In-Vehicle Auditory Interactions” have undergone a rigorous peer-review process where the manuscripts have been reviewed by, at least, two reviewers each. Finally, eight position papers were selected for publication in the compilation of workshop papers and for presentation/discussion at the workshop, held on July 6th, 2015 in the 21st International Conference on Auditory Display (ICAD2015).

Wolf and Nees discuss a taxonomy of in-vehicle auditory interactions, which can be used to categorize diverse situations for in-vehicle auditory interactions: vehicle and driver conditions, collision avoidance, infotainment & monitoring automation. Moreover, they propose to apply continuous auditory displays using naturalistic soundscapes in the vehicle context and identify its limitations.

Similarly, Aghdaei and Riener propose to use inattentively perceived auditory cues to increase driving performance to induce efficient and economic driving. For example, they envision the use of specific mood-inducing music pieces or auditory icon type sounds (e.g., mimicking engine sounds) to change a driver's emotional and cognitive mental state. In the complex, attention-demanding driving situation, using implicit auditory displays can lead to better driving without adding to the driver’s workload.

Powell and Lumsden attempt to use target matching auditory displays in multiple vehicle situations, leveraging their preliminary research on accelerated racing. Given that there are many successful applications of target matching auditory displays for training and learning in arts, sports, and military domains, this approach looks promising.

Gable and Walker review their work about the application of auditory displays for vehicle infotainment systems, especially auditory menus. They wrap up the problem spaces and provide several solutions. Given that the traditional
auditory research has focused on collision warnings and, at most, used simple earcons, this is a new attempt to use of various sounds for the vehicle tertiary tasks.

Sontacchi, Frank, and Höldrich discuss the use of active sound generation (ASG) for the silent hybrid/electric vehicles to improve drivability, driving experience, and attractiveness. Based on the analysis and (re-)synthesis of the engine sounds, the authors obtained positive assessment results of their modified sounds compared to the original sounds. ASG could be actually used for electric vehicles or vehicles with small engines, considering multiple driver variables, such as gender, age, driving style, or cultural background.

Bazilinskyy suggests an experiment of using auditory interfaces for take-over requests in automated vehicles, which is one of the most critical research topics in Human Factors nowadays. The author attempts to conduct a study using the combination of various auditory cues - earcons, auditory icons, and speech sounds. Their future research using trimodal interfaces - visual, haptic, and auditory - would also be of interest to the community.

Morimoto has developed a software library for in-vehicle auditory displays using SuperCollider. This library is expected to provide the simplification of the tasks, the unified interface (real-time and non-real-time), and sound synthesis templates. This type of library will help other researchers to more systematically implement their own in-vehicle auditory displays and ultimately contribute to making a standardized developing protocol.

Finally, Meschtscherjakov, Prötter, Laminger, Wilfinger, Trösterer, and Tscheleig describe an evaluation framework to assess user experience in in-vehicle auditory interactions. Even though researchers could further develop and add vehicle-specific elements, this framework can be a good starting point to create the next version. Moreover, this framework can easily be used to conduct a short questionnaire study as they demonstrated in their paper.

Conclusion

Workshop papers bring up various discussion topics on in-vehicle auditory interactions, including a taxonomy of in-vehicle auditory interactions, sonification strategies (e.g., continuous soundscapes, implicit auditory displays, and target matching auditory displays), specific application areas (e.g., infotainment menu navigation, augmentation of drivability in electric vehicles, take-over requests in automated vehicles), and research frameworks for implementation (a software library for in-vehicle auditory displays) and evaluation (questionnaire factors). We believe that this multifaceted approach could contribute to integrating and maturing the automotive research community and the auditory research community. In conclusion, we greatly appreciate the work of all the authors, participants, and reviewers contributing to making this fruitful workshop possible.

Note

Workshop papers and presentation slides will be available at the ICAD2015 webpage (http://iem.kug.ac.at/icad15/) and the workshop webpage (https://sites.google.com/a/mtu.edu/icad2015-in-vehicle-auditory-interactions-workshop/home).

Peculiarities

This workshop is a joint effort of two groups of workshop organizers. The original workshop titles were 1“In-vehicle auditory interactions” and 2“Auditory displays in the car for supporting fuel efficient driving: About unobtrusiveness and gamification.”
Workshop Papers

“IN-VEHICLE AUDITORY INTERACTIONS”
SOUNDSCAPES FOR IN-VEHICLE TECHNOLOGIES

Katie Anna E. Wolf
Princeton University
Princeton, NJ, USA
kewolf@princeton.edu

Michael A. Nees
Lafayette College
Easton, PA, USA
neesmlafayette.edu

ABSTRACT

Auditory displays have been implemented with some success in a variety of in-vehicle technologies. Most of these applications have used discrete alerts and warnings, which present complications with false alarms and annoyance. Some research has suggested that on-going, continuous information displays may mitigate some of the problems with discrete alerts and warnings. We argue that well-designed naturalistic soundscapes offer a potentially useful means of continuous auditory display for in-vehicle technologies. We present examples of use scenarios and demonstrations of candidate sound manipulations for conveying information with soundscapes. We argue that soundscapes could be used in a variety of in-vehicle technologies, including displays of vehicle and driver conditions, collision and hazard warnings, infotainment systems, and automated systems.

1. INTRODUCTION

Auditory displays have been deployed with some success in various in-vehicle applications, most commonly in the form of auditory warnings and alerts (for a review, see [1]). Warnings and alerts are usually distinct, brief sounds that direct attention to a discrete system event that requires an immediate response (e.g., an impending collision).

A number of design complications emerge with discrete auditory alerts and warnings. Annoyance is a serious concern with auditory displays [1], [2]. Discrete alarms maybe especially problematic in this regard (see [3]). In high numbers, warnings can interfere with other audio communications and induce negative affect and stress [3][4]. Warnings are appropriate for infrequent events, as a proliferation of warnings, especially for non-critical events, results in annoyance and disuse—characterized by disabling or ignoring the warnings [5][6]. The primary source of overuse of alarms seems to be false alarms—alarms that are triggered frequently in the absence of a critical event that requires immediate action [7].

Consider, for example, a crash avoidance system that alerts the human operator to potential dangers. Given that crashes have extremely negative consequences, the designer may choose to err on the side of triggering alerts liberally at the hint of a possible collision rather than risk failing to alert the driver to a possible collision. We know, however, that the base rate of accidents is low, so most alerts will be false alarms (see [8]). Sorkin and Woods argued that false alarms result in less than optimal human monitoring strategies whereby the operator either (1) samples—or attends to—only some of alerts; or (2) attends to all alerts, but with reduced attention. Sorkin and Woods suggested providing the human operator continuous information about the system status rather than binary information triggered by a preset status criterion. Later research [9][10] found benefits of such approaches, dubbed likelihood alarm systems. As such, some of the shortcomings of discrete warnings triggered by system failures could be overcome with displays that provide on-going, information to the user [11].

The auditory modality is well-suited for presenting on-going peripheral displays that provide continuous information about the state of the system processes [12]. Data sonification, soundscapes, and ambient auditory displays are tenable candidate approaches to continuous auditory displays. Watson and Sanderson, for example, demonstrated the value of continuous sonification for monitoring variables relevant to anesthesiology during simultaneous secondary tasks [13], and Seagull et al. [14] showed that continuous auditory displays allowed for adequate monitoring performance. Though auditory monitoring performance was worse than conditions monitoring visual displays, the auditory monitoring condition allowed for considerably better performance on a concurrent visual tracking task that overshadowed the advantages of visual monitoring.

Annoyance, however, remains a major concern with continuous auditory displays. As Kramer [2] described, “…it is a familiar experience of people working in AD that a sonification will be running and it becomes sufficiently annoying that we just turn it off to take a break. Likewise, overly simple, intrusive, or simply unpleasant auditory computer interfaces are turned off, even if they have some utility” (pp. 52). This concern may be particularly relevant for in-vehicle technologies. It is unclear whether drivers and passengers would accept on-going sonifications, and this problem may increase to the extent that continuous auditory displays are perceived to be artificial and intrusive. Sanderson, Anderson, and Watson [15] presented a simple model that related possible system and operator states to outcomes with continuous auditory displays. Ideally, system states and operator states align to accomplish system goals in two ways: (1) the continuous auditory display is relegated to peripheral attention during normal or routine system operations, and this backgrounding (see [2]) of the auditory display allows the human operator to proceed with secondary cognitive tasks with little interference; and (2) the continuous auditory display attracts the focus of attention during abnormal system states; or (2) the auditory display attracts the focus of attention and interferes with other cognitive tasks during routine or normal system states. With continuous auditory displays, a delicate balance must be achieved such that the sounds are subtle enough to be relegated to peripheral attention, yet pronounced enough to call attention to themselves during critical events. To this end, continuous auditory displays that synthesize more naturalistic sounds warrant consideration for use for in-vehicle technologies.
2. THE CASE FOR SOUNDSCAPES

Mauney and Walker [16] presented a prototype system that represented stock data continuously by modulating changes in synthesized natural sounds such as rain, cicadas, and crickets. These artificial soundscapes were designed to display on-going feeds of data in the auditory periphery (also see [17]). Vickers et al. [18] supported soundscapes as a viable auditory interface for process monitoring and situational awareness of networks, primarily because they have the potential to communicate information while also being less fatiguing than other auditory interfaces. Work continues in this area toward developing models for translating data to customizable soundscapes [19], but perceptual and usability evaluations seem to have been limited to the brief initial interviews conducted by Mauney and Walker.

3. PROTOTYPE SOUNDSCAPES AND USE SCENARIOS FOR IN-VEHICLE TECHNOLOGIES

Similar to Wolf et al. [19], we conceptualize soundscapes as a hierarchy of sounds where each soundscape is broken into groups that contain all of the sound samples (i.e. recordings) from the same acoustical source. For instance, all of the recordings from a frog would be in the frog sound group. Additionally, they used terminology such as interval and instant sounds to represent sounds that are played continuously throughout a soundscape and those that occur at discrete intervals respectively. We will refer to these sounds as continuous and event-based. Using this terminology, we present design ideas and prototypes of natural soundscapes that could be used in the varying scenarios for in-vehicle technologies.

3.1. Alerts for Vehicle and Driver Conditions

When alerting a driver of vehicle and driver conditions (i.e. speed alerts, energy consumption, etc.), a system might only seek to inform the driver of an ongoing process, which may or may not require the driver to take immediate action (or any action at all). We can present the display in the form of a continuous soundscape that can be brought to the attention of the driver when they are interested in the information, but that can also fade into the background when not needed.

We present a number of soundscape examples that demonstrate the continuous passive conveyance of information. At a very basic level, a soundscape might include only a sound group or two to represent information. For example, changes in speed could be represented by changes in the gain of a recording (streamGainDemo.wav), changes in the number of layers of a recording to adjust the density of a sound (streamLayersDemo.wav), or changes in the playback rate of a recording (streamRateDemo.wav). Additionally, we can include multiple layers of a soundscape by adding sounds groups that do not convey information, but are there simply to create a more aesthetic appeal (cricketsWithOtherSounds.wav). On the other hand, these layers could be used as alerts to augment the information represented by the continuous sounds. In example (cricketsWithAlerts.wav), we use increasing gain of the stream to represent continuously updating data (e.g. speed), while specific events in the data (e.g. going above target speeds) trigger different event-based sound groups to play, such as the thunder, birds, and flies. When we include these alerts, we can randomize the selection of the sample used (if there are multiple samples from the same sound group), to allow for less predictability in the soundscape, making it potentially less fatiguing. These types of alerts differ from the discrete auditory alerts discussed in Section 1 as they fit into the context of the soundscape (assuming that the sound groups are taken from the same environmental location) and may therefore be less annoying.

The main goal with using soundscapes in this scenario is to provide information to the driver that they can passively observe without having their attention negatively drawn to the sounds in an obtrusive manner. It would be interesting to explore whether these sounds could positively affect the driver by creating a more relaxed driving environment.

3.2. Collision Avoidance and Hazard Warnings

Warnings used for collision avoidance and driving hazards typically indicate an unsafe driving scenario that requires the driver to take immediate action [1]. We can use the continuous feedback in the form of soundscapes to provide users with a passive, non-fatiguing way to monitor their situation, yet confront them with urgent information when the risk of a collision or a driving hazard becomes great.

We propose a few methods for presenting this urgent information using soundscapes. First, the sounds could gradually take on an unnatural aesthetic that we believe would bring the soundscape to the forefront of a user’s attention. In one of our examples (frogsHighRate.wav), the playback rate of a recording of a frog is adjusted higher than the typical range to produce an unnatural pitch. There are several other ways in which we can create these “unnatural” sounds in order to bring the soundscape to the users attention (e.g. applying filters and other modulations). In addition to unnatural sounds, we can introduce sounds that are out of place in that sound environment to indicate to the driver that conditions are changing. For instance, in a forest soundscape with a stream and crickets, the sound of a dolphin is out of place (cricketsWithDolphin.wav). One concern with this type of alert is similar to the concerns of the discrete auditory alarms presented in Section 1, mainly that they may be annoying and the user will turn them off.

Another way we can present urgent information using soundscapes is by adding more layers to the soundscape to give an indication that the risk of a hazard or collision has increased. Example (cricketsWithFrog.wav) demonstrates the addition of the frog sound on top of the crickets to indicate a change in the data. We can even have these added sounds be “unnatural”, as in example (cricketsWithUnnaturalFrog.wav).

In this scenario, the main focus is to provide clear information to the driver in moments of high urgency, while otherwise being unobtrusive and potentially relaxing, as in the previous scenario. While soundscapes have great potential for passive monitoring, there are many ways that we can use sound to intrude on that passivity and alert users that they need to take immediate action. It will still need to be investigated whether these manipulations and intrusions for this scenario work better than other sounds presented in previous work.

3.3. Infotainment

Similar to alerts for vehicle and driver conditions, alerts for various infotainment sources (e.g. email and social network) should allow for passive monitoring of information to allow
the driver to focus on the primary task of controlling the vehicle. Wolf et al. [19] have presented a prototype for sonifying data from the social network Twitter using soundscapes, with a particular focus on allowing the end-user to play a role in the design of the sonification. Their examples include having a particular event-based sound group represent a specific person (i.e. the nightingale bird chirp representing a tweet or an email from a particular colleague), or continuously monitoring a particular trend in the data. Examples for this scenario are very similar to those in Section 3.1. For instance, instead of monitoring the speed of the vehicle, we could monitor the volume of tweets for a particular trending topic.

This leads to another interesting application of soundscapes for in-vehicle technologies, which is that they can provide a way to encrypt and anonymize the data. The driver may know that the nightingale means that an email was received, but others in the vehicle may simply perceive it as an aspect of the soundscape.

3.4. Monitoring Automation Status

Finally, we suggest that soundscapes may be able to promote situation awareness during periods of automation in self-driving vehicles. Even during highly automated driving scenarios, the vehicle operator may need to resume manual control from the system in the event that automated systems falter (see, e.g., [20]). Norman [21] argued that difficulties with automation could be alleviated by “continual feedback about the state of the system, in a normal natural way” (pp. 8). Indeed, research has shown that continuous feedback about the status of the system results in better performance in resuming control from automation. Helfdin et al. [22] displayed a global measure of the autonomous driving system’s uncertainty using a visual meter. Performance metrics suggested that participants with knowledge about the automation status were better able to resume manual control of the vehicle and also better able to safely engage in secondary tasks (also see [23]). Seppelt and Lee [24] observed similar performance benefits when adaptive cruise control (ACC) systems—an example of partial automation—provided users with on-going information about the automation’s status. Further, people seem to be more forgiving and willing to continue to use imperfect automation if they are aware of the reasons for automation failures [25]. Research also has shown that human operators are more trusting and accepting of automation that provides them with feedback and information about the actions being executed by the automation [26]. As such, various combinations of the soundscape examples described above could be implemented to display data from sensors and algorithms in automated driving systems to keep the driver “in the loop” regarding the status of automated systems.

4. CHALLENGES WITH USING SOUNDSCAPE FOR IN-VEHICLE TECHNOLOGIES

While soundscapes offer a possibility for a continuous auditory display that could be useful for in-vehicle technologies, there are some unique aspects of soundscapes that should be considered in the development of these displays. As soundscapes consist of “real-world” sounds, there could be an intersection between the sounds in the soundscape and informative incidental sounds in the real world. For example, rainfall representing in-vehicle data could be difficult to perceive if it was actually raining outside. This is a difficulty of using soundscapes for any auditory display, and work by Wolf et al. [19] seeks to overcome these sorts of difficulties by allowing users to play a more direct role in the creation of auditory displays. Additionally, as with any in-vehicle auditory display, it is important that the sounds in the soundscape do not interfere with a user’s ability to detect potential problems with the car itself.

5. FUTURE DIRECTIONS AND CONCLUSIONS

Soundscapes may offer a useful approach to auditory display for in-vehicle technologies, as they may be able to overcome some of the limitations of discrete auditory alerts and warnings as well as other, less naturalistic forms of continuous sonification. More research is needed, however, as perceptual-cognitive and usability of evaluations of soundscapes have been limited to this point. Direct comparisons of soundscapes to other forms of display, including visual displays, discrete auditory displays, and other forms of continuous auditory display are warranted. Further, given the dearth of research on soundscapes, more research is needed to determine best practices in the design of soundscapes.

6. REFERENCES


All of the sound samples are from freesound.org. The full attribution is listed in the table below.

<table>
<thead>
<tr>
<th>Sound</th>
<th>Sound Title</th>
<th>User</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>stream</td>
<td>Distant stream 1</td>
<td>Adam_N</td>
<td><a href="https://www.freesound.org/people/Adam_N/">https://www.freesound.org/people/Adam_N/</a></td>
</tr>
<tr>
<td>nightingale</td>
<td>nightingale</td>
<td>viaaico2013</td>
<td><a href="https://www.freesound.org/people/viaaico2013/">https://www.freesound.org/people/viaaico2013/</a></td>
</tr>
<tr>
<td>frog</td>
<td>Frogs</td>
<td>juskiddink</td>
<td><a href="https://www.freesound.org/people/juskiddink/">https://www.freesound.org/people/juskiddink/</a></td>
</tr>
<tr>
<td>thunder</td>
<td>Distant Thunder</td>
<td>sean.townsend</td>
<td><a href="https://www.freesound.org/people/sean.townsend/">https://www.freesound.org/people/sean.townsend/</a></td>
</tr>
<tr>
<td>wolf</td>
<td>Wolf Howl</td>
<td>Gorgoroth6669</td>
<td><a href="https://www.freesound.org/people/Gorgoroth6669/">https://www.freesound.org/people/Gorgoroth6669/</a></td>
</tr>
<tr>
<td>owl</td>
<td>Tawny Owls 2</td>
<td>Benboncan</td>
<td><a href="https://www.freesound.org/people/Benboncan/">https://www.freesound.org/people/Benboncan/</a></td>
</tr>
<tr>
<td>fly</td>
<td>flys-l</td>
<td>galeku</td>
<td><a href="https://www.freesound.org/people/galeku/">https://www.freesound.org/people/galeku/</a></td>
</tr>
<tr>
<td>crickets</td>
<td>crickets</td>
<td>FreethinkerAnon</td>
<td><a href="https://www.freesound.org/people/FreethinkerAnon/">https://www.freesound.org/people/FreethinkerAnon/</a></td>
</tr>
</tbody>
</table>
INFLUENCE OF INATTENTIVELY PERCEIVED AUDITORY STIMULATION ON DRIVING PERFORMANCE

Khashayar Aghdæi

University of Applied Sciences Technikum Wien, Intelligent Transport Systems, Hochstraedtplatz 6, 1200 Vienna, Austria
Khashayar.aghdaei@technikum-wien.at

Andreas Rienner

Johannes Kepler University Linz, Institute for Pervasive Computing, Altenberger Str. 69, 4040 Linz, Austria
riener@pervasive.jku.at

ABSTRACT

Rapid technological advancements in recent years have brought up a lot of driving assistance systems of various types. While most of these systems, such as ABS, ESP, etc., automatically work in the background, there exist also a class of assistance systems that aim at directly supporting or influencing the driver in his different tasks, for example by trying to change its driving behavior. The primary objective of this work is to deeply investigate the potential of behavior changing auditory interfaces in different driving settings (speed limits at different levels, different driving scenarios & conditions, etc.). Particular attention is directed towards the power of auditory cues that are inattentively (or unconsciously) perceived by the driver, and that result in increased road safety or a more efficient style of driving (regarding economic measures such as fuel consumption or CO₂ emissions). Our expectations for the workshop, in order to come closer to our target, are to discuss the general feasibility of our approach/experimental setting and in more details questions like the types of sounds to use (i.e., artificial vs. natural), how to generate realistic engine sound/ambient noise, which type of music ((style; title/interpret)) to use to induce certain states of mood at the driver and last but not least to find partners for cooperation in the design of the auditory interface.

Keywords: Driving simulation, OpenDS, Auditory information, Engine sound & ambient noise, Inattentive perception.

1. INTRODUCTION

It has been observed that the private car is the place where a person frequently listens to music, and this source of audio has normally a positive influence on the driver (e.g., relaxation after a hard working day) [1]. However, the driver is most of the time also involved in other auditory tasks which are normally distractive (for example, making phone calls, communicating with passengers, paying attention to other sound information such as honking of other vehicles, traffic information in the radio, ambient noise, etc.). This in-turn increases the risk of getting involved in an incident due to an increased reaction or braking response time [2]. In addition, it is also projected that this kind of distraction has a negative effect on driving efficiency/economy (by losing anticipatory driving capabilities). These factors are, amongst a few others, the reason why auditory feedback (of critical driving situations) is very uncommon in vehicles today.

On the other hand, it is also known that perceived sound do has an impact on the way someone drives, and this circumstance should actually be utilized in this project for experimenting with music and other targeted sound cues to force the driver to change his/her driving behavior. The aim of this project is thus to provide evidence that sound/music has an effect on driving behavior (related to road safety or driving efficiency/economy). The behavior of driving is equated here to vehicle speed and movement trajectory, similarly as used by [1].

1.1. Road safety

In 2014, 47,670 people were injured in road traffic accidents in Austria and another 430 were killed [3]. This is only the seventh part compared to the 2,948 people killed in 1972 [4], but still these numbers are too high. The long term goal is to reach 0 traffic death (cf., “Vision zero” the European Commission’s commitment to a long-term visionary objective of abolishing road deaths and serious road traffic injuries [5]). To achieve this goal, the EU follows seven priority objectives: 1) education and training, 2) enforcement, 3) safer infrastructure, 4) safer vehicles, 5) use of modern technology, 6) emergency and post-injury services, and 7) the safety of vulnerable road users. The actual project can be seen as contributor to point 5) in the list above. By using specific pieces of music or sound cues, one of the major goals of this work is to contribute to a further reduction of traffic deaths or injuries. This could be achieved, for example, by using relaxing pieces of music (classics, etc.) to calm down a stressed or nervous driver [1] or aggressive music (rock, stimulating artificial music) to wake up a tired, exhausted driver [6]. It should be noted that music perception depends on the individual person and this fact needs to be considered in the design of the system, i.e., selection of music pieces to induce certain behavior. It is said that this type of information is perceived inattentively or unconsciously, as it does not contain any activity or task for the driver.

1.2. Driving economy

A second aspect to be looked at in this work is driving efficiency or economy. The economy of driving is to a large extent determined by personal driving preferences such as the general speed of traveling and the acceleration/braking behavior of a driver (=personal driving profile: in its simplest form dichotomous aggressive vs. conservative). Efficiency/economy of driving can be at least to some degree controlled with auditory stimuli. It is known, for example, that engine sound and ambient noise somehow “encode” driving speed (i.e., a particular engine sound and/or volume of ambient noise is implicitly connected to a certain speed of driv-
The association engine sound/ambient noise to driving speed might be used to implicitly “control” (or alter) the actual speed of driving. To give one concrete example: By “replaying” engine sound originating from a higher speed in the passenger cabin while driving at lower speed, the driver should be implicitly forced to slow down [7].

1.3. Summary
This project is investigating how specific pieces of music or artificial sound can be used to increase road safety and positively influence driving efficiency/economy. The aims are twofold: First, music or sound should be employed to change the mental state of a driver (calm down an aggressive driver or wake up a tired driver) and, thus, enhance on road safety. Second, artificially generated or prerecorded engine sound and ambient noise should be used to unconsciously influence drivers’ speed perception and, thus, let them (implicitly) adapt their driving speed.

The rest of this paper is structured as follows: The next section summarizes related work in the broader field of this work, section 3 explains in detail the methods and experimental setting used to test the impact of auditory information on driving behavior. Section 4 speculates about potential results of the study and specifies our expectations for the workshop, section 5 finally concludes the paper and gives some indications for the next steps in this project.

2. RELATED WORK
Engine noise or other acoustic feedback perceived by the driver inside the car is used for self-monitoring of driving performance and, according to [7], very important in the steady speed maintenance task. Drivers underestimated their driving speed to a greater extent when no acoustic feedback was present and tended to drive faster in this situation. Driving speed information derived from the engine noise was helpful when the drivers had to accurately keep a steady speed. In a recent breaking response study [2], authors found that phone conversation during driving had a higher negative impact on reaction time as compared to other potential interference factors like talking to the passengers. Age and gender are the two main substantial factors with respect to the effect on driving performance [8]. Furthermore, a driving study comparing hand held phones with hands free devices suggests that there is no difference in the reaction times [2], however, hand held mobile devices resulted in a reduced average driving speed, which is, according to the authors, caused by the attempts to compensate for the increased mentality demand. Using hands free sets, this compensation is not effective [9], [10].

Driving in difficult or tedious situations like as in congested traffic or waiting for a long time at the red light is said to increase anger and aggression and this can be tracked, for example, using cardiovascular measures like blood pressure (BP) or heart activity (ECG). Music or sound can be used to counteract. A study by Fairclough and colleagues [1] shows that low tempo music can reduce the systolic blood pressure regardless of the state of valance (pleasant or unpleasant, sad or happy). Lower blood pressure is an indicator for relation. On the other hand, fast tempo and preferred music led to an increase in the heart beat rate. This effect was used by [6] to prevent drowsiness in a longitudinal driving study. But it is not only music or sound perceived attentively that has the potential to change driver behavior or mental state; similar as shown by Riener et al. [11] with the haptic channel, it should be even possible to use unconsciously (or subliminally) perceived stimuli to change the style of driving or state of a driver.

3. METHODS AND SETTING
Lot of studies, as pointed out above, have substantiated that different types of music (e.g., low-tempo: classic vs. high-tempo: rock), employed in different driving situations (congested traffic, longitudinal driving study, etc.), can influence the driver’s mental state and, in succession, show also effects on the style of driving. The actual project builds-up on this previous knowledge and defines several settings for user or driver testing. Therefore, we are using a modified OpenDS driving simulator to run different variants of a lane change task (LCT) experiment with added speed regions as shown in Figure 1. One focus of a scientific experiment will be the impact of natural vs. artificial engine sound and ambient noise on driving, in another setting the effect of music (relaxing vs. aggressive) on speed-keeping will be investigated. More settings will be explained below.

Figure 1: Schematic of the lane change task with different speed regions added.

3.1. Research approach and research questions
A number of research hypotheses should be investigated in the frame of this project.

H1: Presence of engine sound and ambient noise results in a better estimation of driving speed.

H2: The type of music (relaxing vs. aggressive) a driver is exposed to has an impact on the average speed of driving and the lane keeping precision.

H3: Increasing/decreasing beats per minute (BPM) of music pieces replayed in the passenger compartment result in higher/lower average driving speed.

In order to be able to provide answers to these hypotheses, the following research questions should be addressed:

1. What is the influence of artificially generated vs. natural engine sound on the speed keeping behavior?
2. Which type of music (style; title/interpret) to use to induce certain states of mood at the driver?
3. How can different types of music (classic/pop/folk/rock) be employed to specifically influence the speed keeping behavior?
4. Is there any relation between the characteristic ‘beats per minute’ (BPM) of music pieces and the estimated speed of driving?
5. What is the influence of dual tasks (such as talking on the mobile phone) while driving on the speed keeping behavior?
6. What is the difference of preferred/disliked music on driving performance?

3.2. Experimental setting and procedure

From a rather generic point of view, the experimental setting to be used in our studies can be described as follows. We will execute different types of simulated driving experiments implemented on top of the OpenDS driving simulator (v3.0). The basic setting is a lane change test (LCT) according to ISO 26022:2010 (3 lanes, constant driving speed of 60km/h) [12]. The standard LCT is modified for this project to include regions with different speed limits (see Figure 1). As we can see in this figure, the overhead signs are continuously shown to the driver and displaying a clearly visible “(x) (x) (x)” pattern on red ground while far away from the sign and change into a specific lane change request pattern (“check mark” on green ground “(✓)”) corresponding to the lane to change to) when approaching the overhead sign. In addition, speed signs (an extension to the original LCT) are presented to the driver at the same time as the lane change request pattern.

Figure 2 below shows the normative path of the driven route (represented by the bold line between the roads) while the thin line indicates the deviation from the normative path of one driver. Calculating the total area between the these two lines in relation to the length of the track results in an error value (deviation error) that covers important aspects of the driving performance of a driver, such as the perception of and reaction to the signs, lane and speed keeping quality, etc. By using an appropriate statistical test (like the t-test) we can finally relate MDEV [12] to our research hypotheses.

The simulated driving scene is presented to the test subjects with a dual-beamer setting. One beamer projects the OpenDS simulator scene (configured to a close-to-realistic size), the other acts as a kind-of “head-up display” (HUD) to show (dependent on the experimental condition) additional information to the driver (speedometer, dual task, etc.). The vehicle in the simulator is controlled by a Logitech G27 racing wheel (including a force-feedback steering wheel, acceleration and braking pedals, switching set to automatic transmission). Audio feedback is delivered to the driver with either stereo speakers placed in front of the beamer wall or via high-quality headphones. We will further use a mobile hand held device to simulate the distortions due to the communication while driving.

3.3. Experimental conditions

In this environment, the central task is to drive as close as possible to a given speed limit, comparing the following different experimental conditions to each other. Going for a full factorial design (FFD), this would give us a total of 2x2x4x2=48 different conditions to evaluate – we will later select a subset to be used for testing:

- Condition 1: Speedometer in the dashboard shown/hidden,
- Condition 2: Engine sound and ambient noise present/absent,
- Condition 3: Different types of loud music predominating (masking) engine sound (off/classic/pop/hardrock),
- Condition 4: Additional dual task (talking on the phone while driving) on/off.

One example of a setting could be condition 1 (dashboard): hidden, condition 2 (engine sound): present, condition 3 (music): off, condition 4 (dual task): off. In this case, the driver does not see a speedometer but has to adapt to different speed limits as shown in the simulator. It is expected that engine noise (after some training) is a useful source of information to reach and retain a certain speed of driving.

3.4. Selection of participants

According to recent statistics, 73% of all people who died in road accidents in Austria were male (in the age group 16 to 29 years, even 87% of people died were male [11]). Following these numbers, persons invited to our user study have to belong to the group of young, male drivers. We further limit the age range to 20-30 years (to allow for at least 2 years of driving experience; driving license issued in Austria at the age of 18). We haven’t decided yet to go for a within- or between-groups setting, and consequently we haven’t fixed the number of participants required to achieve statistically significant results (the final number of experimental conditions to go for will also influence this decision).

3.5. Course of the experiment

We will start by handing out a briefing document followed by a warm-up session on the simulator to allow subjects to get familiar with it. After the main experiment (as outlined before), every participant has to answer both an individual (demographics data, experiment-specific qualitative data) and a standardized questionnaire. We will use here the NASA TLX to assess the subjective mental workload perceived during the experiment. NASA TLX

Figure 2: Lane keeping performance: Deviation of one driver (thin line) from the optimal track (bold line) for the modified lane change task (LCT).
requires participants to rate their perceived levels of mental, physical and time demands as well as their effort, performance and frustration during that task on a 20 level Likert scale. The effect of driving performance on NASA TLX scores will be analyzed with separate independent sample t-tests for each of the 6 items as suggested by [13]. According to the final setting of the study, parts of the questionnaires might be repeated between different tasks (post-task tests). Finally, test subjects are debriefed and dismissed.

3.6. Our expectations on the workshop

Our expectations for the workshop are to discuss with the other participants about

1. The general feasibility of our approach/experimental setting,
2. The types of sounds (engine sound/ambient noise) to use (i.e., artificial vs. natural)
3. The generation of realistic sound sources (computer generated, recordings in the real),
4. Distractions caused by the use of mobile devices,
5. The type of music (low-tempo: classic /high-tempo: hardrock) to use to induce certain states of mood at the driver.

Last but not least, we also hope to find partners cooperating with us in the design of the auditory interface.

4. CONCLUSION

The primary objective of this project is to investigate the potential of behavior changing auditory interfaces in different driving settings (speed limits at different levels, different driving scenarios & conditions, etc.). Particular attention is directed towards the power of auditory cues that are inattentively (or subliminally) perceived by the driver, and that have potential to 1) increase road safety or 2) result in a more economic/efficient style of driving.

As such, the actual project can be seen as a contributor to the European Commissions “Vision Zero” – a commitment to a long-term visionary objective of abolishing road deaths and serious road traffic injuries [5].

5. REFERENCES


TARGET MATCHING AUDITORY DISPLAYS TO SUPPORT IMPROVED DRIVER SKILLS AND VEHICLE COORDINATION

Nicholas Powell
School of Engineering & Applied Science, Aston University, Birmingham, UK, B4 7ET
powelln@aston.ac.uk

Jo Lumsden
School of Engineering & Applied Science, Aston University, Birmingham, UK, B4 7ET
j.lumsden@aston.ac.uk

We have developed a novel real-time auditory display system for accelerated racing driver skills acquisition in which auditory feedback provides drivers with concurrent sensory augmentation and performance feedback using a novel target matching design. In essence, real-time, dynamic, tonal audio feedback representing lateral G-force (a proxy for tire slip) is delivered to one ear whilst a target lateral G-force value representing the ‘limit’ of the car, to which the driver aims to drive, is panned to the driver’s other ear; tonal match across both ears signifies that the ‘limit’ has been reached [1]. We believe that the principle of target matching auditory displays also offers potential improvement within the mainstream driver domain, wherein providing auditory representations of both target attainment and current performance has the capacity to support optimized driving. We propose that potential applications of such an auditory feedback model might include, but not be limited to, vehicle performance (e.g., speed) or navigation (e.g., compass heading) within commercial vehicles. We further propose that a modified version of our motorsport-optimized feedback system could be integrated within high performance road vehicles to allow drivers to adapt to the nuances of a new vehicle effectively and to allow such drivers to maximize the fun of driving such vehicles but safely and legally, without exceeding the limits of the car as a consequence of inexperience or misjudgment.

Speed is often a factor in road traffic accidents [2]. Traffic calming and speed reduction measures such as variable speed limits and speed advisory warnings in accident black spots have all been introduced in an attempt to reduce speed-related accidents, yet they are only as effective as drivers’ willingness and capacity to effectively observe them [3]. The same is true for speed recommendations intended to regulate traffic flow and reduce the environmental impact of congestion. We suggest that an audio target-matching in-vehicle display could simultaneously advise drivers on the active speed limit at any given point in time (where he is exceeding the limit set) and his current driving speed, encouraging him to reduce his speed until tonal match is achieved. We believe this would support a much wider scope for active traffic management and help drivers appreciate the current speed limit and their real-time conformance without the need to repeatedly refer to speedometers and infrequent road signs. Such recommendation-based designs may be accepted due to their capacity to improve safety and fuel economy but we recommend that drivers should be able to customize recommendations to ensure that such auditory displays are not considered either intrusive or obtrusive.

In a similar vein, we propose that target matching-based audio feedback could be used effectively to set distance and directional targets to which drivers could ‘aim’ when navigating; we believe this would reduce drivers’ reliance on visual sat-nav displays and their irregular voice commands.

Evaluation of in-vehicle auditory displays for delivery of feedback about potentially critical performance data for mass consumption will be a complex undertaking. It will be necessary to validate the conceptual mapping of feedback to corrective actions, the timeliness with which both the feedback and user response to the feedback is affected, and to determine the intuitiveness and usability of the feedback design as well as to weigh this against the risk of erroneous interpretation. Displays with an anticipated single polarity (such as speed reduction) may be more readily accepted.

Despite our anticipation of great potential for audio target-matching displays within vehicles, there are also a range of issues to overcome in the delivery of such displays. Delivering such a display depends on reliable and cost effective sensing methods and supportive legislation. Would smart driver feedback systems such as those we have proposed here fall into the same category of manufacturer liability as autonomous commercial vehicles, or would they be provided without safety guarantees, simply to assist drivers where possible? This raises a potentially contentious issue: does the potential to save many lives outweigh the risk of manufacturer liability should a failure occur?

Another similarly contentious issue related to the provision of displays such as those we have proposed relates to user autonomy and enforcement; if real time recommendations of the nature we suggest become common place, pressure to enforce adherence may follow. While we offer no opinion on this issue at this juncture, we expect this to become a polarizing debate. Existing telematics systems typically only transmit averages of driving data for risk ratings rather than enforcement and this compromise appears to be acceptable to drivers who retain their agency while benefitting from reduced risk and premiums. Perhaps this model will also hold for our proposed audio feedback protocols, which have the capacity to not just reduce risk, but to enhance driver performance and even enjoyment.

REFERENCES


AppliYing Spindex and Spearcon Auditory Cues for List Navigation in the Vehicle Context: A Review

Thomas M. Gable  
Georgia Institute of Technology,  
Sonification Lab,  
654 Cherry Street, Atlanta, Georgia 30332, USA  
thomas.gable@gatech.edu

Bruce N. Walker  
Georgia Institute of Technology,  
Sonification Lab,  
654 Cherry Street, Atlanta, Georgia 30332, USA  
bruce.walker@psych.gatech.edu

Abstract
Navigating lists is a common task performed while driving, either on interfaces native devices, those built into the car, or nomadic devices, those brought into the vehicle by the driver. This action of navigating a menu on either type of device can be very visually demanding and researchers have attempted to decrease the distraction this can cause from driving through the use of auditory menus. However, the use of basic auditory menus for menu navigation brings about a number of issues and therefore advanced auditory cues (i.e. spearcon and spindex cues) have been developed and tested within the driving context with promising results. The following paper summarizes the reasoning behind the use of spearcon and spindex auditory cues, reviews the research done in the driving context with the cues, and discusses practical applications of the use and ongoing research regarding the cues.

1. Introduction
Navigating lists is a common task on many electronic devices. Whether it is a short list of options in a main menu or a long alphabetical list such as songs or contacts, the action of navigating a list can be necessary when dealing with modern technologies. These visual menus are designed to be primary tasks when used, not in tandem with other visual tasks. However, most electronics we use require interaction with lists in order to perform tasks we need to complete, and electronic devices are being increasingly used in the vehicle. This presents an issue in cases where the user is engaged in a visually demanding task, while also attempting to complete an interaction with an electronic device. In order to address this issue of two competing visual tasks Nees and Walker [1] suggest the application of auditory cues.

2. Auditory Menus

2.1. Theoretical support
The use of auditory information to help guide people through menus has been done in many contexts, including driving. Through the application of auditory information drivers can interact with a multimodal secondary display instead of simply visual information that competes directly with their primary driving task. Multiple resources theory (MRT), a model of human information processing, supports the use of auditory information in these instances as it predicts a decrease in total workload with this addition of sound to enhance or use in place of a visuals display while completing a dual task situation [2]. This prediction of lower workload is made in MRT through the distinction of separate resources that are available for each modality of information that the human brain processes. These limited amounts of resources for each modality can be overtaxed within each modality, which are processed at the same time, creating overwhelming workload and decreased task performance. Alternatively, by spreading out task loads over multiple modalities less demand of the available resources over those modalities is required and therefore workload is lower. This means that a driver could perform the visual task of driving with all of their visual resources while interacting with an auditory menu system with their auditory resources more easily than doing both tasks at the same time with visuals alone. It should be noted that while MRT supports this approach to using auditory menus it is not the only theory of information processing in the literature. By using the theory one must also accept the other assumptions that MRT brings forth and these should all be considered when applying the theory to guide design of interfaces (See Nees and Walker [1] for a more in depth discussion of these theoretical considerations).

2.2. Related research
This theoretical support for the use of auditory menus and the drive to make interacting with menus while driving a safer task has led many researchers to investigate the use of them within the context of the vehicle. Results of many studies investigating drivers’ abilities to use menu-based interfaces within the car show the need for these auditory menu interfaces through significant negative impacts on driving related performance measures when using visual-only interfaces such as decreased driving performance [3], increased dwell time off the driving task [4, 5], higher subjective workload [5, 6], and decreased hazard detection performance than when only driving [4, 7]. However, when participants used auditory menu interfaces in these studies researchers would see improvements in these measures as compared to the visuals-only data including increased driving performance [7, 8], increased dwell time on the driving task [4, 5], lower workload [5, 6] and better hazard detection performance [7]. While these increases in performance related to driving support the use of auditory menus in the driving context there is one caveat. In most of these studies the type of auditory menu feedback applied was that of speech or text-to-speech (TTS) cues. Although these cues do not take any time to learn and are therefore easy for novice users, these types of cues can take a long time to receive feedback from the interface. Some of these research studies reflected these issues in their results through the significantly longer time it took...
participants to complete the tasks when using the auditory menus than the visuals-only interactions [8, 9]. While this may seem like an acceptable tradeoff upon initial glance this longer completion time has been found to decrease the use of auditory menu interfaces when the completion of the secondary task is emphasized and participants are given a choice of what interface to use [10]. To address the issue of increased interaction times due to the slow nature of TTS interfaces and ensure users will not decrease their use of these promising interfaces a number of solutions have been put forward, one of which has been spindex and spearcon cues.

3. SPINDEX AND SPEARCON CUES

3.1. Introducing spindex and spearcon cues

Spindex and spearcon auditory cues are both types of non-speech cues used to enhance TTS. While other types of non-speech audio cues exist (see [1] for an in depth discussion) and also attempt to solve the issue of slow feedback, spindex and spearcons are based on speech, making them easier to learn than other types of non-speech cues. Spindex (i.e., speech index) cues are short sounds based on the pronunciation of the first letter of a menu item [11]. While also a very short auditory cue based on speech, a spearcon (i.e., speech earcon) is a phrase or word that has been speed up to the point where it is no longer recognizable as a word but simply as a sound [12]. Spindex cues are particularly useful in alphabetical lists, while spearcons are often well suited for well-known menus, since when using the cues you actually learn what the short sound represents. Both cues were developed for application either by themselves or with each other to enhance TTS, although they could also be used independently if desired. Most importantly though, the cues were developed to be interruptible, so that users can move through lists very quickly, only slowing down to listen longer as they get closer to their target items and need to be more accurate.

3.2. Research findings

Research behind the use of advanced auditory cues has been very promising. Outside of the vehicle context the cues have been investigated within a mobile phone song list search task via kinetic flicking, wheeling and tapping. The results of this study showed that participants were significantly faster at the search task and had lower subjective workload when using the spindex or spearcon auditory cues than when using the visuals only for the same input methods [13]. Applied within a driving context results point in similar directions. When interacting with an infotainment unit to perform a long list searching task and a driving-like task, participants were found to have significantly better primary task performance and lower subjective workload when using the spindex or spearcon cues, either with TTS or in conjunction together with TTS, than a visuals-only interface [14]. In addition results showed that in both the spindex plus TTS interaction, and the combined spindex, spearcon, and TTS condition, participants were faster at selecting the menu items than the visuals only interaction, and that participants preferred these two auditory cue combinations over the visuals-only condition [14]. Supporting results were found in a study on a mid-fidelity simulator [6]. A decrease in the standard deviation of steering wheel angle was seen when participants used any of the variations of the spindex or spearcon cues (paired with TTS or with each other and TTS) as compared to a visuals-only interaction, hinting at better driving performance. Faster selection time within the spindex plus TTS interaction as compared to visuals-only was also found, as was lower total subjective workload when using either the spindex plus TTS or spearcon plus TTS cue types as compared to the visuals only interactions [6]. Both of these studies also showed that participants had significant increases in search task performance between two time blocks of interaction, showing they were faster as they gained experience [14]. These studies made large contributions to understanding the performance effects of using spindex and spearcon cues, however, they did not confirm if the use of the cues was also allowing drivers to keep their eyes on the road. To answer this question another study was done, focusing on the measurement of eye movements during interaction. The study had participants use the spindex and spearcon auditory cues as well as in the visuals-only interaction while searching the same type of lists in previous studies, but on a mobile phone [5]. Results of the study found that when participants used the spindex plus TTS cues they had significantly higher percent time eyes on the road than when they use the visuals-only interaction method, confirming that the spindex cues were increasing participants’ time looking at the road when searching long lists [5].

4. DISCUSSION

4.1. Ongoing research

This previous work into spindex and spearcon auditory cues shows the positive impact that they can have in a menu-searching context while performing a demanding primary task such as driving. As can be seen in some of the later studies, the use of spindex cues may be better for novices on long lists than spearcons, and spearcons possibly used for shorter menus as they were designed for. However, it is the opinions of the authors that with enough experience the use of spearcons could also become as fast as spindex for long lists. In discussing the impact of practice it should be noted that one of the next stages of research should be the investigation of extended practice with these types of auditory cues. As stated earlier initial results after enough experience showed significant improvements in search task time when using the spindex and spearcon cues and the authors would expect this to be even greater with extended practice. This much cannot be said for the use of visuals-only interfaces as research has found that many of the negative effects of using visual interfaces for menu navigation while driving do not seem to go away with practice [4]. Another factor that should be considered in future studies is that of situation awareness and the impact of the cues on hazard detection. While research has found that participants have increased time with their eyes on the road when using spindex cues as compared to visuals-only interactions this does not necessarily mean this difference actually impacts driver safety. Through measuring situation awareness and hazard detection when participants are using the different
types of cues researchers could determine the possible driving safety impacts of using the auditory cues within a driving context.

4.2. Implications

The application of spindex and spearcon auditory cues show that their use could help to decrease negative impacts of menu navigation within the vehicle and other contexts where a visually demanding primary task is being undertaken. Future work will continue to explore the impacts of these cues, both within the menu navigation space and in other circumstances. For example, these types of cues could also be applied to give drivers other information that could be necessary while driving. In particular, the use of spearcons within in-vehicle displays could allow designers to give drivers more informative alerts, warnings, or other types of feedback. The use of these cues should continue to be explored in the vehicle context as they have much to offer.

5. ACKNOWLEDGMENTS

Portions of this work are supported by a National Science Foundation Graduate Research Fellowship (DGE-1148903).

6. REFERENCES


IN-CAR ACTIVE SOUND GENERATION FOR ENHANCED FEEDBACK IN VEHICLES WITH COMBUSTION ENGINES OR ELECTRIC ENGINES

Alois Sontacchi, Matthias Frank, and Robert Höldrich

Institute of Electronic Music and Acoustics
University of Music and Performing Arts
8010 Graz, Austria
sontacchi@iem.at

ABSTRACT

This contribution presents active sound generation (ASG) for interior sound enhancement to assist or improve sound feedback in either down-sized combustion engines or electric engines. It reports evaluation results from two studies about the description of engine sounds and the influence of sound feedback on driving behavior. Preliminary results indicate gender-related differences and we draw the conclusion that various possible differences in sound quality assessment depending on gender, age, habit, and cultural background are worth reviewing in future research and require further discussion.

1. INTRODUCTION

Electric cars do not generate as much sound as vehicles with conventional combustion engines and thus set the stage for quieter streets. However, the low exterior sound level yields potential risks for other traffic participants due of the reduced detection distance [1]. For this reason, legal administrations regulated the requirements for the exterior sound of hybrid and electric cars, e.g. minimum sound power [2], and proposes Active Sound Generation (ASG). Furthermore, they prohibit ASG that could be confused with bird’s twittering or sirens [1]. Still, there seems to be a wide range of possibilities for exterior sound design that is consistent with legal regulations and we think that manufacturers and customers of premium and sports cars will not give up manufacturer-individual sound characteristics. This holds true for the interior sound. Although most customers prefer quiet cars, load-dependent loudness adaptation is an important factor for the acceptance of interior sound [3] and for driving feedback. However, preference ratings for vehicle sounds revealed intercultural differences depending on the specific spectral balance [4]. Studies about interior car sounds from combustion engines and electric engines (without ASG) describe the latter as barely sporty [5]. Sportiness in electric cars appears to require ASG, where perceived acceleration can even be increased by pitch shifting [5]. What is more, ASG can be applied for down-sized combustion engines that are lacking the feedback from the familiar interior sound due to the reduced number of cylinders [7, 8].

This work was supported by the projects AAP (http://www.comet-aap.at/) and ASD (http://comet-asd.at/), which are funded by Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and the departments 3 and 14 of the Styrian Government. The Austrian Research Promotion Agency (FFG) conducted the funding under the Competence Centers for Excellent Technologies (COMET, K-Project), a program of the above-mentioned institutions.

In this contribution, we present possibilities to learn from the interior sound of traditional vehicles with combustion engines and how to apply this knowledge in order to retain or enhance sonic feedback capabilities of vehicles with either down-sized combustion engines or electric engines. Gender-related aspects of this feedback are discussed based on previous research studies within our two large-scale collaborative research projects AAP (advanced audio processing) and ASD (acoustic sensing & design, ongoing). The paper is arranged as follows: Section 2 introduces the analysis procedure to capture the deterministic parts of an combustion engine sound. Subsequently, section 3 presents three different re-synthesis approaches. In section 4, existing descriptive sound parameters are discussed and first observable gender differences are stressed. Section 5 revisits a research study of a down-sized engine car equipped with an connectible sound enhancement system. Differences in driving behaviors between male and female subjects are revealed. Finally, the paper is summarized and a discussion on feedback capabilities of ASG in a wider sense is suggested.

2. ANALYSIS OF ENGINE SOUNDS

This section presents an analysis system that allows for a complete or arbitrarily precise recreation of combustion engine sounds [9]. The whole system is based on engine orders (multiples of the engine speed in rpm). For each engine order, it employs parameter sets consisting on rpm-dependent amplitude, amplitude modulation, and phase modulation. A full description of an engine sound consists of three parameter sets with 128 equally spaced sampling points along the whole rpm-range: full-load, proportional load, and recuperation. Depending of the engine load, the system interpolates between the three parameter sets. The parametric approach is a prerequisite for fast responding ASG. Real-time parameter changes are controlled manually during the sound design process or via standardized protocols in the final application. The adaptation to future protocols is easily possible. Microphone recordings are analyzed using second generation Vold-Kalman filters [10, 11]. These filters are designed to track signals of a known structure among residual signals of a different structure. In our case, we know that the engine sound comprises harmonics of the engine’s firing frequency, i.e. engine orders. The filtering process yields a time-domain bandpass signal for each tracked order individually. For each bandpass signal, amplitude, amplitude modulation, and phase modulation are calculated. The order-based analysis considers only the deterministic signal components and thus reduces disturbing noise that is caused by other sound sources, e.g. wind and tread noise.
3. SYNTHESIS OF ENGINE SOUNDS

The following synthesis approaches are implemented in the visual programming language Pure Data (pd). As pd is open-source, it can be compiled for arbitrary computer platforms, such as low-power computers, providing fast adaptation to customer needs. Exemplarily, we have successfully tested the synthesizers on a Raspberry Pi.

3.1. Additive Sinusoidal Synthesis

The synthesizer provides up to 128 oscillators simultaneously and is thus able to synthesize up to the 32nd engine order in steps of quarter orders. Each of the oscillators has exactly the same structure and is parametrized by frequency, gain, amplitude modulation, and phase modulation. Exemplarily, Fig. 1 shows the spectrogram of a full-load run-up recorded in the car interior. The synthesis employs only full and half engine orders. The difference between the original recording and its re-synthesis is limited to the low-frequency rolling noise of the tires on the dynamometer. As this difference includes only the actually unwanted part of the recording, the clear synthesis of the engine sound succeeded.

3.2. Subtractive Filtering of Textures

Replacing the oscillators from the above-mentioned sinusoidal approach by narrowband filters allows for the application of arbitrary sound textures as base material. These sound textures can be any recorded or generated broadband sounds. Using white noise as texture results in a similar, however livelier synthesis compared to the sinusoidal approach. More sophisticatedly, and in a controlled manner, textures can be created by condensing audio material [12], such as natural ambient sounds, instruments, or even entire musical pieces.

3.3. Harmonic Textures

Sound textures can be created which directly contain the analyzed harmonic structure or parts of it. Each of the three defined load conditions (full-load, proportional load, and recuperation) of the engine is represented by its own texture. Intermediate load conditions are handled by weighted mixing of these textures. Real-time dependency on engine speed/vehicle velocity is achieved by pitch shifting with adjustable slope.

4. VERBAL DESCRIPTION OF ENGINE SOUNDS

In addition to the order-based analysis and in order to design sounds for different car classes, we examined a database of 261 cars from AVL-List GmbH in Graz. This database contained not only the interior sound recordings but also 29 synchronously captured descriptive parameters, such as loudness, roughness, and articulation index.

Principle component analysis of the 29-dimensional parameter space including full-load run-ups of all 261 cars yields a 2-dimensional subspace that already explains 93\% of the variance, cf. Fig. 2. All cars exhibit a distinct transition from luxury to sport for increasing rpm. This allows for linear interpolation over rpm at a given state. Moreover, car classes can be easily distinguished by their offset and range.

A preliminary perceptual validation (9 cars within a defined rpm range that sample the principal components as equal as possible) confirms the ranking along the luxury-sport-axis. Applying the repertory grid technique [13, 14], the subjects were presented triplets of stimuli (excerpts of interior full-load run-up recordings). Within each triplet, the subjects had to identify the most different stimulus and verbally describe this difference. Subsequently, each subject rated all 9 stimuli on her/his individual attribute scales. The individual descriptions exhibit a high degree of similarity and agree with the objective parameters. However, a detailed investigation reveals gender-dependent descriptions: Only male subjects named positive energy-related terms, such as powerful, sporty, whereas female subjects focused on manifold terms for sound coloration, such as low bass level, tonality.
5. EVALUATION OF SOUND ENHANCEMENT FOR DOWN-SIZED ENGINES

In [8], we evaluated the effect of sound enhancement for down-sized engines in a blind test scenario using the additive sinusoidal synthesis approach. Ten people drove a test track once with the original two-cylinder sound and on a different date with the four-cylinder sound enhancement without being informed about the changes made to the car.

The subjects found that the tip-in and tip-out responded better in the car with the four-cylinder sound and they perceived the sound as better. The recorded engine speeds also show that the subjects shifted into the 4th gear at a lower rpm with the four-cylinder sound. For the other gears, there are no significant differences between the engine speeds because the first three gears are mainly used in city traffic, where gear shifting depends on outer circumstances and the 5th gear has hardly been used. The shift into the 4th gear however mainly depends on personal decisions. The result of the test thus supports the hypothesis that the sound of the car is a major cue for the gear shift. Interestingly, the driving behavior of female subjects tended to be less influenced by the sound modification. However, they recognized the modified interior sound and assessed it similarly positive as male subjects did.

6. CONCLUSION AND OUTLOOK

Based on the existing regulations, it is obvious that electric cars must be equipped with ASG for exterior noise at low driving speed. Still there is a wide range of possibilities for sound design that is consistent with these regulations and goes beyond mere vehicle alert signals. We suggest the application of familiar characteristics known from combustions engines in order to maintain the tradition of manufacturer-individual and utilization-fitted sounds. This does not necessarily mean to simulate the existing sounds of combustion engines, it rather provides the possibility to gradually draw on already established experiences. For example, the harmonic structure of a Porsche 6-cylinder boxer engine is a unique sound character that could be transferred to future Porsche electric cars. Beyond that, we suggest ASG to design and enhance the interior noise in order to assist drivability and further improve experience of vehicles with down-sized engines or electric engines. The design of interior and exterior noise can both benefit from the experience and knowledge of combustion engine sound. Gradually abstractions can be achieved using 3 different synthesis approaches that are based on the analysis of combustion engine sounds: The sinusoidal approach provides an accurate re-synthesis. Replacing these oscillators by narrow-band filters allows for the application of arbitrary sound textures as base material. Alternatively, these sound textures can carry the harmonic structure themselves and rpm-dependent pitch shifting can be applied. Further effects, such as loudness increase and amplitude modulation, can be added to evoke specific impressions, e.g. sportiness and acceleration.

The results of our studies revealed gender-related differences in the description of engine sounds and the effect on the driving behavior. Thus, we suggest that feedback capabilities of ASG should be discussed on a wide basis including gender, age, habit, and cultural background.

7. ACKNOWLEDGMENT

The authors thank AVL-List GmbH for providing the VOICE database [15].

8. REFERENCES


USE OF AUDITORY INTERFACES FOR TAKEOVER REQUESTS IN HIGHLY AUTOMATED DRIVING: A PROPOSED DRIVING SIMULATOR STUDY

Pavlo Bazilinskyy  
Delft University of Technology,  
Department BioMechanical Engineering,  
Delft, The Netherlands  
P.Bazilinskyy@tudelft.nl

Bastiaan Petermeijer  
Technische Universität München,  
Lehrstuhl für Ergonomie,  
Garching, Germany  
petermeijer@ife.mw.tum.de

Joost de Winter  
Delft University of Technology,  
Department BioMechanical Engineering,  
Delft, The Netherlands  
J.C.F.deWinter@tudelft.nl

ABSTRACT

Highly automated driving can potentially provide enormous benefits to society. However, it is unclear what types of interfaces should be used for takeover requests during highly automated driving, in which a driver is asked to switch back to manual driving. In this paper, a proposal for a driving simulator study on the use of six auditory signals during such takeover requests is outlined. The auditory signals to be tested in the experiment are based on the results of an online international survey previously conducted by the authors. The experiment will involve 24 participants performing a secondary task, and the takeover scenario will be represented by an accident in the middle lane of a three-lane freeway. The time margin prior to takeover will be 7 s. The driving time between subsequent takeover requests will be 2 to 3 min. The application of the results of the proposed study as well as plans for future studies are presented in the last section.

1. INTRODUCTION

There are five levels of automation for on-road vehicles: (1) manual driving, (2) driver assistance, (3) partially automated driving, (4) highly automated driving, and (5) fully automated driving [1]. In the 1990s, Adaptive Cruise Control (ACC), a technology that controls the longitudinal motion of a vehicle, made driver assistance a reality. Further advancements introduced in the 2000s laid ground for partially automated driving, where drivers are no longer required to manually control the lateral movements of the car but still have to keep their eyes focused on the road and/or occasionally touch the steering wheel. The current focus of the scientific community is directed at the fourth level, highly automated vehicles, where the driver no longer needs to keep his/her attention to the road and can remove the hands from the steering wheel. However, even in highly automated driving (HAD) the vehicle cannot control all situations, and the driver may be asked to take back manual control by means of a so-called takeover request (TOR). The time between issuing a TOR and the required moment of transition of control is a critical variable in the design of such automated driving systems [2], [3]. Fully automated driving (FAD) is envisioned to be the final iteration of automated driving, where the vehicle will control the entire task of driving.

We have previously conducted a study that analyzed anonymous textual comments regarding fully automated driving extracted from three online surveys with 8,862 respondents from 112 countries (males represented 74% of the sample, mean age of the participants was 32.6 years) [4]. A crowdsourcing task was created and 69 workers were requested to assign each of 1,952 comments to at least one of 12 predefined categories: positive and negative attitude to automated driving, enjoyment in manual driving, concerns about trust, reliability of software, and readiness of road infrastructure. The public opinion was found to be heterogeneous. A positive attitude towards automated driving was identified in 29% of 1,050 meaningful comments, whereas 18% of the comments were classified as ‘negative attitude towards automated driving’.

1.1. Auditory interfaces for takeover requests in highly automated driving

Our proposed research study investigates the potential of auditory feedback in automated driving. Present day cars often come with Advanced Driver Assistance Systems (ADAS), which offer assistance with driving and monitor the environment for the detection of road safety risks. Most of such systems can provide auditory warnings to the driver.

We have previously conducted an international online-based survey to investigate the opinion of 2,000 participants from 96 countries on the usage of auditory interfaces in modern and future vehicles [5]. The participants reported their attitudes towards two existing auditory driver assistance systems, a parking assistant (PA) and a forward-collision warning system (FCWS), as well as towards a futuristic augmented sound system (FS) intended for fully automated driving. The respondents were generally positive towards the PA and FCWS. The willingness to use them was rated as 3.93 and 3.82, respectively on a scale from 1 = disagree strongly to 5 = agree strongly. The respondents tolerated the FS. The mean of willingness to use it was 3.04 on the same scale from 1 to 5.

Auditory warning signals will not be required in FAD, since, by definition, the automation will be able to keep control of the vehicle in all possible conditions. Our previous study [5] proposed an experimental setup aimed at the auditory 3D representation of the environment outside of a vehicle, which
could be used in FAD for comfort and entertainment. The respondents of the survey were not particularly positive about the FS, possibly because they could not envision the system, or because of the lack of experience with such a system.

Results in [5] also showed that a female voice is the most preferred feedback for the support of TOR in highly automated driving. The female voice was perceived as the third most preferred warning signal for supporting TORs by the participants in [6] after a head-up display with a green icon, and a green icon on the dashboard.

The participants in the survey [6] were also asked to select the most urgent and the most annoying auditory warning signal from six options:

1. One beep
2. Two beeps
3. Horn sound
4. Bell sound
5. Female voice saying “Take over please”
6. Male voice saying “Take over please”

The horn sound was judged to be the most annoying while the male voice saying “Take over please” was considered the least annoying signal. The male voice was found to be the most urgent signal, and the bell sound was seen as the least urgent auditory warning signal.

1.2. Aim of the proposed experiment

The results from the authors’ survey studies [5], [6] will be validated in a driving simulator experiment. Specifically, urgency and annoyance of the six auditory warning signals will be evaluated, and the preferences of the participants for the type of warning signal for TOR in HAD will be polled in a questionnaire after the completion of the simulator experiment. The hypothesis that the female voice is the most preferred auditory signal, as was shown in [5], will be tested. An additional hypothesis that the first eye gaze after receiving a TOR will be directed to the side of the source of such cue will be evaluated.

2. EXPERIMENT ON THE USE OF AUDITORY INTERFACES FOR TAKEOVER REQUESTS DURING HAD

The experiment will be conducted with a Green Dino simulator at the Delft University of Technology. A TOR in the form of an auditory warning signal will support switching from HAD to manual driving.

All participants will be required to have a driver’s license. In the experiment, at least 24 participants, all students of Delft University of Technology, will be exposed to the six auditory signals listed in Section 1.1 during TORs.

Both directional (right and left) and non-directional cues will be provided via speakers in the simulator. The participants will be asked to rank the warning signals based on their urgency and other attributes to be defined in the later stages of the planning. The brake and steering reaction times of the participants will be also measured. The driver’s head and eye movements will be recorded with a non-obtrusive eye tracking system.

2.1. Secondary task

In order to ensure that the participants will have their eyes off the road, the participants in the experiments will be asked to perform a secondary task. A screen will be placed on the right side of the steering wheel during HAD. The participants will be required to perform the Surrogate Reference Task (SuRT, [7]) on the secondary screen.

2.2. Scenario

A within-subject, repeated measures design will be used. The participants will experience 18 TORs (six warning signals and three requests for each, left/right/non-directional). The sequence of TORs will be counterbalanced. The participants will drive a scenario similar to that in Gold et al. [8].

Figure 1 illustrates that the takeover scenario will be represented by an accident on the middle lane of a three-lane freeway. The driver will have the option to either stop on his/her lane by braking, or to swerve to the left or to the right lane. To make a lane change possible, the other lanes will not be occupied by other road users. Participants will be asked beforehand to use the mirrors and perform a shoulder check prior to making lane changes. At the time of the TOR, the stationary vehicle will appear 233 m in the middle lane. At a speed of 120 km/h this implies a takeover request time of 7 sec.

![Figure 1: Takeover scenario.](image)

The driving time between TORs will be 2 to 3 min. At the moment of the TOR the participant will have to take over control by turning the steering wheel or applying the brakes. Either of these actions will disengage the automation. The experiment will take approximately 1 hour per participant. Two breaks will be planned per participant.

2.3. Procedure and instructions

At arrival at the driving simulator laboratory, the participant will sign a consent form, explaining the purpose and procedures of the experiment. After signing the consent form the participant will be asked to fill out a general questionnaire about his/her driving behavior, demographic information, and general opinion about TOR modalities. Before entering the simulator, the participant will be reminded that he/she can stop the experiment at any time. Next, the participant will be asked to enter the simulator and start a training trial, which will take 2 min to complete. The participant will be familiarized with the SuRT task, learn how to disengage the automation, and experience a non-directional TOR. A second questionnaire with items on preferences for the warning signal for TOR in HAD, urgency, and annoyance of the signals will be given to the participant after finishing the scenario.
**2.4. Dependent measures**

We will record the reaction time and the takeover quality of the participants. The following measures will be calculated:

1. **Mean and SD reaction time first gaze reaction:** The time from the warning signal to the moment that the eyes of the participants first move away from the secondary task.
2. **Mean and SD reaction time road fixation:** The time from the warning signal to the moment that the eyes of the participants are directed back on the road.
3. **Mean and SD reaction time hands on steering wheel:** The time from the warning signal to the moment that the hands are back on the steering wheel.
4. **Mean and SD reaction time intervention:** The time from the warning signal to the moment that the driver uses either the brake or steering wheel.
5. **X- and Y-trajectories of the drivers, which allow inferring take-over quality smoothness [3], [8].**
6. **Number of safety checks performed:** The number of areas of interest (e.g., side mirror or blind spot) that the drivers checked before making a maneuver.
7. **Mean and SD of the Time-to-collision (TTC) to the stationary vehicle during TOR.**
8. **Mean and SD of the longitudinal and lateral accelerations during TOR.**
9. **Mean and SD of the steering wheel reversal rate [9].**

**3. RESULTS OF THE EXPERIMENT AND FUTURE STUDIES**

The results of the proposed experiment will be analyzed to derive recommendations for the development of a user interface for supporting TORs during HAD. These recommendations will be used in the preparation of a series of follow-up studies. One of these follow-up experiments will be conducted together with three researchers involved in the HFAuto project [10], at TUM München, the University of Southampton, and the Swedish National Road and Transport Research Institute (VTI). It will conducted during 2016 and incorporate a three-modal auditory/haptic/visual interface.

The first author started his PhD in August 2015. The goal of the author’s PhD research will incorporate two online surveys, a number of driving simulator experiments, and, possibly, field studies, resulting in the creation of an auditory interface that (in combination with a haptic and/or visual interface) will be capable of supporting takeover requests in HAD.

**4. ACKNOWLEDGMENT**

The authors are involved in the Marie Curie ITN: HFAuto (PITN-GA-2013-605817).

**5. REFERENCES**


A SOFTWARE LIBRARY FOR IN-VEHICLE AUDITORY DISPLAY

Yota Morimoto
Institute of Sonology
The Hauge, The Netherlands
yotamorimoto@gmail.com

ABSTRACT

The author develops a software library for in-vehicle auditory display projects. The library is developed using SuperCollider, an open-source cross-platform programming environment for sound synthesis and music composition. The features which cover different scenarios of in-vehicle auditory display projects are explained and short examples are provided.

1. INTRODUCTION

Since 2013 the author has worked for in-vehicle auditory display projects: from sonification of driving operation and vehicle performance data, to unobtrusive auditory notifications and multichannel auditory gamification. The present software library implemented in SuperCollider [1] is shaped over the course of the projects and covers different scenarios of in-vehicle auditory display. It aims primarily at leveraging SuperCollider’s sound synthesis capability [2] in in-vehicle context.

The following sections explain the three key features of the library: common tasks simplified; unified interface to both real-time and non-realtime data; and programatically generated sound synthesis templates.

2. COMMON TASKS SIMPLIFIED

Reading and writing driving operation and vehicle data are simplified. The interface to writing audio file (both realtime and non-realtime) is also greatly simplified (one method call). Mapping data to destination range (with adjustable curvature) is made easier and normalization can be performed within each channel of data or across channels of data [3]. Commonly used signal processing techniques (differentiation, filtering, DC offset removal, moving average, etc.) are implemented.

3. UNIFIED INTERFACE TO BOTH REALTIME AND NON-REALTIME DATA

A common interface to reading both realtime and non-realtime data is implemented. One only needs to specify the source (serial, network, human interface device or text data) and the sound synthesis part needs no change.

The above sources correspond to: realtime serial data from CAN (Controller Area Network) bus; realtime OSC (Open Sound Control) message from a machine over wired/wireless network. HID (Human Interface Device) data from such devices as Logitech Driving Force GT (to simulate driving operation); or CSV files or alike. All realtime data can be recorded for later use or analysis.

4. PROGRAMATICALLY GENERATED SOUND SYNTHESIS TEMPLATES

Over 30 sound synthesis templates (FM, additive, subtractive, granular synthesis, sound file player etc.) and 15 post-processing effectors (multiband EQ, various filters, reverb etc.) are implemented. These can simply be called for the project by name reducing the time-consuming task of sound synthesis programming. The templates are programatically generated and are thus readily extendable.

5. SHORT EXAMPLES

How to read recorded CSV data:

```plaintext
data = Data.new;
data.readCSV("data.csv");
```

or realtime data from CAN bus, OSC or HID:

```plaintext
data.readCAN;
data.readOSC;
data.readHID;
```

to sonify using sine tone sweeping (assuming freqRange defined):

```plaintext
'sweep'.play(data.map(freqRange));
```

Multichannel data expands to a corresponding number of synthesis nodes. It is possible to accesses each channel of data and map individually:

```plaintext
'sweep'.play(
  data[0].map(freqRange), data[1].map(ampRange)
);
```

For a earcon type of display, playing sound conditionally ('condition' synth performs analysis and triggers 'soundfile' player):

```plaintext
'condition'.play(data);
'soundfile'.playCondition;
```

6. REFERENCES

User Experience Evaluation Model for Auditory In-vehicle Systems

Alexander Meschtscherjakov, Nicole Peterer, Arno Laminger, David Wilfinger, Sandra Trösterer, Manfred Tscheligi

Christian-Doppler-Laboratory “Contextual Interfaces”
Center for Human-Computer Interaction, University of Salzburg
Sigmund, Haffner-Gasse 18, 5020 Salzburg, Austria
firstname.lastname@sbg.ac.at

ABSTRACT
In this paper we present an evaluation model to assess the users’ experiences (UX) when listening to an auditory in-vehicle system. The model consists of seven UX dimensions, which we consider to be relevant factors: presence, usability, aesthetics, pleasure, acceptance, self-expression, and status. We present the rationale for this model along with a perceived sound-quality questionnaire. We outline a user study in which we used the questionnaire and discuss the applicability of the model to assess perceived UX towards in-car sound in general.1

1. INTRODUCTION
In general, we can distinguish between two types of in vehicle sounds: natural induced sounds and artificially created sounds. Natural induced sounds, when driving in a vehicle, include the sound of the engine, the tires, wind, or the sound of a car door when being closed. Artificially created sounds include the music from the in-car entertainment system, the voice of the navigation system, and sounds induced by the car to inform the driver such as alarms, the sound when an indicator signal is turned on, or the voice of a remote person on the intercom. Artificially created sounds are played via the vehicle’s speaker system. These systems allow us to create sounds that create a specific effect inside the user. For example, a tone alerting the driver to fasten the seatbelt might be designed to be annoying.

In order to modify these sounds in a fine tuned way the quality of the in-vehicle audio system plays an important role. When evaluating the quality of such sound system, so far, either sound experts have been consulted or study participants had to know a certain sound vocabulary in order to evaluate the perceived sound of a system. Questionnaires, for example, include items such as “How satisfied are you with the playback of high frequencies?”, which are hard to answer for laymen [2].

2. USER EXPERIENCE EVALUATION MODEL
We have developed a model, which aims at evaluating the perceived user experience (UX) with regard to sound qualities of a driver or passenger in a vehicle. The model consists of seven dimensions we consider to be relevant when assessing perceived UX qualities in the automotive domain. The model was developed in an iterative process involving four automotive HMI experts. The dimensions were generated based on previous work we have done [3]. In this work we summarize and discuss a set of in-car field studies we have conducted which focus on UX of drivers and passengers while driving in a vehicle. Based on these in-situ studies we have identified several UX factors, which are relevant in the automotive context. These factors were also inspired by other UX evaluation models and frameworks presented at the COST294-MAUSE affiliated workshop [1].

Apart from the model, we have formulated a car sound UX questionnaire consisting of seven items. The items were used in a user study to compare two assess the perceived UX towards two in-car audio systems. A more detailed description of the design process can be found in [4]. The seven UX dimensions, which we consider to be relevant factors, are presence, usability, aesthetics, pleasure, acceptance, self-expression, and status:

- **Presence** refers to the feeling of being in close spatial proximity of the sound source.
- **Usability** refers to the quality of the music or oral text in terms of understandability.
- **Aesthetics** targets at the sensory quality of the sound.
- The dimension **pleasure** refers to the listener’s feeling of enjoyment when listening to the sound.
- **User acceptance** targets at evaluating the listener’s willingness to use the system in the future.
- The dimension **self expression** refers to the system’s ability to represent the user’s values.
- **Status** refers to the user’s feeling of the sound system as a status symbol.

Based on this model we have built a questionnaire with one item for each of these dimensions. The items were partially derived from literature and partially self formulated. Table 1 includes the seven dimensions of the model as well as corresponding questionnaire items. The items are supposed to be rated on a Likert scale. For the following user study we have utilized a 6-point Likert scale from disagree=0 to agree=5.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Item</th>
</tr>
</thead>
</table>

Table 1: Dimensions of the perceived audio UX evaluation model including corresponding items.

1 This position paper is based on our work Measuring Sound Experience with In-Vehicle Speaker Systems, which we presented at the 4th International Workshop on Perceptual Quality of Systems in 2013 [4], which focuses on the comparison of the sound quality of in-car audio systems. This paper summarizes the results of [4] and discusses its applicability to evaluate perceived UX towards in-car sound in general.
<table>
<thead>
<tr>
<th>Presence</th>
<th>When I hear sound with this system, it feels as if I was where it was recorded (e.g., concert hall).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>I can understand music and language through the system very well.</td>
</tr>
<tr>
<td>Aesthetics</td>
<td>It feels good to listen to the sound through the system.</td>
</tr>
<tr>
<td>Joy of Use</td>
<td>It is pleasurable to listen to sound through the system.</td>
</tr>
<tr>
<td>User Acceptance</td>
<td>In my car, I would use the same system to listen to sound.</td>
</tr>
<tr>
<td>Self Expression</td>
<td>The system sounds exactly how it suits my taste.</td>
</tr>
<tr>
<td>Status</td>
<td>I would be proud if I had the system in my car.</td>
</tr>
</tbody>
</table>

3. USER STUDY

In order to evaluate the feasibility of the questionnaire we conducted a pilot study. The aim of the study was to compare two in-car audio systems. In the following we describe the setup of the user study in order to highlight the applicability of the questionnaire to evaluate perceived sound UX.

Overall, 114 subjects (84m, 30f) with a mean age of 37 years participated in the study. They were asked to fill out the above-mentioned questionnaire while sitting in a parked car together with some demographic questions. The questionnaire had to be filled out twice – one for each audio system, which was installed in two different Volkswagen Golf 6. The order was counter balanced. In order to get a good impression on the sound quality we prepared 6 audio tracks, each lasting for 30 seconds. These sound samples included different music styles (classic, folk, rock, hip-hop), a children’s song as well spoken news and traffic information. We believe that it is crucial to provide different kinds of audio samples in order to get a good impression on the perceived sound quality regardless of the personal preferences of the subjects.

A factor analysis on the seven items of the questionnaire indicates that the questionnaire is a single-factor construct. Internal consistency was shown with a Cronbach’s α coefficient. A detailed description of the quality criteria can be found in [4].

The results of the study allowed us to compare the two systems. When combining factors in the questionnaire, System II (M=4.36, SD=1.43) was rated higher than System I (M=4.79, SD=1.42), although no significant differences were found. Interestingly, for each single dimension System II was rated higher than System I. The factors usability and pleasure were rated highest for both systems; presence and status were rated lowest. Again, more details can be found in [4].

4. DISCUSSION

When reflecting on the study we can conclude that the questionnaire was working well. Participants reported that it was rather easy for them to answer the questions. Regarding the comparison of the two systems we could not find any significant differences. A reason for that could be that both system were rather high-end sound systems, which in fact did not differ very much. Another shortcoming of the study was, that it took place in a parked car with the engines turned off. In a real world setting there will always be natural induced sounds. Nonetheless, the questionnaire is a very efficient method for a quick assessment of UX with in-car sound systems.

Regarding the proposed dimensions, we have to acknowledge that they are neither complete nor mutual exclusive, but they provide a solid overview of dimensions relevant for the assessment of perceived user experience when listening to an in-car sound system.

We are aware that not all factors may be relevant for auditory displays; thus, the model needs to be adapted to the type of auditory display in focus. In order to evaluate, for example, the perceived quality of artificial sounds of an electric vehicle we believe that a model may consists of the following dimensions:

- **Presence** to assess the quality of the sound.
- **Usability** to assess the usefulness of the sound (e.g. to warn pedestrians to about an approaching vehicle.
- **Aesthetics** to assess the sensory quality of the sound.
- **Pleasure** to assess the listener’s feeling of enjoyment.
- **User acceptance** to assess the user’s willingness to use the system in the future.
- **Self expression** to assess the ability of the sound to represent the user’s values, for example, in cases when the user can personalize the sound.
- **Status** to assess the user’s feeling of the sound as a status symbol, again in cases when the user can personalize the sound.

Concluding, we believe that there is a need for such a model in order to be able to assess users’ experiences when listening to in-car sounds.

5. ACKNOWLEDGMENTS

The financial support by the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development and AUDIO MOBIL Elektronik GmbH is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

6. REFERENCES

In-vehicle Auditory Interactions Workshop