AUDITORY DISPLAYS FOR AUTOMATED DRIVING - CHALLENGES AND OPPORTUNITIES

Pontus Larsson
Interactive Sound Quality
Volvo Car Group
Göteborg, SE-405 31, Sweden
pontus.larsson.3@volvocars.com

Justyna Maculewicz
User Experience Centre
Volvo Car Group
Göteborg, SE-405 31, Sweden
justyna.maculewicz@volvocars.com

Johan Fagerlönn
RISE Interactive
Piteå, SE-941 63, Sweden
johan.fagerlonn@ri.se

Max Lachmann
Pole Position Production
Galtabäcksvägen 11
Bromma, SE-168 55, Sweden
max@pole.se

ABSTRACT

The current position paper discusses vital challenges related to the user experience design in unsupervised, highly automated cars. These challenges are: (1) how to avoid motion sickness, (2) how to ensure users’ trust in the automation, (3) how to ensure usability and support the formation of accurate mental models of the automation system, and (4) how to provide a pleasant and enjoyable experience. We argue for that auditory displays have the potential to help solve these issues. While auditory displays in modern vehicles typically make use of discrete and salient cues, we argue that the use of less intrusive continuous sonic interaction could be more beneficial for the user experience.

1. INTRODUCTION

The interest in automated road vehicles has been ever-increasing during the past few years. Reasons for the hype around Automated Driving (AD) may be its potential to bring positive societal effects in terms of reduced environmental impact, improved traffic safety, and more efficient mobility [1]. In addition to this, and regarding the specific appeal to the drivers/users, AD technology may allow people to be more productive, comfortable and relaxed during their daily commutes and other travels [1, 2].

Cars currently on the market offer low levels of automation still requiring human supervision, but we will likely see highly automated cars in the near future [2]. In fact, Waymo is already offering “robotaxi” solutions today, albeit in a limited setting [3]. Users of such vehicles are now considered passengers rather than drivers. Therefore, these AD vehicles bring possibilities to create completely new types of experiences for users. However, they may also introduce new types of problems, such as motion sickness and a lack of trust in the automation.

Sound may be a suitable medium for forming the user experience in AD vehicles. Sound can provide information to the users even if their eyes are closed, inform users continuously and subconsciously, and efficiently affect their emotional state.

In the current position paper, we will discuss the use of sound for the purpose of reinventing the in-car user experience when we go from manually driven- to highly automated vehicles. We will present a set of challenges that we consider vital to this area of research and innovation from the perspective of the automotive industry. The paper is intended to form a foundation for further research activities within our recently initiated research project “Sonic Interaction In Intelligent Cars” (SIIC) [4] and builds on initial investigations within this project. The paper also builds on knowledge from our workshop held at ICAD 2018 with the same title [5]. The intentions of the 2018 workshop were to build a new community for interactive sounds for AD that bridges the auditory display community with the automotive user interface community, and to discuss and exchange ideas within the field of AD as well as to explore promising directions for future work. Hopefully, the current paper will also inspire continued work in the directions set out by the 2018 workshop.

2. BACKGROUND

Driving Automation is currently one of the big trends within the automotive industry today along with electrification and new types of mobility services and solutions. A range of car manufacturers currently offer SAE (Society of Automotive Engineers' Level 2 or 3) AD solutions.
Engineers - an automotive standardization body) Level 1-2 automation functions in their cars. With these functions, the driver still has the responsibility to supervise the automation and take over driving when needed [6]. In other words, the driver cannot perform secondary tasks such as reading - or even take their eyes off the road - while this type of low level automation is active. The next generation of automated cars aimed at reaching Level 3 or 4 automation [6] are currently being developed. With Level 4 (L4) automation engaged, the driver no longer has to supervise the automation and automation will not rely on the drivers’ ability to take over driving when the automation reaches its operating boundaries - which potentially makes it safer than L2-3 automation [7].

L4 automation is also highly desirable from the user perspective since the user can truly make use of the time freed up by the automation [8].

L4 automation or “unsupervised AD” (as we will refer to it hereafter) provides new ways of interacting with cars and the entire experience of them may be drastically different compared to that of a traditional, manually-driven car. For example, when drivers are placed in a new role where they have focused mainly on supervised AD, while there is a lack of work focusing on the possibilities and challenges involved in unsupervised AD.

Within supervised AD, it has been found that trust, feeling of safety and acceptance can be influenced by the users’ interaction with and experience of the car via its human-machine interfaces and it is likely that this will be possible also for unsupervised AD [16,17]. An as of yet rather unexplored area of research is to use sound, or sonic interaction, as a means of communication between the user and the AD car. Sound has unique advantages in comparison to e.g. pure visual communication. For instance, since our hearing is omnidirectional, sound can convey information no matter where the user has his/her visual focus. This feature of sonic interaction may prove to be especially useful in an unsupervised AD context where the user might have his/her visual focus anywhere (he/she might be e.g. reading a book or looking at the passing landscape) and not at the vehicle’s visual displays [18]. For example, the study by Gang et al. [18] approaches the topic of trust in AD with an auditory solution presenting necessary information through spatially located abstract earcons. Their goal was to present desired information without causing alarm or compelling people to act.

Also, it is well-known that sound can easily catch attention and change the emotional and physiological state of the driver [19]. These properties make sound suitable as warning signals and alarms, which is one of the most common types of sonic interactions in cars today (e.g. collision alerts, belt reminder etc.). Sound design requires a lot of consideration especially when it comes to such warning sounds, which should result in appropriate and sometimes quick reactions [19-23]. For unsupervised AD, these types of attention-grabbing sounds triggered by discrete events may not be as useful as for manual driving and supervised AD since quick driver (user) reactions are not likely to be requested by the AD system. The traditional type of discrete sounds may also be too intrusive and annoying - and especially so when the user is really not involved in driving the car.

An alternative type of sonic interaction design deals with the manipulation of sounds which are already part of the soundscape. For instance, Fagerlönn, Lindberg and Sirkka [24] investigated the possibility of manipulating the sound from the in-vehicle radio to provide early warnings. Similarly, Nylänen, Lopez and Toulson [25] investigated the usefulness of various strategies, such as manipulating music content, to help drivers keep a steady speed.

Yet another design alternative to adding discrete and salient auditory cues is continuous sonic interaction, which builds more proactive interaction between a car and a user, rather than just presenting information in a reactive manner. Continuous sonic interaction in the current project application refers to an auditory display that, based on continuous input signals (obtained from control actions by the users or other input signals), provides concurrent auditory information about the resulting state or response of the system [26]. Cars with functions for unsupervised AD are equipped with an abundance of sensors and related processing units which enable them being able to react properly to the surrounding. The streams of data from the sensors could however also be used for creating such continuous sonic interaction for the cars’ users, possibly making the AD system more transparent, intuitive, useful and engaging. An example of this type of auditory display from the AD domain was suggested by Bazilinskyy, Larsson & de Winter [27] who in an on road study explored using a subtle, continuous sound for which the level was mapped to the distance to other, leading vehicles and another sound which was in similar manner continuously informing about the ego vehicle’s lateral position in the lane.

As opposed to the traditional discrete auditory signals, the continuous type of auditory display (sometimes also referred to as sonification) is believed to be better matched to humans who have evolved to act and control their environment in continuous fashion, and the auditory responses in everyday interactions tend to involve nuanced feedback depending subtly on human actions [26]. Continuous interaction also promotes closed-loop relationship which creates a higher level of perceived cooperation between human and machine, which can improve the understanding of the machine and have the potential to increase a user’s sense of engagement [26]. Still, despite having these potential benefits, this type of sonic interaction likely needs to be carefully designed in order for it to be perceived as pleasant and enjoyable - in turn a prerequisite for its perceived usefulness and user adoption.

300
Thus, there may be new ways of using sound in unsupervised AD, but exactly which roles, if any, will sound play in this novel type of human-machine interaction? In the next sections we will go into detail of the aspects and challenges of the user experience in unsupervised AD cars that we foresee will be of interest to UX designers from the automotive industry perspective and how auditory displays possibly could play a role in regards to meeting these challenges.

3. SOUND TO REDUCE MOTION SICKNESS

One of the main arguments for autonomous vehicles is that they will allow users to spend their time in a more productive way. For example, the Concept 26 by Volvo Cars [8] suggest that the driver should, when having delegated the driving task to the autonomous driving function, be able to either relax or “create” (meaning: make calls, write emails or watch films and TV shows, etc.). Similar ideas are envisioned by other companies’ designers, e.g. in the InMotion concept by NEVS [39], where the seats and interior can be arranged for either privacy, social interaction or work-related presentations, or the Volvo 360c concept which has similar features for socializing, working or even sleeping [40].

However, there is a growing concern that the possibilities for performing non-driving-related in-car activities will be limited due to the risk of motion sickness (kinetosis) [30, 31]. Motion sickness is a condition characterized by symptoms of nausea, dizziness, fatigue and other types of physical discomfort [30]. According to [30] there are three main factors leading to the condition of motion sickness; conflict between visual and vestibular inputs, loss of control over one’s movements, and the reduced ability to anticipate the direction of movement - and all these factors may be present in an unsupervised AD scenario. There are basically two different ways of reducing motion sickness: 1) Allow occupants to anticipate the future motion trajectory and 2) Avoid incongruent self-motion cues. This implies that if it would be possible to cancel out the vestibular signals caused by the vehicle motion, when a passenger is looking down, that would reduce motion sickness.

Solutions to the motion sickness problem have been suggested by e.g. Waymo [32], Uber [33], and researchers at UMTRI [34]. UMTRI’s and Uber’s solution suggests providing the user with artificially generated stimuli that would reduce sensory conflicts and giving the user cues to be able to predict the car’s movements – in Uber’s case primarily a combination of visual-, haptic-, and airflow stimuli. While these types of stimuli could be efficient in reducing visual/vestibular conflict, the required displays would likely be quite expensive and cumbersome to integrate in a production vehicle. Using sound reproduced by an audio system, readily available in most production cars today, would be a more practical display. Uber suggests giving the user audio prompts (speech or tonal) or visual indications of upcoming maneuvers that could inform the user when to look up to avoid visuo-vestibular conflict. While this solution could potentially be efficient in reducing motion sickness it could also be quite annoying and thus result in a poor user experience.

Instead of providing audio prompts, one could in a more continuous fashion sonify the car’s maneuvers slightly in advance to them happening so that the passengers know when to look up at the road. This solution would thus have the same effect as the proposal by Uber referred to above, but could be better from a UX perspective and feel more natural - e.g. as an enhanced engine sound. We label this solution the Sonic Ghost Mode, (in analogy to ghost mode used in computer games, see Figure 1) and note that it may be useful also for improving trust since it informs the user of the automation’s intentions (see next section for more elaborations on this matter). For examples on how this may sound as well as other examples related to AD sonification, see [4].

Figure 1: Visual ghost mode in Rallisport Challenge 2 [35]. In this case, the intention is to visualize the difference between two race runs.

Given that humans can experience self-motion as a result of being exposed to certain types of sound [36, 37], such sound synchronized with the car’s movement could be another potential solution to the motion sickness problem. The idea would be that the added sound compensates for the lack of visual stimuli when the user is looking down to read a book or similar and vestibular stimulation is prominent. A similar idea is presented in [38].

Yet another possibility would be to use sound to reduce or even cancel the vestibular signals. This may seem far-fetched but the idea is based on the findings that both air conducted (AC) and bone conducted (BC) sound can affect the vestibular system [39]. For example, in an experiment referred to in [39] a constant 500 Hz tone was shown to cause postural deviation. If sound can be delivered to the passengers when that cause a vestibular response opposite to the response caused by the car, it may result in a reduced total vestibular signal (similar to active noise cancellation). If that sound is presented when passengers look down it could potentially reduce the visual-vestibular conflict and thus also motion sickness.

4. SOUND TO INCREASE TRUST AND ACCEPTANCE

Trust can be defined as “the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” [40]. In the case of vehicle automation, the agent would be the vehicle itself or the part of the vehicle that the user identifies as responsible for the automation. In the lower levels of automation (up to level 3 [6]) the most severe trust-related risk is “overtrust”, i.e. when a person believes that the automation has better capabilities than it actually has. For higher levels of automation overttrust is less of a problem since automation in these levels by definition never should rely on a driver's intervention. The most obvious trust-related risk for high level (L4/5) automation vehicles is that of under-trust - and that people will
not use it due to the fact that they do not trust them. Recent studies have found that many people would be afraid of riding in an automated vehicle [41], and a way to increase trust is through the design of the vehicle itself and its user interface.

For example, including human-like features in the interface, anthropomorphism, has been shown to increase trust in automation [15]. Anthropomorphic features can guide users in their assessment of whether a machine is dangerous. By expressing human intelligence, friendliness and care, the design may increase the feeling of trust. Figure 2 shows an example where anthropomorphism has been used in the external vehicle design to amplify trust during interaction with self-driving cars - another similar concept can be found in [42].

Research has been shown that user’s ability to estimate the predictability of the machine’s behaviours affects trust [44]. A study by Helldin et al [17] showed e.g. that visual representation of a low level automated car’s uncertainty (i.e. how sure it is of its ability to drive automatically) leads to a better calibrated trust. For a high level automation vehicle, it is reasonable to believe that it would be perceived to be more capable of driving by itself when it seems able to think and sense its surroundings than when it just gives an impression of “mindless machinery” [15].

Examples of when this type of representation is shown visually to the backseat passengers in the driverless shuttle service that is being piloted currently can be found in e.g. [45]. Similar type of information can of course be given through other modalities than visual ones. Given that passengers potentially will not, or even would not like to, have their eyes directed to one display, it seems more reasonable to use an auditory display to increase trust.

It is however not obvious what the best type of auditory information would be. One could speculate that event-driven audio prompts (e.g. chimes) are efficient in informing the driver of the vehicle’s action and abilities but will quickly become annoying since they may occur quite often.

Speech messages could provide rich information and also give anthropomorphic features to the automation. Considering the fast development of speech assistant technology today, verbal interaction will most likely have a role in future highly automated cars. But again, providing information about ordinary actions through speech may be too intrusive. Furthermore, designing anthropomorphic features using non-verbal sound is certainly a possibility, which was recently demonstrated by Collins and Dockwray [46]. We therefore hypothesize that a more continuous and subtle sonification of the car movements, intentions and abilities would be more efficient in inducing the appropriate degree of trust in the user.

Trust can be seen as part of the wider scope of user acceptance – naturally also crucial for user adoption of automated vehicles [14]. Among many things, acceptance is contingent on ease of use, usefulness and enjoyment [14] and we believe that sound can play a role in increasing acceptance of autonomous vehicles by making them more useful and comfortable. Consider for example the use case of stop-and-go traffic (low speed queueing) or other situations where the vehicle brakes and/or accelerates frequently which may cause the user to look up to see what is happening. If sonification would provide subtle information to the user on what is going on in traffic as in the Sonic Ghost Mode described in previous section, the user would not be triggered to look up each time he/she experiences sudden motion cues. Or, as we discussed in previous section, the sonification could also aid the user in knowing when to look up in order for him/her to avoid motion sickness. Therefore, e.g. performing eyes-off-road visual tasks may be perceived as more comfortable and less annoying with sonification added. There are several other similar situations when sonification could aid in “relaying” information about the driving scenario to the user without being overly intrusive or annoying (change of route, roadworks ahead, time to handover etc.) which in turn could increase comfort and the perceived usefulness of automation.

5. SOUND TO IMPROVE USABILITY AND MENTAL MODELS OF AD

To be able to predict possible actions and their consequences, users create mental models of the situations which they partake in. If a systems’ behaviour corresponds to user expectations, encapsulated in a mental model, it heightens trust and provides a more positive user experience [47]. When a user approaches a new system, he or she builds a mental model based on previous experiences which might not be applicable in these new situations. However, with a proper user interaction design a user can be provided with a level of information, which could help to build more appropriate mental models and foresee system’s behaviour. Therefore, transparent interfaces adapted to the mental system of the user are a prerequisite for the user to be able to develop necessary situation and system awareness in interactions with the automated system [48].

A use case that has already been identified as critical during lower levels of automation is handover of control from the AD system to the human. Supporting the user in this situation by creating awareness of the system’s state and providing the user with a correct mental model of the system is however also important for higher levels of automation as long as the technology allows for multiple levels of automation (i.e. it will obviously not be important in vehicles where only one level of AD is present, such as in “robotaxi” vehicles, e.g. [45]). In unsupervised AD, the transition to manual driving can be challenging to handle since the intended “driver” may be in very different states ranging from being asleep to fully aware of the traffic situation. According to Strabala et al. [49], to perform successful handover one needs to agree that handover will happen, establish timing of the handover and decide how the process will be performed.
Based on this, we hypothesize that the handover situation requires some preparation to transfer a driver from non-driving towards the driving context. The vehicle-driver interaction should gently prepare a driver for handing over control to the car or vice versa. The cooperation between a car and a driver should be seen as partnership and handovers in both directions should happen by mutual agreement and in the right moment. The preparation probably needs to be adaptive to the current state of the intended driver. In some cases, the driver will need minutes or more to prepare, while other situations may require just a few seconds. In a previous research project Methods for Designing Future Autonomous Systems (MODAS) [50], a driver interface for AD was designed with professional drivers. When possible, the system provided information about upcoming handovers hours in advance during the driving route (see Figure 3).

Figure 3. A design concept from the project MODAS [50]. A visual timeline (blue line) presented at the top of the windscreen indicates when a handover may be necessary.

Sonic interaction accompanying the handover process could be a useful part of a supportive multimodal user interface and help the user in developing a correct system/automation mode awareness and mental model of the system. Using sonic interaction, the user can perceive the information given by the user interface even with eyes closed and allows the user to keep his/her eyes on the road/traffic during the time of the actual handover.

6. PLEASANT AND ENJOYABLE SOUND

Even though a particular sound design might be highly useful (e.g. it reduces motion sickness, increases trust, induces correct mental models etc.), it is likely that the sonification and the AD system as a whole needs to be aesthetically pleasing and induce a sense of joy-of-use [51] to make the user engage the AD system for an extended period of time, and for the user to prefer sonification over more traditional means of signalling (visual displays, traditional sound chimes etc.). Apart from the advantages identified earlier in this proposal, continuous sonification could also be used in to induce certain moods [52] in similar ways as is being done within cinema and computer game sound design. Sound can in this way be used to soothe the user and make them simply enjoy the ride. Moreover, a gradually built up sonic atmosphere can be used to gradually increase attention and awareness of the user when, for example, the autonomous drive is about to reach its operational domain limits and the user is supposed to take over driving. Continuous sonification also makes the automation user interface more responsive and adaptive to user behaviour/reactions which could make the whole experience more balanced and pleasant.

While the field of designing efficient traditional sounds for in-car applications is quite well understood, research regarding how to design the above-described adaptive sonification-based displays that people enjoy using is scarce [53].

In line with what is suggested by [53], we believe that employing ideas and methods from the areas of Design Thinking and User Centered design could be one way of improving the overall user experience of auditory displays. Using guidelines and praxis from the art of sound design for movies or computer games could be another way to understand how appealing, aesthetically pleasing continuous sonic experiences intended for automotive information displays should be designed. For example, [54] proposes a simple method for evaluating computer game soundscapes and devises design guidelines for how to heighten immersion and reduce listening fatigue - these might be applicable to an in-car context as well.

7. CONCLUSIONS

The development in automation has the potential to completely redefine the usage and user experience of road vehicles, especially when the technology allows unsupervised driving. In this paper we have presented a set of design challenges that are central to facilitating the successful introduction of highly-automated cars. These challenges are: (1) counteract motion sickness, (2) increase users’ trust and acceptance, (3) improve usability and support the formation of accurate mental models, and (4) provide a pleasant and enjoyable experience. Furthermore, we argue that the utilization of auditory displays is a promising way to meet these challenges. However, while auditory displays in vehicles typically make use of discrete and salient cues, we argue that the use of less intrusive continuous sonic interaction can be a more successful strategy to facilitate a positive user experience. This will be investigated in recently started project [4] and the sonication solutions will be evaluated with users in a virtual environment and in a test car. This car has systems installed that enable experiences of high-level automation in realistic traffic environments. This is made possible by a “Wizard of Oz” setup, where a test leader/driver monitors the vehicle and can make corrections if necessary without the test person’s awareness [55].

We hope that the project’s challenges, along with the arguments supporting them, can inspire other researchers and practitioners to engage in the research and development of new types of auditory displays for self-driving vehicles.

8. ACKNOWLEDGMENT

The work presented in this paper is part of the currently ongoing project Sonic Interaction in Intelligent Cars (SIIC), funded by the Swedish partnership programme Strategic Vehicle Research and Innovation (FFI), D. Nr. 2018-02730.

9. REFERENCES


The 25th International Conference on Auditory Display (ICAD 2019) 23–27 June 2019, Northumbria University


