Improving the User Experience with Spatial Auditory Displays

Scott Pennock

QNX Software Systems (Wavemakers), Vancouver, British Columbia, Canada, Email: spennock@qnx.com

Introduction

For most non-music applications today, users perceive sound coming from the *same* location within the vehicle. Spatial Auditory Displays (SADs) cause users to perceive sound coming from *different* positions in 3-dimensional space. The positions could be somewhere inside the vehicle, or slightly outside the vehicle.

A standard auditory display and SAD are illustrated in Figures 1 and 2, respectively. The vehicle shown has a Speech Recognition (SR) system that is used to both automatically dial phone numbers and control the climate control system of the vehicle. Figure 1 shows that prompts for both the phone and climate control systems are localized at the same position on the standard auditory display. Figure 2 shows that these same prompts are localized at different positions on the SAD—each position corresponding to a physical location the application is associated with.

The purpose of this paper is to explore the ways in which SADs can be used to improve the user experience of current and future applications. For the examples given an ideal SAD is assumed. This is done to show what is possible. Some of the examples assume technology that exists, but that may not be implemented yet. It should also be noted that even with an ideal SAD there are perceptual limits to what can be achieved in terms of spatial resolution.

The following sections will discuss different aspects of SADs. The next section will discuss how SADs improve the user experience. This will be followed by a section that identifies current and future Next Generation (NG) car applications that could benefit from SADs. Then, there is a section that identifies acoustic cues used by listeners to localize sound. After that, some of the different types of SADs are described. The last section will then summarize conclusions.

User Experience improvements

SADs improve the user experience by increasing task performance and user satisfaction.

Better task performance

SADs improve task performance by reducing judgement errors, decreasing response times, increasing speech comprehension, and reducing driver distraction.

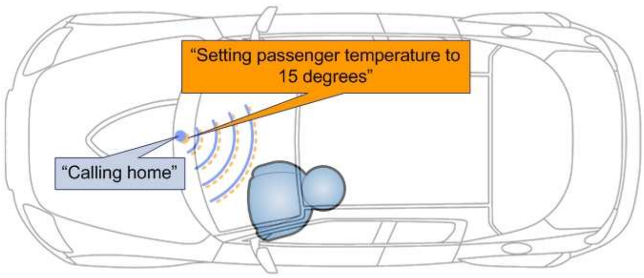


Figure 1: Prompt localization using standard auditory display for phone and climate control applications.

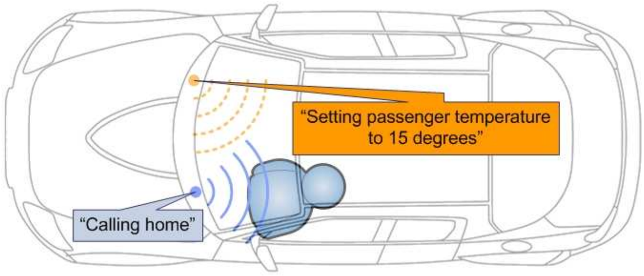


Figure 2: Prompt localization using SAD for phone and climate control applications.

The types of judgement errors that are reduced include Application/Object/Event identification errors, location errors, and navigation manoeuvre errors.

Application/object/event identification errors occur when users think the sound came from an application, object, or event other than the one it actually came from. Often users eventually figure out which object the sound is associated with, but only after increased delays in interacting with the application/object/event. These errors also cause increased effort. SADs help reduce these types of errors because spatial location is a reliable cue to the identity of the sound's source.

Location errors occur when the user's judgement of a sound source's location is wrong, or significantly different than, it's true or intended location. For example, the location of an emergency vehicle may be incorrectly judged if the driver's window is open and there are nearby building reflections. SADs help reduce these types of errors by better controlling the acoustic cues delivered to the user.

Navigation manoeuvre errors occur when the driver fails to correctly execute a navigation manoeuvre. SADs help reduce these types of errors in two ways. First, they provide redundant information on the location of the manoeuvre.

Instead of just hearing a prompt that says “turn right in 100 meters”, distance information can be conveyed through localization of the prompt itself. Second, SADs can convey trajectory information about the manoeuvre. For example, a “slight right” can actually be conveyed using apparent motion of the prompt’s location.

Response times are reduced in a couple of ways. First, the time it takes to identify which application the sound came from is reduced. Second, the time it takes to identify the location of an object is reduced [1]—which is very important if the object happens to be an emergency vehicle nearby or another car on a collision path!

Speech comprehension is also better in noisy vehicle environments [2]. This is because SADs can be used to localize speech at a different location than noise sources (e.g., tire, engine, etc.). This allows users to better separate the speech signal from vehicle noises. This effect is also known as “the cocktail party effect”.

Better speech comprehension will also reduce driver distraction [3]. This is because the speech comprehension task uses fewer shared cognitive resources such as working memory and attention.

Higher user satisfaction

User satisfaction is higher because SADs make applications easier to use and increase their realism.

Applications are easier to use for a few reasons. First, they help users keep better track of all the applications. Spatial location is a mnemonic device that helps users remember application items by associating them with a spatial location. This also facilitates simultaneous interaction with multiple applications. Second, they make identification of the current application/object/event effortless. The human auditory system automatically processes acoustic localization cues to identify position, and users know that any sound coming from this location belongs to the associated application/object/event. Third, they naturally convey information within an application. This could be the physical location of objects/events handled by the application (e.g., emergency vehicle, next navigation manoeuvre, etc.) or more abstract representations such as menu structure of a speech interface.

The realism of applications is also increased by SADs. For speech communications, there is a better “sense of presence“ created by hearing the far end talker crisply localized directly in front instead of a smeared spatial image localized somewhere down in the footwell. For other applications, objects and events are mapped to their physical location just like in real life.

NG Car applications

The NG car will enable many new applications. This increase in applications has the potential to confuse and frustrate users. The following sections describe how SADs can be used to improve the user experience for several current and NG car applications. These vehicle applications are grouped into 3 types: Inside vehicle, Outside vehicle, and Network-based communications. They differ based on where the objects, events, and data that they interact with are located.

Inside vehicle communications

Speech Recognition (SR) command and control

The SR command and control application uses Automatic Speech Recognition to control vehicle related functions. A common SR system is sometimes used for a diverse set of controls (e.g., phone, HVAC, radio, etc.) which can confuse the user as to which command was recognized and what is being controlled. SADs can improve the user experience by localizing prompts from each control at a separate location (as described in introduction). This helps identify which control the prompt is associated with and provides immediate feedback that the command was recognized successfully.

Passenger communications

The passenger communications application reinforces the talker’s speech using the vehicle audio system. Sometimes passengers have trouble hearing each other due to noise or because the talker is projecting speech away from the listener (e.g., driver talking to rear seat passenger). Passenger communications systems help solve these problems, but can cause listeners to hear the talker’s speech coming from a different location. SADs can help improve the user experience by ensuring sound reinforcement is done in such a way that listeners hear the talker’s speech coming from the same location as their actual position.

Vehicle-to-object communications (inside)

The NG car will allow users to connect and interact with many devices (i.e., objects) within the vehicle—using both wired and wireless communications. Wired objects might include a USB drive plugged in to the dashboard or a MP3 player connected to the rear seat entertainment system. Wireless objects might include a Bluetooth enabled phone or WiFi laptop. The Vehicle-to-object communications application handles interaction with these objects. Auditory interaction may take the form of non-verbal prompts, verbal prompts, or speech communications. Keeping track of all the objects that are currently connected will be a challenging task for the user. It will also be difficult figuring out which object the current interaction belongs to. SADs improve the user experience by helping users keep track of all objects. They also allow users to quickly and easily identify the current object with low effort.

Find objects

This application is similar to the paging function on some cordless phones. It plays non-verbal or verbal prompts over the SAD such that they are localized at the physical location of the object. SADs improve the user experience by helping users find objects within, and possibly outside, the vehicle. A problem with the paging function found in cordless phones is that it is not always possible to have the object produce an audible sound (e.g., no speaker, in the trunk, etc.). SADs do not have this issue.

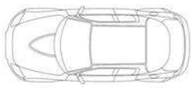
Outside vehicle communications

Collision avoidance

The NG car will be able to track the positions and trajectories of nearby vehicles. The Collision avoidance application will detect situations where there is a high risk of a collision and alert drivers so that corrective action can be taken to avoid it. SADs decrease reaction times and judgement errors so that there is a better chance of avoiding the collision.

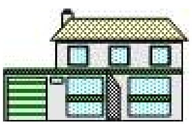
Vehicle-to-object communications (outside)

The NG car will also be able to wirelessly connect and interact with objects outside the vehicle. The Vehicle-to-object communications application handles interaction with these objects. Examples of such objects, and how SADs improve the user experience, are given below:

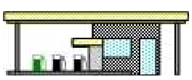


Other vehicle—Speech communications can be established with nearby vehicles that are within sight. SADs allow users to quickly and easily identify the vehicle which is sending or accepting call requests. They also help track the other vehicles locations and make it clear which vehicle speech is coming from.

Roadside sensors—Roadside sensors will detect dangerous road conditions (e.g., icy road, deer, etc.) and notify vehicles that are passing by. SADs will allow users to more quickly and easily respond to potential hazards by localizing associated audible alerts at the location of the hazard.



Home network—Interaction with home networks will include exchanging data (e.g., music, videos, etc.), controlling home functions (e.g., lighting, security, etc.), and talking with home occupants. SADs will cause prompts and speech communications to be heard as coming from the home. This will improve the user experience by making it clear that prompts and speech are coming from the home network instead of other applications.



Roadside stations—The NG car will interact with different types of roadside stations (e.g., toll booth, checkpoint, etc.). SADs will improve the user experience by helping identify where the roadside station is (it may not always be obvious) and keeping associated sounds separate from other applications.

Emergency vehicle alert and tracking

Sirens usually do a good job of alerting the driver that an emergency vehicle is in the area. However, drivers are often confused as to the location and trajectory of the emergency vehicle. Worse yet, drivers sometimes localize the sound incorrectly and think the emergency vehicle is coming from the wrong direction! These localization errors are mainly due to misleading localization cues (e.g., building reflection, cracked window, etc.). However, they are also the result of limitations of the human auditory system for localizing sounds from certain directions (e.g., directly in front or behind). SADs help overcome these issues by providing more control over localization cues presented to the driver. Drivers can be alerted and start tracking emergency vehicles even before they would normally hear the acoustic waveform from the siren. Misleading cues due to multi-path issues can be prevented by presenting strong and controlled acoustic cues over the SAD. Ambiguous localization cues (e.g., vehicle directly behind) can be corrected by slightly shifting the rendered position to produce more robust localization cues for the driver.

Real-time local traffic

The Real-time local traffic application alerts drivers to nearby traffic congestion and helps them avoid it. SADs allow users to identify multiple congestion points with less effort and navigate around them successfully. Perhaps more importantly, it reduces driver distraction. As described earlier, it does this by using lower level perceptual mechanisms to identify location that do not use as much working memory and attention.

Network-based vehicle communications

Phone

The phone application provides speech communications with other people connected to the telephone network. SADs make it easier to understand speech in noisy environments, and reduce driver distraction. They can also be used for multi-party conference calls to localize each party from a different location. This makes it easier to determine who the current talker is.

Navigation

The Navigation application uses GPS to provide turn-by-turn route guidance. Auditory displays are important because they allow drivers to keep their eyes on the road and are more effective at vigilance tasks [4] such as alerting the driver to a route manoeuvre. SADs improve the user experience over a standard auditory display in several ways:

- They make it clearer where the route begins

- They provide additional information on the location of the next manoeuvre
- They provide more information about the manoeuvre that is to be executed such as trajectory information
- When the route is left they allow drivers to more quickly determine how to get back on route

Vehicle-to-internet

The NG car will have access to the internet. SADs can improve the user experience by causing web-based speech and audio to sound like it is coming from the associated visual display.

Information services (e.g., weather, stocks, etc.)

The Information services application provides general information of interest to the driver. This may include weather, stocks, or vehicle information (e.g., trip ETA, tire pressure, etc.). SADs improve the user experience by localizing informational prompts at the associated visual display.

Route traffic avoidance

The Route traffic avoidance application identifies traffic congestion on scheduled or predicted routes and alerts drivers. SADs improve the user experience by communicating the location of congestion using prompts that are localized where the congestion is. This requires less effort on the part of the driver and reduces driver distraction.

Spatial cues used by listeners

SADs work by delivering spatial acoustic cues to users that are used by the human auditory system to localize sounds. Therefore, knowledge of the acoustic cues used by listeners is important when implementing a spatial auditory display. An adequate discussion of this topic is beyond the scope of this paper. However, below is a list of such acoustic cues:

- Onset disparities
- Inter-aural Time Differences (ITDs)
- Inter-aural Level Differences (ILDs)
- Spectrum
- Overall level
- Direct-to-reflected sound ratio
- Deltas from reference objects

Different types of SADs

Localization accuracy of the SAD will depend on its implementation. Below is a list of different vehicle audio system implementations in descending order of accuracy:

- Large array of loudspeakers
- Headphones or two loudspeakers with cross-talk cancellation
- Surround sound system (e.g., 5.1 audio)
- Stereo signals with balance and fade control
- Mono signal with balance and fade control
- Single loudspeaker using monaural cues

The signal processing used with the SAD is just as important, if not more important, than the hardware configurations described above in providing accurate rendering of target positions.

Conclusions

The explosion of new applications in NG cars will increase the complexity of user interfaces. This increased complexity can confuse and frustrate users if user interfaces are not properly designed. SADs will prove to be an important feature of NG cars because they reduce driver distraction, increase task performance, and increase ease of use. In essence, SADs improve the user experience!

References

[1] Begault, D. R., Pittman, M. T. (1996). Three-Dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays. *The International Journal of Aviation Psychology*, 6(1), 79-93.

[2] Levitt, H. and Rabiner, L. R. (1967). "Binaural release from masking for speech and gain in intelligibility". *Journal of the Acoustical Society of America*, 42, 601-608.

[3] Pennock, S., Hetherington, P. (2009). Wideband Speech Communications: the Good, the Bad, and the Ugly. *Audio Engineering Society 36th International Conference,* 2-4 June 2009, Dearborn, Michigan, USA.

[4] Colquhoun, W. P. (1975). Evaluation of auditory, visual, and dual-mode displays for prolonged sonar monitoring in repeated sessions. *Human Factors*, 17, 425-437.