8

Navigating Robot Sonification: Exploring Four Approaches to Sonification in Autonomous Vehicles

Richard Savery, Anna Savery, Martim S. Galvão and Lisa Zahray

DOI: 10.1201/9781003320470-8

8.1 Intro

Autonomous cars are one of the most rapidly growing areas of robotics, and may be the first robots to be fully integrated into society [5]. Autonomous cars are also one of the few robotic platforms that have clear applications in personal, industrial and commercial settings. For these reasons, deep consideration of how we interact with autonomous cars is crucial, with extensive research already existing, focusing on areas such as ethics [14], communication methods [20], environmental impact [19] and many others.

Sound has played a significant role for enhancing the driving experience and improving safety. This can be through roles such as safety alerts, including seat belt warnings, or notifications for crossing lanes. More recently, sonification of car data has emerged as a growing area of research, due to rich layers of available real-time data and growth in production of autonomous vehicles [15].

In this chapter, we will explore the concept of sonification and its application in the context of autonomous cars, examining the benefits, challenges, and future directions of this emerging field. We present four different approaches, three using simulations of cars and one creating a sonification from a real world dataset. Our goal was to describe divergent approaches and cover a range of future possibilities when sonifying car data.

The first project, Unity Algorithmic Music, utilizes the Unity game engine to create a three-dimensional driving simulator on a dirt road with surrounding grass regions and mountains. The sound design is based on the cycle of fifths chord progression in C major, with two instruments assigned to the chords and the subdivision speed determined by the car's speed. The second project, Unity Deep Learning, uses the same Unity engine for car simulation, but instead uses a bidirectional Long short-term memory (LSTM) network, to map movements from the car to melodies, all performed through a virtual string orchestra. The third project, Unreal MetaSounds, is built in the Unreal Game Engine and employs Unreal's MetaSounds framework to create a data-driven synthesizer whose parameters are determined by the car's relationship to its environment. The final project, Sonification with Beads, uses the Berkley Deep Drive dataset to process violin samples and create compositions based on user interactions with the dataset.

Autonomous car sonification provides a unique opportunity to convey data and driving information to passengers. While the four projects presented in this chapter showcase different approaches to sonification, they collectively demonstrate the potential for sonification in the context of autonomous cars. More broadly, this chapter describes the potential for sonification in robotic systems, and how data taken from robot systems can be used to create compelling auditory feedback.

8.2 Related Work

8.2.1 Sonification

Sonification focuses on the use of sound to represent data or information. It is often described as the us of "nonspeech audio to convey information... for purposes of facilitating communication or interpretation" [11]. In Walker's theory of sonification, they describe four key purposes for sonification: alarm, status, art and entertainment, and data exploration [23]. Alarm sonifications can be used in many settings, such as to report patient heart and oxygen rates [10], or monitoring business processes in factories [9]. Sonification for status updates can focus on supplying background information, such as network traffic flow [3]. For art and entertainment, sonification can be used for musical installations [21], or to augment sports, such as backing tracks for soccer [17]. Data exploration sonification can aid both novice analysis of data [22] or provide new insights to those already familiar with data [8].

8.2.2 Robot Sonification

Robot sonification involves translating robot movements, sensor data or other robot data sources into audio feedback. This sonification is used for a range of purposes, including all categories from Walker's theory of sonification. For example, robot sonification has been used to improve social communication for children with Autism Spectrum Disorder (ASD), incorporating sonification to foster emotional and social communication [26]. Frid et al. conducted two experiments to investigate the perception of mechanical sounds produced by expressive robot movement and blended sonifications thereof. They found that blended sonification can successfully improve communication of emotions through robot sounds in auditory-only conditions [7]. Zahray et al. designed and evaluated six different sonifications of movements for a robot with four degrees of freedom to improve the quality and safety of human–robot interactions, recommending that conveying information in a pleasing and social way is important to enhance the human–robot relationship [25]. Other research has investigated the perception of synthesized robot sounds and the materiality of sound made by a robot in motion. This research found that participants preferred more complex sound models for the sonification of robot movements and that sound preferences varied depending on the context in which participants experienced the robot-generated sounds [13].

8.2.3 Autonomous Vehicle Sonification

A recent paper listed multiple key challenges and opportunities for the expanding use of autonomous car sonification. These challenges focus on the user experience design of highly automated cars, including avoiding motion sickness, ensuring trust, supporting accurate mental models, and providing an enjoyable experience. The authors propose that auditory displays could help address these issues, and suggest that continuous sonic interaction may be more effective than traditional discrete cues [12]. Specific studies have focused on a range of issues, one study aimed to investigate whether auditory displays can enhance users' trust in self-driving vehicles. Results suggest that the designed auditory display can be useful in improving users' trust and received high scores in terms of user acceptance, providing implications for the interaction design of self-driving cars and guiding future auditory display research [6]. Other research has described potential applications for monitoring passenger breathing to enhance the driving experience [16].

8.3 Unity Algorithmic Music

The project described in this section uses Unity as a game engine to create a 3D world from the perspective of a car driver, including a dirt road, grassy areas, and mountains. C# scripts control the game logic, allowing for acceleration, deceleration, and gear changes using key presses. The game sends OSC messages to a MaxMSP patch, which sends MIDI notes to Ableton to produce sounds. The sound mapping follows a cycle of fifths chord progression in C major, changing chords every measure at a constant tempo of 90 BPM in 4/4 time. The car's speed controls the subdivision speed of the chords, with faster speeds resulting in faster subdivisions. Turning the car or pressing the brakes produces additional musical effects, with a total of four instruments used in the sound design.

8.3.1 Technical Details

The project uses Unity as a game engine, rendering a three dimensional world viewed from the perspective of the driver's seat of a car. The terrain includes a dirt road surrounded by bumpy grass regions and mountains. During the videos, the car drives over both the road and grass regions. An example frame of the environment is shown in Figure 8.1.

Scripts controlling the game logic were coded in C#. The car accelerates while the gas key is held until reaching a maximum speed, and loses speed slowly when the gas key is not pressed until reaching zero speed. While the brakes key is pressed, the car loses speed quickly until reaching zero speed. Gravity affects the acceleration of the car, i.e. the car will roll down a hill in a case where there are no brakes or gas counteracting it, and accelerate slower up a steep hill when gas is pressed. The gear of the car can be toggled by a key press between reverse and drive. The left and right arrow keys turn the car in the corresponding direction, and the steering wheel visually turns to match.

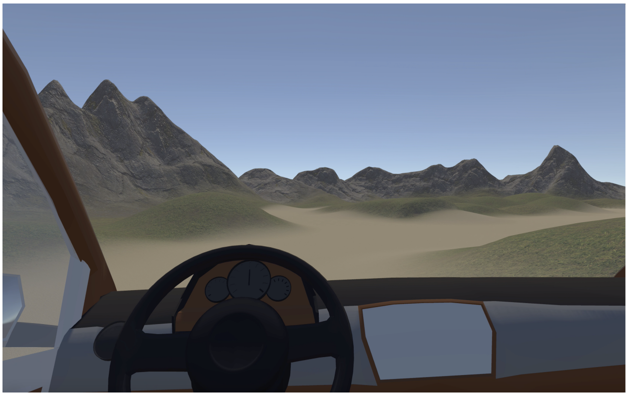

FIGURE 8.1
Video frame from unity project.

The Unity game sends OSC messages to a patch coded in MaxMSP, a visual programming language for music, that details the speed, gear, turn direction, pressed state of the brakes, and pressed state of the gas. The MaxMSP patch processes these inputs and sends MIDI note messages to Ableton, a digital audio workstation, which produces the sounds. The Ableton project contains four instruments customized from the "Ambient and Evolving" presets.

8.3.2 Sound Mapping

The sounds during game-play follow the circle of fifths chord progression in C major, changing chords every measure at a constant tempo of 90 BPM in 4/4 time. Two instruments are assigned to the chords. One instrument plays the bass note of each chord, sustaining it for the full measure. The other plays randomized notes from two octaves of the chord (a choice of six notes total) at four different possible subdivision speeds: sixteenth notes, eight notes, quarter notes, or half notes. This subdivision is determined based on the speed of the car, where a faster car speed maps to a faster subdivision speed. The instruments were designed so that, when the car is in drive, the sustaining instrument has a longer note release time and more reverb than the instrument playing the subdivided notes. When the car is in reverse, these two instruments are swapped, giving the more sustained and reverb-containing instrument the faster-moving part.

Each time the steering wheel is used, a quarter note is played on a third instrument. A left turn plays a low octave C, and a right turn plays a higher octave C. While the brakes are pressed, a fourth instrument plays a sustained note. The brake instrument's sound was designed to be less strongly pitched, acting more as a stutter effect on top of the more musical sounds of the other instruments.

8.4 Unity Deep Learning

The next approach to robot sonification focused on using the same Unity simulation from the previous section with deep learning musical generation. A primary goal was to create varying, aesthetically pleasing music, while still representing the data from the car. For this reason we chose to map features from the car movement to broader musical features such as tempo and pitch contours.

8.4.1 Technical Details and Sound Mapping

At the basis of this sonification approach is a bidirectional LSTM to generate musical patterns. The network is inspired by Piano Genie [4], a neural network model that maps 8 buttons to 88 piano keys in a way that allows non-experts to play music by simply pressing combinations of the buttons. Both Piano Genie and our revised model were trained on the International Piano-e-Competition dataset, a collection of virtuoso performances on piano. Our model converts numbers in the range of 1 to 8 into MIDI sequences, where the important learning in the system is the contour and relation between each number. Constantly rising inputs from 1 to 8 will lead to rising patterns, and multiple notes input at the same time will result in chords.

With this in mind our sonification process mapped different elements of the data to the range of 1 to 8, which created the pitches used in the piece, while the rhythms were placed based on events in the data. Table 8.1 shows each mapping. Tempo for the sonification is mapped directly to the speed of the car, with the range of the tempo between 80 and 140 BPM. To play back sounds we use Kontakt 7, and the 8Dio Century Strings virtual instruments, which provides a range of violin, viola, cello and double bass samples.Throughout the sonification there is an underlying accompaniment, varying between single quarter notes for parked, or chords when in drive and reverse. For park, drive or reversing different timbres are used for the strings. Park mode uses arco string staccato, while drive uses arco string legato. Reverse plays the same harmonic and rhythmic material as drive however uses string tremolo for playback. When the car is turning left or right, melodies are added in pizzicato strings, with the melody ascending for left and descending for right.

8.5 Unreal MetaSounds

The project described in this section engages with autonomous vehicles at the world-modeling stage, using dynamic audio processes as a means for conveying

TABLE 8.1
Summary of unity deep learning mappings.

Parameter	Output
Speed	Tempo, mapped between 80 to 140 BPM
Park	Play Park Music, Change Timbre
Drive/Reverse	Play Drive/Reverse, Change Timbre
Left/Right	Play Melody Sequence
Brakes	Double Notes on Flute

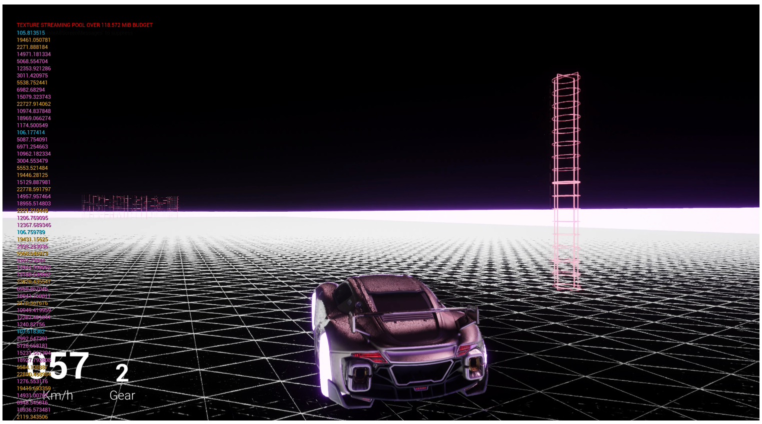

FIGURE 8.2
Video frame from unreal engine project.

information about the vehicle's environment to a human user/driver. Rather than replacing other streams of vehicle information, the process of sonifying the vehicle's relationship to its environment serves as a novel interface for audio synthesis, with parameters such as orientation, speed, and distance to obstacles mapped to synthesis parameters such as frequency, delay time, and amplitude. As the driver navigates the virtual space, a collection of sounds indicates both the vehicle's relationship to other objects as well as key pieces of information regarding the vehicle's current state. Some of these sonifications are constant (e.g. vehicle orientation and speed) while others occur only when corresponding events trigger them (e.g. approaching a wall or crossing the north-south axis). Figure 8.2 shows a screen capture of the Unreal gameplay.

8.5.1 Technical Details and Sound Mapping

The project employs Unreal Engine 5 to construct a virtual environment in which a test vehicle is driven around a simple obstacle course. Using MetaSounds (see Figure 8.3), Unreal Engine's built-in audio engine, several

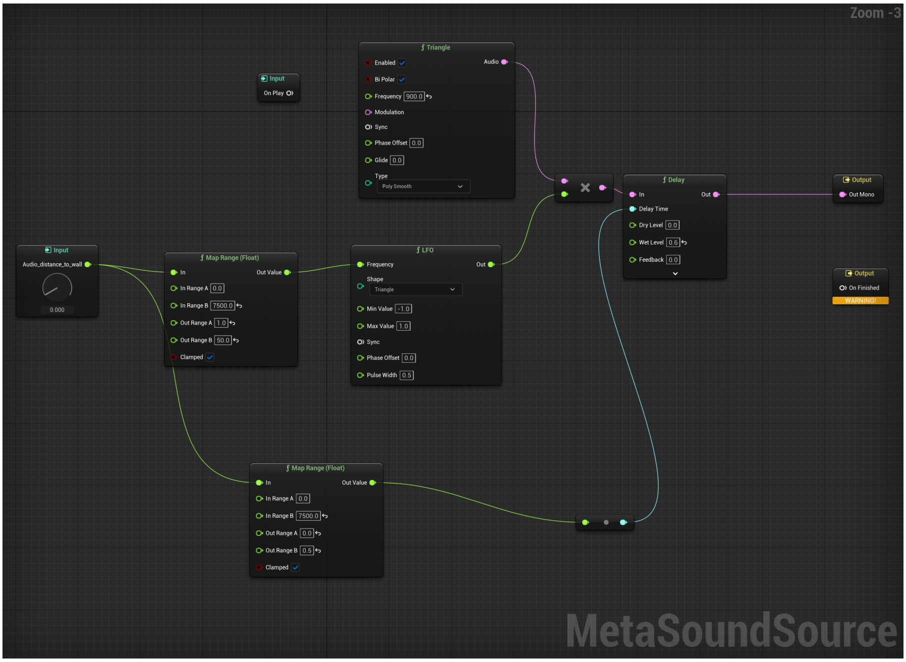

FIGURE 8.3
MetaSoundSource file from unreal engine project.

audio mappings are created to convey information about the player's vehicle and its relationship to the virtual environment:

8.5.1.1 Vehicle Orientation

North and South Markers: Two distinct sonic cues are used to signal that the driver/vehicle has crossed north or south during a rotation. The cue indicating a northern crossing is played an octave higher than the cue for the southern crossing.

Present Heading: Two constantly playing sine wave oscillators are used to indicate the current heading of the vehicle. Both oscillators play at the same frequency whenever the vehicle is facing due north or south. As the vehicle steers away from the north-south axis, the frequency of one of the oscillators changes accordingly. The resulting difference creates a beating pattern, which is used to indicate how far the vehicle has veered from the north-south axis.

8.5.1.2 Vehicle Speed

A square wave oscillator is used to modulate the amplitude of a triangle wave oscillator, which has its frequency mapped to the current speed of the vehicle.

As the vehicle accelerates, the frequency of the triangle wave rises, as does the modulation rate of the square wave. The result is a sound that is timbrally distinct from other indicators, and which the player can reference alongside orientation to determine vehicle speed and heading.

8.5.1.3 Object Identification

Two basic object types are defined in this prototype: walls and poles. Each wall is identified by a triangle wave oscillator set to a fixed frequency while each pole is identified by a pulsing burst of white noise.

8.5.1.4 Object Localization

Spatial: Walls and poles are treated as sound-emitting objects in Unreal Engine, with spatialization employed to place them at the corresponding location in the player's listening field. As the player changes the orientation of the vehicle, the perceived location of walls and poles changes accordingly. Likewise, the vehicle's distance from each sound-emitting object determines the attenuation of the sound source according to a falloff curve that extends out from the boundaries of the object. In the case of walls, this means that attenuation of the source follows a capsule-shaped spread rather than a spherical spread originating from the center point of the object. The falloff is designed to provide drivers with enough notice to avoid striking objects while also maintaining some degree of isolation from nearby objects.

Distance and acceleration: The distance between the vehicle and an object is mapped to both a low-frequency oscillator (modulating the amplitude of the wall's triangle wave oscillator) and the delay time of a delay module placed after the modulated triangle wave. When the vehicle is moving toward a wall, the delay time decreases, which leads to an increase in the perceived frequency being emitted by the wall. The opposite is true as the vehicle moves away from a wall, creating a Doppler effect from the player's perspective. The low-frequency oscillator likewise follows a direct relationship between frequency and distance between the vehicle and a wall. When the vehicle is positioned immediately adjacent to a wall, for example, the sound is heard as distinct pulses, whereas when the vehicle moves away from the wall the frequency of the amplitude modulation increases until the sound is no longer distinguishable as pulses.

8.6 Sonification with Beads

Computer vision and synthesized audio in autonomous vehicles are implemented to serve specific purposes, such as obstacle detection and user feedback. This project intends to employ data from autonomous vehicles within the scope of an interactive visual installation, where meaning is created through direct, real-time manipulation of audio samples as well as visual representation of the Berkley Deep Drive dataset. This is done using Processing software as well as an external library called Beads, which allows for sonification of the images.

8.6.1 Technical Implementation

The Berkley Deep Drive (BDD100K)[1] provides an extensive driving video and image dataset for studying and training autonomous vehicles within a wide range of diverse driving conditions [24]. The purpose of the construction of the 100k dataset was to improve on the limitations of existing visual content used for multitask learning for autonomous driving correlating to real-world computer vision applications required to execute tasks of various complexities [24].

This robust dataset, collected from over fifty thousand rides across New York, San Francisco Bay Area, and other regions [1] contains 40 second videos of diverse driving conditions with varying times of day and weather, ranging from city streets, residential areas and highways. "The frame at the 10th second of each video is annotated for image classification, detection, and segmentation tasks" [1]. These segmenations include; "lane marking, drivable area, full-frame semantic and instance segmentation, multiple object tracking, and multiple object tracking with segmentation" [24].

The concept for the author's project stemmed from a desire to provide humans a way to better understand computer vision in an interactive and artistic way. A single image from the dataset was used to explore the relationship between static images and dynamic interplay through data sonification and visual interaction. Only object bounding boxes were explored as a data source due to their simple coordinate system represented as floating point numbers, which were accessible for both, visual representation and audio synthesis.

The author chose to work with processing[2] as a software environment and the external Beads library[3] for sonifying the data. The audio is made up of five solo violin samples improvised and recorded by the author, reflecting the mood of the image. A looping bass line plays with the initiation of the processing sketch. The other samples respond to user cursor movement across the sketch display window. Five object bounding boxes (*box2d* object in JSON) object coordinates were chosen for visual display. When the user's cursor interacted with each "box", an audio sample assigned to those coordinates is initiated and manipulated based on the mouse movements (see Figure 8.4).

[1] https://www.bdd100k.com/
[2] https://processing.org/
[3] http://www.beadsproject.net/

FIGURE 8.4
User interaction with bounding boxes.

8.6.2 Beads

In order to facilitate playing multiple audio samples concurrently and one at a time, five sample players were created, one for each audio track, with all sample players continuously playing on a loop. A Glide object was then used to control the Gain values of each sample player. Since the bass line audio sample was not manipulated by the user, there was no need to create a Glide object to control its Gain value. When the user's mouseX and mouseY coordinates entered any of the displayed rectangle (object bounding box) values, the gain value of the corresponding audio sample would change from 0.0 to 0.5 and then back to 0.0 once the cursor went outside those values. Figure 8.5 shows an extract of Processing code using Beads library.

The x and y coordinates of each bounding box were then used to manipulate each corresponding track in real-time, using cursor movement. The first bounding box (x1: 751.147935, y1: 267.019365, x2: 771.111999, y2:285.735677) calculates the relationship between current vertical mouse location (mouseY) and previous vertical mouse location (pmouseY). When the cursor is moved downward within the boundaries of that rectangle, the audio sample is played in reverse with the playback rate slowed down by 285.735677 divided by the current vertical mouse coordinates (y2/mouseY, or box2d[3]/mouseY). This is shown in Figure 8.6.

Another bounding box shared coordinates with two other boxes. This overlapping data was sonified by triggering three audio samples simultaneously, representing the act of three distinct boundaries crossing and interacting.

```
  if (mouseX > box2d1[0] && mouseX < width-10 && mouseY > box2d1[1] && mouseY < box2d1[1]+box2d1[3]) {

  if (pmouseY < mouseY) {
    player2.setLoopType(SamplePlayer.LoopType.LOOP_BACKWARDS);
    rateValue.setValue(box2d2[3]/mouseY);
  } else
  {
    rateValue.setValue(1);
  }
  tint(200);
  println(box2d1[0]);

  gainGlide2.setValue((float) 1.);
} else
{
  gainGlide2.setValue((float) 0.);
}
```

FIGURE 8.5
Extract of Beads code in processing.

```
if (pmouseY < mouseY) {
  player2.setLoopType(SamplePlayer.LoopType.LOOP_BACKWARDS);
  rateValue.setValue(box2d2[3]/mouseY);
```

FIGURE 8.6
Real-time audio processing using BDD data.

Rectangle x1:680.025953, y1:283.240168, x2:696.246757, y2:301.956478 was used with a granular sample player, allowing the user to turn the assigned audio track into grains of layering looping sounds (see Figure 8.7).

8.6.3 Visual representation

The chosen background image from the JSON dataset was 4bdd2193-546f8f2c.jpg. To avoid visual clutter, only the first four bounding boxes for that image were used. In order to provide visual feedback for the user, unfilled rectangles representing the object bounding boxes were drawn, with data directly from the JSON file appearing within each rectangle as the user entered their coordinates. A convolution effect [18] was used to alter the background image whenever the cursor was within the boundaries of any of the rectangles to signify the concept of clarity of focus on the context within the bounding box and blurriness of the peripheral vision outside its boundaries.

```
if (pmouseY < mouseY || pmouseX < mouseX) {
  grainSizeValue.setValue((float)box2d3[0]+box2d3[1]+box2d3[2]+box2d3[3]);
}
```

FIGURE 8.7
Granular processing using all four bounding box coordinates.

In this project, the author focused on sonifying only one parameter available within the extensive BDD dataset within a single image out of the available 100k. The coordinates of the object bounding boxes provided an abundance of room for musical and user interaction creativity. Building on the ideas explored in this project, a further expansion on the possibilities of sonifying the robust data from the BDD dataset could lend itself to a potentially large scale series of interactive works exploring the relationship between computer vision, sound synthesis, autonomous car training data and people's perception of this concept.

8.7 Discussion

8.7.1 Balancing Aesthetics with the Conveyance of Information

With any effort at sonifying data, there is inevitably some degree of information loss. Whether using vehicle speed to control an oscillator's frequency or mapping steering wheel movements to tempo and pitch contours, abstraction is a necessary step for translating sensor information into the audio domain. Even if the resulting mapping is direct (e.g. high speeds mapped to high frequencies), it is impossible to determine the exact speed of a vehicle by listening to an oscillator's output. Rather, we can follow the shape or outline of the speed over time, and in doing so gain an abstracted understanding of a vehicle's speed via sound.

This balance between fidelity and intelligibility can also be thought of as a trade-off between conveyance of information and aesthetics. An ideal sonification effort would maximize the amount of information conveyed while avoiding over-saturating the user's sensory input with a deluge of information. According to Csikszentmihalyi, the upper limit of human informational bandwidth is around 120 bits per second, or about double the amount of attention required to carry on a conversation [2]. To avoid exhausting the user's information processing faculties, a successful sonification effort would need to strike a balance between conveying the maximum amount of relevant information when necessary and allowing the user to relax their attention when possible.

Planned future development of the Unreal MetaSounds project includes incorporating this kind of attentional bandwidth management system through a background/foreground sonification model. Drawing from ambient music compositional strategies, this model will highlight important changes in sensor data through distinct sonic cues while otherwise relegating sonified vehicle data to a background process that is still present but only at the attentional periphery of the operator. In such a model, the trade-off between aesthetics and the conveyance of information is handled on a moment-by-moment basis, wherein normal operation is signified by a highly abstracted and attentionally undemanding soundscape while important events trigger a momentary shift in the sonification process to convey the maximum amount of relevant information as quickly as possible.

8.7.2 Adding Liveliness to Static Imagery

With the availability of such rich datasets as the BDDK, translating non human elements such as computer vision and robot sounds into an artistic work can provide new human insights into the world of robotics and machine learning. Although exploring each of the 100,000 images within the dataset is perhaps an unreasonable task, a set of multiple images across a variety of terrains, driving conditions and times of day would illustrate a representation of how a computer sees the world and what is significant for an autonomous vehicle. Multiple parameters from the data could be translated onto the static images in an interactive way, using both sonification and visualization of the accompanying data. In this way, a person interacting with the visual installation could hover over any area of an image seeing it react in a way that represents the robot car. Such approaches could be explored through mapping data to changes in volume or playback speed, depending on which elements of the image are being physically explored. Perhaps an obstacle could trigger a rise in volume or a night time drive would illuminate the road. Another approach would be to use multiple instances of the same image, focusing on specific objects within the data, such as lane markings or traffic lights. Finally, relating back to the entirety of the data, if such a task could be automated, then this would further bridge the gap between human and machine.

8.8 Conclusion

This chapter has explored four approaches to sonification of autonomous car data. While there are many other areas of research and potential mappings of car data to audio, our goal has been to describe just some of the possibilities of sonification and robotics broadly. Further research in this area can lead to the development of more sophisticated sonification systems that can adapt to changing driving conditions and driver preferences. While these systems may be used to augment a driver's awareness of their vehicle's surroundings, they may simultaneously function as generative compositional models that employ processes of abstraction to transform raw data into higher-level signifiers. In the case of the latter, it is possible to imagine a musical model in which both direct sensor inputs and sequences of events at the vehicle level may be mapped to musical gestures, rendering a vehicle's data streams as parametric inputs to an algorithmic composition system.

Likewise, the incorporation of autonomous vehicle training data sets into musical works offers the opportunity for human intervention into what is otherwise a largely machine-oriented endeavor (albeit one often aimed at transporting human passengers). By making the images and sensor information contained in these data sets the primary source material for interactive musical works, the performer and composer are reintegrated into fold. In doing so, they reiterate the uncertainty and improvisatory nature through which autonomous vehicles make sense of the world.

Bibliography

[1] Pedro Azevedo. BDD100K to YOLOv5 Tutorial: https://medium.com/@pedroazevedo6/bdd100k-to-yolov5-tutorial-213e4a67d54b.

[2] Csikszentmihalyi, M. *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi.* Springer Netherlands, 2014.

[3] Debashi, M., and Vickers, P. Sonification of network traffic flow for monitoring and situational awareness. *PloS One 13*, 4 (2018), e0195948.

[4] Donahue, C., Simon, I., and Dieleman, S. Piano genie. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (2019), pp. 160–164.

[5] Dos Santos, F. L. M., Duboz, A., Grosso, M., Raposo, M. A., Krause, J., Mourtzouchou, A., Balahur, A., and Ciuffo, B. An acceptance divergence? media, citizens and policy perspectives on autonomous cars in the european union. *Transportation Research Part A: Policy and Practice 158* (2022), 224–238.

[6] Fagerlönn, J., Larsson, P., and Maculewicz, J. The sound of trust: sonification of car intentions and perception in a context of autonomous drive. *International Journal of Human Factors and Ergonomics 7*, 4 (2020), 343–358.

[7] Frid, E., and Bresin, R. Perceptual evaluation of blended sonification of mechanical robot sounds produced by emotionally expressive gestures: Augmenting consequential sounds to improve non-verbal robot communication. *International Journal of Social Robotics 14*, 2 (2022), 357–372.

[8] Grond, F., and Hermann, T. Interactive sonification for data exploration: How listening modes and display purposes define design guidelines. *Organised Sound 19*, 1 (2014), 41–51.

[9] Hildebrandt, T., Mangler, J., and Rinderle-Ma, S. Something doesn't sound right: Sonification for monitoring business processes in manufacturing. In *2014 IEEE 16th Conference on Business Informatics* (2014), vol. 2, IEEE, pp. 174–182.

[10] Janata, P., and Edwards, W. H. A novel sonification strategy for auditory display of heart rate and oxygen saturation changes in clinical settings. *Human Factors 55*, 2 (2013), 356–372.

[11] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., et al. The sonification report: Status of the field and research agenda. report prepared for the national science foundation by members of the international community for auditory display. *International Community for Auditory Display (ICAD), Santa Fe, NM* (1999).

[12] Larsson, P., Maculewicz, J., Fagerlönn, J., and Lachmann, M. Auditory displays for automated driving-challenges and opportunities.

[13] Latupeirissa, A. B., Panariello, C., and Bresin, R. Probing aesthetics strategies for robot sound: Complexity and materiality in movement sonification. *ACM Transactions on Human-Robot Interaction* (2023).

[14] Lin, P., Abney, K., and Jenkins, R. *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence.* Oxford University Press, 2017.

[15] MacDonald, D. Designing adaptive audio for autonomous driving: an industrial and academic-led design challenge. Georgia Institute of Technology.

[16] Morimoto, Y., and van Geer, B. Breathing space: Biofeedback sonification for meditation in autonomous vehicles. Georgia Institute of Technology.

[17] Savery, R., Ayyagari, M., May, K., and Walker, B. Soccer sonification: enhancing viewer experience. In *International Conference on Auditory Display (25th: 2019)* (2019), Georgia Tech, pp. 207–213.

[18] Shiffman, D. *Learning Processing: A Beginner's Guide to Programming Images, Animation, and Interaction.* Morgan Kaufmann, 2009.

[19] Silva, Ó., Cordera, R., González-González, E., and Nogués, S. Environmental impacts of autonomous vehicles: A review of the scientific literature. *Science of The Total Environment* (2022), 154615.

[20] Thorvald, P., Kolbeinsson, A., and Fogelberg, E. A review on communicative mechanisms of external hmis in human-technology interaction. In *IEEE International Conference on Emerging Technologies and Factory Automation* (2022).

[21] Tittel, C. Sound art as sonification, and the artistic treatment of features in our surroundings. *Organised Sound 14*, 1 (2009), 57–64.

[22] Tünnermann, R., Hammerschmidt, J., and Hermann, T. Blended sonification–sonification for casual information interaction. Georgia Institute of Technology.

[23] Walker, B. N., and Nees, M. A. Theory of sonification. *The Sonification Handbook* (2011), 9–39.

[24] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., and Darrell, T. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 2636–2645.

[25] Zahray, L., Savery, R., Syrkett, L., and Weinberg, G. Robot gesture sonification to enhance awareness of robot status and enjoyment of interaction. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (2020), IEEE, pp. 978–985.

[26] Zhang, R., Jeon, M., Park, C. H., and Howard, A. Robotic sonification for promoting emotional and social interactions of children with asd. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts* (2015), pp. 111–112.