SOUNDSCAPE MAPPING: A TOOL FOR EVALUATING SOUNDS AND AUDITORY ENVIRONMENTS

Iain McGregor, Gregory Leplâtre, Phil Turner and Tom Flint

Edinburgh Napier University,
School of Computing,
Edinburgh, United Kingdom
{i.mcgregor, g.leplatre, p.turner, t.flint}@napier.ac.uk

ABSTRACT

This paper describes a soundscape mapping tool, and provides an illustration of its use in the evaluation of an in-car auditory interface. The tool addresses three areas: communicating what people are listening to, showing how soundscapes can be visualized, and demonstrating how the approach can be used by a designer during the evaluation of an auditory display. The strengths and limitations of this approach are discussed and future work identified.

1. INTRODUCTION

Robare and Forlizzi [1] highlighted a 'lack of design theory' with regard to guidelines for sound design within computing. Despite the dramatic increase in the number of products which replay sound in the last ten years or so, there has been relatively little improvement in how we understand the listeners' experiences. Designers need to consider the context of use, as applications might be used in a wide variety of environments [2]. While product design often explores listening [3], the same cannot be said of the development of auditory displays. This is due, in part, to the relative paucity of formal techniques to measure a design's impact. Available techniques are limited to simple noise pollution measurements [4], the elicitation of interpretations from listeners [5], and 'object-orientated' descriptions [6]. The soundscape mapping tool we present here is designed to evaluate auditory displays in their intended context of use [7]. Our empirical approach to the evaluation of these displays is to study them *in situ* by first, eliciting people's auditory experiences; and then visualising these soundscapes for ease of comparison.

This paper reports the illustration of the soundscape mapping tool through the evaluation of an in-car auditory display. Our interest in evaluating this audio-only interface was in understanding the effect of different auditory contexts on its effectiveness. A small car was chosen as it represented a contained environment that travelled through more complex external auditory environments. In order to provide a consistent experience for all of the participants it was tested in a simulated environment:

- travelling through a busy city centre at rush hour with speech radio playing.
- travelling through a busy city centre at rush hour with both speech radio playing and the auditory display.
- stationary in a quiet location with only the auditory display.

By comparing the findings from these three different contexts we can be confident that both the method and tool are reliable and robust and that they yield ecologically valid results.

2. METHOD

We created a tool for the classification and visualization of soundscapes, that can be used during the evaluation of augmented auditory environments. This tool is based on the results of a series of three studies. The first was an experimental elicitation of concurrent verbalizations by 40 listeners where listeners were asked to describe their auditory environment. The responses were transcribed and coded in order to discover which attributes were important to listeners when describing sound [8]. The second was a questionnaire survey completed by 75 audio professionals where they described the attributes of sound that were important to sound designers [9]. The third study was a soundscape mapping tool based on published methods where 18 listeners' experiences of a shared auditory environment (open–plan office) were compared. The tool was used to represent the experiences of individuals, as well as subsets of users (regular, intermittent and new) of the workspace [10]. The version of the soundscape mapping tool reported here has three distinct phases: capture, classification, and visualization.

- Capture involved the creation of a schematic of the car, recording the sound field, and transcribing the sound events directly from the surround sound recording.

- Classification was conducted by the participants who first listened to the recordings, and then were questioned about the audible attributes of a series of sound events.
- Visualisation involved the creation of a series of annotated soundscape maps based around the physical context of the study.

2.1. Capture

A 20:1 schematic of the car was created, cells were added to the perimeter to facilitate the annotation of external sound events (see Figure 1). A fifteen minute recording was made of the car driving through the city centre, in order to create a consistent soundfield for listeners. This recording was made using a custom eight-channel surround system and then augmented during the experiment with the auditory interface. Eight omni-directional microphones were affixed in suspension mounts inside the car, at approximately head height, and fed into four DAT recorders (see Figure 2).

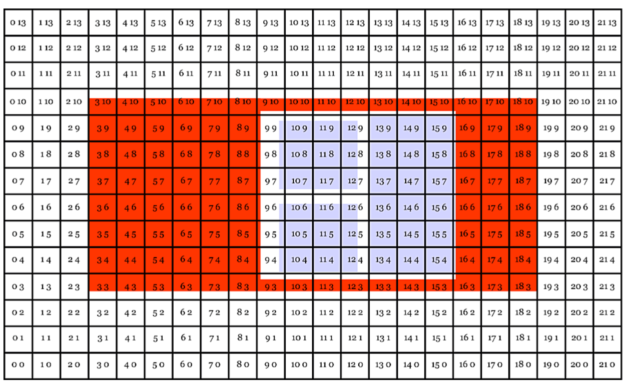


Figure 1: Simplified aerial view of car with grid, red = bodywork, blue = seats

Calibration was achieved by a method borrowed from the film industry, the driver first read off the display of an SPL meter located on the passenger seat illustrating the slow sensitivity peak dbC level, then the driver clapped their hands. The short peak acted as the starting point for the recording, allowing all 8 tracks to be synchronized during the capture process. A handclap by the driver completed the recording, this confirmed whether any of the tracks had drifted during the time period. Each track was subsequently transferred to a Pro Tools LE system, in order to provide a consistent auditory backdrop for the auditory interface.

Sound event transcription included source, action, start time, end time and location. Table 1 contains examples of these. Location was calculated using the perceived central point from the surround sound recording, and notated using x-y coordinates according to the grid. If a sound event moved in relation to the car, the start and end points were documented. Start and end times were also established from the recording, these were rounded down to the nearest second within which the event occurred. In order to reduce the number of events which listeners had to classify, sound events which had the same source, action and location were grouped together.

Figure 2: Microphone placement prior to final positioning and calibration, for surround sound recording

All of the captured material was passed to the designer (the second author) so that he could create the auditory display. The designer decided to limit the interface to only three auditory warnings to reduce the cognitive load on the listener. After creating the design he overlaid the new sounds on to the eight channel surround sound recording. This allowed him to control the level, incidence, duration and (perceived) spatial location of each warning. The designer also provided a written description of the different auditory warnings for the listeners' reference. These warnings included, braking distance, dead angle and email message. This final augmented version of the surround sound recording was then split into three versions, one for each simulated environment.

Event	Source	Start	End	x	y	x	y
Engine Idle	Engine	00:00:00	00:00:21	6	6		
Male Speech	Radio	00:00:00	00:00:13	11	8		
Windscreen Wiper	Car	00:00:02		21	9		
Passing Car	Car	00:00:04	00:00:06	9	13	21	13
Siren	Ambulance	00:00:06	00:00:08	21	12		

Table 1: Example sound event transcription

2.2. Classification

A classification was created based on the findings from previous studies [8, 9, 10]. Table 2 holds these ten distinct attributes each with three options. The first six attributes were derived directly from the comparison between audio practitioners and listeners. In the case of *type* rather than specify whether a source was natural or artificial, choices were confined to speech, music or sound effect, with the last representing all sounds which are neither speech or music. *Material* relates to the substance which gives rise to the sound, either gas, liquid or solid, whilst the *interaction* specifies the nature of the sound's generation whether it was impulsive, intermittent, or continuous. *Temporal* reflects the total length of the sound event (short, medium or long) separate to its interaction; *spectral* applies to its pitch (high, mid and low); and *dynamics* to its volume (loud, medium or soft).

Type	Category
Speech	Spoken language
Music	Performed composition
Sound effect	Audible events and actions
Material	Matter
Gas	Airborne
Liquid	Fluids
Solid	Objects
Interaction	Action
Impulsive	Explosion/drip/impact
Intermittent	Whooshing/splashing/scraping
Continuous	Blowing/flowing/rolling
Temporal	Duration
Short	Brief
Medium	Neither long nor short
Long	Extended
Spectral	Pitch
High	High pitch/frequency Treble
Mid	Medium pitch/frequency Alto
Low	Low pitch/frequency Bass
Dynamics	Volume/Loudness
Loud	High volume *Forte*
Medium	Medium volume/level
Soft	Quiet *Piano*
Content	Relevance
Informative	Relevant information
Neutral	Neither relevant nor irrelevant
Noise	Irrelevant/unwanted
Aesthetics	Beauty
Pleasing	Beautiful
Neutral	Mediocre
Displeasing	Ugly
Clarity	Quality
Clear	Easy to hear and comprehend
Neutral	Neither easy nor difficult to hear
Unclear	Difficult to hear and comprehend
Emotions	Feelings
Positive	Acceptance, Anticipation, Joy, Surprise
Neutral	No emotional content
Negative	Anger, Disgust, Fear, Sadness

Table 2: Sound event classification

Establishing whether a sound is informative within an auditory interface has always been important [11], and here the *content* is classified as informative, neutral or (just) noise. Noise being defined, in this case, as an unwanted or undesired sound, rather than unpleasant [12].

Barrass and Frauenberger [13] referred to the importance of the balance which must be struck between the aesthetic and the informative when creating an auditory display. Our earlier work has also been shown that a sound's aesthetics are integral to its functional effectiveness within an auditory display [14]. For this study our treatment of *aesthetics* has been to reduce them to pleasing, neutral and displeasing, rather than the more commonly used terms of harsh, warm, or bright (the latter terms being rather esoteric and requiring 'critical listening skills' [15]).

Clarity applies to the intelligibility of a sound and is rated according to whether it is clear, neutral or unclear, although in professional practice it is normally described as either poor or good. *Emotions*, which in this case are considered in terms of positive, neutral or negative, are not normally associated with sound design, although Johannsen [16] argues that if a sound has been 'well-designed' appropriate emotions should be evoked.

For this small illustrative study, 10 volunteers from the staff and students within the University participated. Each of the participants was familiar with the inside of a car and with driving, and had no known hearing impairments. Each candidate sat in the centre of eight compact loudspeakers and four sub bass units (see Figure 3). Each speaker location corresponded to the equivalent position of an omni-directional microphone during the recording. This ensured that all of the timings for the audio cues remained consistent, making it a more accurate spatial representation of the interior of the car.

Figure 3: Surround sound reproduction apparatus

Each listener participated individually. They were first asked to read a set of guidelines and invited to ask any questions that they might have. They then listened to the three sounds created by the designer while consulting the printed descriptions. Whilst this meant that that the listeners were primed, which created a risk of a higher rate of recognition, it was necessary for them to have an understanding of the meaning of the sounds as all of other sound events were potentially familiar. The presentation of the second, third and fourth recordings were pseudo-randomised in order to help mitigate the effects of fatigue and the learning effect. After the first sequence participants were asked to use the outline of the car (overlaid with a grid) to record and classify their experiences. Participants were questioned after the replay of the recording so that their responses closely reflect what they had been listening to. Once all of the responses had been elicited, descriptive statistics were applied to them. Aggregated coordinates were derived by using a median rather than a mean, so as to reduce the effect of outliers skewing the data. We also adopted the heuristics that if 50+% of the subjects were aware of a sound event then it was included in the combined map.

2.3. Visualization

Servigne *et al*. [17] have suggested that 'graphic seminology' would be appropriate for displaying sounds, proposing that smiling faces overlaid onto a map could be used to display participant's preferences. And in this spirit we created a set of symbols in order to visualize the listeners' experiences. These symbols may be found in Figure 4.

Each sound event was given a code by the first author and the combination of shapes, colours and symbols were overlaid onto the grid according to the x-y coordinates provided by the participants. If two or more sound events had identical coordinates then they were spaced evenly across the cell so that they remained visible. For ease of interpretation the grid, numbers and interior of the car were removed. The outline of the car was retained in order to provide some indication of orientation and scale.

Sound *type* was represented through either: a series of letters for speech, quavers for music, or a loudspeaker symbol for sound effect. The material was illustrated through the border colour, cyan magenta and yellow (CMY) which were applied to the spectral representation. This allowed colour values to be absolute in both printed and onscreen forms. The interaction was depicted using border dashes, impulsive had short dashes, whilst intermittent had longer, and therefore fewer dashes, whilst continuous was a single dash with no gaps. This approach was chosen so that it visually suggested the length of the sounds' interaction. *Temporal* attributes represented using a fill gradient, a radial gradient was used to suggest a short event, which visually is associated with a droplet falling on to a liquid. A medium event was portrayed with a linear gradient which suggested a more gradual change, and a long event was a solid colour which implied that there was either none or minimal change. The gradient started with the spectral fill colour and then progressed to a pure white and then back to the original fill colour. Fill colour was used for the spectral attribute, red was used for high, green for mid and blue (RGB) for low following the practice of auditory professionals [18].

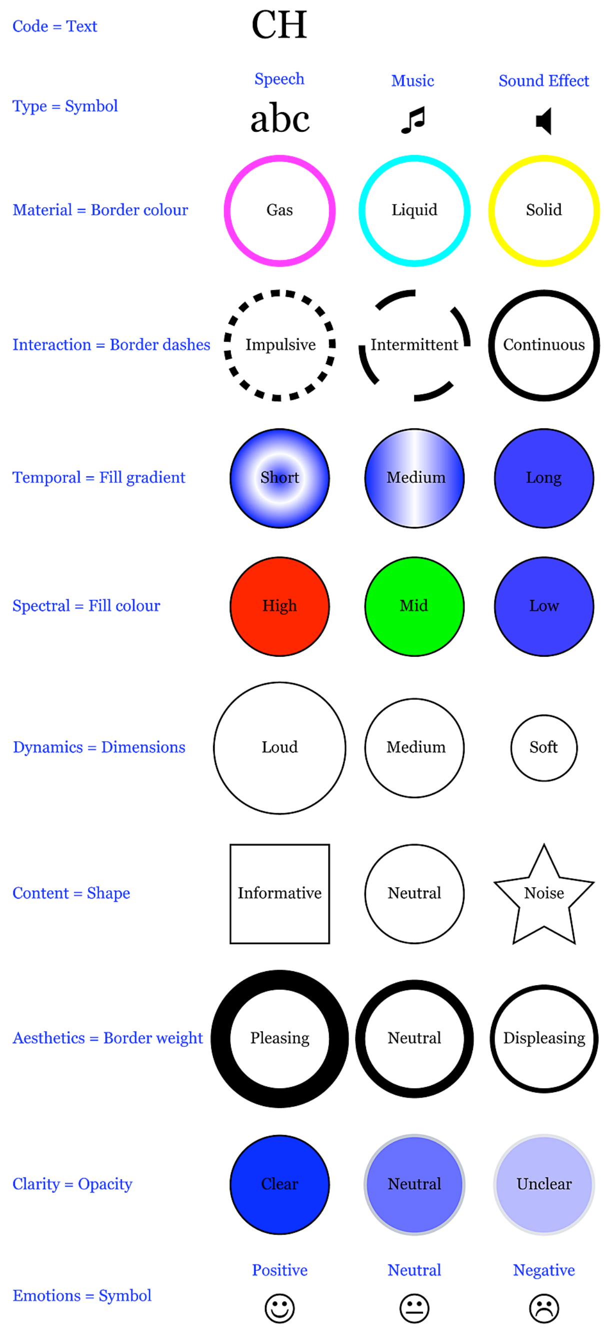


Figure 4: Visualisation key

Dynamics were illustrated using the scale of the shape, a soft sound was half the size of a medium one, and a loud sound event was 1.5 times the size of the medium and three times that of the soft. A square was used to signify informative, a circle for neutral and a star for noise. The three distinct shapes do not share any stroke angles, making it easier to differentiate between them when sound events are overlapped. Aesthetics were denoted by border weight, pleasing was represented with a thick line which was double the width of the neutral and four times the size of the displeasing. The *clarity* of a sound event was shown through the opacity of the shape, clear (=100% opaque), neutral (=66%) and unclear (=33%). Finally, 'emoticons' were used to represent positive emotions (a smile), neutral for neutral and a frown for negative.

3. RESULTS

The recording was relatively simple to transcribe, participants appeared to find the sound events straightforward to classify. The visualisations yielded informative results that showed clearly what participants were listening to.

3.1. Capture

Within the five minutes of audio recording 157 separate audible events were notated, these were identified as having been generated by 49 different sources. Sources such as the car's radio generated more than one type of sound event, so by grouping together sound events according to their source and the event it was possible to reduce the total down to 65. This was augmented by the designer with a further 3 sound events which were grouped together as a simple auditory display.

Sound events were generated from the car under study (28), passing vehicles (28), people (5), a dog and some scaffolding. Within the car, the engine passing through different states (idling, accelerating, cruising and decelerating) was recorded, as well engaging and releasing the handbrake, changing gear and a wide range of vibrations. There were 11 distinct types of sound from the radio, these were split into speech, music and laughter. Outside of the car 27 different vehicles were noted along with a siren, vehicle passes, brake squeals, indicators and windscreen wiping. The remaining sounds included screaming, talking, rustling of clothes, barking and scaffolding being struck.

Regarding the spatial cues, all of the sounds associated with the car could be identified to specific points within the outline of the car. The majority of the passing vehicles were located on the driver's side, which is at the top of the map, whilst most of the stationary vehicles were found to the rear of the car, which corresponds to the right hand side of the illustration. There were few sound events on the passenger's side and in front. The discrepancy to the paucity of sound events on the passenger's side can be partially explained by the comparatively low level of sounds on the pavement, when compared to the much louder vehicles. The shortage of audible sound events at the front of the car is most probably due to masking associated with the car's engine, which was constantly running throughout the recording.

This list only represents what could be heard on the recording, many more sounds would have been present but were either masked or inaudible due to the method of capture. All notes were made listening to the multi channel recording at the original sound pressure level, rather than over amplifying to enhance barely audible sources. This was done so that it replicated the conditions of the original journey as well as the reproduction levels which participants would have experienced.

3.2. Classification

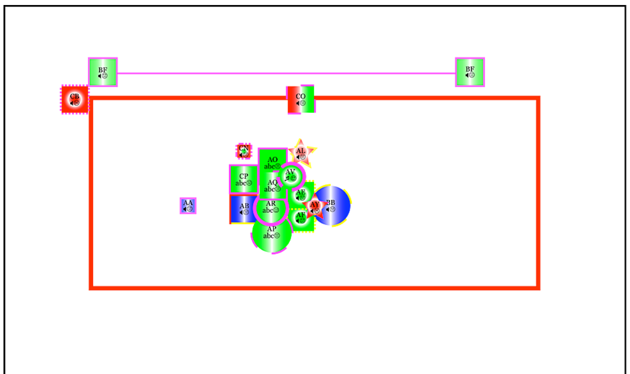
Participants were aware of an average of 30% of the sound events with a range of 38% to 21%. An average of 25% of all of the sound events from the car were heard by the participants the first time they heard the recording compared to 29% for the second. With the auditory display, the average was 94% for the first exposure, compared to 91% for the second, which might be due to habituation, but the difference is too small to draw conclusions from.

Overall there was a high level of awareness for the sounds associated with the car's engine and its handbrake, whereas the other sources such as internal vibrations, and indicating went comparatively unnoticed, except for when all of the wheels passed over a bump together. On the radio the first male voice was discerned, whereas the second, and its associated chanting, was missed. Two out of the three female voices, again on the radio, were identified, as was the interference from a mobile phone, but only one of the pieces of music was attended to. The group laughter was also generally missed, despite being the last thing that was present on the recording. Only two passing cars, and one passing bus were detected, which participants partially explained by the overwhelming urge to listen to the conversation from the radio, even when they were experiencing the identical content for a second time. When listening to the three sound events from the auditory display all of the participants were aware of all of the sounds. When they were listened to in context, then four out of the ten no longer recalled the braking distance cue, and even the designer was unaware of it, despite having added it into the recording himself.

Listeners found it hard to accurately recollect where a sound originated, but were much more comfortable with its orientation in relation to their listening position, although there were the occasional front to back errors. This is not surprising as problems with spatial discrimination are well documented, particularly when the source is not directly in front of the listener [19]. For the classification as a whole there was an average consistency of 80% between individual attributes, with a range of 67% - 98%. The average response of the ten participants was also compared to the combined classification which showed that apart from the *interaction* there was a good level of correspondence between the two sets of figures.

3.3. Visualization

A total of 36 maps were created. Each participant provided classifications for three maps, the car on its own, the isolated auditory display, and the car augmented with the auditory display. The aggregated (combined) classifications were also mapped in the same manner as the individuals' (see Figure 5). In addition it was possible to create a fourth map which represented the auditory display as experienced in context, but isolated from the auditory backdrop.



Code	Event	Source	Code	Event	Source	Code	Event	Source
AA	Engine Idle	Engine	AP	Male 2 Speech	Radio	BF	Passing Car	Car 2
AB	Engine Accelerate	Engine	AQ	Female 1 Speech	Radio	CB	Horn	Car 14
AE	Handbrake Released	Handbrake	AR	Female 2 Speech	Radio	CN	Braking distance	Auditory Display
AF	Handbrake Engaged	Handbrake	AV	Music 1	Radio	CO	Dead angle	Auditory Display
AL	Seat Creak	Driver's Seat	AY	Mobile Phone Interference	Radio	CP	Message	Auditory Display
AO	Male 1 Speech	Radio	BB	Bumps	All wheels			

Figure 5: Visualisation of soundscape for car and auditory display by combined participants

Only sound events which participants stated that they were aware of were included on the maps, otherwise they were omitted. An issue arose when sound events occupied the same coordinates. If their clarity was classified as being neutral or unclear then it was possible to overlap them quite tightly, whilst ensuring that the relevant information was still clearly visible, this was due to their partial opacity. But if all of the sound events were considered to be clear, and therefore opaque, then the amount of overlapping was minimal, as any area that was occluded was therefore no longer visible. Whilst this created problems with accurate positioning on the relevant coordinates, it did visually make it easier to see distinct clear sound events as they occupied a larger area. In contrast clusters of neutral or unclear sound events were visually more complex due to their cluttered nature. A simple solution to allow the inclusion of more sound events within a single grid would be to scale all of the attributes of the shapes down. Monmonier [20] recommended that symbols are moved 'slightly apart' to decrease the amount of overlap, and if this is not possible, then an inset at a larger scale could be used for the crowded area. The code and the type and emotions symbols were always kept at the same scale (8 pt) and opacity (100%) which made them easier to locate and identify.

The maps clearly show the listeners' awareness of sounds located in front and, to a lesser extent, the sides of the listeners. Sound events which were located to the side were normally moving, whilst those in front were almost always stationary. The use of CMY for borders and RGB for fills meant that any combination, even a continuous gas long high sound event which had a continuous magenta border with a solid red fill was clearly legible. Where this does not work as well as hoped was when a sound was classified as displeasing, the thin nature of the border width made it difficult to read the material and interaction, without the ability to zoom. This could be partially rectified by increasing the overall scale of the borders, so that the thinnest is at least 2 points, which is currently the size of the neutral condition.

Shape and size were easy to identify, even when partially occluded due to their symmetrical nature, which meant that the entire symbol does not have to be visible in order to identify its shape. Smaller soft sound events were layered on top of larger loud ones, and semi opaque unclear sounds appeared slightly washed out compared to the stronger colours of the clear ones. When comparing maps it is easy to see what a participant or group are paying attention to, and how this differs from individual to individual. Figure 6 shows the designer's map for the auditory display and the participants' combined responses in situ with the vehicle pre-existing auditory environment subtracted.

It can be seen that the spatial cues have been identified, albeit with slight variation, the email message and the braking distance alerts have remained in front of the driver, but reversed, and the dead angle has been discerned as originating from the right, but not as far back as the designer intended. The type has remained consistent for the braking distance and dead angle, both being considered to be sound effects, but the message has only been classified as speech, rather than a combination of speech and sound effects. This suggests that the sounds contained within the message are passing unnoticed. The material, which in this case was gas, remains constant, whereas the dead angle is perceived as being intermittent rather than impulsive. This shows that the dead angle is thought to be more of a whooshing sound rather than an explosion, which is also possibly due to a close association with the sound which a passing vehicle makes, this is also borne out through the alert being thought to be temporally medium in length rather than short.

The pitch for the two alerts were judged to be high mid rather than just high and the dynamics for the braking distance was considered to be soft but still clear. All of the events were classified as informative and aesthetically neutral, as well as emotionally neutral. It can be seen that the participants experienced the auditory display in context in a manner similar to the designer's intentions.

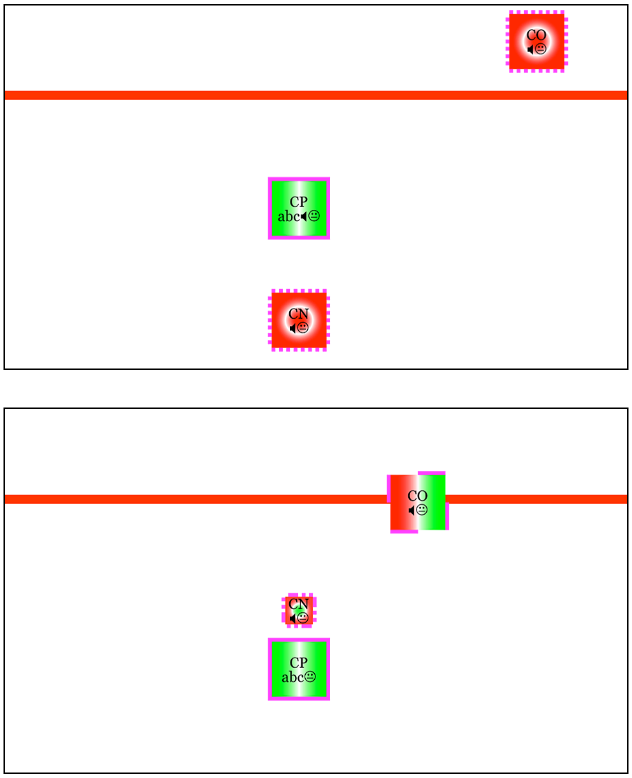


Figure 6: Magnified areas (identical coordinates)of designer's (top) and combined participants' (bottom)soundscape map for the auditory display in context with vehicle sound events subtracted (CN = Braking distance, CO = Dead angle, CP = Message)

3.4. The designer's comments

The designer found this method to be a quick and useful way of interpreting the data. He did, however, identify the need to include *height* channels. There were some other general comments as to the conduct of the studies themselves, observing that for longer duration soundscapes it would be useful for listeners to make notes, interruptions could also be used for longer experiments. He requested a confidence rating for each individual icon, as well as an electronic version where information about how the values were derived was displayed in a side table, on mouse-over of the relevant icon. He also suggested giving the designer a choice of classification scale, as sometimes looser is more appropriate. Some attributes might be better with more categories such as spectral and dynamics, whilst others would suit less, as in informative, where the neutral option could be dropped so that the decision is binary. The inclusion of spatialisation in the form of coordinates was deemed to be appropriate.

The labels used within the classification may require some fine tuning. He found the *temporal*, *spectral* and *dynamics* attributes to be context dependent, but relevant. The issue with the *temporal* attribute is that where a sound event could considered to be high in relation to its source, such as a high note on a cello, which is essentially a bass instrument, or a high tone from a male voice which might be considered to be low pitched in overall terms. It was also suggested that practical examples such as a female voice for the high category might be more helpful than the current examples of 'high pitch/frequency treble'.

With respect to *content*, the need for neutral option was queried and a request for a greater degree of granularity scale with possibly five or seven choices specifying the degree of information, such as moderately informative, informative, highly informative and so on. The use of the term noise was considered to be too ambiguous, noise could be considered as irrelevant and annoying. It was suggested that noise was changed to uninformative for consistency. The description was judged to be imprecise, as the information could be relevant but unwanted, this could easily be improved by removing the term unwanted. This attribute was regarded as the most important for the purpose of interface evaluation, especially with reference to answering the question of how informative it was.

Aesthetics were judged to be relevant, but like content, it would be more useful to have a more discriminating scale. With regards to the descriptions, mediocre was considered to be displeasing rather than neutral, and it was felt that the neutral state did not require a description at all. *Clarity* was regarded as pertinent, and like *type*, *material*, *interaction*, *spectral* and *dynamics* had the correct number of options, at three. Both the terms and descriptions were judged to be suitable. The classification of *emotions* could allow a greater degree of granularity, and the descriptors should be refined. Annoyance is not captured in the descriptor as a negative emotion, and it was queried as to whether surprise and anticipation were positive emotions. Concern was raised about the possibility of aesthetics cancelling out the emotions. There was a tendency for pleasing sounds to be classified as positive, This was even more evident for neutral aesthetics and neutral emotions, but was not the case with displeasing and negative emotions which only coincided fifty percent of the time.

Almost all of the methods of visualizing the attributes were regarded as effective, two suggestions for changes were made. The first was to amend the gradient associated with the temporal attribute so that only a radial gradient was used and that its size varied according to the length of the event. A short event would have a smaller area where the gradient was applied, whilst a long event would have a correspondingly larger area. This would allow for a linear scale as well as addressing the issue of the linear gradient sometimes being difficult to see in conjunction with a low level of opacity. The spectral representation might also be changed from three distinct colours to a continuous scale, in order to allow a greater degree of granularity.

4. CONCLUSIONS

This paper provides an illustration of the use of a soundscape mapping tool. It also showed that the tool could potentially be used by designers for the evaluation of sounds and auditory environments. The process of mapping has allowed a four dimensional auditory environment to be captured in two dimensional form, allowing ease of comparison between a designer's expectations and listeners' experiences. It also represents the effect of listening rather than hearing, where it is clear what is being attended to, and what has become habituated or has been ignored. With the car it was evident that sounds emanating from beyond the rear of the vehicle fell into this latter category, whereas those in front of or immediately surrounding the driver fell into the former. The relevance of sounds were also shown so that unwanted elements such as mobile phone interference and the driver's seat creaking could be silenced or masked, but other sounds such as the engine idling or accelerating, and the handbrake being engaged and disengaged should remain clearly audible as they were considered informative. The next stage of the research is to ask a range of sound designers to use the tool within their professional practise, and then query them about both the attributes and the visualization. This will help establish the tool's suitability for evaluating sounds and auditory environments.

5. REFERENCES

[1] P. Robare and J. Forlizzi, "Sound in Computing: A Short History," Interactions, vol. XVI, pp. 62-65, 2009.

[2] S. Barrass and C. Frauenberger, "A coummunal map of design in auditory display," in Proceedings of the 15 International Conference on Auditory Display, Copenhagen, Denmark, May 18-22, 2009 Copenhagen, Denmark: ICAD, 2009.

[3] S. Bech and N. Zacharov, Perceptual Audio Evaluation. Chichester, West Sussex: Wiley, 2006.

[4] American National Standards Institute., "Acoustical Terminology," American National Standards Institute, New York S1.1-1994, 1994.

[5] J. A. Ballas and J. H. Howard Jr, "Interpreting the Language of Environmental Sounds," Environment and Behaviour, vol. 19, pp. 91-114, January 1987.

[6] G. Bohme, "Acoustic Atmospheres," Soundscape, vol. 1, pp. 14-18, Spring 2000.

[7] M. O. Watson and P. Sanderson, "Designing for Attention With Sound: Challenges and Extensions to Ecological Interface Design," Human Factors,, vol. 49, pp. 331-346, 2007.

[8] I. McGregor, G. Leplatre, A. Crerar, and D. Benyon, "Sound and Soundscape Classification: Establishing Key Auditory Dimensions and their Relative Importance " in ICAD 2006 London: Department of Computer Science, Queen Mary, University of London, 2006.

[9] I. McGregor, A. Crerar, D. Benyon, and G. Leplatre, "Establishing Key Dimensions for Reifying Soundfields and Soundscapes from Auditory Professionals," in ICAD 2007, 2007.

[10] I. McGregor, A. Crerar, D. Benyon, and G. Leplatre, "Visualising the Soundfield and Soundscape: Extending Macaulay and Crerar's 1998 Method," in Proceedings of the 14th International Conference on Auditory Display Paris: IRCAM, 2008.

[11] W. W. Gaver, "Auditory Icons: Using Sound in Computer Interfaces," Human-Computer Interaction, pp. 167-177, 1986.

[12] P. Radomskij, "Measurement of noise," in Noise and Its Effects, L. Luxon and L. Prasher, Eds. Chichester, UK: John Wiley and Sons Ltd., 2007, pp. 13 - 43.

[13] S. Barrass and C. Frauenberger, "A coummunal map of design in auditory display," in Proceedings of the 15 International Conference on Auditory Display, Copenhagen, Denmark, May 18-22, 2009 Copenhagen, Denmark: ICAD, 2009.

[14] G. Leplâtre and I. McGregor, "How to Tackle Auditory Interface Aesthetics? Discussion and Case Study" in Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, S. Barrass and P. Vickers, Eds. Sydney, Australia: International Community for Auditory Display (ICAD), 2004.

[15] W. Moylan, The Art of Recording: Understanding and Crafting the Mix. London: Focal Press, 2002.

[16] G. Johannsen, "Auditory Displays in Human–Machine Interfaces," Proceedings of the IEEE, vol. 92, pp. 742-758, April 2004.

[17] S. Servigne, R. Laurini, M.-A. Kang, and K. J. Li, "First Specifications of an Information System for Urban Soundscape", IEEE International Conference on Multimedia Computing and Systems, vol. 2, pp. 262-266, 1999.

[18] D. Gibson, *The Art of Mixing: A Visual Guide to Recording Engineering and Production*, 2nd ed. Boston: Artist Pro Publishing, 2005.

[19] W. D. Grantham, Spatial Hearing and Related Phenomena. In B. C. J. Moore (Ed.) *Hearing* (2nd edition), London: Academic Press, 1995, pp. 297-345.

[20] M. Monmonier. *Mapping it out: Expository Cartography for the Humanities and Social Sciences*. London: University of Chicago Press, 1993.