doi:10.1093/iwc/iwx007

# A Model of Nature Soundscape for Calm Information Display

Bin Yu, Jun Hu*, Mathias Funk and Loe Feijs

*Industrial Design Department, Eindhoven University of Technology, Laplace 32, 5612 AZ Eindhoven, The Netherlands*
**Corresponding author: J.Hu@tue.nl*

**Calm technology has been stressed in designing the interaction with information, especially in ubiquitous computing, peripheral interaction and ambient display. Inspired by the research on calm technology and model-based sonification, we aim to build a model of nature soundscape for supporting calm information display. A three-layer structure is proposed for construction of the nature soundscape. The structure includes seven acoustic parameters. By setting each of seven acoustic parameters into three levels, seven groups of soundscape samples were created and evaluated in an experiment with 20 participants. Each participant was exposed to 21 soundscape samples to assess each sample regarding seven perceptual attributes through a rating scale. Based on the results, a perceptual model is proposed to link the acoustic parameters of individual nature sounds and the perceptual attributes of the nature soundscape. The developed model offers the designers and practitioners a new tool to utilize nature sounds in the design of the auditory display which could support the calm technology.**

RESEARCH HIGHLIGHTS

- A three-layer nature soundscape structure
- A perceptual model of nature soundscape
- Peripheral display with nature sounds
- Calm information display through the user's perception of nature soundscape

*Keywords: calm technology; nature sounds; soundscape; auditory interface; sonification model; ambient display*

*Editorial Board Member: Phil Turner*

*Received 22 December 2016; Revised 13 March 2017; Editorial Decision 10 April 2017; Accepted 8 August 2017*

## 1. INTRODUCTION

In the field of human–computer interaction (HCI), non-speech audio is widely used in the user interface to communicate information (Buxton, 1989; Csapó and Wersényi, 2013). An auditory display may offer many potential benefits in specific scenarios. For instance, it can present information that is hard to discern visually (Neuhoff *et al.*, 2002), display for the visually impaired (Jagdish *et al.*, 2008) or complement to visual output in the situations that the user's vision is occupied (Gable *et al.*, 2013). Audio is a useful medium to maintain awareness of activities taking place around us. Therefore, it is often used for ambient display (Ishii *et al.*, 1998; Mynatt *et al.*, 1997) and peripheral interaction (Bakker *et al.*, 2015; Cohen, 1993). Since auditory displays liberate the users from visual focus, they may also improve the users' comfort and facilitate their calmness with the information (Annerstedt *et al.*, 2013; Ratcliffe *et al.*, 2013).

The design of auditory display has traditionally addressed the questions around how to present the information in a form that is easy to understand and efficient to use. However, with the overwhelming amount of information coming to us, another question has often been asked and discussed recently: how could the technologies calmly inform us without overburdening us? Ubiquitous computing and information technologies keep people easily informed all the time. The increasing bandwidth and enriching channels of information increasingly engage our attention and keep us farther away from the sense of calmness. Therefore, besides the effectiveness of information delivery, increasing efforts are being made to facilitate the user's calmness. Calm Technology (Mark and Brown, 1997) was formulated by Weiser and Brown in 1996. It suggests that the information display should 'engage both the center and the periphery of the user's attention, and in fact moves back and forth between the two.' A 'calm' auditory display shifts the information from the user's focus of attention to the auditory periphery, and, the users can attune to the information without explicitly attending to it.

In auditory displays, the sonification approach plays a critical role in transforming the data into the right form of audio signal. According to Hermann (Hermann and Ritter, 1999), five sonification approaches are mainly used for acoustic data presentation: *Audification*, *Auditory icons*, *Earcons*, *Parameter-Mapping (PM)* and *Model-based sonification (MBS)*. *Audification* is transforming the data directly to the audible domain, where a time series directly controls the audio signal amplitude (Sandell, 1996). The term *'Auditory icons'* was coined by Gaver (Gaver, 1986) and has been embodied in several designs of his, such as the *SonicFinder* (Gaver, 1989), *SharedARK* (Gaver *et al.*, 1991) and *ARKola bottling plant* (Gaver *et al.*, 1991). The *auditory icons* exploit everyday sounds to convey information in HCI. The semantic link between the attributes of everyday sound-producing events and attributes of computer events makes the *auditory icons* less annoying and easier to be learned by the users. *Earcons* (Blattner *et al.*, 1989) are using synthesized tones or sound patterns as audio messages to represent specific events and convey information. Compared to *Auditory icons*, *Earcons* are more abstract and often used in combination to produce a complex audio message. *Earcons* are not only designed for interacting with computers but used more widely with a longer history, such as alert signals from the emergency broadcast system. The users may need a longer learning process to build the relationship between the *Earcons* and their represented meanings.

In those interactive systems that communicate high-dimensional data, *PM* is a more common approach to convey information or perceptualize data (Hermann, 2008). In PM sonification, the data values are directly mapped to acoustic attributes of a sound, such as the duration, pitch, loudness, position and brightness. In other words, the data 'play' an 'instrument' by manipulating the parameters of a synthesizer. One of the advantages of PM is multivariate representation. Different data variables can be mapped to different acoustic parameters concurrently to produce a complex sound. Thus, many data dimensions can be listened to at the same time. However, because the sound pattern is associated with the data structure, the sounds that are produced in direct mapping might often be unpleasant. It is also difficult to predict the user's perception of the data-controlled sounds produced in a multivariate PM. These problems were reported in Smith *et al.* (1994) and indicated by Barrass and Kramer (1999) and Hermann (2008).

In our previous practice (Yu *et al.*, 2015), we designed an auditory interface for a heart rate variability biofeedback system which is used for stress management and relaxation. The heartbeat intervals are mapped to the variations of rhythm in MIDI notes. The results of the user experiment showed that the created auditory display was a good alternative to the standard graphical feedback. However, regarding the user experience, it received a lower score on the 'comfort of use.' We found that directly mapping the variations in data to the rhythmic variations was somewhat arbitrary. In our case, PM approach could transform the data into the sounds effectively but was difficult to shape a relaxing and pleasant user experience during an extended period of use. Some users even reported more anxiety with the audio feedback, which was recognized as 'relaxation-induced anxiety' (Heide and Borkovec, 1983).

Hermann *et al.* introduced *MBS* in 1999 (Hermann and Ritter, 1999). It employs more complicated mediation between data and sound rendering by introducing a virtual 'sound generating model,' whose properties are linked to the data. The sonification model acts as a 'virtual instrument,' whose 'material structure' defines the sound properties and 'underlying physics' defines the modulation of the output sounds. The MBS is commonly designed to enhance user interaction, which involves 'interacting with data-driven virtual acoustic objects' (Hunt *et al.*, 2004). For instance, a virtual sound object was developed and could be 'played' by the movements of the upper limbs for biofeedback training (Maes *et al.*, 2010).

Besides being mapped to the data in the sonification, the audio signal itself can be an 'active' stimulus contributing to the user experience during HCI. For instance, a piece of music may induce the autonomic relaxation, but a short high-pitched tone may cause an alert adversely. Music signal can play in stimulating the imagination (Lundqvist *et al.*, 2009) and boost moods (McCraty *et al.*, 1998). Nature sounds can also powerfully induce positive emotional states (Ulrich *et al.*, 1991), help in calming down (Alvarsson *et al.*, 2010; DeLoach *et al.*, 2015) and sustain the attention (Kaplan, 1995). In some specific applications for rehabilitation, stress management, relaxation practice and healthcare, these auditory contents are frequently applied to the auditory interfaces for facilitating the user's calmness and relaxation. For instance, Harris *et al.* (2014) developed an auditory display of breathing signal by adjusting the quality of a music recording to promote relaxation. Bergstrom *et al.* (2014) developed a musical interface which presents the user's physiological state by adjusting the musical tempo and volume.

Nature sounds are among 'everyday sounds' around us. When we are outdoors in a garden or the woods, we hear the sounds of birds whistling. It does not usually take too much for us to adapt to these sounds. Besides the ability to foster the experience of calmness and relaxation, the nature sounds have another advantage for auditory display as they are intuitive, familiar and may be understood quickly and learned easily. Thus, nature sounds are often used in ambient displays and peripheral interactions by creating a 'calm' sonic environment, which can engage the periphery of our attention to grab the presented information. For instance, Eggen and Van Mensvoort (2009) used bird sounds in a peripheral display to communicate information about the activities in the office. *AmbientROOM* (Ishii *et al.*, 1998) modulates the volume and density of bird and rainfall sounds to present the number of unread email messages and the value of a stock portfolio. *Audio Aura* (Mynatt *et al.*, 1997) used seagull cries and beach birds as auditory cues to provide office workers with relevant information such as the availability of colleagues.

The value of everyday sounds used as *Auditory icons* lies in their associated meanings. In the above studies (Eggen and Van Mensvoort, 2009; Ishii *et al.*, 1998; Mynatt *et al.*, 1997), the nature sounds were used individually as a special type of 'earcons,' like 'musical tones.' They have been given a renewed meaning in specific contexts to communicate the targeted information, such as the number of unread email messages by the volume and density of birds (Ishii *et al.*, 1998). The changes of an individual nature sound seem to be meaningless to us, but the changes of the 'soundscape' shaped by multiple changing nature sounds can 'inform' us through our intuitive perception of the 'immersed acoustic environment,' such as calmness, pleasantness, loudness, eventfulness and familiarity. We think the perceptual and emotional attributes of soundscape can present information naturally and meaningfully. In our view, it might be better when possible to link the dataset to the attributes of the soundscape in the interface instead of directly to the parameters of individual sounds.

In this study, we take the inspiration from the *MBS* and propose a new approach of using nature sounds for calm information display. Different from the sonification models focusing on the design of the 'virtual instrument,' a nature soundscape (NS) model is developed as a 'virtual natural environment.' The developed NS model offers the sound designers not only a framework to create a coherent soundscape, but also a means to present information in a calm way by linking data to perceived attributes of the overall soundscape. This study is divided into two parts: designing the 'structure' of NS and establishing the 'underlying relations' between the acoustic parameters of individual nature sounds (interfacing with data) and listener's perception of the whole NS (interfacing with a human).

## 2. CONSTRUCTING AN NS

According to Schafer (1993), the 'soundscape' refers to the unique experience of inhabiting an acoustic environment with emphasis on the individual's sensation and perception of different types of sounds. Since then, the term 'soundscape' has been used extensively to describe an 'acoustic environment' about the acoustic resources within a given area. NSs have been studied in many fields, ranging from urban design (Yang and Kang, 2005), monitoring of the wildlife (Pijanowski *et al.*, 2011) and auditory display in public space (Eggen and Van Mensvoort, 2009). A central topic spanning across these fields is the informational aspect of the soundscape, either extracting information from a recorded soundscape or convey information by creating a new one. In this study, we focus on the latter.

In our view, NSs have a great potential in supporting HCI for both informative and experiential goals. NS may refer to both the natural acoustic environment consisting of various natural sounds, and also the listener's perception and experience of sounds heard as an environment. An NS may consist of various sounds including animal vocalizations, the sounds of weather and other natural elements. As each sound element can be a possible information carrier, an NS can present multi-channel of information simultaneously. For instance, Hermann *et al.* (2003) combined the sounds of the wind, rainfall, thunder and frog as an auditory weather forecast presenting various channels of weather information. Moreover, a rich diversity of nature sounds in a coherent context may also create an acoustic environment which can be experienced to be pleasant, calm and relaxing.

### 2.1. Structure of NS

The NS that arises from a real landscape tends to be very complex, varying spatially and temporally. It is difficult to exploit a real recording of a natural environment for information display. Therefore, instead of the realism of the synthesized soundscape, we focus on building a controllable 'virtual natural environment' with limited sound components and investigating about the human perception of the 'acoustic environment' regarding the attributes of the 'NS.'

We propose a practical structure to describe the structural hierarchy of an NS. Based on Pijanowski *et al.* (2011), our working definition of NS is'the collection of biological and Gss that emanate from a natural environment.' Thus, the NS in this study does not include the 'anthrophony' which caused by humans, only focuses on the 'biophony' and 'geophony' created by nature including biology and geography. According to Krause (1987), 'Biophony' describes the composition of sounds created by organisms and 'geophony' describes non-biological ambient sounds occurring at a site. Based on studies of Schafer (1993), the sound components in a soundscape can be classified into three types: *keynotes*, *signals* and *soundmarks*. The *keynote* sound is the tonal center of a soundscape such as the sound of the running water by a riverside. The *signals* sound is the informational sounds that appear infrequently and separately. A *soundmark* is a unique sound to an area, like an audible 'landmark.'

In this study, we simplify the composition of an NS as a three-layer structure, consisting of geophysical sound (Gs), biological sound (Bs) and climatic sound (Cs), see Table 1. Gs reflects the geographical features at a site. It serves as the *keynote* sound, which shapes the basic scenario of an NS. Bs

Table 1. The parameters and perceptual attributes in three-layer framework of the NS.

| | Sound layer | Classes of sounds | Sound source | Audio signal | Parameters | Example |
|---|---|---|---|---|---|---|
| Cs | Climatic sound | Ambience noise | Single | Continuous | Volume | Wind |
| Gs | Geophysical sound | Keynote sound | Single | Continuous | Volume | Water stream |
| Bs | Biological sound | Signals sound | Multiple | Discrete | 1. volume; 2. density; 3. type variation; 4. rhythm variation; 5. direction variation | Silvereye, wren, greenfinch, collared dove, cuckoo |

serve as the *signal* sounds, reflecting natural events and processes. It may consist of a diverse array of nature sounds produced by mammals, birds, amphibians and insects. A soundscape can also be described in terms of Hi-Fi and Lo-Fi based on the ambience noise level (Schafer, 1993). We consider the ambience noise as an independent component, which influences the perception of the soundscape regarding the Hi-Fi and Lo-Fi. Thus, Cs refers to the ambience noise created by the climate such as the wind, rain or shore noise. This simplified structure helps in the selecting and mixing various nature sounds. The resulting soundscape can be one of the many instantiations of the class of 'NSs.' Fox example, in our experiment, the sample soundscape of 'forest' was developed with the combination of the leave rustle (Cs), the murmur of a brook (Gs) and birdsongs (Bs).

### 2.2. Parameters of NS

Each NS has complex properties based on different biological and geographical features. According to Pijanowski *et al.* (2011), a soundscape possesses four measurable properties: acoustic composition, temporal patterns, spatial location and acoustic interactions. These properties are usually measured and analyzed for getting information about a soundscape ecology. In this study, we do the reverse that we select the controllable acoustic parameters based on these properties. The acoustic interactions in the NS vary widely according to animal activities. For practical reasons, we only address the componential, temporal and spatial properties. The composition of NS is associated with various acoustic parameters including frequency, amplitude and type of nature sounds. The temporal pattern of NS is mainly reflected by certain biological events. The spatial location refers to the direction and distance of the sound source.

Based on the above structure, in an NS, the sound selected for the Cs layer is a type of natural white noise, which is continuous and from a single source. In the Gs layer, only one Gs will be selected as the keynote, and it is from a fixed sound source. Therefore, for both Cs and Gs sounds, the *volume* is the only acoustic parameter to be adjusted. The Bs layer is comprised of various Bss from multiple moveable sources, such as birds, frogs and insects in the forest. Compared to Cs and Gs, the sounds in the Bs layer are discrete, and the sources might be 'moving around.' Therefore, more parameters regarding the temporal patterns and spatial location properties are selected.

For the Bs layer, besides the *volume*, the second parameter is *density*, which determines the basic time interval between two successive sound playings. A higher *density* shortens the time interval between the Bs sounds. The other three parameters of the Bs layer are mainly about the dynamics of the Bs sounds; they are the variations of sound type, rhythm and direction. The *type variation* determines how many types of the Bss will be 'activated' for playing. A higher *type variation* means that, for each playing, the sound source will be selected from a wider range of 'sound library'; with the same *density*, more types of Bs sounds will occur in the soundscape. The *rhythm variation* is the range of variation in the basic time interval which is determined by the parameter of *density*. A higher *rhythm variation* means that the Bs sounds will occur more unevenly, with more flexibility. All Bs sounds can be played through mono or stereo channels, which create directionality, perspective and space. The *direction variation* determines the proportion of the Bs sounds presented through the stereo left or right channels. A bigger *direction variation* will lead to a more real stereo surround quality.

In summary, we propose seven parameters distributed in different layers: Cs *volume*, Gs *volume*, Bs *volume*, Bs *density*, Bs *rhythm variation*, Bs *direction variation* and Bs *types variation*. We assume that by controlling one or more of these parameters, the listener's perception of the soundscape will be influenced.

### 2.3. Attributes of NS

In addition to the structure and acoustic parameters of the NS, understanding the user perception of an NS is also essential in the design of the auditory display. The listener's perception of multiple mixed sounds in a coherent context may create the sensation of experiencing a particular acoustic environment. Many studies have been conducted to assess and understand the perception of soundscapes (Coensel and Botteldooren, 2006; Raimbault *et al.*, 2003). In these studies, the assessment of the soundscape involves more perceptual and emotional measures rather than just identification and description of the sound sources. Various attributes of soundscapes emerged in the assessments, such as pleasantness, loudness, eventfulness, familiarity and sound dynamics. Coensel and Botteldooren (2006) suggested that the calmness and pleasantness might be a result of multiple other attributes such as loudness, eventfulness, familiarity, the dynamics of the sounds and the factors related to the spatial characteristics and the spectrum or timbre of the soundscape.

An NS is normally assessed by using a semantic differential. Table 2 shows the most common attributes selected for describing the perception of an NS. Based on the results from Coensel and Botteldooren (2006) and Raimbault *et al*. (2003), loudness is the most important attribute which is about the strength of a soundscape. The attribute of richness describes the sound diversity in an NS. Next, the attributes of steadiness and spatial impression are related to the sound dynamics regarding the temporal balance and the spatial localization. Naturalness evaluates the degree of realism of the soundscape environment. The attributes of calmness and pleasantness are selected to assess the appreciation and user experience of the soundscape. In this study, a specific perceptual attributes rating scales was designed with seven questions for evaluating the listener's perception of the NS, as shown in Table 2.

## 3. USER EXPERIMENT

A user experiment was conducted to understand the relationships between the acoustic parameters and the user perceptions of the NS. Based on the proposed NS structure, we created an NS to investigate how we can influence the user perception of the soundscape through the modulation of the acoustic parameters. By setting each of seven acoustic parameters into three levels, 21 soundscape samples were created. In a within-subjects experiment, each participant was exposed to seven groups soundscape samples and completed the soundscape rating scale for each sample. The independent variable is the level of the parameters (low, moderate and high), and the dependent variables are seven perceptual attributes of the soundscape measured by the soundscape rating scales.

### 3.1. Subjects

Twenty participants took part in the study through informed consent procedures. All participants were volunteers. They were randomly selected from a variety of undergraduate and graduate classes. The eleven males and nine females ranged in age from 22 to 33. All participants reported no history of diagnosed hearing impairments. Participants who were the trained listeners, either through professional audio training or music education were excluded from the study. The participants were unaware of the specific aims of the study, the modulation and the predicted effects of the different samples.

**Table 2.** Selected perceptual attributes of the NS.

| | Perceptual attributes | Rating scales |
|---|---|---|
| 1 | Loudness | quiet (1) vs. loud (5) |
| 2 | Richness | deserted (1) vs. lively (5) |
| 3 | Steadiness | unsteady (1) vs. steady (5) |
| 4 | Spatial Impression | closed (1) vs. open (5) |
| 5 | Naturalness | artificial (1) vs. natural (5) |
| 6 | Calmness | irritating (1) vs. calming (5) |
| 7 | Pleasantness | unpleasant (1) vs.pleasant (5) |

### 3.2. NS samples

Based on the results of our previous user survey (Yu *et al.*, 2016), 'Forest' is one of the most pleasant nature theme among the other scenes, such as ocean and grasslands. The moderate complexity also makes the 'forest' soundscape malleable and controllable. Therefore, we selected the nature sounds from the forest as the auditory contents for constructing the soundscape in this experiment. After analyzing the recorded soundscapes of the real forest, we created the soundscape consisting of wind sound as the Cs, a water stream for Gs and several types of birds (i.e. silvereye, wren, greenfinch, collared dove and cuckoo) for Bss. These birdsongs were selected as they are rated as the most likely to help people relax and recover from mental fatigue (Ratcliffe *et al.*, 2013). Seven groups of soundscape samples are created with the same audio contents. Within each group, one acoustic parameter is modulated into different level while other parameters are set to the default value (moderate level), see the Table 3. The volume of audio sources is normalized firstly and then modulated into different decibel value with a software synthesizer.

### 3.3. Procedure

All participants were tested individually in a small testing room furnished with a recliner chair, rug, lamps and audio equipment. All sound samples were played through an acoustic noise canceling headphones (*Bose, QuietComfort* 25). The participant was seated in the recliner with comfort and read the instruction before the experiment. Each participant listened to seven groups of soundscape samples in a randomized order. For each group, the order of samples was also randomized. The researcher started to play the samples one after another in the first group. After listening to each sample, the participant was asked to judge upon the attributes of what they hear with a rating scale. After each one group had been completed, the participants had 15 s to finalize their answers and hear a 30-s piece of music as a washout period.

### 3.4. Data analysis

The one-way ANOVA was carried out within each group to understand if there is a significant influence of each acoustic parameter on the user perception of the soundscape. A

Pearson correlation coefficient was computed to assess the relationship between the acoustic parameter and perceptual attributes of the soundscape.

## 4. RESULTS

### 4.1. The rating scale on attributes of soundscape within each group

Figure 1(a–c) shows the ratings on the NS attributes with a different level of Cs, Gs and Bs volume. As shown in Fig. 1(a), with a high Cs volume, the loudness of the NS is significantly higher than the one with the low and moderate Cs volume. Moreover, the pleasantness and calmness of the soundscape are reversed. In the second group, the loudness of the soundscape with high Gs volume is significantly higher than the one with moderate Gs volume. The loudness of the soundscape with moderate Gs volume is significantly higher than the one with the low volume. The other attributes like steadiness, spatial impression, naturalness, pleasantness and calmness are all reversed. Figure 1(c) shows that a high Bs volume leads to a significantly higher rating on the loudness. Conversely, the spatial impression, naturalness, pleasantness and calmness significantly decrease when the Bs volume is high. As shown in

**Table 3.** Parameter setting for creating the samples for the experiments.

| | | | Parameters level | | |
|---|---|---|---|---|---|
| Layers | Sound selection | Parameters | Low | Moderate | High |
| Cs | Wind | Volume | 0 dB | 5 dB | 10 dB |
| Gs | Water | Volume | 0 dB | 5 dB | 10 dB |
| Bs | Birds | Volume | 0 dB | 5 dB | 10 dB |
| | | Density | 10 sounds/min | 20 sounds/min | 30 sounds/min |
| | | Type variation | 1 type | 3 types | 5 types |
| | | Rhythm variation | ±0% | ±20% | ±40% |
| | | Direction variation | 100% Mono | 50% Mono<br>50% Stereo | 100% Stereo |

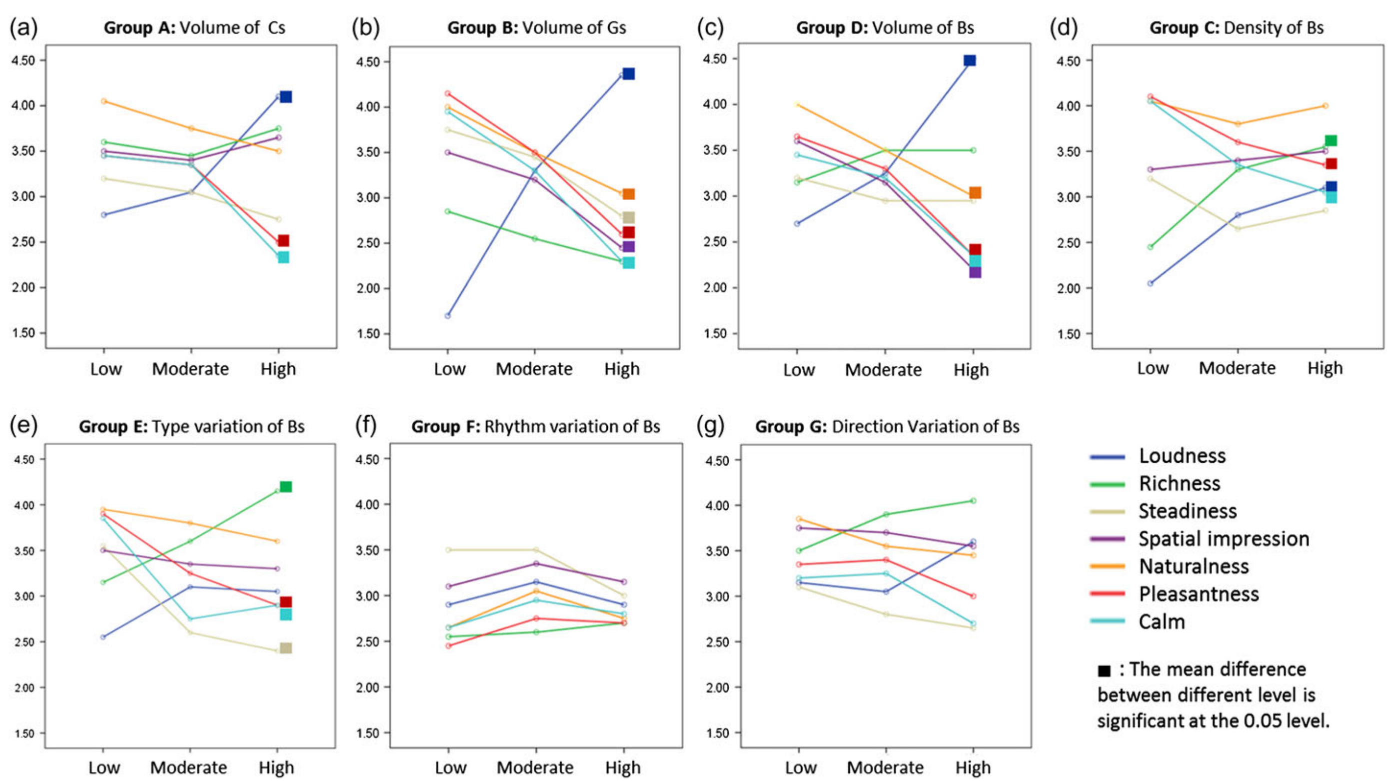


**Figure 1.** The rating scale on attributes of soundscape within each group with different parameter levels.

Fig. 1(d), with a high density of Bs, the loudness and the richness of the soundscape are significantly higher than the ones with a moderate and low Bs density. The pleasantness and calmness of soundscape are reversed. As shown in Fig. 1(e), a high variation of Bs type leads to a significantly higher richness, but a lower steadiness, pleasantness and calmness of the soundscape. Figure 1(f and g) shows that the listener's perception of the NS was not changed significantly with different levels of rhythm variation and direction variation.

### 4.2. The correlations between the acoustic parameters and perceptual attributes

Table 4 presents the results of correlation analysis between acoustic parameters and perceptual attributes. Regarding the loudness, there is a strong positive correlation between the loudness and Gs volume ($r = 0.854$, $P \leq 0.001$). There is a moderate positive correlation between the loudness and Cs volume ($r = 0.545$, $P \leq 0.001$) and Bs volume ($r = 0.677$, $P \leq 0.001$). There is a weak positive correlation between the loudness and Bs density ($r = 0.453$, $P \leq 0.001$).

Regarding the richness, there is a weak positive correlation between the richness and density of Bs layer ($r = 0.406$, $P \leq 0.001$), and type variation of Bs layer ($r = 0.384$, $P = 0.002$). Regarding the steadiness, there is a weak negative correlation between the steadiness and Gs volume ($r = -0.372$, $P = 0.003$), and type variation of Bs layer ($r = -0.398$, $P = 0.002$). Regarding the spatial impression, there is a weak negative correlation between the spatial impression and Gs volume ($r = -0.365$, $P = 0.004$) and Bs volume ($r = -0.431$, $P = 0.001$). Regarding the naturalness, there is a weak negative correlation between the naturalness and Gs volume ($r = -0.357$, $P = 0.005$).

The calmness is negatively correlated with Cs volume ($r = -0.408$, $P \leq 0.001$), Gs volume ($r = -0.647$, $P \leq 0.001$), Bs volume ($r = -0.425$, $P = 0.001$), Bs density ($r = -0.461$, $P \leq 0.001$) and type variation of Bs layer ($r = -0.375$, $P = 0.003$). The pleasantness is negatively correlated with Cs volume ($r = -0.404$, $P \leq 0.001$), Gs volume ($r = -0.647$, $P \leq 0.001$), Bs volume ($r = -0.445$, $P \leq 0.001$), Bs density ($r = -0.346$, $P \leq 0.001$), type variation of Bs layer ($r = -0.409$, $P = 0.001$).

### 4.3. The NS Model

Table 5 illustrates the relationships between the acoustic parameters and the perceptual attributes as a model. The model is built with the three-layer NS structure, which could guide us to construct the NS samples. All perceptual attributes show a hybrid relationship with multiple parameters across the layers. Only first five parameters show a correlation with the user perceptions of the NS. For the Cs and Gs layer, the volume is the only one parameter to control, and for the Bs layer, there are three parameters: volume, density and type variations. These acoustic parameters can be regarded as the input of the model, interfacing to the dataset. Seven attributes of the soundscape are viewed as output to interface to the listener's perceptions and experience. We can conclude that for the NS developed with the NS structure, there is evidence that the loudness is strongly related to the volume of three sound layers and the density of Bs sounds. The richness is related to the density and type variations of bio-sounds. The steadiness is related to to the volume of Gs layer and density and type variations of

**Table 4.** The correlations between the acoustic parameters and perceptual attributes of NS.

| | | Perceptual attributes | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Layers | Parameters | Loudness | Richness | Steadiness | Spatial Impression | Naturalness | Calmness | Pleasantness |
| Cs | Volume | $r = 0.545$ | $r = 0.071$ | $r = -0.184$ | $r = 0.055$ | $r = -0.201$ | $r = -0.408$ | $r = -0.404$ |
| | | $P \leq 0.001$ | $P = 0.589$ | $P = 0.16$ | $P = 0.678$ | $P = 0.123$ | $P \leq 0.001$ | $P \leq 0.001$ |
| Gs | Volume | $r = 0.854$ | $r = -0.254$ | $r = -0.372$ | $r = -0.365$ | $r = -0.357$ | $r = -0.647$ | $r = -0.601$ |
| | | $P \leq 0.001$ | $P = 0.05$ | $P = 0.003$ | $P = 0.004$ | $P = 0.005$ | $P \leq 0.001$ | $P \leq 0.001$ |
| Bs | Volume | $r = 0.677$ | $r = 0.138$ | $r = -0.106$ | $r = -0.431$ | $r = -0.313$ | $r = -0.425$ | $r = -0.445$ |
| | | $P \leq 0.001$ | $P = 0.292$ | $P = 0.421$ | $P = 0.001$ | $P = 0.015$ | $P \leq 0.001$ | $P \leq 0.001$ |
| | Density | $r = 0.453$ | $r = 0.406$ | $r = -0.139$ | $r = 0.079$ | $r = -0.025$ | $r = -0.461$ | $r = -0.346$ |
| | | $P \leq 0.001$ | $P \leq 0.001$ | $P = 0.29$ | $P = 0.549$ | $P = 0.847$ | $P \leq 0.001$ | $P = 0.007$ |
| | Type variation | $r = 0.225$ | $r = 0.384$ | $r = -0.398$ | $r = -0.075$ | $r = -0.134$ | $r = -0.375$ | $r = -0.409$ |
| | | $P = 0.084$ | $P = 0.002$ | $P = 0.002$ | $P = 0.566$ | $P = 0.307$ | $P \leq 0.003$ | $P \leq 0.001$ |
| | Rhythm variation | $r = 0$ | $r = 0.058$ | $r = -0.175$ | $r = 0.020$ | $r = -0.030$ | $r = -0.065$ | $r = -0.114$ |
| | | $P = 1$ | $P = 0.658$ | $P = 0.18$ | $P = 0.880$ | $P = 0.820$ | $P = 0.623$ | $P = 0.385$ |
| | Direction variation | $r = 0.206$ | $r = 0.244$ | $r = -0.181$ | $r = -0.076$ | $r = -0.147$ | $r = -0.231$ | $r = -0.172$ |
| | | $P = 0.114$ | $P = 0.061$ | $P = 0.166$ | $P = 0.564$ | $P = 0.263$ | $P = 0.076$ | $P = 0.188$ |

**Table 5.** The model of NS.

| | | Perceptual attributes | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Layers | Parameters | Loudness | Richness | Steadiness | Spatial Impression | Naturalness | Calmness | Pleasantness |
| Cs | Volume | ++ | | | | | – | – |
| Gs | Volume | +++ | | – | – | – | – – | – – |
| Bs | Volume | ++ | | | – | – | – | – |
| | Density | + | + | | | | – | – |
| | Type variation | | + | – | | | – | – |

+, positive; –, negative; +++/– – –, strong; ++/– –, moderate; +/–, weak.

bio-sounds. The spatial impression and naturalness are related to the volume of Bs and Gs layer. The calmness and pleasantness are related to all five parameters in the model.

## 5. DISCUSSION

In our view, nature sounds can both inform and create calm. Firstly, nature sounds can support clam technology with its subtleness and naturalness. Calm technology (Mark and Brown, 1997) aims to maintain the user's awareness of the displayed information without overburdening. Like other everyday sounds used as *Auditory icons*, nature sounds are intuitive, familiar and tend to engage the periphery of people's attention. NS can create ambient awareness. For example, NSs can be applied to ambient displays in public space. The man-made NS is mixed with the real soundscape in the space, and the nature sounds could respond to the input data source or adapt to the inhabitants in the space. The inhabitants can be aware of the information through general feelings toward the acoustic environment without taking them out of their environment or task. The slow changes in the perceptual attributes of soundscape require a small amount of attention. Therefore, the NS model-based interface is suitable to communicate status. The richness and steadiness of the NS can be manipulated to present some slow-changing status information, such as the temperature of CPU or the stress level of an office worker.

Secondly, nature sounds can be a desirable audio content to enhance the user's calmness and relaxation with the interfaces, especially in the applications for rehabilitation, stress management, relaxation practice and healthcare. Most nature sounds are pleasant and have a therapeutic effect due to its ability to foster the experience of calmness and relaxation. In previous studies (Eggen and Van Mensvoort, 2009; Ishii *et al.*, 1998; Mynatt *et al.*, 1997), the auditory displays were created with nature sounds by *Auditory icons* and *PM*, in which the data was directly linked to the parameters of the individual sounds. These audio displays are effective in information delivery, but few of them focused on creating a 'calm' soundscape. In this study, we developed a model of NS through an empirical study. The model helps to select, organize and manipulate the sounds within a 'nature theme,' and provides a means to manipulate the acoustic parameters of certain nature sounds and finally generate a 'soundscape' with more calmness and pleasantness.

As suggested by Blattner *et al.* (1989), Eggen (2016) and Gaver (1993), people can perceive the sounds at different levels. Beyond a basic-level auditory event, people can also hear more complex, structured combinations of basic-level events and perceive these combinations as the overall attributes or characteristics of the environment. MBS provides a possibility to 'edit' these attributes at higher semantical levels as stressed by Hermann and Hunt (2005). In this study, the NS model aims to link the data to the overall attributes of the soundscape, enabling the auditory display to be manipulated at 'perceptual' or 'experiential' levels. Thus, the listener can extract information by holistically listening to the NS, and also zoom in into a specific sound for more detailed information. The sound perception at different levels allows the audio display to move easily between the periphery and the center of our attention. In an NS-model-based auditory display, individual sounds can be chose to be 'expressive' and 'functional' by retaining close mapping between data and specific acoustic parameters. For instance, a certain *type* of bird sound (i.e. cuckoo) can indicate a discrete data event (i.e. an outlier), and the *volume* of wind sound can represent a continuous flow of data. These detailed sounds communicate explicitly in the center attention of the listener. Moreover, the data can also control several sounds jointly to shape the soundscape perceptually to be discriminable and inform the listener calmly in the periphery, for example, the *richness* of the whole soundscape conveys some supplementary information.

This idea has been explored in term of 'ecology of sounds.' Gaver *et al.* designed an ecology of *auditory icons* for *ARKola factory*, where a number of sounds worked together to convey information about a complex, demanding simulation task (Gaver *et al.*, 1991). As Gaver *et al.* (1991) puts it in his paper: 'an ecology of sounds can be designed that can be heard together as an overall plant noise or attended to separately to obtain information about individual machines.' From an ecological perspective, the individual sounds are not created and manipulated in isolation but as part of a sound ecology so that the listeners could experience all sounds as a unity. Therefore, in the design of auditory display with multiple sounds, the coherence and consistency should always be stressed. In this study, the concept of NS provides a coherent context for selecting, tuning, modifying and mixing various nature sounds in such a way that the resulting soundscape can be perceived and experienced as a harmonious sonic environment.

In the proposed NS structure, the Bs layer consists of various discrete Bss (birds songs in this study) and has five control parameters. The volume and density are related to the strength and frequency of the sounds. The other three parameters are related to dynamics of the soundscape. Instead of modulating the type, rhythm or direction of the individual sound, we use the variation of these sound properties as the parameters. The results show that only the changes on the type variations can be perceived regarding the richness and steadiness of the soundscape. The rhythm and direction variations of different levels were difficult to detect. As reported in the studies about 'audio interface', timbres of different instruments are subjectively easy to tell apart (Brewster *et al.*, 1993). Therefore, timbre is a common audio property being used in the design of *Earcons*. In other studies (Feige, 2009; Hoggan *et al.*, 2009; Yu *et al.*, 2015) about the rhythm-based interfaces, while the sounds maintain a constant timbre and direction, and the rhythmic pattern of sounds could successfully communicate information to listeners. In this study, timbre (the type of bird sounds) and direction were not fixed but varied in the moderate (default) range; we think this explains the reason why the rhythm variations are hard to detect in the soundscape. Therefore, we suggest that when there is only one type of Bs sound being activated, the rhythm and direction variations might be an effective parameter to influence the steadiness and spatial impression.

Regarding the use of NS for auditory display, we think there are still some issues need to be addressed. Firstly, the man-made NS should well match the context of use. To design an appropriate NS for public auditory display, the indoor/outdoor environment, the inhabitants' presence and the activities taking place in the public space should all be considered by designers. The mismatch between the 'man-made' indoor NS and the existing sonic environment could seriously hamper reaching a state of immersion in the environment. Secondly, the perceptions and preferences of certain 'genres' or sound ecologies might be influenced culturally and differentiated individually. Therefore, the factors of the listeners or inhabitants in a public space should also be considered in the selection of the nature sounds. Thirdly, according to the different emphasis on informative or experiential goals, the interface may need to constantly fine-tune the perceptual and experiential qualities of the acoustical mapping from data to the sounds. Thus, the NS can become functional when it is fed a dataset, or become decorative, as a natural and beautiful augment to the acoustic environment, when no dataset needs to be presented. Lastly, as most nature sounds are familiar to ordinary people, it is suggested to get the end-user actively involved in the design and evaluation of the NS regarding its usability and experiential qualities (Eggen *et al.*, 2017).

This study still has certain limitations. Firstly, as a guideline for design, the proposed three-layer NS structure could be refined with more details regarding the selection of audio content. In this NS structure, the Cs and Gs layer only has one sound source, and for Bs layer, only one type of species is involved (bird in this study). The reduced complexity might have a major impact on the perceptual and experiential qualities of the NS. Secondly, in this study, only one NS was constructed as the experimental material. In future research, the NS model could be evaluated and improved with more NS samples created with different nature sound content. Thirdly, for each acoustic parameter, only three levels were tested in our experiment. We suggest that the parameters should be tested with more levels, which might conclude with a linear relationship between the acoustic parameters and perceptual attributes of the soundscape.

## 6. CONCLUSION

In this study, we propose an NS model linking between the acoustic parameters of individual nature sounds and the perceptual attributes of the whole soundscape. The correlations between the acoustic parameters and the human perception of the NS can be used as an interface between the data and specific information in different context. The NS model offers the designers and practitioners a new tool to utilize nature sounds in the design of auditory displays, which could support the calm technology and enhance the user experience. Specifically, we believe the proposed model may contribute the fields of sonification and HCI in two ways. Firstly, the NS MBS offers new means for ambient display, in which data could be used to drive an adaptive acoustic environment. The NS MBS may put the auditory display in the periphery, occupy a small amount of attention and communicate information in a natural and elegant way. Secondly, the NS model can also be used in the design of non-speech audio interfaces to create calm and induce relaxation of the users.

## ACKNOWLEDGEMENTS

The authors would like to thank the members of Designed Intelligence group of Industrial Design department in Eindhoven University of Technology (TU/e) for their supports and assistance in the experiment. We would also like to thank the China Scholarship Council (CSC) for their support on the project.

## FUNDING

This research has been supported by the China Scholarship Council.

## REFERENCES


Alvarsson, J.J., Wiens, S. and Nilsson, M.E. (2010) Stress recovery during exposure to nature sound and environmental noise. Int. J. Environ. Res. Public Health, 7(3), 1036–1046. https://doi.org/10.3390/ijerph7031036.

Annerstedt, M., Jönsson, P., Wallergård, M., Johansson, G., Karlson, B., Grahn, P. and Währborg, P. (2013) Inducing physiological stress recovery with sounds of nature in a virtual reality forest—results from a pilot study. Physiol. Behav., 118, 240–250. https://doi.org/10.1016/j.physbeh.2013.05.023.

Bakker, S., van den Hoven, E. and Eggen, B. (2015) Peripheral interaction: characteristics and considerations. Pers. Ubiquitous Comput., 19(1), 239–254. https://doi.org/10.1007/s00779-014-0775-2.

Barrass, S. and Kramer, G. (1999) Using sonification. Multimed. Syst., 7(1), 23–31. https://doi.org/10.1007/s005300050108.

Bergstrom, I., Seinfeld, S. and Arroyo-Palacios, J. (2014) Using music as a signal for biofeedback. Int. J. Psychophysiol., 93(1), 140–149.

Blattner, M., Sumikawa, D. and Greenberg, R. (1989) Earcons and Icons: their structure and common design principles. Hum. Comput. Interact., 4(1), 11–44. https://doi.org/10.1207/s15327051hci0401_1.

Brewster, S.A., Wright, P.C., & Edwards, A.D.N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems, pp. 222–227. ACM. https://doi.org/10.1145/169059.169179

Buxton, W. (1989) Introduction to this special issue on nonspeech audio. Hum. Comput. Interact., 4(1), 1–9.

Cohen, J. (1993). 'Kirk here:': using genre sounds to monitor background activity. In INTERACT'93 and CHI'93 conference companion on Human factors in computing systems, pp. 63–64. ACM. https://doi.org/10.1145/259964.260073

Csapó, Á. and Wersényi, G. (2013) Overview of auditory representations in human-machine interfaces. ACM Comput. Surveys, 46 (2), 1–23. https://doi.org/10.1145/2543581.2543586.

De Coensel, B. and Botteldooren, D. (2006) The quiet rural soundscape and how to characterize it. Acta Acust. United Acust., 92 (6), 887–897.

DeLoach, A.G., Carter, J.P. and Braasch, J. (2015) Tuning the cognitive environment: sound masking with 'natural' sounds in open-plan offices. J. Acoust. Soc. Am., 137(4), 2291–2291. https://doi.org/10.1121/1.4920363.

Eggen, B. (2016) Interactive soundscapes of the future everyday life. Peripheral Interaction. pp. 239–251. Springer International Publishing, Switzerland, https://doi.org/10.1007/978-3-319-29523-7_11.

Eggen, B., Van Den Hoven, E. and Terken, J. (2017) Human-centered design and smart homes: how to study and design for the home experience? Handbook of Smart Homes, Health Care and Well-Being, pp. 83–92. Springer International Publishing, Switzerland, https://doi.org/10.1007/978-3-319-01583-5_6.

Eggen, B. and Van Mensvoort, K. (2009) Making sense of what is going on 'around': designing environmental awareness information displays. Awareness Systems, pp. 99–124. Springer, London, https://doi.org/10.1007/978-1-84882-477-5_4.

Feige, S. (2009) Can You Feel It? – Using vibration rhythms to communicate information in mobile contexts. IFIP Conference on Human-Computer Interaction, pp. 800–803. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-03655-2_87.

Gable, T.M., Walker, B.N., Moses, H.R., & Chitloor, R.D. (2013). Advanced auditory cues on mobile phones help keep drivers' eyes on the road. In Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 66–73. ACM . https://doi.org/10.1145/2516540.2516541.

Gaver, W. (1986) Auditory icons: using sound in computer interfaces. Hum. Comput. Interact., 2(2), 167–177. https://doi.org/10.1207/s15327051hci0202_3.

Gaver, W.W. (1989) The sonicfinder: an interface that uses auditory icons. Hum. Comput. Interact., 4(1), 67–94.

Gaver, W.W. (1993) What in the world do we hear?: An ecological approach to auditory event perception. Ecol. Psychol., 5(1), 1–29. https://doi.org/10.1207/s15326969eco0501_1.

Gaver, W.W., Smith, R.B., O'shea, T., Xerox, R. & Europarc, C. (1991). Effective Sounds in Complex Systems: The ARKola Simulation. In Proceedings of the SIGCHI Conference on Human factors in Computing Systems, pp. 85–90. ACM.

Harris, J., Vance, S., Fernandes, O., Parnandi, A., & Gutierrez-Osuna, R. (2014). Sonic respiration: controlling respiration rate through auditory biofeedback. In Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems, pp. 2383–2388. ACM. https://doi.org/10.1145/2559206.2581233.

Heide, F.J. and Borkovec, T.D. (1983) Relaxation-induced anxiety: paradoxical anxiety enhancement due to relaxation training. J. Consult. Clin. Psychol., 51(2), 171–182. https://doi.org/10.1037/0022-006X.51.2.171.

Hermann, T. (2008) Taxonomy and definitions for sonification and auditory display. Proceedings of the 14th International Conference on Auditory Display (ICAD 2008). IRCAM, Paris, France.

Hermann, T., Drees, J.M., & Ritter, H. (2003). Broadcasting auditory weather reports - a pilot project. In Proceedings of International Conference on Auditory Display 2003 (ICAD 2003), pp. 208–211. Georgia Institute of Technology.

Hermann, T. and Hunt, A. (2005) An introduction to interactive sonification. IEEE Multimed., 12(2), 20–24. https://doi.org/10.1109/MMUL.2005.26.

Hermann, T. and Ritter, H. (1999) Listen to your data: model-based sonification for data analysis. Adv. Intell. Comput. Multimed. Syst., 8, 189–194.

Hoggan, E., Raisamo, R., & Brewster, S.A. (2009). Mapping information to audio and tactile icons. In Proceedings of the 2009 international conference on Multimodal interfaces, p. 327. ACM. https://doi.org/10.1145/1647314.1647382.

Hunt, A., Hermann, T., & Pauletto, S. (2004). Interacting with sonification systems: closing the loop. In Proceedings. Eighth International Conference on Information Visualisation, pp. 879–884. IEEE. https://doi.org/10.1109/IV.2004.1320244.

Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., & Yarin, P. (1998). ambientROOM: integrating ambient media with architectural space. In CHI 98 conference summary on Human factors in computing systems, pp. 173–174. ACM. https://doi.org/10.1145/286498.286652

Jagdish, D., Sawhney, R., & Gupta, M. (2008). Sonic Grid: an auditory interface for the visually impaired to navigate GUI-based environments. In Proceedings of the 13th international conference on Intelligent user interfaces, p. 337. Gran Canaria, Spain. ACM.

Kaplan, S. (1995) The restorative benefits of nature: toward an integrative framework. J. Environ. Psychol., 15(3), 169–182. https://doi.org/10.1016/0272-4944(95)90001-2.

Krause, B. (1987) Bioacoustics, habitat ambience in ecological balance. Whole. Earth. Rev., 57, 14–18.

Lundqvist, L.-O., Carlsson, F., Hilmersson, P. and Juslin, P.N. (2009) Emotional responses to music: experience, expression, and physiology. Psychol. Music, 37(1), 61–90. https://doi.org/10.1177/0305735607086048.

Maes, P.-J., Leman, M., & Lesaffre, M. (2010). A model-based sonification system for directional movement behavior. In Interactive Sonification Workshop (ISon).

Mark W., & Brown J.S. (1997). The Coming age of Calm Technology. Beyond Calculation, 75–85.

McCraty, R., Barrios-Choplin, B., Atkinson, M. and Tomasino, D. (1998) The effects of different types of music on mood, tension, and mental clarity. Altern. Ther. Health. Med., 4(1), 75–84.

Mynatt, E.D., Back, M., Want, R., & Frederick, R. (1997). Audio Aura: Light-weight audio augmented reality. In Proceedings of the 10th annual ACM symposium on User interface software and technology - UIST'97, pp. 211–212. ACM. https://doi.org/10.1145/263407.264218

Neuhoff, J.G., Wayand, J. and Kramer, G. (2002) Pitch and loudness interact in auditory displays: Can the data get lost in the map? J. Exp. Psychol., 8(1), 17–25. https://doi.org/10.1037/1076-898X.8.1.17.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H. and Pieretti, N. (2011) Soundscape ecology: the science of sound in the landscape. Bioscience, 61(3), 203–216. https://doi.org/10.1525/bio.2011.61.3.6.

Raimbault, M., Lavandier, C. and Bérengier, M. (2003) Ambient sound assessment of urban environments: field studies in two French cities. Appl. Acoust., 64(12), 1241–1256. https://doi.org/10.1016/S0003-682X(03)00061-6.

Ratcliffe, E., Gatersleben, B. and Sowden, P.T. (2013) Bird sounds and their contributions to perceived attention restoration and stress recovery. J. Environ. Psychol., 36, 221–228. https://doi.org/10.1016/j.jenvp.2013.08.004.

Sandell, G.J. (1996) Auditory display: sonification, audification, and auditory interfaces Gregory Kramer. Music Percept., 13(4), 583–591. https://doi.org/10.2307/40285703.

Schafer, M. (1993) The Soundscape: Our Sonic Environment and the Tuning of the World. Inner Traditions/Bear & Co, Vermont.

Smith, S., Levkowitz, H., Pickett, R.M., & Torpey, M. (1994). System for psychometric testing of auditory representations of scientific data. In International Conference on Auditory Display 2008 (ICAD 1994), pp. 217–225. Georgia Institute of Technology.

Ulrich, R.S., Simons, R.F., Losito, B.D., Fiorito, E., Miles, M.A. and Zelson, M. (1991) Stress recovery during exposure to natural and urban environments. J. Environ. Psychol., 11(3), 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7.

Yang, W. and Kang, J. (2005) Soundscape and sound preferences in urban squares: a case study in sheffield. J. Urban Des., 10(1), 61–80. https://doi.org/10.1080/13574800500062395.

Yu, B., Feijs, L., Funk, M., & Hu, J. (2015). Designing auditory display of heart rate variability in biofeedback context. In International Conference on Auditory Display (ICAD 2015) (pp. 294–298) Georgia Institute of Technology, Vermont.

Yu, B., Hu, J., Funk, M., & Feijs, L. (2016). A Study on User Acceptance of Different Auditory Content for Relaxation. In Proceedings of the Audio Mostly 2016 on - AM'16, pp. 69–76. ACM. https://doi.org/10.1145/2986416.2986418