PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Evaluation of preferred levels of natural sounds in situ environment through an augmented reality device

Bhan, Lam[1]; Joo-Young, Hong[1]; Zhen-Ting, Ong[1]; Kenneth Ooi[1]; Woon-Seng, Gan[1]; Samuel, Yeong[2]; Irene, Lee[2]; Sze-Tiong, Tan[2]

[1] School of Electrical & Electronic Engineering, Nanyang Technological University, 639798, Singapore

[2] Building & Research Institute, Housing & Development Board, 738973, Singapore

ABSTRACT

Many studies have investigated the effect of natural sounds on improving soundscape quality under controlled laboratory conditions. The laboratory settings are, however, limited in reflecting real-life settings, resulting in low ecological validity. In situ experiments can provide a realistic representation of the real-life settings, which can guarantee high ecological validity. Recently, augmented reality (AR) technology provides an avenue for virtual augmentation of visual and audio elements in the real environment. Therefore, this study investigates the preferred levels of two types of natural sounds (bird and stream) in a noisy in situ environment using an AR head-mounted device (HMD). Two locations in an outdoor residential area, with background traffic noise levels varying between 60 to 70 dBA, were selected for the in situ experiment. Participants were instructed to adjust the volume of the natural sounds playing via the loudspeakers of the AR HMD to their most preferred levels at each location while considering the perceived loudness of traffic and overall soundscape quality in terms of pleasantness of the overall sound environment. Based on the results of the experiment, the effects of types of natural sounds and background noise levels on determining the preferred levels of natural sounds were examined.

Keywords: Soundscape, Augmented Reality, Natural Sounds, Ecological Validity

1 INTRODUCTION

In the soundscape approach, pleasant natural sounds are employed as design elements to mask traffic noises in outdoor areas [1]. Various studies have investigated the effects of natural sounds (e.g., water sound and birdsongs) for enhancing acoustic environment [2–6]. It has been found that sound level of natural sounds is one of the important acoustic factors to design desirable soundscape. Some have found that similar or 3 dB lower sound pressure levels (SPLs) of natural sounds are preferred in combination with urban noises [7,8].

However, most of the previous studies on the effects of natural sounds have been conducted in controlled laboratory conditions. Laboratory experiments allow for the design and control of experimental variables, which can yield more consistent results and allow for repeatability. However, laboratory settings may be limited in reflecting real-world conditions, resulting in low ecological validity, and limiting the generalization of findings to real-life scenarios. In contrast, in situ experiments can provide a realistic representation of the real-life settings, which can guarantee high ecological validity [9–12]. Since in situ experiments does not allow for direct control of independent variables such as natural sound levels or type of natural sound in the field, these methods have been primarily used to characterize the existing acoustic environment [13–15]. Recently, the consumerization of augmented reality (AR) technology has enabled the virtual augmentation of visual and audio elements in the real environment [9]. Thus, with an AR device, natural sounds can now be introduced and controlled, and their preferred sound levels can be investigated in realistic situations to validate the findings of previous laboratory studies.

Hence, this study investigates the preferred levels of two types of natural sounds (bird and stream) in a noisy in situ environment using an AR head-mounted device (HMD). Preferred natural sound levels may be affected by the characteristics of not only the surrounding noise, but also the natural sound type itself. Thus, in situ experiment was designed to consider varying traffic noise levels and type of natural sounds pair with visual elements representing the sound sources.

2 METHOD

2.1 Experimental Design

As shown in Figure 1, two locations were selected in the vicinity of the student hostels that are flanked by an expressway in order to vary background traffic noise levels. Since location B was closer to the expressway than the location A, A-weighted SPL at location A was approximately 4–5 dB louder than that at location B. Throughout the experiment, the sound levels experienced by the participants were captured using a calibrated binaural microphone (Brüel & Kjær TYPE 4101-B, Denmark) and a portable data acquisition system (HEAD acoustics SQobold, Germany). To evaluate the augmentation of a natural audio-visual stimuli, participants were also fitted with an untethered AR HMD (Microsoft Hololens, USA), as shown in Figure 2.

At each location, the traffic levels were firstly evaluated for approximately one minute in terms of the perceived loudness and overall pleasantness. Subsequently, a natural sound (i.e. stream sound or sparrow call) along with a hologram (i.e. generic speaker or bird/fountain) were projected through the AR HMD. Since the traffic sound levels are fluctuating, the traffic sound levels were re-evaluated before each natural sound and hologram pair. Hence, in each test case the participants were chronologically instructed to: (I) evaluate the traffic scene, (II) adjust the sound levels of a natural sound that is augmented to the same traffic scene to the most desirable level, and (III) evaluate the combined scene. These three steps were repeated a total of 4 times (2 natural sounds × 2 visual stimuli) in a randomised order at each location.

(a) (b)

Figure 1 – Two outdoor locations where the soundscape evaluations were conducted: (a) location A: pedestrian walkway beside a busy road parallel to an expressway behind the greenery in the background, and (b) location B: an open area facing a small section of a minor road parallel to an expressway behind the greenery in the background.

Figure 2 – Equipment setup for the in situ AR experiment

2.2 Stimuli

Two natural sounds were used in this preliminary trial, a sparrow call and a stream sound which were evaluated as preferable natural sound in a previous study [16]. The audio stimuli were spatialised to emulate a point sound source emitting from the hologram position located 2 m away at zero azimuth and elevation (Microsoft Windows SDK Spatializer Plugin, USA). The natural sound stimuli were presented via the downward-firing speakers of the AR HMD.

In step II of the evaluation described in 2.1, the sound levels of the natural sounds were adjusted via a Bluetooth keyboard. Participants were allowed to adjust the normalised sound levels from 0 (silence) to 100 (loudest), where 100 represents the 1-min equivalent sound level of approximately 80 dB for both stimuli.

To obtain the contribution of the natural sound stimuli in the absence of the ambient sounds, participants were brought to a quiet room after the outdoor evaluation. All the natural sound stimuli were reproduced at the participants' chosen levels whilst ensuring the downward-firing speakers of the AR HMD were at the same position as measured during the outdoor evaluations.

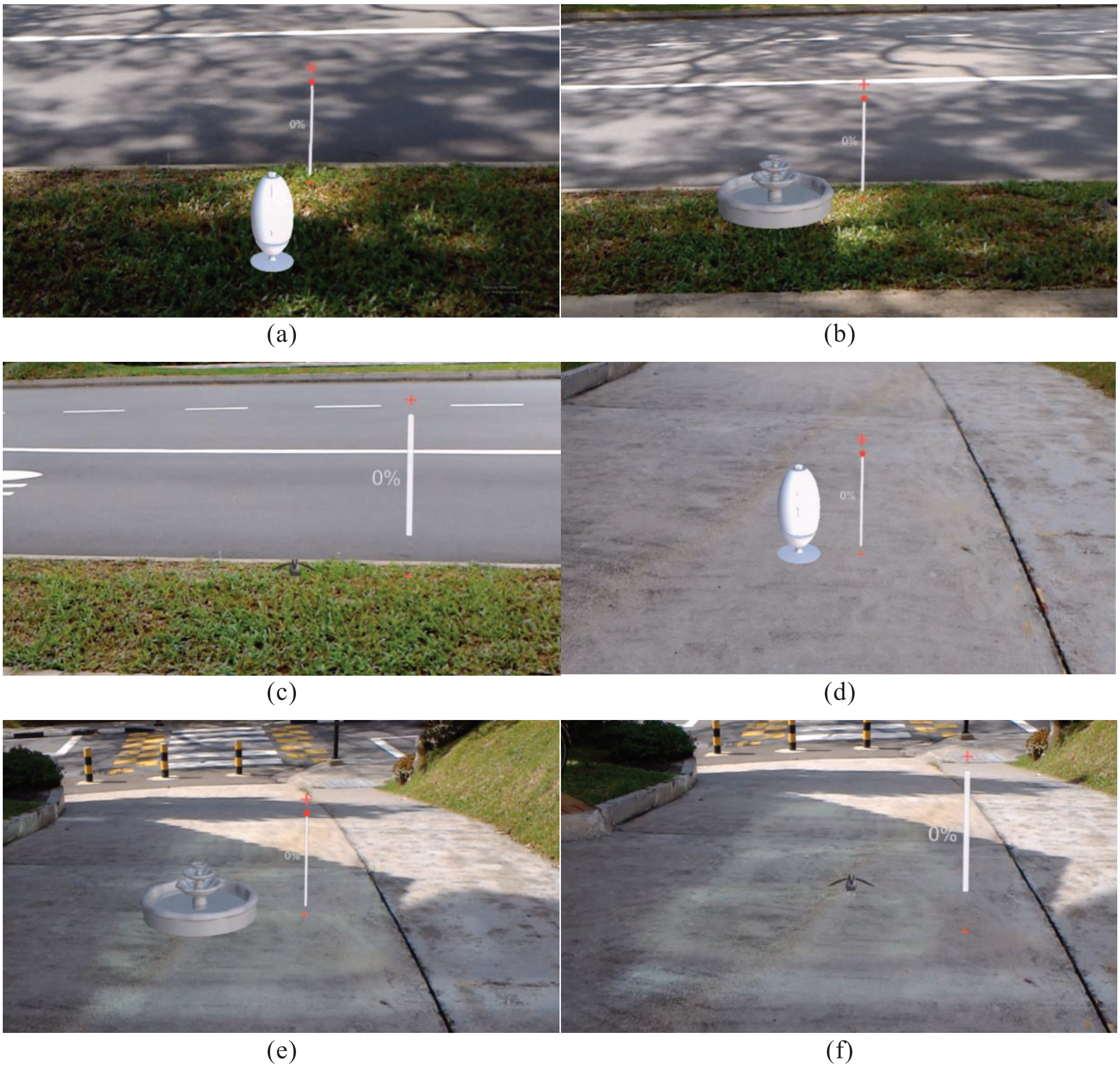


Figure 3 – Actual rendering of the 3D holograms through the augmented reality headset of (a) a speaker, (b) a fountain, and (c) a bird, at location A. Renders of the 3D holograms of (d) a speaker, (e) a fountain, and (f) a bird, at location B.

The natural sounds were presented to the participants in conjunction with holographic projections that mimic the presence of "real" sound sources in the outdoor environment. Two types of holograms were employed in this study, a generic sound source (i.e. an ovoid speaker) and a corresponding natural sound source (i.e. water fountain hologram projecting water sounds, a sparrow hologram projecting sparrow calls). The real-world render as captured through the AR HMD of the generic, fountain and bird hologram in location A are shown in Figure 3 (a) to (c), respectively. The same holograms projected through the AR HMD in location B are shown in Figure 2 (d) to (e), respectively.

2.3 Procedure

For this preliminary study, three participants with normal-hearing (0.125, 0.5, 1, 2, 3, 4, 6, and 8 kHz average pure-tone thresholds < 20 dB HL) were recruited for this study. Formal ethical approval was granted by the institutional review board of Nanyang Technological University (NTU) for this study (IRB-2018-02-024). Informed consent was obtained from each participant prior to the start of the experiments.

For each of the four natural sound and hologram combinations at each location, the participants were instructed to provide subjective assessments in three steps, as described in 2.1. In step I, the participants evaluated the overall soundscape quality and perceived loudness of the traffic sound for a duration of 1 min. The ambient environment during the 1-min evaluation was captured through the binaural microphone worn by the participant. This 1-min capture represents the equivalent sound pressure level before the natural sound augmentation.

Immediately after, in step II, a randomly chosen natural sound and hologram pair was projected through the AR HMD to the participant. Participants were instructed to adjust the sound levels of the natural sound to their most comfortable levels while considering the overall soundscape quality and perceived loudness of the traffic sound.

Lastly, in step III, the participants were instructed to re-evaluate the overall soundscape quality and perceived loudnesss of the traffic sound for 1 min. Similar to step I, the 1-min evaluation of the ambient environment was captured with the binaural microphone on the participant. Hence, the signal-to-noise ratio for each test case is determined by

$$SNR = L_{eq,1\text{ min, traffic}} - L_{eq,1\text{ min, traffic + natural sound}} \quad (1)$$

where, $L_{eq,1\text{ min, traffic}}$ refers to the 1-min equivalent A-weighted sound pressure level in dB captured during step I, and $L_{eq,1\text{ min, traffic + natural sound}}$ refers to the 1-min equivalent A-weighted sound pressure level in dB captured during step III.

3 RESULTS

Figure 4 (a) and (b) show the mean $L_{eq,1\text{ min}}$ values of traffic, preferred bird and water sound levels as a function of location A and B, respectively. Mean background noise levels ($L_{eq,1\text{ min}}$) at location A and B were 66.0 dB (SD=1.8 dB) and 73.2 dB (SD=1.9 dB) showing approximately 7 dB difference between the locations. As shown in Figure 4(a), the mean preferred level for the birdsong was ~73 dB at both locations A and B. Meanwhile, the mean preferred level for water sound was ~70 dB, which is slightly lower than that of the birdsong as shown in Figure 4(b). There were no significant differences between the generic (i.e., speaker) and representative (i.e., bird/fountain) holograms.

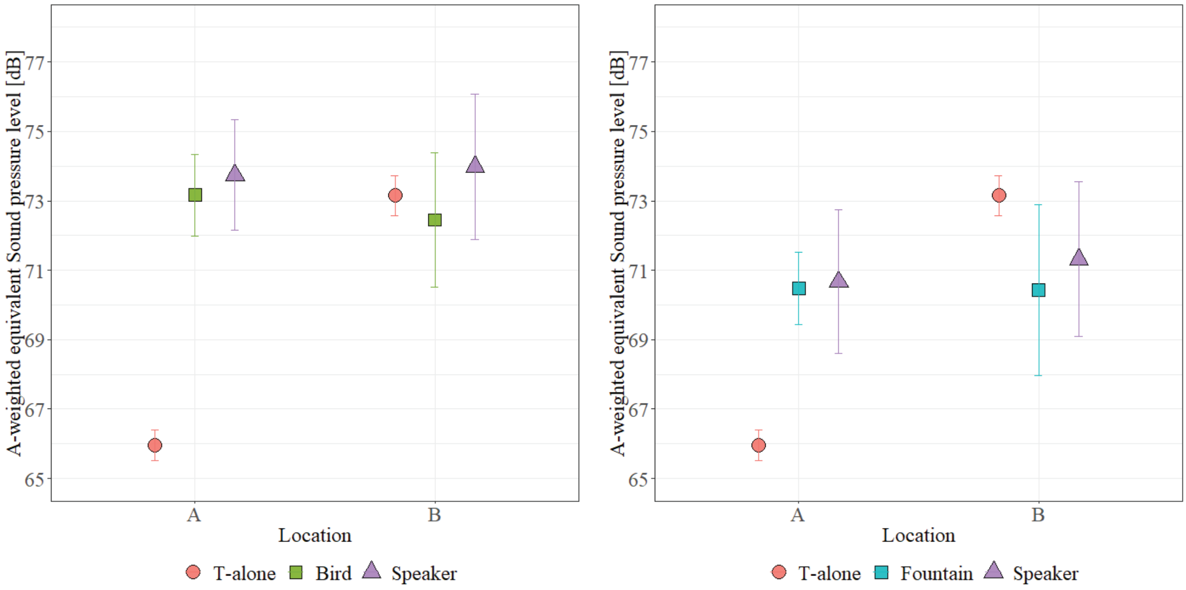
The preferred SNRs as a function of the traffic noise levels are shown in Figure 5. There were significant negative correlations between the preferred SNRs and traffic noise for both birdsong and water sounds. This demonstrates that as the background noise level increases, participants preferred lower natural sound levels than the background noise levels and vice versa.

The prediction models for the preferred SNR for the birdsong and the water sound were obtained from simple linear regression analyses as shown in equations (2) and (3). The coefficient of determination (R^2) for the models of the birdsong and water sounds were 0.66 and 0.58 (p < 0.001).

$$SNR_{bird} = -0.98\, L_{eq,1\text{ min, traffic}} + 71.81, R^2 = 0.66, p < 0.001 \quad (2)$$

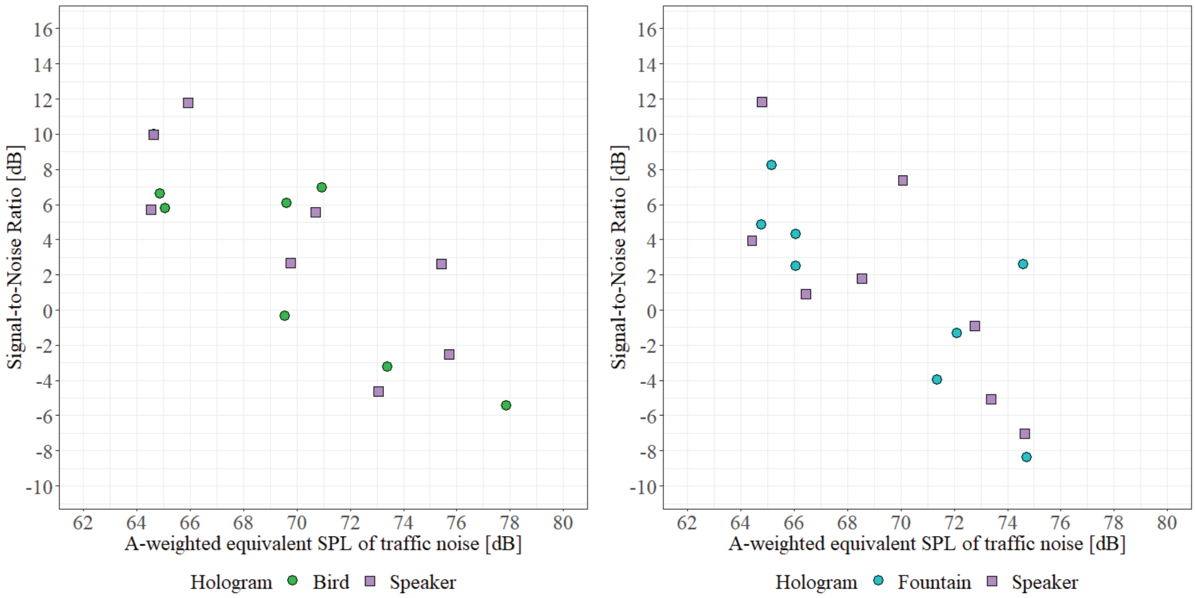
$$SNR_{water} = -1.07\, L_{eq,1\text{ min, traffic}} + 75.44, R^2 = 0.58, p < 0.001 \quad (3)$$

where, $L_{eq,1\,\text{min, traffic}}$ refers to the 1-min equivalent A-weighted sound pressure level.



(a) (b)

Figure 4 – Mean A-weighted equivalent sound pressure levels of traffic noise and masker sounds as a function of the location (a) birdsong and (b) water sound. Error bars indicate standard errors.



(a) (b)

Figure 5 – Signal-to-Noise Ratio as a function of traffic noise levels: (a) birdsong and (b) water sound. Error bars indicate standard errors.

4 CONCLUSIONS

Although there were limited subjective responses in this study, a clear negative correlation was observed between the signal-to-noise ratio and the sound pressure levels of the traffic noise for both types of natural sounds. Observations at higher traffic noise level seem to agree with previous studies where a lower level of water or bird sound was preferred [6,7]. However, at lower traffic sound levels, the trend deviates from previous studies. It is also worth noting that there was no significant difference between the generic and natural visual representations of the sound source, which suggests that integration of speakers into the urban environment is a viable soundscape design option. However, these observations are based only on the perception of perceived loudness and pleasantness. Hence, this experiment should be expanded to include rating responses of perceived loudness and pleasantness from a larger pool of participants.

ACKNOWLEDGEMENTS

This material is based on research/work supported by the Singapore Ministry of National Development and National Research Foundation under L2 NIC Award No.: L2NICCFP2-546 2015-5. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Singapore Ministry of National Development and National Research Foundation, Prime Minister's Office, Singapore. Joo Young Hong is currently supported by the 2017 Lee Kuan Yew Postdoctoral Fellowship (Grant no. 50025000).

REFERENCES

1. Kang, J.; Aletta, F.; Gjestland, T.T.; Brown, L.A.; Botteldooren, D.; Schulte-Fortkamp, B.; Lercher, P.; Van Kamp, I.; Genuit, K.; Fiebig, A.E.; et al. Ten questions on the soundscapes of the built environment. *Build. Environ.* **2016**, *108*, 284–294.
2. Galbrun, L.; Calarco, F.M. a Audio-visual interaction and perceptual assessment of water features used over road traffic noise. *J. Acoust. Soc. Am.* **2014**, *136*, 2609.
3. Hong, J.Y.; Jeon, J.Y. Designing sound and visual components for enhancement of urban soundscapes. *J. Acoust. Soc. Am.* **2013**, *134*, 2026–36.
4. Nilsson, M.E.; Alvarsson, J.; Rådsten-Ekman, M.; Bolin, K. Auditory masking of wanted and unwanted sounds in a city park. *Noise Control Eng. J.* **2010**, *58*, 524.
5. De Coensel, B.; Vanwetswinkel, S.; Botteldooren, D. Effects of natural sounds on the perception of road traffic noise. *J. Acoust. Soc. Am.* **2011**, *129*, EL148-L153.
6. Hong, J.Y.; Lam, B.; Ong, Z.; Gupta, R.; Gan, W.; Chong, S.H.; Feng, J. Appropriate levels of natural sounds to enhance soundscapes in urban areas. *INTER-NOISE NOISE-CON Congr. Conf. Proc.* **2017**, 49–53.
7. Jeon, J.Y.; Lee, P.J.; You, J.; Kang, J. Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds. *J. Acoust. Soc. Am.* **2010**, *127*, 1357–1366.
8. Galbrun, L.; Ali, T.T. Acoustical and perceptual assessment of water sounds and their use over road traffic noise. *J. Acoust. Soc. Am.* **2013**, *133*, 227–237.
9. Hong, J.Y.; He, J.; Lam, B.; Gupta, R.; Gan, W.-S. Spatial Audio for Soundscape Design: Recording and Reproduction. *Appl. Sci.* **2017**, *7*, 1–21.
10. Aletta, F.; Kang, J.; Axelsson, Ö. Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plan.* **2016**, *149*, 65–74.
11. Guastavino, C.; Katz, B.F.G.; Polack, Jean-Dominique; Levitin, D.J.; Dubois, D. Ecological Validity of soundscape reproduction. *Acta Acust. united with Acust.* **2005**, *91*, 333–341.
12. Hong, J.Y.; Lam, B.; Ong, Z.-T.; Ooi, K.; Gan, W.-S.; Kang, J.; Feng, J.; Tan, S.-T. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Build. Environ.* **2019**, *149*, 1–14.
13. Hong, J.Y.; Jeon, J.Y. Relationship between spatiotemporal variability of soundscape and urban morphology in a multifunctional urban area: A case study in Seoul, Korea. *Build. Environ.* **2017**, *126*, 382–395.
14. Jeon, J.Y.; Hong, J.Y.; Lee, P.J. Soundwalk approach to identify urban soundscapes individually. *J. Acoust. Soc. Am.* **2013**, *134*, 803–812.
15. Axelsson, Ö.; Nilsson, M.E.; Hellström, B.; Lundén, P. A field experiment on the impact of sounds from a jet-and-basin fountain on soundscape quality in an urban park. *Landsc. Urban Plan.* **2014**, *123*, 46–60.
16. Hong, J.Y.; Lam, B.; Ong, Z.T.; Gupta, R.; Gan, W.S. Suitability of Natural Sounds To Enhance Soundscape Quality in Urban Residential Areas. In Proceedings of the 24th International Congress on Sound and Vibration; London, UK, 2017.