Perceived width evaluation on interpolated line sources in a virtual urban square*

1st Chunyang Xu
IEDE, The Bartlett
University College London
London, UK
chunyang.xu.18@ucl.ac.uk

2nd Huan Tong
IEDE, The Bartlett
University College London
London, UK
huan.tong.18@ucl.ac.uk

3rd Jian Kang
IEDE, The Bartlett
University College London
London, UK
j.kang@ucl.ac.uk

Abstract—The aim of this study is to investigate the relationship between the perceived width and different interpolation setups of a line source, and whether it is possible to reduce interpolations of a line source in a virtual open urban space. In order to auralise sound sources efficiently and accurately, the subjective evaluation was carried out in accordance with Virtual Reality (VR) with spatial audio. The results show that when the audible angle between adjacent interpolation points of a line source is lowered to 1°, the perceived width of the line source will significantly improve under VR experience. When the angle is less than 1°, the variation in points interpolation does not significantly affect the perceived width. This improvement may enhance the immersion during subjective evaluation, thus creating a more realistic experience.

Index Terms—sound environment, virtual reality, interpolation, subjective evaluation, auralisation

I. INTRODUCTION

We are surrounded and enveloped by a wide variety of sounds every day. These sounds are produced by different people or objects in diverse environments. Different types of sounds vary in frequency contents, spatial positions, and source volume in urban spaces with various geometric shapes. Depending on the acoustic performance of different urban spaces, such as urban squares and streets, many studies have been carried out to simulate the urban sound environment by various methods [1, 2, 3, 4]. These studies parametrically investigate the characteristics of sound propagation in urban spaces in terms of sound level or frequency.

For sound reproduction, the subjective perception of human ears is an important reference for judging the quality of auralisation design. Soundscape researchers, therefore, place a primary priority on the perception of urban sounds, evaluating the entire sound environment as a whole [5, 6, 7]. A number of studies have been carried out on sound playback methods and signal processing during propagation combined with subjective evaluation, e.g., refs. [8, 9, 10, 11, 12]. In the field of environmental acoustics or soundscape, few studies combine simulation with subjective evaluation of the design of sound sources. A spatial audio technique for area and volumetric sound sources was developed in 2016 on the basis of Head-Related Transfer Function (HRTF), and Schissler et al. [13]

utilised Monte Carlo projection to sample sound sources with the use of orthonormal basis functions to replace the analytical solutions in the source projection function and the HRTF.

In urban sound environment reproduction, many sound sources are considered to be point sources in the far-field [14], and their geometry is simplified to a point in Cartesian coordinates. For urban open spaces, however, the near-field experience of these sound sources is equally important as well. When planning a square park, to place a fountain of a certain length, for instance, designers typically consider the aesthetic characteristics and public engagement first, and then, they need to consider how this sound can fit into the overall soundscape within a close or far distance. The simulation of such sound sources in urban spaces is therefore an ongoing issue in auralisation research. Line sources are one of the common non-point sources in the audio industry. The study of the subjective impressions, especially the perceived width, of such non-point sound sources in urban spaces will allow researchers or engineers to reproduce the acoustic environment more accurately. The introduction of VR also offers an immersive way to verify the ecological validity of auralised sounds [15].

At this point, the issue is whether we can rationally interpolate this kind of line sources so that the sound matches the vision in the near-field under a virtual experience. This study aims to investigate the effect of different interpolations of line sources on perceived width under virtual environments through subjective evaluation.

II. METHODS

A. Scene visualisation and animation

In order to assess the perceived width of line sources under virtual experience, an urban space was pre-defined to place various sound sources. A garden square of 6,400 m² (80 m×80 m) was created in virtual reality as shown in Fig. 1. The buildings and other installations were modelled using SketchUp Pro 2018. This kind of configuration is one of the classic rectangular garden squares, and it is partially paved and partially grassy.

All models were imported into the game engine (Unity). Unity was used to synthesise animation and visual rendering. To provide an immersive experience, a limited number of characters were placed in the square with their individual

Fig. 1. The view of the reproduced garden square.

activities attached with animation. These extra characters would not get close to the user location, and would keep a distance of more than 30 m away from the user. To further enhance immersion, the trees and grass in the square sway slightly with the wind. The lighting condition was set to a rational solar zenith angle and illuminance according to an appropriate geometrical location.

B. Auralisation and VR synthesis

To investigate the perceived width of different sounds in the square, several typical sounds were chosen, including the voices of a group of people talking (voice), sound from a water blade (water), and construction noise (construction). The reproduced models corresponding to these sounds are shown in Fig. 2. Because of England's lockdown regulations on non-essential fieldwork, the royalty-free sounds were chosen from Adobe Audition Sound Effects rather than in situ sound recordings. These sounds have a high signal-to-noise ratio. The sample rate of the three sounds is 44.1 kHz, and the depth is 16 bit.

To explore the effect of line source interpolation on subjective evaluation, the line source needed to be interpolated with discrete point sources. The assumed length of the line sound source is 10 m. There are three cases in this study, i.e., 5, 21, and 101 point sources to simulate a sound source. The mean distances between the point sources for these three cases are 2.5 m, 0.5 m and 0.1 m.

There are two observation points at 5 m and 30 m away from the sound source, as shown in Fig. 3, and both observation points are moving in parallel to the line source at a speed of 0.5 m/s for a total time of 10 s, passing symmetrically through the centre of the source. An approximation of audible angles for these two observation points was made. For the near case of 5 m, the angle between two adjacent point sources for the observer (henceforth to be referred to simply as 'spatial resolution') is 22.5°, 4.5° and 0.9° for the three interpolations. For the far case of 30 m, the spatial resolution is around 4.7°, 0.9° and 0.2° for the three interpolations.

According to different numbers of point sound sources, the volume for these sounds at the receiving location was
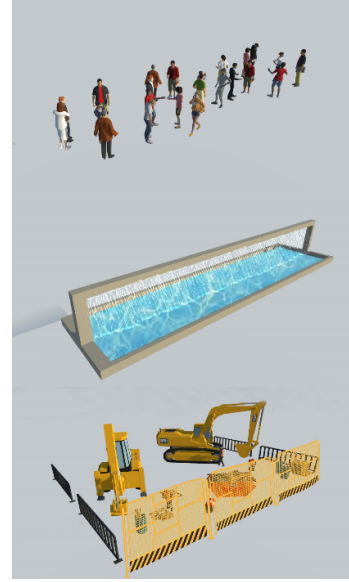

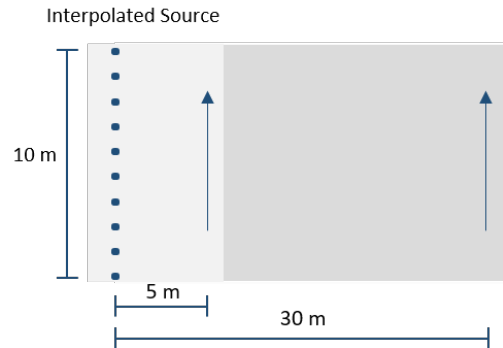Fig. 2. Reproduced people, water blade, and construction site (from top to bottom)

Fig. 3. Two observation distances at 5 m and 30 m to the interpolated source

calibrated to the same level. The participants could turn their heads, but they could not experience the spatial displacement due to first-order Ambisonics. The whole VR videos were recorded on a high-performance desktop PC (Inter Core i7-9700k, 32 GB of RAM and NVIDIA GeForce RTX 2080) to ensure high resolutions of recordings. The recorded videos consist of 30 frames per second.

C. Subjective test

Thirty-four participants took part in the subjective evaluation. They conducted the subjective evaluation in a poorly lit, quiet room. They downloaded the videos, incorporating the audio, on their phones. The participants heard the audio using a wired earphone connected to their smartphones. Through a VR cardboard, they could watch the VR videos.

Each video with one interpolated sound lasts 10 seconds. It was easy to hold the lightweight VR cardboard, and the participants could give their subjective ratings on their laptops or tablets after taking off the VR cardboard. They were asked

to rate the perceived width (narrow—wide) in a structured questionnaire for each sound with the question: How wide is this sound in this environment? The questionnaire adopted the continuous rating scale that has been widely used in subjective evaluation studies of the acoustic environment [8, 16, 17]. Each participant watched 18 videos (three interpolations, three sounds, and two distances) in total. The sequence of different interpolations was randomised. The participants were not informed how many points were applied to each sound.

III. RESULTS

The subjective ratings of width for three types of sounds are shown in Fig. 4. When the distance from the sound source is at 5 m, the width ratings show a visible increase for three types of sounds with 101 points. When the distance from the sound source is at 30 m, the rating difference between 5 and 21 points is noticeable. Based on the analysis of Mauchly's test of sphericity, the results of different interpolations satisfied the sphericity hypothesis (Mauchly's W= 0.952, Sig.= 0.070). Thus, the within-subjects effect of interpolations is considered to be statistically significant (F= 19.256, Sig.< 0.001) when sphericity is assumed.

To investigate the differences between the interpolations, the pairwise comparison was made through the repeated measures ANOVA shown in Table I. When the distance is at 5 m, the rating differences between 21 and 101 points are significant for all types of sounds (speech Sig.=0.004, water Sig.=0.002, and construction Sig.=0.020). With consideration of the spatial resolution of sound, the rating difference between 4.5° and 0.9° is significant. The rating differences between 5 and 101 points show statistical significance for speech (Sig.=0.007) and water (Sig.= 0.004). When the distance is at 30 m, the rating differences between 5 and 101 points (4.7° and 0.2°) are significant for all types of sounds (speech Sig.=0.034, water Sig.=0.005, and construction Sig.= 0.002). In addition, the rating differences between 5 m and 30 m are also significant under 5 points for all types of sounds (speech Sig.=0.050, water Sig.=0.009, and construction Sig.=0.004). The rating differences between 5 m and 30 m are also significant under 101 points for speech (Sig.=0.013) and water (Sig.=0.008). Some participants responded that they could distinguish the location of individual point sources of some videos, and these sounds are not sufficiently immersive in a VR experience. In psychoacoustics, the Minimum Audible Angle (MAA) is about 1° [18]. For the present study, where the observer was moving, the spatial resolution was dynamic and approximate. The results of the evaluation are still consistent with the MAA. When the audible angle is less than 1°, the perceived width for such line sources does not vary significantly under virtual urban environments.

IV. CONCLUSIONS

This study interpolated line sources in an outdoor sound environment and assessed the perceived width of different sounds through subjective evaluation under VR. The results of this study reveal that: when the points of the interpolated
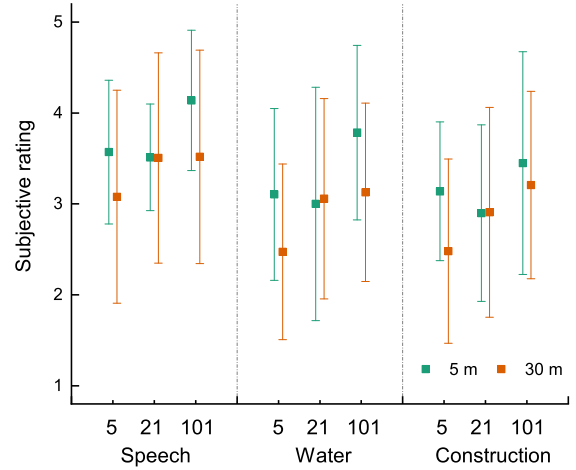

Fig. 4. The view of the reproduced garden square.

TABLE I
PAIRWISE COMPARISONS FOR PERCEIVED WIDTH RATINGS UNDER DIFFERENT INTERPOLATIONS, DISTANCES AND SOUND TYPES

Sound	Point	Mean Difference	Std. Error	Sig.
Distance=5 m				
Speech	5*21	0.058	0.225	0.798
	5*101	-0.569*	0.203	0.007
	21*101	-0.627*	0.210	0.004
Water	5*21	0.105	0.257	0.685
	5*101	-0.680*	0.225	0.004
	21*101	-0.785*	0.237	0.002
Construction	5*21	0.241	0.226	0.290
	5*101	-0.311	0.231	0.182
	21*101	-0.552*	0.232	0.020
Distance=30 m				
Speech	5*21	-0.426	0.225	0.063
	5*101	-0.439*	0.203	0.034
	21*101	-0.012	0.210	0.953
Water	5*21	-0.584*	0.257	0.026
	5*101	-0.655*	0.225	0.005
	21*101	-0.071	0.237	0.765
Construction	5*21	-0.427	0.226	0.063
	5*101	-0.727*	0.231	0.002
	21*101	-0.299	0.232	0.202

source form a sufficiently small audible angle (<1° in this study), the perceived width of the line source will significantly improve under VR experience. This increase in the perceived width may result in a more immersive experience for the virtual sound experience.

Overall, this work suggests that sufficient points of an interpolated line source should be ensured to avoid discriminating interpolation points. When reproducing such non-point sources of sound, the adjacent points in interpolated sources should keep an angle to the observer of less than 1°, so that the entire source can be perceived as a whole. When the audible angle is less than 1°, the perceived width for such line sources does not vary significantly under virtual urban environments. The result offers positive guidance on soundscape evaluation, game design, urban planning and other industries involved in sound design and auralisation under VR.

REFERENCES

[1] J. Kang, "Numerical modelling of the sound fields in urban streets with diffusely reflecting boundaries," *Journal of Sound and Vibration*, vol. 258, no. 5, pp. 793–813, 2002.

[2] ——, "Sound propagation in interconnected urban streets: a parametric study," *Environment and Planning B: Planning and design*, vol. 28, no. 2, pp. 281–294, 2001.

[3] ——, "Sound propagation in street canyons: Comparison between diffusely and geometrically reflecting boundaries," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1394–1404, 2000.

[4] ——, "Numerical modeling of the sound fields in urban squares," *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3695–3706, 2005.

[5] ——, *Urban sound environment*. CRC Press, 2006.

[6] J. Kang, F. Aletta, T. T. Gjestland, L. A. Brown, D. Botteldooren, B. Schulte-Fortkamp, P. Lercher, I. van Kamp, K. Genuit, A. Fiebig *et al.*, "Ten questions on the soundscapes of the built environment," *Building and Environment*, vol. 108, pp. 284–294, 2016.

[7] J. Kang and B. Schulte-Fortkamp, *Soundscape and the Built Environment*. CRC press, 2016.

[8] C. Xu and J. Kang, "Soundscape evaluation: Binaural or monaural?" *The Journal of the Acoustical Society of America*, vol. 145, no. 5, pp. 3208–3217, 2019.

[9] ——, "Simplification of reflection orders in virtual soundscapes through a subjective evaluation," in *23rd International Congress on Acoustics*. Deutsche Gesellschaft für Akustik (DEGA eV), 2019, pp. 6432–6437.

[10] Y. Smyrnova, Y. Meng, and J. Kang, "Subjective evaluation of urban soundscape auralisation based on combined ray-tracing and radiosity (crr) model," *Proc. of the Institute of Acoustics (IOA)(UK)*, vol. 30, no. 2, 2008.

[11] D. Murphy and S. Harriet, "Auralisation of finite difference time domain simulations of sonic crystal noise barriers in an urban environment," in *10th European Congress and Exposition on Noise Control Engineering, EuroNoise*, 2015.

[12] C. Calleri, L. Shtrepi, A. Armando, and A. Astolfi, "Evaluation of the influence of building façade design on the acoustic characteristics and auditory perception of urban spaces," *Building Acoustics*, vol. 25, no. 1, pp. 77–95, 2018.

[13] C. Schissler, A. Nicholls, and R. Mehra, "Efficient hrtf-based spatial audio for area and volumetric sources," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 4, pp. 1356–1366, 2016.

[14] R. Mehra, L. Antani, S. Kim, and D. Manocha, "Source and listener directivity for interactive wave-based sound propagation," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 4, pp. 495–503, 2014.

[15] C. Xu, T. Oberman, F. Aletta, H. Tong, and J. Kang, "Ecological validity of immersive virtual reality (ivr) techniques for the perception of urban sound environments," *Acoustics*, vol. 3, no. 1, pp. 11–24, 2021. [Online]. Available: https://www.mdpi.com/2624-599X/3/1/3

[16] L. Jiang and J. Kang, "Combined acoustical and visual performance of noise barriers in mitigating the environmental impact of motorways," *Science of the Total Environment*, vol. 543, pp. 52–60, 2016.

[17] ——, "Effect of traffic noise on perceived visual impact of motorway traffic," *Landscape and Urban Planning*, vol. 150, pp. 50–59, 2016.

[18] D. W. Grantham, B. W. Hornsby, and E. A. Erpenbeck, "Auditory spatial resolution in horizontal, vertical, and diagonal planes," *The Journal of the Acoustical Society of America*, vol. 114, no. 2, pp. 1009–1022, 2003.