Evaluating the User in a Sound Localisation Task in a Virtual Reality Application

Adrielle Nazar Moraes
Dr Ronan Flynn, Dr Niall Murray
Department of Electronics and Informatics, Athlone Institute of Technology
Athlone, Ireland
email: a.nmoraes@research.ait.ie, rflynn@ait.ie, nmurray@research.ait.ie

INTRODUCTION

Virtual reality technology is fast becoming the accepted platform for delivering high quality immersive audio-visual experiences [1]. It has the potential to enable truly personalised experiences due to its adaptability and flexibility. Although a lot of the discussion on VR has focused on the visual aspect of the experience, but audio also plays a decisive role in terms of user immersion [2]. With the advancement of audio technologies, it is possible to render and spatialise audio with a high level of quality, matching human resolution to discriminate sounds in terms of depth, elevation and lateralisation [3]. Spatialised audio is very important in industries like gaming and for communication tasks, and in environments that require humans to constantly give attention to a specific sound in space [4][5][6]. In this paper, an immersive VR spatial audio application is presented. It enables us to evaluate the ability of users to specify or localise the source of a sound. An integrated sensing system continuously collects relevant data from the user in order to fully understand how to quantify and evaluate spatial auditory skills from a quality of experience (QoE) perspective. QoE gives insight into a user's state and behaviour. To perform a detailed QoE evaluation of the listening task, implicit and explicit metrics were collected from the user. Data collected from this QoE evaluation gives an insight into a user's abilities to localise sound sources in VR, and also provides information on behaviour and effort (workload) in performing the task.

EXPERIMENTAL METHODOLOGY

The virtual environment was designed using a Unity game engine (version 2018.2.15f1) with a Steam Audio set of generic HRTFs to render the spatialised audio. The immersive HMD used was the HTC VIVE with Tobii eye tracking integrated. One of the experiment aims was to evaluate user interaction with the system. This involved comparing two interaction methods: eye gaze pointing (GP) and pointing with a controller (WP). The immersive VR sound localisation task consisted of three testing phases (fig. 1), inspired by the LISN test:

a) Target only.
b) Target + distractor (same location).
c) Target + distractor (±90º apart from target).

The user was required to select the sound source a total of 72 times (24 trials for each testing phase). The experiment protocol was designed to guarantee that each sphere will reproduce a sound 3 times in total, one for each testing phase.

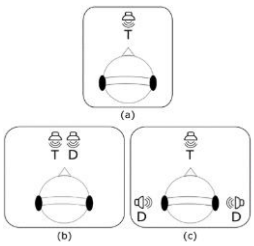

Figure 1 :Test phases

In order to evaluate localisation blur and front-back confusions, the order and location of the sound stimulus is randomized. The sound stimulus is a 2.5 s duration white noise presented at a constant sound pressure level of 55 dB SPL. The distractor sound stimulus consists of a low-frequency, 1,500 Hz pure-tone sine wave presented at the same SPL as the target stimulus. During the experiment, gaze data is sampled at a rate of 120Hz. In addition, the listener's performance data includes: the number of selection attempts; the number of correct selections; the informed location of the target; the real position of the target and the completion time for each sound source.

RESULTS

The results from Table I represent the average error angles for each group along the testing phases. This table shows that both of the groups had lateralisation errors (angle error < 45°). Additionally, these values indicate that participants missed the target by one (15°) or two spheres (30°). The results from the table also indicate that front-back confusions were not frequent across the tests, with the majority of localisation errors classified as localisation blur.

Table I: Average error angles for each group (eye gaze pointing (GP) and wand pointing (WP)) across testing phases

	Testing Phase 1	*Testing Phase 2*	*Testing Phase 3*
GP	30 ±22°	31 ±22°	35±22°
WP	37±14°	28±17°	33±18°

The time-to-complete values for the test were evaluated in order to give insight into the learning curve for the localisation task. An interesting result is related to the trend of the completion time mean values across the three conditions. Since the complexity is increasing across the phases, it was expected that the time to complete each phase would increase. However, this number has decreased, a possible indication that listeners learned how to perform the localisation task.

CONCLUSION

The use of 3D audio in a VR application produces a realistic and more immersive representation of the surrounding by enhancing the level of realism and awareness. Furthermore, the use of combined visual and auditory cues can improve the ability to localize a sound source.

From the QoE perspective, this application can provide more detailed information regarding the sound localization task, and its contributions can be applied within the health context, adding relevant data to the existent assessment and intervention methodologies for auditory disorders.

ACKNOWLEDGENTS

This Research is supported by Science Foundation Ireland and the ADAPT Centre. It is co-funded by the European Regional Development Fund under Grant Number 12/RC/2106.

REFERENCES

[1] C. Keighrey, R. Flynn, S. Murray, S. Brennan, and N. Murray, "Comparing User QoE via Physiological and Interaction Measurements of Immersive AR and VR Speech and Language Therapy Applications," in Proceedings of the on Thematic Workshops of ACM Multimedia 2017, 2017, pp. 485–492.
[2] T. Walton, "The Quality Of Experience Of Next Generation Audio: Exploring System, Context And Human Influence Factors," Newcastle University, 2018.
[3] J. Broderick, J. Duggan, and S. Redfern, "The Importance of Spatial Audio in Modern Games and Virtual Environments," in 2018 IEEE Games, Entertainment, Media Conference (GEM), 2018, pp. 1–9.
[4] G. Llorach, G. Grimm, M. M. E. Hendrikse, and V. Hohmann, "Towards Realistic Immersive Audiovisual Simulations for Hearing Research: Capture, Virtual Scenes and Reproduction," in Proceedings of the 2018 Workshop on Audio-Visual Scene Understanding for Immersive Multimedia, 2018, pp. 33–40.
[5] O. Rummukainen, S. Schlecht, and A. Plinge, "Evaluating Binaural Reproduction Systems from Behavioral Patterns in a Virtual Reality—A Case Study with Impaired Binaural Cues and Tracking Latency," in Audio Engineering Society Convention 143, 2017.
[6] R. Kilgore, M. Chignell, and P. Smith, "Spatialized audioconferencing: What are the benefits?," Cent. Adv. Stud. Collab. Res., pp. 135–144, 2003.

THE TIMES THE SUNDAY TIMES GOOD UNIVERSITY GUIDE 2020 INSTITUTE OF TECHNOLOGY OF THE YEAR

AIT Research