2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 978-1-6654-4057-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/VRW52623.2021.00071

Revisiting Audiovisual Rotation Gains for Redirected Walking

Andreas Junker, Carl Hutters, Daniel Reipur, Lasse Embøl*
Aalborg University

Niels Christian Nilsson, Stefania Serafin †
Aalborg University

Evan Suma Rosenberg‡
University of Minnesota

ABSTRACT

In this paper, we present a psychophysical study exploring how spatialized sound affects perceptual detection thresholds for rotation gains during exposure to virtual environments with varying degrees of visibility. The study was based on a 2×3 factorial design, crossing two types of audio (no audio and spatialized audio) and three degrees of visibility (low, medium, and high density fog). We found no notable effects of sound spatialization or visibility on detection thresholds. Although future studies are required to empirically confirm that vision dominates audition, these results provide quantitative evidence that visual rotation gains may be robust to auditory interference. Furthermore, they suggest that rotation gains may be useful even when the virtual environment offers very limited visibility.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality;

1 INTRODUCTION

Redirected walking (RDW) makes it possible to present large virtual environments (VEs) in a comparatively small tracking space. This can be accomplished by up- or downscaling walking users' physical movement by applying translation, rotation, curvature, and bending *gains*. Regardless of the type of gain, manipulations should remain unnoticeable to the user. For this reason, several studies have explored detection thresholds for different types of gains (for a recent review see [1]). While most of this work focused on RDW using visual gains, it has been demonstrated that blindfolded users can be redirected using auditory feedback [4]. Nevertheless, the combination of visual and auditory feedback for RDW remains largely unexplored. Moreover, a recent study [2] explored whether audio that is spatially aligned with a corresponding visual source reduces detection thresholds for rotation gains. No notable differences were found between conditions involving no audio, static audio, and spatialized audio. The authors speculate that the absence of an effect may be attributed to visual dominance. This explanation is plausible considering that a visually complex VE and a relatively sparse soundscape were used.

The current study sought to estimate detection thresholds for audiovisual rotation gains in a VE with a rich soundscape and varying degrees of visibility. This paper presents the following contributions: (1) quantitative evidence indicating that spatialized sound does not affect users' ability to detect visual rotation gains; and (2) results indicating that rotation gains potentially can be deployed even when the VE offers very limited visibility.

2 METHODS AND MATERIALS

To explore how detection thresholds for rotation gains are affected by spatialized audio under conditions with varying degrees of visibility, we performed a within-subjects study based on a 2×3 factorial design, crossing two types of audio (no audio and spatialized audio) and three degrees of visibility (low, medium, and high density fog). Specifically, the two types of audio were *no audio* (A_0), where no audio and only visual information was presented; and *spatialized audio* (A_1), which involved a spatialized soundscape. The three degrees of visibility were *low density fog* (F_0), where everything was visible except from distant objects, such as mountains; *medium density fog* (F_1), where only nearby objects were visible; and *high density fog* (F_2), where only the vegetation closest to participants and the ground directly beneath them were visible. Figure 1 shows the three degrees of visibility from the participants' perspective.

The audiovisual stimuli were presented using an *Oculus Rift CV1*. The VE consisted of ruins on a mountaintop, with a river running from a smaller waterfall to a larger waterfall through the ruins, crossed by a stone bridge. The soundscape was created using Unity's native sound spatialization, and it included ambient sound, a wood chopper, a river, two waterfalls, and a flock of birds.

To determine whether the participants' thresholds for detecting rotation gains varied across the six conditions, we performed a user study. Twenty participants (12 males, 8 females) were recruited from the student body at Aalborg University Copenhagen, with ages ranging from 21 to 34 years (M=24.7, SD=3.2). All participants gave written informed consent and reported having normal or corrected-to-normal hearing and vision. The method of constant stimuli in a two-alternative forced-choice (2AFC) task was adopted from Steinicke et al. [5]. Specifically, participants were exposed to a series of rotation gains and had to report whether their virtual rotation was faster or slower than their physical rotation. In the current study, participants had to perform a series of rotations on the spot. The first trial began when participants had put on the virtual reality headset and stated that they were ready. As seen in Figure 1, at the beginning of each trial, participants faced a cyan sign with an arrow indicating the direction they should rotate—the direction of the arrow was randomized for each trial. Participants should rotate until facing a red stop sign. Upon facing the stop sign, participants stated (verbally) whether they perceived their virtual rotation to be *faster* or *slower* than their physical rotation. Each virtual rotation was 90°. The next trial was

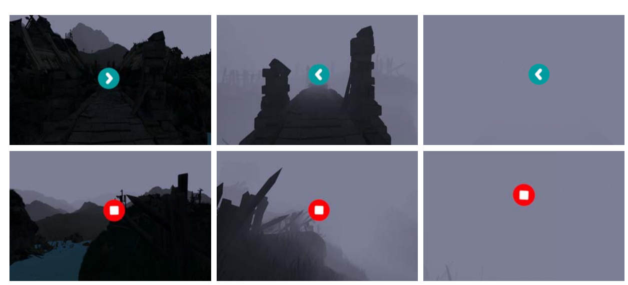

Figure 1: The VE during exposure to low density fog (F_0, left), medium density fog (F_1, middle), and high density fog (F_2, right). The top and bottom rows show the appearance of the signs shown in the beginning and end of each trial, respectively.

*e-mail:{ajunke16, chutte16, dreipu16, lsaren16}@student.aau.dk
†e-mail:{ncn,sts}@create.aau.dk
‡e-mail:suma@umn.edu

978-1-6654-4057-8/21/$31.00 ©2021 IEEE
DOI 10.1109/VRW52623.2021.00071

initiated once the participant's response was registered. The visual appearance of the signs was not affected by the fog. The rotation gains were varied between trials; the same range of gains was used as in previous work on detection thresholds for rotation gains [2, 5]. Specifically, gains ranged from 0.5 to 1.5 in increments of 0.1. Each gain was repeated twice, resulting in a total of 132 trials (22 for each condition) presented in a randomized order. After completing all trials, participants filled out a questionnaire including items asking about the usefulness of the audio, the effect of the fog, and their experience of the rotation gains.

3 Results

Data were analyzed using the same approach as previous work on detection thresholds for rotation gains [2, 5]. That is, for each condition, we determined a point of subjective equality (PSE) at which participants could not determine whether their virtual rotation was faster or slower than their physical rotation. Moreover, the detection thresholds for rotation gains smaller than the PSE were defined as the points, where participants responded "faster" in 25% of the trials (DT_{low}). Similarly, the detection thresholds for rotation gains higher than the PSE were defined as points, where participants responded "faster" in 75% of the trials (DT_{high}). The detection thresholds for each of the six conditions and the PSE are summarized in Table 1.

Regarding questionnaire responses, four participants found the addition of audio useful. However, two stated they found it useful because it contributed to the feeling of presence in the VE. Eight participants did not find the addition of audio useful as they did not notice it. Six participants expressed it was not useful as they were too focused on completing the trials and only paid attention to the visuals. Two participants responded the audio was distracting them from completing the tasks. Regarding fog, two participants found it confusing, seven participants did not really notice it, five participants felt it added nothing to the experience, and five participants found it more difficult to evaluate their rotation when the density of fog was high. Two participants stated the addition of fog made the rotations less sickening.

Finally, all participants preferred when the virtual rotation felt faster than their physical rotation. When asked to explain this preference, the primary reasons were that faster rotations elicited less cybersickness compared to slower rotations, and the faster rotations did not restrict virtual movements, introduce additional physical load, nor reduce completion times.

4 Discussion

We hypothesized that the addition of rich, spatialized soundscapes would decrease users' ability to detect rotation gains in VEs, especially during exposure to scenarios where visuals provide limited spatial information. If true, the results should have revealed lower DT_{low} and higher DT_{high} for the conditions including sound, particularly concerning the condition involving dense fog. However, as apparent from Table 1, this pattern is not present in the current results. Instead, we found similar detection thresholds across all six conditions. Thus, concerning the effect of spatial audio, our findings are consistent with the ones reported by Nilsson et al. [2], who attributed the absence of an effect to visual dominance. It is noteworthy that the results did not reveal a pronounced effect of varying visibility on detection thresholds. However, the condition with high-density fog did involve some lamellar optic flow, as the participants were able to see the closest vegetation and the ground directly beneath them. This arguably lends more credence to the claim that visual information heavily dominates proprioceptive and vestibular cues during this type of discrimination task. Notably, some participants remarked they almost exclusively focused on the start and stop signs when performing the task, and for that reason, barely noticing the audio or the fog. Previous work relying on the same methods, including similar signs, found that it was impossible to discern if virtual rotations were faster or slower than physical rotations when participants were placed on a flat grey plane devoid of any objects [3]. Thus, these signs did not provide enough information to make accurate judgments. Nevertheless, future work should explore whether participants rely on visual signs when discriminating between real and virtual rotations, e.g., by replacing signs with audio cues.

In conclusion, the results do not provide unequivocal evidence that spatialized audio cannot affect users' ability to detect RDW based on gains. However, the results corroborate past work indicating that vision is likely to dominate audition during exposure to rotation gains [2]. Moreover, even though some virtual self-motion cues appear to be necessary [3], our findings indicate that rotation gains can be deployed imperceptibly even when visuals are impoverished, or the VE offers very limited visibility. Importantly, this does not imply that spatialized sound, or soundscapes more broadly, should not be considered important for users' experiences (i.e., spatialized sound may be central to eliciting presence), and purely auditory redirection remains relevant to scenarios where walking users are momentarily or permanently deprived of visual feedback [4]. Finally, it is also worth noting that a higher fidelity sound engine, for example using personalized head-related transfer functions (HRTFs), might affect the results.

Table 1: The point of subjective equality (*PSE*) and lower and upper detection threshold (DT_{low} and DT_{high}), for each of the the six conditions.

Condition	$\mathbf{DT}_{low}$	PSE	$\mathbf{DT}_{high}$
A_0F_0	0.77	0.95	1.13
A_0F_1	0.78	0.93	1.09
A_0F_2	0.81	0.96	1.11
A_1F_0	0.78	0.95	1.12
A_1F_1	0.75	0.91	1.09
A_1F_2	0.80	0.94	1.09

References

[1] N. C. Nilsson, T. Peck, G. Bruder, E. Hodgson, S. Serafin, M. Whitton, F. Steinicke, and E. S. Rosenberg. 15 years of research on redirected walking in immersive virtual environments. *IEEE Computer Graphics and Applications*, 38(2):44–56, Mar. 2018. doi: 10.1109/mcg.2018.111125628

[2] N. C. Nilsson, E. Suma, R. Nordahl, M. Bolas, and S. Serafin. Estimation of detection thresholds for audiovisual rotation gains. In *2016 IEEE Virtual Reality (VR)*, pp. 241–242. IEEE, New York, NY, Mar. 2016. doi: 10.1109/vr.2016.7504743

[3] A. Paludan, J. Elbaek, M. Mortensen, M. Zobbe, N. C. Nilsson, R. Nordahl, L. Reng, and S. Serafin. Disguising rotational gain for redirected walking in virtual reality: Effect of visual density. In *2016 IEEE Virtual Reality (VR)*, pp. 259–260. IEEE, New York, NY, Mar. 2016. doi: 10.1109/vr.2016.7504752

[4] S. Serafin, N. C. Nilsson, E. Sikstrom, A. De Goetzen, and R. Nordahl. Estimation of detection thresholds for acoustic based redirected walking techniques. In *2013 IEEE Virtual Reality (VR)*, pp. 161–162. IEEE, IEEE, New York, NY, 2013.

[5] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics*, 16(1):17–27, Jan. 2010. doi: 10.1109/tvcg.2009.62