The Interaction of Visual and Auditory Cues to Linear Self-Motion Perception

Bill Kapralos\textsuperscript{a}, Andrew Hogue\textsuperscript{a}, Agnieszka Kopinska\textsuperscript{b}, and Saad Khattak\textsuperscript{a}

\textsuperscript{a}Faculty of Business and Information Technology, Health Education Research Technology Unit, University of Ontario Institute of Technology, Oshawa, Ontario, Canada. L1H 7K4

\textsuperscript{b}Department of Psychology, York University, Toronto, Ontario, Canada. M3J 1P3

ABSTRACT

We are capable of estimating the magnitude of our own self-motion and the relative motion of other objects as we move in our natural environment. This perception is based on information arising from several sensory modalities including visual, auditory, and physical motion. The perceived distance of self-motion is over-estimated when using visual or physical motion cues solely or even in conjunction, although judgments are more accurate when both cues are available. Despite the potential contribution of dynamic auditory localization to the perception of self-motion, few studies have examined the effects on auditory cues on the perception of self-motion. This paper describes a study that builds upon previous work that examined the interaction of auditory and visual cues in the perception of self-motion. Preliminary experimental results indicate that self-motion perception is more accurate when both auditory and visual cues are present.

1 INTRODUCTION

We are capable of estimating the magnitude of our own self-motion and the relative motion of other objects as we move about in our natural environment. This perception is based on information arising from several sensory modalities including visual, auditory, and physical motion. In general, the perceived distance of self-motion is over-estimated when using visual or physical motion cues solely or even in conjunction, although judgments are more accurate when both cues are available [1,2,3,4]. In other words, the distance we perceive ourselves to have moved is greater than the distance we have actually moved.

Many studies have examined the perceptual effects that visual and vestibular information may have in determining the perceived distance of linear self-motion (see [5] for a review). Harris et al. [1] examined the roles of visual and vestibular cues on the perceived distance of passive linear self-motion. When optic flow was the only cue available, participants overestimated their self-motion towards a previously shown target; they indicated that they reached a previously shown target when in fact they only moved about half the way to the target. During physical motion in the dark, participants once again perceived that they had traveled farther than they actually did. When presented with both physical motion and optic

\textsuperscript{a} Email address: bill.kapralos@uoit.ca
flow, participants’ accuracy improved and perceived self-motion was much closer to the actual distance traveled.

Despite the importance of dynamic auditory localization, few studies are available regarding the effects that audio cues may have on the perception of self-motion (but see [6,7,8,9,10,11,12,13]). The available studies have shown that auditory stimuli alone can evoke a sense of self-motion in a stationary observer both for linear motion [10,11,12,13] and angular motion [7]. While there is substantial literature regarding the perception of sound source intensity, it is not known if humans can use this information to gauge their self-motion. Kapralos et al. [3,14] and Zikovitz and Kapralos [15] assessed the relative roles of physical motion cued by sound on the perceived relative motion of an observer and a sound source. Similar to self-motion cued by vision alone [1], motion cued by audio alone resulted in perceptual gains (perceived motion/actual motion) as high as five at low accelerations but was closer to unity for the higher accelerations. For the largest acceleration examined, auditory cues presented with simultaneous physical motion cues were associated with more accurate estimates of self-motion than when either auditory or physical motion cues were presented alone [3].

In this paper we describe the preliminary results of an experiment that builds upon the study by Kapralos et al. [3] by examining the interaction of auditory cues with visual cues in the perception of self-motion. More specifically, we examined the ability to judge the distance of self-motion when simulating the movement away from a stationary sound source following one of five pre-defined motion profiles in the presence of optic flow (visual) cues and decreasing sound intensity cues. The results were compared to the results of previous experiments that examined self-motion perception in the presence of either visual (optic flow) cues alone [1], or ii) auditory cues (decreasing sound source intensity) alone [3].

Just as the addition of auditory cues increase the accuracy of self-motion perception when combined with physical motion cues as opposed to physical motion cues alone, the addition of auditory cues increases the accuracy of self-motion perception over visual cues alone. The results of the experiment have potential implications for the designers of immersive environments that aim to simulate self-motion.

1.1 Paper Organization

The remainder of the paper is organized as follows. Greater details regarding the experimental procedure are provided in Section 2. Experimental results in addition to a discussion regarding the results are presented in Section 3. Finally, concluding remarks and plans for future work are provided in Section 4.

2 PROCEDURE

2.1 Participants

All participants were unpaid volunteers and were researchers from the University of Ontario Institute of Technology. Participants reported no history of auditory disease or disorders. None of the participants reported any difficulties in hearing the stimuli or completing any of the tasks. Three participants completed the experiment. The experiment abided by the University of Ontario Institute of Technology Research Ethics Board policies for the ethics review process for experiments involving human participants.

2.2 Auditory Stimulus

The auditory stimulus consisted of a broadband white-noise signal sampled at a rate of 44.1 kHz. The noise was band-pass filtered using a 256-point Hamming windowed FIR filter with low and high frequency cut-offs of 200 Hz and 10 kHz respectively. A broadband
sound was used since sound source distance estimates are more accurate for broadband sounds [16,17,18]. The auditory stimuli used in this work were identical to the auditory stimuli considered by Kapralos et al. [3,14] to ensure meaningful comparisons of the results could be made. All auditory stimuli were output with a pair of AKG Acoustics A240 headphones.

2.3 Visual Stimulus

The visual stimulus in the form of optic flow consisted of movement (in a “first-person” perspective) down a textured corridor (see Figure 1(a) below for an example of the corridor). Visual stimuli were displayed in a mono (i.e., stereo vision was not considered) on a 15” laptop monitor. The visual stimuli used here were similar to the visual stimuli used in the experiment conducted by Harris et al. [1]. However, rather than using a textured wall pattern as done here, they used a striped wall (see Figure 1(b)) to imitate a series of previously conducted experiments that consisted of physically moving participants through a hallway at in a building that had similar striped walls.

![Figure 1. Visual stimuli of (a) the experiments described here, and (b) the experiments conducted by Harris et al.[1].](image)

2.4 Experimental Procedure

Participants were informed about the required task prior to the start of the experiment by one of the experimenters and were given several test trials before the start of the experiment to ensure that they understood their task. Participants were seated and remained stationary in front of a laptop computer monitor for the duration of the experiment (a total of 100 trials). At the start of each trial, participants were presented with a visual target at one of four target distances (1, 2, 3, or 4 m away from the starting position) along the virtual hallway (see Figure 1(a)). The target remained visible for five seconds after which, participants were presented with the visual and auditory stimuli simulating movement of the participant from the starting position along the hallway at one of five rates of acceleration: 0.0125, 0.0250, 0.050, 0.100, and 0.200 m$^2$s$^{-2}$.

The task of the participant in each trial was to judge when they perceived that they reached the location of the previously presented target. They indicated this by pressing a key on a gamepad controller. Pressing of the button indicates the end of the trial. The next trial began after participants pressed a key on the gamepad controller. Essentially, five rates of acceleration were tested over four target distances and each acceleration/distance combination was tested five times for a total of 100 trials. The ordering in which trials were
presented to the participants was randomly chosen. Approximate experiment duration was 20 minutes and carried out in a single session.

3 RESULTS AND DISCUSSION

3.1 Data Analysis
The perceived distance (i.e., the distance to the target) is the distance that the subject perceived they had moved (i.e., 1, 2, 3, or 4 m). The “actual distance” is the simulated distance moved and is measured from the starting position to the position corresponding to the point at which the subject pressed the response button. For each subject, the perceived distance moved (i.e., distance to the target) was plotted as a function of the actual distance moved. The ratio of perceived distance / actual distance is the perceptual gain [1] (see Figure 2 for an example). A perceptual gain of one is obtained if the perceived distance equals the actual distance. If a subject traveled farther than the target distance to achieve the perception of having moved beyond the target, the resulting perceptual gain will be less than one (actual distance > perceived distance). If a subject traveled less than the target distance this resulted in a perceptual gain greater than one (actual distance < perceived distance).

Figure 2. Perceptual gain example.

3.2 Experimental Results
Figure 3 plots average actual (simulated) distance moved versus perceived distance (i.e., distance to the target) moved for each experimental condition (i.e., perceptual gain). The slope of each resulting line is equal to 1/perceptual gain and is less than one in all cases indicating that perceptual gain was consistently greater than one. Included in this plot are the results of the experiment conducted by Kapralos et al. [3] which investigated self-motion perception using auditory cues alone (participants remained stationary and presented with auditory stimulus whose intensity was decreased to simulate the intensity that would occur if the participant was moved away from a stationary sound source at one of the five rates of acceleration considered here). Also included are the results of the experiment conducted by Harris et al. [1] that investigated self-motion perception in the presence of visual cues alone (participants remained stationary in front of a computer monitor looking down a virtual hallway and presented with optic flow cues to simulate motion along the hallway).
As shown in Figure 3, the combination of auditory and visual cues (“audio + vision”) provided the more accurate self-motion perception while auditory cues alone (“audio only”) was the least accurate although, perceptual gain decreased with increasing acceleration (i.e., self-motion perception accuracy increased with increasing acceleration). For the “audio + vision”, condition examined here, perceptual gain were essentially the same for the first two accelerations examined (perceptual gain of 1.21 and 1.23 for the 0.012, and 0.025 ms\(^{-2}\) accelerations respectively) but then decreased with increasing acceleration although to a lesser degree. A repeated measures ANOVA was performed and revealed a significant difference between the “audio + vision” and the “vision only” (P=<0.001) and “audio only” conditions (P=0.026). This is consistent with previous results whereby self-motion perception accuracy becomes more accurate as the number of sensory cues is increased [1,3].

### 4 SUMMARY

Here we have presented the preliminary results of an experiment conducted to test the interaction of auditory and visual motion cues to self-motion perception. Preliminary results indicate that the addition of auditory cues increases the accuracy of self-motion perception considerably when combined with visual motion cues as opposed to visual cues alone. Future work will include testing of more participants in order to verify the preliminary results presented here and to allow for meaningful statistical analysis to be performed. Future work will also include adding an “acoustical target” in addition to the visual target presented to the participants at the start of each trial. The acoustical target will provide participants with a reference to the intensity of the sound at the target position thus potentially allowing them to make better (more accurate) use of the auditory cues.

### 5 ACKNOWLEDGEMENTS

The financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged.

### 6 REFERENCES


