Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnline CONFERENCE PAPER ARCHIVE
IFAC-PapersOnLine 49-19 (2016) 531–536

Sonifying the location of an object: A comparison of three methods

Pavlo Bazilinskyy, Wessel van Haarlem, Hashim Quraishi, Coen Berssenbrugge, Jasper Binda, Joost de Winter

Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, the Netherlands (Tel: +31152787891; e-mail: p.bazilinskyy@tudelft.nl)

Abstract: Auditory displays are promising for informing operators about hazards or objects in the environment. However, it remains to be investigated how to map distance information to a sound dimension. In this research, three sonification approaches were tested: Beep Repetition Rate (BRR) in which beep time and inter-beep time were a linear function of distance, Sound Intensity (SI) in which the digital sound volume was a linear function of distance, and Sound Fundamental Frequency (SFF) in which the sound frequency was a linear function of distance. Participants ($N = 29$) were presented with a sound by means of headphones and subsequently clicked on the screen to estimate the distance to the object with respect to the bottom of the screen (Experiment 1), or the distance and azimuth angle to the object (Experiment 2). The azimuth angle in Experiment 2 was sonified by the volume difference between the left and right ears. In an additional Experiment 3, reaction times to directional audio-visual feedback were compared with directional visual feedback. Participants performed three sessions (BRR, SI, SFF) in Experiments 1 and 2 and two sessions (visual, audio-visual) in Experiment 3, 10 trials per session. After each trial, participants received knowledge-of-results feedback. The results showed that the three proposed methods yielded an overall similar mean absolute distance error, but in Experiment 2 the error for BRR was significantly smaller than for SI. The mean absolute distance errors were significantly greater in Experiment 2 than in Experiment 1. In Experiment 3, there was no statistically significant difference in reaction time between the visual and audio-visual conditions. The results are interpreted in light of the Weber-Fechner law, and suggest that humans have the ability to accurately interpret artificial sounds on an artificial distance scale.

Keywords: road safety, driver support, auditory display, human-machine interface, driving simulator, detecting elements

1. INTRODUCTION

Auditory displays can be of value in a broad spectrum of applications, especially in situations where visual feedback is restricted, when the visual system is overburdened, or when the message is short and calls for immediate action (Stanton and Edworthy, 1999). Adding auditory feedback to a human-machine interface may shorten visual search times and reduce the workload compared to using vision only (Perrot et al., 1990; Wickens, 1984).

Usually, auditory feedback takes the form of short warning signals (Patterson, 1982; Stanton and Edworthy, 1999). For example, auditory warnings are used in blind spot monitoring and forward collision warning systems in modern cars (Bazilinskyy et al., 2015; Jamson et al., 2008).

Auditory feedback can also be used to perceptualize objects or activity in the environment, a method which is called sonification (Hermann et al., 2011). One of the earliest known applications of sonification is an optophone. The device, used by the blind, was developed in 1913; it scans text and generates time-varying chords of tones to identify letters (Capp & Picton, 2000). One of the most successful examples of sonification is the Geiger counter, in which auditory clicks are produced to represent ionization events. The Geiger counter was developed in the early 1900s, and is still used today to measure the level of radiation in the environment (Knoll, 2010). An auditory pulse-oximeter, a device similar to the Geiger-counter, was used in hospitals in the United States in 1980's. It generated a tone that varied in pitch based on the level of oxygen in patient's blood (Kramer et al., 1999). Spain et al. (2007) investigated the implications of the use of sonified feedback during a patient monitoring task. They found that a short inter-pulse time contributes to a higher level of perceived urgency.

Sonification is also useful in the field of data analysis, in which case it is sometimes called audification or auditory graphing (Flowers, 2005). During the Voyager 2 space mission, the control encountered a problem when the spacecraft was going through the rings of Saturn. The unexpected behaviour could not be explained by means of a visual analysis of the data. When the data was played through a music synthesizer, a 'machine gun' sound was heard, leading to the conclusion that the problem was caused by collisions with electromagnetically charged micrometeoroids (Barrass and Kramer, 1999; Kramer et al., 1999).

Sensory substitution of visual information may be of value in supporting persons in locomotion tasks (e.g., Hussain et al., 2014). As early as 1936, De Florez suggested that pilots of

2405-8963
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.614

aircrafts can benefit from the support of sonified instruments in so-called "blind flying" (De Florez, 1936). Parseihian et al. (2012) studied the mapping of the sonified distance to the actual object's location, and developed a sonified device for visually impaired persons. In the automotive industry, the parking sensor of a modern car is another example of the use of sonification, where an increasingly frequent beep is emitted to indicate that the car approaches an object. Although a parking sensor is a successful demonstration of sonification, it remains to be investigated which sonification method is the most effective for conveying information about distance or the degree of hazard.

Haas and Edworthy (1996) showed that sounds producing the highest level of perceived urgency are sounds of a high beep rate, a high intensity, and high frequency. This suggests that each of these three dimensions may be intuitive for sonification purposes. A review article of 179 publications related to sonification of physical quantities concurs that pitch (frequency), loudness (e.g., volume, intensity), and duration (e.g., beep time, inter-beep time) are the most often used auditory dimensions for sonification (Dubus and Bresin, 2013). Sanders and McCormick (1987; as cited in Stanton and Edworthy, 1999) on the other hand suggested that the auditory discrimination power of humans is rather limited, and contended that humans can identify only 2 to 3 levels of sound duration, 4 to 5 levels of sound intensity (at a given frequency), and 4 to 7 levels of sound frequency. Zahorik (2002) and Loomis et al. (1998) found that participants consistently underestimated the distance in auditory distance perception tasks. Thus, more fundamental research into the topic of mapping of given auditory cues to the distance needs to be conducted.

As mentioned above, beep time, intensity, and frequency are primary sonification dimensions. The aim of this study was to investigate which of these three sonification dimensions allows a person to most accurately indicate the location of an object. Participants completed two experiments; the first experiment involved one-dimensional distance estimation, whereas the second experiment involved the localization of an object in a two-dimensional plane. The participants were presented with sounds without visual feedback, and subsequently had to click on the screen to locate the object. In an additional Experiment 3 we sought to determine whether directional auditory feedback improves reaction times compared to visual-only feedback.

2. METHOD

Apparatus. The research was conducted using a computer program created with the Unity game engine (version 4.6.1f1). Razer Electra headphones were used.

Auditory feedback. Three types of auditory feedback were tested. The first type was Beep Repetition Rate (BRR), in which the beep time was linearly related to distance with respect to the bottom of the screen. For the closest distance (bottom of the screen), the beep time and inter-beep time were 0.05 s (i.e., 10 beeps per second). For the farthest distance (top of the screen), the beep time and inter-beep time were 0.55 s (i.e., 0.91 beeps per second). BRR resembled the feedback in a parking sensor, in that it 'beeps' faster as you are closer to an object. In the BRR condition, the sound volume was 100%, and the frequency of the beeps was 460 Hz. The volume of the laptop computer was set so that 100% sound volume generated by the software was regarded as loud but not uncomfortable.

Second, we tested Sound Intensity (SI), where the volume intensity was linearly related to the distance to the object. The volume was 0% at the top of the screen and 100% at the bottom of the screen. The frequency of the sound was 460 Hz.

Third, we tested the Sound Fundamental Frequency (SFF), where the frequency of the sound was linearly related to the distance. The frequency was 1,076 Hz at the bottom of the screen and 184 Hz at the top of the screen. The volume of the sound was 100%.

Participants. Twenty-nine persons (8 females) participated in the experiment. Most participants were students and employees of Delft University of Technology, and were on average 29.6 years old (SD = 15.7 years). None of the participants had a hearing disorder or used hearing aids.

Procedure. The participants conducted three experiments in the following order: Experiment 1: Distance estimation, Experiment 2: Distance and angle estimation, and Experiment 3: Reaction time. In Experiment 1 and Experiment 2, the participants completed three sessions, each session with a different sound condition (BRR, SI, SFF). To neutralize the effects of a learning curve, we randomized the order of the three sound conditions. We did, however, have the same order for the sessions in Experiments 1 and 2, to prevent participants from experiencing the same sound method right after each other. In Experiment 3, the participants completed two sessions: No Sound and Sound, in randomized order. Each session consisted of 10 trials. Accordingly, each of the participants completed 80 trials in total (30 in Experiment 1, 30 in Experiment 2, and 20 in Experiment 3). The three experiments are explained below.

Experiment 1: Distance estimation. In the first experiment the participant heard a sound, equally loud in both ears of the headphones. The duration of the sound was 1.0 s for SI and SFF, and 3 beeps for BRR. The participant had to locate the object as accurately as possible by clicking on the screen. Immediately afterwards, the participants were shown the chosen location (cyan square) and the actual location of the object (red square), as well as an absolute distance error score expressed as a percentage shown in the left top of the screen. The experiment was preceded by a short automated demonstration in which the participants were presented with 11 sounds from low to high intensity (0%, 10%, …, 100%), together with a corresponding red square on the screen from top to bottom.

Experiment 2: **Distance and angle estimation**. The second experiment was the same as the first, but this time the participant had to locate the object in a two-dimensional plane (Fig. 2). Not only the distance but also the azimuth angle had to be estimated. To represent the angle, we used the volume per ear linearly mapped from the azimuth angle. If the sound volume was 100% in both ears, the object was in