Applied Acoustics 139 (2018) 82–92

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Localization of a virtual wall by means of active echolocation by untrained sighted persons

David Pelegrín-García[a,b,*], Enzo De Sena[b,d], Toon van Waterschoot[b], Monika Rychtáriková[a,c], Christ Glorieux[a]

[a] Laboratory of Acoustics, Division Soft Matter & Biophysics, Dept. Physics & Astronomy, KU Leuven, Belgium
[b] STADIUS-ESAT, Department of Electrical Engineering, KU Leuven, Belgium
[c] STU Bratislava, Faculty of Civil Engineering, Dept. of Building Structures, Radlinskeho 11, Bratislava 810 05, Slovak Republic
[d] Institute of Sound Recording, University of Surrey, Guildford GU2 7XH, UK

ARTICLE INFO

Keywords:
Human echolocation
Virtual acoustics
Masking
Sound localization

ABSTRACT

The active sensing and perception of the environment by auditory means is typically known as echolocation and it can be acquired by humans, who can profit from it in the absence of vision. We investigated the ability of twenty-one untrained sighted participants to use echolocation with self-generated oral clicks for aligning themselves within the horizontal plane towards a virtual wall, emulated with an acoustic virtual reality system, at distances between 1 and 32 m, in the absence of background noise and reverberation. Participants were able to detect the virtual wall on 61% of the trials, although with large differences across individuals and distances. The use of louder and shorter clicks led to an increased performance, whereas the use of clicks with lower frequency content allowed for the use of interaural time differences to improve the accuracy of reflection localization at very long distances. The distance of 2 m was the most difficult to detect and localize, whereas the furthest distances of 16 and 32 m were the easiest ones. Thus, echolocation may be used effectively to identify large distant environmental landmarks such as buildings.

1. Introduction

The sense of hearing provides relevant information for spatial perception [1]. Audition is particularly important for blind people [2], who lack visual stimuli to build spatial representations of their surroundings. Some blind people have learnt to echolocate [3], i.e. to detect and localize obstacles and environmental features based on the reflections they produce in response to self-generated sounds (active echolocation), typically oral clicks [4], or even to ambient noise [5] (passive echolocation). Echolocation, initially called facial vision because it was believed that sensation arose from pressure sensors on the skin [6], is in fact a purely auditory phenomenon [7]. Sound reflections, or echoes (if perceived as a separate event from the direct sound), arrive at an echolocator with variable attenuation, delay, interaural level and time differences and spectral cues which they exploit [8,9] to infer information about the distance [10,11], angular location [12], size [13], shape [14] and texture [13] of the boundary at which they were generated. At short distances, it is possible to use coloration cues (i.e. a change in the tonal character), arising from the interaction of direct and reflected sounds [15] to detect the latter. However, it is highly unlikely to use coloration cues to detect reflections with delays longer than 10 ms (or arising from obstacles at distances further than 2 m) [16,17]. Localization of reflections is particularly precise due to a partial inhibition of the precedence effect—a collection of phenomena that makes it possible to localize an original sound source in a room despite all reflections [18]—during echolocation [19,20]. This technique represents an active perception mode [21], meaning that the perception of auditory space integrates the auditory sensation at different positions and orientations with the vestibular and proprioceptive feedback [22], and thus head movements are crucial [23] for effective mobility and related tasks like shape perception [14]. In this way, echolocation contributes to enhance the auditory spatial localization of blind people [24].

In addition, echolocation has benefits on the independence of functional echolocators (i.e. people who use echolocation in daily life), namely better mobility in unfamiliar places and access to better salaried jobs [25]. Thus, echolocation is highly relevant for the rehabilitation of persons who have lost sight. Despite its benefits, echolocation is not yet a widespread technique and much research is devoted to understanding the degree to which early and late blind people can profit from echoic

* Corresponding author currently at: Widex A/S, Nymøllevej 6, 3540 Lynge, Denmark.
E-mail address: david.pelegringarcia@kuleuven.be (D. Pelegrín-García).

https://doi.org/10.1016/j.apacoust.2018.04.018
Received 6 September 2017; Received in revised form 4 April 2018; Accepted 13 April 2018

information [26]. In the present study, we focus on sighted persons without prior experience in echolocation, and as such their performance can be considered similar to that of potential candidates in rehabilitation programs of orientation and mobility.

Acoustic Virtual Reality (AVR) systems that account for head orientation are regarded useful for the acquisition of auditory space maps [27], for evoking sensations arising in echolocation [10,22] and for the conduction of psychoacoustic tests [20]. When using an AVR system for the conduction of psychoacoustic or active echolocation tests, non-auditory cues such as wind, temperature, or other tactile cues are decoupled from the auditory stimulation. In addition, ambient noise, which could lead to effective passive echolocation in reality, can be minimized or studied separately from active echolocation. In real life, these cues are an integral part of the environment perception and contribute to the detection of obstacles. Furthermore, arbitrarily large surfaces can be simulated with the AVR system at any position. Because large surfaces reflect sound more efficiently than small obstacles, and because the background noise in an AVR is negligible when used in a silent anechoic room, the range of distances where obstacles can be detected with echolocation is expected to be much larger using an AVR system than in other real life experiments.

Sighted participants have been able to learn basic echolocation tasks using an AVR system [11,20]. Therefore, this kind of systems shows a potential to explore effective learning strategies in echolocation, gain further knowledge about its psychophysical mechanisms, and study the relationship between performance and the self-generated oral signals [11,17]. When using an AVR system, the simulated acoustic condition is defined by the Oral-Binaural Room Impulse Response (OBRIR) [28], which contains information on the propagation of sound, including all relevant reflections between a reference point close to the mouth and the entrance of the closed left and right ear canals.

In a previous study [17], we used a static AVR system to determine the detection thresholds for a single reflection up to distances of 16 m, by artificially varying the strength of the reflection in relation to the strength of the direct sound. This relation between strengths was quantified by means of the reflected-to-direct level difference (RDLD) [29].[1] The results closely followed the signs of forward masking—a characteristic of the human auditory system for which one sound can render inaudible another fainter sound closely delayed. When using oral clicks, thresholds improved with increasing distance due to the increased temporal separation between direct and reflected sound, which caused less forward masking. Louder signals resulted in lower (i.e. better) detection thresholds due to the non-linear behavior of masking [17]. These results suggest that it might be possible to use echolocation in a range of distances longer than reported in several previous studies of echolocation with sighted participants in real settings (e.g. 55–65 cm [30], 62–130 cm [31] or less than 1.8 m [7,32]) or in laboratory settings with sighted and blind participants (up to 2 m [9,33]), when using larger obstacles in an otherwise reflection-free environment with low background noise. This was already explored by Schrnich et al. [11] by measuring distance discrimination thresholds in echolocation up to 6.8 m using an AVR. In the present study, we explore echolocation for detection and localization of large obstacles in a range of distances broader than in previous studies, from 1 m to 32 m.

In [17], the task was to discriminate one out of three intervals which was different from the other two. A question that arises is whether untrained sighted individuals would also be able to use echolocation in a more realistic task than the one in [17]. In the present study, we develop and use an AVR system that accounts for head rotations in the horizontal plane. We use this system to simulate a simple outdoor echolocation exercise in which a person has to orient herself towards a simulated large wall (or building) at different distances on a reflecting floor, without competing reflections from other surfaces nor background noise. Thus, participants in the present experiments had firstly and implicitly to detect the reflecting object, and only then, to identify the direction of arrival of the reflection.

Beyond knowing whether untrained sighted individuals can echolocate, we aim at determining the most difficult distance conditions to detect and localize a large obstacle (e.g. a wall) and how performance relates to features of the emitted signals. This knowledge can serve as an inspiration for determining good clicking strategies based on parameters such as intensity, duration, bandwidth and frequency content, and in addition, to point out the most difficult distance conditions in the training of echolocation. We hypothesize that, in absence of background noise and reverberance, the most difficult conditions are those in which the reflection level of the wall is near the audibility thresholds for a reflection determined in [17]. Easier conditions may be those where the reflection is well above threshold. Moreover, in the distance conditions where the reflection level is near the audibility thresholds, it is expected that localization accuracy decrease [34].

2. Method

2.1. Participants

The 21 participants in the experiments, labeled S1 to S21 hereafter, were 7 female and 14 male sighted persons, between 22 and 48 years old, with normal hearing (HL < 20 dB from 250 Hz to 8 kHz following audiometric screening) and without previous experience in echolocation. They participated on a voluntary basis and did not receive any compensation for the participation. Informed consent was given by all participants and ethics approval was granted for this research by the Medical Ethics Committee at UZ KU Leuven (number B322201317883).

2.2. Stimuli

Stimuli were designed to replicate a simple outdoor echolocation exercise in which participants had to identify the direction of a large wall (e.g. a building) and turn towards it. Virtual flat concrete walls (of dimensions 10 m × 10 m) at distances d of 1, 2, 4, 8, 16 and 32 m on a concrete floor (see Fig. 1) of 100 m × 100 m were simulated through streaming convolution of the OBRIRs characterizing these scenarios with the oral clicks produced by the participants themselves. The following sections give more insight on the computation of the OBRIRs and on the actual stimuli generation via streaming convolution.

2.2.1. OBRIR calculation

The reflections of the wall and the floor were simulated with the room acoustics simulation software CATT-Acoustic™v9.0c. A binaural receiver was placed at the middle point in between the ears, and a source simulating the mouth with the average frequency-dependent directivity pattern of the human voice [35] was placed 0.1 m in front of the receiver and pointed away from it. Such a pattern, on a frequency-band basis, is similar to the one produced by echolocation clicks [36]. The receiver was always pointing towards the source. Both source and receiver were placed at a height of 1.5 m from the floor. In separate calculations, the wall was placed at each of the six different distance conditions (d = 1, 2, 4, 8, 16 and 32 m) from the receiver. These distances were chosen to represent a large range of distances, as in a previous study [17], and were logarithmically spaced because the reflected level decays approximately 6 dB per doubling of distance. At each distance d, simulations were performed for 24 reference orientations θ_0 of source/receiver, always rotating the source position around the fixed receiver, at intervals $\Delta\theta$ of 15°.

The wall and the floor had an absorption coefficient of 0.01 at

[1] Note that the RDLD is unambiguously defined for a single reflection. In the case of multiple reflections, there could be an ambiguous interpretation. On the one hand, the RDLD could be referred to a single reflection of interest by filtering out all other competing reflections. On the other hand, the RDLD could take into account all existing reflections to calculate the strength of the total reflected sound. In this study, we use the latter interpretation.