Binaural Navigation for the Visually Impaired with a Smartphone

Lee Tae Hoon  
NUS High School  
h0910080@gmail.com

Manish Reddy Vuyyuru  
NUS High School  
wildtangles@gmail.com

T Ananda Kumar  
NUS High School  
h0910139@gmail.com

Simon Lui  
Singapore University of Technology and Design  
simon_lui@sutd.edu.sg

ABSTRACT

We aimed to determine the feasibility of binaural navigation and demonstrate piloting of blindfolded subjects using only recorded (virtual) binaural cues. First, the localization accuracy, sensitivity to angular deviations and susceptibility to front-back confusion of the subjects was measured using setups modified from previous works on localization capacity. Afterwards, a software prototype that seeks the relevant binaural cues from a database to keep a subject on a predetermined path by determining the subject’s head orientation and coordinates was developed on the Android platform. It was scored by the root-mean-squared deviation (RMSD) of blindfolded subjects from the route. Experimental results show that precise piloting of blindfolded subjects is achievable using only virtual binaural cues. We also discuss the growing potential for more sophisticated uses of binaural media.

1. MOTIVATION

The ability to deduce the location of a sound source from auditory cues (localization) is a behavioral instinct developed upon birth[1]. By recruiting the visual cortex in the context of localization, the visually impaired process auditory cues better than people without any visual impairment[2]. However, outdoor navigation via environmental cues is still remarkably difficult for them. Since audio navigation systems for the visually impaired rely primarily on verbal instructions instead of spatial audio localization[3][4][5], the investigation was initiated to explore and develop the feasibility of recorded (virtual) binaural cues for indoor and outdoor navigation.

2. BACKGROUND

Three main auditory cues are responsible for auditory localization - interaural intensity differences (IID), interaural time differences (ITD) and monaural spectral cues (MSC). IID and ITD account for sound localization laterally (left, ahead, right). The sound received at the contralateral ear is shadowed by the head and results in IID. The cue is more profound for higher frequencies where the wavelength of the sound is smaller than the distance between the two ears (>1600Hz). The arrival time delay from the physical separation of the ears on the head results in ITD. The cue is more profound for sounds of lower frequencies (<1600Hz). The sound stimulus was played from the reference angles 75°, 55°, 35° and -65° at a radius of 1.5m from the origin. The angles were purposefully picked from the frontal azimuth since subjects instinctively turn to face sounds coming from the back before attempting to pinpoint the source. Subjects, unaware of the reference angles, were tasked to point a laser at the source after exposure to each audio stimulus. The laser then lands on a sheet of paper arched into a semi-circle of radius 1.8m from the origin positioned at the height of subjects’ torsos to capture localization data. The subjects were

4. ABSOLUTE LOCALIZATION

4.1 Method (Real Sound Source)

The audio stimulus was played from the reference angles 75°, 55°, 35° and -65° at a radius of 1.5m from the origin. The angles were purposefully picked from the frontal azimuth since subjects instinctively turn to face sounds coming from the back before attempting to pinpoint the source. Subjects, unaware of the reference angles, were tasked to point a laser at the source after exposure to each audio stimulus. The laser then lands on a sheet of paper arched into a semi-circle of radius 1.8m from the origin positioned at the height of subjects’ torsos to capture localization data. The subjects were
blindfolded and fastened to a chair. Head movements, tracked by a laser, were damped with a neck brace. Wrists were immobilized with a device resting on a series of armrests arranged around the subjects at chest level so that the laser pointer was aligned with the direction the arms were pointing.

Every subject underwent a session consisting of 30 trials. In every trial, the audio stimulus, repeated 5 times with a time delay of 1.10s to allow time for processing, was played from the speaker from the 5 angles in a randomized order. Subjects then pointed a laser at the perceived sound source and the position the laser landed on the arched sheet was marked and the angular deviation from the actual sound source measured. At the end of the 30 trials, all 5 reference angles were tested 30 times each.

4.2 Method (Virtual Sound Source)

Instead of a real sound source (i.e. speakers), the stimulus is retrieved from a subject’s personal spatial audio database and delivered via in-ear stereo earphones to preserve the auditory cues. The stimulus, setup and task are otherwise identical to the previous experiment method as in Section 4.1.

4.3 Results and Discussion

No statistically meaningful difference in the localization of real versus virtual sound sources is observed. However, single sample t-test statistics show there is sufficient evidence at the 1% significance level to conclude that the angular deviation of the perceived source from the real sound source is significant. The most likely culprit, a lack of coordination when pointing at the perceived source, cannot be solely held accountable as the sample standard deviation of signed errors (SSD SE) increases as the absolute angular distance from the nasal axis increases. This systematic behavior of SSD SE hints at a psychoacoustic nature of subjects to perceive sounds away from the source. The localization accuracy is greatest at the nasal axis and decays with increasing unsigned angular distance from the axis. The conclusion is in agreement with previous works on human auditory localization for real sound sources[7].

Auditory localization in a navigation task is a continuous process. Sounds presented in a free field are localized by instinctively turning to face the perceived source which recursively shifts the source closer to the nasal axis for accurate localization. Thus, the localization errors (<15°) far from the nasal axis and errors (<5°) close to the nasal axis are within reasonable ranges for binaural navigation.

5. RELATIVE LOCALIZATION

Localization errors close to the nasal axis are investigated. A relative localization paradigm[8] was employed to determine sensitivity to small angular deviations of a sound source about the axis. Only virtual sound sources were used.
5.1 Method

$1\degree$ intervals were marked from $10\degree$ to $-10\degree$ at a radius of 1.5m from origin. Subjects’ spatial audio databases were updated with binaural recordings from each of the 21 points. S1 denotes the reference angle, $0\degree$ and S2 any of the 21 points. Subjects are played S1 followed by S2 through in-ear stereo earphones and tasked with determining the position of S2 relative to S1 (right, left, on). By doing so, the errors from a lack of coordination which are inherent to an absolute localization paradigm were eliminated. The responses were categorized into 3 types - Type 0 when subjects perceived change or absence of change correctly and stated the correct position of S2 relative to S1, Type 1 when subjects perceived absence of change when there is a change and Type 2 when subjects perceived change but stated the wrong position of S2 relative to S1. Each subject underwent 30 sessions, with each session involving all 21 recordings played as S2 in a randomised order.

5.2 Results and Discussion

Figure 7. Percentage of Response Types against Value of S2 (Experiment 2)

Angles beyond $1\degree$ and $-2\degree$ inclusive had over 75% Type 0 responses. The absence of Type 2 responses shows there’s no front-back confusion at the nasal region. Subjects were able to distinguish and state relative direction of S2 coming from angles as small as $1\degree$ to $2\degree$ relative to S1. Thus accurate navigation ($1\degree$-2$\degree$ resolution) is achievable around the nasal axis and by extension, through the continuous localization process as explained in Section 4.3, outside the nasal region.

6. FRONT-BACK CONFUSION

6.1 Method

MSC can be too subtle for listeners resulting in front-back confusion. Thus, its extent and its effect on the feasibility on binaural navigation was studied next. The reference angle, S1, is set to $-90\degree$ and S2 can take any of 61 values from $-60\degree$ to $-120\degree$ at 1$\degree$ intervals. The experimental method is otherwise identical to that detailed in Section 5.1.

6.2 Results and Discussion

Figure 9. Percentage of Response Types against Value of S2 (Experiment 3)

Subjects faced significant confusion determining the direction of S2 within the range $-66\degree$ to $-106\degree$. The presence of MSC was confirmed via temporal Fourier transform. These cues were not picked up by the subjects resulting in front-back confusion. Subjects still performed well, with at least 75% responses of Type 0, outside the zone. Thus, with the continuous localization process as explained in Section 4.3, the 40$\degree$ region of uncertainty on either side should not significantly impair binaural navigation.

7. PROTOTYPE

Following the feasibility study, an android application was developed. Given a pre-determined route, the user’s current coordinates and head orientation, the application seeks the relevant binaural cue from the user’s spatial audio database for playback via stereo in-ear earphones at fixed time intervals. The pre-determined route was broken down into a series of straight-line paths. A virtual source is placed at the nearest junction of the straight-line paths, breaking the navigation down into a series of localization tasks. The user’s coordinates were manually fed to the application but high-precision GPS solutions are available to automate the process. The head orientation was determined with a compass. Due to time constraints, only one complete spatial audio database was generated. Subjects were blindfolded and tasked to navigate the route by following the binaural cues. The path taken was recorded every 2 steps and scored by measuring the root mean squared deviation (RMSD) from the pre-determined route.

23 of 30 subjects kept close to the path (min RMSD: 366mm, mean RMSD: 530mm) and reached the end in a timely fashion. While moving to new straight-line sections, movement
planned for an off-the-shelf solution to the hardware requirements for a fully functional navigation application in the outdoor environments of Singapore.

**ACKNOWLEDGMENT**

This work is supported by the SUTD-MIT International Design Center Grant.

(IDG31200107 / IDD11200105 / IDD61200103)

9. REFERENCES


