

A Novel Sound Localization Experiment for Mobile Audio Augmented Reality Applications

Author:
Mariette, Nick

Publication details:
Advances in Artificial Reality and Tele-Existence
pp. 132-142
978-3-540-49776-9 (ISBN)

Publication Date:
2006

Publisher DOI:
http://dx.doi.org/10.1007/11941354_15

License:
https://creativecommons.org/licenses/by-nc-nd/3.0/au/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/44729 in https://unsworks.unsw.edu.au on 2024-03-28

A Novel Sound Localization Experiment for Mobile Audio Augmented Reality Applications

Nick Mariette

Audio Nomad Group, School of Computer Science and Engineering
University of New South Wales, Sydney, Australia
nickm@cse.unsw.edu.au

Abstract. This paper describes a subjective experiment in progress to study human sound localization using mobile audio augmented reality systems. The experiment also serves to validate a new methodology for studying sound localization where the subject is outdoors and freely mobile, experiencing virtual sound objects corresponding to real visual objects. Subjects indicate the perceived location of a static virtual sound source presented on headphones, by walking to a position where the auditory image coincides with a real visual object. This novel response method accounts for multimodal perception and interaction via self-motion, both ignored by traditional sound localization experiments performed indoors with a seated subject, using minimal visual stimuli. Results for six subjects give a mean localization error of approximately thirteen degrees; significantly lower error for discrete binaural rendering than for ambisonic rendering, and insignificant variation to filter lengths of 64, 128 and 200 samples.

1 Introduction

Recent advances in consumer portable computing and position sensing technologies enable implementation of increasingly sophisticated, light-weight systems for presenting augmented reality (AR) and mixed reality (MR) environments to mobile users. Greater prevalence of this technology increases the potential for more common usage of AR/MR as a form of ubiquitous computing for information and entertainment applications. Furthermore, audio-only AR/MR applications allow for less encumbered use than visual AR/MR applications, since the output device is a set of headphones, which are less intrusive and more familiar to the general public than visual devices such as the head mounted display (HMD).

The concept of audio augmented reality, proposed at least as early as 1993 [1], is to present an overlay of synthetic sound sources upon real world objects that create aural and/or visual stimuli[1] [2]. Also in 1993, even before the completion of the Global Positioning System (GPS), the concept was proposed to use GPS position tracking in a personal guidance system for the visually impaired, by presenting the user with

[1] In this paper, the augmentation of real *visual* stimuli with virtual sound will be considered audio AR, although existing definitions of AR and MR are not clear with regards to cross-sensory stimuli for the real and virtual components of the user's environment [2].

virtual sound beacons to guide their travel [3]. Since then, several outdoor, GPS-based audio AR implementations have been built as fairly bulky packages, for example, backpack-based systems [4], [5], or roll-around cases [6]. In 2001-2004, the indoors LISTEN project [7], had high resolution, sub-decimetre tracking, and further reduced the worn system to passive tracking signal emitters and headphones, with remote tracking and spatial audio rendering. A substantial collection of these projects and other relevant literature is reviewed in [8].

With cheap digital compasses, powerful portable computers, lightweight consumer GPS receivers (and soon, Galileo European Satellite Navigation System receivers), implementation of affordable portable outdoors audio AR systems is possible. Potential applications include personal tourist guides, location-based services and entertainment, or even navigation for the visually impaired. However, despite this burgeoning potential, little evaluation has occurred on the usability and perceptual performance afforded by these systems. Subjective testing of mobile audio AR systems has often been limited to simply verifying functionality. Some examples of evaluations in the literature are discussed in the next section.

In a separate field of research, the human ability to localize real and synthetic spatial sounds has been extensively studied via laboratory-based perceptual experiments, yielding fine-grained results on the effects of many factors. These experiments tend to neglect ecological factors relevant to the audio AR situation, where the creative content designer intends synthetic sounds to be perceived as co-incident with real audible and visible objects in uncontrolled environments. In AR, as in the real world, with simultaneous, distracting ambient stimuli from other foreground and background objects, it is important that people can maintain direct or peripheral awareness of aural and visual object positions while moving.

The present experiment is designed to evaluate perception quality afforded by practical mobile audio AR systems, such as "Campus Navigator" – a tour guide demonstrator being built by the Audio Nomad research group[2]. Firstly, the experiment verifies that a pedestrian user can localize synthetic binaural spatial audio in relation to real stationary visible objects, and indicate their judgment by walking. Secondly, it examines binaural rendering factors' effects on localization errors, informing software design decisions that balance perceptual performance with limited speed of portable computing. Further, the experiment controls for the effects of latency and accuracy of position and orientation information by using static, pre-rendered spatial audio stimuli with a mobile subject response method. Finally, validation of the novel response method, by cross checking results against similar laboratory experiments, sets precedence for using similar response methods in future AR audio localization experiments.

2 Background

Few examples of sound localization research incorporate ecological validity to the AR setting by including interaction via body translation motions (not just head-turns),

[2] http://www.audionomad.com

and/or multimodal stimuli. In 1993, Cohen et al [1] presented a very early AR study verifying two subjects' ability to successfully co-localize a virtual binaural sound source with a real sound source received via telepresence from a robot-mounted dummy head. Since then, limited evaluation has occurred for many audio AR projects, up to and including the sophisticated LISTEN project of 2001-2004 [7]. Following, is a brief discussion of selected experiments with quantitative evaluations.

Cheok et al [9] used a visual AR environment to assess 3D sound's impact on depth and presence perception, and audio/visual search task performance, showing all three performance indicators improved using 3D sound. Ecological validity to the mobile, outdoors AR setting is limited due to the HMD visuals and tethered position and head orientation tracking, confined to a 3x3 metre area. Also, the performance metrics of depth judgment rate, task performance time and questionnaire results do not compare easily with traditional sound localization experiments.

Härmä et al [8] described the use of their wearable augmented reality audio (WARA) system for "preliminary listening tests". The subject is seated in a laboratory with stationary head position, and is requested to indicate whether a test signal was virtual or originated from a loudspeaker placed out of sight. Results showed subjects could not discriminate between virtual and real sound sources, with audio rendering using individualized head related impulse responses (HRIRs). Relevance to mobile AR is limited by lack of subject interaction via head-turns or position translation.

Walker and Lindsay [10] presented an investigation of navigation efficiency with respect to waypoint beacon capture-radius in an audio-only virtual environment. The use of navigation performance tasks to study the effect of rendering system factors was novel, yet relevance to mobile AR is limited due to only implementing the auditory modality, the lack of subject motion interaction, and a purely virtual environment. Yet, subject tasks might be successfully transferred to mobile AR studies.

Loomis [3] presents the subjective sound localization research most relevant to the mobile AR setting, based on the Personal Guidance System for the visually impaired. One study examines distance perception [11], using a novel outdoors subjective method of "measurement of perceived distance using perceptually directed action", whereby subjects' judgments were indicated via the open-loop spatial behaviour of pointing at the perceived location of the auditory image while moving. Loomis' research bears strong relevance to the present work, although to best of the author's knowledge the study of angular localization has not occurred.

Having discussed applied AR studies incorporating 3D sound, we will briefly address relevant laboratory-based research on fundamental human sound localization ability. Experiments are often designed for precision with respect to specific, often artificial factors (e.g. stimuli frequency spectrum), rather than ecological validity to a particular application environment. Three relevant research topics are: studies on localization precision, multimodal stimuli, and head-turn latency.

Localization precision afforded by binaural 3D sound rendering methods may be compared to baseline localization ability of about one degree minimal audible angle (MAA) in the horizontal plane [12]. This research provides a basis for expected performance, subjective experimental methods and associated performance measures such as mean localization error or response time to localize brief sound stimuli.

Strauss and Buchholz [13] compared localization error magnitudes for amplitude panning and first order ambisonic rendering methods to a six-channel hexagonal speaker array. With head movements permitted (allowing more accurate localization due to the closed perception-action feedback loop), the mean localization error was 5.8 degrees for amplitude panning (AP) and 10.3 degrees for ambisonic rendering (Ambi). Without head movements, mean errors were 15.8 degrees (AP) and 18.5 degrees (Ambi). The present study uses virtual amplitude panning and ambisonic rendering for binaural output, by replacing speakers with convolution by HRIR pairs.

One multimodal aspect is the "ventriloquist effect" [14], identified as a visual bias on sound localization during simultaneous presentation with visual objects. Larsson et al [15] also noted higher level cognitive effects of improved presence, focus, enjoyment and faster completion of navigation tasks for virtual visual environments augmented with correlated auditory stimuli. These results inform the decision to trial the visual/motile response method in the present experiment. Future experiments will investigate how multimodal perception might mitigate system latency limitations.

System latency to head-turns is known to affect localization ability for real-time binaural spatial audio. Brungart et al [16] discovered that system head-turn latency is detectable above 60-80 milliseconds for a single sound source, or above 25ms when a low-latency reference sound is present, as per the case of virtual sound sources augmenting real sources. The present study notes this result by using an experimental design that controls for position/orientation latency effects to isolate the rendering method effects. Using static pre-rendered virtual sources and requiring subjects to respond by moving relative to a static visual reference object, the experiment exhibits infinite latency to head orientation and position. Future experiments will re-introduce latency, studying its effects on localization and task performance in AR settings.

3 Experimental Procedure

The experiment was performed in a flat, open, grassy space, in clear weather conditions during daylight hours. To date, six volunteers (all male, aged in their 20s) have performed the experiment. Subjects wore/carried a system comprised of: a set of headphones; a position tracking system mounted at the centre back of the waist; and a portable computer running custom experiment software that displayed a graphical user interface, played sound stimuli and logged user positions.

The positioning system, a Honeywell DRM-III [17], combines an inertial navigation system (INS), a GPS receiver, pedometer, digital compass and barometric altimeter (that can all be individually activated/deactivated), with optional Kalman filtering and a serial RS232 interface. Stated INS position accuracy is 2-5% of distance traveled and the compass is accurate to within one degree. A feasibility study by Miller [18] using the DRM-III, suggests that positioning accuracy varies significantly according to usage factors such as stride length variation. We executed a preliminary performance test, obtaining the most accurate positioning for small distances (tens of metres) by using only the INS and digital compass. It was also necessary to request subjects to move only in the direction their torso was facing, never sideways, only changing direction by on-the-spot rotation.

Other equipment included Sennheiser HD485 headphones (an economical, open backed, circumaural design) and a Sony Vaio VGN-U71 touch-screen handheld computer with a Pentium M processor, running Windows XP Service Pack 2. Present experiment software is not computationally taxing, however this powerful portable platform will be necessary for future experiments employing real-time binaural rendering. The DRM-III interfaced to the Vaio with a Keyspan USB-Serial interface.

3.1 Subject task and instructions

The experiment configuration (Fig. 1) used a camera tripod as the visual reference object, placed at the end of a straight, fifteen-metre *reference line* from the *base position*. Each subject listened to 36 binaural stimuli and responded to each by walking forward until the tripod position corresponded to the perceived auditory image position. For example, if the sound seemed to be located 45 degrees to the *right*, the subject walked to the *left* until the tripod was positioned 45 degrees to their right. Subjects were asked to keep their heads parallel to the reference line when making localization judgments, achieving this by fixing their gaze on a distant object past the tripod in the direction of the reference line. Subjects were also asked to judge source distance, and advised that all stimuli matched a tripod position in front of them – thereby avoiding the occurrence of front-back confusions.

The experiment user interface is simple, with only two buttons (Fig. 2). For each stimulus, the subject begins at the base position, facing the tripod, and clicks the green (first) button to start the sound. The stimulus plays for up to 50 seconds, during which the subject walks to match the tripod position with the perceived auditory image position. Clicking the red (second) button stops the stimulus, and the subject returns to base, ready for the next sound. After the final stimulus, the subject records a walk from the base position to the tripod, capturing a reference track.



Fig. 1. Experiment layout with base position, tripod and 15m reference line.

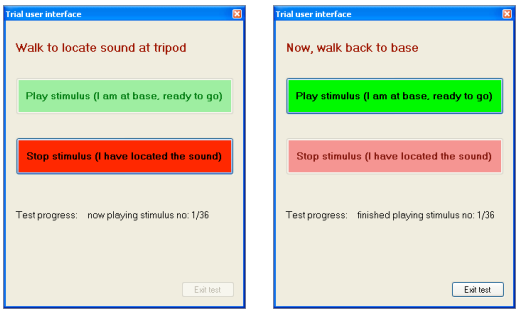


Fig. 2. Graphical interface used by the subject to run the experiment.

Experiment software resets the INS position when the green *play* button is clicked and records the subject's position 4 times per second until the red *stop* button is clicked. For each subject, stimuli order is randomized, avoiding bias effects such as progressive fatigue during the experiment.

For each test, 37 position log files are recorded, representing 36 stimuli tracks and one reference track. The stimuli play order is also recorded. Subsequent data analysis is performed in Matlab using several custom scripts.

3.2 Binaural Stimuli and Factors

A single, mono white noise sample was processed in Matlab into 36 binaural stimuli, each created using a different combination of three factors: azimuth angle, filter length and rendering method. The process used HRIRs taken directly from subject number three in the CIPIC database [19]; the subject chosen arbitrarily due to lack of literature recommending a single preferable set.

Filter length was chosen because a tradeoff exists between the need for fast computation (requiring shorter filters), and high rendering quality (requiring longer filters). An optimal rendering system would use the shortest possible filters that don't significantly affect perceptual performance. Three different HRIR filter lengths were obtained: the 200-sample originals and new 128 and 64-sample versions, created by truncating the tail using a rectangular window.

Two rendering methods were used: discrete binaural rendering and a virtual ambisonic method. Discrete rendering simply convolved the source audio with the appropriate left and right-ear HRIRs of each length and azimuth angle. The virtual ambisonic method, adapted from [20], multiplied the source by a panning vector to become a four-channel b-format signal, subsequently decoded via a twelve-channel "virtual speaker array" of twelve HRIR pairs, resulting in the final binaural output.

Rendering method was a focal point because ambisonic rendering is more computationally efficient, scaling at a much lower rate per sound source than discrete rendering. However, localization accuracy afforded by first-order ambisonic rendering is expected to be lower than for discrete rendering [13].

Ambisonic rendering requires a constant computation load equivalent to five HRIR convolutions to convert the b-format signal into binaural, with only four additional multiply-accumulate (mac) operations per sound source to create the b-format signal. In comparison, discrete rendering requires two HRIR convolutions per sound source, with two mac operations to mix in each additional source.

A further ambisonic advantage is that the intermediate b-format signal can be easily rotated relative to listener head-orientation at a stage between mixing mono sources to b-format and rendering to binaural. A distributed rendering architecture becomes possible where many sources are mixed to b-format on a powerful, capacious server, the b-format streams wirelessly to a computationally limited portable device that rotates it with head-turns, and renders to binaural as close as possible (with lowest latency) to the orientation sensor. Since perceptual quality is significantly affected by latency to head-turns [16], the ambisonic method is preferable if it has insignificant effects on localization ability.

Each combination of factors (three HRIR filter lengths and two rendering methods) was used once to generate stimuli at six azimuth angles: -65, -35, -15, 10, 25 and 45 degrees from the median plane. Stimuli were amplitude normalized across, but not within, rendering methods. Nevertheless, stimuli amplitude should only affect distance perception, which is not analyzed in this paper.

4 Results Analysis and Discussion

Each subject's raw track data was imported into Matlab and matched to corresponding stimuli factors using the play order record. For each stimulus, perceived direction and localization errors are calculated and tabulated with respect to subject, stimulus azimuth, filter length and rendering method.

Fig. 3 shows all six subjects' raw tracks, rotated for display so the reference track runs due north, from the base position to the tripod. We can see that movement style is fairly individual to each subject, some honing their localization in a piece-wise manner, correcting many times (e.g. subject 1), while others choose a single direction at the outset and walk until they achieve localization (e.g. subject 5). Subject 4 appears to have made gross position misjudgments (or misunderstood instructions), having crossed from one side of the reference line to the other for two localizations.

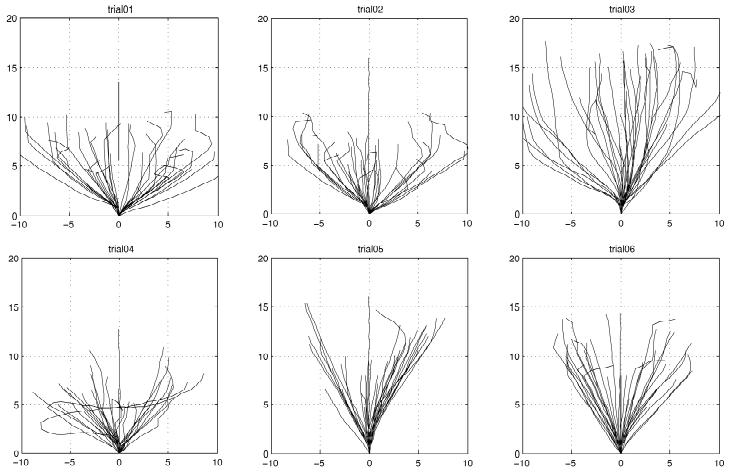


Fig. 3. Raw position tracks for each subject (both x and y axes in metres).

For each subject's set of raw tracks, we assume the tripod to be located exactly 15 metres from the base point, in the reference track direction. Thus the perceived distance and direction of each localization judgment can be calculated as a vector from each recorded stimulus track terminal to the assumed tripod position. We know that every recorded track includes INS positioning errors, but the actual reference line is a measured 15 metres. While the recorded reference track length may not be precisely 15m, assuming the tripod position avoids summing INS positioning errors for stimulus and reference tracks, which are likely to be uncorrelated due to different types of movements that created them. Thus the recorded reference track is used for its angular heading and as a basic reality check for correct tracking.

A one-way ANOVA test across subjects, with a post-hoc multiple comparison using Tukey's Honestly Significant Difference (HSD) (Fig. 4) showed that Subject 3 had significantly different mean absolute azimuth error from all other subjects ($F(5,190)=8.1$, $p<0.001$). With Subject 3's data removed, the same tests (now for $p<0.05$) show no significant difference between remaining subjects (Fig. 5).

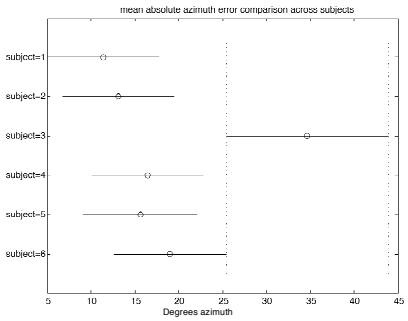


Fig. 4. Multiple comparison test of mean absolute azimuth error for six subjects. Five subjects have significantly different marginal means to Subject 3 (p<0.001).

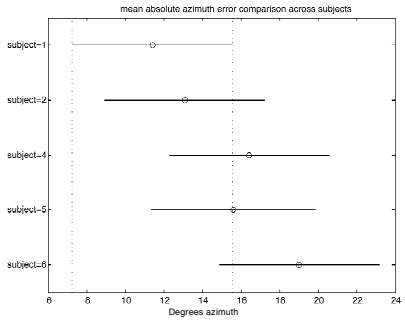


Fig. 5. Multiple comparison test of mean absolute azimuth error for five subjects after removing Subject 3. No subjects have significantly different marginal means (p<0.05).

Cross-checking with notes taken during the experiment, Subject 3 mentioned a high rate of front-back confusions and did not follow instructions to keep the tripod positioned in front (necessary to control for this type of confusion). Tracks in Fig. 3 confirm that Subject 3 often moved to the far end of the reference line. Due to this significant difference, Subject 3's results are removed from all subsequent analyses.

Next, we present scatter plot analyses of the remaining subjects' perceived azimuth, across single factors (Fig. 6), and factor pairs (Fig. 7). The ideal response would be points on a diagonal line, with perceived and actual azimuth values matching exactly. The results show more accurate localization for discrete rendering than ambisonic rendering, and a general agreement between perceived and intended azimuth for all factors, verifying that all subjects achieved some degree of correct localization using the novel mobile, multimodal response method.

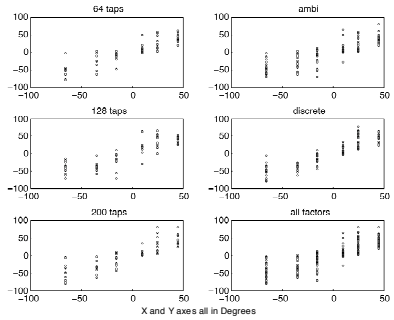


Fig. 6. Scatter plot analysis of perceived azimuth by single factors for all subjects: filter length on left; rendering method and all factors on right. X-axis is intended azimuth, Y-axis is perceived azimuth, both in degrees.

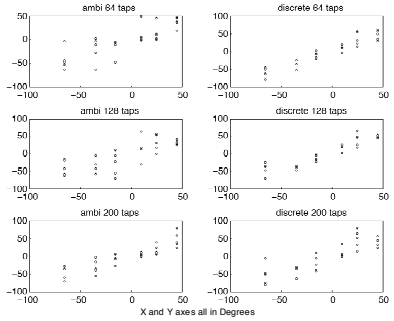


Fig. 7. Scatter plot analysis of perceived azimuth by paired factors for all subjects: filter length varies top to bottom; rendering method varies left to right. X-axis is intended azimuth, Y-axis is perceived azimuth, both in degrees.

Fig. 8 presents a three-way ANOVA test of mean absolute azimuth error for the 5 remaining subjects, across factors of azimuth (a reality check), rendering method and filter length. Significant effects are observed due to azimuth ($F(5,152)=3.16$; $p<0.01$); render method ($F(1,152)=6.84$; $p<0.01$); and interaction between render method and filter length ($F(2,152)=3.66$; $p<0.05$). For reference, "ambi render?" is a label for rendering method, set to 1 for ambisonic, 0 for discrete rendering. The question arises why azimuth significantly affects the mean azimuth error, even though it has insignificant effect in combination with any other factor, as should be expected.

Analysis of Variance

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
azimuth	2405.8	5	481.15	3.16	0.0096
ambi render?	1039.6	1	1039.61	6.84	0.0098
filter length	189.2	2	94.6	0.62	0.5382
azimuth*ambi render?	1064	5	212.8	1.4	0.2276
azimuth*filter length	1253.3	10	125.33	0.82	0.606
ambi render?*filter length	1113.2	2	556.61	3.66	0.028
Error	23116.3	152	152.08		
Total	30152.6	177			

Fig. 8. Multi way ANOVA test of mean azimuth error for 5 remaining subjects, across factors: azimuth, render method and filter length.

A post-hoc multiple comparison test using Tukey's HSD (Fig. 9) reveals that only stimuli at -65 degrees azimuth have a significant effect on mean absolute azimuth error ($p<0.05$). This is the greatest absolute angle, so the larger mean error might be explained by these stimuli requiring the most subject movement, causing greater position tracking errors. This angle also positions the tripod furthest into the subjects' peripheral vision, maximizing the likelihood of aural/visual localization mismatch. No other stimulus angle has a significant effect on mean absolute azimuth error, so we shall accept this reality check to hold.

A final post-hoc multiple comparison test using Tukey's HSD (Fig. 10) shows the significant effect of rendering method on mean absolute azimuth error, ($p<0.05$).

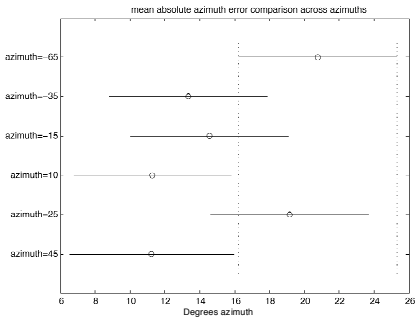


Fig. 9. Multiple comparison of mean azimuth error across azimuths, for 5 subjects. Two groups have marginal means significantly different from azimuth = -65 ($p<0.05$).

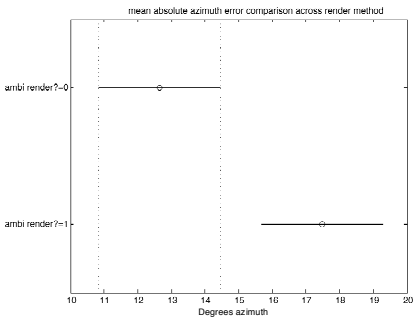


Fig. 10. Multiple comparison of mean azimuth error across rendering methods, for 5 subjects. "ambi render?=0" is discrete rendering, "ambi render?=1" is ambisonic rendering. They have significantly different marginal means ($p<0.05$).

Results show a mean absolute azimuth error of 13 degrees for discrete rendered stimuli, versus approximately 17.5 degrees for ambisonic rendered stimuli. These values correspond closely to results of Strauss and Buchholz' experiment for subjects seated in a laboratory, localizing sounds rendered to a hexagonal speaker array [13]. For subjects with unrestricted head movements, their experiment produced a mean azimuth error of 5.8 degrees for amplitude panning and 10.3 degrees for ambisonic rendering. Amplitude panning is equivalent to discrete binaural rendering for sounds aligned to the speaker directions, indicating a relevance to these results. Error magnitude differences between the Strauss and Buchholz results and the present results might be attributed to: use of non-individualized HRIRs, INS position tracking errors and the subject response method being less precise. Nevertheless, our novel methodology is validated by a reasonable mean absolute azimuth error of 13 degrees, with discrete panning affording better localization than ambisonic panning.

5 Conclusion

Preliminary results are presented for an outdoors sound localization experiment using static, pre-rendered binaural stimuli to study the effect of HRIR filter length and ambisonic or discrete binaural rendering on angular localization errors. A novel response method was employed, where subjects indicated the perceived sound source location by walking to match the auditory image position to a real visual object.

Results for 5 subjects show a mean absolute azimuth error of 13 degrees for discrete rendering – significantly better than 17.5 degrees error for ambisonic rendering. This variation according to rendering method compares well with other researchers' results for static laboratory experiments. HRIR filter lengths of 64, 128 and 200 samples show no significant effect on azimuth error.

The results validate the novel outdoors experiment and subject response method designed to account for multimodal perception and subject interaction via self-motion, both often ignored by traditional sound localization experiments. Thus, the novel methodology presented can be considered more ecologically valid for studying perceptual performance afforded by mobile audio AR systems.

Acknowledgments. Audio Nomad is supported by an Australian Research Council Linkage Project with the Australia Council for the Arts under the Synapse Initiative.

References

1. Cohen, M., Aoki, S., and Koizumi, N. Augmented Audio Reality: Telepresence/AR Hybrid Acoustic Environments. in IEEE International Workshop on Robot and Human Communication. (1993).
2. Milgram, P. and Kishino, F., A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems, (1994). E77-D(12).
3. Loomis, J.M. Personal Guidance System for the Visually Impaired Using GPS, GIS, and VR Technologies. in VR Conference. (1993). California State University, Northridge.

4. Holland, S., Morse, D.R., and Gedenryd, H. AudioGPS: Spatial Audio in a Minimal Attention Interface. in Proceedings of Human Computer Interaction with Mobile Devices. (2001).
5. Helyer, N. (1999-2001), Sonic Landscapes Accessed: 22/8/2006 http://www.sonicobjects.com/index.php/sonicobjects/more/sonic_landscapes/
6. Rozier, J., Karahalios, K., and Donath, J. Hear & There: An Augmented Reality System of Linked Audio. in ICAD 2000, Atlanta, Georgia, April 2000. (2000).
7. Olivier Warusfel, G.E. LISTEN - Augmenting Everyday Environments through Interactive Soundscapes. in IEEE VR2004. (2004).
8. Härmä, A., et al. Techniques and Applications of Wearable Augmented Reality Audio. in AES 114TH Convention. (2003). Amsterdam, The Netherlands.
9. Zhou, Z., Cheok, A.D., Yang, X., and Qiu, Y., An Experimental Study on the Role of Software Synthesized 3D Sound in Augmented Reality Environments. Interacting with Computers, (2004). 16: p. 989.
10. Walker, B.N. and Lindsay, J. Auditory Navigation Performance Is Affected by Waypoint Capture Radius. in ICAD 04 - The Tenth International Conference on Auditory Display. (2004). Sydney, Australia.
11. Loomis, J.M., Klatzky, R.L., and Golledge, R.G., Auditory Distance Perception in Real, Virtual and Mixed Environments, in Mixed Reality: Merging Real and Virtual Worlds, Ohta, Y. and Tamura, H., Editors. (1999): Tokyo. p. 201-214.
12. Grantham, D.W., Hornsby, B.W.Y., and Erpenbeck, E.A., Auditory Spatial Resolution in Horizontal, Vertical, and Diagonal Planes. Journal of the Acoustical Society of America, (2003). 114(2): p. 1009-1022.
13. Strauss, H. and Buchholz, J., Comparison of Virtual Sound Source Positioning with Amplitude Panning and Ambisonic Reproduction. The Journal of the Acoustical Society of America, (1999). 105(2): p. 934.
14. Choe, C.S., Welch, R.B., Gilford, R.M., and Juola, J.F., The "Ventriloquist Effect": Visual Dominance or Response Bias? . Perception & Psychophysics, (1975). 18: p. 18, 55-60.
15. Larsson, P., Västfjäll, D., and Kleiner, M. Ecological Acoustics and the Multi-Modal Perception of Rooms: Real and Unreal Experiences of Auditory-Visual Virtual Environments. in International Conference on Auditory Display. (2001). Espoo, Finland.
16. Brungart, D.S., Simpson, B.D., and Kordik, A.J. The Detectability of Headtracker Latency in Virtual Audio Displays. in International Conference on Auditory Display. (2005). Limerick, Ireland.
17. Point Research, DRM-III Oem Dead Reckoning Module for Personnel Positioning. (2002): Fountain Valley, California.
18. Miller, L.E., Indoor Navigation for First Responders: A Feasibility Study. (2006), National Institute of Standards and Technology.
19. Algazi, V.R., Duda, R.O., Thompson, D.M., and Avendano, C. The CIPIC HRTF Database. in Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics. (2001). Mohonk Mountain House, New Paltz, NY.
20. Noisternig, M., Musil, T., Sontacchi, A., and Höldrich, R. A 3D Real Time Rendering Engine for Binaural Sound Reproduction. in International Conference on Auditory Display. (2003). Boston, MA, USA.