

A COMPARISON OF HEAD-TRACKED AND VEHICLE-TRACKED VIRTUAL AUDIO CUES IN AN AIRCRAFT NAVIGATION TASK

Douglas S. Brungart¹, Brian D. Simpson¹, Ronald C. Dallman²,
Griffin Romigh¹, Richard Yasky³, John Raquet⁴

¹Air Force Research Laboratory, WPAFB, OH

²General Dynamics Advanced Information Systems, Dayton, OH

³Nasa Langley Research Center, Hampton, VA

⁴Air Force Institute of Technology, WPAFB, OH

douglas.brungart@wpafb.af.mil

ABSTRACT

Since the earliest conception of virtual audio displays in the 1980's, two basic principles that have guided their development have been 1) that virtual audio cues are ideal for providing information to pilots in aviation applications; and 2) that head-tracked virtual audio displays provide more accurate and more intuitive directional information than non-tracked displays. However, despite the obvious potential utility of spatial audio cues in the cockpit, very little quantitative data has been collected to evaluate the in-flight performance of pilots using virtual audio displays. In this study, sixteen pilots maneuvered a general aviation aircraft through a series of ten waypoints using only direction cues provided a virtual audio display system. Each pilot repeated the task twice: once with a virtual display slaved to the direction of the pilot's head, and once with a virtual audio display slaved to the direction of the aircraft. Both configurations provided audio cues that were sufficient for successful aircraft navigation, with pilots on average piloting their aircraft to within 0.25 miles of the desired waypoints. However performance was significantly better in the plane-slaved condition, primarily due to a leftward bias in the head-slaved flight paths. This result suggests how important frame of reference considerations can be in the design of virtual audio displays for vehicle navigation.

1. INTRODUCTION

In the early to mid 1980's, when virtual audio displays were first conceived as a way to provide intuitive spatial information to listeners wearing headphones, aircraft cockpit displays were among the earliest applications envisioned for them. Spatial audio cues were seen as a way to reduce pilot workload by offloading some visual information from the chronically overcrowded control panels of modern aircraft. They were also seen as a way to enhance situational awareness by providing information about the locations of objects or events outside the pilot's visual field of view. Because of these potential benefits, virtual audio displays were included as one of the key components of the influential Super Cockpit program at WPAFB in the mid 1980's [1].

However, despite the obvious potential advantages of virtual audio displays in aviation applications, very little quantitative data has been collected to evaluate how useful spatial audio cues might be in actual flight operations. In the early 1990's, three pilots evaluated the utility of a virtual audio display in a modified TAV-8B

VTOL aircraft [2]. The results showed a subjective decrease in target acquisition times, a subjective improvement in speech intelligibility, a subjective increase in situational awareness, and a subjective decrease in pilot workload. Similar subjective improvements were found in follow-on studies conducted with the TAV8-B and with a modified OV-10 aircraft. A more recent study conducted by our laboratory evaluated how well three pilots were able to maneuver a Beechcraft King Air twin-engine aircraft in the direction of a spatialized audio navigation beacon. The results showed a mean error of approximately 9° between the audio beacon location and the final aircraft heading. [3, 4]. These studies all suggest that virtual audio displays can and do provide some performance benefits to pilots flying actual aircraft. However, the limited number of data points available makes it difficult to quantify how large these benefits might be in normal flight operations.

In this study, we were interested in determining how effectively virtual audio cues could be used to convey georeferenced spatial information in operational flight tasks. Although a number of possible response metrics were considered, in the end we decided to base the experiment on a navigational flight task that required pilots to steer their aircraft through a series of fixed georeferenced virtual audio "waypoints." This response metric allowed an assessment of the true operational effectiveness of the audio cues, including any directional errors that might be have been introduced due to inaccuracies in the headtracking device and any discrepancies between the perceived and actual orientations of the audio waypoints relative to the aircraft. We were also interested in determining how important dynamic head-tracking cues might be for obtaining good performance from a spatialized audio navigation display. In most virtual audio display applications, dynamic head-tracking cues play an absolutely critical role in helping listeners distinguish between sounds that fall on the same "cone-of-confusion" (i.e. are at the same lateral angle) [5]. Thus, without head-tracking, most listeners will experience a large number of "front-back" confusions about the locations of virtual sound sources [6]. This is especially true in applications of virtual audio displays where listeners are forced to rely on "generic" Head-Related Transfer Functions (HRTFs) rather than customized HRTFs measured on their own ears [7]. However, the practical reality is that the cost of an interactive virtual audio display is now driven *almost entirely* by the cost of the headtracking device: virtual audio display rendering can be done at little or no cost with open-source software running on a general purpose PC, but the



Figure 1: Cirrus SR-22 aircraft used for the flight tests.

cost of a high-quality headtracker suitable for laboratory still costs many hundreds of dollars, and a more rugged headtracker suitable for use in an aircraft cockpit is even more expensive. Thus we felt it was important to compare the performance of a high-fidelity head-tracked virtual audio display to the level of performance that could be achieved with a lower-cost “plane-tracked” display that updated the position of the sound on the basis of the current heading of the aircraft (which could be obtained at no cost from the onboard aircraft instrumentation).

2. METHODS

2.1. Participants

Sixteen male pilots were recruited by the NASA Langley Research Center (NLRC) to serve as subject pilots. The pilots were screened to meet the following six criteria:

- Possess a current private or commercial pilot certificate with instrument rating
- Have performed at least three takeoffs and landings within the last 90 days
- Possess a current medical flight physical
- Have fewer than 3000 hours of total flight time
- Possess a high performance endorsement or experience
- Be at least 18 years of age

2.2. Aircraft

The flight tests were conducted with a single-engine, four-seat general aviation aircraft (Cirrus SR-22) stationed at the NASA Langley Research Center (Figure 1). In addition to the normal IFR cockpit instrumentation, the aircraft was equipped with a custom air data and attitude/heading reference system (ADAHRS) that provided real-time measurements of a wide variety of in-flight variables including the aircraft’s GPS location, altitude, velocity and orientation. It was also equipped with an on-board windows-based PC, which was used as the control computer for this experiment. For the purposes of this study, the aircraft was reconfigured



Figure 2: Configuration of orientation tracking device in the head-slaved (left panel) and plane-slaved (right panel) conditions of the experiment. In each case, the location of the IMU unit has been circled.

in a three-seat configuration, with the subject pilot in the front left seat, the safety pilot in the front right seat, and the experimenter in the right rear seat. The left rear seat was replaced with a customized pallet that allowed the experimenter to access the specialized audio equipment that was installed for this experiment.

2.3. Audio Display System

The spatial audio cues used in the experiment were generated by the Sound Lab (SLAB) audio rendering package developed at the NASA-AMES research center [8]. The SLAB software was loaded with a set of 128-point HRTFs that were measured on the ears of one of the authors (DSB) at Wright-Patterson Air Force Base. In order to allow the subject pilot to converse with the test pilot and the experimenter, the stereo output of this audio display system was mixed with the monaural output of the aircraft intercom system using a 4-channel analog mixer (Allen & Heath Phoenix). The resulting stereo signal was presented to the listener via ANR stereo headphones (DRE-6500).

2.4. Orientation Tracking System

The orientation tracking system used in the experiment was a custom GPS-aided MEMS system that was developed at the Air Force Institute of Technology (AFIT) [4]. This system consisted of a microelectromechanical system (MEMS) inertial measurement unit (IMU), a Global Positioning System (GPS) receiver (Garmin Industries), and a PC-based embedded computer system. The system differed from traditional inertial head-tracking systems in that it used acceleration information derived from the GPS system, rather than a digital compass or tilt meter, to correct for drift in the MEMS gyros. This means that the system may slowly drift in azimuth over prolonged periods of straight-and-level flight, but this drift is automatically corrected as soon as the aircraft experiences a significant amount of lateral acceleration from a turn. The tracking system was deployed differently in the two different tracking conditions of this experiment, as shown in Figure 2: In the *head-slaved* condition (left panel), the MEMS sensor was attached to the top of the subject pilot’s headset and it reported the orientation of the pilot’s head in world coordinates. In the *plane-slaved* condition (right panel), the MEMS sensor was rigidly affixed to the

center console of the aircraft and it reported the orientation of the aircraft in world coordinates.

2.5. Stimuli

The spatialized auditory navigation cues used in the display were generated by convolving pre-recorded waveforms of a male talker speaking the word “waypoint” with the appropriate HRTF required to shift its apparent location to the desired destination waypoint of the aircraft. Two different methods of spatialization were employed. In the “head-slaved” condition, the auditory stimuli were updated in real time in response to changes in the orientation of the subject pilot’s head. In the “plane-slaved” condition, the auditory stimuli were updated in real time in response to changes in the position and orientation of the aircraft. In both cases, “ground track” information from the ADAHARS was used to correct the directional cue to account for the effects of any crosswinds in the area. This information was used to determine the true aircraft heading that was required to point the aircraft’s “ground track vector” (i.e. the net forward motion of the aircraft plus the effects of any head, tail, or cross winds) directly towards the location of the desired waypoint. Thus, in the plane-slaved condition, the navigational cue sounded like it was directly in front of the listener when the aircraft itself was headed directly toward the waypoint. Similarly, in the head-slaved condition, the navigational cue sounded like it was directly in front when the listener’s head was pointed in the direction the plane needed to be pointed in order to travel directly to the waypoint.

The vocal effort level of the spoken “waypoint” cue also modified to provide the pilot with information about the distance to the next waypoint. Initially, when the waypoint was more than 3500 m away, the audio cue was presented at a loudly shouted level, conveying the impression of a relatively distant sound source. As the pilot approached the waypoint, the overall level of the audio cue remained the same, but the vocal effort level of the talker systematically decreased, until it reached a quiet conversational voice when the pilot was within 650 m of the waypoint¹. If the pilot missed the waypoint, the vocal effort level again increased. Previous experiments conducted in our laboratory have shown that vocal effort cues of this type can very effectively convey information about the absolute distance of a virtual sound source [9].

2.6. Procedure

The data for each subject pilot were collected in two test flights, each lasting a total of approximately 1.5 hours. Prior to the first test flight, each subject was instructed on the protocol of the experiment, and familiarized with the task and audio display by interacting with a custom flight simulation on a laptop computer that simulated what would be encountered during the flight test. Each subject also participated in a localization pre-test to verify they could accurately localize virtual sounds in the horizontal plane. Once in the air, the subject was given in-flight familiarization with the displays and controls of the aircraft, as well as additional exposure to the auditory display. The aircraft was then flown to a GPS

¹A total of six voice levels were employed, with the production level to the talkers voice (as measured with a microphone 1 m from the mouth) dropping from 90 dB to 60 dB in 6 dB steps as the distance to the waypoint passed boundary distances of 3500 m, 2000 m, 1500 m, 1000 m, and 650 m.

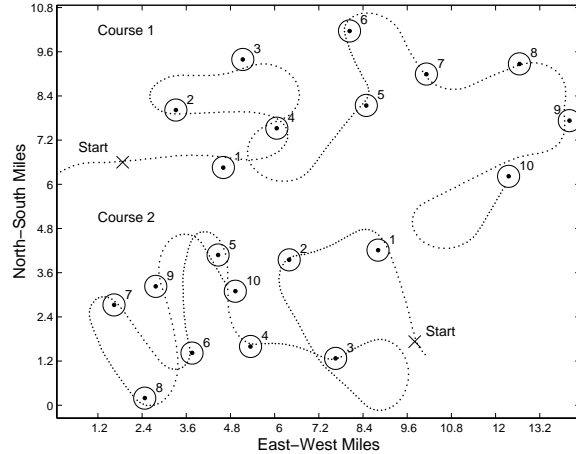


Figure 3: Examples of waypoint navigation for a typical plane-slaved flight through Course 1 (top) and a typical head-slaved flight through Course 2 (bottom). The circles show the 500 m distance limit that had to be met before a listener could proceed on to the next waypoint.

location designated as the start point of one of two different navigation courses. At this point the auditory display was enabled, the subject was given control of the aircraft, and formal data collection was initiated.

The subject pilot was then required to maneuver the aircraft through a navigation course consisting of 10 waypoints using only the auditory display to determine the direction of, and distance to, each waypoint. A waypoint was acquired if the aircraft was flown to within a 500 m tolerance range. Once the waypoint was acquired, the audio display immediately discontinued the directional audio cue associated with that waypoint and began playing the directional audio cue associated with the next waypoint in the 10-point course. If the waypoint was missed, the subject would have to fly out and do a procedural turn to attempt to re-acquire the waypoint. Two different, but equivalent courses were developed for the experiment (As shown in Figure 3), and the tracking conditions were balanced so that half the pilots flew the first course in the head-slaved condition, and half flew the first course in the plane-slaved condition. All of the flights were conducted in Visual Meteorological Conditions (VMC) with the plane traveling approximately 140 KIAS at an altitude ranging from 2000-8000 ft.

3. RESULTS

Figure 3 shows the flight path taken by the aircraft in two typical trials in the experiment. The top curve shows a flight path through the first course in the plane-slaved condition. In this course, the subject pilot was initially given control of the aircraft at the left side of the panel. The bottom curve shows a flight path through the second course in the head-slaved condition. In that course, the subject pilot initially received control of the aircraft at the right side of the panel.

As these flight courses illustrate, most of the subject pilots had little difficulty navigating either of the two flight courses. In most

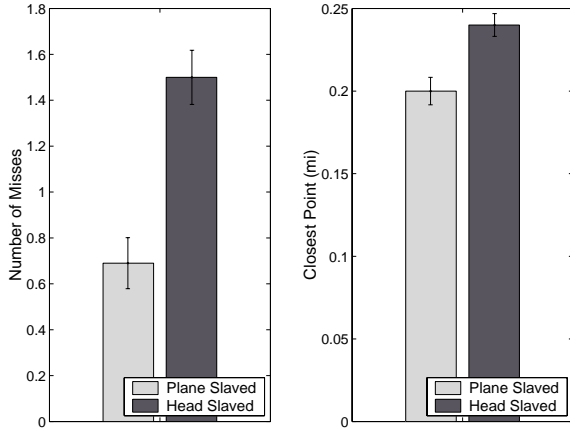


Figure 4: The left panel shows the average number of missed waypoints in each 10-point navigation course completed in each condition. The right panel shows the mean point of closest approach to the waypoints in each condition (in miles). The error bars show one standard error of the mean in each condition.

cases, they hit the waypoints directly, and used the corresponding change in the audio cue to immediately turn their aircraft in the direction of the next waypoint. However, in some cases, the subjects came close but did not quite reach the 500 m barrier required to hit the waypoint and move on to the next point in the course. When that happened, the pilots had to fly far enough past the waypoint to be able to loop back in a 180° turn and reach the waypoint from the other direction. This occurred on the path from point 9 to point 10 in the top panel of the figure, and in the path from point 2 to point 3 in the bottom panel of the figure. For the purposes of this paper, we will refer to incidents where the subject pilots flew past the waypoint but were further than 500 meters at the point of closest approach as “misses”.

One metric that can be used to evaluate overall performance in the navigation task is the average number of “misses” that occurred in each 10-point course. This is shown in the left panel of Figure 4. In both cases, the percentage of misses was relatively modest. However, it is clear that overall performance was significantly worse in the head-slaved condition (with misses in 15% of the trials) than in the plane-slaved condition (with misses in only 7% of the trials).

Of course, the selection of 500 m as the threshold distance required to hit a waypoint is a relatively arbitrary one. An alternative way to rate performance is to look at how close the pilots actually got to the waypoint before they flew past it, regardless of whether they actually hit the 500 m boundary required to qualify as a hit. This “point of closest approach” value is shown by the right panel of Figure 4. By this metric, we say that 1) the pilots were on average able to get even closer to the waypoint than the 500 m threshold value in both tracking conditions; and 2) the difference in performance between the head-slaved and plane-slaved condition was much smaller by this metric than by the “miss” metric, suggesting that the pilots in the plane-slaved condition often were very close to making the 500 m threshold value when they technically missed the waypoints.

An alternative metric of performance is the average total time required to complete the 10-point navigation course. The left pair

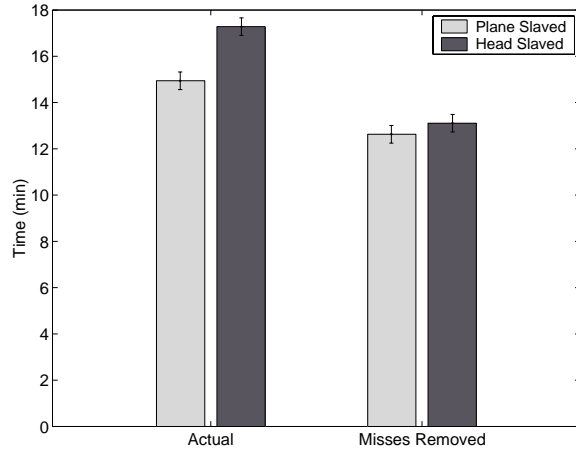


Figure 5: Average total time required to complete 10 waypoint navigation course. The left bars show the actual total time required. The right bars show the average total time required with the misses removed.

of bars in Figure 5 shows the raw data, which the entire time spent on the navigation course including any time spent recovering from missed waypoints. On average, the subject pilots spent about 15 minutes navigating the course in the plane-slaved condition and 17 minutes navigating the course in the head-slaved condition of the experiment. Of course, much of the extra time in the head-slaved condition was spent recovering from the extra misses that occurred in that condition. Each missed waypoint required quite a lot of time to fly out, turn around, and come back, so these misses could potentially cause a large skew in the data. An alternative way of analyzing the data is to count only the time spent from the arrival of one waypoint to the point of closest approach to the next waypoint. This analysis is shown in the right set of bars in Figure 5. Under this analysis, there really is no difference between the head-slaved and plane-slaved condition. Thus we can conclude that most, if not all, of the difference in performance between the head- and plane-slaved conditions was the result of an increased number of trials where the subject pilot *slightly* missed the 500 m threshold around the waypoint.

4. DISCUSSION

The results of this experiment show that virtual audio cues can be effective for providing georeferenced spatial information, and in particular navigational information, to general aviation pilots. To put the results in perspective, keep in mind that the pilots in this task had absolutely no information about where the waypoints were located other than what was provided to them by the virtual audio cue. They were just arbitrary points in the sky with no visual or other reference points to help identify them. Furthermore, they were generated with a prototype low-cost orientation tracking system that must have introduced at least a few degrees of error into the direction of the audio cues, and they were provided to subject pilots who had, at most, a few hours of experience listening to a virtual audio display. Yet the pilots were able use the audio cues to maneuver an aircraft flying at a speed of 70 m/s to pass within 350

m of these arbitrary points on average, and within 500 m of them nearly 90% of the time. We believe that this is truly a remarkable achievement, and one that clearly demonstrates the potential utility of audio displays for providing navigation information to pilots in GA aircraft.

However, the results of the two tracking conditions were a somewhat contrary to our initial expectations. Although the subject pilots performed well in both conditions, overall performance in the task was clearly superior in the plane-slaved condition. At first glance, this result seems inconsistent with the large body of literature that has shown that virtual audio localization performance is greatly improved by the addition of head-slaved dynamic localization cues, particularly in terms of the number of front-back confusions. However, this result is not so surprising when one considers the front-back confusion really has little or no effect in a navigation task like this one, where the pilots were instructed to respond to a directional audio cue by turning the aircraft in the direction of the sound source. In such a task, a pilot hearing a sound source in the right hemisphere will respond by turning the plane to the right, regardless of whether the sound is correctly localized at the two o'clock position or incorrectly localized at the four o'clock position. And as soon as the aircraft starts to turn, the true location of the sound source will become almost immediately apparent. Thus, it is not really surprising that the reduction in front-back confusions that normally occurs in a head-slaved virtual audio display did not generally translate into a significant improvement in navigation performance in this experiment. However, it should be noted that there is one special case where front-back cues do matter: the case where a new navigation waypoint occurs directly behind the aircraft. In this case, a front-back confusion would cause the pilot to assume the aircraft is already on the same correct course, and to fly directly in the opposite direction of the waypoint forever. This did in fact happen at one waypoint in the plane tracked condition of this experiment, and the pilot in that case flew far off course until the voice distance cues made it apparent that the aircraft was getting further away rather than closer to the waypoint. In any real-world implementation of a plane-slaved navigation system, it is clear that some kind of audio warning will have to be used to alert the pilot when a new navigation cue is located directly behind the aircraft.

This argument can explain why head-slaved performance was no better than plane-slaved performance in the navigation task, but it still cannot explain why head-slaved performance was substantially worse. The average distances of closest approach in the two conditions, shown in the right panel of Figure 4, were pretty similar, which suggests that poorer performance in the head-slaved condition was the result of the pilot very slightly missing the 500 m threshold required to hit the waypoint. This raises the question of whether there was a systematic pattern of errors that could explain the relatively poor performance observed in the head-slaved condition. Figure 6 shows an analysis of the misses in each condition as a function of the direction of the miss: i.e. whether the plane flew by with the waypoint to the right of the pilot or whether it flew by with the waypoint to the left of the pilot. In the plane-slaved condition, these two types of misses were almost evenly balanced (five versus six). But in the head-slaved condition, virtually all of the misses occurred with the waypoint to the right of the pilot. Thus, it seems that the poorer performance in the head-slaved condition was the result of a systematic tendency for the pilots to steer the aircraft to the left of the correct heading required to reach the waypoint.

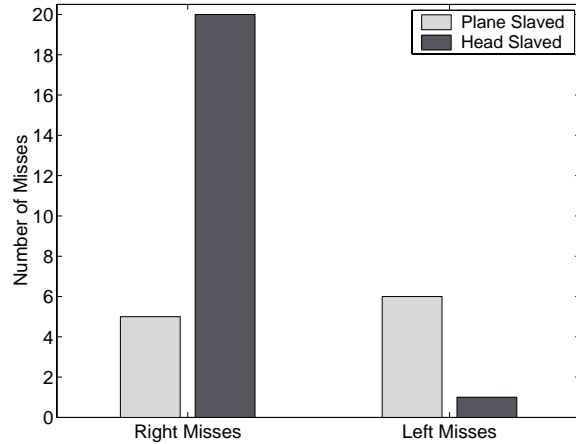


Figure 6: Analysis of misses in terms of the direction of the waypoint relative to the pilot when the plane flew past its point of closest approach.

If the head-slaved and plane-slaved conditions were conducted with different tracking devices, it would be easy to attribute this result to a bias in the position measurement of the head-slaved sensor. However, in this experiment the same tracking device was used in both conditions- the sensor was simply affixed to the aircraft rather than the headset in the plane-slaved conditions.

There is, however, one plausible explanation that could account for this result. One underlying assumption in the head-slaved condition is that the pilot's frame of reference in the cockpit is aligned with the centerline of the aircraft. In an aircraft with side-by-side seating, there may be a fundamental error with this assumption, especially in an aircraft like the SR-22 where the control panel is tilted slightly to wrap around the pilot in the left (command) seat. In such cases, it is quite possible that the pilot will perceive "straight ahead" not as the centerline of the aircraft, but rather as the line of sight from the left seating position to the hub of the propeller. In this experiment, such a shift in frame of reference would have caused the subject pilots to fly the aircraft slightly to the left of the desired heading in the head-slaved conditions, which would explain the waypoint passed by on the right side in 19 out of 20 of the missed trials. Frame of reference problems such as this one have largely been ignored in the auditory display literature, but it is clear that they could be critically important in all situations where an audio cue is used to provide information to an operator about the desired orientation of a vehicle rather than the relative location of an object.

5. SUMMARY AND CONCLUSIONS

In this experiment, we have demonstrated that virtual audio cues can be used quite effectively to provide navigation information to the pilots of general aviation aircraft. We have also shown that these spatialized navigation cues can be just as effective, or perhaps even more effective, in a display that spatializes the sound based on the heading of the aircraft rather than the direction of the listener's head. This is a very important finding, because it suggests that a functional auditory aircraft navigation display could be commercialized for a few hundred dollars rather than the few

thousand dollars it would cost to build a system with a headtracker. However, we should note that some caution should be used in extrapolating these results to other kinds of virtual audio cues that could be used in the cockpit. Front-back confusions generally don't matter in a task where the pilot will always steer the aircraft towards the next cue, but they do matter in cases where the pilot is not expected to maneuver the aircraft in response to the cue (as would be the case for almost any cue provided to enhance situational awareness) or when the direction of the maneuver might depend on the specific location of the cue (as might occur in a collision avoidance display). In these cases, there is reason to believe that head-slaved virtual audio cues will continue to be superior to those updated solely on the basis of the orientation of the aircraft.

6. REFERENCES

- [1] T.A. Furness, "The super cockpit and its human factors challenges," *Proceedings of the Human Factors Society 30th Annual Meeting.*, pp. 48-52, 1986.
- [2] R.L. McKinley, M.A. Ericson, and W.R. D'Angelo, "Three dimensional audio displays: development, applications, and performance," *Aviation, Space, and Environmental Medicine*, vol. 74, no. 9, pp. 937-46, 2003.
- [3] B.D. Simpson, D.S. Brungart, R. Dallman, J. Joffrion, and M. Presnar, "Spatial audio as a navigation aid and attitude indicator," 2005.
- [4] J.M. Joffrion, J.F. Raquet, and D.S. Brungart, "Boon: Head tracking for audio using a gps-aided mems imu," *Inside GNSS*, vol. 1, 206.
- [5] H. Wallach, "On sound localization," *Journal of the Acoustical Society of America*, vol. 10, pp. 270-274, 1939.
- [6] F.L. Wightman and D.J. Kistler, "Resolution of front-back ambiguity in spatial hearing by listener and source movement," *Journal of the Acoustical Society of America*, vol. 105, pp. 2841-2853, 1999.
- [7] E.M. Wenzel, M. Arruda, D.J. Kistler, and F.L. Wightman, "Localization using non-individualized head-related transfer functions," *Journal of the Acoustical Society of America*, vol. 94, pp. 111-123, 1993.
- [8] J.D. Miller, "Slab: a software-based real-time virtual acoustic environment rendering system," in *Proceedings of the International Conference on Auditory Display (ICAD 2001)*, Espoo, Finland, July 29-August 1, 2001.
- [9] D.S. Brungart and K.R. Scott, "The effects of production and presentation level on auditory distance perception of speech," *Journal of the Acoustical Society of America*, vol. 110, pp. 425-440, 2001.