AUDITORY LOCALIZATION CUE SYNTHESIS AND HUMAN PERFORMANCE

Mark A. Ericson 1st Lt, USAF
Armstrong Aerospace Medical Research Laboratory (AAMRL/BBA)
Wright-Patterson Air Force Base, OH

Richard L. McKinley
Armstrong Aerospace Medical Research Laboratory (AAMRL/BBA)
Wright-Patterson Air Force Base, OH

ABSTRACT

The Armstrong Aerospace Medical Research Laboratory Biological Acoustics Branch (AAMRL/BBA) is developing directional audio technology for headphone applications. A laboratory demonstration model of a real-time, digital auditory localization cue synthesizer is integrated with a head position tracker, A/D and D/A converters, and a host computer to form a directional audio display (DIRAD). The DIRAD presents sounds over headphones that are perceived to be outside of the listener's head and relatively easy to localize. Possible military applications of directional audio technology include integration with radar warning receivers, voice communication systems and collision avoidance systems. Natural audio presentation of critical information may improve situational awareness and reduce pilot workload.

The auditory localization cue synthesizer board is a modular unit that requires analog and digital interfacing, a host PC with a serial (RS-232) data port, a monaural audio input, a Polhemus digitizing position tracker and a pair of headphones. The synthesizer design is based on advanced digital signal processing technology. Two Texas Instruments TMS-320C25 processors provide the localization cue synthesis while a TMS-32020 processor controls the input-output functions.

Performance of the DIRAD was measured using a panel of ten subjects. Five male and five female subjects responded to randomly positioned stimuli consisting of wide band pink noise, male speech bandlimited to 3.5 KHz, female speech bandlimited to 6 KHz and octave band noises from 125 Hz to 8 KHz. The listeners were instructed to respond by turning their head and body to face the direction of the sound source. Data was collected for 72 different positions and the overall performance were a mean magnitude error of 4.8 degrees and a mean response time of 4.4 seconds.

INTRODUCTION

Flying a high performance aircraft is a highly complex task, especially in combat situations. In addition to controlling the aircraft, the fighter pilot maintains communication with his wingman, initiates attacks on enemy targets and hopefully avoids being hit. The HUD, communications, radar warning receiver, master caution panel and other inputs can often overload the fighter pilot with visual and auditory information.

For example, a radar warning receiver provides a warning tone over the headset and the missile's azimuth location on a CRT screen. After the pilot recognizes the audio warning for an incoming missile, he redirects his attention to the radar warning receiver CRT. He reads the missile's direction in azimuth from the 2-D CRT and translates the direction onto the 3-D frame of reference of the aircraft. He visually acquires the missile plume and begins evasive maneuvers, time permitting.

One possible improvement to the existing radar warning receivers could be to implement a DIRAD in line with the present equipment. A DIRAD would present that same radar warning tone over headphones, only now the tone would sound like it was coming from the direction of the approaching missile. The pilot would determine the direction of the missile almost instantaneously without ever having brought his eyes into the cockpit. The time to visually acquire an object would be greatly reduced by the enhancement of directional audio information. This has been demonstrated in laboratory experiments (Perrot, 1988). In Perrot's experiment, subjects were able to locate visual targets in azimuth outside of their initial field of view 500-1000 ms faster when aided with a correlated directional auditory cue.

DIRAD has the capability of providing directional information to audio signals over headphones. For example, a pilot currently monitors his wingman's communication's, ground-to-air communications, airborne command
post communication, TACAN, etc., all mixed together in his headset. DIRAD can improve the pilot's ability to simultaneously monitor several different communication sources by spatially separating the sounds as at a cocktail party. Collision avoidance and navigation aid applications of a DIRAD may also be possible. A DIRAD would enhance the auditory information displayed to the pilot, resulting in improved pilot situational awareness and reduced target acquisition times.

The Biological Acoustics Branch has established an inhouse facility to research human performance using directional audio displays and to develop techniques for presentation of directional audio information over headphones. The facility has been used to collect directional cue data, develop a laboratory demonstration model DIRAD and measure the moving head localization performance of humans using natural and synthesized cues.

The current study was designed to measure the speed and accuracy of humans in localizing auditory signals in azimuth without visual cues. The experimental paradigm was the same for the control and test conditions, i.e., natural and synthesized cues. The auditory stimulus was presented from 24 random directions in each session. Each session was repeated 3 times for a total of 72 locations. The locations were chosen to distribute the stimuli evenly around the circle by picking 4 locations within each of six 60° sectors (see figure 1). Speed and accuracy means were calculated and recorded for each trial.

OBJECTIVES

The objectives of the experiments were to measure human auditory localization performance for natural and synthesized cues, and by comparing the two data sets, determine the feasibility of providing directional audio information over headphones. A goal was established to achieve human auditory localization performance of less than 30° error with synthesized cues from the DIRAD.

APPROACH

A facility was built to collect directional cue data and develop a laboratory demonstration model DIRAD. The presentation of directional audio information over headphones was based on two directional cues: the head related transfer function (HRTF) and the interaural time delay (ITD). HRTF and ITD cues were collected at 1° spacings in azimuth. Digital FIR filters of the 360 HRTF's were generated, and the HRTF and ITD codes were implemented in look-up tables on the DIRAD's TMS320C25 processors. These two cues enabled the listener to localize sounds over headphones, i.e., determine the direction of the sound source.

The DIRAD was developed in-house as a laboratory demonstration model. The localization cue synthesizer, analog and digital interface components were wire-wrapped to allow ease of modifications. The breadboards included off-the-shelf A/D's, D/A's, DSP and TTL chips. Software for the DSP processors provided head coupled, directional cue presentation in azimuth. The DIRAD was incorporated into the localization performance measurement facility.

The facility was used to measure human auditory localization performance of directional cues presented in a natural environment and over headphones. A ring of loudspeakers presented the natural directional cues. Either a single loudspeaker was used to provide a real source or two loudspeakers were phased together to create "phantom" sources at 1° increments. The DIRAD provided the synthesized directional cues over headphones. A Polhemus head tracker measured the subject's head position response for both conditions. Response time was measured as the duration from stimulus presentation until the subject indicated the source was directly in front of him/her.

Human auditory localization performance was collected using the facility. Subjects from the general population participated in the study. The subject's task was to respond to the presentation of a directional auditory stimulus by facing the perceived direction of the sound as quickly and as accurately as
possible. The performance criteria were the mean magnitude error, mean direction error and mean response time. A statistical analysis of the data was done to compare the two conditions and, thereby, determine the feasibility of presenting directional auditory information over headphones.

APPARATUS

An anechoic chamber and the control room adjacent to the chamber house all of the equipment used in the electro-acoustic measurements. The anechoic chamber is part of the AAMRL Biodynamics and Bioengineering Division facilities. All six interior surfaces are covered with 4 feet deep sound absorbing wedges that provide 60 to 120 dB attenuation from 50 to 10,000 Hz. The setup for the natural cue localization performance measurements is shown in figure 2.

The natural cues are displayed to the subjects via a ring of loudspeakers. The loudspeaker ring equally positions twenty-four 4.5" diameter loudspeakers around a 7' radius circle at the ear level of a listener. The structure is constructed of 4" dia. aluminum rods to minimize weight and reflective surfaces.

The DIRAD is composed of the localization cue synthesizer board, digital interface board, analog interface board, head position tracker, host computer and headphones (see figure 3). The synthesizer board connects with the digital interface board (DIB) and analog interface board to communicate with the peripheral digital and analog equipment, respectively. The host PC communicates with the DIB over an RS-232 serial data bus and the head tracker provides the real head angle information to the digital interface board via a 16 bit parallel bus. The digital board sends relative head angle indexes over two separate 16 bit parallel I/O ports to the right and left processors on the synthesizer board. The processors lookup the corresponding FIR filter coefficients and encode the audio signal with the directional cues. The analog interface board performs the A/D and D/A conversions for the synthesizer board. The right and left side audio signals from the analog interface board are amplified over a standard pair of headphones. The DIRAD provided the synthesized cues in the human performance experiments.

The DIRAD meets the following design and performance criteria. The DIRAD operates in real time, provides auditory direction information relative to the listener's head position and provides auditory cues of 1° resolution in azimuth. The processors are clocked at 40 MHz in order to calculate a 179 tap FIR filter every 1 ms. The audio input signal is sampled at 40 KHz to provide a 10 KHz bandwidth at the output of the analog interface board.

SUBJECTS

Ten subjects (five males and five females) were selected from a contractor maintained pool (Systems Research Laboratory) of volunteer subjects. These subjects were monetarily compensated for their participation. Each subject underwent an audiometric hearing test prior to participation in the study. The hearing threshold levels of the subjects were not greater than 15 dB at any frequency of a standard audiometric test from 500 to 8000 Hz (Air Force Regulation 160-43, Attachment 4, June, 1987).

RESULTS

Human performance was measured by the following criteria: (1) mean magnitude error (MME), (2) mean direction error (MDE), and (3) mean response time (MRT). The means of 72 trials were calculated for each stimulus and subject. The MME was defined as the absolute measure of the average difference in degrees between the actual location of the stimulus and the subject's perceived location of the stimulus. The MDE was defined as the arithmetic measure of the average difference, right or left, between the actual location of the stimulus and the subject's perceived location of the stimulus. The MDE indicates the directional bias of the listener by averaging the right and left direction errors for 72 stimuli. For example, a +7° error averaged with a -3° error results in a MDE of +2° (to the right). The MST was defined as the time from stimulus onset to the subject's response indicating the perceived location of the stimulus.

The MME's of the subject's were 6.0° (SD=1.3) for the natural (control) condition and 4.8° (SD=1.6) for the DIRAD condition. Individual subject means ranged from 2.3 to 13.0 degrees. The MME in sector 1 (330-29°) was 4.26°, significantly lower (P=.02) than the other five sectors (range 5.68 - 7.21°) in the natural condition. The 4 KHz stimulus in this sector was especially difficult for subjects to localize due to a lack of directional cues in that frequency range. Humans localize sounds in the 4 KHz octave band with fewer directional cues than at other frequencies. ITD cues above 1.5 KHz are not effective for localization by the human hearing system. Pinnae cues are most effective in localizing sounds above 8 KHz. The MME as a function of condition and stimulus is shown in figure 4.

The calculated MDE's of the subjects were +1.2° (SD=2.8) in the natural condition and +1.0° (SD=2.3) with the DIRAD. Individual
subject means ranged from -12.1 to +10.2 degrees. These data show that the subjects responded more precisely with the DIRAD (1") than with natural cues (1.3"). Overall, the subjects responded to the right side of the source location slightly more than the left side. MDE as a function of condition and stimulus is shown in figure 5.

The MRT's were 3.5 seconds (SD=1.1) in the control condition and 4.3 seconds (SD=1.6) with the DIRAD. Individual mean response times ranged from 1.7 to 7.7 seconds. The subjects required 0.8 seconds more time with the DIRAD to acquire the direction of the stimulus. The longer response time observed with the DIRAD may have been due to the head tracker's relatively slow 54 Hz update rate, because laboratory measurements have shown that humans can move their heads at a much greater rate (more than 200 degrees per second). The MRT as a function of condition and stimulus is shown in figure 6.

The multiple comparison tests (Tukey and Bonferroni) of the sectors for mean response times indicate no significant differences (MSD = .25 and .26, respectively) between the MRT's of sectors 3 (90-149") and 5 (210-269") and between sectors 2 (30-89") and 6 (270-329"). All other pairwise comparisons for sector showed significant differences with the highest MRT for sector 3 (150-209", mean = 4.8 seconds) and the lowest for sector 1 (330-29", mean = 3.4 seconds). These data followed expectations since it took longer for the subject to move towards the right side or halfway around (see figure 7). A summary of MME, MDE and MRT for both conditions is shown in Table 1.

**TABLE 1**

<table>
<thead>
<tr>
<th>Condition</th>
<th>MME (SD)</th>
<th>MDE (SD)</th>
<th>MRT (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATURAL</td>
<td>6.03 (1.3)</td>
<td>1.31 (2.8)</td>
<td>3.51 (1.1)</td>
</tr>
<tr>
<td>DIRAD</td>
<td>4.84 (1.6)</td>
<td>0.98 (2.3)</td>
<td>4.34 (1.6)</td>
</tr>
</tbody>
</table>

**DISCUSSION**

The data show that humans perform equally well with the synthesized cues as well as with the natural cues. Accuracy with both cue varied among listeners. This psycho-acoustic phenomena may be due to a lack of visual cues accompanying the auditory source during the period that the stimulus was in the listener's field of view.

This study demonstrates that synthesized directional cues can be authentically presented over headphones. Now that the feasibility of simulated auditory localization has been demonstrated, much interest is generated in the possible applications of this technology.

Synthetic localization cues can be used to provide localization perception in environments where they do not occur naturally. Humans are three dimensional creatures and are sometimes forced to operate with one or two dimensional cues. The addition of three dimensional sound in some areas could enhance operator awareness and task performance.

Synthetic localization cues could dramatically increase the effectiveness of systems in aircraft such as the radar warning receivers. By providing directional audio cues of a missile's location relative to the aircraft, the pilot would interpret the information as he would in a natural environment. Experimental data needs to be collected to measure human performance with the present system and with the integration of a DIRAD.

Directional audio information from a DIRAD could have other applications in the cockpit. A DIRAD could present voice communications from a wingman, other aircraft or ground stations to appear from the direction relative to the aircraft. The pilot could thereby monitor several spatially separated audio channels simultaneously. Head-on collision avoidance application of a DIRAD may also be possible. A DIRAD could present any audio signal to appear from any direction in azimuth and utilize another modality of information input to the pilot.

Besides in a cockpit, the DIRAD may have applications in the presentation of multiple near field sources at a C post. Unlike monaural audio signals, directional audio signals are spatially separated and can be discerned among a group of signals. A DIRAD could enhance the C operator's discrimination of messages among multiple simultaneous sources.

In the localization performance experiments, most of the subjects perceived the auditory stimulus from the DIRAD to have good spatial qualities. However, the sound source location appeared to be elevated while in front of the listener for both the natural and synthesized cues. The degree of elevation varied among listeners. This psycho-acoustic phenomena may be due to a lack of visual cues accompanying the auditory source during the period that the stimulus was in the listener's field of view.

A simple experiment could be done to test this hypothesis. In this experiment, a visual cue would be coupled to the auditory stimulus'
position on the horizon. The subject would be instructed to turn and face the source location with their eyes open. If the subjects perceived the stimulus to remain on the horizon, then the hypothesis would be correct. If the source still appeared to elevate in front, then other hypotheses would have to be formulated and tested.

PLANS

The present breadboard version will be fabricated onto a PC board before flight simulator testing is initiated. Simulator testing will provide data on the feasibility of incorporating a DIRAD with radar warning receivers.

The AAMRL/BBA is continuing research and development of directional audio technology for military applications. The addition of elevation information on the synthesizer board is nearly complete. Experiments are planned to measure moving head and fixed head auditory localization performance in elevation, dynamic tracking and multiple source localization.

SUMMARY

The performance experiments demonstrated the ability of humans to localize sounds over headphones. Overall mean magnitude errors of 6.0 degrees were measured in the control condition and 4.8 degrees with the DIRAD. Mean response times of 3.5 seconds in the control condition and 4.3 with the DIRAD were measured. The performance experiments showed that the DIRAD could consistently point humans to well within 20° of azimuth using auditory localization cues over headphones. The DIRAD performance indicates the application of directional audio information to cockpit radar warning receivers may be feasible.
FIGURE 3

HUMAN AUDITORY LOCALIZATION PERFORMANCE
MEAN MAGNITUDE ERROR BY STIMULUS
NATURAL  DIRAD

FIGURE 4
HUMAN AUDITORY LOCALIZATION PERFORMANCE
MEAN DIRECTION ERROR BY STIMULUS

![Graph showing mean direction error by stimulus.](image)

HUMAN AUDITORY LOCALIZATION PERFORMANCE
MEAN RESPONSE TIME BY STIMULUS

![Graph showing mean response time by stimulus.](image)
HUMAN AUDITORY LOCALIZATION PERFORMANCE
MEAN RESPONSE TIME BY SECTOR

NATURAL

DIKAD

FIGURE 7