Proceedings of the Ninth International Symposium on Aviation Psychology The Ohio State University, Columbus Ohio (1997)

AUGMENTED TCAS ADVISORIES USING A 3-D AUDIO GUIDANCE SYSTEM

Durand R. Begault, Ph.D.†
Elizabeth M. Wenzel, Ph.D.
William B. Lathrop†
†San Jose State University
Flight Management and Human Factors Division
NASA Ames Research Center
Mail Stop MS 262-2
Moffett Field, CA 94035-1000
db@eos.arc.nasa.gov

ABSTRACT

The use of spatialized ("3-D") audio techniques for head-up auditory displays in commercial aircraft has been investigated at NASA Ames Research Center since 1989. Our previous aeronautic application studies have shown a significant improvement in target acquisition time when a 3-D audio system was used for aurally guided search (from 0.5 to 2.2 sec faster, depending on experimental condition). This paper reports on a full-mission simulation using 10 crews in a within-subjects study at the NASA Ames Crew-Vehicle Systems Research Facility (CVSRF) Advanced Concepts Flight Simulator. We studied whether the existing Traffic alert and Collision Avoidance System (TCAS) could be improved by using 3-D sound as an "added feature," as opposed to our previous investigations where 3-D audio was used in lieu of a visual map display of traffic. The verbal alert for the TCAS advisory was presented through a stereo headset so that the perceived lateral spatial location of the sound corresponded to the location of the trafficout-the-window. Scenario software generated 49 targets corresponding to aircraft along a SFO-LAX flight route; each crew ran this route twice, once with and once without the addition of a 3-D audio cue to process the TCAS advisory. Across all ten crews, the results showed a significant differencein target acquisition time between the two conditions, favoring the 3-D audio TCAS condition by 307 msec; there was no significant differencein the number of targets that were acquired. Six of the ten crews had shorter acquisition times using 3-D audio (ranging from 225–749 msec) while one crew showed a shorter acquisition time without 3-D audio (420 msec). Questionnaire data showed a favorableattitude from most crews towards augmenting the current TCAS interface with a 3-D audio system.

BACKGROUND: AURALLY-GUIDED VISUAL SEARCH

Commercial aircraft in the US are equipped with Traffic alert and Collision Avoidance System (TCAS) equipment that allow other similarly equipped aircraftto see their location on a radar map display. The system indicates four warning levels for situational awareness, ordered in terms of decreasing time-until-impact: "other traffic,""proximate traffic,""traffic advisories," and "resolution advisories." The resolution advisory also gives instructions for evading an accident, such as "climb...climb." The concern of this study is focused on the TCAS advisory only. Currently, a TCAS advisory warning activates a verbal announcement over the cockpit loudspeaker, a loud "traffic....traffic." The pilot determines the location of the aircraftin question both by looking out the window and referring to a visual radar display map.

The spatial audio processing of an alert signal results in a two-channel binaural signal ("3-D audio") for a pre-determined, virtual location (see Begault, 1994; Wenzel, 1992 for overviews). 3-D audio techniques impose the interaural spectral, intensity, and time differencesof head-related transfer functions (HRTFs), the primary cues for auditory localization. With the use of a two-channel warning system and stereo 'on-the-ear' headsets, the perceived auditory spatial location of the "traffic....traffic" alert can be placed at positions corresponding to the relative azimuth and elevation of out-the-window traffic(see Figure 1). The goal of the current experiment was to reduce the time needed to acquire the trafficby use of "aurally-guided visual search," effectivelyby reducing the head-down time necessary to find the traffic on the TCAS map display.

Aurally-guided visual search using a 3-D audio warning signal can be thought of as a type of "head-up auditory display." The advantages are similar to a visual head-up display in that the pilot can acquire information while keeping the eyes out the window. Additionally, there is a reduction in the time needed to acquire information, corresponding to the extra time required with a visual display in moving the head downwards and then back up.

Two experiments were previously conducted at NASA Ames Research Center to determine the advantage of 3-D audio techniques for acquiring out-the-window traffic. In the first study (Begault, 1993), a between-subjects design was used, with six of twelve crews hearing one-ear (monotic) playback of the TCAS advisory, and the other six receiving spatialized cues over stereo headphones. No TCAS map display was supplied for either group. Only seven positions were used for azimuth cueing (corresponding to clock positions from 9 to 3 o'clock), all at eye level elevation. In addition, the spatialized audio stimuli were exaggerated in relationship to the visual stimuli by a factor of up to 2 (e.g., visual targets at 15 degrees azimuth would correspond to spatialized stimuli at 30 degrees azimuth). This was done to maximize the range of interaural time differences,thereby causing the sensation of leftward stimuli to be maximally differentiatedfrom rightward stimuli. Results of the study found a significant reduction in visual acquisition time when spatialized sound was used to guide head direction (4.7 versus 2.5 sec).

In the second TCAS experiment (Begault & Pittman, 1996), a between-subjects design was also used, with half of a total of ten crews using a head-down visual map display with standard TCAS symbology and overhead speaker annunciation, and five crews using 3-D audio TCAS presentation with no map display. Results showed a significant differencein target acquisition time between the two conditions, favoring the 3-D audio TCAS condition by 500 msec (2.6 versus 2.1 sec). No exaggeration factor between audio and visual stimuli was used, in order to determine if a more literal cueing between interaural time differencesand visual position would be effective. Another difference between the first and second experiment was that three categories of HRTF elevation cues were provided— "eye level," "above" (alert spatialized to 36° above eye level), and "below" (36° below eye level). Multivariate data analysis and post-experiment interviews from this second experiment suggested that elevation cues were not effectivelyutilized by the pilots; a verbal cue would probably have been better (e.g., "traffic, high"). Successful perception of elevation cues from HRTFs tends to be more difficult than azimuth or externalization cues (Wenzel, Arruda, Kistler & Wightman, 1993). This is because cues to azimuthal location involve interaural level and time differences, while elevation is only cued by spectral differences

STIMULI AND SUBJECTS

The experiments took place at NASA Ames' Crew-Vehicle Systems Research Facility, under full-mission simulation conditions, within a generic "glass cockpit" flight simulator. Each of the 10 flight crews evaluated consisted of current captains and first officers from the same company and aircraft type. Scenario software generated 49 visual traffic stimuli ("targets") along a flight path from San Francisco to Los Angeles, primarily during the cruise portion of the flight (see Figure 2). This resulted in a target occurring approximately once every 40-90 sec; crews were informed in that the number of targets encountered would be artificially large. A within-subjects design was used, where each crew flew the same SFO-LAX run twice with the same pattern of targets, once with a normal TCAS display and once with 3-D audio processing of the verbal "traffic.....traffic" alert. The ordering of these two runs was arranged so that five crews flew the 3-D audio run first, and the other five crews flew it second. Crews were not informed in advance that the same targets would appear on each SFO-LAX run. Otherwise, they were briefed about normal procedures, departure and arrival routes. About four hours of simulator training occurred in the morning, and the two experimental runs occurred in the afternoon. Runs were terminated shortly after landing at LAX.

The simulator used in this experiment had a 3-channel, 4-screen display that limited the available out-the-window field-of-viewfor each pilot. Figure 3 shows the available field-of-viewfor the Captain; the First Officer's view would be the mirror image of this figure. The center channel screen enables each pilot a field-of-viewextending to approximately ±25° azimuth. In addition, 2 side screens fed by the other 2 channels gave each pilot a unique side field-of-view that extended the total field-of-view to approximately ±52° azimuth. The immediate range of the vertical field-of-view is from approximately -13° to +16°, but can extend from -18° to +20° with head and body adjustments.

The out-the-window positions of the targets patterned a matrix of 7 azimuths by 3 elevations. The azimuth distribution of the 49 targets was as follows: 66% at ±48°; 11% at ±35°; 11% at ±22°, and 11% at 0°. The elevation distribution of targets was: 33% at 4,000 feet above own ship (about 15° elevation); 33% at 2,500 feet below own ship; 16.5% at 1,000 feet above own ship; and 16.5% at 1,000 feet below own ship. For both conditions, a computer generated multiple moving symbols depicting aircraft on the TCAS map display. The symbols appeared at pseudo-random positions, and eventually one would be elevated to advisory status for target acquisition evaluation, while the remaining symbols would eventually vector off the display. The corresponding 3-D audio alert used for each azimuth is shown at the top of Figure 3.

The simulation was computed to have 20 mile night time visibility. All trafficstimuli ("targets") were set at a distance of 3 miles. The positions of all targets were designed to be held at a constant distance and incidence angle, independent of the movement of the pilots' aircraft. This made the target appear as a flashing dot of light similar to that seen out the cockpit window of a real aircraft. Overall, this procedure eliminated any variability between crews as a function of aircraft movement. The alternative of modeling targets with normal ballistic trajectories would have been unacceptable because the angle of incidence and size of the target would have varied for each crew. The dependent variables were: (1) the time interval between the appearance of a visual target in conjunction with an aural advisory and the verbal response from a crew member indicating acquisition of the target; and (2) the number of targets acquired. The crew members were instructed to call out verbally when they had visually acquired the aircraft outside the window (a consistent utterance, such as "got it!"). Acquisition time (the differencebetween the time the visual target was generated and the beginning of the verbal utterance) was derived from video tapes by an observer who was not aware of the design of the experiment. Each verbal acquisition increased the count for the number of targets acquired. The acquisition time was determined by the time code generated on the video tapes (accuracy within 1 video frame). A target was considered "missed" if it was not acquired within a 10 sec window. In addition, a questionnaire was given at the end of the experiment to determine each pilot's preferencefor the audio display system. Answers were indicated on a Likert scale, "strongly disagree–disagree–neither disagree or agree–agree–strongly agree." The questions for which statistically significant consensus was obtained are described in the results section. A hearing questionnaire and an audiogram was also given to the pilots.

RESULTS

An analysis of variance revealed no significant differencein the number of targets acquired between conditions across all crews. Figure 4 shows the equivalent result for acquisition time. A significant difference ($P = .03$) of 307 msec was obtained: The mean acquisition time for the standard TCAS ("map only") condition was 2.2 sec (SD = 0.93), but acquisition time was decreased to 1.8 sec (SD = 0.71) when the map was augmented with 3-D audio. Six of the ten crews had shorter acquisition times using 3-D audio (ranging from 225–749 msec) while one crew showed shorter acquisition time without 3-D audio (420 msec). The differencein acquisition time for the remaining three crews was insignificant. A chi-square test performedon the questionnaire data showed statistical significance ($p = .01$) for the "agree" response to the following questions: "The 3-D audio traffic advisories were useful in helping me locate traffic out of the window", and "The 3-D audio alert might be useful for TCAS RAs as well." Additional data from the questionnaires and audiograms will be reported in a future publication.

DISCUSSION

The advantage of aurally guided visual search is in line with previous results where stimuli presentation occurred under more controlled conditions (Perrott, Sadralodabai, Saberi & Strybel, 1991). Although 307 msec is a modest improvement, it does suggest that, in an operational setting, an aural 3-D TCAS display may be desirable in addition to a standard TCAS display. It should be pointed out that the experimental conditions were optimized for ease of target acquisition, due to the lack of surfacelights on the ground and the relatively low workload present during the phases of the flight that were evaluated. A greater difference might be found between the two conditions if workload were higher, and/or if the visual field-of-view contained "noise" from multiple aircraftor city lights. The 3-D audio technology should be inexpensive to implement, and the integration of 2-channel, lightweight noise-canceling headsets into the cockpit should offer a hearing conservation benefit to the aviation community.

ACKNOWLEDGMENTS

Many thanks are due to the team at the Crew-Vehicle Systems Research Facility (CVSRF). Support for this research was in part provided by a grant from NASA Ames Research Center to the San José State University Foundation (NCC-2-327). Thanks also to Joel Miller and Rick Shrum of Sterling Software.

REFERENCES

Begault, D. R. (1993). Head-up Auditory Displays for Traffic Collision Avoidance System Advisories: A Preliminary Investigation. Human Factors, 35(4), 707-717.

Begault, D. R. (1994). 3-D Sound for Virtual Reality and Multimedia. Cambridge, MA.: Academic Press Professional.

Begault, D. R., & Pittman, M. T. (1996). Three-dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays. International Journal of Aviation Psychology, 6(1), 79-93.

Perrott, D. R., Sadralodabai, T., Saberi, K., & Strybel, T. Z. (1991). Aurally aided visual search in the central visual field: effects of visual load and visual enhancement of the target. Human Factors, 33(4), 389-400.

Wenzel, E. M. (1992). Localization in virtual acoustic displays. Presence: Teleoperators and Virtual Environments, 1, 80-107.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using non-individualized head-related transfer functions. Journal of the Acoustical Society of America, 94(1), 111-123.

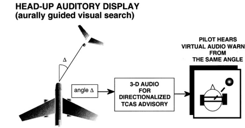

FIGURE 1. The use of 3-D audio as a head-up auditory display for aurally-guided visual search.

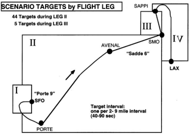

FIGURE 2. The route used during the experiment. Most of the targets occurred under low workload conditions, with a minimum of on-ground visual distractions from lights.

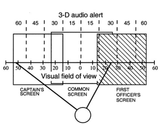

FIGURE 3. The horizontal field-of-view in the simulator, from the perspective of the left seat (Captain's position). The numbers within the dashed lines show the mapping between visual azimuths and the specific azimuth position of the 3-D sound cue that was used for the alert

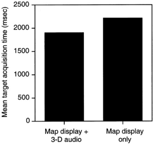

FIGURE 4. Mean acquisition times for 464 targets acquired under the map display (2.20 sec; *SD=.93, SE=.04)* and the map plus 3-D audio (1.89 sec; *SD=.71, SE=.03)* conditions, across 10 crews. A significant difference*(p = .03)* in target acquisition time was found, favoring the 3-D audio TCAS condition by 307 ms.