APPLICATIONS OF VIRTUAL AUDIO

Mark Ericson, William D’Angelo, Eric Scarborough, Steve Rogers, Philip Amburn, and Dennis Ruck

1 Armstrong Laboratory
2 Air Force Institute of Technology
Wright-Patterson Air Force Base, OH

ABSTRACT

Technology for electronically simulating spatial sound over loudspeakers and headphones has matured in the past few decades to facilitate many new applications of virtual audio. Although binaural sound reproduction has been around since 1881 (Sunier, 1986), applications were limited by the physical apparatus needed to simulate a particular listening environment. The "dummy head" method developed in the 1930s required a manikin to be instrumented with microphones near the entrances to the ear canals and immersed in a desired listening environment, such as a concert hall. The sound fields at the manikin’s ears were displayed either over headphones or loudspeakers to a remote listener. Some new technologies, which include electro-magnetic head trackers and digital signal processors, have made it possible to synthesize spatial sound without physical system requirements, such as the "dummy head" method. These advancements have enabled scientists and engineers to better observe, analyze, and synthesize spatial hearing cues. Electronic simulation of directional and distance auditory cues has greatly expanded the areas of application of virtual audio. Some potential aerospace applications include monitoring spatially separated speech communication signals to increase understanding, navigating by an auditory beacon, and acquiring visual targets with the aid of directional audio signals. Potential non-aerospace applications include navigational aids for the blind and enhanced virtual reality simulation for entertainment and education. In the future, more applications may use virtual audio technology to display spatial auditory information.

INTRODUCTION

Auditory localization is the perception of a sound source originating from a particular location in space. Auditory localization cues are a combination of directional and distance cues. The cues attributed to directional hearing are found almost entirely by head related transfer functions (HRTFs) correlated with head motion. The HRTF is the difference between a sound field from an auditory event in free space and the sound field from the same auditory event at the entrance to the ear canal. HRTFs contain time of arrival and frequency information that correspond to a particular angle of incidence from an auditory event to the listener. Distance cues as described by Coleman (1963) and Blauert (1983) include (1) intensity, (2) atmospheric absorption, (3) head motion, (4) reverberation, (5) binaural time differences, (6) binaural spectral differences, and (7) reflections. Reflections as found with the precedence effect have a biasing effect on the sound source's perceived direction and distance (Hafer, 1992). The amount of bias toward the echoes (4-5 degrees) appears to be independent of the source-echo separation (8-40 degrees) and delay (1-30 ms) of the reflection. Other cues such as bone conduction and
prior knowledge of the sound source have an undetermined effect on one’s distance perception. Virtual audio is the perception of being immersed in a listening environment different from the actual one in which a listener is physically located. The perception is achieved by synthesizing spatial hearing cues, such as HRTFs, reflections, and reverberation, and encoding them onto a given waveform. The goal of virtual audio technology is to create the illusion that a listener is in a particular acoustic environment. The previously described direction, distance, and possibly other cues are needed to create the illusion.

Recently, several digital electronic systems have been developed to synthetically generate auditory localization cues. When acoustic signals are digitally encoded with these cues, the resulting binaural sound is sometimes called "3-D sound", "3-D audio" or "virtual audio". The main advantages of virtual audio compared to previous methods are the independence from a physical setup and a sense of presence created by the interaction between the listener and the sound source within a virtual acoustic environment. Another benefit of virtual audio is that sounds can be created electronically that are either difficult or impossible to make with a physical system. Many researchers use virtual audio technology to investigate the effects of spatial hearing cues on auditory perception and to explore the processes of the auditory system (Durlach, 1991). Control over the generation of directional and distance auditory cues makes many applications possible. Contained in this paper are three sections on virtual audio applications. The first section is a review of techniques to physically reproduce spatial sound. The second section discusses current virtual audio applications. The third section describes possible future applications. In each section, loudspeaker and headphone systems are described as used in aerospace, entertainment, and research applications.

OBJECTIVE/PURPOSE

The objective of this paper is to describe past, present, and future applications of virtual audio technology. The purpose is to provide a broad perspective of virtual audio for someone who is new to the field and is interested in using the technology in a practical application.

PHYSICAL SPATIAL SOUND REPRODUCTION

LOUDSPEAKERS (Aerospace)

3-D sound sources have been physically generated over loudspeaker arrays to evaluate the potential of virtual audio in aerospace applications. Folds (1990) studied the applicability of using directional audio in a cockpit simulator. Directional audio was found to be beneficial when a person was required to attend to more than three simultaneous spatial tasks. In a later study by Folds (1993), a benefit of using directional audio was found when four to six simultaneous directional tasks required attention. Loudspeakers are an ideal means to present virtual audio because of their small size and high signal separation characteristics.

LOUDSPEAKERS (Non-Aerospace)

Some researchers have used an array of loudspeakers to simulate spatial sound (Blauert (1983) and Sakamoto (1982)). Spatial sound, as with localized sound, consists of directional and distance cues as generated in the listening environment. Applications of this research include the reproduction of concert hall effects for listening pleasure. Blauert (1983) found that at least twenty individual sound sources are necessary to accurately simulate a typical concert hall environment.

HEADPHONES (Aerospace)

The first application of binaural sound in the military was for locating artillery targets and trench digging operations during World War I. In both cases, the interaural time delay of the signals (air and ground) were amplified to increase the directional acuity of the listener to provide super-localization.
abilities (Durlach, 1991). The artillery locating device was called a pseudophone, and similarly, the ground device was called a geophone. However, the invention of RADAR made the use of pseudophones obsolete. In World War II, hydrophones were installed on submarines to determine the location of submarines. Mudd and McCormick (1960) and Ericson (1991) have investigated the possibilities of using virtual audio in cockpit applications. Virtual audio was evaluated as an aid to help alleviate workload on the visual modality. A directional stimulus was presented aurally, visually, and simultaneously in a target acquisition task. Response times were found to be similar in all three conditions.

Recently, the dummy head method was implemented in a tele-operated vehicle (TOV) to improve the remote human operator's situational awareness (Smith, 1993). The TOV was equipped with a dummy head and microphones at the location that a human operator would normally be located. The dummy head was rotated with a hydraulic actuator slaved to the head of the human operator to provide head motion cues. Although spatial sound was successfully presented to the human operator, unwanted noise and slight lag times of the systems degraded the operator's feeling of presence in the TOV.

HEADPHONES (Non-Aerospace)

Binaural sound was first reproduced over headphones in 1881 at the Paris world fair by Clement Ader. A pair of microphones were separated by a baffle to simulate the mass of a head. Signals from the two microphones were transmitted over a pair of headphones to provide binaural sound to a remote listener. Although spatial sound was successfully presented to the human operator, unwanted noise and slight lag times of the systems degraded the operator's feeling of presence in the TOV.

HEADPHONES (Non-Aerospace)

Binaural sound was first reproduced over headphones in 1881 at the Paris world fair by Clement Ader. A pair of microphones were separated by a baffle to simulate the mass of a head. Signals from the two microphones were transmitted over a pair of headphones to provide binaural sound to a remote listener. The binaural sound technique was first applied to radio broadcasting of concert hall performances by the Berlin Opera house in Germany and by a radio station WPAJ, in New Haven, Connecticut in 1925. WPAJ transmitted the left signal over one radio at one frequency and the right channel over a second radio at another frequency. When the signals were played over separate loudspeakers, cross-talk from opposing left and right signals degraded the effect. The effect was optimized when the signals from two radio receivers were played over separate circumaural earphones. The effect was quite pleasant. Listeners to the radio program would get the impression of being in the concert hall while remaining in their homes. However, the expense of two transmitting antennae and two radio receivers prohibited the widespread application of binaural sound. When stereo techniques were realized, binaural sound over radio met its demise. "Bell Labs" was one of the first groups to use the dummy head technique to listen to binaural sound (1932). A manikin named Oscar consisted of a torso, head, and ears, with a microphone at each cheek. The manikin encoded sounds with the naturally occurring directional and distance cues. A listener wearing a pair of headphones heard the sound field at the location of the manikin. The "dummy head" technique has become popular as a measurement device.

Advancements in manikin and microphone technology enabled other applications of virtual audio. Three analysis applications have been developed: (1) auditoria design, (2) automotive noise analysis, and (3) binaural hearing research. Concert hall acoustical simulation can be economically beneficial when designing a concert hall or auditorium. Blauert (1983) made a one-hundredth scale model of a concert hall which was instrumented with a down- scaled manikin inside the model. With this apparatus, concert hall designers were able to measure the acoustics of the hall when materials were changed and when structures in the hall were rearranged. Head Acoustics, in cooperation with the University of Aachen in Germany, developed a noise analysis system using digital signal processing technology called the binaural-analysis-system (BAS). The outputs from a pair of microphones inside a manikin's ears are processed by a digital computer and displayed in the time and frequency domains to the user. The signals provided by the manikin are much closer to what a person would hear than that provided by a single microphone. The "Aachen
Head" system has gained acceptance in the automotive industry to identify the sources and actual listening levels of automotive sounds. The Knowles Electronics Manikin for Acoustic Research (KEMAR), instrumented with a pair of Zwislocki middle ear couplers and Bruel & Kjaer half inch microphones, has become standard equipment for psychoacoustic experiments.

Researchers have investigated the effects of the physical size and shape of heads and pinnae on HRTFs and perception using manikins and live subjects. Kuhn (1977) measured the frequency dependence of interaural time delays. Shaw (1974) developed a mathematical model which correlated various aspects of the pinna with the spectral components of the HRTFs. Bronkhorst and Plomp (1988) have studied the effects of magnitude and phase components of the HRTF on directional speech intelligibility in noise. Kulkarni (1992) studied the effect of reverberation of different environments on localization and externalization perception. Advancements in miniature microphones have facilitated recordings in the ear canals of live subjects.

Two applications of virtual audio using miniature microphones have been realized: (1) multi-media audio/video recordings and (2) head related transfer function (HRTF) measurements (Genuit and Burkhard, 1992). Binaural microphones available from Core Sound in Teaneck, New Jersey greatly enhance the audio recording capability of any video camera equipped with stereo audio inputs. One caution is that the wearer of the microphones must keep his/her head aligned with the camera to prevent conflicting motion of the audio and visual images. In another application of miniature microphones, Wightman (1989) measured the HRTFs of humans at 144 directions: 24 azimuths at six elevations. In the laboratory, these data were used to compare directional hearing perception with one's own HRTFs and with other HRTFs.

A different technique for measuring HRTFs was realized using half inch microphones in the ear canals of a manikin (McKinley, 1988). In this technique, the interaural time delay was measured separately from the magnitude portion of the transfer function. Measurements were made for 272 directions at geodesic vertexes of about fifteen degrees separation. Also, a high resolution set of HRTFs were collected at one degree spacing in azimuth to avoid interpolating between points. A detailed description of other HRTF measurement techniques have been documented by Genuit and Burkhard (1992). HRTFs have been implemented in several virtual audio systems to encode directional cues onto sound sources.

VIRTUAL SIMULATION

LOUDSPEAKERS (Aerospace)
The encoding of spatial hearing cues over loudspeakers in aircraft has been nearly non-existent due to limited available space, high noise levels, and degrading reverberation effects in cockpits. Despite the deleterious listening environment, Tobias (1972) demonstrated that some benefit can be gained by encoding speech signals with phase information. An array of three loudspeakers was used to spatially separate speech messages and speech intelligibility was measured in a light aircraft. A twelve to fifteen percent improvement in intelligibility, which corresponds to an effective two dB decrease in the noise level, was measured for the speech out-of-phase versus in-phase condition. Licklider (1948) found a 3-12 dB effective decrease in the noise masking level from changing the relative phase of the signal and noise in an anechoic environment. Use of loudspeakers to incorporate spatial hearing cues has been in the interest of the entertainment industry to make audiences more a part of the event by immersing them in the sound field.

LOUDSPEAKERS (Non-Aerospace)
Several techniques have been implemented to simulate spatial sound over loudspeakers, but none have been fully successful in creating the illusion of being fully immersed in another listening environment. The difficulty lies in making the illusion interactive with the listener over a large area with a limited number of
loudspeakers in various environments. Entertainment industries are interested in spatial sound simulation in the cinema and home music sound systems.

Some success was achieved by Damaske (1971) with the true reproduction of all directional information by the stereophony (TRADIS) process. The TRADIS technique reduced the cross-talk of the left and right signals by transmitting a correction of cross-talk from the right side with the left signal, and vice versa. The spatial image of the sound appeared to be extremely good, but only in a small area in a particular room with a specific arrangement of the loudspeakers. Unfortunately, the spatial image would not work well over large areas. Other techniques have achieved spatial imaging over larger areas, but with less accuracy.

Quadrophony included lateral delays of the back two loudspeakers to increase the spatial sound image. The effect was not convincing enough to justify the cost of the quadrophonic mixer and two additional loudspeakers. The surround Sound system from Hughes has been more widely accepted in cinemas and home entertainment systems. Phasing between loudspeakers in addition to the lateral delays helps to create a better sense of immersion than the quadrophonic system.

Reinforcement of the lateral reflections in auditoria has been achieved with electronic means to improve the "liveliness" of the acoustics. An auditorium is instrumented with an array of microphones and loudspeakers and is controlled by a central processing unit. Varying degrees of success have been reported with such a system. Reproduction of binaural sound over headphones using synthetic means is in some sense easier than over loudspeakers because the left and right signals are nearly completely separated.

HEADPHONES

At the late 1980's, three virtual audio systems for headphone presentation had been developed. All the systems incorporated HRTF's and head motion coupling to present various sound sources from localized points around, above and below the listener. In contrast to the entertainment industry's requirement for spatial sound to envelop the listener, aerospace applications generally demand that sounds originate from a specific point in space. Three groups have developed virtual audio systems that differ in design and capability.

(Aerospace)

The Air Force's Armstrong Laboratory system is based on Texas Instruments TMS-320 series digital signal processing chips and can throughput about 1000 HRTF head position updates per second. The system is compatible with several electromagnetic head trackers from Polhemus Corporation and Ascension Technologies. However, the state-of-the-art in a head tracker technology consists of processing speeds in the range of 30 to 100 Hz, which creates a noticeable lag in the location of a sound source during moderate head rotation. The Air Force system can be configured in software for either high resolution (one degree, 360 points) in the azimuth only mode, or for interpolated high resolution in the combined azimuth and elevation mode. Two audio inputs can be processed simultaneously in the laboratory version, and up to four inputs can be processed in the military version. The laboratory systems have been used in flight simulators, and for researching speech communications and target acquisition applications. The military version has been successfully flown in an AV-8B Harrier aircraft demonstrating spatially separated communications and target acquisition applications (McKinley, 1993).

Several virtual audio generators were developed by Gehring Research Corporation (GRC). The most recent version, the AL-204, was capable of presenting up to four sound sources from any direction by interpolating between measured HRTFs at ten degrees in azimuth and eighteen degrees in elevation. Based on TMS-320 processors, the AL-204 system had similar processing power as the Air Force system. Several Gehring systems were sold to simulator and University facilities. However, only one reference on this system was found (Valencia, 1990).
Another virtual audio system, called the Convovotron, was developed by NASA-Ames and Crystal River Engineering. Unlike the other two systems, the Convovotron is based on hard-wired multiply and accumulate processing chips. In contrast, convolutions can be calculated much more quickly, but the head position update rate is slower. Initially, the Convovotron was developed for applications in space environments, such as the Shuttle. It was used on the ground in an experiment with the VIEW helmet mounted display system. Virtual audio was added to the virtual vision of the VIEW device to improve the sense of presence in performing several tasks in the virtual environment (Wenzel, 1990). Both the Gehring and Crystal River systems have been incorporated into commercial products for non-aerospace applications.

HEADPHONES (Non-Aerospace)

Applications for virtual audio outside the aerospace industry are beginning to emerge. Areas include improved virtual reality simulation, spatial audio mixing of music, and navigational aids for the blind. Both the GRC and Crystal River systems have been incorporated into commercial products for non-aerospace applications.

FUTURE APPLICATIONS

Virtual audio technology may benefit many human operators, who work in environments of reduced situational awareness. These environments are sometimes created at the human/machine interface by one or two dimensional displays of three dimensional information. An example would be trying to monitor several radio communications over a single channel headset. In other circumstances, too much information is presented through one sensory channel, thereby overloading the operator. By spatially presenting some visual information through the auditory modality, an operator may be able to perform a task more efficiently (Folds, 1990). Many high workload tasks exist that have reduced spatial display formats.
airports. Pilots would have improved situational awareness by knowing the relative location of other aircrew members, control towers, and other aircraft, the latter of which may help to avoid air-to-air collisions.

In the military, similar environments exist in aircraft, submarines, and armored vehicles, which could be improved with virtual audio technology. Threat warning signals could be spatially displayed to more quickly alert personnel of imminent danger. The intelligibility of speech signals would improve while listening to spatially separated communications. Navigation towards an auditory beacon may make orienting, aerial refueling and landing tasks easier. Virtual audio displays have many potential applications outside of the aerospace industry.

NON-AEROSPACE

As with the aerospace applications, low situational awareness environments may benefit from virtual audio technology (Hamilton, 1992). Two diverse areas include training and entertainment. Training applications range from education of architectural design, medical diagnosis, and enhancing virtual reality simulations. Entertainment areas include arcade (video) games, spatial music reproduction for the cinema and home entertainment, and others. Many communications technologies could be improved with virtual audio techniques. Some include hearing aids (Yost, 1987), fleet dispatching, and teleconferencing. Development of such commercial products will probably be driven by technology and consumer demand.

SUMMARY

In summary, several applications of virtual audio have been described as to their method and purpose. Past, present, and future applications were discussed for aerospace, research, and entertainment industries. Techniques for creating virtual audio, over loudspeakers and headphones were described. Virtual audio systems are more flexible than previous physical simulations and can improve low fidelity spatial environments. Technological advances of computers and position trackers have enabled the development of spatial audio display systems. A high potential exists for virtual audio applications in aerospace and non-aerospace environments.

REFERENCES


McKinley, Richard, to be published in NAECON 93 proceedings.


Smith, David C., Telephone conversation, 1993.


