

REALISTIC 3D SOUND SIMULATION IN THE VIRTTEX DRIVING SIMULATOR

Mike Blommer, Jeff Greenberg
Ford Motor Company
P.O. Box 2053, MD 3135
Dearborn, MI 48124
Phone/FAX: 1-313-621-8197
mblommer@ford.com

Submitted July 15, 2003

ABSTRACT

Relatively inexpensive interactive 3D sound simulation systems can now be realized in driving simulators because of recent advances in binaural recording, PC-based sound simulation, and professional audio hardware components. Binaural recordings of sounds at the driver's position are easily acquired, largely due to increased focus by vehicle OEMs on customer satisfaction with vehicle sound quality. The recent explosion in popularity of PC-based multi-media games has led to cost-effective hardware and software methods to simulate effects such as Doppler shift, reverberation, and spatialization. The ever-expanding types of professional audio hardware components now make it easy to filter, mix, and route multi-channel audio data.

Combining the advances in all of these areas has led to the development of a realistic interactive 3D sound system for the VIRTTEX driving simulator at Ford Motor Company. This paper describes in detail the sound database, the software programming environment, and the hardware components used to create the 3D sound system. Characterization of how well the simulated acoustic environment matches an actual vehicle acoustic environment is also presented. This characterization is important since a realistic simulated acoustic environment allows for conducting a variety of studies, including meaningful Noise, Vibration, & Harshness studies, in driving simulators.

INTRODUCTION

A major goal in driving simulation is to immerse the driver into a meaningful virtual world. The success of this immersion depends on the realism of many sensations, including 3D sound. It is preferred to render 3D sound using loudspeakers instead of headphones because drivers do not wear headphones in real vehicles. Different methods to render 3D sound using loudspeakers include crosstalk cancellation (1,2), stereo-dipole methods (3,4), and surround sound. Headphone playback is also used in order to avoid the complex acoustic field in the vehicle cabin. Binaural recording and equalization techniques can be used with each of these methods to further enhance the perception of a 3D acoustic environment. The upgraded sound system used in VIRTTEX is capable of using both headphone playback of binaural recordings, and playback of binaural recordings over multiple loudspeakers using a binaural equalization technique (5-7). The binaural equalization technique is analyzed in this paper for the possibility of conducting Noise, Vibration, & Harshness (NVH) studies in VIRTTEX since NVH sounds are typically recorded with a binaural head at the driver's location in a real vehicle. Besides conducting specific NVH studies, this playback technique is advantageous for general 3D sound rendering since many vehicle sounds (e.g., powertrain, road, and wind noise) are easily recorded with a binaural head in a vehicle of interest.

Implementing the binaural equalization technique in the VIRTTEX simulator involves using "off-the-shelf" audio hardware that supports up to 8 channels of high-bandwidth data. This multi-channel audio data is transferred over 45 meters from a readily available PC system to the motion base of the VIRTTEX simulator using a single fiber-optic connection that minimizes noise interference. The simulated effects of Doppler shift, reverberation, and spatialization, as well as playing monophonic sounds to specific speakers (e.g., sounds for warning or navigation systems) are generated by treating multi-channel audio output as pairs of stereo outputs and using a free and user-friendly Application Programming Interface (API) that provides an interface to Microsoft® DirectSound®.

The major advantage of this system is its ability to render realistic 3D sound at a reasonable cost (less than \$10k US for the entire system including computer, software, audio equipment, and speakers). The free API easily generates 3D effects based on customizable software that receives information from the main vehicle model computer. Thus,

the system described here provides an audio system that can be customized to fit the needs of most any driving simulator.

The following sections provide a brief overview of the VIRTTEX simulator along with a detailed description of the sound-rendering system. In particular, analysis of equalizing the acoustic field in the simulator buck and the resulting crosstalk is presented. This analysis provides objective data for assessing the possibility of conducting NVH studies in VIRTTEX using the binaural equalization technique.

VIRTTEX DRIVING SIMULATOR

VIRTTEX is a large, 6 degree-of-freedom (DOF) moving base driving simulator. A vehicle buck is mounted inside a 24-foot diameter dome, with the driver's eye-point located at the center of the dome viewing volume. Five projectors are used to project a 180°x40° front and a 120°x25° rear image onto the inside surface of the dome. The inside surface of the dome is covered with a high-gain (4.5:1) coating to provide a bright image at the drivers eye-point location.

OVERVIEW OF RENDERING 3D AUDIO

The methods currently used to render audio in driving simulators varies greatly (e.g., 8-12). Some simulators play sounds over speakers without regard to the acoustics of the listening environment, whereas other simulators use headphone playback in order to avoid the complex acoustic field inside the vehicle cabin. The process of rendering interactive 3D audio can be broken down into recording/synthesizing sounds, possibly modifying the sounds based on the current state of the driving simulator (e.g., vehicle speed, engine rpm, Doppler effects for exterior objects, etc), and then playback of the sounds to the listener. In order to better describe the VIRTTEX sound system and compare it to other rendering methods, a brief overview of rendering 3D audio is presented in the following subsections.

Recording/Synthesizing Sounds

Sounds can be recorded or synthesized using a variety of techniques. Binaural recordings at the driver's position in a vehicle cabin will better represent what a driver actually hears. The binaural head recording contains most characteristics of the acoustic filtering by a listener's head, torso, and pinnae (13). These characteristics include diffractions, reflections, and linear distortions of the incoming acoustic waves, which result in amplitude and phase modifications of the waves as a function of both frequency and spatial location. The human auditory system "decodes" these amplitude and phase modifications into spatial locations of the sound sources.

By measuring these head-related transfer functions (HRTFs) from a spatial location to the ear canal of the right and left ears, one can also incorporate binaural effects when synthesizing sounds. Merely filter the synthesized sound by a set of HRTFs that correspond to desired spatial locations.

Interactive Audio Effects

Generating interactive effects such as echo, reverberation, and Doppler shifts are fairly standard in software today. For example, many Windows®-based software applications generate these effects using Microsoft® DirectSound®. These effects can greatly enhance the perception of 3D sound, with an example being the Doppler shift due to a passing vehicle.

Acoustic Environment and Sound Playback

Sounds are typically played back over a set of loudspeakers or headphones. For loudspeaker playback, there are various techniques to "expand" the acoustic image in an attempt to render 3D audio. The most straightforward technique is to use many loudspeakers and play sounds from the appropriate speaker in order to create the 3D perception. A reduced version of this technique is surround sound, which typically involves four speakers placed about the listener and one "center-channel" speaker.

For two-speaker playback, crosstalk cancellation or stereo-dipole techniques are used to expand the acoustic image. A pictorial representation of crosstalk cancellation is shown in Figure 1. The loudspeakers are usually placed 30-60 degrees to each side of the listener. The goal is to eliminate the crosstalk from the right speaker to the left ear

(and the left speaker to the right ear) so that the loudspeakers essentially act as a set of headphones. That is the right channel of the stereo sound goes only to the right ear, and the left channel goes only to the left ear. Simple models are typically used to represent the transfer functions from the speakers to the ears. Drawbacks for crosstalk cancellation include a relatively small "sweet spot" (i.e., where the expansion effect works well), and also insensitivity to front/back source location.

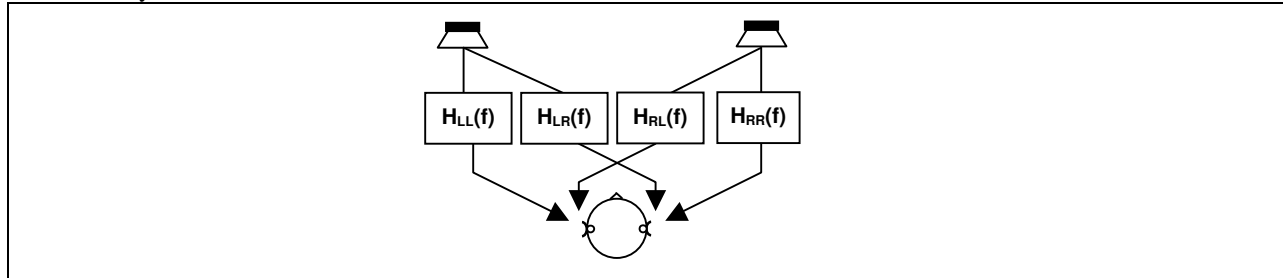


Figure 1. Crosstalk cancellation

The stereo-dipole technique is somewhat similar to crosstalk cancellation, except the speakers are placed 5-10 degrees apart and in front of the listener. Stereo-dipole playback usually has a larger sweet spot than crosstalk cancellation, but is also insensitive to front/back source location.

Headphone playback avoids dealing with the usually complex acoustic environment associated with loudspeaker playback, but also has drawbacks. These include internalization of sounds (i.e., sounds appear to emanate from inside the listener's head), and the fact that sound sources move with the listener's head movements unless some type of head-tracking device is used.

All of these playback methods can be enhanced with the use of HRTF equalization. That is, the standard implementation of these playback methods assume a point-receiver model for each ear of the listener and do not take into account the acoustic filtering of the head, torso, and pinnae. By including HRTF equalization in lieu of the point-receiver models, the expansion effect is improved. However, the computation requirements can be greatly increased with HRTF equalization, which may not make it appropriate in all interactive audio environments.

BINAURAL EQUALIZATION TECHNIQUE

The binaural equalization technique uses standard loudspeaker playback in combination with HRTF equalization. The basic idea is to place a binaural head at the driver's location in the simulator buck, and measure the transfer function consisting of the right simulator speakers, simulator buck cabin, and right ear of the binaural head. The inverse of this transfer function is calculated and programmed into equalizers corresponding to the right simulator speakers. A similar process is carried out for the left simulator speakers. Then the driver will experience realistic vehicle sounds when binaural recordings are played over the simulator sound system. Figure 2 shows an overview of the vehicle recording and simulator equalization process. For simplicity, only the process for the front right speaker is shown.

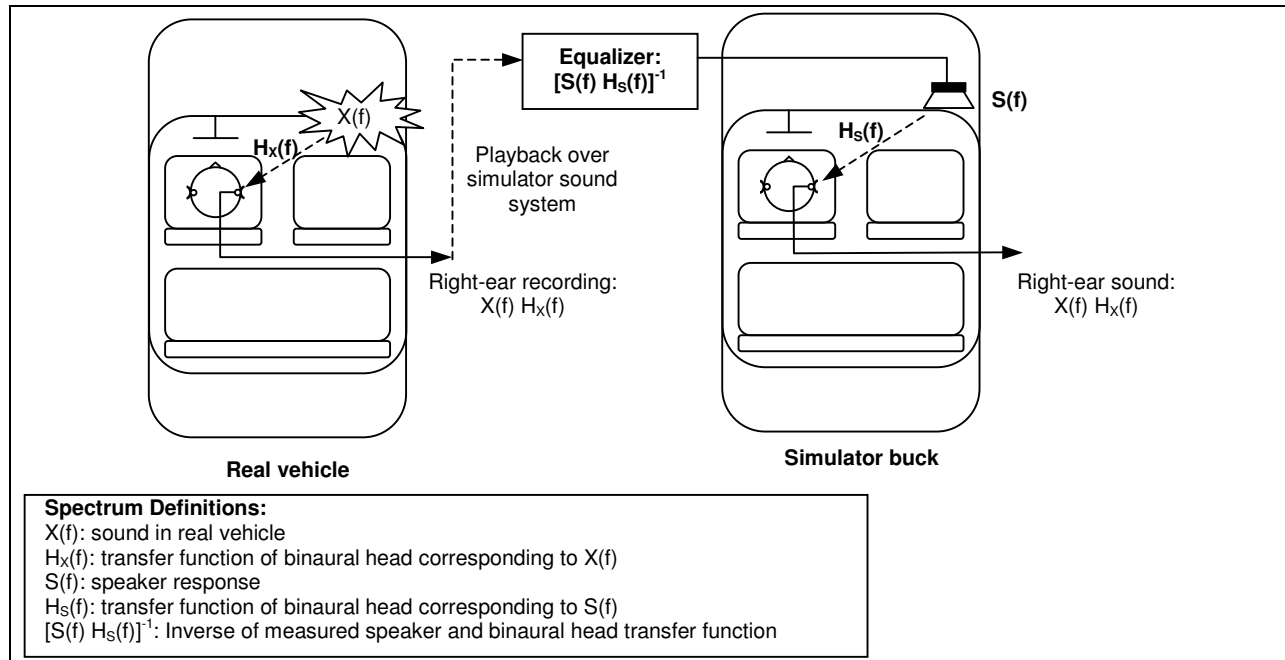


Figure 2. Overview of vehicle recording and simulator equalization process.

SOUND RENDERING

The primary sounds sensed by drivers of real vehicles are associated with engine/powertrain, wind, and road/tire noises. These sounds are a function of the following parameters, or vehicle states: vehicle speed, engine rpm, engine load, and road conditions. In addition, drivers of real vehicles experience other continuous or transient background sounds (e.g., expansion joints, rumble strips, car horns, etc.). These background sounds are a function of scenario states. All sounds are rendered in the VIRTTEX simulator by

1. Receiving vehicle and scenario state information from a real-time I/O computer
2. Selecting appropriate pre-recorded sounds from a database for powertrain, wind, road, and background noise
3. Appropriately pitch-bending and blending selected sounds to generate sounds corresponding to the vehicle and scenario states
4. Playing sounds over the simulator buck speaker system or headphones

Figure 3 shows how vehicle and scenario state information is sent to a designated sound-rendering PC. The following subsections provide more details on the sound-rendering subsystems.

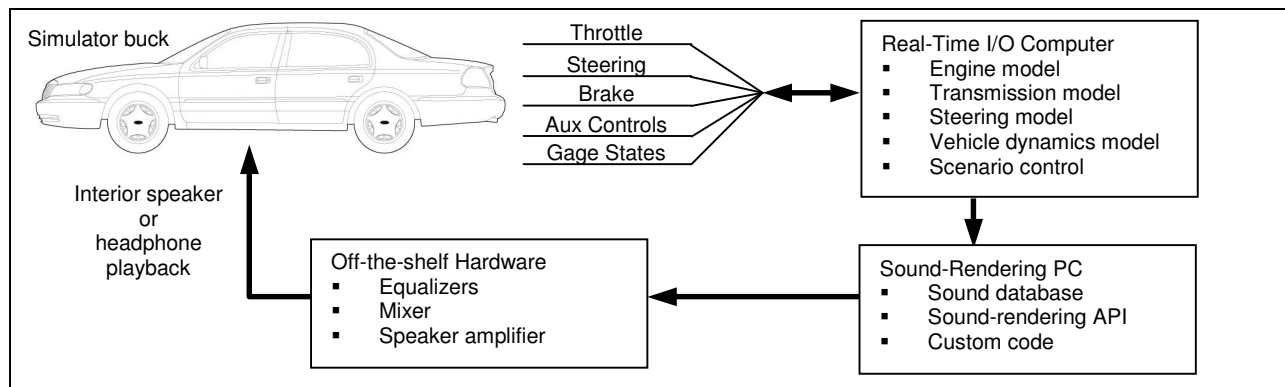


Figure 3. Vehicle and scenario state information input to the sound-rendering PC.

Sound Database

Powertrain, wind, and road sounds are recorded in a real vehicle cabin using a binaural head in the driver's location. These binaural recordings are made as a function of engine rpm/load and vehicle speed, and stored in a database. The wind and road sounds can be made separately or simultaneously. Separate measurements can be made in a wind tunnel and on chassis rolls, respectively. Simultaneous measurements can be made by turning the engine off at various speeds on different test track surfaces. Sounds measured in 16 kph (10 mph) increments and of 10-20 seconds in duration are sufficient for realistic road and wind noise simulation. Simulation at speeds between the recorded speeds is done through appropriate sound-level mixing of the database sounds (see Software subsection). Each database sound is prepared as described in the APPENDIX so that looping on the sound during playback does not produce any clicks or pops.

Powertrain sounds can be recorded on chassis rolls at a number of rpm/load conditions. The rolls and tires are selected to minimize their contribution to the powertrain recording. Alternatively, powertrain sounds can be synthesized. The VIRTTEX sound database currently contains both types of sounds at increments of 500 rpm and various load conditions. Genuit and Bray (7) recommend 10 different load conditions in order to accurately simulate all powertrain conditions. Similar to the wind and road sounds, the powertrain sounds are 10-20 seconds in duration and prepared to prevent clicks/pops during looping playback. Pitch-bending and sound-level mixing are used to simulate powertrain sound at rpm/load conditions that are not in the database.

Background sounds (e.g., rumble strips, car horns, etc.) are recorded either with a binaural head or monophonically. Appropriate HRTFs or panning algorithms (see Software subsection) are used if the sounds are recorded monophonically.

Software

Many sound-rendering APIs currently exist for Windows-based PCs (14-17). Additionally, there are complete products available for 3D sound rendering (18-20). Requirements of the API for the VIRTTEX simulator are: (i) low-cost, and (ii) user-friendly programming interface to Microsoft® DirectSound®. An API that meets these requirements is FMOD (14). It includes support for 3D sound, MIDI, MP3, recording, obstruction/occlusion, DSP effects (e.g., pitch-bending, panning, and variable playback rates), spectral analysis, user-created samples and streams, and synchronization. The FMOD dynamic linked library (DLL) currently allows for stereo sound, but multiple instances (i.e., multiple sound cards) can be run simultaneously to generate multi-channel sound.

Custom software is built around the FMOD API, which receives vehicle and scenario state information from the real-time I/O computer at a rate of 60Hz and outputs sound to designated channels. Transient sounds are generated by receiving timing information from the vehicle and scenario states. Continuous sounds are generated by appropriately looping sounds from the database, as described in the following subsection.

Pitch-bending and Level-mixing Database Sounds

For simplicity, consider simulating powertrain sound at an rpm R such that $R_1 < R < R_u$, where R_1 and R_u correspond to rpm conditions in the database. The powertrain sound in VIRTTEX is simulated by mixing pitch-bent and level-adjusted versions of the sounds corresponding to R_1 and R_u . These sounds are pitch-bent up and down, respectively, in "real-time" to R using the FMOD API. Level-mixing is carried out through a "constant-power" method. If $x(t)$ denotes the desired sound at R rpm, and $x_1(t)$ and $x_u(t)$ denote the neighboring database sounds, then $x(t) = A x_1(t) + B x_u(t)$. A and B are chosen to preserve power, i.e., $A^2 + B^2 = 1$. A monotonic function is used for B that increases from 0 to 1 as R increases from R_1 to R_u . The corresponding value for A is $(1 - B^2)^{0.5}$.

Hardware

The hardware system is based on relatively inexpensive "off-the-shelf" pro-audio components that interface with a Windows-based PC. Figure 4 shows the layout of the hardware components and Table 1 describes the hardware components. Components are located either in the VIRTTEX control room or on the motion platform. Audio data between these two locations is transferred over a 45 m (150 ft) optical cable with associated transmitter and receiver.

There are speakers located at the four "corners" of the simulator buck, along with a subwoofer and a seat shaker. The seat shaker is used to provide contextual high-frequency (> 20 Hz) vibration to the driver. Dedicated speakers are used for the front speakers, whereas the simulator buck rear speakers are used for both 3D sound rendering and output from the buck's audio system. The equalizers can be bypassed if desired, and the sounds played over headphones to guarantee accurate binaural reproduction. However, as mentioned previously, a drawback with headphone playback is that the driver loses some sense of immersion in the virtual world since drivers do not wear headphones in real vehicles.

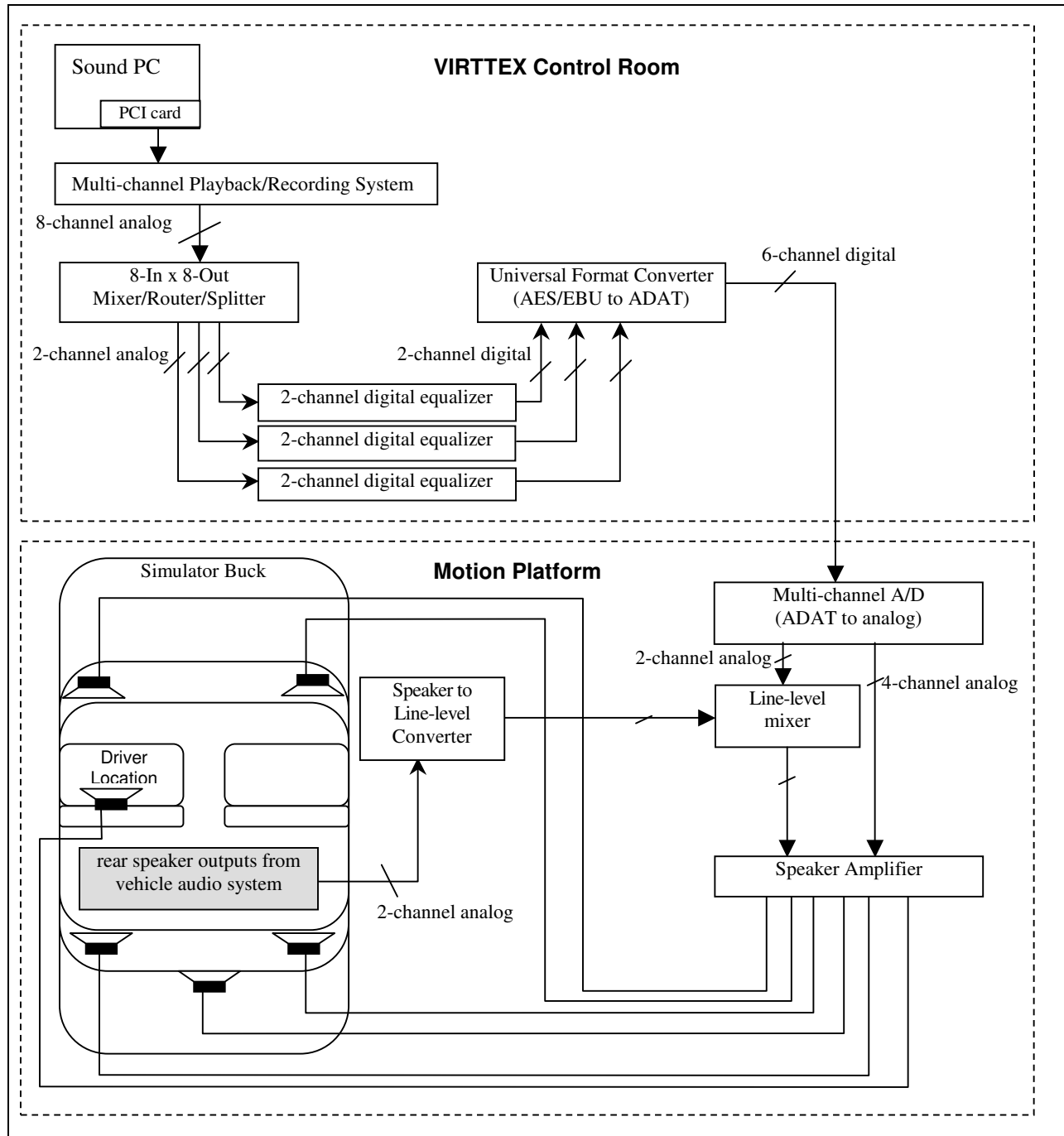


Figure 4. Layout of sound system hardware.

Table 1. Description of sound system hardware.

Item	Description
Multi-channel Playback/Recording System	Echo Layla 24 8-channel sound card
8-In x 8-Out Mixer/Router/Splitter	Shure Auxpander
2-channel Digital Equalizer	Behringer Ultracurve PRO DSP 8024
Universal Format Converter	RME ADI-8 DD
Multi-channel A/D	RME ADI-8 PRO
Speaker Amplifier	Rane MA 6S
Rear speakers	Simulator buck speakers (currently Volvo S80)
Front speakers	Monsoon Planar Media 9 LCD Flat Panel Speaker
Seat Shaker	Aura Bass Shaker AST-2B-4

EQUALIZATION OF SPEAKERS, SIMULATOR BUCK, AND BINAURAL HEAD

All transfer functions for equalization were measured using a maximum-length sequence (MLS) technique (21-23), a common method for measuring acoustic system transfer functions. *N*-point MLSs have the property that their circular cross-correlation is 1.0 at zero-lag and $1/(N-1)$ at all other lags – almost an ideal impulse function. Thus, by measuring a system response to an MLS and then cross-correlating the MLS to the measured response results in a very accurate estimate of the system impulse response.

The responses of thirty, 16383-point MLSs were recorded at a sampling rate of 44.1 kHz, and the result averaged to calculate each impulse response. A 16384-point FFT was applied to each impulse response to calculate the corresponding frequency response. As a characterization of the measurement system itself, Figure 5 shows the transfer function associated with the analog-to-digital (A/D) – speaker amplifier – D/A system, which is flat (± 3 dB) from 25 Hz to more than 10 kHz.

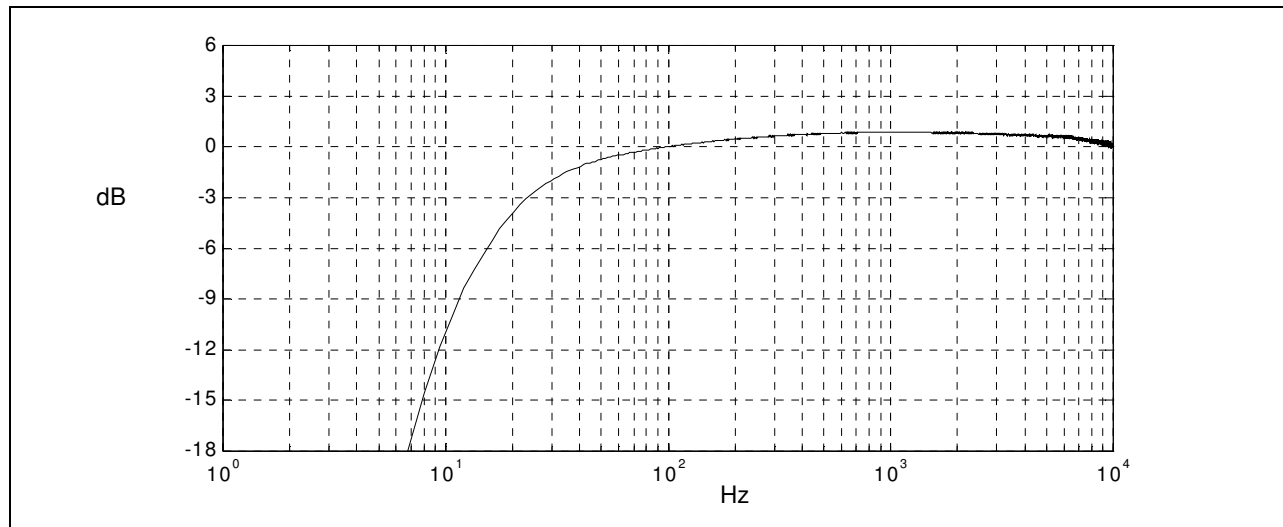


Figure 5. Frequency response of the measurement system consisting of D/A, speaker amplifier, and A/D.

The simulator buck was equalized for use with the binaural equalization technique described above in BINAURAL EQUALIZATION TECHNIQUE. Recall that a binaural head is located at the driver's location and transfer function measurements are made from each speaker to both the driver outboard ear (DOE) and driver inboard ear (DIE). It is desired to have a flat frequency response consisting of the passenger-side speakers, simulator buck cabin, and the DIE. Similarly, it is desired to have a flat frequency response consisting of the driver-side speakers, simulator buck cabin, and the DOE.

Figure 6 and Figure 7 show the 1/3-octave responses of the driver-side speakers to DOE, and passenger-side speakers to DIE, respectively. Notice that these responses vary by more than 20 dB, which must be flattened to within +/-3 dB by the equalizers.

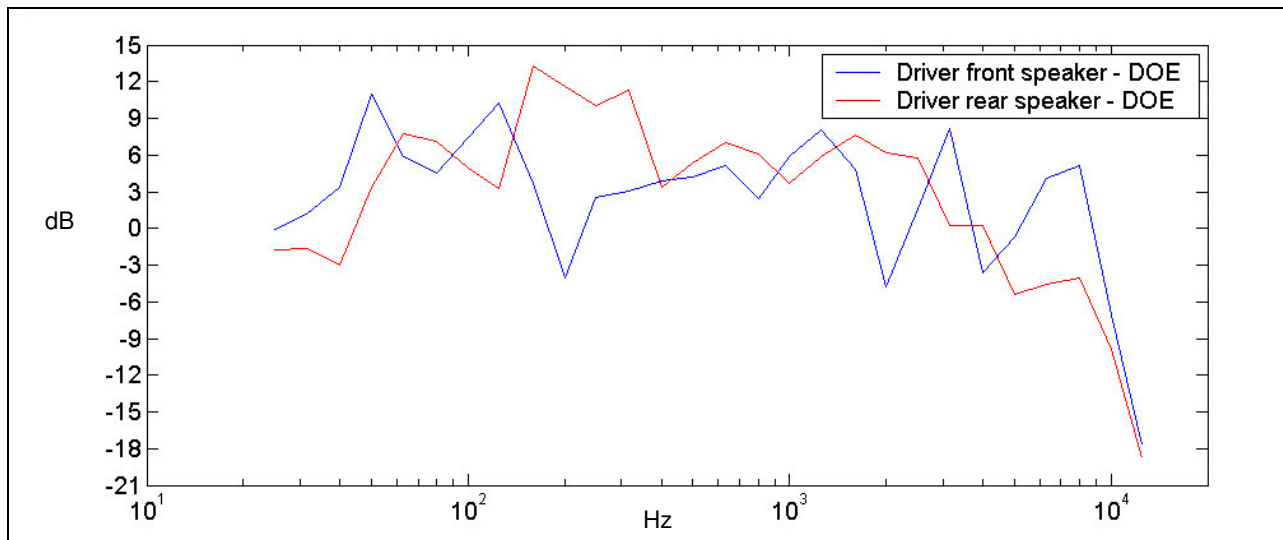


Figure 6. 1/3-octave responses corresponding to driver-side speakers and DOE.

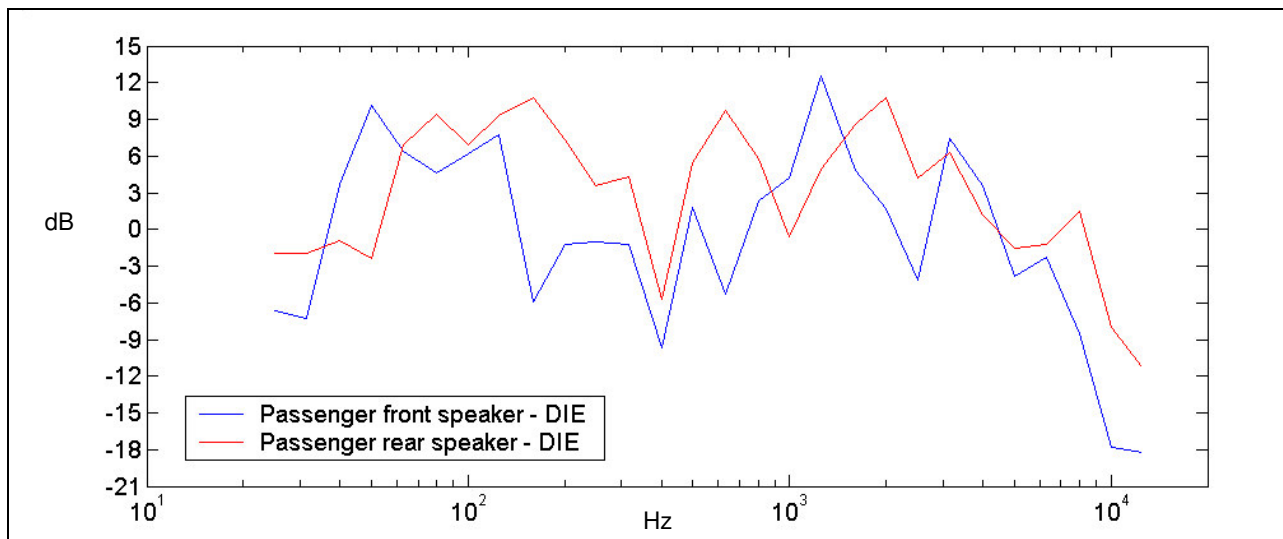


Figure 7. 1/3-octave responses corresponding to passenger-side speakers and DIE.

Figure 8 and Figure 9 show the *equalized* 1/3-octave responses of the driver-side speakers to DOE, and passenger-side speakers to DIE, respectively. That is, these responses include an equalizer in the measurement chain with equalizer gains set to flatten out the responses of Figure 6 and Figure 7, respectively. The equalized frequency responses are within +/-3 dB from 40 Hz to 10 kHz. Equalized frequency responses are larger than +/-3 dB where the speakers produce relatively little acoustic energy.

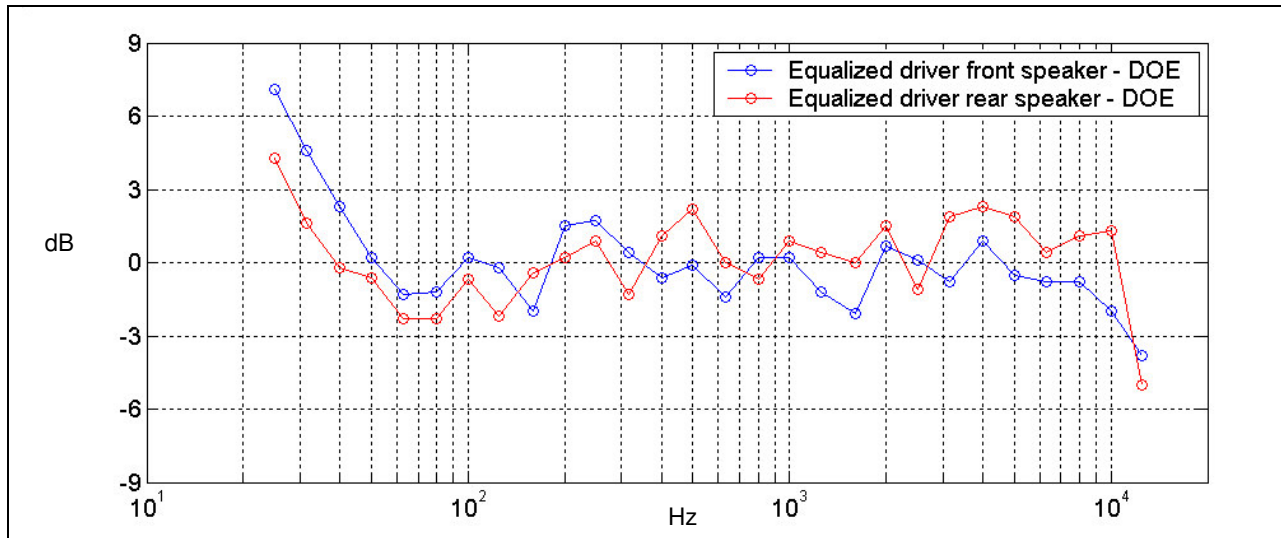


Figure 8. Equalized 1/3-octave responses corresponding to driver-side speakers and DOE.

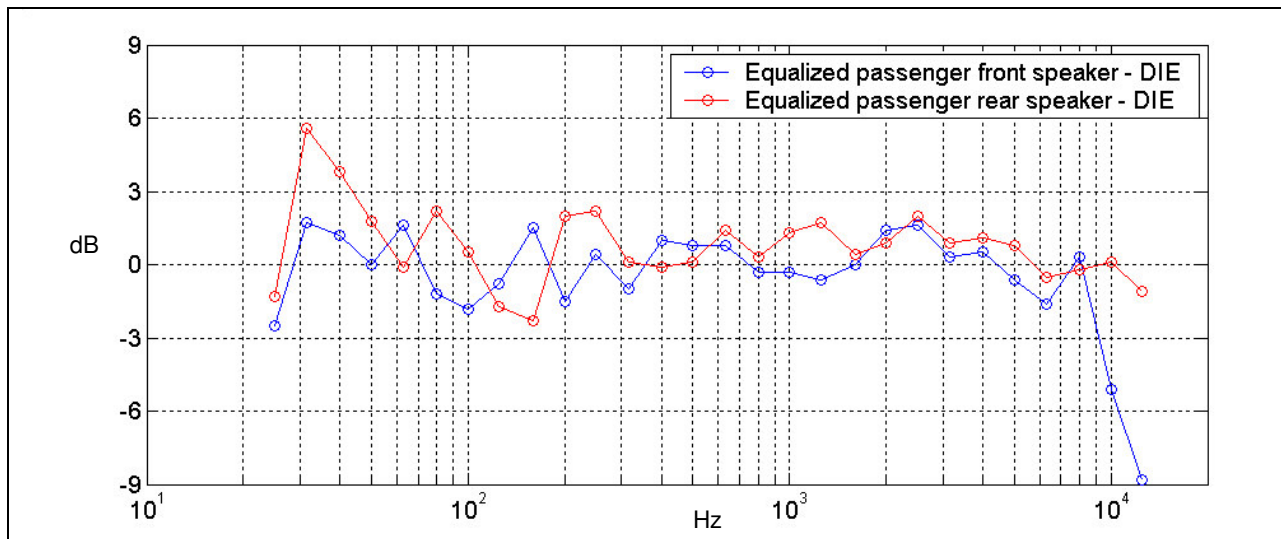


Figure 9. Equalized 1/3-octave responses corresponding to passenger-side speakers and DIE.

Pure time delay associated with each transfer function is easily calculated from the impulse response. This delay is defined as the time where the absolute value of the impulse response first achieves 5% of its maximum absolute value (24-25). Table 2 lists the time delay for each transfer function. These delays are programmed into the equalizers so that there are not any relative time delays between any of the transfer functions listed in Table 2.

Table 2. Time delays associated with the listed transfer functions.

Transfer Function	Time Delay [ms]
Driver front speaker – DOE	5.17
Driver rear speaker – DOE	5.99
Passenger front speaker – DIE	5.99
Passenger rear speaker - DIE	6.19

CROSSTALK

Crosstalk for the binaural equalization technique is defined as the acoustic energy that reaches the DOE from the passenger-side speakers, and the acoustic energy that reaches the DIE from the driver-side speakers. Crosstalk is reduced with the binaural equalization technique by relying on acoustic shadowing effects of the driver's head (13). That is, the head provides acoustic attenuation, or "shadowing" from the passenger-side speakers to the DOE, and from the driver-side speakers to the DIE.

Figure 10 and Figure 11 show the crosstalk in the sound-rendering system, after the equalization described above. The figures correspond to the passenger-side speakers to DOE, and driver-side speakers to DIE, respectively. These responses should be significantly attenuated from 0 dB for minimized crosstalk.

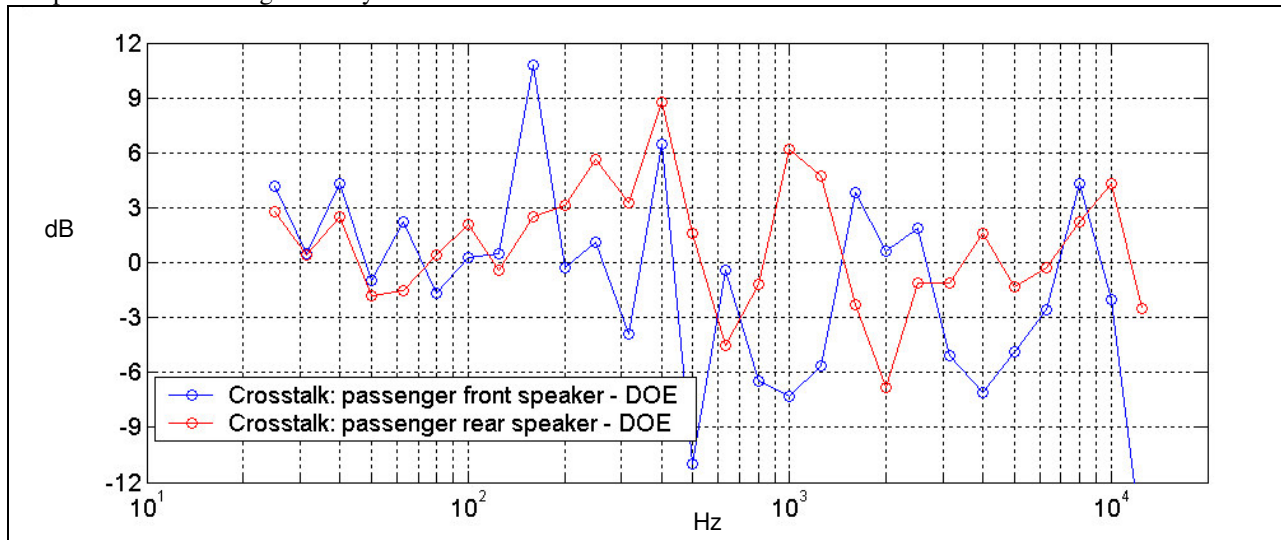


Figure 10. Crosstalk 1/3-octave responses corresponding to passenger-side speakers and DOE.

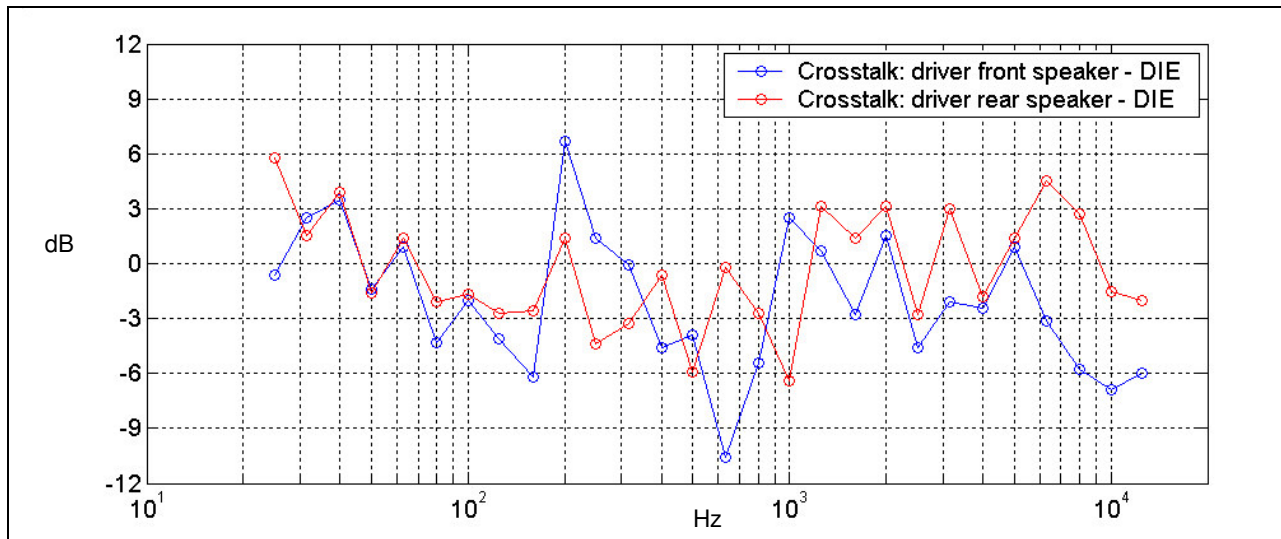


Figure 11. Crosstalk 1/3-octave responses corresponding to driver-side speakers and DIE.

There is significant crosstalk compared to the results shown in (6) for an anechoic chamber. This is expected since the simulator buck is an enclosed environment with many reflecting surfaces. However, (6) also indicates that directionality of binaural sounds is still relatively accurate in driving simulators using the binaural equalization technique.

DISCUSSION

Informal listening by experienced listeners of binaural recordings indicates that directionality of binaural sounds played over the simulator sound system is significantly improved over the condition that bypasses all equalization. Since the crosstalk in the loudspeaker system is significant, it is likely that interaural time differences (ITDs) (*I3*) in the actual binaural recordings provide additional directional cues to overcome the crosstalk. However, comparison of the binaural equalization technique to headphone playback indicates that directionality is reduced and some spectral coloration of sound sources is introduced due to the crosstalk in the loudspeaker system. Crosstalk cancellation techniques could be used to reduce or eliminate the crosstalk, but at the expense of a relatively small sweet spot. Since drivers of different heights, and thus, different listening positions, use the VIRTTEX driving simulator, it was decided not to pursue these techniques.

Based on the above observations, the binaural equalization technique is not recommended for experiments or tests where very accurate acoustic playback is required, e.g., for NVH studies. In this case, it is recommended that the binaural recordings be played back over appropriately equalized headphones.

CONCLUSION

A relatively inexpensive and interactive 3D sound system for the VIRTTEX driving simulator at Ford Motor Company has been described in terms of its sound database, software, and hardware systems. The binaural equalization technique is used to render 3D sound from binaural recordings. An analysis of various transfer functions indicates that while digital equalizers can be used to appropriately equalize the simulator buck cabin in real-time, there is significant crosstalk in the system that reduces the perception of sound directionality. Informal listening evaluations indicate that while the binaural equalization technique does provide reasonable directionality of sound sources, it should not be used for extremely accurate playback of binaural sounds. Instead, playback of binaural sounds over headphones should be used.

REFERENCES

1. Atal, B., Hill, M., and Schroeder, M., "Apparent Sound Source Translator," U.S. Patent 3,236,949, Feb. 22, 1966.
2. Bauck, J., "A Simple loudspeaker array and associated crosstalk canceller for improved 3D audio," *J. Audio Eng. Soc.*, Vol. 49, No.1/2, Jan./Feb. 2001, pp. 3-13.
3. Kirkeby, O., Nelson, P., and Hamada, H., "The 'Stereo Dipole' – A Virtual Source imaging system using two closely spaced loud speakers," *J. Audio Eng. Soc.*, Vol. 46, No. 5, May 1998, pp. 387-395.
4. Nelson, P., Kirkeby, O., and Takeuchi, T., "Sound fields for the production of virtual acoustic images," *J. of Sound and Vibration*, Vol. 204, No. 2, 1997, pp. 386-396.
5. Xiang, N., Genuit, K., and Gierlich, H., "Investigations on a New Reproduction Procedure for Binaural Recordings," Presented at the 95th Convention of the Audio Engineering Society, Oct. 7-10, 1993, New York.
6. Krebber, W., Gierlich, H., and Genuit, K., "Auditory virtual environments: basics and applications for interactive simulations," *Signal Processing*, Vol. 80, No. 11, Nov. 2000, pp. 2307-2322.
7. Genuit, K., and Bray, W., "A Virtual Car: Prediction of Sound and Vibration in an Interactive Simulation Environment," *Proc. Of SAE Noise & Vibration Conf.*, Traverse City, MI, May 2001, paper 2001-01-1474.
8. Colinot, J., and Belay, G., "A Multifunctional Sound Generator For The PSA's Sherpa Driving Simulator," *Proc. of DSC2000*, pp. 165-179.
9. Allman-Ward, M., Williams, R., Cockrill, M., Distler, H., Crewe, A., and Heinz, T., "The Interactive NVH Simulator As A Practical Engineering Tool," *Proc. of SAE Noise & Vibration Conf.*, Traverse City, MI, May 2003, paper 2003-01-1505.

10. Caird, J., "The University Of Calgary Driving Simulator (UCDS): Age-Related Intersection Crash Contributors And Intelligent Transportation System (ITS) Countermeasures," http://www.eng.ucalgary.ca/Civil/C-TEP/research/research_UCDS.htm.
11. Goetchius, G., Ketelhut, C., Smallwood, B., and Eaton, C., "Subjective evaluation of NVH CAE model predictions using an operator-in-the-loop driving simulator," *Proc. of SAE Noise & Vibration Conf.*, Traverse City, MI, May 2001, paper 2001-01-1590.
12. Heidet, A., Warusfel, O., Vandernoot, G., Saint-Loubry, B., and Kemeny, A., "A Cost Effective Architecture For Realistic Sound Rendering In The Scanner II Driving Simulator," *Proc. of the 1st HCTSC*, Iowa City, Iowa, November 2001.
13. Blauert, J. *Spatial Hearing. The psychophysics of human sound localization*. MIT Press, Cambridge, MA, USA, 1997.
14. FMOD: Music and Sound Effects System. <http://www.fmod.org>
15. AudioWorks2™. <http://www.multigen-paradigm.com/products/runtime/vega/modules/audioworks.shtml>
16. BASS Sound System. <http://www.un4seen.com>
17. Open Audio Library (AL). <http://www.openal.org>
18. Huron Digital Audio Convolution Workstation. <http://www.lake.com>
19. SMx Digital Audio System. <http://www.simphonics.com>
20. HEAD Acoustics H3S. <http://www.headacoustics.com>
21. Rife, D., and Vanderkooy, J., "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc.*, Vol. 37, No. 6, June 1989, pp. 419-444.
22. Borish, J., and Angell, J., "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.*, Vol. 31, No. 7, July/August 1983, pp. 473-487.
23. Borish, J., "Self-contained crosscorrelation program for maximum-length sequences," *J. Audio Eng. Soc.*, Vol. 33, No. 11, November 1985, pp. 888-891.
24. Blommer, M. *Pole-zero modeling and principal component analysis of head-related transfer functions*. Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, MI, 1996.
25. Sandvad, J., and Hammershøi, D., "Binaural auralization. Comparison of FIR and IIR filter representations of HIRs," presented at 96th convention of *Audio Eng. Soc.*, Amsterdam, 1994.

APPENDIX

Creating Click-Free Looping Sounds

Consider extracting sound from a relatively long recording in order to create a click-free looping sound. A common but difficult and tedious method is to manually select beginning and end points such that they do not create a relatively large discontinuity when played back in a looped mode. An alternative and much easier method is used for the VIRTTEX database sounds. The method is shown pictorially in Figure 12, and described below:

1. Extract sound of desired duration (usually 10-20 seconds) from the recording.
2. Define a blending duration, DUR. DUR = 20 msec works well.
3. Gate the onset and offset of the extracted sound with a cosine-squared window of duration DUR.
4. Split the extracted sound into two halves. Label the two sounds as SND1 and SND2.
5. Swap the play order of SND1 and SND2 and mix them such that their gated portions overlap. This will result in a smooth transition between the two sounds without any loudness change or clicking.

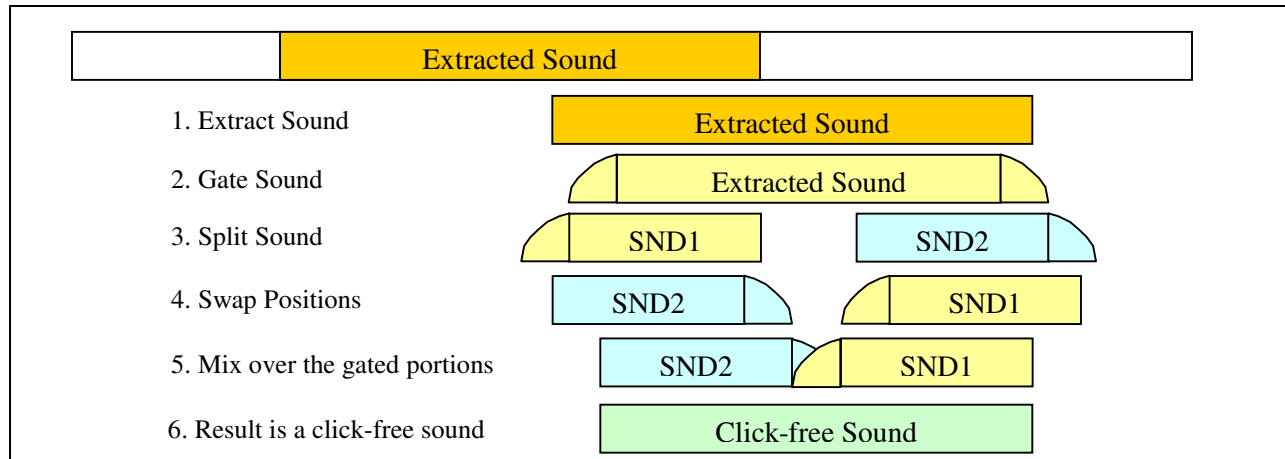


Figure 12. Process used to create click-free sounds for looped playback.