【研究室紹介】

研究室紹介

●研究室紹介●

会津大学の人工的な音響の研究

マルチメディアシステム学講座

ウイリアムL. マーテンス
公園マイケル

0. Introduction
(はじめに)

Virtual acoustic research at the University of Aizu is conducted mainly through two groups led by the authors: The "Cyberspatial Audio Group," led by the first author (Martens), focuses on human spatial hearing and virtual acoustic rendering technology, while the "Spatial Media Group," led by the second author (Cohen), is primarily focussed upon hypermedia interfaces. The two groups are engaged in collaboration that connects applied research on auditory display technology with prototype applications of that technology. The needs of those applications sometimes drive the research, but new research results also drive the development of new applications. Several of these applications (such as the "Internet Chair" [7][8][3]) were described in the authors' review in the September 2001 issue of this journal [4], but this review will concentrate on the virtual acoustic research conducted within the Cyberspatial Audio Group [2][6].

1. Virtual Acoustic Research
(人工的な音響の研究)

To understand the range of virtual acoustic research projects undertaken by the Cyberspatial Audio Group, the scope of virtual acoustic simulation required for cyberspatial audio applications should be described. As will be explained below, the scope is quite broad, including aspects of sound generation, transformation, and display. Before presenting a more detailed explanation of virtual acoustic simulation, however, the field of application under investigation by the group will first be distinguished from the broader field of sound synthesis and processing in general. The next subsection thus attempts to provide this distinction.

1.1 What is Cyberspatial Audio?
(「サイバースペーシャルオーディオ」とは?)

The "Cyberspatial Audio Group" at the University of Aizu is engaged in research and development of new virtual acoustic rendering technology for human inhabitants of the shared synthetic worlds that networked computer users occupy. Such synthetic worlds constitute an alternative reality that has come to be termed "cyberspace," which, stretched by human imagination, spans a wide range of possibilities. Concepts of cyberspace portrayed in popular media are dominated by immersive graphic imagery, but immersive auditory imagery is arguably the more important component when telepresence is desired. It is argued, for example, that users of everyday, telephone-based conferencing (電話会議) are at least partially present in cyberspace whenever they enter the audio-mediated, shared mental spaces provided by conventional telephony. Of course, mobile telephony is increasingly reliant upon the internet, making it possible to consider augmented, computer-mediated audio teleconferencing (コンピューターを使った電話会議)

Auditory display technology that attempts to provide users with a satisfying experience of virtual acoustical space is termed here "cyberspatial audio." Cyberspatial audio is distinguished from the broad range of spatial audio applications in a number of important ways that help to focus our group's research. Most significant is that cyberspatial audio is most often designed to be responsive to user inputs. In contrast to non-interactive auditory displays, cyberspatial audio display systems typically allow active exploration of the virtual environment in which users find themselves. Thus, at least some portion of the audio presented in a cyberspatial environment must be selected, processed, or otherwise rendered with minimum delay relative to user input. Besides the technological demands associated with realtime virtual acoustic rendering, the type and quality of auditory experiences supported can also very different from those associated with passive displays that do not support interactive sound processing. The research projects described in this review share the goal of developing efficient means for rendering virtual acoustical effects in order to minimize computational cost while maximizing audibility and identifiability of those effects. The desired result is typically a digital signal processing model (ディジタル信号処理のモデル) of low computational cost that can produce distinct auditory spatial images associated with identifiable virtual acoustic events, as verified through systematic perceptual evaluation.

1.2 Scope of Virtual Acoustic Research
(人工的な音響研究の範囲)

The scope of virtual acoustics under investigation by the Cyberspatial Audio Group includes the entire chain of events that starts with an abstract acoustical model of a virtual sound source, and ends with the human perception of virtual acoustic events associated with the behavior of that virtual sound source, as it sonically interacts with other modeled objects in a modeled virtual acoustic environment. The simulated chain of events must include a model of perception as well as a model of the physical space within which the listener is situated, since cyberspatial audio is always intended to produce distinct subjective experiences of virtual acoustic events. To fail to do so is analogous to displaying the results of graphic object simulation without regard for human visual sensitivities and capacities (potentially computing details that are undetectable by human observers, or conversely underestimating acuity).

The approach taken by the Cyberspatial Audio Group is to divide as much of the virtual acoustic simulation as possible into sequences of linear transformations, and then to study independently the perceptual response to variations in these transformations. Because these transformations are linear, they may be connected in cascade combination (縦続結合) without complicating interactions. Each transformation is implemented as a transfer function (TF) (伝達関数) via a DSP module. In applying these TFs to audio signal processing tasks required for realistic simulations of virtual acoustics, these transformations typically attempt to capture how the sonic behavior of simulated acoustic elements depends upon the position in space from which that behavior is observed. Clearly, sound that is generated in a virtual acoustical space should depend upon the location of the listener within that space. But virtual acoustical spaces can include anisotropic sound sources (radiating sound differently in different directions). And if there is a large object between a sound source and the listener, the sound received by that listener should be modified accordingly, providing an unambiguous clue to the listener that the source has been occluded, and not simply attenuated (as in [36]). This is to say that the spatial auditory image should present some convincing change in quality that, while not necessary realistic, affords recognition of sonic occlusion. (The term affords is used here in the Gibsonian sense [10].) Typically under investigation in any given study are only one or two of the following basic transformations.

1.3 Five Basic Transfer Functions
(5 つの基本的な伝達関数)

- Source Transfer Function (STF)
(音源の伝達関数)

— the spatially-dependent acoustical transformation of sound radiated from a sound source. Also called the "3D spatial radiation characteristic" for the sound source, this TF is exemplified by the change in a human speech signal observed as the angular orientation of the talker varies.

- Occlusion Transfer Function (OTF)
(妨害の伝達関数)

— the acoustical response of a sonically obstructing object interposed between two spatial positions, typically between a sound source and receiver (sink). The size, position, and orientation of the obstruction affects the magnitude and phase response of this TF (see [26] for details).

- Reflection Transfer Function (RTF)
(反射の伝達関数)

— the acoustical response of a sonically reflective object that generates a single, discrete reflection of its incident sound. As in the case of the OTF, the size, position, and orientation of the obstruction affects delay and gain in a frequency-dependent manner (again, see [26] for details).

- Enclosure Transfer Function (ETF)
(囲いの伝達関)

— the acoustical response of an enclosed space excited at a given location and measured at a given location. It captures the spatial configuration of talker and listener (source and sink) in relation to the boundaries of the enclosure (walls, ceiling, floor, etc.).

- Directional Transfer Function (DTF)
(方向の伝達関)

— the acoustical response to incident sound measured at the listener's ears. It is most often represented as a filter response describing the transformation of a sound as it arrives at the ear of the receiver from a given direction and distance, in which case the DTF is refered to as a "Head-Related Transfer Function" (HRTF). Note that this TF may also capture range-dependent variation in the DTF that is otherwise independent of range when the source is further than one meter from the listener [9].

In the next sections of this review, the methods and results of some representative research projects are briefly summarized, and the text assumes familiarity with the terms (transfer function names) defined above.

2. Research Projects
(研究プロジェクト)

2.1 Acoustic Research Facilities
(音響研究の設備)

The cyberspatial audio research space, also known as the Acoustical Measurement Laboratory (音響実験室), consists of two adjacent rooms, the University's large anechoic chamber (無響室) and an adjacent sonically-damped laboratory containing all the computers and electronic equipment required for acoustical measurement and recording. The anechoic chamber is also used in conducting subjective listening experiments, described in more detail below.

The lab is equipped with many specialized microphone systems, including the B&K HATS (Head And Torso Simulator) [5], a standard binaural microphone system (dummy head manikin) that can be used in two-channel encoding of spatial sound. The system also enables head-related acoustic measurements using a Type 2012 B&K Audio Analyzer. For example, the

B&K HATS has been used to measure HRTFs at high spatial density for nearby sound sources (located within arm's reach). Figure 1 shows the configuration of loudspeaker and HATS for one measurement from a set of 2520 such measurements filling the 3D space surrounding the dummy head (in contrast to conventional measures made most often at a fixed single range).

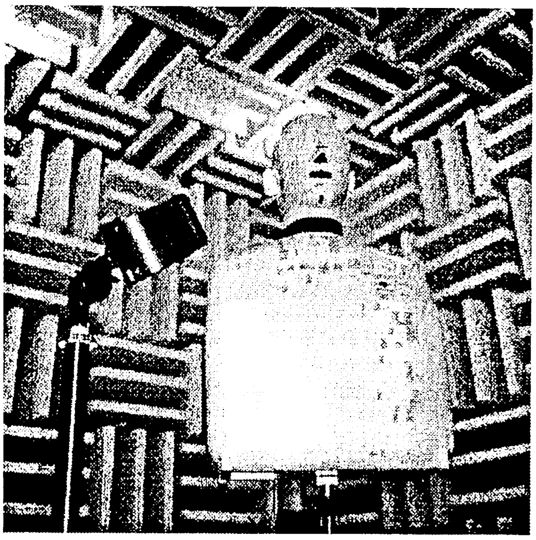

Figure 1: The B&K HATS (Head And Torso Simulator) configured with a Bose Acoustimass Cube (the CS-6J satellite loudspeaker) in the University of Aizu's anechoic chamber (無響室).The HATS dummy head system was rotated in 5 deg. increments using the B&K 5960 Controllable Turntable, and the loudspeaker was positioned at 7 elevation angles and 5 ranges, all within 1 m of the center of the head.

Measurements have also been made for the DTFs of other microphone systems, such as a human-head-sized spherical binaural system [21] and the sphere-mounted, four-microphone array that provides audio inputs for "HERO," a mobile HEaring teleRObot. "HERO" is designed to track sonic targets on the basis of computational auditory scene analysis (CASA). In addition to research on human spatial hearing, we are engaged in collaborative research with Prof. Jie Huang on 3D robotic hearing for the telerobot [13] (shown in the authors' previous JVRSJ review [4]) In related research, the telerobot is represented by its avatar (proxy) in cyberspace as part of a project studying mixed and augmented audio reality [39].

Many other acoustical systems have been measured in the University's anechoic chamber, and the results of those measurements have proven quite valuable in research and development of new virtual acoustic rendering technology. For example, the 3D directional dependence of the sound radiated from a clarinet was measured and captured in a set of STFs for spatial sound reproduction [40]. Also, the acoustic transformations associated with occluders and reflectors has been studied as a function of the position and angular orientation of the sonically-obstructing surface relative to the direct path between a sound source and receiver (sink). The primary goal for these measurements was the identification of the most salient acoustical features of the actual acoustical phenomena of interest [25]. Based upon hypotheses regarding which acoustical features are most salient, a simple OTF and RTF filtering model was designed to capture those features very efficiently [26]. In a subsequent study [20], the perceptual importance of the simulated acoustic features was tested in systematic subjective evaluation experiments, described below.

2.2 Subjective Evaluation Experiments (主観評価実)

Subjective evaluation experiments (using human listeners) executed by the Cyberspatial Audio Group can be classified according to whether headphone- or loudspeaker-based spatial sound reproduction is employed. In the group's headphone-based spatial sound reproduction experiments, measured or modeled HRTFs are typically employed in an effort to create externalized auditory imagery for the virtual sound sources under study. In the group's loudspeaker-based spatial sound reproduction experiments, the listener's actual HRTFs are typically allowed to transform the sound arriving from one or more loudspeakers, typically located within the anechoic chamber. Occasionally, an attempt is made to cancel out the listener's HRTFs so that novel DTFs may be evaluated for loudspeaker reproduction. This short review presents only a few representative examples of the types of subjective listening experiments often executed by the group. First, a study of headphone-displayed virtual sources at close-range is presented.

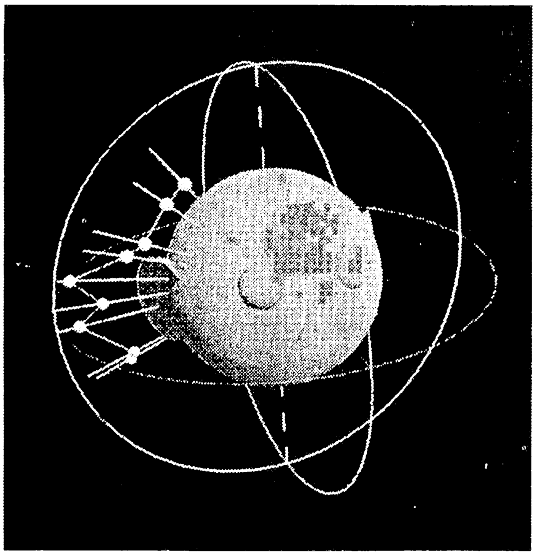

Figure 2: One subject's average judgments of the spatial location of four short speech sounds that were convolved with nine HRTFs measured at elevations ranging from 40 deg. below to 40 deg. above ear level, but always at 90 deg. azimuth. The subjects listened to these stimuli via eardrum-equalized STAX SR Λ ear-speakers [16] and indicated their apparent spatial location by adjusting the length and direction of a line segment in a computer graphic display under mouse control. The scale for apparent distance was in terms of head-radius units, and the reference circles in the plot define the sphere of two head-radius units centered on the reference head. Note that stimuli judged to be above ear level were also judged to be more frontally located than stimuli judged to be below ear level.

Figure 2 shows a graphic representation of source localization judgments for HRTF-based processing of virtual sources that vary in apparent azimuth, elevation, and range (see figure caption for details). Typically, personal, headphone-based auditory spatial display technology either fails to project (externalize) the auditory image of a virtual source to a location in the listener's auditory representation of the surrounding space (i.e., no externalization means no "out-of-head localization" [33]), or the source may be well externalized, but projected to a location at some greater distance from the listener, most often via the inclusion of a significant amount of reverberation that is easily detectable by the listener. Though all the average judgments shown in Figure 2 are externalized, the range of the sources varied with elevation angle. The goal of many of the related experiments executed by the Cyberspatial Audio Group, succinctly put, has been to confirm the effectiveness of an efficient means to control the range of an externalized virtual sound source, and enable placement so close to the listener's ear that it enters a listener's "personal space" [34].

When the externally-projected auditory image of a virtual sound source enters the listener's "personal space," a psychological boundary is crossed that potentially carries special meaning to users in particular applications such as teleconferencing in shared virtual acoustic environments. If such an audio transformation were properly engineered (both validated and calibrated perceptually), a spoken message could be made to sound as if it were whispered into the ear of the recipient, letting them know, for instance, that the message was intended for them in confidence (providing what has been termed a "whisper function"). Most recent results in this area have focussed upon control of source range with independent control over loudness [22][23]. The details of the audio signal processing are beyond the scope of this review, but have not changed appreciably since first described in 1984 by Kendall and Martens [15]. Only the signal processing components that control range for sources very close to the listener's ear are a novel contribution of the current research project. Related listening experiments have provided invaluable guidance in deploying appropriate audio signal processing technology for typical telecommunication scenarios. A full description of these user tests is also beyond the scope of this review, but the methods and results of these tests have been described by the first author (Martens) elsewhere [28].

Studies of virtual source localization have become quite common over the last ten years, but studies of other perceptual results of virtual acoustic rendering are much less common. New exploratory studies of the most salient dimensions of virtual acoustic rendering have begun (such as [29]), but more focussed confirmatory studies have also been executed. For example, studies of the identification of sonic occlusion on the basis of the changes in perceptual characteristics of a

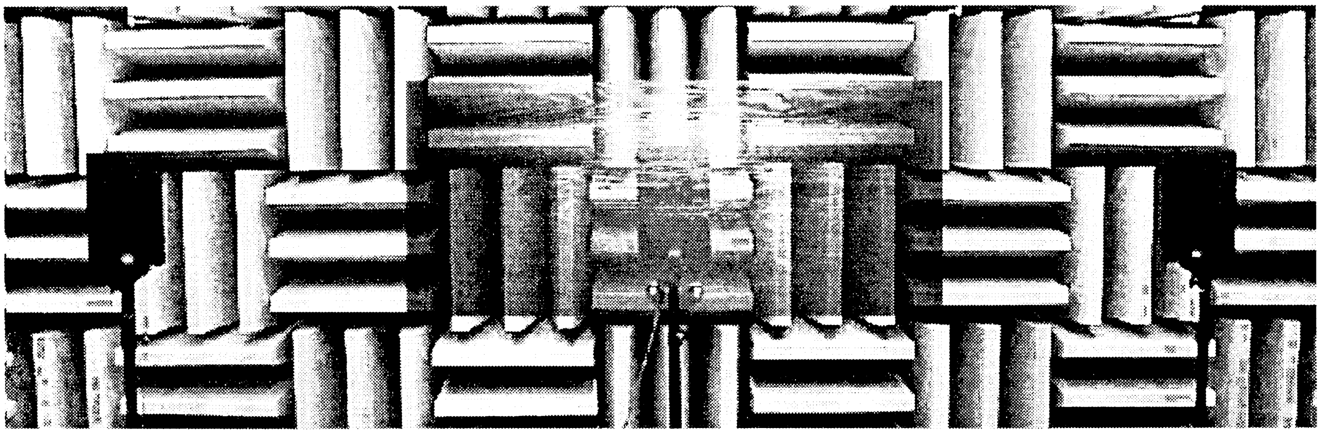

Figure 3:The three-loudspeaker array used for stimulus presentation in theperceptual study of simulated occlusion effects. A semi-transparent "virtual" occluder is pictured in front of the middle loudspeaker, which presented the test sound signal; the two lateralized loudspeakers presented spatialized reverberation.For this investigation, as the parameters of the simulated occlusion were varied, the spatiotemporal distribution of reflections comprising the reverberation was held constant. This very nearly matches the actual physical case, since the primary consequence of a simulated occluder interposed between source and receiver is the modification of the direct sound component that is propagated along a relatively direct path from source to receiver.

virtual source are practically non-existent. Though studies of how to simulate such occlusion effects can be found [38], systematic study of human ability to identify such simulated effects is limited indeed (see, for example, [20]). In recent studies, the likelihood of judging whether a virtual source is occluded or not was determined in the presence of spatialized reverberation while the level of sound source was allowed to vary from trial to trial. Figure 3 shows the configuration of three loudspeakers employed in these studies. The task was to identify whether the center loudspeaker in the anechoic chamber was placed behind an invisible "virtual" occluder, graphically rendered as semi-transparent in Figure 3. The results of these studies have lead to the development of a refined OTF model for simulating occlusion effects [26].

The Cyberspatial Audio Group continues with this type of virtual acoustic research, as the Masters research projects of students within the group are gradually addressing more and more of the issues that remain unresolved. For example, one current project (described in a recent ICAT proceedings paper [35]) asks listeners to judge which of six walls seems to be missing from a multichannel reproduction of simulated room reverberation (using a 3D ETF). The anechoic loudspeak array for this study is configured as a 4.2 channel reproduction system (where the ".2" means that two subwoofers were used). The four satellites in this system act as "acoustic windows" for mid- and high-frequency stimulation, and the two subwoofers are placed on either side of the listeners to allow low-frequency listener envelopment [11]. Listener envelopment (termed LEV) is one of the distinct perceptual attributes of auditory spatial imagery, defined as the sense of feeling surrounded by sound (音に包まれた感じ).

Using a realistic-sounding wall-reflection model (captured as a frequency-dependent RTF), various rooms were simulated. In the experiment, many paired comparisons were presented, the first stimulus of the pair being the standard (having all walls present) and the second one having identical values on all parameters except that one of the walls was missing from the simulation. The results generally showed that the location of missing walls is more easily identified in simulations containing higher numbers of reflections, and that the reduction of low frequency content in the reflections

1: The Japanese language expression provided here for LEV was taken from [31], in which it is reported that listeners can readily distinguish between LEV and auditory source width (ASW) (みかけの音源の幅). A short English-language version of this paper was also published [32]. As walls are selectively removed, changes in lev can allow listeners to identify differences in the simulated reverberant space.

also made identification less difficult. Other studies attempt to answer important questions about how to optimize multichannel sound reproduction based upon the extent of the variation in auditory spatial imagery supported by various loudspeaker configurations. For example, two studies have been completed that question whether only a single subwoofer channel is needed in two-channel and multichannel stereophonic sound reproduction systems, or whether two subwoofers are better than one. The conclusion reached was that presenting two low-frequency signals rather than one had an identifiable impact upon spatial impression (広がり観). Results revealed that the two most salient perceptual attributes underlying the subjective differences between the two reproduction modes were perceived auditory source width (ASW) (みかけの音源の幅) and perceived auditory source distance (ASD) (みかけの音源の距離).

Other subjective listening experiments have been executed in the "Synthetic World Zone" (人工世界ゾーン) of the university's Multimedia Center (マルチメディア センター), where a hemispherical array of 15 loudspeakers is co-located with a wide-angle 3D stereoscopic image theater (3D 立体映像 シアター). The audio portion of this spatially immersive display system is based upon a 3D acoustic model (立体音響のモデル), and created using the PSFC (Pioneer Sound Field Controller), which was described in the March 1998 issue of TVRSJ [1].

A recently completed research project in this space included the multidimensional psychophysical calibration of some of the display parameters of the PSFC [12]. As accurate control of virtual source range was confounded by variations in both the liveness parameter and in overall PSFC channel volume, an empirical approach was employed to derive a Look-Up Table (LUT) inverting average range estimates obtained from a group of human subjects who listened to a set of virtual sources (short spatialized speech samples).

2: Note that an additional, alternative spatial sound reproduction system has recently been installed in the "Synthetic World Zone" that combines four Roland RSS-10 units to drive eight loudspeakers. Also note that Roland Corporation has licensed the sound spatialization technology developed and patented by the first author (Martens) and colleague G. Kendall [14].

2.3 Acoustic Events Modeling (アクースティックイベントモデリング)

All of the subjective evaluation experiments (主観評価実験) described above are leading up to an improved general model for describing acoustic events that takes human perception of those events into account. Along with related work in other labs, most notably that of Tohyama (see, for example, [37]), this research is contributing to a human-centered foundation for virtual acoustic rendering [18]. In contrast to previous approaches which have their basis in physical models, such as the first author's proposed sound "nodes" [17] for audio services in the Virtual Reality Modeling Language (VRML), this approach is more consistent with the goals of Tohyama's Acoustic Events Modeling Language (AEML) [37].

3. The Future of Cyberspatial Audio (サイバースペーシャルオーディオの未来)

The future of cyberspatial audio, as seen from our group's perspective, is the continuation of the shift from the currently poor auditory interface to cyberspace, to more and more immersion in computer-mediated virtual acoustic worlds. As the audio component of popular media becomes more interactive, it will also come to support telepresence and highly realistic auditory imagery. Over 35 years ago, Marshall McLuhan promulgated the revolutionary statement "the medium is the message" and ushered in the "Age of Information" (a term also coined by McLuhan). His book "Understanding Media: The Extensions of Man" [30] begins with the recognition that technological tools shaped by humans for human use have in fact shaped their human users into inhabitants of what has come to be called cyberspace:

"During the mechanical ages we had extended our bodies in space. Today, after more than a century of electric technology, we have extended our central nervous system itself into a global embrace, abolishing both space and time." ([30], p. 3)

Besides making the world a smaller place, however, virtual acoustic rendering technology will change the "form" of the audio medium significantly, so that not just the features expected in natural environments are supported, but also augmented cyberspatial audio features [27][39] will enrich in new ways the experiences of users situated within the collective virtual worlds that constitute cyberspace.

References

[1]天野 克巳, 松下 文雄, 柳川 博文, 公園 マイケル, イエンツ ヘルダー, ウイリアム L. マーテンス, 木場 美治, and 東山 三樹夫.- A Virtual Reality Sound System Using Room-Related Transfer Functions Delivered Through a Multispeaker Array: The PSFC at the University of Aizu Multimedia Center.TVRSJ: 日本バーチャルリアリティ学会論文誌 (Trans. of the Virtual Reality Society of Japan), 3(1):1-12, March 1998.ISSN 1344-011X.

[2]マイケル 公園, 伊苑津 ヘルダー, and ウイリアム L. マーテンス. 技術. J. Acous. Soc. Jap., 55(10):730-731, October 1999.ISSN 0388-2861; www.u-aizu.ac.jp/~mcohen/welcome/publica tions/JASJ-reviewJ.ps.

[3]公園 マイケル and 佐々 健太. In Proc. WESTPRAC VII: 7th Western Pacific Regional Acoustics Conf., pages 321-324, 熊本, October 2000.

[4]公園 マイケル and ウイリアム L. マーテンス.、JVRSJ: J. of the Virtual Reality Society of Japan, 6(2):52-57, September 2001. ISSN 1342 6680; www.u-aizu.ac.jp/mcohen/welcome/publications/spatial-media.ps.

[5]Brüel and Kjær. Head and torso simulator, type 4128.

[6]M. Cohen, J. Herder, and W. L. Martens. Cyberspatial audio technologies. J. Acous. Soc. Japan (E), 20(6):389-395, 1999.

[7]M. Cohen.A design for integrating the internet chair and a telerobot. In Proc. IS2000: Int. Conf. on Information Society in the 21 Century, pages 276-280, Aizu-Wakamatsu, Japan, November 2000. IPSJ, IEICE, IEEE.

[8]M. Cohen.The Internet Chair. IJHCI: Int. J. of Human-Computer Interaction, 2001. www.u-aizu.ac.jp/~mcohen/welcome/publications/ic4.ps. In press.

[9]R. O. Duda and W. L. Martens.Range dependence of the response of an ideal rigid sphere.J. Acous. Soc. Amer., 105(5):3048-3058, 1998.

[10]J. J. Gibson.The Senses Considered as Perceptual Systems. Houghton-Mifflin, Boston, MA, 1966.

[11]D. Griesinger. The psychoacoustics of apparent source width, spaciousness andenvelopment in performance spaces. Acta Acustica, 83:721-731, 1997.

[12]K. Honno, W. L. Martens, and M. Cohen. Psychophysically-derived control of source range for the Pioneer Sound Field Controller. In Proc. 110 Conv. of the Audio Engineering Society, Amsterdam, Netherlands, May 2001. Preprint 5313.

[13]J. Huang, T. Supaongprapa, I. Terakura, F. Wang, N. Ohnishi, and N. Sugie. A model based sound localization system and its application to robot navigation.Robotics and Autonomous Systems (Elsevier Science), 27(4): 199-209, 1999.

[14]G. Kendall and W. L. Martens.Spatial reverberation. U.S. Patent 4,731,848, 1988. (also European patent specification 0 207 084 B1).

[15]G. S. Kendall and W. L. Martens.Simulating the cues of spatial hearing in natural environments. In Int. Computer Music Conf., pages 111-126, Paris, 1984. Computer Music Association.

[16]W. L. Martens. Directional hearing on the frontal plane: Necessary and sufficient spectral cues. PhD thesis, Northwestern University, 1991. Department of Psychology

[17]W. L. Martens.Proposed VRML 1.x audio specification, Audio Advisory Board to the VRML Architecture Group, Sep. 15 1995. http:// www.u-aizu.ac.jp/~wlm/audprop.html.

[18]W. L. Martens.Human-centered design of spatial media display systems. In International Workshop on Human Interface Technology, pages 1-12, Aizu-Wakamatsu, Japan, October, 1998.

[19]W. L. Martens.The impact of decorrelated low-frequency reproduction on auditory spatial imagery: Are two subwoofers better than one? In Proc. Audio Engineering Society 16th Int. Conf. on Spatial Sound Reproduction, pages 67-77, Rovaniemi, Finland, 1999.

[20]W. L. Martens.Efficient auralization of small, cluttered spaces: Simulating sonic obstructions at close range. In S. Kuwano and T. Kato, editors, Proc. 7th Western Pacific Regional Acoustics Conf., pages 317-320, Kumamoto, Japan, Oct., 2000. Acous. Soc. Japan.

[21]W. L. Martens. Pseudophonic listening in reverberant environments: Implications foroptimizing auditory display for the human user of a telerobotic listening system. In Proc. IS2000: Int. Conf. on Information Society in the 21 Century, pages 269-275, Aizu-Wakamatsu, Japan, November 2000. IPSJ, IEICE, IEEE.

[22]W. L. Martens. Binaural range display with indepen-

dent control over loudness: Presenting nearby whispers in the dark. In Proc. ICAT: Int. Conf. Artificial Reality and Tele-Existence, pages 100-109, Tokyo, December 2001.

[23]W. L. Martens. Psychophysical calibration for controlling the range of a virtual sound source: Multidimensional complexity in spatial auditory display. In Proc. Int. Conf. on Auditory Display, pages 197-207, Espoo, Finland, 2001. ICAD, Keynote address.

[24]W. L. Martens. Two-subwoofer reproduction enables increased variation in auditory spatial imagery. In International Workshop on Spatial Media, pages 86-97, Aizu-Wakamatsu, Japan, October 2001.

[25]W. L. Martens and J. Herder. Perceptual criteria for eliminating reflectors and occluders from the renderingof environmental sound. In Proc. 137th Meeting, Berlin, 1999. the ASA and the 2nd Conv. of the EAA.

[26]W. L. Martens, J. Herder, and Y. Shiba.A filtering model for efficient rendering of the spatial image of an occluded virtual sound source. In Proc. 137th Meeting, Berlin, 1999. the asa and the 2nd Conv. of the EAA.

[27]W. L. Martens and A. Yoshida. Augmenting spoken telecommunication via spatial audio transformation. The Journal of Three Dimensional Images, 14(4):169-175, 2000.

[28]W. L. Martens and A. Yoshida. Psychoacoustically - based control of auditory range: Display of virtual sound sources in the listener's personal space. In Proc. IS2000: Int. Conf. on Information Society in the 21 Century, pages 288-294, Aizu-Wakamatsu, Japan, November 2000. IPSJ, IEICE, IEEE.

[29]W. L. Martens and N. Zacharov Multidimensional perceptual unfolding of spatially processed speech I: Deriving stimulus space using INDSCAL. In Proc. 109th Conv. of the Audio Engineering Society, Los Angeles, Sept. 2000. Preprint 5224.

[30]M. McLuhan. Understanding Media: The Extensions of Man. The MIT Press, Cambridge, Massachusetts, 6th edition, 1994. (first published in 1964).

[31]M. Morimoto, H. Fujimori, and Z. Maekawa.みかけの音源の幅と音に包まれた感じの差異 J. Acous. Soc. Japan, 46(6):449-457, 1990. [Discrimination between auditory source width and envelopment].

[32]M. Morimoto and Z. Maekawa. Auditory spaciousness and envelopment. In Proc. 13th Int. Congress on Acoustics, pages 215-218, Belgrade, Yugoslavia, 1989.

[33]N. Sakamoto, T. Gotoh, and Y. Kimura.On "out-of-head localization" in headphone listening. J. Audio Eng. Soc., 24:710-715, 1976.

[34]R. Sommer. Personal Space - The Behavioral Basis of Design. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1969.

[35]K. Suzuki and W. L. Martens. Subjective evaluation of room geometry in multichannel spatial sound reproduction: Hearing missing walls in simulated reverberation. In Proc. ICAT: Int. Conf. Artificial Reality and Tele-Existence, page 9, Tokyo, December 2001.

[36]T. Takala and J. Hahn. Sound rendering.Computer Graphics, 26(2):211-220, July 1992.

[37]M. Tohyama, M. Kazama, and H. Yanagawa. Acoustic events modeling for rendering 3D sound and perception. In Proc. World Multiconference on Systemics, Cybernetics, and Informatics, volume VI, pages 444-448, 2000.

[38]N. Tsingos and J.-D. Gascuel. Fast rendering of sound occlusion and diffraction effects for virtual acoustic environments. Proc. Audio Engineering Society 104 Conv., May 1998. Preprint No. 4699.

[39]Yasuhiro Yamazaki, Michael Cohen, Jie Huang, and Tomohide Yanagi. Augmented audio reality: compositing mobile telerobotic and virtual spatial audio. In Proc. ICAT: Int. Conf. Artificial Reality and Tele-Existence, Tokyo, December 2001.

[40]S. Yasuda. Modeling spatial radiation characteristics of the clarinet. Master's thesis, University of Aizu, 1998.