

to correct for the effects of platform and target motion, and for ocean instabilities. The processing is tuned to the specific environment by setting the cross-correlation window size based on the two-way group arrival structure of the channel. The aliasing that results from the undersampled aperture is reduced to acceptable levels through the appropriate use of cross-range matched field processing. Images formed by processing the data from a forward-look synthetic aperture measuring 2200 wavelengths with a mean target distance of 9.5 km are presented.

### Contributed Paper

11:30

**4aUW9. Array measurements and characterizations with sources in motion.** William M. Carey (Dept. of Aersp. and Mech. Eng., College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, wcarey@bu.edu)

Traditional processing has treated space and time as independent random variables with assumed ergodicity; however, dynamic-noise fields are produced by moving sources. In these cases, can velocity and range information be used to increase gain, cancel noise, and improve resolution? High-resolution arrays have measured the dynamic ambient noise fields at low frequencies, over long time intervals and at basin/margin spatial

scales, and have resolved/tracked shipping. This paper summarizes experiments with arrays which used source-relative velocity as a processing variable, that is, space-time processing. Results from a closed-basin shipping dominated noise field showed the relative velocity acted as a filter that improved the signal-to-noise ratio for sources with known relative velocity and reduced it for sources having different relative velocities. Results from a shallow water experiment with a vertical array, moving source, and the Bessel transform were shown to determine the wave number spectrum and were comparable to the modeled spectrum in both amplitude and phase. Thus the contention that improved array ship-noise measurement and cancellation is possible if the velocity is known may have merit, especially if modal variations can be adaptively modeled.

FRIDAY AFTERNOON, 2 JUNE 2000

AMERICAN ROOM, 2:30 TO 4:50 P.M.

### Session 4pAA

## Architectural Acoustics: Computer Modeling and Measurement Techniques for Large Room Acoustics

Paul T. Calamia, Chair

*Kirkegaard & Associates, 801 West Adams Street, 8th Floor, Chicago, Illinois 60607*

Chair's Introduction—2:30

### Contributed Papers

2:35

**4pAA1. An echoic room model for acoustic simulations by non-reflection boundaries using the Bergeron method.** Hidemaro Shimoda (Inst. of Technol., Shimizu Corp., 3-4-17 Etchujima Koto-ku, Tokyo 135, Japan, shimoda@sit.shimzu.co.jp)

An analytical model of anechoic rooms by distributed equivalent circuits is constructed in the computer using the Bergeron method [Shimoda *et al.*, "Analysis of sound fields in rooms using Bergeron's method," Trans. I.E.C.E. Japan, 72 A, 1, 1–11 (1989); English transl., 1989 Electronics and Communications in Japan, Part 3, Vol. 72(12), Scripta Technica, Inc. (Wiley, New York, 1990)]. In this analytical method, the non-reflection boundaries are modeled by gradually increasing sound transmission losses in specific regions near boundaries just like a sand beach exposing sea waves. Also, the sound transmission characteristics or the sound pressure distributions in the room are investigated. The model can be applied to the sound scattering analysis by an arbitrary object.

2:50

**4pAA2. From auditorium ray acoustics to wave acoustics: The next giant step.** Herman Medwin (OCEANAC Assoc., 4021 Sunridge Rd., Pebble Beach, CA 93953)

For at least five decades, architectural acousticians have used straight-line, infinite-frequency rays to determine which regions of an auditorium areinsonified by a sound source and which are shadowed. In the next decade, leading acoustical designers of auditoriums will turn to predictions of the timing and amplitude of the different frequencies which partially reflect and bend around real auditorium components such as proscenium arches and boxes. The frequency dependence of wave reflection from rigid, finite surfaces has been developed by Clay *et al.* [J. Acoust. Soc.

Am. 94, 2279 (1993)]. The calculation of the frequency dependence of the diffraction component, which some have called the BTM technique, is a digitally formulated, finite-wedge, frequency-spectral interpretation of the classic Biot–Tolstoy (1957) theory for impulse scatter from an infinite wedge. The BTM technique, which we described first at an ASA meeting in 1978, has been applied to shadowing by highway noise barriers [Medwin, J. Acoust. Soc. Am. 69, 1060–1064 (1981)] and to multifrequency reverberation from randomly rough, ocean surfaces [Keiffer and Novarini, in *Computational Acoustics*, edited by D. Lee *et al.* (1990), Vol. 1, pp. 67–81]. The procedure is explained in detail in Medwin and Clay [*Fundamentals of Acoustical Oceanography* (Academic, New York, 1998)]. The justification and analytical extension of this method have recently been formulated by Svensson [J. Acoust. Soc. Am. 106, 2331–2344 (1999)], who has made the program available on his web site and has suggested applications to auditoriums. The physical interpretation of the BTM technique will be described in terms of the spectral and temporal behavior of finite plates and wedges as modules of more complex components of auditorium surfaces.

3:05

**4pAA3. Subjective relevance of objective measures for spatial impression.** Lily M. Wang and Anders C. Gade (Dept. of Acoust. Technol., Tech. Univ. of Denmark, Bldg. 352, DK-2800, Lyngby, Denmark)

Several objective measures have been proposed to describe the feeling of spatial impression in concert halls, including Lateral Energy Fraction (LF) and Interaural Cross-Correlation Coefficient (IACC). However, previous studies have shown that LF and IACC values did not highly correlate with each other at individual seat positions in real halls [J. S. Bradley, J. Acoust. Soc. Am. 96, 3525–3535 (1994)]. To investigate the listener

envelopment aspect of spatial impression further, subjective paired-comparison tests have been run using signals which have various values for LF, early IACC (from 5–80 ms), late IACC (from 80 ms–1 s), and late lateral relative sound level (GLL). Another proposed measure, called Inter-aural Level Fluctuations (IALF), has also been included, which is based on the rate of change over time of the level difference between the ears. The binaural test signals were generated by the room acoustics simulation program ODEON, using four different musical motifs, and were presented to test subjects via headphones. Results from the experiments are presented. [Work supported by ASA Hunt Postdoctoral Research Fellowship.]

3:20

**4pAA4. Measurements of the subjective acoustical parameters on an orchestra platform.** Jin Yong Jeon (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

This paper deals with the relationship between the subjective room-acoustical parameters and the objective properties of the sound field on an orchestra platform. A set of acoustical parameters is investigated and subjective preference tests are conducted both on a real platform and a simulated platform of a concert hall. Parameters EDT, RT, C80, D, Ts, 1-IACC (or LEF) are used for assessments of the quality of the stage acoustics. Advantages and limitations of computer simulations versus field experiments are discussed and both experiments are reviewed, in which the EDT/RT and spatial impressions are investigated. From the subjective preference tests, it is found that musicians form their stage acoustic judgments based on their perception of reverberance, clarity, envelopment, loudness, warmth, brilliance, and overall impression. In addition, the relative reverberance EDT/RT is revealed to have high correlation with the subject-averaged judgments and the spatial impression indices.

3:35

**4pAA5. Music-masked maximal-length sequences for auditorium acoustic measurements.** Ning Xiang and Charles Sabatier (Natl. Ctr. for Physical Acoustics, Coliseum Dr., Univ. of Mississippi, University, MS 38677, nxiang@olemiss.edu)

Maximal-length sequence (M-sequence) measurement technology has gained acceptance within the architectural acoustics community for quite some years. Owing to the high noise immunity and computational efficiency of the fast M-sequence transform, acoustic qualities in auditoria can be determined quickly and with satisfactory precision. The present work investigates the use of music signals to mask M-sequence excitations. The inherent cross-correlation mechanism of the fast M-sequence transform makes it possible to determine room impulse responses with a positive music/M-sequence ratio (expressed in dB). The feasibility, limitations, and potential applications for using music-masked M-sequences in taking auditorium impulse responses measurements will be discussed in the presentation.

3:50

**4pAA6. Acoustic design for University of Southern California, Annenberg School of Communications, lecture hall remodel.** Fiona Gillan (Arup Acoustic, 901 Market St., Ste. 260, San Francisco, CA 94103) and Nathan Sevener (Arup Acoustic, Los Angeles, CA 90064)

In 1998, the redesign of the 250-seat lecture hall at the USC Annenberg School of Communications was undertaken to update the aesthetics of the 1974 design and meet the modern functional requirements of the University. Primary uses of the hall were to be lectures, audiovisual presentations, and panel discussions. The existing fan-shaped space had a concrete ceiling, plastered concrete side walls, and a concave rear wall with a slatted wood acoustic finish. The existing reverberation time (RT60) was 0.95 s at mid-frequency, but there was a sense of over-reverberance and a lack of speech clarity. This was attributed to focusing by the curved rear wall and to excessive bass reverberation. The acoustic design targets included a modest reduction of the mid-frequency RT60, a significant reduction of the bass reverberation, and mitigation of the ef-

fects of focusing, as well as a maximum noise level of NC30 for the new HVAC system. Simple materials, including gypsum board wall linings, fabric wrapped panels, and acoustic foam in inconspicuous areas, were used to achieve cost effectiveness. Acoustics and architecture were fully integrated for efficiency of design. The project, completed in 1999, was an acoustic and architectural success. [Architectural Design by Siegal Diamond Architecture.]

4:05

**4pAA7. Some acoustic considerations in the use of Kerto-wood in the construction of Sibelius Hall.** Christopher A. Storch (Artec Consultants, Inc., 114 West 26th St., 9th Fl., New York, NY 10001-6812)

Sibelius Hall is a 1300-seat concert hall and congress center which was built by the City of Lahti, Finland, with the sponsorship of the Finnish wood industry. The wood industry's support in this construction project was conditional based on using wood in new and experimental ways, not only as a finish material, but as a structural material. A proprietary product called Kerto-wood was used in a wide variety of structural and finish applications. Sand-filled Kerto-wood panels 300 mm thick were used in the slanted exterior walls. Details of the sand-filled panels' construction, the joining details, and the transmission loss testing of the assembly will be discussed. Prefabricated Kerto-wood elements were also used to construct the ceiling, balconies, and reverberation chamber doors. Discussion of the reflective properties will also be presented.

4:20

**4pAA8. Adjustment of reverberation in the rehabilitation of theaters. Its application to the architectural patrimony of theaters in Andalusia, Spain.** Angel Luis Leon, Jaime Navarro, and Juan Jose Sendra (School of Architecture of Seville, Av. Reina Mercedes, 2 Sevilla 41012, Spain, leonr@arquitectura.us.es)

One hundred years after its formulation by W. C. Sabine, the reverberation time continues to be, we feel, the one most useful indicator for architects when undertaking the rehabilitation of theaters in which the main objective is to obtain good acoustic conditions. The reverberation time is easily quantifiable during the planning phase, thus proving to be of great interest in the process of theater design. The adjustment of reverberation times in a theater to the different frequencies at intervals regarded optimum is considered to be an indispensable determinant, though not in itself sufficient, to obtain a good acoustics. This adjustment is intimately related to the decisions proper to the work of the architect in interior design, such as the election of surface materials or the volume of the locale. The authors of this study, who have participated in the rehabilitation of 35 theaters in Andalusia as acoustic consultants, propose to relate the results of their work with the different proposals of the principal researchers in room acoustics, about optimum values of reverberation, with the aim of obtaining a series of conclusions to offer to professionals of rehabilitation.

4:35

**4pAA9. Analysis of spatial acoustic qualities from listening in rooms.** Martin A. Gold (321 ARCH, Dept. of Architecture, Univ. of Florida, Gainesville, FL 32611, mgold@ufl.edu)

Statistical analysis was conducted on data from in-room listening studies to evaluate the effects of spatial impression and auditory source width as independent variables with regard to overall acoustic impression. The listening studies were conducted with groups of listeners (30+) in the rooms under evaluation. Subjects in each circumstance were asked to evaluate the acoustic environment using a bi-polar semantic rating scale for the acoustical qualities of overall impression, loudness, clarity, reverberance, treble response, bass response, spatial impression, auditory source width, echoes, and background noise. Ratings were taken in a

variety of listening environments ranging from outdoors to a 2000-seat auditorium to elicit a range of qualitative assessments. Both live music and recorded music and speech sources were evaluated in the study. The study also includes a comparison of in-room ratings. Results indicate that, in

most cases, it is difficult to statistically isolate independent spatial variables with regard to overall room ratings in complex environments. In specific room locations, these variables appear to be statistically independent.

FRIDAY AFTERNOON, 2 JUNE 2000

CANADIAN ROOM, 2:30 TO 5:10 P.M.

## Session 4pAO

### Acoustical Oceanography: Fine Scale Ocean Processes, Bubbles and Internal Waves

Grant B. Deane, Chair

*Marine Physical Laboratory, Scripps Institution of Oceanography, University of California—San Diego, La Jolla, California 92093-0238*

Chair's Introduction—2:30

#### Contributed Papers

2:35

**4pAO1. Nonlinear sound scattering by subsurface bubble clouds in sea.** Lev A. Ostrovsky (Univ. of Colorado, CIRES/NOAA ETL, 325 Broadway, Boulder, CO), Zinovy Kluzek (Polish Acad. of Sci., Sopot, Poland), Alexander M. Sutin, Irina A. Soustova, Alexander I. Matveev, and Andrey I. Potapov (Russian Acad. of Sci., Nizhny Novgorod, Russia)

This presentation is based on a series of field experiments performed in a shallow sea (Baltic Sea). Investigated were bubble clouds of two types: one created in a motorboat's jet, and the other existing naturally in the subsurface layer. A two-frequency signal (30 and 35 kHz) was radiated from below (at a 15-m depth) and received upon backscattering at sum and difference frequencies (65 and 5 kHz). The intensity of the scattered signal averaged over five subsequent pulses was measured. Theoretical consideration is given for reverberation at sum frequency and coherent propagation of the difference-frequency wave with subsequent reflection from the surface. The method is based on integration of fields nonlinearly scattered by single bubbles. A good qualitative and, in many cases, quantitative agreement was obtained between theory and experiment. A possibility to use nonlinear scattering for bubble distribution evaluation is also discussed.

2:50

**4pAO2. High-frequency propagation through the wave boundary layer.** David Farmer, Svein Vagle (Inst. of Ocean Sci., 9860 W. Saanich Rd., Sidney, BC V8L 4B2, Canada, farmerd@dfo-mpo.gc.ca), and Grant Deane (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0230)

The surf zone presents one of the most extreme and challenging environments for the study of shallow water propagation where the complexities of a sloping beach, wave amplification, wave breaking, and the presence of dense bubble clouds combine to block or modify acoustic signals in a variety of ways. Propagation measurements at 12 and 100 kHz reveal sporadic drop-outs during which the signal can be abruptly blocked for 1–5 min or may be weakly detected but with strong modulation by the surface waves. Bubbles introduced to the water column through breaking are subject to persistent vertical mixing through the turbulent wave boundary layer which becomes progressively more energetic closer inshore. The bubble suspension is further modified by buoyancy and dissolution effects. Different acoustic frequencies are attenuated by bubbles of different radii, which in turn respond differently to the competing influences of turbulent mixing, buoyancy, and dissolution. We explore the implication of these effects on high-frequency propagation with a model that includes bubble injection, vertical mixing by the wave boundary layer, and bubble cloud

evolution. The predictions are compared with observations of high-frequency propagation, wave amplification, wave breaking, and bubble distributions. [Work supported by the US ONR.]

3:05

**4pAO3. A maximum likelihood estimation algorithm for tracking two multipaths.** Alan Barton and Daniela Di Iorio (Dept. of Marine Sci., Univ. of Georgia, Athens, GA 30602)

A high-frequency (67 kHz) acoustical scintillation experiment was carried out in the northern entrance to Hood Canal, Puget Sound, WA. This experiment made use of a 4-transmitter and 4-receiver array configured as a T-shape. The two-dimensional feature of this array was designed to measure both along-channel small-scale properties as a result of advection and vertical properties as a result of refraction from temperature/salinity stratification. Sampling was carried out at 5 Hz for all transmitters so as to understand the turbulent effects of the medium on acoustic scattering. This experiment tested if acoustical scintillation techniques can be extended to long distances (2380 m) and in deep (100 m) coastal channels having strong tidal currents and temperature/salinity stratification. With the longer path lengths and stable stratification, acoustic propagation resulted in multipath arrivals which were separable for most of the measurement period. In order to study the oceanographic effects on the direct path signal at 20-m depth and the upward refracted signal into the surface mixed layer, a maximum likelihood estimation algorithm is developed that tracks the two paths and calculates amplitude, phase, and travel time for each. [Experiment was carried out by SAIC of Bellevue, WA, and student support was from the University of Georgia.]

3:20

**4pAO4. The time-domain behavior of coupled resonator scatterers.** C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004, cf@nrlssc.navy.mil)

In previous work [J. Acoust. Soc. Am. **98**, 1178–1190 (1995)] a coupled oscillator formalism was introduced for describing collective resonances, scattering, and superresonances of multiple air bubbles in a fluid. Subsequent time-domain investigations, of the impulse response of coupled systems, have disclosed the conditions which exactly determine whether the ensemble behavior is properly described using: (a) a multiple scattering paradigm or (b) a collective modal methodology. The important criteria are the Q of the individual scatterers, and their typical separations in the medium. For highly damped or sparse systems, e.g., scattering from loose schools of swimbladder fish, or from a gassy seabed containing entrained bubbles, the scatter counting approach may be applicable. For

more strongly coupled systems, e.g., a dense cloud of resonating bubbles in the water column, the individual scattering events are typically not resolvable and a modal approach must be used. The result has implications for both volume and bottom-scattering applications. Theoretical and numerical analysis will be presented. [Work supported by ONR.]

3:35–3:55 Break

3:55

**4pAO5. Wideband determination of source functions using the Faran elastic sphere scattering model.** C. Feuillade, R. W. Meredith (Naval Res. Lab., Stennis Space Center, MS 39529-5004, cf@nrlssc.navy.mil), N. P. Chotiros (Univ. of Texas, Austin, TX 78713-8029), and C. S. Clay (Univ. of Wisconsin—Madison, Madison, WI 53706)

During August 1998, a bottom scattering tank experiment was performed at ARL-UT to measure wideband reverberation from multiple objects (e.g., cobbles and pebbles) placed on a sediment simulation of the sea floor. As a first step in this experiment, a procedure was adopted to calibrate the source transfer function by measuring reflected signal pings from the water surface (which has a spike impulse response), with the transducer directed vertically upwards. Subsequent analysis showed that the strength of the reflections may have induced a nonlinear receiver response in the transducer. Fortunately, time-domain scattering measurements had also been made with a stainless steel ball suspended in the water column. These data were used to study the source drive by deconvolving the theoretical impulse response of a steel sphere, obtained via the Faran scattering model, from the scattered time signals. Both the characteristics of the experimental source signal and another, low-power, source measured later were examined. Results show the critical importance of accurate input–output system calibrations for wide signal bandwidth sound scattering research, and have implications for boundary and volume scattering applications. Experimental data and theoretical numerical analysis will be presented. [Work supported by ONR.]

4:10

**4pAO6. Simultaneous measurements of the bubble size distributions and radiated noise from open-ocean whitecaps.** Grant Deane and Dale Stokes (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238, grant@mpl.ucsd.edu)

It is now widely acknowledged that wind-induced oceanic ambient noise comes from ringing bubbles formed by breaking waves at the ocean's surface. The numbers, sizes, creation rates, and spatial distribution of these bubbles are all important in determining the acoustic radiation signature of a breaking wave. Results from an experiment to measure the bubble size distribution within open-ocean whitecaps and simultaneously record the burst of sound radiated by forming whitecaps will be presented. The bubble size distribution measurements were made with an optical instrument mounted within 0.5 m of the ocean's surface beneath a surface-following frame. The instrument is capable of resolving the time-evolving bubble size distribution on millisecond time scales and centimeter length scales within a whitecap. The simultaneous optical and acoustical measurements are important for modeling the sound of breaking waves in the open ocean and the surf zone. [Work supported by ONR and NSF.]

4:25

**4pAO7. Results of an acoustic normal-mode analysis of SESAME II data, and conclusions on internal wave activity.** Jacob George and Robert L. Field (NRL Code 7185, Stennis Space Center, MS 39529)

An acoustic normal-mode analysis has been used to examine the effect of internal waves (IWs) both at the receiving array and along the path of acoustic transmission in SESAME II data (year 1996; off the coast of

Scotland). Power transfer from lower to higher modes has been observed to occur only intermittently. The sound-speed profiles (SSPs) at the receiving array during these occurrences are found to be nearly identical. Fourier analysis has revealed that a prominent peak in the frequency band 39–47 cycles/day in the acoustic reception at 800 Hz corresponds to a similar peak in the same frequency band in the thermal data at the receiver. That coincidence was investigated through normal-mode analysis, IW time delay between two nearby vertical arrays of thermal sensors, and a cylindrical wavefront model. The results show that the two peaks could not have been caused by the same IW. Therefore, the IW responsible for the acoustic peak should have originated farther north. [Work supported by ONR and NRL.]

4:40

**4pAO8. Internal waves and temperature inversion in the East China Sea.** Tao Wang (Eng. Sci. 104-44, California Inst. of Technol., Pasadena, CA 91125), Tian-Fu Gao, De-Jun Jiang, and Yun-Peng Zhang (Chinese Acad. of Sci., Beijing 100080, PROC)

Several shallow-water acoustics experiments were performed in the Zhoushan offshore area of the East China Sea from 1996 to 1998. The water depth is about 55 m at the mooring locations. Temperature chains sampling once per 6.4 s or 1 min provide the oceanographic data. One clearly sees the long-wavelength internal tides. The dominant component of the internal wave field is the semidiurnal tide of 10-m height. High-frequency internal waves with 10–15 min duration and 2-m amplitude ride on top of the internal tide. These rank-ordered waves are likely internal solitons, which is validated by recent SAR images. The asymmetry of the initial and final states indicates that the solibore description is somewhat applicable to this case, but more evidence is needed. The inversion phenomenon of temperature and salinity frequently occurs in the long and narrow East China Sea area off the coast of Jiangsu and Zhejiang. Temperature inversion is noted via the measurements and shows yearly variations. Its structure and generation mechanism are discussed. [Work supported by NSFC.]

4:55

**4pAO9. Higher-order spectral analysis of shallow-water internal waves.** Tao Wang (Eng. Sci. 104-44, California Inst. of Technol., Pasadena, CA 91125) and Tian-Fu Gao (Chinese Acad. of Sci., Beijing 100080, PROC)

Higher-order spectral analysis is a useful means for investigating nonlinear properties of shoaling random surface waves and a wide range of other phenomena, but has been considered to be fruitless for internal waves based on some studies in which the energies are represented by the deep-ocean GM model. However, it has been increasingly recognized that nonlinearity manifests itself in various shallow-water internal wave phenomena. The classical phase-blind power spectra contain information in the second-order statistics of a signal, and only suffice for describing Gaussian processes. Consequently, there are three main reasons for using higher-order spectra: to extract information due to deviations from Gaussianity, to recover correct phase properties, and thus to detect and quantify nonlinearity in time series. Existing techniques of bispectral analysis are particularly useful for studying quadratic interaction, the lowest order of nonlinear interactions. In this paper, it is found that internal waves on the continental shelf could produce a characteristic bispectral signal of sufficient strength to be verified by observations, and that real and imaginary parts of bispectra represent contributions to vertical and horizontal asymmetry of the internal wave elevations from each frequency pair, respectively. [Work supported by NSFC.]

## Session 4pMU

## Musical Acoustics: Musical Instruments and Scales

James P. Cottingham, Chair

Department of Physics, Coe College, Cedar Rapids, Iowa 52402

## Contributed Papers

2:30

**4pMU1. 2DOF model applied to the violin octet A0 and A1 cavity modes.** George Bissinger (Dept. of Phys., East Carolina Univ., Greenville, NC 27858)

An optimized 2DOF model for the A0 and A1 cavity modes [E. A. Shaw, *J. Acoust. Soc. Am.* **87**, 398–410 (1990)] has been used to estimate these frequencies for the Hutchins–Schelleng violin octet. The 2DOF model is notable because it predicts coupling between the A0 and A1 cavity modes [G. Bissinger, *J. Acoust. Soc. Am.* **104**, 3608–3615 (1998)]. This coupling reduces the volume dependence of the A0 mode. 2DOF predictions of A0 frequencies are compared with measurements made on the four largest members of the octet during the course of rib-height modifications [C. M. Hutchins, *J. Acoust. Soc. Am.* **92**, 639–650 (1992)]. Comparisons with new A0 and A1 measurements for the octet will also be made. [Work supported by NSF DMR9802656.]

2:45

**4pMU2. Acoustics of mandolins.** David Cohen (J. Sargeant Reynolds Community College, Richmond, VA 23261) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

We compare the normal modes of vibration and sound radiation from two Gibson F-type mandolins, one with an elliptical sound hole and one with *f*-holes. The instruments were constructed by the first author, using carved plates and other parts supplied by the Gibson company. Normal modes of vibration were determined with the help of electronic TV holography. Air cavity modes were determined by imbedding the instruments in sand to immobilize the plates. Normal modes in a mandolin are somewhat akin to those observed in guitars, although the coupling between plates and the enclosed air is weaker in mandolins.

3:00

**4pMU3. An objective method for determining soundboard material quality.** Mark French (Robert Bosch Corp., Farmington Hills, MI 48335-1672) and John G. Cherng (Univ. of Michigan—Dearborn, 4901 Evergreen Rd., Dearborn, MI 48128)

For a stringed instrument of the guitar or violin family, the dynamic characteristics of the soundboard are extremely important in determining the sound quality of the finished instrument. Since an exceptional instrument can be worth several tens of thousands of dollars, the importance of getting such a basic step correct are obvious. This paper describes our efforts to describe the desirable characteristics of a guitar soundboard blank using what is currently some of the highest quality test equipment available. Our goal is to provide objective standards that can be implemented in an industrial guitar-making process to help ensure the highest quality possible sound quality in the finished instruments. Formal modal tests on a range of soundboard blanks ranging from inexpensive ones intended for use by those just learning their craft to master grade blanks suitable for the highest quality instruments have been performed. The subjective ratings used to grade these blanks with objective test results are correlated.

3:15

**4pMU4. Modal analysis of a new steelpan: The ping.** Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115), Uwe J. Hansen (Indiana State Univ., Terre Haute, IN 47809), Felix Rohner, and Sabina Schärer (Panart, Bern, Switzerland)

The ping, which has the same playing range as a tenor steelpan, is a new instrument in the steelpan family. Its playing surface is sandwich hardened steel having a uniform thickness of 0.9 mm, and its skirt is of stainless steel 1.25-mm thick. Each note area has a small elliptical dome at the center to provide stability and a durable strike point. We compare the modes of vibration of the ping to those of a tenor steelpan. In both pans, the (0,1) mode of each note is tuned an octave above the fundamental, and the (1,0) mode is tuned a twelfth (three times the frequency) above the fundamental in the lower notes. In the ping, it is possible to tune a twelfth throughout more of the range, and the frequency remains unchanged for a longer time during hard playing. Radiation of sound by the skirt is compared.

3:30

**4pMU5. Acoustics of rototoms.** Thomas D. Rossing and Juyong Kwon (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Rototoms are single-membrane drums that can be quickly tuned by rotating the drum. The sound spectra and modal shapes recorded holographically show membrane modes quite similar to those observed in other drums, such as tom-toms, snare drums, and timpani. The sound is dominated by the fundamental radiated bathe (0,1) mode of vibration. The frequency ratio of the various partials changes with tension over the playing range of each drum, resulting in a changing timbre.

3:45

**4pMU6. Lip and throat configuration in flute playing.** Joon-Hee Beth Hwang<sup>a)</sup> (Faculty of Health Sci., The Univ. of Sydney, Australia) and William Thorpe (The Univ. of Sydney, East St. Lidcombe, NSW 2141, Australia)

This study reports on simultaneous measurements of lip aperture and glottal aperture during varied performance tasks by each of a group of six expert flute players. Lip aperture was recorded using a video camera, and glottal aperture with a nasendoscope. Video images were selected from stable portions of the recording and digitized, and measurements of vocal fold and lip opening were made from the display. The results showed that there was a general trend of increasing lip aperture with increase in loudness and decrease in lip aperture with increasing pitch. Similarly, glottal aperture increased with increase in loudness and decreased with increase in pitch. The ratio of glottal to lip aperture area generally decreased with increase in pitch, the glottal aperture always being greater than the lip aperture. Although the results of this study generally agree with those of previous studies [J. W. Coltman, *J. Acoust. Soc. Am.* **40**, 99–107 (1966), **44**, 983–992 (1968), N. H. Fletcher, *ibid.* **57**, 233–237 (1975), S. Mukai, *J. Otolaryngol. Jpn.* **92**, 260–270 (1989)], players showed considerable individual variation. This suggests that various compromises in technique are possible among expert players. <sup>a)</sup>Present address: 7205 Mentor Ave., #C102, Mentor, OH 44060.

4:00

**4pMU7. The acoustics of a symmetric free reed coupled to a pipe resonator.** James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The Asian free-reed mouth organs employ symmetric free reeds mounted in resonating pipes, with the reed vibration strongly coupled to the pipe resonance. The sheng, the sho, and the khaen use a single pipe for each reed, constructed so that the pipe resonance frequency is fairly close to the natural frequency of the reed. The playing frequency is typically slightly above the resonant frequencies of both the reed and the pipe. The free-reed pipe with finger holes (known in China as the bawu) employs a pipe resonator of variable effective length in which both the pipe resonance and sounding frequency are normally well above the natural reed frequency, resulting in a striking change in tone quality. The operation of these instruments has been studied both experimentally and theoretically, with particular attention to the coupling of the reed vibration with the pipe resonator. Experimental measurements include both studies of reed vibration and impedance measurements of the pipes. In general, the experimental results can be shown to agree with the predictions of simple theoretical models.

4:15

**4pMU8. The influence of upstream geometry on the activation pressure of free and restricted reed configurations.** Marius O. Vermeulen, Jabus A. Wessels (Dept. of Medical Physiol., P.O. Box 19063, Tygerberg, 7505, Cape Town, South Africa, jaw@gerga.sun.ac.za), and Theodore W. von Backstrm (Univ. of Stellenbosch, Matieland, South Africa)

This paper describes a computationally efficient algorithm for the automatic calculation of reed activation pressure. The method was evaluated using both free and restricted reed configurations which were not connected to any external air column. It is shown how the reed activation is

influenced by intrinsic factors such as rounded and minute burrs on the edge of the reed as well as extrinsic factors such as upstream geometry of the experimental apparatus. The introduction of constrictions in the flow path upstream from the reed resulted in marked changes in the activation pressure which was highly dependent on the length of the constriction. Expansions of the same magnitude, however, had less influence on the reed activation pressure.

4:30

**4pMU9. A complexity measure for musical scales.** Alpar Sevgen (Dept. of Phys., Bogaziçi Univ., Bebek 80815, Istanbul, Turkey, sevgena@boun.edu.tr)

Equally tempered scales with  $N$  semitones and  $M$  notes and interval structures  $\mathbf{n} = \{n_1, n_2, \dots, n_M\}$ , where  $n_k$  is the number of semitones between the notes  $t_k$  and  $t_{k+1}$ , possess the following properties: Each distinct interval structure  $\mathbf{n}$  corresponds to a multiplet of  $N$  scales. Members of a multiplet can be labeled by a set of integers  $\{c\}$ , modulo  $N$ , called scale labels. Each scale label is the *difference* between the number of sharps and flats occurring in that scale and is unique within the multiplet if  $N$  and  $M$  are relative primes. This labeling does not differentiate between different scale structures. To do this, *complexity* is introduced as the *sum* of the number of sharps and flats occurring in a scale. For  $N = 12$  and  $M = 7$ , out of 462 possible scale structures, the major scale and its cyclical permutations, called modes, have the minimum complexity which allows the practical use of the key signatures in music. Complementary scales where notes and no notes are interchanged have the same complexity. The minimum and maximum complexity scales occupy the opposite ends of the energy spectrum under the force laws  $\pm n^\alpha$  ( $\alpha \neq 0$ ), between the notes of a scale.

FRIDAY AFTERNOON, 2 JUNE 2000

SPANISH ROOM, 1:30 TO 5:15 P.M.

### Session 4pPA

## Physical Acoustics: Outdoor Sound Propagation and Acoustic Seismic Coupling

Gregg D. Larson, Chair

*School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332*

### Contributed Papers

1:30

**4pPA1. Systematic investigation on acoustic-to-seismic responses of landmines buried in soil.** James M. Sabatier and Ning Xiang (Natl. Ctr. for Physical Acoust., Coliseum Dr., Univ. of Mississippi, University, MS 38677, sabatier@olemiss.edu)

Recently, acoustic-to-seismic coupling has been successfully applied to landmine detection [Sabatier and Xiang, *J. Acoust. Soc. Am.* **105**, 1383 (1999); **106**, 2143 (1999)]. When airborne sound penetrates the surface of ground it is refracted towards the normal. If a landmine is buried below the surface of an insonified patch, the transmitted waves will be scattered or reflected, resulting in increased ground surface vibrational amplitudes. These distinct acoustic-to-seismic coupled vibrational changes are sensed using a scanning laser Doppler-vibrometer (LDV) device. To better understand this mine detection phenomenon, the present work is a systematic investigation of the acoustic-to-seismic response to different types of mines in different soil types and at different burial depth has been conducted. [This work is supported by U.S. Army Communications-Electronics Command.]

1:45

**4pPA2. Air acoustic sensing of seismic waves.** Gregg D. Larson, James S. Martin (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Waymond R. Scott, Jr., and Cheng Jia (Georgia Inst. of Technol., Atlanta, GA 30332)

Propagation of elastic waves in damp, compacted sand involves pressure, shear, and Rayleigh waves. The associated seismic surface displacements can be detected by sensing the acoustic pressure immediately above the surface. Propagation speeds are very low in sand. The high wave numbers of seismic displacements are, therefore, evanescent in air. Thus, the acoustic pressure can only be measured well within a seismic wavelength of the surface. Planar near-field acoustic holography techniques can then be used to back-propagate these signals and calculate surface displacements. Measurements have been made using a laboratory experimental model to investigate the potential of using this technique to detect buried land mines. The experimental model utilizes a surface-coupled transducer to generate elastic waves in a sand-filled tank, which simulates the earth. The microphone and a radar system were used to independently

measure the surface displacements. Data taken with both sensors compare well and exhibit the signature of a buried inert antipersonnel mine. For a 100–800-Hz incident pulse, the mine signature can be seen in the raw microphone data when the height of the microphone is less than 3 cm. Holographic signal-processing techniques will be investigated to increase the allowable height for the microphone. [Work supported by ARO.]

2:00

**4pPA3. Electric arc source for high-frequency seismic measurement.**

James S. Martin, Gregg D. Larson, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, james.martin@me.gatech.edu), and Waymond R. Scott, Jr. (Georgia Inst. of Technol., Atlanta, GA 30332-0250)

An electric arc source was designed to study high-frequency seismic surface wave propagation. The noncontact nature of this source made it feasible for use in synthetic aperture transmit arrays. The transmit signal, which was not linearly controllable, was found to be predominantly in the 1- to 4-kHz band with a Gaussian spectrum. This is an octave below the simultaneously generated air acoustic pulse. The source was used to create a synthetic line array in conjunction with a stationary receiver. The experiment was conducted in a sand-filled tank. Surface wave speeds in the range of 80 m/s were measured on the resulting seismograms. Significant dispersion occurred in the propagating waveform at distances both near and far from the source. Surface wave arrivals were discernable over 1 m from the source and compressional head waves could be observed within 30 cm. The data were in good agreement with lower frequency measurements made by other techniques. The surface wave generation was studied and found to be a combination of the surface interaction of the arc and the air acoustic interaction. Strong hysteresis was observed in the first arcing event. Later, the signal was smaller but sufficiently stable for averaging. [Work supported by ARO.]

2:15

**4pPA4. Seismic/electromagnetic system for landmine detection.**

Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250), Gregg D. Larson, James S. Martin, and Peter H. Rogers (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A system has been designed for the detection of buried landmines. The system uses a stationary seismic source in conjunction with a movable displacement sensor that is based on an 8 GHz CW radar. The sensor measures the surface displacement by analog demodulation of the radar signal, which is reflected from the soil surface and modulated by the surface motion. The sensor is not in direct contact with the soil surface and is, therefore, capable of interrogating surface motion immediately above a buried mine. This configuration provides the dual advantage of removing half the seismic propagation path that would be encountered with a classical pulsed echo technique and detecting localized fields that would not propagate to a remote receiver. The system has been used in the laboratory to image inert antipersonnel mines and simulated antitank mines buried in damp compacted sand. Signal processing in the wave number domain provides significant improvement in the contrast between mine-related and background motion. The simplest detection cue for antipersonnel mines was found to be low-frequency resonances of their trigger mechanisms. These responded to seismic excitations with substantial local displacement. The resonances made these mines easily discernable from buried clutter such as rocks and sticks. [Work supported by ARO.]

2:30

**4pPA5. A three-dimensional model for elastic waves in the ground.**

Christoph T. Schroeder, Kangwook Kim, and Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., 777 Atlantic Dr., Atlanta, GA 30332, christoph.schroeder@ee.gatech.edu)

A three-dimensional finite-difference time-domain model for elastic waves in the ground has been developed and implemented on a massively parallel computer. The model is based on the three-dimensional equation

of motion and the stress–strain relation, from which a first-order stress-velocity formulation is obtained. The boundary between the soil and the air is modeled as a free surface. A perfectly matched layer is implemented at the remaining grid edges to absorb the outward traveling waves. The numerical model has been developed as part of a project in which elastic and electromagnetic waves are used synergistically to detect buried landmines. The numerical model is being used to study the interaction of the elastic waves with the buried mines. To verify that the model accurately predicts the mine–wave interaction, the eigenfrequencies of various solid bars and plates are determined numerically and compared to analytical solutions. Currently, the model is being refined to incorporate loss within the bulk medium. Results will be shown for various landmines buried in both loss-free and lossy ground. [Work supported by ARO and ONR.]

2:45

**4pPA6. Detection of land mines in fluid-saturated unconsolidated soil: Numerical modeling.**

Yanqing Zeng and Qinghuo Liu (Dept. of Elec. and Comput. Eng., Duke Univ., Durham, NC 27708)

Because of the strong interactions of waves with the solid grains and the fluid in the pore space, it is more appropriate to model the soil as a fluid-saturated unconsolidated material than a single-phase elastic medium. A three-dimensional finite-difference method for modeling acoustic waves propagating in fluid-saturated unconsolidated soil has been developed. Instead of the conventional elastic wave equations, Biot's equations are used for the poroelastic model. Based on the strain–stress relationship, Biot's equations are reformulated into a first-order hyperbolic system which is equivalent to strain-velocity formulation. A perfectly matched layer (PML) is used to absorb outgoing waves at the truncated boundary of an unbounded medium. The numerical method is validated by comparing numerical results to an analytical solution. Models of a land mine buried in fluid-saturated unconsolidated soil are developed. The numerical method is used to study the interaction of acoustic waves with the buried mines. Comparison of these results is made with those for a buried land mine in a conventional single-phase elastic soil model.

3:00–3:15 Break

3:15

**4pPA7. Acoustically induced slow dynamics in nonlinear mesoscopic elastic materials.**

Alexander M. Sutin (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030), Paul A. Johnson, and James A. TenCate (Los Alamos Natl. Lab., Los Alamos, NM 87545)

We have known about slow dynamics in rock due to continuous wave excitation drive for several years (<http://www.ees4.lanl.gov/nonlinear>). TenCate, Smith, and Guyer (see abstract, this meeting) have recently discovered that both the elastic modulus and the wave dissipation display log time recovery in granular solids, and that it may be thermally or mechanically induced. Much to our surprise, we have discovered that a CW or broad-frequency band acoustic source can also induce slow dynamical response. This response was observed as a variation of the ultrasonic probe wave amplitude, the resonance frequency, and Q factor after the action of a pump wave. The slow time recovery took place in materials such as powdered metals, damaged materials, concrete, and rocks as well. The variations of material properties due to the action of pump waves lead to transient amplification and an obscuration of CW probe waves. The observed behavior may be more universal than was first thought. The results have potential implications to many topics, including laboratory wave studies, earthquake strong ground motion, elastic waves emanating from a point source, damage detection, and manufacturing processes. [Work supported by Stevens and by the Department of Energy: Office of Basic Energy Sciences.]

4p FRI. PM

**4pPA8. Quantifying a new form of acoustic creep in rocks and concrete.** Eric Smith, James A. TenCate (Earth and Environ. Sci., Los Alamos Seismic Res. Ctr., Los Alamos Natl. Lab., Los Alamos, NM 87545, desmith@arlut.utexas.edu), and Robert A. Guyer (Los Alamos Natl. Lab., Los Alamos, NM 87545)

A new form of nonlinear elastic response, which we call “slow dynamics,” has recently been found to scale universally in all sandstones, limestone, and concretes we have tested. Slow dynamics is unique among known creep processes; the elastic modulus and inverse loss tangent drop in response to harmonic acoustic stress, so the response violates the symmetry of the source. We show here that the slow dynamic recovery rate is a function of temperature, consistent with a thermal activation model for creep requiring a spectrum of energy barriers. We also show that short-term modulus drop, with slow dynamic recovery, accompanies temperature change of either sign, again violating the source symmetry. These two modes of activation are compared quantitatively, and suggest a common origin for slow dynamics in disequibrated internal strain fields. [Work supported by the Department of Energy: Office of Basic Energy Sciences and IGPP, the Institute for Geophysics and Planetary Physics at Los Alamos.]

**4pPA9. Examination of surface roughness of porous soil via acoustic backscatter.** Michael L. Oelze and James Sabatier (NCPA, Univ. of Mississippi, University, MS 38677, mloelze@olemiss.edu)

Work has been done to determine acoustically the pore properties of soils, such as porosity, permeability, and tortuosity. Early methods used probe microphones to determine the soil pore characteristics by looking at the attenuation and phase change of sound propagating through the soil pores. The attenuation and phase change of the sound were then related to a surface impedance that incorporated the soil pore properties. Seeking noninvasive means of finding these properties, forward scattering or acoustic level difference experiments were performed. It was seen that sound propagated over porous surfaces is reduced by the surface impedance and by the surface roughness. Theories have been developed to take into account the reduction in acoustic signal due to the finite surface impedance. Difficulties have arisen in separating the signal loss due to impedance and the signal loss from roughness scattering. Acoustic backscatter has been used to study the roughness statistics of soil surfaces in an effort to determine the effects of roughness apart from the surface impedance effects. The roughness statistics obtained via acoustic backscatter are then compared with alternate nonacoustic methods for examining surface roughness.

**4pPA10. High-angle propagation modeling in the atmosphere.** Xiao Di (Appl. Res. Lab., Pennsylvania State Univ., State College, PA 16804) and Kenneth E. Gilbert (Univ. of Mississippi, University, MS 38677)

Many important situations in outdoor sound propagation involve high-angle propagation, for example, elevated sources and ground receivers, or vice versa. For such cases, we have compared two parabolic equation models: the GFPE model of Gilbert and Di and an atmospheric version of the RAM algorithm of Collins. Although the GFPE model, which is based on a split-step Fourier method, gives exact results for a homogeneous atmosphere (constant sound speed and constant density), it was initially thought to be less accurate for an inhomogeneous atmosphere than the RAM algorithm, which is based on high-order Padé approximations. Surprisingly, it was found that for cases with sufficiently weak atmospheric gradients, the GFPE is considerably more accurate at high angles than the atmospheric RAM model, given equal computational effort with both models. With increased computational effort (i.e., higher-order Padé expansions and shorter range steps), the RAM model gives high-angle accuracy comparable to the GFPE. As the atmospheric gradients steepen, the accuracy of the RAM model exceeds that of the GFPE. Nevertheless, for many typical atmospheric conditions, the GFPE offers sufficient accuracy

and considerable computational savings. In this paper, we compare the GFPE and RAM models with benchmark calculations from a fast-field program (FFP) and discuss practical guidelines for using the GFPE and RAM models for high-angle propagation.

**4pPA11. Detection and tracking of broadband airborne sources: An “acoustic fence” approach.** Jay E. Williams, Chad M. Williams (Miltec, Inc., NCPA, Coliseum Dr., University, MS 38677, jwilliams@mil-tec.com), and Kenneth E. Gilbert (Univ. of Mississippi, NCPA, University, MS 38677)

An “acoustic fence” concept is discussed as an advance detection system for approaching subsonic airborne vehicles. Each “post” of the acoustic fence is an array of three closely spaced microphones at right angles to each other on the ground. The posts themselves are spaced approximately 1 kilometer apart. It is shown that the azimuthal and elevation angles for an airborne vehicle can be accurately determined from simple cross correlation of the broadband time series measured at a single array (i.e., at one post). With two arrays, the vehicle track can be determined by finding the stationary points of an iterative nonlinear equation. The equation is of the form  $\tau_p = f(\tau_e)$ , where  $\tau_e$  and  $\tau_p$  are, respectively, the estimated and predicted acoustic travel-time differences for two adjacent arrays. The stationary points, i.e., where  $\tau_p = \tau_e$ , allow the vehicle track to be computed directly from triangulation. Examples are discussed using both synthetic data and actual field measurements.

**4pPA12. Sound propagation through thermal turbulence near a convex surface.** Ph. Blanc-Benon, J. Wasier, and D. Juvé (Ecole Centrale de Lyon, LMFA UMR CNRS 5509, BP 163 69131 Ecully Cedex, France, acous@ec-lyon.fr)

Making use of an acoustic analogy, we study the sound propagation above a convex surface, to simulate the propagation of sound in an upward-refracting medium and to investigate the scattering of sound by turbulence into an acoustic shadow zone. An experimental study was achieved under well-controlled laboratory conditions. A heated grid is placed horizontally in a large anechoic room and the mixing of the free convection plumes above the grid generates a homogeneous isotropic random thermal field. The convex surface is simulated by a vertical cylinder. Experimental data are obtained by varying both the frequency of the acoustic source and the distance of propagation. Measurements will be presented for a rigid convex surface. In this paper, we concentrate on the mean sound-pressure levels and the intensity fluctuations. Experimental values will be compared with results of numerical simulations based on a wide-angle parabolic code.

**4pPA13. A fast method for deducing ground-impedance parameters from measurements of excess attenuation spectra.** Shahram Taherzadeh (Faculty of Technol., The Open Univ., Milton Keynes MK10 9BS, UK, s.taherzadeh@open.ac.uk)

An efficient numerical method for indirect deduction of ground-surface impedance [J. Acoust. Soc. Am. **105**, 2039–2042 (1999)] is used to deduce surface impedance of absorbing grounds and to obtain estimates for the effective flow resistivity of the ground. One-parameter and two-parameter impedance models are used for the parameter-fitting purposes. It is shown that the two-parameter model can be modified to give a better agreement with the measured data at grazing angles.

5:00

**4pPA14. The estimation of excess attenuation of acoustic beam and its application.** Naixian Pan (Dept. of Geophys., Peking Univ., Beijing 100871, PROC)

A theory of the excess attenuation for a conical beam is presented. The equation of attenuation coefficient shows the attenuation contributed by the item of Cv square is larger than that of Ct square about five times at least in the atmosphere. The theory shows that a significant difference of excess attenuation between emitting wave and back scattering on the same path. Excess attenuation and absorption both are taken into account in

sodar data processing and the Ct square measured by sodar has a reasonable agreement in statistics with that measured by a filament thermometer. Based on the quantitative measurement of Ct square, sodar was used in an experiment whose aim was estimating the seeing quality of a candidate observatory in China in 1997. The seeing measured by sodar has a better correlation with the seeing measured by a differential image motion monitor in our experiment than that in other experiment in Russia (1988), in which the attenuation of sound had not been considered. The value of the excess attenuation is estimated to be the same order as the absorption in a typical clear day [E. H. Brown and S. F. Clifford, J. Acoust. Soc. Am. **60**, 788–794 (1976)].

FRIDAY AFTERNOON, 2 JUNE 2000

STATE ROOM, 2:30 TO 4:30 P.M.

### Session 4pPP

## Psychological and Physiological Acoustics: Auditory Modeling and Physiology

Neal F. Viemeister, Chair

*Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455*

### Contributed Papers

2:30

**4pPP1. Directional microphone theory.** Patrick C. Fournier (Prescription Design Hearing Aid Ctr., 2311 Tenth Ave. S., Great Falls, MT 59405)

An alternate to the cochlear amplifier and resonant tectorial membrane models, this theory models the cochlea as a type of directional microphone. The main source of cochlear nonlinearity is a mechanical canceling of more intense signals from the interaction of the cochlea's various diaphragms. These interactions allow both the inner and outer hair cells to focus a wide range of environmental amplitude into the relatively invariant encoded signal that is sent to the auditory brain. This cancellation of intense signals allows concurrent weak signals in the environment to be perceptively enhanced as compared to the more intense concurrent signals. This model explains Yates dual-wave mode in the cochlea, how and why the outer hair cells are stimulated at low intensities, and how the cochlea performs a fast Fourier transform. Loudness growth charts recalculated in pascals clearly show the enhanced perception of soft signals in comparison to intense ones. Detection of pure-tone signals delivered in a sound field of narrow-band white noise showed equivalent intensity differences for both impaired and nonimpaired ears for perceived loudness levels when calculated in pascals. Expected focus or enhancement of normal soft to intense sound perception was recreated artificially for the impaired ears.

hand, various experiments demonstrated that modulations applied to spectrally separated carriers can interfere, indicating that (prior to the decision stage) amplitude modulation is not processed independently across carrier frequency in these conditions. In the present study, modulation detection and modulation masking conditions were examined where the signal and the masker modulation were applied either to the same carrier or to spectrally separated carriers. Simulations are presented based on an envelope power spectrum model [Ewert and Dau, J. Acoust. Soc. Am. (submitted)], which calculates the envelope power in the transfer range of a modulation filter tuned to the signal-modulation frequency. Within the model, different strategies of combining across-channel envelope information, as well as the influence of peripheral filtering and compression, are considered. [Work supported by the Deutsche Forschungsgemeinschaft.]

3:00

**4pPP3. Modeling the temporal response of auditory nerve fibers with a leaky integrator.** Kyle T. Nakamoto, Rita L. Smith, and Bruce G. Berg (Dept. of Cognit. Sci., Univ. of California, Irvine, CA 92612)

A leaky integrator model (e.g., bandpass filter, followed by half-wave rectification and a low-pass filter) describes the temporal response of auditory nerve fibers presented with multicomponent, periodic sounds. It is shown that the spectra of the leaky integrator output closely resemble the spectra of period histograms obtained from published studies. Five parameters define the model: a time constant defining the cutoff frequency of the low-pass, one-pole filter, a parameter representing internal noise, and three parameters describing the initial asymmetric roex(p) filter (center frequency, low- and high-frequency skirts). For high intensities, an algorithm was introduced to represent saturation of the fiber. In most cases, the model accounted for more than 90% of the variance. Parameter estimates are consistent with known characteristics of auditory nerve fibers: (1) the low skirt of the initial filter was shallower than the high-frequency skirt, (2) increases in intensity were accompanied by shifts in the center frequency of the initial filter, and (3) bandwidth of the initial filter increased with intensity. With respect to the nonlinearity, square wave and full wave rectification were found to be implausible alternatives.

2:45

**4pPP2. Modeling across-channel processing of amplitude modulation.** Stephan Ewert, Jesko Verhey, Birger Kollmeier (Carl von Ossietzky Universitaet Oldenburg, AG Medizinische Physik, D-26111 Oldenburg, Germany, se@medi.physik.uni-oldenburg.de), and Torsten Dau (Boston Univ., Boston, MA 02215)

In typical modulation detection experiments, a sinusoidal signal modulation is applied to broadband noise as the carrier. In such conditions, the carrier envelope at the output of different peripheral auditory channels might be considered as nearly uncorrelated while the signal modulation is coherent across channels. Models of auditory processing often combine signal information optimally in terms of signal detection theory, assuming independent "observations in different auditory channels." On the other

**4pPP4. The phases of auditory-nerve fiber responses to tones: Dependence on stimulus frequency and intensity.** Andrei N. Temchin and Mario A. Ruggero (Hugh Knowles Ctr. and Inst. for Neurosci., Northwestern Univ., 2299 North Campus Dr., Evanston, IL 60208-3550)

As stimulus intensity increases, the phases of responses to tones of squirrel monkey auditory-nerve fibers [ANFs; Anderson *et al.*, *J. Acoust. Soc. Am.* **49**, 1131–1139 (1971)] and guinea pig inner hair cells [IHCs; Dallos, *Hearing Res.* **22**, 185–198 (1986)] undergo lags and leads, respectively, at frequencies lower and higher than the characteristic frequency (CF) and remain relatively constant at frequencies near CF. A similar pattern holds for basal sites of the basilar membrane (BM) in several species. In contrast, tectorial-membrane (TM) vibrations at apical sites of the chinchilla cochlea exhibit phase shifts of opposite polarity [Rhode and Cooper, *Aud. Neurosci.* **3**, 101–121 (1996)]. To explore this inconsistency between neural and mechanical behavior at apical cochlear regions, we investigated the phases of responses to tones of ANFs recorded in deeply anesthetized chinchillas. For stimulus intensities lower than 70 dB SPL, responses of most ANFs (regardless of CF) exhibited phase shifts similar to those observed in squirrel monkey ANFs, guinea pig IHCs, and basal BM sites in several species. Thus, the inconsistency between neural and TM mechanical data cannot be ascribed to species differences or a dependence of phase behavior on CF. [Work supported by NIH Grant DC-00419.]

## 3:30

**4pPP5. Insight into the diagnosis of conductive hearing loss using wideband acoustic reflectance measurements of surgically modified human temporal bones.** Antonio Miller (Mimosa Acoust., Inc. and the Penn State Grad. Prog. in Acoust., 382 Forest Hill Way, Mountianside, NJ 07092, ajm249@psu.edu)

There is mounting evidence to suggest that wideband acoustic impedance/reflectance measurements will aid in the clinical diagnosis of various middle ear pathologies. In preparation for clinical trials, wideband pressure reflectance measurements (100 Hz to 8 kHz) were made looking into the ear canals of three human temporal bones which were surgically modified to simulate five possible pathological conditions [unpublished data of J. Allen, G. Ball, S. Nishihara, and R. Goode]. The experimental procedure, data, and analysis are presented for the different modifications. The unmodified, air-filled tympanic cavity data characterize the baseline losses for the entire system. A fluid-filled tympanic cavity characterizes the effect of fluid loading on the tympanic membrane, simulating the condition of acute and chronic otitis media (middle ear infection). The fixed stapes condition simulates the effect of otosclerosis (calcification of the stapes footplate). Disarticulated stapes and the severed stapedius tendon conditions are both symptoms associated with chronic otitis media or cholesteatoma (middle ear cysts). Finally, the fixed malleus condition can be a warning sign of otitis media, but can easily be misdiagnosed as otosclerosis. It is hoped that this work will provide a physical basis for the interpretation of wideband pressure reflectance data collected in future clinical trials.

## 3:45

**4pPP6. Is hearing high cochlear (revisited)? Frequency tuning and intensity thresholds in acoustically stimulated, myogenic vestibular-evoked potentials (MVEP).** Neil P. McAngus Todd (Dept. of Psych., Univ. of Manchester, Manchester M13 9PL, UK, todd@fs4.psy.man.ac.uk), Fred W. Cody, and Jon Banks (Univ. of Manchester, Manchester M13 9PL, UK)

There has amassed considerable evidence that the sacculus, an organ of hearing in lower vertebrates, has retained some acoustic sensitivity throughout phylogeny. In humans, myogenic vestibular-evoked potentials

(MVEPs) may be obtained from motoneurons innervated by the vestibulo-spinal tract, particularly from the cervical region of the spinal cord. The MVEP has been studied principally as a noninvasive clinical tool for evaluation of normal otolith vestibular function, since traditional nystagmographic methods only really assess canal function. However, we have been interested in what “auditory” function acoustic sensitivity sacculus may have in the perception of sound. The results of a series of experiments are reported which may be summarized as follows. (1) MVEP has a frequency tuning [Todd, Cody, and Banks, *Hear. Res.* (in press)], with a maximum response frequency between 300 and 350 Hz and a band-width of about 3 octaves. (2) MVEP can be obtained to natural acoustic stimuli [Todd and Cody, *J. Acoust. Soc. Am.* **107**, 496–500 (2000)], such as dance music, above about 90 dB SPL. (3) MVEP can be obtained to continuous sounds. Given the above then, there are a number of possibilities where acoustic sensitivity of the sacculus may play a role in human perception.

## 4:00

**4pPP7. Acute effects of nicotine on physiological responses of the auditory systems of nonsmokers.** Ashley W. Harkrider (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 S. Stadium Hall, Knoxville, TN 37996, aharkrider@yahoo.com) and Craig A. Champlin (Univ. of Texas, Austin, TX 78712)

The role of cholinergic mechanisms in the auditory system was investigated by assessing the acute effects of nicotine, a cholinergic drug, on aggregate responses within the auditory pathway. In a single-blind procedure, evoked potentials (auditory brainstem, middle-latency, long-latency, steady-state responses) and electroencephalograms were measured from 20 normal-hearing, nonsmokers (10 male) under two conditions (nicotine, placebo). After the drug session, plasma tests indicated a subject’s nicotine concentration. Results indicated that transdermal administration of nicotine to nonsmokers does appear to acutely affect the transmission of acoustic information. The measures most affected by nicotine administration likely originate from the midbrain and cortex and are believed to involve processing of auditory information related to vigilance and the screening of sensory input. The effects of nicotine on evoked potentials or the electroencephalogram do not appear to be dependent on the sex of the subject or the hemisphere from which the responses are measured. [Work supported by AAA and NIDCD.]

## 4:15

**4pPP8. Standing waves in the brain and bone conduction hearing.** Martin Lenhardt, Alan Madsen, James Genova, and Howard Neal (Prog. in Biomed. Eng., Virginia Commonwealth Univ., Richmond, VA 23298-0168, lenhardt@hsc.vcu.edu)

Bone conduction hearing studies have a long history; however, in the last 25 years the topic of bone conduction hearing has been virtually abandoned. Delivering sound, including speech, to the head has certain advantages in high-noise communication. Difficulty arises in localization since bone conduction is a binaural perception, at least for tones under 10 kHz. At higher frequencies a robust lateralization of tones was observed for stimuli of 20 digital tones in the range of 4–18 kHz and 20–34 kHz. Inverting the listener’s body altered perception and transmission by bone conduction. Some tones in areas of resonance were altered in amplitude when the body was inverted, presumably due to increased cerebral spinal fluid pressure. Modeling of the skull as a sphere with rigid boundary conditions generally predicted direct observations. These findings also account for the observation of perception jumping from ear to ear with changes in frequency during ultrasonic bone conduction listening.

**Session 4pSA****Structural Acoustics and Vibration: Tire Vibration**

J. Stuart Bolton, Chair

*School of Mechanical Engineering, Purdue University, 1077 Ray W. Herrick Laboratories,  
West Lafayette, Indiana 47907-1077***Invited Papers****2:00****4pSA1. Wave-number domain representation of tire vibration.** J. Stuart Bolton and Yong-joe Kim (1077 Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077)

In the work to be described here, wave-number decomposition techniques have been used to study tire vibration. In the experimental component of this work, a tire was driven radially at a point on its treadband. Measurements of the resulting radial treadband vibration were made at approximately 200 points around the treadband circumference by using a laser Doppler velocimeter. From an inspection of the resulting space-frequency data, it was possible to identify the frequency ranges in which the tire responded either modally or nonmodally. Further, by performing a circumferential wave-number decomposition of the space-frequency data, the propagation characteristics of the wave types that contributed to the response of a tire could be determined. It was observed that a small number of circumferentially propagating waves combine to control the response of a tire in both low- and high-frequency ranges. The cut-on and propagation characteristics of these waves are closely related to propagating waveguide modes of flexural wave-bearing systems. Together, these results are consistent with a curved waveguide model of a tire. It is therefore suggested that a membrane-like waveguide model combined with a propagating wave representation of the tire vibration may offer an efficient way of representing tire vibration.

**2:30****4pSA2. Two-dimensional experimental analysis of tire vibration.** Jacob Klos, Robert J. Bernhard, and Paul Cvar (1077 Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077)

This paper presents an experimental analysis of the velocity response of a treadless tire in the two-dimensional wave-number domain. The out-of-plane velocity response was measured on a stationary tire using a Doppler laser vibrometer for a two-dimensional 2-cm by 2-cm grid of points over the frequency band from 70 to 1500 Hz. From the data, three distinct regions of behavior can be identified: simple standing wave patterns, waveguide behavior, and nearly cylindrical wave propagation. Using 2-D wave-number reduction techniques, the measured spatial velocity distribution at each frequency was transformed into 2-D wave-number space. Different wave types in the tire carcass can be identified including the dominant flexural and in-plane waves (due to Poisson's strain). At frequencies below the ring frequency of the tire cross section, a qualitative comparison is made to published results for 2-D dispersion curves of cylinders. At higher frequencies, a comparison is made in the wave-number domain, to a coupled plate model of the tire carcass and models of wave behavior in inhomogeneous plates. A qualitative explanation of the behavior of the tire is identified from the comparisons.

**3:00****4pSA3. A one-dimensional wave model for a pneumatic tire.** R. J. Pinnington (Inst. of Sound and Vib. Res., Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK, rjp@isvr.soton.ac.uk)

A wave equation for one-dimensional waves out of plane waves of an infinite tire belt are derived. The dynamics of the belt are regarded as those of a Timoshenko beam under tension, that is, uniformly supported by the sidewall stiffness. The sidewall is modeled here as a static stiffness controlled by the internal air pressure and the geometry of the tire cross section. The fourth-order differential wave equation is solved yielding two wave numbers as a function of frequency. These describe all the wave types of the belt. Below about 100 Hz the sidewall stiffness controls the deformation and there is only motion local to the excitation. At higher frequencies

waves propagate around the belt with four different controlling mechanisms, tension, bending, shear, and longitudinal stiffness. The tire is made circular by the application of the boundary conditions, which is mainly a uniform force applied on a line across the width. The input and transfer mobilities of the belt are output, showing standing wave behavior below about 500 Hz, above which there are only propagating waves and near-field waves as on an infinite belt.

### 3:30–3:45 Break

### 3:45

**4pSA4. Modeling of tire/road interaction.** Wolfgang Kropp (Dept. of Appl. Acoust., Chalmers Univ. of Technology, S-41296 Gothenburg, Sweden)

In order to reduce tire/road noise a deep understanding of the noise generation mechanisms involved is required. The key to reach this understanding is a correct model of the interaction between a tire and a rough road. This includes the modeling of normal and tangential contact between tire and road and the modeling of the wave propagation in the tire. A tire model is presented, which includes the radial, lateral, and tangential motion of the tire. The tire is modeled as a double-layered plate where the two layers are formulated by the general field equations including external tension. The tire model also takes into account the local deformation of the rubber surface due to external forces. This allows for solving directly the contact problem applying Green's functions (i.e., impulse response functions) of the vibrating body. The model also takes into account the coupling between adjacent points at the surface of the tire. Numerical results will be presented for both the tire model and contact model.

## Contributed Papers

### 4:15

**4pSA5. Tire tread block noise experiments including effects of pavement porosity.** Sarah Jones, Dean Karnopp, and Nesrin Sarigul-Klijn (Transportation Noise Control Ctr. (TNCC), Dept. of Mech. and Aero. Eng., Univ. of California at Davis, Davis, CA 95616-5294)

A tread block impact device has been developed to record the pressure pulse signature generated by a single tread block impacting a surface. High-speed video images are used to relate the impact progression with the instantaneous acoustic signal. The acoustic pressure pulses generated by this device can be associated with the changing volume velocity of the air displaced as the tread block approaches the surface. A simple monopole source can be used to predict the acoustical pulse that is radiated from impact. Experimentation and theoretical calculations will be correlated, as well as discussion of previous research. Noise due to a sequence of blocks on a tire can be inferred from single pulse results and the sound power level can be examined as a function of the speed of the tire. By using impact surfaces with known area removal and, hence porosity, experimental relationships between the pressure pulse signature and impact surface characteristics can be determined. By evaluating the frequencies generated by tread block impacts and the effect of porosity on noise reduction, tuned pavements which could attenuate the most irritating frequencies could be developed. The theoretical and experimental models may prove useful in suggesting means of reducing this component of traffic noise.

### 4:30

**4pSA6. Predicted ground effects for the sound fields near a reflecting wall.** Raymond S. H. Tang and Kai Ming Li (The Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Kowloon, Hong Kong, mmkml@inet.polyu.hk)

In many environmental noise prediction schemes, an empirical correction factor of 2.5 or 3.0 dB(A) is to be added for the calculation of noise 1 m in front of a reflecting facade. In this paper, theoretical and experi-

mental studies have been conducted to examine the validity and accuracy of such approximations. The theoretical analysis is based on the representation of the sound waves in their integral forms. By imposing the impedance boundary conditions on the absorbing ground and the vertical wall, it is possible to extend the classical Weyl–Van der Pol formula for inclusion of the effect of a reflecting facade. As expected, the analytical solution is composed of four terms (a source and three image sources) as a result of the absorbing ground and the reflecting facade. Experimental measurements have been conducted indoors to validate the theoretical predictions. It has been demonstrated that the simple empirical correction factor is not adequate in most cases for the assessment of the facade noise levels above an absorbing ground. [Work supported by the Hong Kong Polytechnic University.]

### 4:45

**4pSA7. Efficiency of barriers for transport noise insulation in the near-field region.** S. A. Rybak (N. N. Andreyev Acoust. Inst., Moscow 117036, Russia, rybak@tarsks.msk.su) and S. A. Kostarev (Lab. of Acoust. and Vib. Tunnel Assoc., Moscow 107217, Russia)

The efficiency of noise-insulating barriers needs to be carefully calculated in the near-field region where the Fresnel's parameter is not valid. The correct calculative method for the diffraction across the barrier is developed, taking into account both the resonance absorption at the boundary of the barrier and its thickness for the low-frequency range. The comparison of the calculus and experiments for the barriers arranged along the transport lines was made and the agreement is sufficiently good.

## Session 4pSC

## Speech Communication: Modeling Speech Processing and Production (Poster Session)

Elliot Saltzman, Chair

*Department of Physical Therapy, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215*

## Contributed Papers

All posters will be on display from 2:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 2:00 p.m. to 3:30 p.m. and contributors of even-numbered papers will be at their posters from 3:30 p.m. to 5:00 p.m.

**4pSC1. Waveguide modeling as a tool for fitting a hearing aid.** Grzegorz Szwoch, Andrzej Czyzewski, and Bozena Kostek (Sound Eng. Dept., Tech. Univ. of Gdansk, Poland, kid@sound.eti.pg.gda.pl)

A physical modeling method is used to develop a model of the acoustical elements of a hearing aid and to examine its properties. This model enables one to design an acoustical element having desired frequency characteristics, thus providing the acoustical system well fitted to the individual patient's needs. The acoustical system of the hearing aid is treated as a waveguide—a long and narrow tube. Modeling methods similar to the ones applied to human vocal tract synthesis are used. The system is divided into several cylindrical sections. Physical phenomena present in the acoustical system (wave reflections, frequency-dependent damping, etc.) can be implemented directly in the model. By adjusting the parameters of the model, one is able to examine acoustical properties of various systems, having different shapes and dimensions. This method may be helpful for the audiologists during the process of hearing aid fitting.

**4pSC2. Using Markov models to assess articulation errors in young children.** H. Timothy Bunnell, Debra M. Yarrington, and James B. Polikoff (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19803)

Digital recordings of children producing the names "Rhonda" and "Wanda," and/or "Toto" and "Coco" were made using the microphone input to a Toshiba laptop computer (16-bit samples, 22 050-kHz sampling rate) with an AKG C410/B head-mounted condenser microphone. These names were associated with animated characters in a mock video game running on the laptop under the control of a Speech Language Pathologist. The children, ranging in age from four to six years, were undergoing speech therapy at the Alfred I. duPont Hospital for Children for one or both of two common articulation errors: /w/ substituted for /r/; and/or /t/ substituted for /k/. The initial segment in each recorded utterance was classified by laboratory staff as either r/w or t/k, and assigned a goodness rating. Discrete Hidden Markov phoneme Models (DHMMs) trained using data recorded from normally articulating children were then used to classify the same utterances and results of the automatic classification were compared to the human classification. Results indicate that appropriately trained DHMMs can provide accurate classification of utterances from children in speech therapy. This technology could support articulation drill on home computer systems as an adjunct to speech therapy. [Work supported by Nemours Research Programs.]

**4pSC3. Speech intelligibility and acoustic characteristics of Mandarin speakers with cerebral palsy.** Jing-Yi Jeng, Gary Weismer, and Ray Kent (Dept. of Communicative Disord. and Waisman Ctr., 1975 Willow Dr., Univ. of Wisconsin-Madison, Madison, WI 53706)

Reduced speech intelligibility is often associated with individuals having cerebral palsy. The purpose of the present study was to investigate the speech intelligibility and acoustic characteristics of three subtypes of cerebral palsy (athetoid, spastic, mixed) among adult speakers of Mandarin. A Mandarin word intelligibility test was constructed using phonetic contrasts known from previous studies to be especially sensitive to the dysarthria associated with cerebral palsy. These included tonal contrasts, whose relationship to speech intelligibility deficits in persons with cerebral palsy is not well understood. Acoustic measures consisted of formant frequencies, segment durations, and measures of tone. Preliminary results show that speech intelligibility of athetoid speakers is often worse than that of spastic speakers, and that there are also reliable acoustic differences between the groups. A multidimensional scaling analysis of the perceptual results and regression models of the relationship of the acoustic to the perceptual variables will also be reported. [Work supported by NIH DC00319.]

**4pSC4. Compensatory articulatory behavior to immediate structural perturbation of the oral cavity.** Masaaki Honda, Tokihiko Kaburagi, and Akinori Fujino (NTT Commun. Sci. Labs./CREST JST 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa 2430198, Japan, hon@idea.brll.ntt.co.jp)

Speech compensation to the structural perturbation of the oral cavity was investigated by using an artificial palate whose thickness in the alveolar region is dynamically changeable. The artificial palate was constructed with a base and thin rubber attached to the base, and the rubber balloon was dynamically controlled by means of an external air flow. The articulatory behavior with the structural perturbation was recorded using an electromagnetic articulographic system. A set of VCV syllables uttered in the carrier sentence, containing the consonants /s, sh, t, n/ and the vowels /a, i, u/, was used for the test. Unexpected perturbation of the palate shape, a change in thickness of 1 to 4 mm, caused significant speech articulation errors for the fricative consonants, fricative consonants /s, sh/ changed to the stop consonants /t, ch/. On the contrary, the stop consonant /t/ was very robust to unexpected perturbation. After five to seven speaking trials with the modified palate shape with or without audio feedback, these articulation errors disappeared and the subject could produce correctly perceived

consonants. Then, the articulatory compensation in the alveolar region significantly appeared in the tongue tip position but was small in the jaw and the lip positions.

**4pSC5. On the correlation between orofacial movements, tongue movements, and speech acoustics.** Jintao Jiang, Abeer Alwan, Patricia Keating (Depts. of Elec. Eng. and Linguist., UCLA, Los Angeles, CA 90095), and Lynne Bernstein (House Ear Inst., Los Angeles, CA 90057)

This study is a first step towards quantifying the relationship between external orofacial movements, tongue movements, and the acoustics (AC) of speech sounds. The database analyzed consisted of six sentences and 23 CV syllables spoken by a male and a female; each utterance was repeated eight times. A Qualysis (optical motion capture system) and an EMA (electromagnetic articulography) system were used to characterize orofacial and tongue movements, respectively; principal component analysis was applied to both types of data to reduce the dimensionality. Acoustic features were represented by linear spectral pairs. The datastreams were aligned using a synchronization technique, and a multilinear regression technique was applied to quantify the correlation between them. The results are in general agreement with the study of Yehia *et al.* [AVSP-1997, pp. 41–44], which examined two sentences by a male talker, but also shows that correlations are highly context and speaker dependent. On average, the highest correlations were found between the optical (orofacial) and EMA (tongue) features. The results and their implications will be discussed. [Work supported in part by NSF KDI award 9996088.]

**4pSC6. Planning and forming articulatory movements from phoneme-specific representatives.** Takesi Okadome and Masaaki Honda (NTT CS Labs./CREST, JST, Morinosato-wakamiya 3-1, Atsugi-shi, Kanagawa pref 243-0198, Japan, houmi@idea.br.ntt.co.jp)

A computational model of sensorimotor planning of speech production is described. The model uses phoneme-specific representatives to predict the trajectories of articulatory movements for continuous speech. As phoneme-specific representatives, the model uses acoustics for the vowels and articulatory features for the consonants. The acoustics are represented by LPC cepstral segments. The articulatory features of the consonants are those of the kinematics of the articulators; that is, the position and velocity of the articulator. The model first recovers the kinematic features from the acoustic segment for the vowels using the acoustic-to-articulatory inverse mapping method [S. Suzuki *et al.*, Proc. ICSLP 6, 2251–2254 (1998)]. Using the kinematic features of the vowels together with those for the consonants, the model determines the minimum-acceleration trajectory that coincides with the extremum of the time integral of the square of the magnitude of acceleration of the articulator [T. Okadome *et al.*, Proc. IEEE SMC 2, 469–474 (1999)]. The model predicts both the qualitative features and quantitative details experimentally observed.

**4pSC7. Demonstrating effects of parameter dynamics on gestural timing.** Elliot Saltzman (Haskins Labs., New Haven, CT and Dept. of Physical Therapy, Boston Univ., 635 Commonwealth Ave., Boston, MA 02215, esaltz@bu.edu) and Dani Byrd (Haskins Labs., New Haven, CT and USC, Los Angeles, CA 90089-1693)

Speech movements have been modeled as goal-directed gestures whose activation-gated dynamics are those of critically damped point attractors. Recent work by ourselves and others has suggested that gestural parameter values are not fixed but rather vary as a function of a gesture's

position in the word and phrase. For example, lowering of gestural stiffness values has been hypothesized to underlie gestural lengthening adjacent to phrasal boundaries. Additionally, evidence exists suggesting that prosodic structure affects the relative timing between gestures. In earlier work, we have hypothesized that slowing at phrase boundaries can be implemented using nonconstriction gestures called pi-gestures (prosodic gestures) that act on the parameter values of all concurrently active constriction gestures in proportion to the activation level of the pi-gesture. In order to understand the potentially complex consequences of these parameter dynamics, we present a series of simulations involving different temporal arrangements of two constriction gestures coordinated with one another and with a prosodic gesture. The within- and between-gesture temporal patterning of the constriction gestures will be evaluated under the influence of pi-gestures of greater and lesser strength (activation) and duration. Simulation results will inform interpretation of natural kinematic articulatory data collected in movement tracking experiments. [Work supported by NIH.]

**4pSC8. Kinematic factors underlying timing in connected speech.** Stephen M. Tasko (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., Washington, DC 20307-5001) and John R. Westbury (Univ. of Wisconsin, Madison, WI)

Current understanding of the kinematic factors underlying speech timing has been largely derived from short, simple speech tasks such as citation words. However, timing analyses based upon acoustic segment durations have shown that timing rules differ for citation words and connected speech. Thus it is not clear how well temporal correlates of citation-style speech movement might generalize to connected speech. A better understanding of the articulatory bases of timing patterns associated with connected speech may aid in the refinement of natural-sounding articulatory-based speech synthesis. This study sought to determine temporal correlates of articulator fleshpoint movements in orally read speech recorded from 20 normal adult speakers of American English. Position histories associated with the tongue blade, tongue dorsum, jaw, and lower lip were segmented into series of movement strokes, and measures of distance, maximum speed, and duration were determined for each stroke. Canonical correlation analysis showed that for all articulators, these three measures were highly interrelated, and that 80% of the variance in movement duration could be accounted for by movement distance and maximum movement speed. These results suggest the possibility of predicting articulator timing based on a small number of kinematic measures. [Work supported by NIH Grant Nos. DC00820, DC03723, and DC03659.]

**4pSC9. Articulatory kinematic and formant frequency changes across vocalic nuclei: Some preliminary analyses.** Gary Weismer and Jing-Yi Jeng (Dept. of Communicative Disord. and Waisman Ctr., 1975 Willow Dr., Univ. of Wisconsin-Madison, Madison, WI 53706)

There are few data relating changes in articulatory position to changes in vocalic formant frequencies. This kind of information is useful for speech-production theory in general and specifically in understanding the relationship in neurogenic speech disorders between movement effects and resulting speech-intelligibility deficits. The purpose of this study was therefore to assemble a fairly large kinematic/acoustic database on normal speakers, for a few words that have proven useful in our studies of the intelligibility deficit in dysarthria. Data were part of the x-ray microbeam database, and for the present study included the words “hail,” “shoot,” and “coat.” The position histories of four tongue pellets and two lip pellets as well as the time histories of the first and second formant frequencies were studied for 45 young adult speakers. The F2 transition extent (change in F2 across a defined time interval) was significantly correlated with some combination of the pellet time histories for each word,

and the magnitude of the correlation showed clear gender effects. Results will be discussed in terms of the linkage between planar kinematic data and acoustic data in predicting speech-intelligibility deficits. [Work supported by NIH DC03723 and DC000820.]

**4pSC10. Effects of central dopaminergic stimulation by apomorphine on producing sentence focus in speech production in Parkinson's disease—a preliminary report.** Q. Emily Wang (Dept. of Commun. Disord. and Sci., Rush Univ., Rush-Presbyterian-St. Luke's Medical Ctr., Chicago, IL 60612), Yi Xu (Northwestern Univ., Evanston, IL 60208), Katie Kompolti, and Christopher G. Goetz (Rush Univ., Rush-Presbyterian-St. Luke's Medical Ctr., Chicago, IL 60612)

It is well documented that when a sentence focus is produced, the  $f_0$ -range of the word under focus is expanded while the  $f_0$ -range of the post-focus words is suppressed. It is also a well-known fact that individuals with Parkinson's disease (PD) exhibit reduced ability in pitch modulation, resulting in a perceptual quality of "monopitch" in connected speech. Although the use of dopaminergic stimulation improves other physical symptoms such as balance and walking in PD, it is unclear whether a similar effect will be found on vocal folds, producing an improvement in patients' ability in pitch modulation. In this study, five nondemented English-speaking individuals with a clinical diagnosis of idiopathic PD (Hoehn and Yahr stage 3 and 4, off) were examined for their ability in pitch modulation in association with Apomorphine, a nonergot dopamine agonist. Two six-word sentences with four sentence focus variations each were used. Subjects randomly produced three repetitions of each focused version at baseline and two clinical conditions: injected with placebo or apomorphine. Perception tests and acoustic analyses were conducted to determine changes in these patients' ability to produce sentence focus. Implications of the findings are discussed in terms of speech motor control.

**4pSC11. A quasi-one-dimensional model for speech production.** Wei Zhao, Zhaoyan Zhang, Steven H. Frankel, and Luc Mongeau (School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

A quasi-one-dimensional model for speech production was developed. The unsteady quasi-one-dimensional Euler equations with temporal and spatial area variation were solved using a fourth-order accurate Runge-Kutta scheme for time integration and a sixth-order compact finite-difference scheme for spatial discretization. A model for wall friction was used, and a model for the pressure loss associated with flow separation at the glottis was derived. The code was validated by calculations of steady isentropic flow through a nozzle and an acoustic standing wave in a tube with uniform cross-sectional area. A simple geometry was then used to simulate the acoustic wave generated in the vocal tract. The results are in qualitative agreement with data from experiments obtained using a dynamic model of the larynx. [Work supported by NIH DCO 3577-02, ROI grant from NICDC.]

**4pSC12. Articulatory speech synthesis and the analysis of boundary conditions.** Jun Huang, Stephen Levinson (Univ. of Illinois, 405 N. Mathews Ave., Urbana, IL 61801, jhuang@ifp.uiuc.edu), Scott Slimon, and Don Davis (Electric Boat Corp., Groton, CT 06340)

In this work, a computational fluid dynamics (CFD) approach was used to model the unsteady fluid flow in idealized human vocal tracts. The speech waveform synthesis was based on the slightly compressible Reynolds-averaged Navier-Stokes (RANS) equations. A K-epsilon turbulence model has been used to represent the effects of turbulence. The

vocal tract geometry was determined from the mid-sagittal cut through the three-dimensional vocal tract shape. The moving vocal tract shapes during the speech articulation were estimated using interpolation technique. The excitation signal for the simulation was accomplished by specifying a time-varying area at the inlet to the bottom of the pharyngeal cavity and a sawtooth waveform was used to represent the area variation in time. Different boundary conditions were applied at the inlet of the vocal tract. One is based on specifying the volume velocity and the other is based upon specifying the particle velocity, respectively. We will further analyze the effects of different boundary conditions using an analogy of transmission line T-network. Finally, we will present the waveform and spectrum of some synthesized voiced sounds based on the above CFD approach and boundary conditions. [This work was supported by Motorola and NSF.]

**4pSC13. Glottal pressure profiles for a diameter of 0.04 cm.** Ronald C. Scherer and Daoud Shinwari (Dept. of Commun. Disord., Bowling Green State Univ., Bowling Green, OH 43403, ronalds@bgnnet.bgsu.edu)

Computer models of phonation often rely on aerodynamic equations for flow through the glottis. Usually the aerodynamic equations have come from empirical work with steady-flow models made from hard material. The equations simplify the pressure-flow-geometry relations through the larynx. The model used here (model M5) has 14 pressure taps on the vocal folds to give rather complete pressure profiles. Pressure profiles will be reported for symmetric glottal shapes, a minimal diameter of 0.04 cm, and nine glottal angles (uniform; convergent, and divergent 5, 10, 20, and 40 deg) for transglottal pressures of 3, 5, 10, and 15 cm H<sub>2</sub>O. The glottis is rectangular in the anterior-posterior direction, with a glottal length of 1.2 cm. Results indicated that the pressure coefficient for the glottal entrance ranged from 1.09 to 1.60 with an average of 1.33 (compared to van den Berg's 1.375). The value decreased as flow increased for any specific glottal angle. The transglottal pressure coefficient ranged from 1.00 to 1.82 with an average of 1.28. The average value for the uniform glottis was 1.63. Values decreased as flow increased. Pressure profiles will be compared to predictions from current glottal aerodynamic equations. [Work supported by NIH Grant 1R01DC03577.]

**4pSC14. Syllable-boundary magnitudes from jaw movement patterns.** Bryan Pardo and Osamu Fujimura (Speech and Hearing Sci., The Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210-1002)

The C/D model [O. Fujimura, J. Acoust. Soc. Jpn. (E) **13**, 39–48 (1992)] describes rhythmic patterns of utterances by amplitude-controlled pulse trains, each pulse representing a syllable or a boundary. The syllable magnitude (pulse height) controls mandibular movement. Assuming a direct proportional relation between the two variables, the syllable magnitude for each occurrence of five, nine, and Pine under various prominence conditions was inferred using Erickson's Pine Street dialogue data [Erickson, Fujimura, and Pardo, Lang. Speech **41**, 395–413 (1998)], which contained repeated corrections of the street address. A computational procedure has been devised to process semi-automatically microbeam data produced by two male and two female speakers. The timing of each syllable pulse was determined using the iceberg method [O. Fujimura, J. Acoust. Soc. Am. **99**]. Syllable durations were determined in proportion to the syllable magnitudes. Boundary magnitudes were represented by the gaps among syllable durations. The syllable and boundary durations thus determined reasonably agreed with acoustic durations. [This work was supported in part by NSF Grant Nos. SBR-951199B, SBR-9809046, and BCS-9977018 (O. Fujimura) and by ATR/MIC, Japan.]

**4pSC15. Effect of vocal fold tissue shear properties on phonation threshold pressure in a physical model of the larynx.** Roger Chan and Ingo Titze (Natl. Ctr. for Voice and Speech, Dept. of Speech Pathol. and Audiol., The Univ. of Iowa, Iowa City, IA 52242, chan@shc.uiowa.edu)

Previous analytical studies have shown that phonation threshold pressure ( $P_{th}$ ), the minimum subglottal pressure required to initiate and sustain vocal fold oscillation, is directly related to the viscous shear properties (dynamic viscosity or viscous shear modulus) of the vocal fold mucosa in vibration [Titze, *J. Acoust. Soc. Am.* **83**, 1536–1552 (1988); Chan, Ph.D. dissertation (1998)]. This relationship was verified empirically with a physical model of the larynx [Titze *et al.*, *J. Acoust. Soc. Am.* **97**, 3080–3084 (1995)], but no biological materials were tested in that study. This study attempted to establish the effects of biologically based tissue shear properties on  $P_{th}$ . Biomaterials with varying viscoelastic shear properties (abdominal subcutaneous fat, 0.01%–1.0% hyaluronic acid, and hyaluronic acid mixed with fibronectin) were incorporated into the artificial vocal fold mucosa of the physical model and self-sustained oscillation was induced. Results showed that no stable phonation could be established for fat, while higher  $P_{th}$  were consistently observed for higher concentrations of hyaluronic acid (with or without fibronectin), in correlation with their differences in viscous shear properties. These findings support the notion that hyaluronic acid could potentially be an optimal surgical bioimplant for the vocal fold mucosa, because it likely facilitates the ease of phonation. [Work supported by NIDCD.]

**4pSC16. Measuring vocal quality with speech synthesis.** Bruce R. Gerratt and Jody Kreiman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab, Los Angeles, CA 90095)

Previous research has demonstrated that listeners do not agree well when using traditional rating scales to measure pathological voice quality. Although these findings may indicate that listeners are inherently unable to agree in their perception of such complex auditory stimuli, an alternate explanation implicates the particular measurement method—rating scale judgments—as the culprit. A novel method of assessing quality—listener-mediated analysis by synthesis—was devised to assess this possibility. In this new approach, listeners explicitly compare synthetic and natural voice samples, and change speech synthesizer parameters to create acceptable auditory matches to voice stimuli. Thus this method replaces unstable internal standards for qualities like breathiness and roughness with externally presented, infinitely variable stimuli, thus overcoming one major hypothetical source of disagreement in rating scale judgments. To assess the reliability of this method, listeners were asked to adjust the signal-to-noise ratio for 13 pathological voices so that the resulting synthetic stimuli matched the original voices as well as possible. In a separate experiment, listeners judged the noisiness of the same stimuli using a traditional visual-analog rating scale. The relative reliability and validity of the two methods of quality assessment will be discussed, as will implications for models of voice quality perception.

**4pSC17. A comparison of linguistic and pathological breathiness using the LF model.** Melissa A. Epstein (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, melissae@ucla.edu)

This study investigated the hypothesis that breathiness used contrastively or to enhance a contrast in a language is as “breathy” as breathiness resulting from vocal pathology. Vowel samples were taken from four male and four female native speakers of English and four male and four female speakers with pathologically breathy voices. In the English data, a single pitch period was extracted and analyzed from an allophonically modal sample (following /ʔ/), a vowel-initial allophonically breathy sample (following /h/), and a vowel-medial allophonically breathy sample.

A representative vowel-medial cycle was selected for speakers with vocal pathology. To determine relative breathiness, vowels were inverse filtered and fit with the LF model of the glottal flow derivative; measurements of open quotient, glottal skew, excitation strength ( $E_e$ ) and the shape of the return phase were made from the LF smoothed pulses. Breathiness occurring allophonically after intervocalic /h/ did not differ significantly on any parameter from the vowels produced by speakers with pathologically breathy voices.  $E_e$  proved to be the best measure of relative breathiness. It distinguished the pathological samples from the normal samples (allophonically modal and allophonically breathy together); it also distinguished the allophonically breathy samples from the allophonically modal samples.

**4pSC18. Recognition of digit strings in noisy speech with limited resources.** Douglas O’Shaughnessy and Marcel Gabrea (INRS-Telecommunications, 900 dela Gauchetiere West, P.O. Box 644, Montreal, QC H5A 1C6, Canada)

Automatic recognition of sequences of spoken digits (e.g., telephone or credit card numbers) can be accomplished with excellent accuracy, even in speaker-independent applications over telephone links. However, even such relatively simple recognition tasks suffer decreased performance in adverse conditions, such as significant background noise or fading on portable telephone channels. If one further imposes significant limitations on the computing resources to be dedicated to a recognition task, then robust, limited-resource speech recognition remains a suitable challenge, even for a vocabulary as simple as the digits. Since connected-digit recognition over telephone lines is a very practical application, the amount of computer resources needed for a given level of recognition accuracy was investigated for different levels and types of acoustic noise. Rather than use a traditional hidden Markov model approach with cepstral analysis, which is computationally intensive and does not always work well under adverse acoustic conditions, simpler spectral analysis was used, combined with a segmental approach. The limited nature of the vocabulary (i.e., ten digits) allows this simpler approach. High recognition accuracy is maintained despite a massive decrease (versus traditional methods) in both memory and computation.

**4pSC19. The application of segment-based orthogonal transform in signal bias removal.** Wern-Jun Wang (Dept. of Commun., Natl. Chiao Tung Univ., Taiwan, ROC), Eng-Fong Huang (Chunghwa Telecommunication Labs., Taiwan, ROC), and Sin-Horng Chen (Natl. Chiao Tung Univ., Taiwan, ROC)

The iterative technique for signal bias removal (SBR) to minimize the effects of corrupted speech with unknown noisy sources was proposed for robust speech recognition. The SBR method was applied by considering a multiplicative spectral bias or equivalently, an additive cepstral bias. Two problems of this approach can be found. One is that it needs sufficient iterations to attain better results and the other difficulty is the lower accuracy of bias estimation. To overcome these drawbacks, a novel approach of SBR incorporated with orthogonal transform is adopted in this study. Instead of the conventional frame-based process, this method is a segment-by-segment approach. The basic idea of this approach is that the segment constituted by any feature element will be modeled using orthogonal transform coefficients. Owing to the characteristics of orthogonal transform, the high-order coefficients are bias-free and only the zeroth order coefficient is bias-corrupted. The high-order coefficients can therefore be used for bias estimation and the real bias will be deduced from the zeroth-order coefficient. Effectiveness of the method was confirmed by an

experiment of multispeaker noisy continuous Mandarin speech recognition. The significant improvement on computation and recognition rate can be achieved by this method.

**4pSC20. Modeling stop-consonant releases for synthesis.** Helen M. Hanson and Kenneth N. Stevens (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144-2500 and MIT Res. Lab. Electron., Cambridge, MA 02139, hanson@sens.com)

This study is part of a project leading to rule-based speech synthesis using the HLsyn synthesizer. In HLsyn, stop-consonant releases are generated by controlling the time variation of a constriction that is formed by the lips, the tongue blade, or the tongue body. In order to generate a consonant release that is perceptually acceptable and that has acoustic characteristics that match those of normal speech, it was found that the trajectory of the consonant release (cross-sectional area versus time) had to exhibit an initial rapid rise, followed by a delay in which the rise was interrupted, followed finally by a final rise. The burst at the consonant release is generated by the airflow through the constriction during the time that the rise is delayed. A model of the time course of this three-stage release of the articulator has been developed, taking into account the role of the intraoral pressure, the tapering formed by the articulator, and the glottal area. The model shows that the delay between the initial and final area increases progressively longer for labials, alveolars, and velars, in agreement with acoustic data on the burst duration. [Work supported in part by NIH Grant Nos. MH52358 and DC00075.]

**4pSC21. Quantization errors in formant estimation.** Gautam K. Vallabha and Betty Tuller (Ctr. for Complex Systems & Brain Sci., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431, vallabha@walt.ccs.fau.edu)

The locations of formants in a speech signal are usually estimated by computing the linear predictive coefficients (LPC) over a sliding window and estimating the peaks in the spectrum of the resulting all-pole LPC filter. The peaks are estimated either by solving for the roots of the LPC polynomial or by computing its DFT and finding the peaks in the magnitude spectrum. Three different sources of errors in this analysis were investigated using synthesized vowels: (1)  $F_0$  quantization: In addition to the expected result that error increases with  $F_0$ , it was found that the absolute error decreases with higher formant frequencies. (2) LPC+root solving: The location and bandwidth of a peak are usually estimated from the location of a pole on the  $z$  plane. It was found that this approximation is invalid if the formants have high bandwidths or if they are either too close or too far from each other. (3) LPC+DFT: In order to compensate for the quantization error introduced by the short-term DFT, a three-point parabolic interpolation scheme is usually employed for better estimation of peak locations. It was found that this compensation scheme is effective only for high-bandwidth formants. [Work supported by NIMH.]

**4pSC22. Pathologies not included: Survey of premillennium speech recognition software.** Renetta Tull (Dept. of Communicative Disord., Trace R & D Ctr., Univ. of Wisconsin—Madison, 1500 Highland Ave., Madison, WI 53705, tull@waisman.wisc.edu)

An October 1999 survey of several leading speech recognition software manufacturers revealed that the speech technology industry is unprepared to serve users with vocal pathologies in 2000. Companies surveyed had neither prepared for consumers with vocal pathologies nor had im-

mediate plans to alter systems to accommodate the pathological voice by 2002. Examples of vocal pathologies include leukoplakia, carcinoma, hyperfunction, polyps, and nodules. While the aforementioned pathologies involve direct impairment to the vocal mechanism (e.g., swollen pharyngeal mucous membranes, vocal fold polyps), neurological disorders such as Parkinsons disease also cause vocal pathologies (e.g., dysarthrias, spasmodic dysphonia). Some companies are beginning to incorporate natural language processing for consumers with aphasia; however, many of the top producers of speech recognition engines have not considered the pathological voice within their disability market. Acknowledging that the speech technology market is still new, speech recognition software developers note that lack of time and manpower prohibits concentration on groups with special needs. This research examines the feasibility of altering linguistic grammars of templates from healthy, prototypical voices to accommodate the phonetic changes that are observed in the pathological voice. Backus-Naur form and Chomskys syntactic structures are used to modify existing grammar templates.

**4pSC23. The perceived pitch of synthesized vowels with alternate pulse cycles.** Xuejing Sun (Dept. of Commun. Sci. and Disord., Northwestern Univ., Evanston, IL 60208, sunxj@nwu.edu)

Speech with alternate pulse cycles (APC), also known as diplophonia, double pulsing, etc., often brings problems to pitch determination algorithms. Although this phenomenon has been studied in light of nonlinear theory recently, to date, no study has been done to measure the perceived pitch of speech with APC. In the present study, a perception experiment was conducted to investigate the perceived pitch of speech with APC. Two synthetic vowels, /a/ and /i/, with fundamental frequencies at 140 and 220 Hz, both amplitude and frequency modulated, respectively, were used as stimuli. The modulation levels ranged from 0 to 0.9. The subjects were asked to judge the pitch of modulated vowels in reference to the synthetic vowels without modulation. The results show that the modulation level affects the perceived pitch in a nonlinear manner, and frequency modulation seems to have greater effects on the perceived pitch than amplitude modulation. In addition, different fundamental frequencies have different effects on the perceived pitch across modulation levels, whereas different vowel types appear to make no significant differences. Based on these results, a pitch determination algorithm was developed, and the performance of the algorithm appeared promising.

**4pSC24. Improved voice activity detection of noisy speech.** Min-sung Koh and Margaret Mortz (School of Elec. Eng. and Computer Sci., Washington State Univ., Spokane, WA 99202-1648, mkoh@eecs.wsu.edu)

This paper reports on the performance of a new algorithm for detecting voice activity in the presence of additive Gaussian noise. The proposed algorithm enhances the noisy speech before determining voice activity. In the speech enhancement, two noise reduction stages are used. The first stage uses Donoho's wavelet-based noise-reduction method. The second stage uses nonlinear spectral magnitude enhancement in the Fourier domain. After signal conversion back to the time domain and renormalization, simple energy thresholds are used to detect the presence of voice activity. Experimental results show that the new method can achieve good performance with preprocessing a speech signal even though a simple energy-based voice-activity detector is used. A new algorithm shows sig-

nificantly better performance when it is compared with the voice-activity detector specified in the ITU standard G.729 Annex B and other methods proposed recently.

**4pSC25. Speech processing using the average localized synchrony detection.** Ahmed M. A. Ali, Jan Van der Spiegel (Dept. of Elec. Eng., Univ. of Pennsylvania, 200 South 33rd St., Philadelphia, PA 19104, ahm@ee.upenn.edu), and Paul Mueller (Corticon, Inc., King of Prussia, PA 19406)

A new auditory-based speech-processing system based on the biologically rooted property of average localized synchrony detection (ALSD) is proposed. The system is a modification to the generalized synchrony detector (GSD) [S. Seneff, *J. Phonetics* **16**, 55–76 (1988)]. It generates a pseudospectrogram of the speech signal by detecting periodicity, while reducing the response to the individual harmonics of the fundamental frequency and the sensitivity to implementation mismatches. This is achieved without sacrificing the frequency resolution. Hence, it presents a more consistent and robust representation of the formants. The system is evaluated for its formant extraction ability while reducing spurious peaks. It is compared with other auditory-based front-end processing systems in the tasks of vowel, stop and fricative classification on clean speech from the TIMIT database, and in the presence of noise. The results illustrate the advantage of the ALS system in extracting the formants and reducing the spurious peaks, while preserving the frequency resolution. They also indicate the superiority of the synchrony measures over the mean rate in the presence of noise. [Work supported by Catalyst Foundation.]

**4pSC26. Relative contributions of passband and filter skirts to intelligibility of bandpass speech: Some effects of amplitude and context.** James A. Bashford, Jr., Richard M. Warren, and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin—Milwaukee, Milwaukee, WI 53201, bashford@uwm.edu)

Recently, Warren and Bashford [*J. Acoust. Soc. Am.* **106**, L47–L52 (1999)] reported over 90% intelligibility for everyday sentences reduced to a 1/3-octave band (center frequency 1500 Hz, slopes 100 dB/octave, peak levels 75 dB). Surprisingly, when the bandpass sentences were partitioned using 2000-order FIR filtering, the rectangular 1/3-octave passbands had only 24% intelligibility, while the filter skirts separated by a 1/3-octave notch had an intelligibility of 83%, despite their severe spectral tilts. The present study found that when monosyllabic word lists were substituted for the sentences, the whole-band intelligibility was 26%, the passband 4%, and the filter skirts 16%. Intelligibility also was measured for 1/3-octave sentences with peak levels ranging from 85 down to 35 dB: The whole band intelligibilities ranged from 90% to 68%, and the filter skirt pairs each had three to four times the passbands' intelligibility. These findings demonstrate the ability to compensate for extreme distortions of the spectral profile of speech. Hence, it is necessary to consider the contribution of frequencies outside the nominal passband as defined by the 3 dB downpoints along the filter slopes when dealing with filtered speech. [Work supported by NIH.]

**4pSC27. Implementation of high-quality text-to-speech synthesis for limited domain applications.** Sunil R. Shukla and Thomas P. Barnwell III (Ctr. for Signal and Image Processing, Georgia Inst. of Technol., GCATT Bldg., 250 W. 14th St. NW, Atlanta, GA 30318, shukla@ece.gatech.edu)

Though current TTS systems can produce unrestricted, intelligible speech using a limited database of prerecorded sounds, the quality of speech is highly artificial. By using larger speech segments, such as a combination of words and diphones, the number of segment boundaries

are reduced drastically, resulting in high-quality TTS that has a humanlike natural voice. Such a system can be useful for limited domain applications such as driving directions, weather advisories, etc. Experiments were conducted to generate natural sounding speech from words and diphones by modifying the pitch and duration of the units to match the prosody characteristics. The Festival TTS system, developed at the CSTR at the University of Edinburgh, was used to aid in the implementation of the experiments. The method for synthesis is pitch-synchronous residual LP. This method allows for producing smooth transitions between segments. A database of residuals and parameters for segments (words and diphones) is prepared. During synthesis, the segments are concatenated and the desired prosody is matched using time-domain modifications of pitch and duration. Using natural prosody, waveforms were synthesized, resulting in close to natural sounding speech. This demonstrates synthesized speech can sound identical to the original speaker, given the correct prosody.

**4pSC28. Robust speech localization using formant phase subtraction.** Parham Aarabi and Vaughan Pratt (Gates 4B, Dept. of Comput. Sci., Stanford Univ., Stanford, CA 94305, parham@stanford.edu)

This paper examines the benefits of a context aware time difference of an arrival-based speech localization system. This approach uses information about the formant structure of the speaker in order to compute a more robust time difference of arrival between two microphones. This ability allows microphone arrays to localize speech in environments with signal-to-noise ratios as low as  $-10$  dB, when the noise source is another speaker, and  $-4$  dB, when the noise source is white Gaussian.

**4pSC29. Comparing single- and two-channel telephone speech enhancement for elderly hearing-impaired listeners.** Amy E. Sheffield, Ashok Krishnamurthy (Dept. of Elec. Eng., The Ohio State Univ., Columbus, OH 43210, sheffield.20@osu.edu), Stephanie Davidson, and Lawrence L. Feth (The Ohio State Univ., Columbus, OH 43210)

Elderly listeners with sensorineural hearing loss have difficulty understanding speech over the telephone. Hearing aids and telephone amplifiers help, but many elderly individuals do not have the access to these devices. One possible solution to the problem is to “preprocess” the speech at the talker before transmitting it to the hearing-impaired listener. The purpose of this research is to compare processing techniques applied at the talker end that can significantly improve the speech understanding of these persons. Intelligibility enhancement of single- and two-channel amplitude compression schemes are compared by objective and subjective means. The articulation index (AI) and the diagnostic rhyme test (DRT) are used as the objective and subjective measures, respectively. In a preliminary study, we used normal listeners with speech that is passed through Moore's [(1999)] hearing loss model. The hearing loss model simulates the effects of loudness recruitment and reduced frequency selectivity as well as elevated thresholds of hearing. In addition, elderly hearing-impaired subjects were tested for a subset of the processing schemes. We present results comparing relative intelligibility of the various processing schemes. [Work supported by a grant from the Ohio State University College of Social and Behavioral Sciences.]

## Session 4pUW

## Underwater Acoustics, Signal Processing in Acoustics and Acoustical Oceanography: Model Based Processing for Sources in Motion II

Peter G. Cable, Cochair

*BBN Technologies, Union Station, New London, Connecticut 06320*

William M. Carey, Cochair

*Aerospace and Mechanical Engineering Department, Boston University, 110 Cummington Street,  
Boston, Massachusetts 02215*

Chair's Introduction—2:00

### Invited Papers

2:05

**4pUW1. A review of matched field processing and moving sources.** A. Tolstoy (ATolstoy Sci., 8610 Battailles Court, Annandale, VA 22003, atolstoy@ieee.org)

Over the past two decades or so there has been considerable work devoted to matched field processing (MFP), with particular emphasis on source localization. Most of that work has assumed a stationary source, a simpler problem in many ways. However, there has also been work done to investigate and even take advantage of source motion. Examples include: early work by Song (1990) and Song and Baggeroer (1990) based partially on research by Shan *et al.* (1985); work by Zala and Ozard (1991) generalized the Linear and Capon processors for source motion; Tran and Hodgkiss (1991) analyzed 200-Hz data for low speed sources. Alternately, one might consider the source to be fixed while the array moves, introducing synthetic aperture techniques as seen in Stergiopoulos and Sullivan (1989). MFP for a moving source can be a more challenging problem than for the fixed source but also offers new and different techniques with much promise. [Work supported by ONR.]

2:20

**4pUW2. Matched field processing with Doppler and the degrees of freedom in an acoustic channel.** Arthur B. Baggeroer (MIT, Cambridge, MA 02139, abb@arctic.mit.edu)

Matched field processing which incorporates source motion, or Doppler, has been examined by several investigators. There are two motivations: (i) long integration times improve signal-to-noise ratios, but this leads to signal gain degradation if the source transits a single resolution cell, and (ii) mainlobes interfere constructively and sidelobes destructively. The MFP Doppler processing can be separated into incoherent and coherent methods. In the former a series of short-term MFP outputs are incoherently combined while in the latter a time-varying replica is used. In both methods the number of degrees of freedom (DOF), i.e., the dimension of the representation basis set, in the acoustic signal is a critical parameter since it determines both MFP resolution and sidelobe properties. The DOF are limited by the minimum between number of sensors and the number of modes propagating in the channel. We examine the concept of degrees of freedom in an acoustic channel and the performance of Doppler-based MFP methods. [Work supported by ONR, Code 321US, Undersea Signal Processing.]

### Contributed Papers

2:35

**4pUW3. Spatial modulation using vertical arrays in an underwater acoustic telemetry channel.** Daniel Kilfoyle, James Preisig (Appl. Ocean Phys. and Eng. Dept., W.H.O.I, Woods Hole, MA 02543, dkilfoyle@whoi.edu), and Arthur Baggeroer (MIT, Cambridge, MA 02139)

Multiple propagation paths between sound sources and receivers in an underwater acoustic channel are a common component of underwater acoustic telemetry. The temporal reverberation that results may adversely affect performance measures such as bit error rate and output signal-to-noise (SNR) ratio. With the use of transducer and hydrophone arrays, however, resolvable paths may also be exploited as additional bandwidth

in a manner completely analogous to frequency bandwidth using a parallel channel model. In this discussion, we describe an experiment conducted in Buzzards Bay, MA, that utilized moving, vertical arrays to spatially modulate and demodulate two acoustic communication signals that fully occupied identical time-frequency spaces. Decomposition of the composite shallow water channel into two, parallel communication channels was accomplished solely with spatial degrees of freedom. Performance is compared with a conventional, single signal transmission. In addition, *in situ* channel impulse response measurements are analyzed to explain the success of the spatial modulation experiment. Performance improvements comparable to those yielded by a 3–6-dB SNR increase are achieved without using additional signal power. We will conclude with a brief discussion regarding the challenges as well as potential applications of the spatial modulation technique.

**4pUW4. Tracking performance of least-squares algorithms in time-varying ocean channels.** James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Adaptive least-squares algorithms are widely used to estimate the state of or compensate for the effects of time-varying acoustic channels in the ocean environment. A special case of such time-varying channels is those through which the signals emitted by moving sources propagate. Current methods of characterizing the tracking behavior of these adaptive algorithms are limited to cases where the algorithm accurately models the time-varying behavior of the channel or for which this behavior has a particularly simple form. A new technique is presented which allows the performance of adaptive least-squares algorithms to be characterized for more general classes of channel fluctuations. The technique shows that the performance can be interpreted in terms of model mismatch. New applications enabled by the technique are presented. [Work supported by ONR.]

3:20

**4pUW5. Motion compensation for a synthetic aperture sonar.** John E. Piper (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001, piperje@ncsc.navy.mil)

Synthetic aperture sonar is capable of producing high-resolution images. The method typically requires many small corrections to the timing of the data to compensate for motion perturbations. A redundant phase center approach is used to calculate these corrections. This approach uses channels that spatially overlap between pings. This allows construction of a correlation function that can be used to estimate motion-induced time lag differences. To improve the accuracy of this method the correlation function is fit to a quadratic function to find its maximum. Experimental data from a variety of at-sea tests is used in this analysis. Results indicate that significant improvements to the image resolution can be gained with this approach.

### Invited Paper

3:35

**4pUW6. Adaptive and synthetic aperture processing structures for high-resolution mine-hunting sonars deploying planar or cylindrical arrays.** Stergios Stergiopoulos and Amar Dhanantwari (DCIEM, 1133 Sheppard Ave. West, Toronto, ON M3M 3B9, Canada)

The developments reported in this paper have been devoted to minimizing the influence of partially directional noise fields as well as maximizing the array gain of mine-hunting sonar systems by means of signal processing. The effort of this study included the design of a generic beamforming structure that allows the implementation of nonconventional signal-processing techniques in integrated active/passive sonar systems. The nonconventional processing schemes of this study are adaptive and synthetic aperture beamformers that have been shown experimentally to provide improvements in array gain for signals embedded in partially correlated noise fields. The reported results show that synthetic aperture and adaptive processing schemes with near-instantaneous convergence can be implemented in mine-hunting sonars deploying 3D arrays. Moreover, practically realizable angular resolution improvements provided by the nonconventional beamformers are equivalent with those provided by the conventional beamformer of a three-times-longer physical aperture and for broadband FM type of signals. The same set of results demonstrates also that the combined implementation of a synthetic aperture and the subperture STMV adaptive scheme suppresses significantly the sidelobe structure of cw pulses for mine-hunting operations. In summary, the present paper attempts to address characterization of source signals detected by planar or cylindrical arrays of mine-hunting sonar systems with 3D image resolution capabilities.

### Contributed Papers

3:50

**4pUW7. Exploiting target motion for active matched-field depth estimation.** Granger Hickman and Jeffrey Krolik (Duke Univ., Box 90291, Durham, NC 27708)

This paper concerns the use of complex multipath propagation for discriminating the depth of a moving target with a low-frequency active sonar. Although matched-field processing techniques, which exploit full-wave modeling of multipath propagation, have been previously demonstrated in passive sonar applications, their use in active sonar has been precluded by uncertainty in the complex backscattering characteristics of the target. In this paper, a matched-field technique is proposed which exploits changes in the complex moving target return seen between consecutive sonar pings at a horizontal receive array. Because the method exploits only changes in the coherent combination of multipath arrivals, it does not require *a priori* knowledge of the target backscattering characteristics assuming its aspect is only slowly changing from ping to ping. In this paper, simulation results are presented which are modeled after a realistic Mediterranean scenario taken from the Shallow-Water Active Classification Experiment performed in November 1995. Matched-field depth estimation simulation results indicate that target depth accuracies of better than 20% of the channel depth can be achieved at ranges of 10 to 15 km. [Work supported by ONR.]

4:05

**4pUW8. Design and development of software tool for high-power sensor array design.** P. M. Joseph (Sharjah College, P.O. Box 5398, Sharjah, UAE, pmjosep@hotmail.com)

The transmitting and receiving characteristics of sonar transducer arrays depend on various factors like its geometry, number of elements, source strength, relative phases of the elements, etc. An array with narrow main beam and low sidelobes is desirable for a variety of applications. But, these are conflicting requirements. One of the requirements can be achieved only at the cost of the other. It is the designer's choice to select the parameters depending on the nature of application. Due to mechanical and electrical limitations, the measured array parameters are found to differ from the theoretically formulated ones, especially at higher drive levels. An array design package has been developed and presented in this paper for predicting the optimum array, based on the requirements of the user, taking into account various high-power electrical and mechanical limitations. In order to make the package more user friendly, the array design software has been split into different modules, each carrying out a specific task like array shape selection, element shape selection, power level selection, mode selection, etc.