Session 4aAAa

Architectural Acoustics: The Technical Committee on Architectural Acoustics Vern O. Knudsen Distinguished Lecture

Steven M. Brown, Chair
Steelcase, Inc., CD-5E-16, P.O. Box 1967, Grand Rapids, Michigan 49501

Chair's Introduction—9:00

Invited Paper

9:05

4aAAa1. Concert hall design: A consultant's perspective and retrospective. J. Christopher Jaffe (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington Ave., Norwalk, CT 06854)

Architectural environments for symphonic performances can perhaps be described as one of the most complex problems in physical acoustics due to the multiplicity of sound source–path–receiver systems present in listening spaces. On the other hand, people performed and enjoyed listening to music in architectural and outdoor spaces long before the development of architectural acoustics as a scientific discipline and well before there were acoustical consultants. This paper describes one practitioner's historical perspective and analysis of the problem as it relates to public concert hall design and musical performance over the last several centuries and how the interrelationship between the two has actually fashioned the traditional symphonic environment. Based on this analysis, a methodology will be discussed to successfully replicate traditional symphonic environments utilizing psychoacoustic rather than geometric guidelines. The results of applying this methodology in a variety of building types over a period of 40 years, e.g., concert halls, recital halls, music pavillions, multiuse halls, etc., will be discussed. Both physical and electroacoustical solutions will be presented to support the author's original analytic premise.

10:00–10:30
Discussion

THURSDAY MORNING, 4 DECEMBER 1997 SENATE/COMMITTEE ROOMS, 10:45 TO 11:45 A.M.

Session 4aAAb

Architectural Acoustics: Book Signing in Honor of ''The Sabines at Riverbank''

William J. Cavanaugh, Chair
Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Chair's Introduction—10:45

John Kopec will attend this session to sign copies of his book on the history of Riverbank Laboratories and the role of the Sabines (Wallace Clement, Paul Earls, and Hale Johnson) in the science of architectural acoustics.

Session 4aAB

Animal Bioacoustics: Biologically Inspired Acoustics Models and Systems II

Whitlow W. L. Au, Chair
Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Chair's Introduction—8:30

Invited Papers

8:35

4aAB1. Temporal and spectral information in echoes for biomimetic object recognition. Roman B. Kuc (Dept. Elec. Eng., Yale Univ., New Haven, CT 06520-8284)

Typical echolocation sounds produced by bats and dolphins have a wide bandwidth. This allows the echoes to contain information in the time and frequency domains. Temporal information includes the evolution of the energy in the perceived echo packet. However, if the initial echoes in a packet are weak, they may not exceed the perceptual threshold at a repeatable time, making temporal comparisons difficult. Spectral cues are more robust in that a time delay in an echo packet does not effect the power spectrum. However, because the emitted beam pattern and the reception patterns of the ears are frequency dependent, spectral comparisons are difficult. This paper describes methods for using temporal cues to aid in the spectral comparisons and methods for using spectral cues to aid in the temporal comparisons. The methods are illustrated using experimental results obtained with a biomimetic sonar mounted at the end of a robot arm that is used to recognize a variety of objects. [This research was funded through a grant from the NSF IRI-9504079.]

9:00

4aAB2. Real-time sonar classification of proud and buried targets using dolphinlike echolocation signals. Reid H. Shizumura (ORINCON Corp., 970 N. Kalaheo Ave., Ste. #C215, Kailua, HI 96734), Whitlow W. L. Au (Hawaii Inst. of Marine Biol., Kaneohe, HI), Gerald C. Moons (ORINCON Corp., Kailua, HI 96734), Paul E. Nachtigall (Hawaii Inst. of Marine Biol., Kaneohe, HI), Herbert L. Roitblat (Univ. of Hawaii, Honolulu, HI), and Robert C. Hicks (ORINCON Corp., Kailua, HI 96734)

A broadband active sonar system using dolphinlike echolocation signals and biologically inspired signal processing algorithms was developed to classify proud and buried targets in real time. A dolphin simulator transducer was attached to a bottom-crawling remotely operated vehicle and used to transmit a dolphin click (120-kHz center frequency, 39-kHz bandwidth, 50-μs duration) through seawater. Reflected target echoes were received and transmitted via cable to a shore-based computer system where the echoes were digitized at 1 MHz and subsequently processed. Two time-frequency representations of the echoes, one based on the short-time Fourier transform and the other on the Morlet wavelet, were processed in a hierarchical neural network system to derive target classifications. Echolocation returns were collected from six objects (cast iron pot, stainless steel sphere, glass jar, concrete tile, and coral rock) that were placed either on the ocean bottom or buried by bottom sediment. Echoes were separated into three target categories: (1) cast iron pot; (2) stainless steel sphere; and (3) a miscellaneous category consisting of the remaining four objects. Following supervised neural network training, the system was able to correctly identify 74%, 97%, and 88% of the category 1, 2, and 3 test echoes.

9:25

4aAB3. The adaptive silicon cochlea. Rahul Sarpeshkar (Bell Labs., Dept. of Biological Computation, Rm. 3L-404, 600 Mountain Ave., Murray Hill, NJ 07974, rahul@physics.bell-labs.com)

Low-power wide-dynamic-range systems are extremely hard to build. The biological cochlea is one of the most awesome examples of such a system: It can sense sounds over 12 orders of magnitude in intensity, with an estimated power dissipation of only a few tens of microwatts. An analog electronic cochlea that processes sounds over six orders of magnitude in intensity, while dissipating less than 0.5 mW, is described. This 117-stage, 100-Hz to 10-KHz cochlea has the widest dynamic range of any artificial cochlea built to date. This design, using frequency-selective gain adaptation in a low-noise, traveling-wave amplifier architecture, yields insight into why the human cochlea uses a traveling-wave mechanism to sense sounds, instead of using bandpass filters.

9:50–10:00 Break

10:00

4aAB4. Biologically inspired SCAT sonar receiver for 2-D imaging. James A. Simmons, Prestor A. Saillant, and Seth Boatright-Horowitz (Dept. of Neurosci., Brown Univ., Providence, RI 02912)

The SCAT (spectrogram correlation and transformation) model of biosonar in bats represents FM broadcasts and echoes as spectrographs with an integration time of 300–400 μs. The model then mimics signal processing by bats with two parallel computational paths: (1) determination of the delay of echo spectrograms using delay lines and coincidence detectors, and (2) deconvolution of the echo spectrum to determine delay separations smaller than the integration-time window [Saillant *et al.*, J. Acoust. Soc. Am. (1993)] to produce high-resolution echo-delay images. The new, binaural version of the SCAT model receives broadcasts and echoes at two ''ears'' and uses binaural algorithms to produce range/azimuth sonar images. The model incorporates several stages of binaural interactions that facilitate different functions including determination of azimuth itself, enhancement of target detection, and separation of adaptive imaging from these other functions. The model produces images that locate the principal glints in complex targets from only a few sonar broadcasts without having to sample many different target aspect angles. [Work supported by ONR and NSF.]

10:15

4aAB5. Target localization aboard a tone emitting robotic echolocation system with mobile receivers. V. Ashley Walker, Herbert Peremans, and John C. T. Hallam (Dept. of A.I., Univ. of Edinburgh, 5 Forrest Hill, Edinburgh EH1 2QL, Scotland, ashley@aifh.ed.ac.uk)

Using only *two* receivers, animals are able to determine the location of a target in *three* dimensions. Echolocating bats, which employ broadband calls, have access to spectral cues that may be used, in conjunction with interaural disparities of intensity (IIDs) and/or time (ITDs), to encode a 3-D target angle. This work addresses the question: what cues might a tone-emitting echolocator employ to determine target azimuth and elevation? To this end, a 6-degree-of-freedom robotic sensor head was built to investigate the direction cues that might be generated by the systematic pinnae scanning movements used by many species of tone emitting bats. The first strategy investigated was the use of receiver motion to rotate the SONAR horizon through discrete IID sampling positions so that a 3-D target angle can be resolved across IID readings taken at two or more receiver positions. Next, the change in amplitude measured by continuously scanning pinnae was used to create temporal cues, which vary systematically with the target angle. In this scheme, angular resolution depends upon scan speed (ms) rather than interaural separation (microsecond). Finally, the use of Doppler shifts to encode the target angle through the cosine law was investigated.

10:30

4aAB6. Estimating acoustic flow parameters for multiple echoes within a biologically motivated signal processing framework. Rolf Müller and Hans-Ulrich Schnitzler (Animal Physiol., Tübingen Univ., Morgenstelle 28, D-72076 Tübingen, Germany, rolf.mueller@uni-tuebingen.de)

It is hypothesized that proportional changes in echo amplitude and carrier frequency might aid obstacle avoidance in cf bats by means of acoustic flow perception. This approach appears to be feasible in principle for single targets, however, obstacle avoidance, unlike prey capture in midair, is invariably prone to constitute a multiple target problem, which has to be addressed in assessing the hypothesis' validity. In order to map a compound excitation pattern formed by superposition of responses to multiple echoes back into target space by means of an acoustic flow approach, instantaneous (carrier) frequency and amplitude together with their first-order derivatives must be recovered individually for each echo to be perceived as a separate entity. The primal sketch serving as a substrate for this operation is impoverished in that it is only preserving response envelopes. At the same time, the system design might capitalize on knowledge about the transfer functions of the filter bank's channels as well as at least a rough guess about the echo spectrum. Consequently, it is attempted to explain compound excitation patterns as a sum of superimposed responses to individual echoes, which are known for any given frequency location to a scaling constant.

10:45

4aAB7. The ecological approach toward ultrasonic perception. Herbert Peremans and John Hallam (Dept. of Artificial Intelligence, Univ. of Edinburgh, Forrest Hill 5, Edinburgh EH1 2Ql, UK)

Navigation seems to automatically entail the construction and use of environment models. In this paper it is shown how, in certain situations, the information provided by an echolocation system contains invariants that can be used directly to control the navigation behavior of an agent without the need for an explicit environment model. In particular, the case of an echolocating agent moving through an opening between two reflecting objects is analyzed. The transmit signal, duration 2 ms, is a frequency modulated sinusoid, hyperbolic frequency sweep from 100 down to 20 kHz. The processing performed on the raw measurement results, i.e., the signals picked up by the microphones, is modeled on the peripheral auditory processing of the mammal brain. The output of the model is an approximation of the short-time amplitude spectrum of the received signal. It will be shown that the control of the agent can be expressed solely in terms of invariant properties of the patterns appearing on the outputs of this model. This approach not only predicts the flight path of bats flying between vertical wires but it also significantly simplifies the control of a mobile robot moving between obstacles as will be shown.

11:00

4aAB8. Emergent collective behaviors in a simplified model of dolphin echolocation and predation. Thomas J. Hayward (Naval Res. Lab., Washington, DC 20375-5350)

A stochastic modeling framework for representing and simulating marine-mammal spatial distributions, collective motions, and vocalization occurrence times was previously developed [J. Acoust. Soc. Am. **101**, 3197(A) (1997)]. Emergent behaviors in these models include clustered and polarized motions simulating collective motions of marine mammals. The present work explores the potential of these models for representing and simulating more complex acoustic emission and collective motion behaviors related to short-range echolocation and predation by dolphins. A simplified predation scenario is constructed by including predator, prey, and nonprey species in the collective motion model. Echolocation is simulated by a simplified, sonar equation-based representation of the backscattered acoustic energy, with different backscattering characteristics for predator, prey, and nonprey individuals. Parameters determining echolocation signal types and collective motion characteristics are then allowed to vary in a random search that favors parameter values leading to higher prey capture rates. Emergent collective behaviors, including echolocation signals and prey pursuit and capture, are then observed in computer-generated video animations. The potential of the models for representating sound generated by groups of marine mammals and for interpreting observed marine mammal collective behavior is discussed. [Work supported by ONR Base funding at NRL.]

Session 4aBV

Bioresponse to Vibration/Biomedical Ultrasound and Physical Acoustics: Lithotripsy, Shock Waves and Bubbles and Tactile Estimation

Thomas J. Matula, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Contributed Papers

8:30

4aBV1. *B*-scan ultrasound monitoring of cavitation acitivity in and around the kidney during shock wave lithotripsy. Robin O. Cleveland, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), David A. Lifshitz, Bret A. Connors, Lynn R. Willis, Andrew P. Evan (Indiana Univ. Med. School, Indianapolis, IN 46202), and James E. Lingeman (Methodist Hospital, Indianapolis, IN 46202)

Acoustic cavitation is thought to play a major role in both stone comminution and tissue injury during shock wave lithotripsy (SWL). Although ultrasound machines cannot directly observe cavitation events, remnant gas bubbles appear as echogenic regions. We treated young pigs in an unmodified Dornier HM3 and used a standard clinical ultrasound machine (Brüel and Kjær 3535) to image the kidney while SWL was being administered. We observed large regions of *in vivo* echogenicity within the renal pelvis (collecting system), which has not been reported before. When the ureter was blocked, up to 30% of the renal pelvis was covered with echogenicity. Clouds of echogenicity progressed through the collecting system with consecutive shots. The echogenic regions in the renal pelvis could be temporarily enhanced when nondegassed fluid was injected through a ureter catheter. Echogenicity was observed on the posterior and anterior surfaces of the kidney and appeared to occur at sites of subcapsular hematomas. The presence of echogenicity is consistent with the claim that cavitation plays an important role in both stone comminution and tissue damage. [Work supported by NIH and the Hunt Fellowship.]

8:45

4aBV2. Collapse of cavitation bubble induced during extracorporeal shock wave lithotripsy. Tetsuya Kodama and Kazuyoshi Takayama (Shock Wave Res. Ctr., Inst. of Fluid Sci., Tohoku Univ., 2-1-1 Katahira, Aoba Ward, 980-77 Japan, kodama@ifs.tohoku.ac.jp)

The interaction of air bubbles attached to gelatin surfaces, rat's livers, or rat's abdominal aortas with underwater shock waves was investigated to clarify the tissue damage mechanism by cavitation bubbles induced during extracorporeal shock wave lithotripsy. The overpressure of the shock wave was 10.2$\pm$0.5 MPa $(n=4)$. The initial bubble radii varied from 0.12–3.06 mm. The subsequent collapse of the bubbles was recorded by a high-speed framing camera. The liver cell damage was histochemically evaluated. The bubble attached to the gelatin or rat's liver surface migrates away from the surface with an oscillatory growth/collapse behavior after the shock wave interaction. The penetration depth of the liquid jet into the gelatin and the diameter of the subsequent damage pit on the surface depend on the initial bubble radius. The bubble near the rat's liver surface or aorta surface tends to show the same behavior as for the gelatin. The elongation and split of the nuclei in the liver parenchyma in the direction of the liquid jet and the increase in the cell density within the circumference of the injured region are histochemically revealed.

9:00

4aBV3. Damage of living tissue cells in water by shock waves and mathematical model. Masaaki Tamagawa (Faculty of Energy Sci., Kyoto Univ., Sakyo-ku, Kyoto 606-01, Japan) and Teruaki Akamatsu (Neyagawa, Osaka 572, Japan)

In this study, the damage experiments of the living tissue cells (red blood cells and cancer cells) [M. Tamagawa and T. Akamatsu, Proc. 20th Int. Shock Wave Symp., Pasadena, USA, 517–518 (1995)], and the experimental model for cells (microcapsules including dye stuff) by plane shock waves are executed. It is shown from the results that using a microcapsule is an effective method for sensing injury to living tissue cells. To explain these phenomena, the living tissue cells are modeled mathematically as one and two spherical elastic shells filled with liquid toward the shock wave. Using stationary and transient analysis of a spherical shell, the dynamic characteristics of the one living tissue cell and two with a mutual interaction are evaluated. The results show that damage to living tissue cells depends on: (1) the elastic modulus of the cell, (2) the bulk modulus of intracellular material, (3) the thickness of the cell membrane, (4) the distance between the cells on shock-induced damage, and (5) the rise time of the shock wave.

9:15

4aBV4. Overpressure reduces acceleration of thrombolysis due to ultrasound. E. Carr Everbach (Swarthmore College Eng. Dept., Swarthmore, PA 19081-1397, ceverba1@swarthmore.edu), Janice White, and Charles W. Francis (Univ. of Rochester, Rochester, NY 14642)

It has been shown that application of cw ultrasound at 1–8-W/cm2 SATA intensities and megahertz frequencies can accelerate the dissolution of blood clots when a clot-lysing agent is present. To determine whether or not a cavitational mechanism is responsible for ultrasound accelerated thrombolysis (UAT), a hyperbaric chamber was constructed that can apply 10 atm of static air overpressure to an *in vitro* exposure apparatus. Five hundred μl fibrin clots overlaid with a clot-lysing agent were exposed to 4-W/cm2 SATA ultrasound at 1 MHz and lysis quantified via release of bound radiolabel as a function of overpressure. More than half the acceleration due to ultrasound was removed, suggesting a bubble mechanism is responsible at least in part for UAT. Transmission and reflection of 20-MHz tone bursts of 0.5 μs duration provide evidence that bubble activity was reduced concomitantly during overpressure. [Work supported by NIH.]

9:30

4aBV5. Bubble collapse emissions suggest mechanism for transient response imaging. E. Carr Everbach (Swarthmore College, Eng. Dept., Swarthmore, PA 19081-1397, ceverba1@swarthmore.edu), Shouping Li, and Thomas R. Porter (Univ. of Nebraska Med. Ctr., Omaha, NE 68198-2265)

Transient response imaging (TRI) is the increase in apparent myocardial contrast with reduced image frame rate (PRF) when perfluorocarbon-filled echo-contrast agents are infused. TRI allows improved estimation of myocardial perfusion, but its mechanism is not known. Recent work sug-

gests that contrast microbubbles are destroyed by imaging at conventional frame rates (30–40 Hz), while less destruction occurs at cardiac-synchronized frame rates. Using a 20-MHz passive cavitation detector, acoustic emissions were recorded from within the myocardium of two anesthetized dogs *in vivo* during ultrasonic imaging with a clinical scanner. Strong correlations among root-mean-square detector output, videodensity, and PRF point to bubble emissions themselves as possibly responsible for the improved contrast. [Work supported by Nebraska AHA.]

9:45–10:00 Break

10:00

4aBV6. Nonlinear effects in HIFU propagation and attenuation in biological tissue. Vera A. Khokhlova, Oleg A. Sapozhnikov (Dept. of Acoust., Phys. Faculty, Moscow State Univ., Moscow 119899, Russia, vera@na.phys.msu.su), and Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105)

For therapeutic applications of medical ultrasound, such as hyperthermia, acoustic surgery, and hemostasis, it is desirable, for precise targeting, to induce effective local absorption of the acoustic energy at a small focal area of the acoustic beam. A source with high-focusing gain can be used to achieve the most effective regime of focusing, in which nonlinear effects increase dramatically very close to the focal point. Sharp shocks are developed and the corresponding absorption rate increases at the focus without extra nonlinear attenuation in the prefocal region. In an earlier presentation [Khokhlova *et al.*, J. Acoust. Soc. Am. **98**, 2944 (1995)], a modified spectral approach was used that enables us to model propagation of nonlinear plane waves, including shocks, in biological tissue. Here, this method is extended for the description of high-intensity focused sound beams. The acoustic pressure field is calculated numerically for specific focusing gains and tissue absorption over a wide range of the source amplitudes. The effect of nonlinearity, focusing gain, absorption on the waveforms, amplification and spatial localization of the intensity, heating rate, and positive and negative peak pressure at the focus is discussed. [Work supported by ONR and CRDF.]

10:15

4aBV7. Remote generation of shear wave in soft tissue by pulsed radiation pressure. Victor G. Andreev, Vladimir N. Dmitriev, Oleg V. Rudenko (Phys. Dept., Moscow State Univ., Moscow, Russia 119899), and Armen P. Sarvazyan (Artann Labs., East Brunswick, NJ 08816, armen@mail.crisp.net)

Remote generation of shear waves in tissues by a focused ultrasonic beam is the basis of a new acoustic method of medical diagnostics, namely, shear wave elasticity imaging [A. P. Sarvazyan, U.S. Patent No. 5,606,971, 1997]. The feasibility of SWEI was demonstrated recently in experiments with optical and NMR detection of ultrasonically induced shear waves in tissue phantoms. In the present study an SWEI system with ultrasonic detection of shear waves was designed and tested. The shear wave was excited inside an inhomogeneous tissue phantom using radiation force generated by the reflection of ultrasound from internal inhomogeneity. A focusing transducer of 6 cm in diameter and with 6-cm focal length was used. The carrier frequency was 2 MHz, and the intensity was varied in range from 10–30 W/cm^2. The rectangular envelope was 0.2 ms in duration. Induced shear motion was evaluated using a 10-μs-long 3-MHz ultrasonic pulse by measuring the time delay for the pulse backscattered by moving inhomogeneities in the tissue. Maximum displacement in the propagating shear wave was over 10 μm, which agrees with theoretical predictions.

10:30

4aBV8. Nonlinear acoustic parameter of trabecular bone. Dimitri M. Donskoy (Davidson Lab., Stevens Inst. of Technol., Hoboken, NJ 07030, ddonskoy@stevens-tech.edu) and Alexander Sutin (Stevens Inst. of Technol., Hoboken, NJ 07030)

This work is an attempt to measure the nonlinear parameter of porous trabecular bone. Two different methods were applied. One method involves an AM ultrasonic signal transmitted through a bone submerged into water. A part of the signal energy transforms into the low-frequency (the frequency of modulation) signal due to the nonlinearity of the sample. The larger the nonlinear parameter of the sample, the more intensive such a transformation. This low-frequency signal is then picked up with a hydrophone and measured. The other method employs the effect of modulation of higher-frequency ultrasound by lower-frequency vibration applied directly to the bone. The same high frequencies and low frequencies were used in both tests, so the independently measured nonlinear parameters of the samples (cut from the same bone) could be compared. High dissipation of ultrasonic waves in trabecular bone in the frequency range above 100 kHz made it difficult to measure the nonlinear effects. The nonlinear parameter in the frequency range below 100 kHz was able to be measured reliably. Thus the nonlinear parameters were measured for three ultrasonic frequencies 26, 37, and 60 kHz modulated by the vibration frequencies 4.6 and 9.8 kHz. *In vitro* determined values of the nonlinear parameter for bovine bone used in these tests were in the range 80–120, which is an order of magnitude higher than for nonporous media (e.g., compact bone and liquid). [Work supported by NASA.]

10:45

4aBV9. Method of pathology recognition in bio-organs by multifrequency ultrasound signals. Lena I. Oboznenko (Acoust. and Acoustoelectron., Dept. of Electron., Ukrainian Natl. Tech. Univ., Kiev, Ukraine 252057)

Ultrasound methods of pathology recognition in bio-organs are proposed. They are based on the application of multichannel sounding signals (or wide band signals). The peculiarities of the proposed methods take into account the specific character of the biomedium in the ultrasound diagnostic. The nuances of the effective application in the ultrasound diagnostic of the following method of recognition are analyzed: application of the natural resonance frequencies of the bio-organs without signal form recollection (vessel, red corpuscles, leucocytes, etc.), by apportionment of complex poles of the Laplas transformation function (R. Prony's method), application of the stepped influence with organ boundary restoration by impulse reaction and by the algorithm of the form restoration with projection (J. Young's method and V. Bojarsky's method, J. Radon transformation), structure methods of the two-dimensional bio-organ form picture (the grammar method), multichannel irradiation of the bio-organ with its phase-frequency characteristic synthesis of three-dimensional bio-organ transformation by its Ess. in LF, MF, HF bands. The numerical data concerning the form compensation of the spherical and elliptical elastic objects and recognition of the red corpuscles and leucocytes in blood vessels are deduced. Frequency band: 1–10 MHz, feeler length: one period, discrete radiation: 1, 2, 4, 6, 8, 10 MHz.

11:00

4aBV10. Tactile estimation of roughness and size. Ronald T. Verrillo, Stanley J. Bolanowski (Inst. for Sensory Res., Syracuse Univ., Syracuse, NY 13244), and Francis P. McGlone (Port Sunlight Lab., Bebington Wirral L63 3J2, England)

Subjects estimated the subjective magnitude of roughness and the size of steel balls by the method of absolute magnitude estimation. Using the index finger and thumb, estimates were made of 11 grades of sandpaper with particle diameters ranging from 16–905 μm. The results describe a

power function and show spatial summation when stimuli are applied simultaneously to two sites (finger and thumb). The same subjects judged the size of nine steel balls ranging in diameter from 0.4–6.4 mm. Estimates were made with combinations of the index finger, volar forearm, and thumb. Again, all of the results were power functions and spatial summation of tactile stimulation was evident. The results are discussed in terms of theoretical considerations and practical application. [Work supported by Unilever and NIH, PO1DC00380].

THURSDAY MORNING, 4 DECEMBER 1997 FORUM ROOM, 8:55 TO 11:35 A.M.

Session 4aEA

Engineering Acoustics: Acoustic Systems Designed for Harsh Environments

Fred C. DeMetz, Chair
Allied Signal Ocean Systems, 15825 Roxford Street, Sylmar, California 91342

Chair's Introduction—8:55

Invited Papers

9:00

4aEA1. Reliable high-temperature electronics for instrumentation and data acquisition. Ben Gingerich (Honeywell Solid State Electron. Ctr., 12001 State Hwy. 55, Plymouth, MN 55441)

Honeywell is developing and has introduced the first products of the HTMOS™ product line, which is targeted for use in systems operating in severe high-temperature environments. The HTMOS™ line of electronic products incorporates Honeywell's oxide-isolated high-temperature process and high-temperature circuit design methodologies. These components are being developed to meet applications in instrumentation, data acquisition, and control where hostile environments require a large temperature range of -55 to $+225$ °C with an extended operating life of 5 years at elevated temperatures. Honeywell has introduced a quad operational amplifier, quad analog switch, an 8-bit microcontroller, and a 32 K$\times$8 SRAM. Additional products are planned for introduction in 1997. This paper will discuss the high-temperature features of the silicon on insulator (SOI) process technology and product design approaches for reliable high-temperature electronic components. The paper will also discuss the existing product performance in a high-temperature environment.

9:25

4aEA2. Pressure-tolerant network interface cards for sensor arrays. John Walrod (Planning Systems, Inc., 21294 Johnson Rd., Long Beach, MS 39560, john_walrod@psilongbeach.com)

Pressure-tolerant ATM and Ethernet network interface cards (NICs) were developed for towed array data acquisition and telemetry. The card functions include sensor signal conditioning, analog-to-digital conversion, network protocol processors, and network interface transceivers. Design and fabrication techniques were developed to achieve space constraints, power constraints, and operation at 2500-psi pressure in oil or saltwater submersion without the use of pressure vessel housings. The development included the testing of candidate electronic components and NIC assemblies for pressure, temperature, and oil tolerance. Components and assemblies which failed testing were analyzed, and design rules were established for future designs. Generally, the elimination or minimization of voids by proper component selection and fabrication was found to be critical. Pressure-tolerant fabrication techniques were developed such as temperature cycle screening, encapsulation in low-viscosity polyurethane, pressurized curing, and compression sealed wires. High-density printed circuit boards and minimalist circuit designs were used to minimize NIC size. Low-voltage components, selective component location, and thermally conductive packaging were used to minimize operating temperatures. Commercially available components and open standard interfaces were used to minimize costs.

9:50

4aEA3. Comparison of a thermally biased PMN-PT transducer and a conventional PZT-based underwater source. Scott A. Hudson (AlliedSignal Ocean Systems, 15825 Roxford St., Sylmar, CA 91342)

Recent research into PMN (lead magnesium niobate)-based electrostrictive materials have led to compositions with strain rates significantly higher than that of piezoelectric ceramics. The large strain rate makes it ideal for use in high-power SONAR transducers. However, unlike piezoelectric ceramics, the properties of PMN vary nonlinearly with temperature. The highest performance PMN compositions exhibit optimal characteristics at temperatures near 85 °C. AlliedSignal has solved the engineering challenges of implementing a high-temperature PMN composition in a SONAR transducer and demonstrated the performance at NUWC Lake Seneca test facility. This paper discusses the properties of high-temperature PMN materials and presents the test results of a PMN transducer compared with a similar transducer which utilizes conventional piezoelectric ceramic. The PMN transducer produced roughly 6 dB higher source level than the PZT transducer. Electromechanical coupling and bandwidth of the two transducers was similar.

10:15–10:30 Break

10:30

4aEA4. Development of an advanced vibratory source for borehole seismology. Bjorn N. P. Paulsson (Paulsson Geophysical Services, Inc., 1300 Beach Blvd., La Habra, CA 90631-6374)

Single-well seismology, reverse vertical seismic profiles (RVSP's) and Crosswell seismology are three new seismic techniques for obtaining much higher resolution images of oil and gas reservoirs than what is currently obtainable with surface seismic techniques. Borehole seismology involves inserting both the source and the seismic receivers in oil or gas wells. In the past these methods have been limited to short distances between source and receivers in shallow wells. In 1997 a new borehole seismic source and receiver system has been deployed with a capability of handling source–receiver spacings of more than 2000 m in rock with a Q value of more than 20, and well depths as great as 6500 m. This advanced borehole seismic data acquisition system consists of a powerful, clamped, swept-frequency, vibratory source which is nondestructive, a multilevel receiver string of clamped, three-component geophones, and an acquisition system for reverse VSP's. The downhole vibrator produces high-quality S waves, with controlled polarization, as well as P-wave direct arrivals and reflections. This paper discusses the design and performance of this system.

10:55

4aEA5. Acoustic systems for well logging at extreme temperature and pressure. W. D. Squire and H. J. Whitehouse (Linear Measurements, Inc., 4174 Sorrento Valley Blvd., San Diego, CA 92121)

Various acoustic systems will be described briefly that have been studied and developed by one or both of the authors for use in well logging. Most of these systems have been developed with the goal of operating at extreme temperature (in some cases as high as 200 °C). A few of them have been developed with the additional goal of operating at extreme pressure (in some cases as high as 20 000 psi). A list of the systems to be discussed follows: a wireless telemetry system to transmit data from the bottom of the well to the surface, utilizing the drill pipe as an acoustic wave guide with zero-order torsional waves on the drill pipe as the data carrier; an oil and gas well passive listening device or an acoustisonde (a stethoscope or, as called by the petroleum industry, a noise tool); a wireless telemetry system to transmit data from the bottom of the well to the surface, utilizing the drill pipe as an acoustic wave guide with acoustic pulses in the drilling fluid (drilling mud) inside the drill pipe as the data carrier, with a data rate greater than existing mud pulse systems; an acoustic borehole televiewer technology with an order of magnitude finer resolution than existing systems; and advanced pressure release materials for use in acoustic transducers at extreme temperature and pressure.

Contributed Paper

11:20

4aEA6. Perspective directions of vector-phase methods application in acoustic-seismic environmental monitoring. Valery B. Mit'ko (Dept. of Marine Information Systems, Electrotechnical Univ. 5, Prof. Popov St., St. Petersburg, 197376, Russia)

Estimation of harmful influence parameters of acoustic-seismic fields is important in environmental complex monitoring to decide such problems as the determination of infrasound harbinger of tsunami, earthquake. Such enormous engineering structures as dams, power stations heat changers, and main pipelines are also infrasound sources and should be under control. Theoretical methods of sound field estimation are noneffective in low and underlow frequencies because of the simplifying assumptions. The mathematical model of vector-phase methods is based, for example, on solving for the integral on the closed surface. It permits one to define vector-kinematic characteristics of an acoustic field and its spatially-temporary distribution in controllable volume. The development of elastic wave-field parameter measurements and the creation of multicomponent receiver parks allow one to carry out background parameter measurement or responses to pulse influences with the purpose of heterogeneity localization in sedimental or radical layers of the researched area with minimal expense of energy and means for nature measurement realization. The presence of spectral composites as a result of local heterogeneities in the research area was established. It was noted that stream of power measurements provide the increasing signal/noise ratio up to 16–20 dB in different conditions.

THURSDAY MORNING, 4 DECEMBER 1997 COUNCIL ROOM, 10:30 A.M. TO 12:00 NOON

Session 4aED

Education in Acoustics: Take Fives—Sharing Ideas for Teaching Acoustics

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Uwe J. Hansen, Cochair
Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Do you hava a novel demonstration, a new laboratory experiment, a favorite video, a recorded sound example, or a new idea for teaching acoustics which you are willing to share with your colleagues? At this session a sign-up board will be provided for scheduling presentations. No abstracts are printed. Presenters are encouraged to have handouts to distribute. Multiple presentations are acceptable (not consecutively). Presentations are limited to 5 minutes. Keep them short! Keep them fun!

Session 4aNS

Noise and Engineering Acoustics: Noise Modeling and Outdoor Sound Propagation

Victor W. Sparrow, Chair
Graduate Program in Acoustics, Pennsylvania State University, 157 Hammond Building, State College, Pennsylvania 16804

Contributed Papers

8:30

4aNS1. The application of geographic spatial analysis in modeling traffic noise propagation away from a freeway. John C. Bennett (Dept. of Geography, San Diego State Univ., 5500 Campanile Dr., San Digeo, CA 92182-4493)

This paper demonstrates the use of a geographic information system (GIS) to perform a spatial analysis of traffic noise away from a freeway. The more traditional analysis procedures used by the Federal Highway Administration (FHWA, 1978 and 1982) employ a noise model that is range limited and not designed to handle large-scale topographic variation near a freeway noise source. A digital elevation map of a portion of a freeway in San Diego County that includes a mesa and valley landscape is incorporated into a GIS where a spatial analysis of ray path geometry, substrate, and vegetation is used to evaluate measured noise levels of traffic. Preliminary results indicate that a GIS model of traffic noise provides a more realistic basis for predicting noise impacts, particularly at longer ranges in areas of variable topography.

8:45

4aNS2. Six-year study of IMSA road racing noise. Gordon Bricken and Christopher Jean (Gordon Bricken and Assoc., Inc., 1621 E. 17th St., Santa Ana, CA 92705-8518)

The Del Mar Fairgrounds, San Diego, hosted the IMSA Grand Prix for 6 years from 1987–1993. The project provides a unique long-term study of professional road racing noise in a unique setting, a narrow valley just east of the coastline and within an environmentally sensitive lagoon. Sponsors were required to perform community noise monitoring at selected locations in the adjacent hills and lagoon. Both human and bird responses were examined. The monitoring documented the hour by hour variations in the maximum and average noise levels at up to 12 separate locations. The results indicated the thresholds which would trigger community response, the sensitivity of the response to diurnal and short-term wind shift, the effect of the inversion layer, the effects on bird patterns, and the statistical distribution of vehicle noise levels for the IMSA class vehicle. Basic findings revealed community response was triggered when levels exceeded 70 dBA Lavg. Downwind levels tended to increase at 1 dBA/mph over 5 mph. When the track was within an inversion, levels increased 10 dBA. The mean vehicle level at 15 m was 105 dBA with a 4-dBA standard deviation.

9:00

4aNS3. Nine-year study of NHRA drag racing noise. Gordon Bricken and Christopher Jean (Gordon Bricken and Assoc., Inc., 1621 E. 17th St., Santa Ana, CA 92705-85181)

Pomona International Raceway, Pomona, California is the premier venue for drag racing sanctioned by the National Hot Rod Association. Community noise was studied for 9 years from 1984–1993. The project provided a unique long-term study of professional drag racing noise at 14 locations in a built-up community adjacent to the starting line. The object was to establish a basis for documenting the effects of noise mitigation's planned for a renovated facility. The study revealed startling diurnal variations in noise, year to year variations, and the wide variety of emission characteristics of racing types. The simultaneous community and trackside locations permitted accurate correlation of events and established the basis for predictive models. The large data base permitted creation of a benchmark location by which the construction mitigation measures could be verified in a statistically valid manner. The study also revealed that the data spread is so wide that only multiple event sampling will provide an accurate picture of the community noise levels.

9:15

4aNS4. Effectiveness of random edge barriers: Further studies. Eric J. Rosenberg and Ilene J. Busch-Vishniac (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712, erosenberg@mail.utexas.edu)

Previously, the idea of introducing a random height fluctuation to the conventional straight edge of a noise barrier was discussed as a way to increase insertion loss and improve acoustic performance. Here, an exhaustive set of experiments is discussed to more accurately discern the effects of the random edge. Experimental results were obtained using a spark source and a plywood barrier with interchangeable metal edges serving as the various edge profiles. Unlike the past experiments, which reported peak-to-peak sound pressures, results were obtained for the energy spectral density. The results indicate that the insertion loss data obtained with a random edge are not solely a function of Fresnel number as they are with a straight edge, and that local effects may dominate when considering the energy measured at a point in space. [This work was supported by a Texas Advanced Technology grant.]

9:30

4aNS5. Approximation to study the acoustic field diffracted in the shadow of a half-infinite barrier by the boundary element method. Antonio Sanchis, Albert Marín, José Romero, and Alicia Giménez (Appl. Phys., Dept., High Tech. School of Industrial Eng., Polytech. Univ., P.O. Box 22012, 46080 Valencia, Spain)

The application of the BEM (direct form) for the external domain, with a lineal combination of Helmholtz's equation integral and their derivative (Burton, Miller) makes possible the calculation of the unique solution for each number of waves. The fundamental characteristics of this model that works with lineal sound source are: (1) The calculation surface is the topographical profile adapted with adjacent plans, discertizated in elements defined by two extreme nodes and one central node. (2) The acoustic impedance of each element is determined starting from the Delany equation. (3) The unknowns are the sound field in each node ϕ_i. (4) A form function ϕ is determined starting from the ϕ_i of each node. (5) Helmholtz's operators are obtained by integrating the ϕ function in each element. (6) By solving the equation system with these operators, the values of the ϕ_i, and the function ϕ, are obtained. (7) By integrating Helmholtz's equation integral starting from ϕ, the value of the sound field in each point of the space is obtained. Finally, the model is applied to several simple barriers for which there are experimental values. The comparison of the obtained outputs shows the validity of the calculation.

9:45–10:00 Break

10:00

4aNS6. The calculation of curved path sound propagation in temperature inversion conditions. Robert L. Bronsdon (Walt Disney Imagineering, 1401 Flower St., Glendale, CA 91221)

Sound propagation models have typically been used to develop noise control solutions for a wide variety of troublesome sources from highways, power plants, and airports to stadiums and outdoor amphitheaters. Almost all of these models assume that there is no variability in the atmospheric properties of the propagating medium but this assumption is almost invariably false. Predictions from these models are usually validated during the day when propagation conditions are favorable, producing lower levels than the models calculate. The problems usually come at night, when temperature and wind variations in the atmosphere create propagation conditions which are not favorable, resulting in much higher noise levels in the community, often on the order of 15–20 dB. This paper describes work which has been undertaken by the Walt Disney Company in an effort to produce a more precise computer model for these unfavorable conditions. The model developed is based on Gaussian beams, a hybrid technique developed to retain some of the fundamental aspects of ray tracing while incorporating some of the more advanced GFPE concepts. The fundamental code, developed by Ken Gilbert and Xiao Di, is being integrated into SOUNDPLAN, a commercially available sound propagation computer program.

10:15

4aNS7. Wide-angle parabolic equation in moving inhomogeneous media. V. E. Ostashev (Dept. of Phys., Box 30001, New Mexico State Univ., Las Cruces, NM 88003-8001, vostashe@nmsu.edu), Philippe Blanc-Benon, and Daniel Juvé (Ecole Centrale de Lyon, 69131 Ecully, France)

The wide-angle parabolic equation has been used widely to predict a sound field in an ocean with variations in the adiabatic sound speed c. The same equations with c replaced by the effective sound speed $c_{\text{eff}}=c+v_x$ has recently been adopted for calculations of sound propagation in a stratified and/or turbulent atmosphere. Here, v_x is the wind velocity component in the direction from a source to a receiver. However, as is shown in the presentation, such a replacement does not allow one to correctly take into account the effects of the medium velocity $\mathbf{v}$ on sound propagation and scattering in a wide-angle approximation. Furthermore, the effects of fluctuations in the density ρ on sound scattering are also omitted in this approach. In the presentation, a correct wide-angle parabolic equation in media with variations in c, ρ and $\mathbf{v}$ and its Padé (1,1) approximation are derived. The difference in numerical predictions of sound propagation in a stratified and turbulent atmosphere, based on the equation derived and that used previously, is demonstrated. [This material is based upon work supported in part by the U.S. Army Research Office under Contract No. DAAH04-95-1-0593.]

10:30

4aNS8. Sound propagation through a turbulent jet: Experiment and theory. V. E. Ostashev (Dept. of Phys., Box 30001, New Mexico State Univ., Las Cruces, NM 88003-8001, vostashe@nmsu.edu), Philippe Blanc-Benon, and Daniel Juvé (Ecole Centrale de Lyon, 69131 Ecully, France)

The coherence functions of a spherical sound wave after passing through a turbulent jet were measured for different values of sound frequencies and mean speed of the jet [Ph. Blanc-Benon, Ext. Rev. Cethedec-Ondes Signal, No. 79-1984 (1984)]. In the presentation, the experimental data obtained are compared with the theoretical predictions based on the previous theories of sound propagation in turbulent media with medium velocity fluctuations [e.g., E. H. Brown and F. F. Hall, Rev. Geophys. Space Phys. **16** (1), 47–110 (1978)] and with the predictions of the theory recently developed [V. E. Ostashev, Waves Random Media **4**, 403–428 (1994)]. The differences between these theories and their results for the coherence function of a spherical wave are explained. It is also shown that the predictions based on the previous theories do not agree with the experimental data. On the other hand, the predictions of the theory developed by V. Ostashev fit the measured coherence functions fairly well. [This material is based upon work supported in part by the U.S. Army Research Office under Contract No. DAAH04-95-1-0593.]

10:45

4aNS9. Focusing of sonic boom noise penetration into a homogeneous wavy ocean: Complex surfaces and wavelength comparisons. Judith L. Rochat and Victor W. Sparrow (Penn State Univ., P.O. Box 30, State College, PA 16804, rochat@sabine.acs.psu.edu)

The last decade has seen a revived interest in the study of sonic booms; this is due to an upcoming new breed of supersonic passenger aircraft along with a heightened awareness of marine environmental issues. Sonic boom noise penetrating into a *flat*, *homogeneous* ocean is a topic several researchers have already addressed. The primary goal of the authors' research is to study the effects of *realistic* ocean features on noise penetration. These features include a wavy ocean surface and bubbles near the surface. This presentation will focus on a wavy ocean surface, specifically on somewhat complex surface profiles and on sonic boom effective wavelength/ocean wavelength comparisons. Results using finite difference simulations indicate that a somewhat complex wavy ocean surface profile slightly augments the underwater pressure field (as compared to a flat surface); this result was also found when studying simple surface profiles. Some of the trends observed, however, were not as obvious as seen when studying the simple surface profile. This prompted a study on how the sonic boom effective wavelength compared to the ocean wavelength affects the amount of augmentation to the underwater pressure field. [Work supported by the NASA Langley Research Center.]

11:00

4aNS10. Methods of correlating measured sound levels from an aircraft noise monitoring system with aircraft operations. Steve Alverson (Harris, Miller, and Hanson, 945 University Ave, Ste. 201, Sacramento, CA 95825)

The Port of San Diego, owner and operator of San Diego International Airport (SAN), maintains a comprehensive aircraft noise control program that is designed to minimize noise impacts from aircraft operations at San Diego International Airport—Lindbergh Field. One element of this program involves continuous monitoring of noise at 24 remote monitoring terminals in the neighborhoods surrounding SAN and the acquisition of aircraft flight track data. This paper explores methods used by the SAN Noise Information Department staff to correlate, using algorithms in the aircraft noise and operations monitoring system (ANOMS), measured aircraft sound levels with actual aircraft operations. The paper will discuss the use of sound level rise and decay times, frequency filters, timing of noise events between monitors, passive secondary surveillance radar (PASSUR) to acquire aircraft flight tracks, and aircraft situation display (ASD) information to identify aircraft only noise events. The paper will also discuss how measured aircraft noise levels are compared to computed aircraft noise levels to verify the accuracy of the annual community noise equivalent level (CNEL) contours. The CNEL contours are submitted to the State Department of Transportation in response to the State of California aircraft noise regulation reporting requirements.

Session 4aPP

Psychological and Physiological Acoustics: Temporally Dynamic Psychophysics and Physiology

Sid P. Bacon, Chair
Department of Speech and Hearing Science, Arizona State University, Box 871908, Tempe, Arizona 85287-1908

Contributed Papers

8:45

4aPP1. Rhythmic masking release: A paradigm to investigate auditory grouping resulting from the integration of time-varying intensity levels across frequency and across ears. Martine Turgeon and Albert S. Bregman (Dept. of Psycho., McGill Univ., 1205 Dr. Penfield Ave., Montreal, QC H3A 1B1, Canada)

The perceptual organization of sequences of short-duration 200-Hz-wide modulated noise bands was investigated. Sequences with one of two possible rhythms were presented over headphones or in a semi-circular array of 13 speakers. The ''rhythm'' was embedded in an irregular sequence of ''maskers'' with the same frequencies and energy. Adding ''flankers'' in other critical bands, synchronous with the irregular maskers released the rhythm from masking. RMR as measured by the correct identification of the rhythm and its perceived clarity was assumed to be contingent upon the perceptual grouping of the maskers and flankers. Using a 2-AFC procedure with a four-point clarity rating for each alternative, the effect of the following relations between the maskers and flankers on their simultaneous grouping was investigated: stimulus onset asynchrony (SOA), frequency separation (ΔF), commonality of microstructure, and angular separation of their sources ($\Delta\theta$). RMR decreased significantly with SOA, $\Delta\theta$, differences in microstructure and with ΔF for spatially separated maskers and flankers. RMR will be compared to other paradigms which have been claimed to result from a cross-spectral binaural analysis of the temporal envelope, such as dichotic CMR and/or of the fine temporal structure such as BMLD [Hall *et al.*, J. Acoust. Soc. Am. **83**, 1839–1845 (1988)].

9:00

4aPP2. Effect of temporal gaps on informational and energy-based masking. Toktam Sadralodabai, Donna L. Neff, and Traci R. Gleason (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

The effect of a temporal gap on detecting a 1000-Hz tone in the presence of 60-dB SPL simultaneous maskers was examined. Ten-component, random-frequency maskers and broadband-noise maskers were used in a 2-AFC adaptive task. Random-frequency components were drawn from 300 to 3000 Hz, excluding a 160-Hz band around the signal. Temporal gaps of 10, 20, 40, 80, and 160 ms were tested, positioned either at the onset, center, or offset of either the signal or the masker. Without gaps, both signal and masker durations were 200 ms. To maintain equal energy across all conditions, level compensation was applied when gaps were employed. For temporal gaps in either the multicomponent or noise masker, masked thresholds consistently decreased as gap duration increased. Gaps in the masker appeared to provide a temporal window for detection of the signal. However, for gaps in the signal, masked thresholds decreased with the multicomponent masker, but remained constant with broadband noise masker. With multicomponent maskers, the gaps appeared to reduce informational masking by perceptually segregating the signal from the masker. With broadband noise maskers, there was little informational masking and therefore the temporal gaps did not improve performance. [Work supported by NIDCD.]

9:15

4aPP3. Reducing the effects of masker uncertainty with harmonicity and onset/offset synchrony. Eunmi L. Oh and Robert A. Lutfi (Dept. of Psych. and Dept. of Commun. Disord., Univ. of Wisconsin, Madison, WI 53706)

Detection thresholds for a tone in an unfamiliar tonal pattern can be greatly elevated under conditions of masker uncertainty [E. Oh and R. A. Lutfi, J. Acoust. Soc. Am. **101**, 3148 (1997)]. Two experiments were conducted to determine whether harmonicity and onset/offset synchrony of masker tones can reduce the detrimental effect of masker uncertainty. Masker uncertainty was introduced by randomly varying the frequencies and amplitudes of masker tones on each presentation. In experiment 1, maskers were composed of 2–49 tones that were multiples of 200 Hz, not including 1000 Hz. Anywhere from 2–10 dB less masking was observed for an inharmonically related signal at 1047 Hz than for a harmonic signal at 1000 Hz. In experiment 2, the onsets and offsets of masker tones covaried or varied independently of each other. As much as 30 dB less masking was obtained for the 1000-Hz tone signal when masker tones covaried in time. The results support the idea that harmonicity and onset/offset synchrony cause the signal to be perceptually segregated from the masker, thereby resulting in less masking.

9:30

4aPP4. Spectral integration and the detection of tones in modulated and unmodulated noise. Sid. P. Bacon, Nicolas Grimault, and Jungmee Lee (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, spb@asu.edu)

The intent of the present study was to determine whether spectral integration is similar in modulated and unmodulated noise. The signal consisted of one sinusoid of a triplet, or all three sinusoids together. Four sets of triplets were used. The masker consisted of one or three 100-Hz-wide bands of noise centered at the sinusoidal component(s) of the signal. They were unmodulated or sinusoidally amplitude modulated (in phase) at a rate of 8 Hz and a depth of 1.0. Thresholds were measured adaptively for each sinusoid in a triplet, and then for all three together, with the relative level of each determined by the threshold differences when measured individually. In unmodulated noise, the improvement in threshold (from one to three sinusoids) was about 2 dB, consistent with a $\sqrt{3}$ improvement. In modulated noise, however, the improvement was often greater than that (5 dB). This greater improvement could not be explained by shallower psychometric functions in the modulated condition, as the psychometric functions for the detection of the individual sinusoids were similar for both backgrounds. Instead, it is likely due to the comodulated nature of the three masker bands. [Work supported by NIDCD.]

9:45

4aPP5. Auditory temporal microstructure: Evidence of under- and over-shoot at onset and offset of a narrow-band noise masker measured in an off-frequency masked detection task. Craig Formby, Sarah H. Ferguson[a) (Div. of Otolaryngol.—HNS, Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201, cformby@surgery2.ab.umd.edu), and Michael G. Heinz (Harvard-MIT, Cambridge, MA 02139)

Detection thresholds were measured as a function of the temporal position of a 5-ms sinusoidal signal presented within one of two 500-ms observation intervals. Both intervals contained a 400-Hz-wide noise masker, logarithmically centered at 2500 Hz, that was gated on at 150 ms and off at 350 ms within each interval. The signal and masker were both gated with 2-ms linear ramps. The level of the signal was tracked adaptively in blocks of forty 2I 2AFC trials to estimate a 70.7% correct detection threshold. Thresholds were measured at 43 temporal signal positions within the observation interval, with detailed sampling around the onset and offset of the masker to assess temporal edge effects. Experiments to date with four listeners have focused on signals presented at 1900 or 2100 Hz. The temporal microstructure of the detection functions has characteristically revealed a complex pattern of under- and over-shoot (i.e., enhanced and diminished detection, respectively). These effects, which have been observed at both the onset and offset of the masker for virtually all masker spectrum levels between $N_0=0$ and 70 dB, may accentuate the temporal edges of the masker and, in general, enhance detection of acoustic onsets and offsets. [a)]Currently at Dept. of Speech and Hearing Sci., Indiana Univ.

10:00

4aPP6. Effects of fine structure on onset behavior of distortion products in humans. Carrick L. Talmadge (Dept. of Phys., Purdue Univ., West Lafayette, IN 47907), Glenis R. Long, and Arnold Tubis (Purdue Univ., West Lafayette, IN 47907)

There is accumulating evidence that the fine structure of distortion product otoacoustic emissions (DPOAEs) in humans originates from an interference between two cochlear sources, one at the ''overlap'' site of the two primaries, and the other at the best frequency for the DPOAE. It is shown that the DPOAE onset behavior measured in humans, after the turn on of an external tone, is consistent with previous evidence for a ''two source'' origin of DPOAE fine structure [reviewed in Mills, J. Acoust. Soc. Am. **102**, 413–429 (1997)]. A variation of Whitehead's methodology [J. Acoust. Soc. Am. **100**, 1663–1679 (1996)] was used to measure the initial DPOAE buildup. The $f1$ primary was on continuously for a given measurement and the $f2$ primary was pulsed (square wave, 50% duty cycle) at a rate of 5 Hz. The DPOAE had an initial buildup, followed either by further growth of the DPOAE, when the DPOAE frequency was near a fine-structure maximum, or by decay of the DPOAE, when the DPOAE frequency was near a fine-structure minimum. This pattern of secondary growth/decay is shown to be further evidence favoring the two-source model of DPOAE fine structure.

10:15–10:30 Break

10:30

4aPP7. The ''center-of-gravity'' effect in dynamic tones. Jayanth N. Anantharaman, Ashok K. Krishnamurthy (Dept. of Elec. Eng., Ohio State Univ., Columbus, OH 43210, jayanth+@osu.edu), and Lawrence L. Feth (Ohio State Univ., Columbus, OH 43210)

Earlier work [e.g., Xu *et al.*, J. Acoust. Soc. Am. **101**, 3149(A) (1997); Chistovich, *ibid.* **77**, 798–805 (1985)] has demonstrated the spectral center-of-gravity (CoG) effect. The studies have largely concentrated on steady-state signals. Recently, Lublinskaya (1996) has investigated the CoG effect in dynamic synthetic speech. In the study presented here, the CoG effect in dynamic two-tone complexes was investigated. The frequencies of the two tones were held constant while their amplitudes were varied linearly: the amplitude of the lower frequency tone was decreased while that of the higher frequency tone was increased. This causes a variation in the spectral CoG, although the locations of the frequency content remain constant with time. Subjects matched the slope of a linear glide (target) having the same center frequency as the two-tone complex (standard) using a double-staircase procedure. The experiment was carried out for various center frequencies (500–2000 Hz) and frequency separations (0.5–8 ERBs). The center frequency was also roved between trials. Matching experiments were also carried out with another linear glide as the standard. The results of these experiments will be presented and the center-of-gravity effect in these dynamic two-tone complexes will be discussed. [Work supported by AFOSR and OSU-GSARA.]

10:45

4aPP8. Discriminability of correlated bursts of reproducible noise. Susan M. Fallon (Dept. of Psych., Univ. of Alaska, 533 East Pioneer Ave., Homer, AK 99603)

Discriminability of pairs of 100-ms bursts of noise was measured using a same–different psychophysical method. In the first condition, the correlation between bursts in a pair was fixed at 1.0 during each same trial. On each different trial, bursts were identical except for τ ms of independent noise located at either the beginning, the middle, or the end of the pairs. The correlation between bursts presented on each different trial was increased from 0.0 by decreasing the duration of τ. The second condition was the converse of the first; the correlation between bursts presented on each different trial was fixed at 0.0 and the correlation between bursts presented on each same trial was decreased from 1.0. As the τ ms of either independent (condition 1) or the perfectly correlated (condition 2) noise is moved from the beginning to the end of bursts, discriminability increases in the first condition and decreases in the second condition. These results suggest that a ''difference'' is a more salient cue in the discrimination task than a ''sameness'' and are discussed in terms of a correlational model. [Work supported by NIH and AFOSR.]

11:00

4aPP9. Sequential interactions in the discrimination of frequency increments co-occurring with irrelevant increments in frequency, duration, or level. Blas Espinoza-Varas and Hyunsook Jang (Commun. Sci. and Disord., Univ. of Oklahoma Health Sci. Ctr., Oklahoma City, OK 73190, blas-espinoza-varas@uokhsc.edu)

Discrimination of target frequency increments co-occurring with irrelevant increments in frequency, duration, or level was studied. *Pairs* (t_1,t_2) of 1500-Hz, 80-ms pure tones separated by 20–120 ms of silence (ITI) were presented in a three-interval, two-alternative, forced-choice (3I/2AFC) task. On each trial, a ''standard'' *pair* of tones (interval 1) was followed by two ''comparison'' *pairs* (intervals 2 and 3), with 500 ms of silence between observation intervals. One (randomly chosen) comparison *pair* contained a target increment in the frequency of either t_1 *or* t_2; the size of the target increment was controlled by adaptive rules. In addition to the target increment, both comparison stimuli contained identical irrelevant increments in either the duration, frequency, or level of either t_1 or t_2; there was no correlation between the observation intervals containing target and irrelevant increments. Listeners had to determine which observation interval contained the target increment. Discrimination of t_1 frequency was degraded by irrelevant increments in t_2 frequency or t_2 level, particularly at the shortest ITI of 20 ms; prolonged discrimination training lessened the effects. Discrimination of t_2 frequency was not affected by irrelevant increments. [Work funded by OCAST Project HR4-064.]

11:15

4aPP10. A recovery model for coincidence neurons. Timothy A. Wilson, Rachod Thongprasirt (Dept. of Elec. Eng., Univ. of Memphis, Memphis, TN 38152), and Aziz A. Khanifar (Vanderbilt Univ., Nashville, TN 37240)

A new model for firings of coincidence neurons in the MSO is presented. The model is similar to one by Colburn *et al.* [Hearing Res. **49**, 335–346 (1990)]: Both are point processes, both have inputs which are impulse-train representations of the firing events of primarylike neurons, both operate on a low-pass filtered sum of those inputs, and both output a sequence of impulses at the event times of the coincidence neuron's firings. In this model, their threshold-crossing neuron model is replaced by a recovery model similar to that modeling the input firings. The coincidence neuron's drive function is a nonlinearly distorted version of the filtered sum of inputs; for simplicity, its recovery function is the same as that of the input neurons. The model's firing rates, vector strengths, period histograms, and interval histograms are compared to data from dog and cat and with similar responses of the Colburn *et al.* model.

11:30

4aPP11. Neural responses to the onset of voicing are unrelated to other measures of temporal sensitivity. Donal G. Sinex and Guang-Di Chen (Arizona State Univ., Dept. of Speech and Hearing Sci., Tempe, AZ 85287-1908)

Voice onset time (VOT) is a temporal cue that can distinguish consonants such as /d/ from /t/. We have previously shown that neurons' responses to the onset of voicing are strongly dependent on their spectral sensitivity. This study examined the relation between neurons' temporal sensitivity, determined from responses to amplitude-modulated tones, and the same neurons' responses to the onset of voicing. Responses to VOT syllables and two types of modulated tones were obtained from low-frequency neurons in the inferior colliculus (IC) of the chinchilla. Both VOT and the modulation period varied from 10–70 ms in 10-ms steps. Neurons that respond selectively to modulated tones might be expected to respond strongly to the syllables whose VOTs match the preferred modulation periods. However, for most neurons the correlation between discharge rate to modulated tones and voicing onset was low. Some neurons exhibited moderate selectivity for certain modulation periods, but this selectivity was usually unrelated or even inversely related to the same neurons' selectivity for VOT syllables. Overall, responses to modulated tones did not account for the responses of IC neurons to the complex temporal cues associated with the onset of voicing. [Work supported by NIDCD.]

11:45

4aPP12. Regular interval stimuli: Are higher-order intervals necessary? William A. Yost and William P. Shofner (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

Many stimuli that contain regular temporal intervals are perceived as having a pitch equal to the reciprocal of the interval. Autocorrelation is one method of determining the interval in these regular interval stimuli (RIS). Autocorrelation analysis has been used as a ''model'' for calculating the pitch and pitch strength of RIS. Autocorrelation is an ''all-interval'' analysis as opposed to the first-order interval analysis that is often used to describe the temporal properties of neural spike trains. Both actual neural responses (from the AVCN of the chinchilla) and those of a simple neural model were measured to determine the order of intervals contained within RIS. In a psychoacoustic experiment, RIS were generated with no first-order intervals or with only first-order intervals, and listeners discriminated between these stimuli and stimuli with no regular intervals. The results of the experiments and the calculations indicate that the pitch and pitch strength of RIS probably depend on only lower-order intervals (no higher than fourth-order intervals). These results will be discussed in terms of theories of pitch and timbre processing. [Work supported by NIDCD, DC-000293.]

THURSDAY MORNING, 4 DECEMBER 1997

SAN DIEGO/GOLDEN WEST ROOMS, 8:30 A.M. TO 12:00 NOON

Session 4aSC

Speech Communication: Vocal Tract Modeling, Speech Technology and Other Topics (Poster Session)

Anders Lofqvist, Chair
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

All posters will be on display from 8:30 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:15 a.m. and contributors of even-numbered papers will be at their posters from 10:15 a.m. to 12:00 noon. To allow for extended viewing time, posters will remain on display until 5:00 p.m.

Contributed Papers

4aSC1. A finite-element template for tongue modeling. Reiner Wilhelms-Tricarico (Res. Lab. of Electron., MIT, Cambridge, MA 02139) and Chao-Min Wu (Univ. of Wisconsin, Madison, WI 53705)

This report describes the design, data structure, and applicability of a refined finite-element template for the human tongue and connected oral structures. The finite-element model relies mainly on data sets from the Visible Human Project of the National Library of Medicine. The model is composed of macroblocks, which represent geometric subsections of the tongue. These blocks represent, in some cases, either individually or in combination, functional subsections, such as individual muscles. For each macroblock (or geometric region) a finite-element mesh can be generated such that the whole of the tongue can be modeled by a mesh of finite-elements, since the subdivision is compatible across block boundaries. The finite element template contains information about muscle tissue distribution of the tongue and velopharynx. Approximate muscle fiber directions are represented as direction fields. The model can be adapted to individual morphology if a set of morphological landmarks can be specified from measurements such as those from MRI data of an individual.

4aSC2. A study of speaker adaptation of continuous density hidden Markov models. Prabhu Raghavan and Chiwei Che (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

Considerable research in the area of speech recognition, in the past two decades, using HMMs has led to successful reduction in word error rates (WER). Owing to their statistical nature, HMMs require large amounts of data to get good estimates of the model parameters. It has been shown that a speaker dependent (SD) model has about 2–3 times less WERs compared to a speaker independent (SI) model. But the heavy data requirement from one speaker is quite unreasonable. Speaker adaptation techniques mitigate this problem by the use of SI data to obtain robust estimates and then use small amounts of SD data to adjust these parameters so that they are optimal to that speaker (in some optimality criterionlike ML). In this paper, the different techniques of adaptation are studied. The salient methods, namely MLLR and MAP, of each type are discussed in detail. A method of constrained unsupervised adaptation, which discards parts of the original hypothesis that are deemed to be incorrect, is proposed. This decision is made on past experience obtained on training data. This decision making module uses a neural network and a CART tree.

4aSC3. Mapping of muscle anatomy on three-dimensional magnetic resonance images of the human tongue based on morphological landmark selection. Chao-Min Wu (Waisman Ctr., Univ. of Wisconsin, 1500 Highland Ave., Madison, WI 53705) and Reiner Wilhelms-Tricarico (MIT, Cambridge, MA 02139)

The purpose of this study is to incorporate accurate muscle morphology in a computational 3-D tongue model based on anatomical literature and MRI data. Previous work [C.-M. Wu and R. Wilhelms-Tricarico, J. Acoust. Soc. Am. **100**, 2659(A) (1996)] defined several landmarks within the 3-D space of the tongue tissue and developed a computational method to map Miyawaki's drawing of muscle fiber directions [K. Miyawaki, Ann. Bul. RILP, Univ. Tokyo **8**, 23–50 (1974)] onto the cross-sectional tongue images from the visible human data. In this study, the same technique is applied to identify a set of corresponding landmarks between Miyawaki's data and MRI scans of a subject's tongue forming a consistent articulatory posture. The geometrical mapping of these two data sets is achieved by the thin-plate spline mapping method. The result shows the sufficiency and robustness of the chosen landmark points to represent a general 3-D structure of the human tongue. This study is expected to provide a computational basis for quantitative comparison of MRI data across subjects and vowel types.

4aSC4. Articulatory based low bit-rate speech coding. Samir Chennoukh, Daniel Sinder, Gael Richard, and James Flanagan (Ctr. for Comput. Aids for Industrial Productivity, Rutgers Univ., P.O. Box 1390, Piscataway, NJ 08855-1390)

A voice mimic system has been designed to achieve an articulatory description of the speech signal based on an analytic description of the vocal tract (VT) parameters as a function of the first two formant frequencies. Articulatory based low bit-rate speech coding using the analytic mapping technique has been demonstrated. The natural speech input is analyzed to obtain the first two resonances of the VT, and the VT shape is estimated from the analytic acoustic-to-articulatory mapping. The shape is modeled using three parameters. Then, the articulatory parameters in addition to the frequency of the vocal cord vibration are transmitted. The receiver restores the VT shape from the articulatory parameters, computes the formant frequencies corresponding to the VT shape and synthesizes the output speech using a formant synthesizer. The bit-rate of the coding is determined from the parameter sampling rate and their digital representation. An intelligible output speech signal has been obtained for bit-rates as low as 624 bits/s using parameter sampling at 48 times/s, 3 bits for each of the VT parameters and 4 bits for the vocal cord vibration frequency. Coding at this bit-rate has been demonstrated on voiced sentences. [Research supported by ARPA DAST 63-93-C-0064.]

4aSC5. An aeroacoustic fricative source model. D. Sinder, M. Krane, and J. Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

A model for fricative sound generation based on aeroacoustic theory is implemented in an articulatory synthesizer. The aeroacoustic model is based upon Howe's reformulation of the acoustic analogy, which allows the source to be specified in terms of the jet formed at a constriction, the shape of the vocal tract walls, and the acoustic loading due to the lip termination. As such, the model uses existing acoustic models for the vocal tract combined with an approximate model for the behavior of the jet based on turbulence theory and experiment. This allows the fricative source strength, impedance, and temporal characteristics to be specified using the lung pressure, the constriction diameter, and the shape of the vocal tract downstream of the constriction. Evaluation of the source model is presented. Computed estimates of fricative source characteristics are presented and compared to source models used by other investigators. A comparison of synthetic fricative sounds generated using these models is also presented. [Research supported by NSF/ARPA IRI-9314946 and ARPA DAST 63-93-C-0064.]

4aSC6. Labeling a speech database with landmarks and features. Jeung-Yoon Choi, Erika Chuang, David Gow, Katherine Kwong, Stefanie Shattuck-Hufnagel, Kenneth Stevens, and Yong Zhang (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

A database of utterances has been labeled for acoustically salient landmarks and for phonetic features. The landmarks locate syllabic nuclei, glide midpoints, and acoustic discontinuities at the narrowing or releasing of consonantal constrictions. The consonantal landmarks are labeled with values of the articulator-free features (e.g., [continuant] and [sonorant]), and values of articulator-bound features are recorded at each landmark based on acoustic evidence near the landmark. These features specify, for example, the major articulator and its placement, as well as the activity of secondary articulators such as the glottis and the soft palate. Conventions have been developed for labeling reduced vowels, syllabic consonants, glottal stops, consonant sequences in which closures or releases are not acoustically marked, etc. The labeled database is being used for two purposes: (1) to provide a documentation of instances in which the acoustic/articulatory representation of an utterance deviates from the underlying feature-based lexical representation and (2) to train and to test signal-processing algorithms that are being developed to locate the landmarks and to identify the features attached to these landmarks. An advantage of this labeling method is that it gives a more fine-grained phonologically-relevant representation of an utterance than a conventional labeling in terms of phonetic symbols. [Supported in part by NIH Grant No. DC02978.]

4aSC7. New parameters and mapping relations for the HLsyn speech synthesizer. Helen M. Hanson, Kenneth N. Stevens, and Robert E. Beaudoin (Sensimetrics Corp., 126 Landsdowne St., Cambridge, MA 02139)

HLsyn is a quasiarticulatory speech synthesizer in which a small set of parameters control a Klatt synthesizer [Stevens and Bickley, J. Phon. **19** (1991)]. Originally, ten parameters were used. In this paper three new physiologically based parameters, together with some additional modifications, are described. The first is a time-varying subglottal pressure parameter, which provides the user with additional control of the voice-source amplitude. It can also be used to turn on voicing, and it has an influence on fundamental frequency. The second parameter is a time-varying percentage change in the compliances of the vocal-tract walls and vocal folds. This parameter can, for example, be employed when synthesizing voiced obstruents: increasing it results in facilitation of glottal vibration and lowering of the fundamental frequency. The third parameter is the time-varying cross-sectional area of a posterior glottal chink. Because it is independent of the area at the membranous folds, significant noise-source amplitudes can be achieved during voiced speech. Thus one can synthesize aspirated voiced stops, as well as breathy speech, in a natural manner. Finally, a default female voice and an intrinsic pitch feature have been added to the system. Examples of copy synthesis where the new parameters play a role are presented. [Work supported by NIH Grant MH52358.]

4aSC8. An HMM-based measure for evaluating the quality of foreign-accented speech. Robert Smith, Kiyoaki Aikawa, and Satoshi Takahashi (NTT Human Interface Labs, 1-1 Hikarino-oka, Yokosuka-shi, Kanagawa 239, Japan)

A new criterion for measuring the quality of foreign-accented speech is proposed. The proposed criterion uses both the absolute and relative quality scores produced by an HMM speech recognizer. An absolute scoring function is determined by the difference of likelihood distributions between the correct phone HMM and an incorrect phone HMM for a speech segment. A relative scoring function is established to discriminate a delta-self value distribution from a delta-sub value distribution. A delta-self value is obtained by the likelihood score difference between the correct model and an incorrect model for a speech segment. A delta-sub value is obtained by the likelihood score difference between two incorrect phone models against the speech segment. The new idea is motivated by the following evidences. It is sometimes found that the absolute score is low, even if the speech segment sounds correct. In that case, however, the relative score is still high for the speech segment. This may happen because the HMM parameters were not chosen for evaluating voice quality. The best speech quality meets a condition that both the absolute and the relative scores are high. A practical integrated measure is the linear combination of the absolute and the relative scoring functions.

4aSC9. Indexical cuing in laughter. Jo-Anne Bachorowski, Moria J. Smoski (Dept. of Psych., 301 Wilson Hall, Vanderbilt Univ., Nashville, TN 37240, bachja@ctrvax.vanderbilt.edu), and Michael J. Owren (Cornell Univ., Ithaca, NY 14853)

Laughter is a frequently produced yet poorly understood vocal signal. Despite the lack of basic knowledge regarding this species-typical vocalization, this signal may be advantageous to study as it provides an opportunity to examine indexical cueing in the absence of linguistic content. The source-filter model of vowel production guided acoustic analysis of laughter recorded from experimentally naive subjects as they watched emotion-inducing film clips in a laboratory setting. Subjects were 48 males and 48 females tested either alone or with a same- or opposite-sex friend or stranger. Laugh acoustics were found to vary by gender, talker, and social context. Males produced predominantly unvoiced, ''grunt'' laughs, whereas female laughter was more variable and included more voiced, long duration ''song'' laughs. Discriminant function analyses revealed that source and filter-related cues were differentially associated with social context and talker identity. Variability in source-related cues, particularly $F0$, more successfully discriminated social context, whereas cues related to the vocal tract transfer function better discriminated among individuals. Females, but not males, had significantly higher formant frequencies when tested with opposite-sex strangers. The overall pattern of results is consistent with recent examinations of indexical cuing in human speech and nonhuman primate vocalizations. [Work supported by NIMH.]

4aSC10. Parsing the contribution of lip and tongue position to fricative noise frequency. D. H. Whalen and Bryan Gick (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The English alveolar and palatal fricatives have, as part of their realization, high-frequency noises that are shaped both by the rounding and protrusion of the lips and the position of the tongue on the hard palate. The acoustic effects are of two main types, influence on the main spectral resonance of the noise, and the frequency and intensity of a second formant resonance within the noise. However, the exact contribution of each of these articulatory factors to the two acoustic effects is not known. The present study used electromagnetometer measures of tongue and lip location and video-based measures of lip aperture, area, and protrusion. Initial data indicate that area of the lip opening correlates significantly with $F2$ resonance, while tongue position correlates with fricative pole. Surprisingly, upper lip protrusion correlates with both the $F2$ resonance and the fricative poles. These preliminary results suggest that both lip and tongue configuration contribute to the shape of the fricative poles, while the $F2$ resonance is primarily dependent on the lip configuration. Implications for perceptual parsing will be discussed. [Work supported by NIH grant DC-02717.]

4aSC11. Predicting pharynx shape from the tongue during vowel production. Min Kang (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Dept. of Linguist., Yale Univ., New Haven, CT 06520), D. H. Whalen and Harriet Magen (Haskins Labs., New Haven, CT 06511),

The shape of the pharynx greatly affects the formants of vowels, yet the measurement of pharyngeal shape is relatively difficult. Using midsagittal MRI images of a male speaker producing eleven English vowels, the predictability of pharyngeal shape during vowel production was assessed. There were two categories of predictors: measured and categorical. Measured predictors were four points on the tongue representing the points likely to be locations for receiver coils in an electromagnetometer system. Categorical predictors were vowel frontness and height, and the typical position of the epiglottis relative to the base of the tongue (as seen in MRI images). Midsagittal distance was measured at 3-mm intervals along the vocal tract length. A stepwise regression using measured and categorical variables gave good predictions for all distances along the vocal tract (r greater than 0.80), except near the uvula. (Using just the categorical predictors resulted in surprisingly good estimations.) Position of the uvula varied across repetitions of the vowels, and so exclusion of the uvula increased predictability in the uvular region. For static vowel patterns in English, then, pharynx shape is fairly predictable from tongue shape. [Supported by NIH grant DC-02717.]

4aSC12. Name retrieval from pronunciation and spelling over the telephone, using HMM modeling and robust directory access. G. Gravier, F. Yvon, G. Etorre, and G. Chollet (ENST, 46 rue Barrault, 75634 Paris Cedex 13, France)

A system to retrieve names in a directory from their pronunciation and their spelling is presented, for telephone quality speech. Most of the previous approaches to letter recognition for spelled names are knowledge-based because of the high phonetic confusability between letters [C. S. Myers and L. R. Rabiner, J. Acoust. Soc. Am. **71**(3), 716–727 (1982)] [Cole *et al.*, Eurospeech'91, 479–482]. To avoid the knowledge-based approach, HMM modeling of speech units (SU) is used with word models for letters and phone models for name pronunciation. The directory is built automatically from a list of names using a grapheme-to-phoneme converter for the name and rules for the spelling, each entry in the directory consisting of several variants for both the pronounced and spelled names. From the acoustic recognition, the corresponding entry in the directory is found using DTW or a tree-search algorithm. Both methods allow insertion and deletion, and the cost for a substitution can either be fixed or determined from the train corpus confusion matrix. On a test database of 5000 names uttered by 3000 speakers ($\sim$ 50 000 SU), 70% of the units are correctly recognized. Preliminary experiments based on spelled names with a DTW lexical search, showed that substitution costs defined using the confusion matrix significantly improved the results.

4aSC13. Stochastic linear regression analysis of speech production data. H. Betty Kollia (Dept. of Logotherapy, TEI Patras, Patra, Greece) and Jay Jorgenson (Dept. of Math., Oklahoma State Univ., Stillwater, OK 74074)

Dynamic models in speech production have been traditionally based on linear regression analysis of articulatory data. The purpose of this work is to introduce a stochastic calculus methodology to the analysis of such data in speech production models. The issue of variability encountered in the analysis and description of articulatory movements during speech may be addressed with the implementation of a stochastic term, thus placing the variability in a controlled probabilistic framework. Upper lip, lower lip, jaw, and velum movements tracked optoelectronically at Haskins Laboratories [Kollia *et al.* J. Acoust. Soc. Am. **98**, 1313–1327 (1995)] during production of a test utterance are described in this manner. The authors believe that this algorithm has wide applicability since it yields confidence interval rather than point value estimates for kinematic data description. [Work supported by NSF.]

4aSC14. Improvement in quality of speech received by a mistuned receiver at suppressed carrier SSB. Shin'ichiroh Kaneko, Jouji Suzuki, and Tetsuya Sjimamura (Dept. of Information and Comput. Sci., Saitama Univ., Shimo-okubo 255 Urawa, 338 Japan)

Single sideband modulation (SSB) is the most efficient speech transmission system from the viewpoint of power consumption and bandwidth occupation. It is widely used in amateur radio, mobile radio, and others. At suppressed carrier SSB(SCSSB), the receiver should tune to the carrier frequency accurately. Manual tuning, however, is very difficult due to the lack of the carrier signal. It fails in fine tuning at the reception and the quality of received speech degrades seriously owing to the shift of speech spectrum. This paper proposes a method to improve the quality of speech degraded at SCSSB. Received speech is transformed into the log spectrum and its autocorrelation function (ACLOS) is calculated. If tuning is carried out correctly, the log spectrum of voiced sound has a harmonic structure. Mistuning gives a spectrum shift on that frequency axis. On the other hand, ACLOS is always a periodic signal that corresponds to the harmonic structure. By calculating a cross-correlation function between ACLOS and the log spectrum, mistuning frequency can be easily estimated. This method is able to estimate the mistuning frequency within the accuracy of 10 Hz, even in noisy channels. After obtaining the mistuning frequency, the shifting spectrum of received speech is carried out and more natural speech is obtained.

4aSC15. The effects of aging and rate on tongue shape variability and on preservation of position in glides. Maryam Jaberi, Maureen Stone (Dept. of Otolaryngol., Univ. of Maryland Med. School, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201, mstone@umabnet.ab.umd.edu), Suzanne Gray (Johns Hopkins Univ., Baltimore, MD 21218), and Elsbieta Slawinski (Univ. of Calgary, Calgary, Canada)

Slowing of neural processes due to aging may cause effects in motor speech production. In a study of /w/, Slawinski [E. B. Slawinski, J. Acoust. Soc. Am. **95**, 2221–2230 (1994)] found a slower prevoicing stage, and a smaller off-glide frequency range for older speakers than younger. The latter finding suggests that older speakers centralize their articulatory gestures. Articulatory gestures may also centralize at faster speech rates. The present study compared midsagittal tongue gestures for /w/, for older versus younger subjects, matched for rate of speech. Ten female subjects (five aged 21–29, five aged 64–71) repeated: /əwVp/ and /əgwVp/ using the vowels /i,a,u/, five times each, while ultrasound video recordings were made. The hypothesis was that coarticulationwould cause different /w/ positions, (higher in /g/ context). If older subjects centralize tongue position for /w/, they might produce an even greater difference, if the high /g/ tongue position facilitated a more extreme /w/. Target frames were digitized and tongue surface edges were extracted [M. Unser and M. Stone, J. Acoust. Soc. Am. **91**, 3001−3007 (1992)]. The distances between /w/ and /(g)w/ were determined using $L2$ norms. For two older and two younger speakers, rate effects were observed and age effects were not.

4aSC16. An automatic speech recognition system using time-delays self-organizing maps with physiological parametric extraction. Jose M. Ferrandez, Daniel del Valle, Victoria Rodellar, and Pedro Gomez (Laboratorio de Comuncacion Oral R. W. Newcomb, Facultad de Informatica, Univ. Politecnica Madrid, C. Montegancedo, Boadilla del Monte, 28660 Madrid, Spain, jmanuel@naranjo.datsi.fi.upm.es)

Physiological parametric extraction uses auditory models as a front end for speech recognition. These last methods assume that if speech signals are coded in the same way that the auditory system does, speech could be later identified showing the main properties that biological systems do: robustness and accuracy. The proposed system consists of a cochlear model implemented by gammatone filterbanks as proposed by Patterson [J. Acoust. Soc. Am. **96**, 1409–1418 (1994)]. This stage will feed a nonlinear mechanical-to-neural transduction module based on the Meddis hair-cell model [J. Acoust. Soc. Am. **79**, 702–711 (1986)], which will compute auditory-nerve firings. Finally, a temporal integration/component–extraction module will integrate neural patterns for identifying the relevant components embedded in the speech signals [characteristic frequency (CF), frequency modulation (FM) and noise burst (NB)], which are shared by human speech and animal sounds for communication. The model adopts a spatiotemporal strategy, which uses temporal information in low CF fibers (phase-locking mechanism) and spatial information for the higher ones. The recognizing module consists in a time-delay self-organizing map, which will capture not only the spectral variability contained in the signal, but also the temporal one, providing better generalization properties. [Work supported by NATO CRG-960053.]

4aSC17. Crime automatic speaker verification and identification system. Igor I. Gorban (Inst. of Math. Machines and Systems, Acad. Glushkov, 42 Ave., Kiev, Ukraine 252187, gorban@olinet.isf.kiev.ua)

A new automatic speaker verification process was presented at the Eighth Annual International Conference on Signal Processing Applications and Technology (ICSPAT '97), San Francisco, CA (I. I. Gorban, Robust Speaker Verification Algorithms). On the basis of these algorithms the crime automatic speaker verification and identification (CASVI) system has been created. The main feature of the system is that it is possible to verify and identify a speaker when the speech records are very corrupted and/or contain extraneous noise. Testing conducted under different conditions has shown that the CASVI system works well when the level of frequency distortions is lower than 35 dB and the level of the signal-to-noise ratio is lower than 12 dB. These results have been obtained for a new modification of the algorithms in which the adaptive threshold verification process is included and the identification algorithms are added. A description of the CASVI system, it's algorithms, the methodology, and the results of the system testing will also be presented at the meeting.

4aSC18. A dynamic neural network model of speech production in the developing child. Daniel E. Callan, Ray D. Kent, and Houri K. Vorperian (Dept. of Commun. Disord., Univ. of Wisconsin, Madison, WI 53706)

Many theories of speech processing propose the existence of motor control systems that utilize invariant vocal tract configurations to specify particular goals of speech production. One problem that challenges these theories is the fact that the associated structures involved with speech production go through a considerable amount of change during development. The same speech goals continue to be achieved during the course of development despite the changes that the vocal tract configuration undergoes. In this paper, speech production in the developing child is modeled by a dynamic neural network that incorporates value-dependent learning based on self-produced auditory stimulation across self-organizing sensorimotor maps. The conversion from articulatory configuration to acoustic signal is worked out by a modified version of the Maeda articulatory model that utilizes developmental parameters. The performance of the neural network was assessed at different points during development by determining the articulation-acoustic output needed to produce various segments of speech in relation to the corresponding trajectory of activation across the neural network. The results are discussed in relation to a dynamic theory of speech processing that can account for developmental change as well as some speech disorders. [Work supported by NIDCD.]

4aSC19. An acoustic model of gesture overlap: Further studies. Gary Weismer and Jeff Berry (Dept. Commun. Disord. and Waisman Ctr., 1500 Highland Ave., Univ. of Wisconsin, Madison, WI 53705)

Previous work from our laboratory has suggested that certain acoustic measures may serve as indices of the extent of overlap for successive articulatory gestures. These experiments have employed speaking-rate variation as a means to vary the extent of articulatory gesture overlap, in a number of different syllable types. Explicit models have been used to predict how rate-induced variation in gesture overlap would be realized in the acoustic output of the vocal tract. Specifically, the $F2$ onset and offset values at CV and VC interfaces, respectively, where C and V are constant, have been shown to be sensitive to the rate variable, sometimes approaching the magnitude of variation associated with CV syllables where C is constant and V is varied (i.e., the classic vowel-induced coarticulation paradigm). However, the results to date appear to be dependent on specific syllables, but the reasons for this specificity are unknown. In the present study, some systematic analyses of vowel effects on the acoustic model of gesture overlap are presented. Results will be discussed in terms of the goodness of the model and theories of coarticulation.

4aSC20. Evaluation of a reiterant force impulse task: A constructive replication. Kate Bunton and Gary Weismer (Dept. of Commun. Disord., Univ. of Wisconsin, 1500 Highland Ave., Madison, WI 53705)

In the current study, characteristics of labial force–impulse tasks were investigated under three speaking conditions using a minipressure transducer. In the task, neurologically normal adults were first asked to read aloud two sentences of different lengths (ten repetitions each) containing bilabial voiced and unvoiced consonants, and were then required to repeat the series of utterances using mimed speech. Mimed speech was produced with subjects holding their breath; subsequently, the experimenter cued subjects to use a normal airflow and the mimed utterances were repeated. Subjects were then trained to produced sequences of labial force impulses modeled on the real-speech utterances by substituting /ba/ or /pa/ for each syllable. The reiterant condition was also recorded as mimed. The goal of the analyses were to (a) compare the timing of the real-speech sequences to the timing of mimed and reiterant versions, and (b) compare the force magnitudes associated with labial stress produced during real speech to mimed and reiterant repetitions. Results will be discussed in terms of the relationship between orofacial, nonspeech motor performance and speech production performance. [Work supported by NIH Award No. T32 DC00042 and R01 DC00319.]

4aSC21. Magnetometer observation of articulation in sitting and supine conditions. Mark K. Tiede, Shinobu Masaki, Masahiko Wakumoto, and Eric Vatikiotis-Bateson (ATR Human Information Processing Res. Labs., 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan, tiede@hip.atr.co.jp)

Magnetic resonance imaging (MRI) has recently emerged as a useful tool for the investigation of the three-dimensional vocal tract shape in sustainable articulatory targets such as vowels and fricatives. However, existing scanning devices require that subjects phonate in a supine position. Previous work [Whalen, J. Acoust. Soc. Am. Suppl. 1 **88** (1990)] compared MRI to cineradiographic data collected in a sitting posture, and reported limited effects of gravity on velar height differences in vowel production. This study used electromagnetometry [Perkell *et al.*, J. Acoust. Soc. Am. **92**, 3078–3096 (1992)] to investigate whether detectable systematic differences in midsagittal tongue posture exist between the two phonation conditions. Subjects fitted with nine transducers (four on the tongue, mandible, upper and lower lips, and two for head correction reference) produced both running speech and sustained vowels mimicking MRI protocol, repeated in sitting and supine postures. Analysis of separately recorded acoustics showed systematic effects of production posture on lower formant bandwidth. Results of comparing tongue height differences between the two conditions will be presented.

4aSC22. MRI observation of dynamic articulatory movements using a synchronized sampling method. Shinobu Masaki, Mark K. Tiede, Kiyoshi Honda (ATR Human Information Processing Res. Labs., 2-2 Hikaridai Seika-cho Soraku-gun, Kyoto 619-02, Japan, masaki@hip.atr.co.jp), Yasuhiro Shimada, Ichiro Fujimoto, Yuji Nakamura (Takanohara Chuo Hospital) and Noboru Ninomiya (Shimadzu Corp.)

The use of magnetic resonance imaging (MRI) in speech research has, typically, been limited to static, sustainable vocal tract configurations due to the long acquisition times required [e.g., Baer *et al.*, J. Acoust. Soc. Am. **90**, 799–828 (1991)]. This work describes the application of a synchronized MRI sampling technique developed for cardiac motion imaging to the visualization of dynamic articulatory movements. Subjects produced 128 repetitions of a target utterance (/tata/ and /kaka/) coincident with audible tone bursts synchronized with gating pulses controlling MRI scanning. During each repetition, 35 scans at 25-ms intervals were acquired, each scan providing one row of frequency-domain data for each of the 35 output frames. After repetitive scans were completed, the full sequence was processed to give a 40-frames-per-second midsagittally oriented movie. Analysis of the resulting movies revealed that tongue shapes for the interconsonantal /a/ varied between the two utterance types, indicating a place-of-articulation effect of consonants on coproduced vowel tongue shape. This suggests that the synchronized sampling method provides a practical technique for investigation of the dynamic characteristics of articulatory movements during short utterances.

4aSC23. Three-dimensional visualization of electropalatography. Masahiko Wakumoto and Shinobu Masaki (ATR Human Information Processing Res. Lab., 2-2 Hikari-dai, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan, mwakumo@hip.atr.co.jp)

Electropalatography (EPG) is a widely used noninvasive technique for monitoring dynamic tongue-palate contact patterns in speech research and speech clinics. However, the conventional two-dimensional (2D) visualization system is incapable of informing individual characteristics of three-dimensional (3-D) palatal shape or the relationship between electrode positions and anatomical landmarks (e.g., teeth). This paper reports a method for visualizing tongue-palate contact patterns on the 3-D palatal shape display. A convex palatal shape model was formed by silicone impression material from an individual dental cast, and electrode positions on the individual-specific artificial palate (RION SP-01, 63 electrodes) were marked on the surface of the convex model. The 3-D geometry and the surface texture information of the convex model were extracted by a laser beam scanner (Cyberware 4020). Then, palatographic data were collected through the EPG system (RION DP-20) and the tongue-palate contact patterns were mapped on the 3-D graphical display of the palate. Preliminary observation of Japanese fricatives revealed vertical differences in constriction between [s] and [ʃ]. Its application to speech production research and speech training will be discussed.

4aSC24. Geometrical display of speech spectra as an aid to lipreading. Eric J. Hunter and William J. Strong (Dept. of Phys. and Astron., Brigham Young Univ., Provo, UT 84602)

A geometrical display of speech spectra intended as an adjunct to lipreading was developed. Spectra were calculated at 5-ms intervals from speech sound pairs ambiguous to lipreaders. The spectra were displayed as sequences of irregular decagons. Human subjects were asked to discriminate between pairs of spectral decagon sequences derived from pairs of ambiguous speech sounds. Subjects were able to discriminate between most of the visual spectral patterns derived from ambiguous sounds. However, spectral patterns associated with the voiced/unvoiced contrast in some stop pairs were not discriminated consistently.

4aSC25. Interarticulator programming in VCV sequences. Anders Lofqvist and Vincent L. Gracco (Haskins Labs., 270 Crown St., New Haven, CT 06511)

This study examined the coordination of tongue body and lip movements in VCV sequences, where the consonant is a bilabial stop /p, b/, and the vowel one of /i, a, u/; only asymmetrical vowel contexts were analyzed. Lip and tongue timing and vowel-to-vowel tongue body kinematics were examined. The movements were recorded using a magnetometer system. Four subjects participated and produced ten tokens of each sequence. The phasing between the onset of the lip closing movement and the onset of the tongue body movement from the first to the second vowel varied across subjects, with no obvious pattern according to consonant voicing or vowel context. Overall, the onset of the tongue movement to the second vowel started earlier relative to the lip closing movement as the length of the tongue movement trajectory increased. During the oral closure, 40%–80% of the tongue movement path was traversed, again with no clear pattern across subjects. Consonant voicing reliably influenced the tongue movement from the consonant release to the second vowel. During this interval, the tongue traversed a longer path for voiceless than for voiced stops, reflecting the longer voice onset time for the voiceless stops. [Work supported by NIH.]

4aSC26. Facial animation using visual polyphones. Åsa Hällgren and Bertil Lyberg (Telia Res. AB, Spoken Lang. Processing, S-136 80 Haninge, Sweden)

Today, synthetic speech is often based on concatenation of natural speech such as polyphones. So far there has mainly been two ways of adding a visual modality to such a synthesis: Morphing between single images or concatenating video sequences. In this study, however, a new method will be presented where recorded natural movements of points in the face are used to control an animated face. During the recording of polyphones to be used for synthesis, the movements of a set of markers placed in the face of the speaker are simultaneously registered in three dimensions. Connected to each acoustic polyphone will be a perfectly synchronized set of dynamic movement patterns, a visual polyphone. These visual polyphones are then concatenated in the same manner as the acoustic polyphones, and used to control the movements of an animated face. In the presented system, demisyllables are used as units both for the acoustic and the visual polyphones and it seems as these units will, in a good way, reflect the coarticulation process in natural speech.

4aSC27. On the meaning and the accuracy of the pressure-flow technique to determine constriction areas within the vocal tract. Xavier Pelorson, Salima Fahas, and Pierre Badin (Institut de la Commun. Parlée, INPG-Univ. Stendhal, 46 Ave. Félix Viallet, 38031 Grenoble Cedex, France, pelorson@icp.grenet.fr)

Since the work of D. W. Warren and A. B. Dubois [Cleft Palate J. **1**, 52–71 (1964)] was published, the ''pressure-flow technique'' has been widely used to estimate constriction areas within the vocal tract. However, two fundamental questions regarding this technique have not been clearly addressed: (1) What is exactly measured (minimum, maximum, or ''mean'' areas)? (2) Which accuracy can be expected from this technique? A theoretical and experimental study based on a vocal tract mechanical model, including various constriction shapes, is presented. Compared with much more complex viscous flow solutions, a simple one-dimensional flow model is shown to yield fair estimates of the areas (within 10%), except for low Reynolds numbers. The empirical head-loss factor $K=0.67$, sometimes used, appears meaningless and is probably due to an experimental artefact. The pressure-flow technique is shown to be relatively insensitive to the exact constriction shape (converging, straight, or diverging), and the estimated area to be close to the minimum area of the constriction. This result can be theoretically rationalized by considering that in all cases studied here the flow separation point is always close to the minimum constriction. Extensions and limits of these conclusions for unsteady flows will also be discussed.

4aSC28. An experimental study of the open-end correction coefficient for side branches within an acoustic tube and its application in speech production. J. Dang (ATR HIP, 2-2 Hikaridai Seikacho Soraku-gun Kyoto, 619-02 Japan), C. Shadle (Univ. of Southampton, UK), and H. Suzuki (College of Sci. and Technol., Tohoku, Japan)

The open-end correction coefficient (OECC) for side branches within an acoustic tube was investigated experimentally, and combined with morphological data of the vocal tract to arrive at OECC estimates useful in speech production. A number of mechanical acoustic models were used to examine the effects of the angle of the branch axis, and the proximity of the main tube around the open end of the branch. The results indicate that the OECC (α_1) depends on the cross-dimension L of the main tract at the branching point and the branch diameter D ($\alpha_1=0.414+0.517\cdot\log_{10}(L/D^2)$, $0.158<L/D^2<6.97$). For side branches connected to the main tract by a narrow neck, the OECC of each end of the neck is determined by $\alpha_2=0.82(1-\xi)/18\xi^2-3.1\xi+1$, where $\xi=$ the ratio of radius of the neck to that of the adjacent section. Using morphological data of the vocal-tract, the range of appropriate OECC values for estimating accurate vocal-tract transfer function was evaluated. The OECC of the nasopharyngeal port was about 0.8 for the oral-tract side, 0.3–0.6 for the nasal-tract side. For the ostia of the paranasal cavities, the value was ~0.8 for the cavity side, ~0 for the nasal-tract side.

4aSC29. A physiological model of the tongue and jaw for simulating deformation in the midsagittal and parasagittal planes. Jianwu Dang and Kiyoshi Honda (ATR Human Information Res. Labs., 2-2 Hikaridai Seikacho Soraku-gun Kyoto, 619-02 Japan)

A physiological articulatory model was developed for simulation of 3-D deformation and movement of speech organs. The model is composed of a midsagittal portion of the tongue and the jaw and hyoid bone that yield rotation and translation. The tongue model consists of a midsagittal layer and two parasagittal layers with 1-cm intervals. The form and muscle geometry of the whole model are extracted from MR images of a speaker. All the soft tissue and rigid organs are modeled by node points with mass and two types of branches to adjacent points: soft branches for muscles and connective tissue between node points, and hard branches for bony organs. The muscles are defined based on an improved Hill's model. Muscle activation signals, in part derived from EMG recordings, serve as the model's control parameters. Preliminary experiments revealed that the model can produce realistic 3-D deformation and movement for vowels and a sustained consonant /r/. Since all the displacements of soft tissue and rigid organs are computed with the same algorithm, the model produces plausible dynamic patterns of the tongue–jaw complex with relatively short computational time in comparison with the finite-element method.

Session 4aSP

Signal Processing in Acoustics: General Signal Processing

Brett Castile, Chair
Orincon Corporation, 9363 Towne Centre Drive, San Diego, California 92121

Contributed Papers

9:00

4aSP1. An application of laser-based ultrasonic nondestructive evaluation using a fiber-optics-based Fabry–Perot interferometer. Robert D. Huber, James V. Candy, Diane J. Chinn, and Graham H. Thomas (Lawrence Livermore Natl. Lab. P.O. Box 808, L-333, Livermore, CA 94550)

Fiber optics lend increased flexibility to laser-based ultrasonic nondestructive evaluation (NDE). In this work, fiber-optic cables are used to transmit light from a laser to the detection site, and then from the detection site to a Fabry–Perot interferometer. The use of fibers allows both the detection laser and interferometer to be placed at a considerable distance from the object under test. A direct line-of-sight of the object from the main equipment is not required, since the fibers may be fed through walls and around obstacles. In addition, by containing the laser light in the fibers, the chance of accidental exposure to powerful laser beams that may otherwise be transmitted through air is decreased. Laser-based ultrasonics is generally less sensitive to traditional contact ultrasonics, and in addition, some light is lost in the coupling of laser light energy into optical fibers, further decreasing the sensitivity; thus the need for signal processing of the received signals is of great importance. In this work, the waveforms obtained using the Fabry–Perot interferometer and the corresponding signal processing performed on the data to enhance the resulting image for NDE are discussed.

9:15

4aSP2. On-line measurement of error-path impulse response for an adaptive ANR headset. George J. Pan and Anthony J. Brammer (Inst. for Microstruct. Sci., Natl. Res. Council, Montreal Rd., Ottawa, ON K1A 0R6, Canada)

A digital, adaptive ANR headset system has been developed and tested in the laboratory, using a filtered-X LMS adaptive control algorithm and a floating-point digital signal processor (DSP). The implementation of the filtered-X LMS adaptation algorithm requires an impulse response (IR) model of the so-called ''error-path,'' which consists of the serial connection of an earphone, earcup acoustics, microphone, and electronic amplifiers and filters. Though the IR model can be measured off-line for laboratory studies using standard measurement procedure, on-line measurement or identification, using the DSP engine and digital signal processing techniques, is considered necessary and advantageous for practical applications. This paper discusses a DSP-based identification technique for the error-path IR model, which employs the well-known pseudorandom binary signal (PRBS), also called the maximum length sequence (MLS) signal, as the input to the error path. In this paper, the measurement method and system will be discussed. As an example, the error-path impulse response and the corresponding ANR results obtained in a laboratory setup will be presented. [Work supported by DCIEM, Toronto, Canada.]

9:30

4aSP3. Experiments on bit-rate reduction for orthophony signals. Ana M. Monsalve Caballero (Inst. of KW, Tech. Univ. of Berlin, Einsteinufer 17, Sekt. EN 8, R 452, 10587 Berlin, Germany)

Orthophony is the art of recording an acoustical sound field from a room and the reproduction of it by means of loudspeakers. For the reproduction of the signals, the simplest method is to use a transference matrix with the signals. The problem of the matrices is the sizes that they have, because they usually present dimensions too difficult to transmit or/and store. A way to solve the problem is using bit-rate reduction. Based on orthophony signals recorded in the studio of the institute of KW (Kommunikationswissenschaft) of the Technical University in Berlin and their corresponding matrices, experiments were accomplished for the reduction using MATLAB since it facilitates especially the use of matrices. The behavior of the different methods with different signal types (for example, speech, music, and noise) was observed and analyzed. The results are presented in this work. [This work was supported by Professor M. Krause, Institute of KW. Tech. Univ. of Berlin, Germany.]

9:45

4aSP4. Voiced excitation functions calculated from micropower impulse radar information. G. C. Burnett, T. J. Gable, J. F. Holzrichter, and L. C. Ng (Lawrence Livermore Natl. Lab., L-437, P.O. Box 808, Livermore, CA 94551, burnett5@llnl.gov)

Efforts underway at the Lawrence Livermore National Laboratory to use newly designed micropower impulse radars (MIR) to measure in real time the excitation function of the vocal tract will be presented. Studies undertaken in collaboration with the University of California at Davis and the University of Iowa with high-speed laryngoscopic cameras, electroglottographs, flow masks, and subglottal pressure transducers have solidified the relationship between the signal returned by the MIR and the voiced excitation function of the vocal tract. As a result, for the first time a transfer function of the vocal tract can be calculated in real time and with unprecedented clarity for voiced speech. This new capability could have significant implications for improvements in speech recognition and speech synthesis processing.

10:00

4aSP5. Comparison of conventional acoustic and MIR radar/acoustic processing of speech signals. T. G. Gable, G. C. Burnett, J. F. Holzrichter, L. C. Ng (LLNL, P.O. Box 808, L437, Livermore, CA 94551, gable2@llnl.gov), and W. A. Lea (Speech Sci. Inst., Apple Valley, MN 55124-0428)

Applications of the micropower impulse radar (MIR) to speech research at the Lawrence Livermore National Laboratory has produced potentially new methods of speech processing. They include the accurate calculation of vocal tract transfer functions, formant, and pitch analysis, and basic phoneme synthesis. These speech parameters have traditionally been in the realm of all-pole LPC calculations. Related research using the MIR radar has supplied an increasingly accurate voiced excitation function, which makes possible transfer function calculations using both poles

and zeros, yielding more accurate formant information and more natural sounding synthesis. This paper compares the newly obtained results with traditional LPC and cepstral approaches and demonstrates the improvements based on experimental data from several male and female subjects. The radar data also allow extremely accurate pitch tracking, which is simpler and more robust than that calculated by traditional means. This information can significantly enhance acoustic-based speech processing by allowing pitch adaptive processing.

10:15

4aSP6. Measurements of normal incidence surface impedance using the single transfer function method with multitone periodic random noise signal. Xavier Olny and Franck C. Sgard (LASH, DGCB URA CNRS 1652, ENTPE—Rue Maurice Audin, 69518 Vaulx-en-Velin Cedex, France, xavier.olny@entpe.fr)

Commonly used signals for determining the transfer function of a linear filter are pure tone excitation, white noise, or a pseudorandom noise (MLS) signal. However, it is known that the accuracy of this measurement depends strongly on the kind of signal used. In order to try to improve the quality of the measurement, multitone periodic random noise signal properties are investigated in the case of the normal incidence surface impedance measurement. Kundt's tubes and free-field impedance results using a single microphone transfer function method are given in a large frequency range (20–4000 Hz) for a multitone periodic random noise signal. These results are compared with those obtained with pure tone, white noise, and MLS signal excitation. Also, the influence of noise immunity is investigated for each type of signal.

10:30

4aSP7. Plane wave reflection coefficient estimation by use of spatial parametric signal modeling. Hyu-Sang Kwon and J. Stuart Bolton (1077 Ray W. Herrick Lab., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077)

The plane wave reflection coefficient of an absorbing infinite plane can be estimated from pressure distributions on two lines parallel to the surface. The measured pressures are decomposed into incoming and outgoing plane waves in the wave-number domain after the Hankel transform is applied to the spatial data. In practice, only a finite number of pressure measurements can be made on the measurement lines. As a result, errors resulting from aliasing and truncation are induced in the wave-number spectra and subsequently the reflection coefficient. The aliasing results from insufficiently small sample spacing, and can be easily controlled. The truncation error results from the limited size of the measurement aperture, which may be difficult to expand for practical reasons. Therefore, extrapolation techniques based on parametric signal modeling have been developed to ''virtually'' expand the measurement aperture, thus minimizing truncation effects. However, some restrictions, particularly on spatial sampling rate, must be considered in order to obtain good estimates of the reflection coefficient. Prony series modeling is one procedure that can be used to model and extrapolate the spatial sound pressure data. Reflection coefficients estimated by using the Prony series approach will be used to illustrate this procedure.

THURSDAY MORNING, 4 DECEMBER 1997 CALIFORNIA ROOM, 8:00 A.M. TO 12:00 NOON

Session 4aUW

Underwater Acoustics: Matched-Field Processing and Time Reversal

Hee Chun Song, Chair

Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0238

Contributed Papers

8:00

4aUW1. Matched-field processing for multitonal source localization. Zoi-Heleni Michalopoulou (Ctr. for Appl. Mathematics and Statistics, New Jersey Inst. of Technol., Newark, NJ 07102)

Recent studies with simulations and real data have shown that multiple frequency information can lead to increased robustness in matched-field estimation. This observation generates the question of how this information should be combined for best results. In this work an estimator is proposed that is suitable for matched-field inversion involving multitonal sources. The processor operates on the basis of an interfrequency spatially coherent framework. Designed for passive problems, it does not require knowledge of the source spectrum, an advantage over other coherent schemes. Moreover, the processor has the additional advantage of a simple implementation. Using synthetic data, the new coherent processor was compared to the conventional Bartlett processor with incoherent averaging across frequencies. The results indicate that the new processor is, in general, preferable to the conventional Bartlett processor. Interestingly, the superiority of the coherent processor is most pronounced when the receiving array undersamples the propagation medium, making the processor attractive for its cost effectiveness. [Work supported by ONR.]

8:15

4aUW2. Range-dependent multitone matched-field processing in SWellEX-3. Phil W. Schey, Newell O. Booth, Richard T. Bachman (NCCOSC RDTE DIV, Code D881, San Diego, CA 92152-5001), and William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0701)

SWellEX-3 was an experimental study of shallow-water propagation and matched-field processing conducted 10 km off San Diego in July, 1994. Data were recorded on a 64-element, 118-m aperture vertical line array anchored in 200 m of water. A cw source emitting tonals from 53–197 Hz was towed over tracks where the water depth varied from 200–75 m. Examples of range-dependent matched-field processing of this multitone signal using the finite-element parabolic equation program for propagation modeling will be presented.

8:30

4aUW3. Evaluation of different techniques in matched-field processing using measured data. Ahmad T. Abawi, Newell Booth, and Phil Schey (Ocean and Atmospheric Sci., Div. NCCOSC RDTE DIV, Code D881, San Diego, CA 92152-5001)

The data from May 1996 Shallow Water evaluation cell Experiment (SWellEx-96) are used to evaluate the performance of commonly used processors in matched-field processing (MFP). The objective of SWellEx-96, which was conducted west of Point Loma in shallow water, was to investigate source tracking using matched-field processing and low-level signal detection. Data were recorded on a 64-element 118-m vertical line array and on an identical 45-deg tilted line array anchored in 200 m of water. A cw source emitting tonals from 49–388 Hz was towed over tracks ranging up to 6 km from the array. Signals were transmitted in group of frequencies. The source level in each frequency group decreased by 3 dB every 3 Hz, thus making it possible to study MFP as a function of the signal-to-noise ratio (SNR). Various figures of merits have been used to evaluate the performance of the Bartlett, the minimum variance distortionless response (MVDR) as well as MVDR with white-noise gain-constraint (MVDRWNC) processors.

8:45

4aUW4. Broadband signal detection: Comparison of vertical aperture arrays using adaptive matched-field processing and a horizontal line array using adaptive plane-wave beamforming. Newell O. Booth, Phil W. Schey (NCCOSC RDTE DIV, Code D881, San Diego, CA 92152-5001), and W. S. Hodgkiss (Scripps Inst. of Oceanogr., San Diego, CA 92152-6400)

Detection and localization of low spectral-level, broadband signals in the presence of surface interference are described and illustrated with examples obtained in the May 1996 Shallow Water evaluation cell Experiment (SWellEx-96). The experiment was carried out west of Point Loma in 200-m water of complex bathymetry. A 118-m vertical line array (VLA) was deployed next to an identical line array tilted at 45 deg (Tilted VLA). A 240-horizontal line array was also deployed nearby. The detection performance of each of the vertical arrays with adaptive broadband matched-field processing (MFP) is measured and compared with the performance of the horizontal array with adaptive broadband plane-wave beamforming.

9:00

4aUW5. Detection of low-level broadband signals using adaptive matched-field processing with horizontal arrays. Newell O. Booth, Phil W. Schey (NCCOSC RDTE DIV, Code D881, San Diego, CA 92152-5001), and W. S. Hodgkiss (Scripps Inst. of Oceanogr., San Diego, CA 92152-6400)

The detection performance of low spectral-level, broadband signals from a submerged acoustic source in the presence of surface interference is described and illustrated with examples obtained in the May 1996 Shallow Water evaluation cell Experiment (SWellEx-96). The experiment was carried out west of Point Loma in 200-m water of complex bathymetry. Adaptive broadband matched-field processing (MFP) was performed on a 240-m aperture horizontal line array. Detection and localization in range, depth, and azimuth is illustrated at all bearings except near the array broadside. The effects of horizontal aperture on MFP robustness and sidelobe levels are also discussed.

9:15

4aUW6. Broadband matched-field processing. Arthur B. Baggeroer (MIT, Cambridge, MA 02139)

Several algorithms have been proposed for broadband matched-field processing; however, the issue of signal coherence across frequency and its implications for the signal coherence has not been treated consistently. If one uses a wide sense stationary random process model for a passive source, then spectral bands are uncorrelated, and one is led to an incoherent sum of energy of the optimal beamformer across frequency bins. Active sources imply a model where there is signal coherence across frequency which can be characterized using the two-frequency coherence function. Here, it is important to preserve phase in the output of the optimum beamformer since it contains the group speed information of the several rays and/or modes of the propagation. Models for transient signals fall in between the passive and active ones. The limited duration of the transient contains multipath/multimode information; however, the signal waveform is usually not known so matched filtering is not possible. The presentation will relate the issues of signal coherence broadband MFP algorithms and provide examples of the performance under different model assumptions.

9:30

4aUW7. Matched-field processing using Bartlett sidelobe interference structures. Aaron M. Thode, W. A. Kuperman, and Gerald L. D'Spain (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0205)

Waveguide invariants [e.g., G. A. Grachev, Acoust. Phys. **39**(1), 33–35 (1993)] predict the local behavior of interference maxima/minima of acoustic intensity in the frequency-range plane. These invariants are independent of the detailed environment of the waveguide. As the Bartlett matched-field processor (MFP) ambiguity surface has a physical interpretation in terms of a time-reversed acoustic field, with the sidelobes analogous to local interference maxima, these invariant concepts are reformulated for application to MFP. The theory predicts that when a series of perfectly matched range-independent Bartlett surfaces computed over small increments of frequency are stacked and then sliced along a constant depth, the resulting sidelobe contours in the range-frequency plane display interference patterns that converge to the correct range. Indeed, these patterns are demonstrated to exist in simulated data, broadband source tows, and blue whale vocalizations recorded off San Miguel Island last year. Using the Radon transform, it is even possible to localize a source at the correct range and depth using only the trajectories of the sidelobes of the MFP surfaces. The robust nature of the range-frequency invariant suggests that incoherent postprocessing based on this concept might yield results insensitive to environmental mismatch. [Work sponsored by ONR.]

9:45

4aUW8. Coherent and incoherent broadband matched-field processing with SWellEX-96 data. Gregory J. Orris, Michael Nicholas, and John S. Perkins (Naval Res. Lab., Washington, DC 20375)

Several approaches to passive broadband matched-field processing have been proposed. In most cases, an incoherent average over a narrow band is quite successful in driving down unwanted sidelobes. The ability of coherent broadband techniques to out perform this simple approach may require additional phase information about the source that, in general, may not be available. In this paper, a fully coherent method that can out perform an incoherent average is derived, although its practical application is problematic. In other words, the simple incoherent average may be the most efficient approach. These ideas are demonstrated using data from the SWellEx-96 experiment. Given the environmental conditions under which these data were taken, the new coherent approach has some physical advantages over other broadband methods since it more efficiently uses variations of range dependence exhibited by the different frequency components of the signal. The net result is that the range, depth, and azimuth ambiguities have comparable spatial extent.

10:00–10:15 **Break**

10:15

4aUW9. Absolute units for acoustic pressure levels: Shades of Carl Eckart and Bob Urick [ASA medalists as Pioneers in Underwater Acoustics, 1973 and 1988, respectively]. F. H. Fisher (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0701)

In *Principles of Underwater Sound* (1946), edited and largely written by Carl Eckart, he showed what acoustic pressure levels in dynes per square centimeter (rms) meant in decibels (air and water) compared with one another and also in watts per square centimeter. In *Principles of Underwater Sound* by Bob Urick, 3rd ed. (1983), Sec. 1.6 deals very clearly with absolute units of intensity with respect to the new standard pressure reference level of 1 (μPa)$=10^{-5}$ dyn/cm^2 (rms) for both air and underwater acoustics; see ''Preferred Reference Quantities for Acoustic Levels,'' ANSI S1.8-1969(R1974). In physical units this intensity amounts to 0.67×10^{-22} W/cm^2. A few examples will be given of difficulties encountered by not including absolute units, most recently with respect to whales and sound sources radiating only 250 W. The inclusion of absolute units would facilitate understanding of acoustics to the scientific, technical, and public communities.

10:30

4aUW10. Near-surface source depth resolution in a shallow-water environment. W. S. Hodgkiss, J. J. Murray (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0701), N. O. Booth, and P. W. Schey (NCCOSC RDTE DIV., Code D881, San Diego, CA 92152-5001)

The depth discrimination capability of matched-field processing provides the possibility of separating surface from submerged sources of acoustic energy. Data collected in 200-m water during SWellEx-96 will be used to illustrate the achieved depth resolution in a downward refracting, shallow-water environment. During the event analyzed, two sources were towed simultaneously. The nominal depths of the deep and shallow sources were 60 and 10 m, respectively. Both projected unique, multitone transmissions with the deep source covering the 50–400 Hz band and the shallow source covering the 100–400 Hz band. In addition, the source ship herself (R/V SPROUL) had a detectable radiated acoustic signature across this same frequency region. The data were received by a 120-m aperture, 64-element vertical array deployed from the R/P FLIP. Broadband adaptive matched-field processing has been carried out on the data. The depth resolution characteristics of each of the three sources of acoustic energy will be presented. [Work supported by ONR, Code 321US.]

10:45

4aUW11. Source localization in shallow water with internal waves. Kwang Yoo and T. C. Yang (Naval Res. Lab., Washington, DC 20375)

Source localization can be a difficult problem in temporally and spatially varying shallow-water environments. The SWARM 95 data indicate that the acoustic fields at 225 and 400 Hz change drastically in the time scale of several minutes. Since it is practically impossible to model the replica field at this time scale, one expects degraded performance for matched field processing (assuming a time-independent replica field). An alternate processing scheme is proposed here. First it is shown that internal waves cause mode conversions and that mode conversions degrade matched-field source localization. This is illustrated for the Garret–Munk internal waves, the solitary internal waves and the combination of both. It is then shown that source localization can be achieved by using only the uncoupled modes. This concept was implemented using matched-beam processing. The beam filter serves to suppress the contribution of the coupled modes. It is shown that this technique significantly improves the localization performance over conventional matched-field processing. [Work supported by the Office of Naval Research.]

11:00

4aUW12. Time-reversal focusing with less than a full water column source array. W. S. Hodgkiss, W. A. Kuperman, H. C. Song (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0701), T. Akal, C. Ferla (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), and D. R. Jackson (Univ. of Washington, Seattle, WA 98105)

Recent time-reversal mirror (TRM) experiments conducted in the Mediterranean Sea [Kuperman *et al.*, J. Acoust. Soc. Am. **101**, 3088(A) (1997)] have demonstrated the spatial focusing and temporal stability of the phase conjugation process over ranges of up to 30 km in a shallow-water waveguide. In the May 1997 experiment, the vertical source-receive array (SRA) consisted of 23 sources and receivers spanning 77 m of a 125-m water column. One series of tests involved a probe source (PS) and 94-m aperture vertical receive array (VRA) located 15 km from the SRA. Various subsets of the SRA sources were used for retransmitting the time-reversed signals received by SRA receivers from the PS. With a single source, the results essentially were the same as those reported previously [Parvulescu and Clay, Radio Elec. Eng. **29**, 223–228 (1965)]. Temporal compression is achieved but vertical spatial focusing is not. By using all 23 sources, the retransmitted field focuses distinctly at PS with levels above and below the focal region suppressed by more than 15 dB. Significant spatial focusing also was achieved with as few as six sources where these levels were suppressed by more than 10 dB. [Work supported by ONR, Code 321US.]

11:15

4aUW13. Comparisons of computed and exact time-reversing array retrofocusing in a shallow-water sound channel. Michael R. Dungan and David R. Dowling (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109)

A time-reversing array (TRA) can retrofocus acoustic energy to the location of its original source even when the acoustic environment is complex and unknown. However, predictions of array retrofocusing performance require an acoustic propagation model for the environment. Results are presented from a study of the sensitivity of TRA retrofocusing to acoustic propagation modeling. Results from a modal sum propagation model and a parabolic equation (PE) code (RAM by M. D. Collins) are compared for a perfect waveguide. The influence of waveguide complexity on retrofocusing is determined by adding range dependence and realistic bottom properties to the computed PE fields. Variations in acoustic frequency and sound source depth are also explored. It is found that acoustic penetration into the bottom primarily influences sound pressure levels but does not strongly influence the size of the TRA retrofocus. In addition, small changes in the waveguide's range dependence are found to have little effect on retrofocus properties. [Work supported by the Office of Naval Research.]

11:30

4aUW14. Variable range focusing in a time-reversal mirror. Hee Chun Song, William A. Kuperman, William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238), T. Akal, C. Ferla (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), and Darrell R. Jackson (Univ. of Washington, Seattle, WA 98105)

Recently, a time-reversal mirror (or phase conjugate array) was demonstrated experimentally in the ocean, which spatially and temporally re-focused an incident acoustic field back to the original position of the probe source. [Kuperman *et al.*, accepted for publication in J. Acoust. Soc. Am.]. Here, this waveguide time-reversal mirror technique is extended to refocus at ranges other than that of the probe source. This procedure is based on the acoustic-field invariant property in the coordinates of frequency and range in an oceanic waveguide [G. A. Grachev, Acoust. Phys. **39**(1), 33–35 (1993)]. This technique has been verified in a second phase conjugate experiment conducted in the Mediterranean Sea in May 1997. The theory behind variable range focusing will be presented along with experimental results demonstrating the ability to shift the focal range ±1 km for a probe source in the vicinity of 15 km from the source–receive array.

11:45

4aUW15. Acoustic time reversal in nonlinear wave propagation. Ibrahim M. Hallaj, Steven G. Kargl (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105), and Paul E. Barbone (Boston Univ., Boston, MA 02215)

The application of acoustic time reversal in nonlinear underwater propagation is investigated theoretically and numerically. An acoustic-wave equation containing thermal and viscous absorption as well as finite-amplitude terms is presented. The behavior of such a system under time reversal is discussed. Theoretical results and numerical examples are presented, showing when and why the propagation is time reversible, and when the equation is no longer time invariant. [IMH and SGK's work is supported by the Office of Naval Research.]

THURSDAY MORNING, 4 DECEMBER 1997 — SHEFFIELD ROOM 8:00 TO 9:30 A.M.

Meeting of the Standards Committee Plenary Group

S1 Acoustics U.S. Technical Advisory Group (TAG) for IEC/TC 29 Electroacoustics and ISO/TC 43 Acoustics

S3 Bioacoustics U. S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics, IEC/TC 29 Electroacoustics, and ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock

S12 Noise U. S. Technical Advisory Group (TAG) for ISO/TC 43/SC1 Noise and ISO/TC 94/SC12 Hearing Protection

An initial meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S3 and S12, to meet in the following sequence on the same day: S12—9:30 a.m. to 11:30 a.m.; S1—2:00 p.m. to 3:15 p.m.; S3—3:30 p.m. to 4:45 p.m.

(Note separate listings for these Standards Committee Meetings in the program)

Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all S Committees, plus a review of the international standardization activities including reports on recent meetings and planning for forthcoming meetings.

Members of S2 on Mechanical Vibration and Shock (and U.S. TAG for ISO/TC 108 and five of its Subcommittees, SC1, SC2, SC3, SC5 and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting, even though the S2 meeting will take place on Wednesday, 3 December 1997 at 9:00 a.m.

The U. S. Technical Advisory Group Chairs for the international Technical Committees and Subcommittees under ISO and IEC, and parallel to S1, S2, S3 and S12 are:

P. D. Schomer, U.S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics and ISO/TC 43/SC1 Noise

H. E. von Gierke, Vice Chair U.S. Technical Advisory Group (TAG) for ISO 43 Acoustics and ISO/TC 43/SC1 Noise

E. H. Berger, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 94/SC12 Hearing Protection

J. Erdreich, U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock

V. Nedzelnitsky, U.S. Chair, Technical Advisory Group (TAG) for IEC/TC 29 Electroacoustics

Meeting of Accredited Standards Committee S12 on Noise

P. D. Schomer, Chair S12, and Chair U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
U. S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

B. M. Brooks, Vice Chair S12
Brooks Acoustics Corporation, P.O. Box 3322, Vernon, Connecticut 06066

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
1325 Meadow Lane, Yellow Springs, Ohio 45387

E. H. Berger, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 94/SC12, Hearing Protection
E-A-R/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Accredited Standards Committee S12 on Noise. Working group chairs will report on their progress for the production of noise standards.

Scope of S12. Standards, specifications and terminology in the field of acoustical noise pertaining to methods of measurement, evaluation and control; including biological safety, tolerance and comfort and physical acoustics as related to environmental and occupational noise.

THURSDAY AFTERNOON, 4 DECEMBER 1997 SENATE/COMMITTEE ROOMS, 1:25 TO 5:00 P.M.

Session 4pAA

Architectural Acoustics: Paul S. Veneklasen Memorial Session

Leo L. Beranek, Cochair
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138-5755

David Lubman, Cochair
D. Lubman & Associates, 14301 Middletown Lane, Westminster, California 92683

Contributed Papers

1:25

4pAA1. Paul S. Veneklasen, Father, mentor. Lee H. Veneklasen (3445 Badding Rd., Castro Valley, CA 94546)

Early recollections included making things, riding bikes, sailing, boy scouts, skiing, many of the usual things that a father does with his son, but my father, Paul was different. He wanted us to learn why and how did it work; was there a better way of doing it? This atmosphere created a desire for learning and innovation. My father encouraged my education and revelled as I approached his field of electron-beam lithography from experiment and the machine shop, much as he approached acoustics. My brother, Mark, a wonderful craftsman, developed a new French Horn, with the help of my father's patient advice and occasional prodding. A few anecdotes illustrate a lifetime of memories.

1:40

4pAA2. Paul S. Veneklasen: A glimpse into his early years at UCLA, 1938–1942. Robert S. Gales (1645 Los Altos Rd., San Diego, CA 92109)

Paul joined the acoustics program of the UCLA Physics Department in 1938, shortly after receiving his M.S. in physics and mathematics from Northwestern University. He was a spirited contributor to Professor Vern Knudsen's weekly seminars dealing with nearly all aspects of acoustics: propagation, architectural, hearing, music, etc. Here, he shared ideas with other students: Dick Bolt, Cyril Harris, Bob Leonard, Walter Rosenblith, Izzy Rudnick, Ludwig Sepmeyer *et al.*, and mentors: Leo DelSasso, Vern Knudsen, and Norman Watson. In 1941 Paul became a key contributor to a new research effort to develop improved hearing protection devices as military acoustics became important in those pre-World War II years. This paper is a brief anecdotal glimpse into this short period of Paul's early professional life.

1:50

4pAA3. Paul S. Veneklasen: The War Years, 1941–1945. Leo L. Beranek (975 Memorial Dr., Ste. 805, Cambridge, MA 02138)

Paul was a senior research associate at the Electro-Acoustic Laboratory (directed by Leo L. Beranek) and the Psycho-Acoustic Laboratory (directed by S. Smith Stevens) during World War II. This paper discusses his contributions during this period to the hearing protection of military personnel and to the development of special electronic equipment for the precision measurement of the properties of sound fields.

2:00

4pAA4. Paul S. Veneklasen, 1945–1996, an overview. Jerry P. Christoff (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

The author joined Paul S. Veneklasen & Associates and the Western Electro-Acoustic Laboratory in 1956, the day after his graduation from UCLA with a B.S. in Applied Physics and a specialization in acoustics. During the next forty years, he was privileged to work side by side with Paul Veneklasen on a variety of projects ranging from speech communication in noise, acoustical model testing, and a wide variety of architectual acoustical commissions. Most people think of Paul in only one field of endeavor, but those who knew him best knew that as a physicist, he would be interested and try anything to make an advance in technology. The paper provides a brief chronology of his major interests, projects, patents, and unpublished ideas.

2:20

4pAA5. Paul S. Veneklasen, contributions to electroacoustic system and design. Jose C. Ortega (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

Paul Veneklasen's interest in acoustics was wide ranging. Research and development in electroacoustics occupied his endeavors in the early years. This paper will present a survey of his work. Topics to be presented include development of a series of condenser microphones and loudspeaker systems at the Altec Lansing Corporation, and development of instrumentation to measure the sound pressure level using the Western Electric 640 AA condenser microphone. Projects of a more unusual nature included the development of a Fabric Acoustimeter or ''Fuzz Feeler'' to measure by acoustic means the ''handle'' of a fabric, development of the tooth and forehead microphones for detecting and transmitting voice signals in very high-level noise enviroments and research and development of the electronic stethoscope. This instrument extended the frequency range of heart sound analysis. Paul's interest and pioneering work in auditorium acoustics led to contributions to the development of instrumentation for use in acoustical scale modeling research.

2:40

4pAA6. Paul S. Veneklasen, contributions to psychological acoustics and speech communication systems. Jerry P. Christoff and Jose C. Ortega (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

Early in the authors' careers working with Paul Veneklasen at the Western Electro-Acoustic Laboratory, the authors were involved with programs dealing with speech communication in high-noise fields and the development of a new ear plug, possibly to replace the venerable V51-R ear defender that was developed by Paul Veneklasen, Vern O. Knudsen, and others during World War II. His later interests were involved with the perception of lateral reflections in rooms, which Knudsen called ''envelopment,'' primarily through research in his patented Auditorium Synthesis demonstration room. Highlights of this work are recounted including the novel forehead, ear, and tooth microphones.

2:55

4pAA7. Paul S. Veneklasen, contributions to enviromental noise control. Jose C. Ortega (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

Paul had a deep and continuing interest in controlling ennviromental noise. He was instrumental in organizing one of the earliest symposiums on urban noise. His ''City Noise—Los Angeles'' paper in *Noise Control Magazine* in July 1956 is filled with prophetic comments on urban noise and its potential effect on people. Paul was a guiding influence on the development of a Model City Noise Ordinance commissioned by the Department of Health of the State of California. Paul not only reported on enviromental noise but was actively engaged in solving noise problems. His work on the development of a ground run-up suppressor system design for the Northrop F-89 aircraft is a text book example of defining, understanding, and the developing methods to reduce enviromental noise from a very high-level noise source.

3:10–3:25 Break

3:25

4pAA8. Acoustical modeling and simulation in the 1960s: Paul Veneklasen's quest for ''envelopment'' and correlations with auditorium design. Jerald R. Hyde (Jerald R. Hyde Acoust., Box 55, St. Helena, CA 94574)

Paul's favorite ''in-house'' project in the 1960's was his auditorium synthesis system. He devised an electro-acoustic laboratory listening environment which simply divided a test sound field, excited by an anechoic recording, into the three basic components called ''direct,'' ''envelopment,'' and ''reverberance.'' The envelopment factor was the summation of early delayed reflections which surrounded the listener at the sides via loudspeakers. Paul was primarily looking for how much envelopment, or lateral early energy, a listener preferred relative to the direct sound level. At around the same time, he was using a nitrogen jet ultrasonic noise source to model test exactly how much lateral early energy was possible. He found that in a conventional hall with the best possible lateral energy design, the amount of envelopmental energy available at most seats was around 3 to 5 dB lower than ''desired.'' Concurrently, Marshall and then Barron extended the previous work of Keet and Kuhl in developing the parameter called ''lateral energy fraction,'' which essentially quantified Veneklasen's envelopmental energy. Computer systems later allowed the actual measurement of typical lateral energy values in concert halls. Veneklasen's findings will be compared to the actual measured values for hall's of well-known quality.

3:40

4pAA9. Paul Veneklasen and the design and engineering of the dampened mild carbon steel stage acoustic shell: A reminicence. George C. Izenour (16 Flying Point Rd., Stony Creek, CT 06405)

Paul and I were introduced to each other in the early 1960s by Vern Knudsen with whom I was then collaborating on Frank Lloyd Wright's last public building, The Grady Gammage Auditorium on the Campus of Arizona State University (1964). My initial collaboration with Paul, with Vern Knudsen looking over both our shoulders, was the multipurpose arena on the campus of Washington State University. Simultaneously with the Washington State Project (1973) there were two other collaborations: (1) the remodeling-restoration of the Eastman theater on the campus of the University of Rochester (1972) and (2) the Hancher auditorium on the campus of the University of Iowa (1973). All told, my three collaborations, with Paul as acoustics consultant and myself as theater consultant, extended over a period of about 5 years (1968–1973). It is our collaboration on the Eastman project that is the subject of this paper.

4:00

4pAA10. Paul S. Veneklasen, contributions to architectural acoustical design, 1960–1996. Jerry P. Christoff and James A. Good (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

Paul Veneklasen will be remembered most for his contributions to architectural acoustics, especially completed, highly acclaimed auditoriums, ranging from the Seattle Opera House and Los Angeles Music Center to Wolf Trap Farm Park to name a few. But he also left a legacy of hundreds of radio, television, and recording studios, and virtually every other building type from office buildings to convention centers and even jet engine test cells and noise suppressors. But less known are his development of orchestra enclosures, pit lifts, multilayer gypsum board construction, damped glass panels, floating floors, and sound doors incorporating magnetic seals and acoustical filters. His research extended from optimizing sight lines and seating arrangements in auditoriums, acoustical modeling, and theater chairs. Highlights of Paul Veneklasen's well-known projects, developments, and research are presented.

4:20

4pAA11. The auditorium measurement chain: A study of the use of auditorium synthesis in the design of concert halls. John J. LoVerde (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

In the early 1960s, Paul Veneklasen began to develop an idea to synthesize an indoor acoustical space coming to fruition with the development and construction of an electroacoustic system called Auditorium Synthesis. This system allows the listener to preview various types of spaces prior to construction [Paul S. Veneklasen, Int. Symposium on Arch. Acoustics, 21–43 (1974)]. Listener preferential settings including source strength (G), early decay time (EDT), early/late energy (C), and lateral energy fraction from the Otis A. Singletary Center for the Arts (Lexington, Kentucky). A comparison of preferred orchestral settings and modeled seat locations will be performed. With impulse response data from the completed concert hall at the Singletary Center for the Arts, a comparison will be shown of newer parameters between the modeled seat locations from Auditorium Synthesis and concert hall data.

4:35

4pAA12. The acoustical modeling of Paul Veneklasen. Gary W. Siebein and Martin A. Gold (Univ. of Florida, Gainesville, FL 32611)

The pioneering acoustical modeling work of Paul Veneklasen has served as an inspiration to much current work in the area, including that conducted at the University of Florida. Paul constructed small models of spaces he was designing and conducted a variety of acoustical tests in them to assist in making design decisions on many innovative rooms. He determined criteria by which to evaluate the acoustic response of the rooms obtained from the models from listening to simulated sounds in his laboratory. These criteria included: providing a strong direct sound throughout the room; providing reflections from overhead within 40 ms or less of the direct sound for clarity and loudness; providing lateral reflections within 80 ms of the direct sound for envelopment; and providing a rich reverberant sound across a full range of frequencies. These criteria are still used today as the basis for interpreting the impulse responses obtained in the design process of rooms. The experiments he engaged in as a part of the process of many rooms involved research and design components including listening, modeling, designing, and building can be traced in concept to the work of Wallace Sabine. This process has become an established approach to acoustical design today.

4:50

4pAA13. Paul S. Veneklasen, mentor and associate. Jerry P. Christoff and Jose C. Ortega (Veneklasen Assoc., 1711 Sixteenth St., Santa Monica, CA 90404)

The authors both arrived at Paul S. Veneklasen & Associates and Western Electro-Acoustic Laboratory in 1956 and 1957, fresh from UCLA's acoustics program which was part of applied physics. In an era when most people change jobs numerous times during their careers, what was the glue that kept the authors there? Security? Technical challenge? Opportunity to work with a truly creative, renaissance person? Probably all of the above. But more than acoustics and business were learned from Paul. A few anecdotes illustrate the affection for Paul and the legacy he left.

THURSDAY AFTERNOON, 4 DECEMBER 1997 — WINDSOR ROOM, 1:00 TO 2:45 P.M.

Session 4pABa

Animal Bioacoustics: Reptile and Amphibian Bioacoustics

Peter M. Narins, Chair
Department of Physiological Science, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095

Chair's Introduction—1:00

Invited Papers

1:05

4pABa1. Temporal processing in water and air: Developmental changes in synchronized response to AM signals across metamorphosis in the bullfrog. Seth S. Boatright-Horowitz and Andrea M. Simmons (Dept. Neurosci., Brown Univ., Box 1953, Providence, RI 02912)

Acoustic limitations on propagation of low-frequency complex communication sounds make the amplitude modulation (AM) rate an important auditory feature for both aquatic larval (tadpole) and partly terrestrial (adult) bullfrogs. Phase-locked responses to AM signals were recorded from neurons in the auditory midbrain during metamorphic and postmetamorphic development. The best modulation frequency (BMF) and maximal significant AM rate are significantly negatively correlated with the developmental stage. Pre- and early prometamorphic tadpoles show phase locking to higher AM rates than postmetamorphic animals. Late prometamorphic (deaf period) tadpoles show a substantial reduction in auditory sensitivity, lower BMFs, reduced bandwidths of modulation transfer functions (MTF), and loss of significant phase locking. By the onset of the metamorphic climax, there is a recovery of auditory sensitivity and phase-locked responding, with late climax tadpoles showing auditory sensitivity similar to that of postmetamorphic

4p THU. PM

frogs. The changes in neural responding across metamorphosis are correlated with morphological changes in the auditory periphery and central auditory nuclei, and match differences in the acoustic environment as the animal shifts from an aquatic to a more terrestrial lifestyle. [Research supported in part by an NIH research Grant NS28565 (AMS) and an NSF Graduate Fellowship (SBH).]

1:25

4pABa2. Sound radiation and postglottal filtering in frogs. Alejandro P. Purgue (Dept. Physiol. Sci., UCLA, Los Angeles, CA 90095, apurgue@ucla.edu)

This work presents a comparison across selected species of several aspects of the mechanism of sound radiation in frogs. Measurements of the magnitude of the transfer function of the radiating structures show that the structures radiating the bulk of the energy present in the call vary depending on the species. Bullfrogs (*Rana catesbeiana*) radiate most of the energy (89% sound level) in their calls through their eardrums. Additionally, the transfer function of the eardrums displays peaks coincident with those observed in the mating and release call of this species. The vocal sac and gular area contribute energy only in the lower band (150–400 Hz). The ears are responsible for radiating additional frequency bands to the ones radiated through the vocal sacs. In *Rana pipiens*, the ears also broadcast a significant portion of the energy present in the call (63% sound level), but the frequencies of the aural emissions are a subset of those frequencies radiated through the vocal sacs. Finally, the barking treefrog (*Hyla gratiosa*) appears to use two different structures to radiate his call. The low-frequency band is preferentially radiated through the lungs while the high-frequency components are radiated through the vocal sac.

1:45

4pABa3. Acoustic tuning mechanisms in reptilian hair cells. Ruth Anne Eatock (The Bobby R. Alford Dept. of Otorhinolaryngology and Communicative Sci., Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, eatock@bcm.tmc.edu)

Tuning by hair cells can be mechanical or electrical. The best-studied examples of each class are from reptilian ears. In the free-standing region of the alligator lizard's cochlea, where hair cells have acoustic characteristic frequencies (CF_a's) that range systematically from 1 to 4 kHz, acoustic tuning can be explained by micromechanical resonances of the hair bundles [D. M. Freeman and T. F. Weiss, Hear. Res. **48**, 37–68 (1990)]. In the turtle cochlea, hair cells have CF_a's from 30 to 700 Hz and the acoustic tuning derives from electrical resonance of the hair cell membrane, reflecting the interplay of capacitive current and several ionic currents [Y.-C. Wu *et al.*, Prog. Biophys. Mol. Biol. **63**, 131–158 (1995)]. Tonotopic variation in CF_a arises from variation in hair bundle height in the lizard and from variation in ion channel kinetics and number in the turtle. Coupled mechanical and electrical tuning [T. F. Weiss, Hear. Res. **7**, 353–360 (1982)] seems likely but has not been demonstrated. Hair bundle height varies with Cf_a in the turtle cochlea, suggesting a micromechanical contribution. Hair cells dissociated from the liquid cochlea resonate electrically, but at frequencies tenfold lower than their CF_a's [R. A. Eatock *et al.*, J. Neurosci. **13**, 1767–1783 (1993)].

2:05

4pABa4. Automated bird songs recognition using dynamic time warping and hidden Markov models. Joseph A. Kogan and Daniel Margoliash (Dept. of Organismal Biol. and Anat., 1027 East 57th St., The Univ. of Chicago, Chicago, IL 60637, joseph@modeln.uchicago.edu)

The performance of two well-known recognition techniques in speech adapted to the automated recognition of bird song units from continuous recordings is studied. The advantages and limitations of dynamic time warping (DTW) and hidden Markov models (HMMs) are evaluated on a large database of songs of two species: zebra finches (*Taeniopygia guttata*) and indigo buntings (*Passerina cyanea*), which have different types of vocalizations and have been recorded under different laboratory conditions. The recognition performance of these methods is also assessed with regard to song signal representation (including representations commonly used in speech recognition), model structure, and sensitivity. Depending on the quality of recordings and song complexity, the DTW-based technique gives from excellent to satisfactory results. Under more challenging conditions, such as noisy recordings or the presence in bird vocalizations of confusing transient calls, HMMs can significantly outperform DTW but requires more training examples. Even though the HMMs perform usually quite well, they often misclassify short transient calls and song units with more variable structure. To address these and other weaknesses of the studied techniques, new approaches to analyze transient and variable bird vocalizations are discussed. [Work supported in part by U.S. Army Research Office and NIH.]

2:25

4pABa5. Desert tortoises (*Gopherus agassizii*) lack an acoustic startle response: Implications for studies of noise effects. Ann E. Bowles and Scott A. Eckert (Hubbs-Sea World Res. Inst., 2595 Ingraham St., San Diego, CA 92109)

The startle reflex is little studied in nonmammalian vertebrates, but in mammals it is often characterized by increase in heart rate. Orienting can also occur, characterized by a brief decrease in heart rate. To determine whether these responses occur in a testudinate, the desert tortoise, heart rate and behavior were measured during exposures to simulated jet overflights (94.6–114.2 dB CSEL) and sonic booms (6–10.5 psf). The best sensitivities of the 14 subjects ranged from 23 dB–50 dB SPL (average 34 dB SPL at 250 Hz) measured using ABR. Initial exposure to simulated jet overflights produced a typical reptilian defensive response: freezing. Tortoises became totally immobile for periods of up to 113 min. The average heart rate showed a gradual decline (7–8%), recovering within 2–4 h. Tortoises oriented when exposed to simulated sonic booms and after repeated exposure to subsonic aircraft noise; orienting produced no detectable change in heart rate. These results suggest that (1) high-intensity transients affect desert tortoises by altering activity patterns, but do not have a direct effect on heart rate and (2) tortoises experience a physiological response that produces protracted freezing. [Work supported by USAF, Edwards Air Force Base, and F-22 SPO.]

Session 4pABb

Animal Bioacoustics: General Topics: Effects of Noise on Animals

Ann E. Bowles, Chair
Hubbs-Sea World Research Institute, 2595 Ingraham Street, San Diego, California 92109

Contributed Papers

2:50

4pABb1. Effects of fixed-wing military aircraft noise on California gnatcatcher reproduction. Frank T. Awbrey and Don Hunsaker II (Hubbs Sea World Res. Inst., 2595 Ingraham St., San Diego, CA 92109)

To test the assumption that high levels of aircraft noise impede bird reproduction, noise analyzers were placed for 1 week in the nesting territory of each of 39 California gnatcatcher pairs on Naval Air Station Miramar. The 1-week average sound levels (7DL) recorded in those nesting territories were then related to the number of nest attempts; number of eggs laid; number of chicks hatched; number of chicks fledged; and number of eggs, chicks, and fledglings per nest attempt. Nest attempts and eggs laid have weak negative correlations ($p=0.14$ and 0.28) with 7DL. That is, the birds may tend to build fewer nests and lay fewer eggs in noisier areas, which is consistent with the common observation that bird nesting is more easily disturbed before eggs are laid than after. None of the other indicators is correlated with sound levels. Once a nest is established, with eggs in it, military aircraft noise has no detectable influence on reproductive performance. Gnatcatchers reproduced in places where 1 HL exceeds 80 dB for several hours every day. If fixed-wing aircraft noise impedes California gnatcatcher reproduction, it is overwhelmed by such factors as disturbance, predation, weather, edge effects, and differences in quality of habitat.

3:05

4pABb2. Response of elephant seals to acoustic thermometry of ocean climate sound transmissions. Daniel Costa, Daniel Crocker, James Gedamke, Paul Webb, Burney Le Boeuf, Danielle Waples, Sean Hayes, and James Ganong (Dept. of Biol., Univ. of California, Santa Cruz, CA 95064)

The hypothesis that northern elephant seals would respond to acoustic transmissions from the ATOC sound source was tested. Elephant seals were chosen because they have the best low-frequency hearing of any pinniped, are abundant, naturally migrate past the Pioneer Seamount, and are deep divers. ARGOS satellite tags provided information on animal location while at sea, while archival tags provided information on swim speed, time, depth, ambient acoustic environment, and ambient sound pressure levels. Instruments were deployed on 14 naturally migrating adult male elephant seals that were expected to swim near the source site, and upon 29 juvenile animals that were translocated and released 1 h prior to transmission. Measured mean intensity of ATOC exposure ranged from 120 to 135 dB for 60–90 Hz ($n=6$) compared to ambient levels of 100–107 dB (60–90 Hz). Animals did not alter return track, diving pattern, or swim speed, did not go to the surface, and often continued to dive closer to the sound source if on the descending segment of a dive. [Work funded by ONR and ARPA.]

3:20

4pABb3. Acoustic interaction of humpback whales and whale-watching boats off Maui, Hawaii. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734) and Marsha Green (Albright College, Reading, PA 19612)

Whale-watching activities during the humpback whale annual winter sojourn in Hawaii have increased steadily over the past years. A concern of many is the possible detrimental effects of whale-watching boats on humpback whales; including the possibility of damage to their auditory system. To address this concern, the noise generated by five different types of whale-watching boats was measured, two with outboard engines, two with inboard diesel engines, and a seagoing vessel, in the near shore waters of west Maui. The first series of measurements was conducted at the peak of the whale season, and the boat noises were partially masked by humpback whales' chorusing (boat noises and whale chorus levels were on the same order of magnitude). A 1/3 octave band spectral analysis indicated whale chorus levels as high as 136 dB re: 1 μPa with no whales in sight. Throughout the day humpback whale chorusing noise was present at similar levels for miles along the same coastline. Measurements taken close to the end of the season had about a 15-dB decrease in the humpback whale chorusing levels. Considering the high level of chorusing by humpback whales, the acoustic effects of the whale-watching boats should be relatively small or negligible.

3:35

4pABb4. Synthesis of distant signatures of underwater explosions for sea mammal hearing studies. Joseph A. Clark and Jane A. Young (CDNSWC, Bethesda, MD 20084 and COMB, Ste. 236, Columbus Ctr., 701 East Pratt St., Baltimore, MD 21202)

Experiments which measure temporary threshold shifts (TTS) in marine mammals produced by loud noises provide one means for determining the effects that underwater explosions might have on sea life in the neighborhood of the explosion [S. Ridgway *et al.*, J. Acoust. Soc. Am. **101**, 3136(A) (1997)]. In order to obtain meaningful results which do not presuppose an understanding of the possible hearing damage mechanisms, the waveforms of the noises used as stimuli in the TTS measurements should reproduce the complexities of the actual sound field experienced by sea mammals at the outer edges of the region within which significant acoustic effects on hearing are expected. Multipath arrivals and significant refractive effects complicate the far-field signals produced by explosive shock tests. For example, a large number of superimposed short impulses are often observed. In this talk a system for synthesizing distant signatures of underwater explosions will be described. The computer-controlled system takes calculated pressure-time waveforms as input and generates an acoustic simulation of the input signal. Preliminary tests to validate the operation of the system will also be reported.

Session 4pABc

Animal Bioacoustics: Animal Bioacoustics Poster Session

Ann E. Bowles, Chair

Hubbs Sea World Research Institute, 2595 Ingraham Street, San Diego, California 92109

All posters will be on display from 4:00 p.m. to 6:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 4:00 p.m. to 5:00 p.m. and contributors of even-numbered papers will be at their posters from 5:00 p.m. to 6:00 p.m.

Contributed Papers

4pABc1. Bat jitter discrimination, adaptive channel equalization, and the critical interval. Richard A. Altes (Chirp Corp., 8248 Sugarman Dr., La Jolla, CA 92037, altes@msn.com) and James A. Simmons (Brown Univ., Providence, RI 02912)

When simulated echoes have a bandwidth between 20 and 80 kHz, bats (*Eptesicus fuscus*) can discriminate between pulse-to-pulse delay jitter and constant echo delay with a sensitivity of 10 ns [Simmons *et al.*, J. Comp. Physiol. A **167**, 589–616 (1990)]. To investigate this amazing acuity, high-pass and low-pass filters were inserted into the channel between the simulated target and the bat. The effect of such filters on jitter acuity was measured. The data are explained by a receiver model that attempts to equalize the channel filter. The equalizer uses a smeared version of the actual channel transfer function, where spectral smearing induces relatively slow cutoff. Spectral smearing occurs in spectrogram (short-time Fourier transform) analysis. Smearing varies inversely with the impulse response duration of the analysis filters. This impulse response duration corresponds to the critical interval, which is frequency dependent for proportional bandwidth filtering. Bat jitter discrimination data imply channel equalization with a smeared version of the actual channel transfer function. This smearing is commensurate with critical intervals that are obtained from other experiments [Simmons *et al.*, J. Acoust. Soc. Am. **86**, 1318–1332 (1989)]. Bats apparently use time-frequency echo representations to perform environmentally adaptive filtering.

4pABc2. Quantitative and qualitative analysis of the ascendent frequency whistles of tucuxi, *Sotalia fluviatilis*, in the Sepetiba Bay, Rio de Janeiro, Brazil. Sheila M. Simao, Luciana D. Figueiredo, and Salvatore Siciliano (UFRRJ, Dept. Ciencias Ambientais, Seropedica, Rio de Janeiro, 23851-970, Brazil)

Sotalia fluviatilis was the only cetacean species observed in all 16 boat surveys (82 h) conducted in the Sepetiba Bay (22° 58′, 44° 02′W), between December 1993 and August 1995. During the boat survey of 5 January, 1995—a very calm day (Beaufort = 1)—it was possible to make good subaquatic sound recordings (30 min) of a group of *ca*. 20 young and subadult dolphins. As there is little bioacoustic information on tucuxi, sonogram analysis was used to characterize the tucuxi's whistles of ascendent frequency. The qualitative analysis resulted in 866 whistles divided in 48 types. Some of these sonograms shapes match the ones presented by other cetacean species sonograms (*Tursiops truncatus, Delphinus delphis, and Globicephala macrorhynchus*). The quantitative analysis was based on: number of harmonics; number of inflexions; beginning, final, maximum, and minimum frequecies; duration; and frequency modulation. Minimum, maximum, and average values, standard deviation, and coefficient of variation were calculated for each type of whistle. The *S. fluviatilis* whistles are extremely diverse (high coefficient of variation values for each type) and frequent (38.5 whistles/min), seeming to have an important role in social relations among these animals. Such whistle diversity and quantities could be explained by the composition of the dolphin group, i.e, it was exclusively formed by young and subadult animals.

4pABc3. Quantitative and qualitative analysis of the descent, up–down, and down–up frequency whistles of tucuxi, *Sotalia fluviatilis*, in the Sepetiba Bay, Rio de Janeiro, Brazil. Tereza C. C. L. Pereira, Sheila M. Simao, and Salvatore Siciliano (UFRRJ, Dept. Ciencias Ambientais, Seropedica, Rio de Janeiro, 23851-970, Brazil)

Sotalia fluviatilis was the only cetacean species observed in all 16 boat surveys (82 h) conducted in the Sepetiba Bay (22° 58′S, 44° 02′W), between December 1993 and August 1995. Previously, *S. fluviatilis* whistles were bioacoustically poorly characterized. The objective of this research was a quali–quantitative characterization of tucuxi's acoustic emissions of descent, up–down, and down–up frequency. The sound analysis was based on sonograms from 30 min of a subaquatic recording obtained on 5 January, 1995. The qualitative analysis resulted in 390 whistles classified in 3 subclasses (descent, up–down, and down–up frequency), 21 types, and 46 subtypes. The quantitative analysis was based on: number of harmonics; number of inflexions; beginning, final, maximum, and minimum frequencies; duration; and frequency modulation. Minimum, maximum, and average values, standard deviation, and coefficient of variation were calculated for each type and subtype of whistle. The *S. fluviatilis* whistles are extremely diverse (high coefficient of variation values for each subclass, type, and subtype) and frequent (38.5 whistles/min), seeming to have an important role in the social relations among these animals. Such whistle diversity and quantities could be explained by the composition of the dolphin group, i.e., it was exclusively formed by young and subadult animals.

4pABc4. Acoustic geographic variation in two populations of Delphinus delphis. Caitlyn Toropova (1264 Poplar Dr., Arcata, CA 95521)

Vocal patterns from two populations of common dolphin (*Delphinus delphis*) were compared. Field observations took place in the Channel Islands, CA and La Paz, Mexico. Whistles were recorded with a Sony HTI hydrophone (range=10–32 000 kHz) and schools were observed for group number and distance to the hydrophone by onboard observers. Sounds and spectrograms were analyzed for mean intensities in the (a) 15- to 20-kHz bandwidth, and (b) the ratio of the intensities from the 15- to 20-kHz band compared to the whole spectrum, 0–22 kHz. Significantly different ($p<0.001$) sound intensity levels were found between the La Paz and Channels Islands populations. When the 15- to 20-kHz band was examined, the values were 104.70±1.73 dB (mean ±S.E.) and 86.80 ±0.90 dB for the La Paz and Channel Islands, respectively. When the 15- to 20-kHz band was looked at as a function of the total energy in the 0 to 22-kHz band, significantly different energy levels were present as well (La Paz=76.2%, Channel Islands=67.9%). Therefore, the two populations were shown to be significantly different ($p=0.001$). This suggests that acoustic geographic variation is a useful tool for distinguishing populations within a species, and raises the possibility of external influences determining vocal patterns.

4pABc5. Vocalization as an indicator for disorders in mammals. Tobias Riede, Guenter Tembrock (Inst. fuer Biol., Abt. Verhaltensphysiologie, Humboldt-Univ., Invalidenstrasse 43, 10115 Berlin, Germany), Hanspeter Herzel (Humboldt-Univ., Berlin, Germany), and Leo Brunnberg (Freie Univ., Berlin, Germany)

As in humans, diseases in animals may lead to characteristic changes of the voice. In particular, multiparametric frequency, time analysis, as well as the amount of nonlinear phenomena (frequency jumps, subharmonics, biphonation, chaos) provide information about the well-being of an animal and the progress of a disease. Vocalization of domestic animals (dogs and cats) in a university hospital and of a macaque species in a zoological garden have been recorded. First results show that diseases lead typically to shortened call length and decreasing frequency modulation in the cat. Using narrow-band spectrogram subharmonics, biphonation, frequency jumps, and deterministic chaos are identified during the recuperation process of a cat with craniocerebellar trauma. The same nonlinear phenomena appear in several dogs with dysphonia following exhaustion of the voice and in a Japanese macaque infant with an unidentified systemic disorder. The ratio of nonlinear phenomena during systemic disorders of an animal shows a species-specific pattern. There was an increase mostly of frequency jumps in the macaque infant, of frequency jumps and chaos in the dog, and of chaos and subharmonics in the cat vocalization. [Work supported by NaFoeG, Humboldt-University Berlin.]

4pABc6. Variations in the vocal repertoire of the manatee in the Caribbean (*Trichechus manatus manatus*). Jose A. Alicca, James Harvey (Moss Landing Marine Labs., P.O. Box 450, Moss Landing, CA 95039-0450), and Antonio Mignucci-Giannoni (Caribbean Stranding Network, San Juan, Puerto Rico 00937)

While different life history parameters for the manatee have been well studied in Florida, little is known about these aspects of manatee biology in the Caribbean. As part of a multifaceted study in Antillean manatee life history, the geographical and individual variability in vocal repertoires of Antillean manatees from Puerto Rico and the Dominican Republic were examined and compared to the repertoire of the Florida manatee. Differences in structural sound characteristics (frequency, duration, and relative power of each harmonic), usage (number of call types), and the associated behavior were studied from recordings of wild and captive animals. Over 100 h of recording were obtained from six different natural localities in Puerto Rico and one natural locality in Florida. Captive manatees were recorded in the Dominican Republic (1), in Puerto Rico (2), and Florida (13), in addition to recording seven semicaptive individuals in Florida. The recording of underwater vocalizations has been suggested as a useful tool for the study of marine mammals distribution. In the turbid waters of rivers, bays, and lagoons where manatees congregate in the Caribbean, these acoustic data might be useful for estimating population distribution and habitat use.

4pABc7. Is the echolocation ability of bottlenose dolphins too good? William A. Friedl (Natl. Defense Ctr. of Excellence for Res. in Ocean Sci., 73-4460 20 Queen Kaahumanu Hwy., Ste. 111, Kailua-Kona, HI 96740, billf@ceros.org)

Odontocete echolocation studies reveal a system with both great plasticity and precision, but knowledge of the animal's day-to-day use of the process is limited. High-speed cine photographs of a trained bottlenose dolphin swimming underwater show that the dolphin's head and rostrum move vertically even when the dolphin swims at modest speed. Does this movement limit or otherwise affect the echolocation precision of freely-swimming dolphins? Results from kinematic studies, combined with echolocation system descriptions from acoustic research, were examined to show if head movements might limit acoustic abilities of swimming dolphins. So far, these speculative approaches are inconclusive. The general synthesis, using disparate data from acoustic and kinematic studies, indicates that head motion associated with modest swimming speed does not degrade echolocation ability. Data from coordinated studies that involve both echolocation and movement at realistic, natural speeds are lacking. Given the difficulty involved in working with naturally swimming dolphins, this data set is likely to remain unfilled, so further conclusions will also remain speculative.

4pABc8. A source–filter model of humpback whale (*Megaptera novaeangliae*) vocal sound production. Eduardo Mercado III (Kewalo Basin Marine Mammal Lab., Univ. of Hawaii, 1129 Ala Moana Blvd., Honolulu, HI 96814, mercado@hawaii.edu)

Acoustic waveforms produced by vocalizing humpback whales are well described by a source–filter model. This model characterizes vocalizations as the output of a time-varying signal generator (the source) passed through a time-varying resonator (the filter). For most humpback whale vocalizations, the source can be modeled as either a quasiperiodic impulse train, white noise, or a combination of these two signal types; the resonator can be modeled as an all-pole filter. The source–filter model is useful because it (a) provides a way to quantitatively characterize humpback whale vocalizations that can be automated and that allows for the ''resynthesis'' of sounds; (b) facilitates the application of advanced speech processing technologies to the analysis of humpback vocalizations; and (c) can potentially provide clues about the actual mechanisms humpback whales use to produce sounds. Currently, it is not known how humpback whales internally produce sounds, although structures in the larynx are thought to be involved. Source–filter-based analysis of humpback vocalizations revealed characteristic regions of frequency enhancement (formants) that may correspond to resonant frequencies of a physiological ''filter.'' Distribution of formants varied with pitch in a manner that appeared similar to pitch-related formant variation in human singers, suggesting that both species use analogous production mechanisms.

4pABc9. A sonobouy array for two-dimensional location of dolphin vocalizations. Eric S. Howarth and R. H. Defran (Cetacean Behavior Lab., Dept. of Psych., San Diego State Univ., 5500 Campanile Dr., San Diego, CA 92182)

A four-sonobouy array was deployed 0.25 km offshore in southern California waters. The array was developed and deployed to passively locate coastal bottlenosed dolphins (*Tursiops truncatus*). Sonobouy mooring sites were selected to overlap areas shown by the Cetacean Behavior Laboratory at San Diego State University to be highly utilized by these coastal dolphins. Arrival-time-difference measurements were used to provide two-dimensional locations of underwater biological sounds such as dolphin vocalizations. Sonobouys were separated by a range of between 367 and 1585 m and the accuracy of source location depended on the relative proximity to the nearest sonobouy. Global Positioning Satellite and theodolite data were used to verify sonobouy locations and, in turn, the acoustic arrival-time differences of on-board calibration noises. Bottlenosed dolphin whistles, including mimics of the calibration sound, were among the biologically identifiable sounds tracked by the array. Other

acoustic signals recorded but not located were snapping shrimp (*Californiensis dentripes*) and sources judged to be fish. Recorded sound levels varied over a considerable range, even on identical hydrophones, and were probably due to the signal transmission distance, shallow-water propagation effects, and diurnal fluctuations in the sources.

4pABc10. Selection of instrumentation for monitoring undersea animal sounds. Daniel R. Raichel (Dept. of Mech. Eng., City College of New York, Steinman Hall, New York, NY 10031)

In order to properly select instrumentation for bioacoustic studies, a list of measurable criteria must be drawn up, but even that may not suffice unless some genuine thought is given to the possibility of mapping additional criteria, such as frequencies beyond the range usually considered. In the planning and design of the experiment, some of the criteria include: frequency and dynamic ranges, the number of channels desired (not necessarily always available!), compatibility of components (as exemplified by the matching of acoustical and electrical impedances), setup of sensing/receiving transducer arrays, mode of recording, statistical methodologies, usable computer prowess, etc. Environment constitutes an extremely important factor from two aspects: (1) It should not be adversely affected by the use and operation of the equipment; and (2) the equipment should be able to withstand rigors of temperature, humidity, vibration, and shock. Provision should be made for instrument calibration at proper intervals to ensure the integrity of obtained data. Cost is another factor that can determine the reliability of the equipment, quality of performance, ruggedness, and ease of operation.

THURSDAY AFTERNOON, 4 DECEMBER 1997 FORUM ROOM, 2:45 TO 4:00 P.M.

Session 4pEA

Engineering Acoustics: Flow Noise, Aeroacoustic Noise

Sung-Hwan Ko, Chair
Naval Undersea Warfare Center, Code 2133, Building 1171/3F, Newport, Rhode Island 02841

Contributed Papers

2:45

4pEA1. Measurement and empirical prediction of bottom surface acoustic pressures on a hover aircraft model. Herbert L. Kuntz, Floyd O. Hickmon III, and Edward P. Feltz (Lockheed Martin Skunk Works, 1011 Lockheed Way, Palmdale, CA 93599-2522, hkuntz@lmco.com)

Acoustic measurements were made on the bottom surface of the 86% scale large-scale propulsion model (LSPM) of an advanced short takeoff and vertical landing (ASTOVL) airplane. The measurements were obtained during fixed hover heights where the lower wing surface varied from 5 ft (airplane's wheels on the ground) through 36 ft above the ground for a wide range of propulsion thrust mixes and airplane control surface configurations. The complete propulsion system provided four exhausts through an integrated unit consisting of an aft located main jet engine, which powered a forward located shaft driven lift fan and was completed with two main engine bleed high-pressure bypass roll jets (in the wings). Flush surface measurements were made at 12 locations on the bottom surfaces and time correlated with the operating propulsion parameters. An empirical acoustic model was made through analyses of the geometry and the data, and was used to predict contours on the bottom of the aircraft. [This work was performed by Lockheed Martin Skunk Works at the NASA-Ames Research Center.]

3:00

4pEA2. Approximate model for sound generation due to unsteady flows in pipes. Michael Krane, Daniel Sinder, and James Flanagan (CAIP Ctr., Rutgers Univ., Piscataway, NJ 08855)

An approximate model for the generation of sound due to unsteady vortical flow in pipes of varying cross-sectional area is described. This model, which is being used in speech synthesis research, is based upon Howe's acoustic analogy. This formulation allows the sound generation to be specified in terms of properties of the unsteady behavior of a vortical flow and the potential flow solution for the pipe that would exist in the absence of vortical inhomogeneities. An approximate model for vorticity formation and evolution is presented, and the potential flow solution is obtained from the axial duct shape. Computed estimates of the time evolution of the sound generated by a confined jet passing through a pipe constriction are presented. These results are compared to experimental data. [Research supported by NSF/ARPA IRI-9314946 and ARPA DAST 63-93-C-0064.]

3:15

4pEA3. Experimental study of cavity resonance suppression methods. J. Scott Johnson and Luc G. Mongeau (Dept. of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

Flow-excited cavity resonance is a common, generally undesirable, sound-generating phenomenon that occurs in many engineering applications. In this study, low-speed cavity resonance suppression techniques were investigated. The focus was on two methods: leading edge spoilers and leading edge air mass injection. Experiments were conducted in a closed section wind tunnel, which included a secondary blowing apparatus for controlling boundary layer thickness. The cavity, attached to the wind tunnel test section, acted as a side-branch resonator. Smoke flow visualization was performed for various operating conditions. For example, a nondimensional spoiler length of 0.156 (based on cavity length) inclined at a 50-deg angle of attack typically provided a 12 dB (re: 20 μPa) reduction of the 140-dB, 124-Hz tonal cavity pulsation. Air mass injection through a perforated array at a rate of 2.8 liter/s yielded a 27-dB reduction. In general, the results suggested that such leading edge suppression techniques yield greater orifice shear layer thicknesses, similar to that of increasing the boundary layer thickness. The thickened shear layer may cause greater vortex diffusion, a key parameter establishing the level of the cavity pressure fluctuations. Diffuse vortices lead to weaker cavity pressure oscillations and lower vortex convection velocities.

3:30

4pEA4. Turbulent flow noise estimate by use of a velocity sensor embedded in a three-layered composite structure. Sung H. Ko (Naval Undersea Warfare Ctr. Div., Newport, RI 02841)

A theoretical model was developed to evaluate flow noise levels received by a velocity sensor that is embedded within an outer decoupler mounted on a low acoustic impedance surface. The model, a three-layer structure, consists of a layer of nonvoided elastomer (outer decoupler), which covers a layer of microvoided elastomer (inner decoupler) backed by an elastic plate. The rubberlike material of the outer decoupler is designed to reduce the flow noise generated by turbulent boundary layer pressure fluctuations, and the low acoustic impedance (soft) material of the inner decoupler is configured to reduce the flexural wave noise generated by the vibration of the elastic plate. The upper surface of the outer decoupler is in contact with turbulent flow (water) and the lower surface of the backing plate is in contact with semi-infinite space (air). This work uses the model of the three-layer composite structure for the analyses of both the flow noise and signal levels received by the embedded velocity sensor. The final noise levels are the equivalent plane-wave levels, which were obtained by subtracting the signal levels from the flow noise levels. [Work supported by the Office of Naval Research, Code 321SS.]

3:45

4pEA5. A refined four-load method for the evaluation of in-duct acoustic source characteristics. Seung-Ho Jang and Jeong-Guon Ih (Dept. of Mech. Eng., Korea Adv. Inst. of Sci. and Technol., Science Town, Taejon 305-701, Korea)

By using the four-load method, one-port source characteristics can be determined in the duct system. In this study, a modification is made to avoid the instability problems occurring in the conventional four-load method. In the present method, the acoustic power spectra are measured at the inside or outside of the duct varying the acoustic loads, and an error function based on the linear time-invariant source model is obtained. It is numerically shown that the method is less sensitive to input errors compared to the conventional four-load method. The two methods are utilized to obtain the source parameters of a loudspeaker and a centrifugal fan using various types of loads. The estimated impedance values are compared with those obtained from the direct measurement method. As a direct measurement method, the multiple sensor method is adopted using three microphones. It is observed that the conventional four-load method results in large errors, whereas the present method yields far better agreement with the actual source characteristics.

THURSDAY AFTERNOON, 4 DECEMBER 1997 COUNCIL ROOM, 1:30 TO 5:30 P.M.

Session 4pMU

Musical Acoustics: Computer Jazz Improvisation

James W. Beauchamp, Chair
University of Illinois, Music and Electrical and Computer Engineering, 1114 West Nevada, Urbana, Illinois 61801

Chair's Introduction—1:30

Invited Papers

1:35

4pMU1. GenJam: An interactive genetic algorithm jazz improviser. John A. Biles (Information Technol. Dept., Rochester Inst. of Technol., 102 Lomb Memorial Dr., Rochester, NY 14623-5608, jab@it.rit.edu)

GenJam is an interactive genetic algorithm that learns jazz improvisation. It uses two hierarchically related populations to represent melodic ideas at the measure and phrase levels. These populations are evolved using tournament selection, single-point crossover, musically meaningful mutation, and replacement with a 50% generation gap. Fitness for the individual measures and phrases is derived from real-time feedback, which is provided by a human mentor while GenJam improvises to the accompaniment of a synthesized rhythm section. GenJam has been used for actual gigs under the billing Al Biles Virtual Quintet, which features the author on trumpet and GenJam on a variety of synthesized instruments, playing a repertoire of over 90 tunes in a variety of jazz, Latin, and new-age styles. Recent enhancements to GenJam include a pitch-to-MIDI capability that allows GenJam to listen to a human soloist, map his four-bar phrases to the GenJam genetic representation for phrases, apply selected mutation operators to these phrases, and play them back in real time as it trades fours or eights with a human soloist. In this way GenJam is truly interactive in performance, as well as during training. The lecture will feature a live demonstration of GenJam.

2:05

4pMU2. Aesthetic considerations for the composition and performance of computer-based jazz. Neil Leonard III (Music Synthesis Dept., Berklee College of Music, 1140 Boylson St., Boston, MA 02215, nleonard@it.berklee.edu)

Computers, critical doubts not withstanding, provide an important way to extend jazz improvisation and compositional resources. Jazz musicians have long explored the use of new technologies and have given a number of recent instruments their first virtuosos and repertoires. Just as jazz had given the saxophone its voice, through musicians like Coleman Hawkins, it also established the validity of the drum set in the work of Baby Dodds among others, and the electric guitar in the breakthroughs of Charlie Christian. I have sought to expand my work as a jazz improviser and composer through the use of interactive systems. For the past 12 years I have presented concerts that included original compositions for saxophone and computer controlled electronics. In the process I have explored several approaches to designing and performing with these systems, and addressed issues of how jazz is evolving. The lecture will include a demonstration of one such composition that uses an interactive system.

2:35

4pMU3. Statistical information for the development of computer models of jazz. Topi V. Jarvinen (Dept. of Musicology, Univ. of Jyvaskyla, P.O. Box 35, FIN-40351 Jyvaskyla, Finland, tjarvine@cc.jyu.fi)

Computer models of jazz improvisation have often been based on modelers' personal musical experiences or on anecdotes about jazz musicians' preferences. While this type of information is essential for any study of jazz, it would seem that it should be supplemented with some more concrete evidence on the underlying principles of jazz. This paper presents the results of a few recent statistical studies on bebop-styled jazz improvisation [e.g., T. Jarvinen, ''Tonal Dynamics and Metrical Structures in Jazz Improvisation,'' Dept. of Musicology, Univ. of Jyvaskyla, Finland]. In particular, it is discussed how bebop musicians employ meter, global tonality, and development of tonality in order to create coherent improvisations. It is argued that these types of results may help researchers design models with greater relevance and more explicit premises. An example of an application of these results is presented in another paper in this session by Petri Toiviainen.

3:05

4pMU4. Understanding and following a jazz improvisation in real time: A statistical approach. Petri H. Toiviainen (Dept. of Musicology, Univ. of Jyvaskyla, P.O. Box 35, FIN-40351, Jyvaskyla, Finland, ptoiviai@jyu.fi)

A system for following an improvised jazz solo and inferring the current location in the chord progression is presented. The system will ultimately work in real time in a MIDI environment: It listens to an improvised jazz melody played on a MIDI instrument and joins in the performance with a synthesized rhythm section. It is comprised of two stages: a beat tracker and a performance analyzer. The beat tracker is based on a continuously adapting nonlinear oscillator, which synchronizes to the beats of the improvised solo by continuously adapting its phase and period. It can follow local variations of tempo and produce expectations of next beats. The performance analyzer attempts to infer the current location in the harmonic progression by comparing the notes played in the improvised melody with statistical distributions of notes obtained from improvisations of well-known jazz musicians [T. Jarvinen, *Tonal Dynamics and Metrical Structures in Jazz Improvisation* (Dept. of Musicology, Univ. of Jyvaskyla, Finland) and trying to find the best match from several possible alternatives. [Work supported by the Academy of Finland.]

3:35

4pMU5. Representing groove: Rhythmic structure in interactive music performance. Vijay Iyer (Ctr. for New Music & Audio Technol., Univ. of California at Berkeley, 1750 Arch St., Berkeley, CA 94709, vijay@cnmat.berkeley.edu), Jeff Bilmes, Matt Wright, and David Wessel (Univ. of California at Berkeley, Berkeley, CA 94704)

The concept of ''groove,'' a musical element particularly operative in African-derived musics such as jazz, may be described as a collectively determined and relatively isochronous pulse that can be inferred from the interaction of a number of interlocking rhythmic groups. A powerful representation for musical rhythm as it appears in groove-oriented contexts has been developed. Implemented in the MAX music-programming environment, the representation encompasses pitch, accent, rhythmic deviations, tempo variation, note durations, and probabilistic processes. The design facilitates ''bottom-up'' combination techniques such as the construction of large musical objects by assembling small musical ''cells'' in series or parallel, and thus it allows improvisatory manipulation of simple rhythmic structures. The richness of control over many meaningful musical quantities distinguishes the representation from those in more common usage, such as music notation programs, sequencers, and drum machines. The implementation supports a variety of creative applications in improvised performance. Rather than attempting to make the system ''sound human'' by forcing it to simulate specific human activities, two alternative approaches are taken: the development of novel, sensitive control structures, and/or the emphasis on the system's active role as a situated participant in an interactive musical environment.

Contributed Paper

4:05

4pMU6. A novice's interface to programming digital synthesizers based on genetic algorithms. Frode E. Sandnes (Dept. Comput. Sci., Univ. of Reading, Reading RG6 6AY, England) and Craig Robertson (Univ. of Edinburgh, Edinburgh, Scotland)

The last decade has seen vast advances in commercial synthesizer developments, both in cost and features. Programming of synthesizers requires patience and detailed knowledge about sound synthesis. Also, the low-cost interfaces to these units are time consuming to access compared to the sliders of analogue synthesizers. Many musicians resort to costly off-the-shelf preprogrammed sound libraries, often without the actual characteristics and textures required. Therefore, a novel approach to programming of digital synthesizers is proposed. The task of programming is defined as an optimization problem, where the optimization function is the musician's rating of the sound, and the function arguments are the actual parameters of the sound synthesis algorithm, such as ADSR, LFO's, etc. These parameters are coded as chromosomes, and a genetic algorithm is used to search for the desired sounds. The interface works by presenting an initially random palette of sounds to the musician, for him/her to evaluate. Based on these evaluations, the genetic algorithm applies standard genetic operators to the chromosomes and over time breeds the desired sounds. Genetic algorithms ensure rapid convergence. This makes the technical details of a synthesizer transparent to the musician while it still allows him/her to explore all of its advanced features.

4:20–4:30 Break

4:30–5:30
Concert Demonstration

Session 4pNS

Noise and Engineering Acoustics: Computational Aeroacoustics

Luc Mongeau, Chair
School of Mechanical Engineering, Purdue University, 1077 Ray W. Herrick Laboratories, West Lafayette, Indiana 47907-1077

Chair's Introduction—1:00

Invited Papers

1:05

4pNS1. Recent simulations in computational aeroacoustics. Philip J. Morris (Dept. of Aerosp. Eng., Penn State Univ., University Park, PA 16802)

Computational aeroacoustics involves the direct simulation of noise production and radiation from unsteady flows. Its application to problems of practical interest has depended on both algorithm development and advances in computer technology. This paper presents some recent computational aeroacoustics calculations in several different areas. In each case parallel computations are performed to improve efficiency. A new methodology for the numerical simulation of high-speed jet noise is described. It is based on a separation of the instantaneous flow variables into time-averaged and disturbance components. The disturbances are obtained from a three-dimensional, time-dependent, compressible flow simulation. Results are presented for the effects of Mach number and temperature on jet noise. In the simulation of broadband noise from rotors and propellers, a key component is the noise generated by the interaction of a vortical gust with an airfoil. Simulations are described for such an interaction including studies of the effect of airfoil shape and loading. Finally, some examples of the simulation of acoustic scattering in both two and three dimensions from complex geometry bodies in nonuniform flow are described. A discussion of the choice of algorithms, including discretization, boundary treatments, artificial dissipation, and steady-state computations, is provided. [Work supported by NASA.]

1:35

4pNS2. A superior Kirchhoff method for aeroacoustic noise prediction: The Ffowcs Williams–Hawkings equation. Kenneth S. Brentner (NASA Langley Res. Ctr., Hampton, VA 23681, k.s.brentner@larc.nasa.gov)

The Lighthill acoustic analogy, as embodied in the Ffowcs Williams–Hawkings (FW–H) equation, is compared with the Kirchhoff formulation for moving surfaces. A comparison of the two governing equations reveals that the main Kirchhoff advantage (namely, nonlinear flow effects are included in the surface integration) is also available to the FW–H method if the integration surface used in the FW–H equation is not assumed impenetrable. The FW–H equation is analytically superior for aeroacoustics because it is based upon the conservation laws of fluid mechanics rather than the wave equation. Hence, the FW–H equation is valid even if the integration surface is in the nonlinear region. This is demonstrated numerically for helicopter rotor applications in the paper. The Kirchhoff approach can lead to substantial errors if the integration surface is not positioned in the linear region (i.e., if the input data are not a solution to the wave equation). These errors may be hard to identify in some cases.

Contributed Paper

2:05

4pNS3. Large-eddy simulation of compressible free jet turbulence applied to computation of exhaust mixing noise. David B. Schein (Northrop Grumman Corp., 9H11/GK, 8900 E. Washington Blvd., Pico Rivera, CA 90660 and Dept. of Mech. and Aerosp. Eng., UCLA) and William C. Meecham (UCLA, Los Angeles, CA, meecham@seas.ucla.edu)

A computational study of free, heated jet flow and resultant far-field sound was performed using large-eddy simulation (LES) and Lighthill's acoustic analogy. A subgrid scale model for small-scale compressible turbulence was developed using a combination of the popular Smagorinsky model and a deductive model. The primary objective is to address large Reynolds number (Re), high subsonic (compressible) flow with realistic geometries more representative of aircraft engine exhausts than typically considered using direct numerical simulation (DNS). Flow field fluctuations are stored over a period of time and used to calculate rms turbulence within the computational domain. The far-field sound and directivity is computed using the time-derivative form of Lighthill's source–integral result formulated in terms of quadrupole sources from the simulated flow field, which is integrated in time and contains the fluctuations set up by the time-varying stress tensor. A simulation for a WR19-4 turbofan engine exhaust ($\mathrm{Re} \approx 2 \times 10^6$ based on exit velocity and diameter) is presented, and propagated jet noise results are compared with experimental acoustics data. Methods to account for effects of source convection and thermal refraction to obtain realistic frequency spectra and directivity are considered.

Session 4pPA

Physical Acoustics: Bubbles, Drops and Particles

Robin O. Cleveland, Chair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

1:30

4pPA1. Mutual interaction of a two-phase turbulent/bubbly submerged water jet with a single-phase jet for the measurement of the enhanced hydrodynamic near-field pressure spectrum. Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

The interaction of mutually perpendicular submerged turbulent water jets is studied from the measurement of the enhanced hydrodynamic near-field pressure spectrum. One jet is a conventional single-phase, free turbulent shear flow circular jet, while the other jet is a two-phase bubbly jet of similar construction. The overlap region is located at four-nozzle diameters (4D) away from each orifice. A small hydrophone located outside the overlap region (4D away from the interaction region and 45 deg from the jet axis) measures the near-field pressure spectrum. Very small gas bubbles are required for this two-phase turbulent jet. They are generated by pressurized carbonated water flowing through a thin circular nozzle plate having an array of small holes. These CO_2 gas bubbles are mixed within the entrance to the tapered jet nozzle. Alternatively, bubbles are generated by pressurized N_2 gas passing through a fritted disk housed in a Buchner funnel located near the nozzle entrance. Spectra are measured and compared for different void fractions and different Reynolds numbers that are based on the jet diameter.

1:45

4pPA2. Acoustic microcavitation at surfaces: New results in deinking of paper. Sameer I. Madanshetty (Mech. and Nuclear Eng., Kansas State Univ., 339 Durland Hall, Manhattan, KS 66506-5106)

Deinking remains an important step in environmentally conscious manufacturing of paper. A novel method based on acoustic microcavitation is used for deinking xerographic ink from paper. Microcavitation evolves microimplosions, which are effective in causing deinking preferentially at ink sites. Acoustic microcavitation is brought about by low megahertz acoustic fields giving rise to micron-sized bubbles that live a few microseconds. In exposing a surface to continuous waves for a defined duration one could obtain cavitation effects in an average, overall sense; the details of nucleation, evolution of inertial events, and the precise interplay of field parameters in effecting cavitation, however, get glossed over. Studying pulsed cavitation using tone bursts at low duty cycles, instead of CW insonification, reveals interesting details of the initiation and evolution of acoustic microcavitation. Experiments indicate that the 2-D context of how a xerographic print is deinked might be useful in inferring some attributes of 3-D pulp deinking. Assuringly, the results indicate that paper fibers are entirely undamaged, and the ink separation leaves them immaculately white, and that acoustic methods for deinking are a viable, chemical free, environmentally responsible means of recycling paper.

2:00

4pPA3. Acoustic radiation force on micrometer-size particles: Size dependency and the efficiency of the superposition method. Kenji Yasuda (Adv. Res. Lab., Hitachi Ltd., 2520 Akanuma, Hatoyama, Saitama 350-03, Japan)

The acoustic radiation force on polystyrene spheres was measured through observation of the sphere movement in a 500-kHz ultrasonic standing wave. As predicted by Yosioka, the linear dependency of the force on the cube of the sphere radius began to fail when the sphere radius was below 5 μm. This failure can be accounted for by the presence of a shell layer that increased the effective radius of the sphere and it suggests that acoustic radiation might be used in handling microspheres smaller than previously thought [K. Yasuda and T. Kamakura, Appl. Phys. Lett. (in press)]. A method for concentrating particles using an acoustic radiation force by the superposition of the higher harmonics on fundamental ultrasound has also been investigated. The efficiency obtained when blood cells are concentrated by using the superposition method should theoretically be 20% better than that obtained when using the sine wave. Although the improvement obtained experimentally was only 5% because the waveform was incomplete, the potential usefulness of the superposition method has been demonstrated. [K. Yasuda, Jpn. J. Appl. Phys. **36**, 3130–3135 (1997)].

2:15

4pPA4. Novel configurations for acoustophoresis. Todd L. Brooks and Robert E. Apfel (Yale Univ., New Haven, CT 06520-8286)

Separation of particles by their acoustomechanical properties can be accomplished by acoustophoresis, which depends on both primary and secondary acoustic radiation forces. In an acoustic standing wave, particles are moved either toward the pressure nodes or the antinodes depending on the relative contrasts in the particle density and compressibility with the surrounding host liquid. Because particles reradiate the sound waves, there are secondary interparticle forces as well. These secondary forces can either hinder or enhance the separation process. If two species which are to be separated attract each other because of interparticle forces, the separation will be compromised. But if several like particles aggregate, then primary forces are much stronger and the separation occurs more rapidly. When flow is superimposed on the system, as is common in practical applications, then drag forces must also be accounted for. In the present work, novel field configurations have been explored, and the trajectories of particles have been computed for a number of frequencies, flow speeds, field strengths, and property contrasts. Based on these simulations, experiments have been designed to optimize the efficiency of acoustophoresis, especially for biotechnology applications. [Work supported by NASA through Grant NAG8-1351.]

2:30

4pPA5. Numerical simulation of superoscillations of a BSA-bearing drop in microgravity. Xiaohui Chen, Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., 9 Hillhouse Ave., New Haven, CT 06520-8286), and Tao Shi (Thomas Jefferson Univ., Philadelphia, PA 19107)

Large-amplitude nonlinear oscillations of a water drop, radius 2.53 cm and aspect ratio 2.10, with surfactant BSA (bovine serum albumin) 1.0×10^{-5} g/ml(1CMC), is numerically simulated based on the boundary integral method. BSA is a globular, large molecular-weight protein. Two surface viscosities, surface dilatational viscosity and surface shear viscosity, as well as the Marangoni effect, are considered. The high-surface viscosities make the rotational velocity the same order as the irrotational velocity in the boundary layer. A boundary-layer method is used to calculate the rotational velocity. When the surface dilatational viscosity is 0.5 sp and the surface shear viscosity is 0.15 sp, the numerical simulation results are in good agreement with the experiment results observed in the space shuttle during the second United States Microgravity Laboratory, USML-2, in October, 1995. The evolution of the drop oscillation for both experiment and simulation is given. The change of the BSA distribution along the surface, and the relative importance of the rotational and irrotational velocities, as the drop oscillates is also shown. [Work supported by NASA through JPL, Contract No. 958722.]

2:45–3:00 Break

3:00

4pPA6. Drop and bubble dynamics investigations on Earth and in low gravity using ultrasonics. Eugene H. Trinh (JPL/Caltech, MS 183-401, 4800 Oak Grove Dr., Pasadena, CA 91109)

Two ultrasonic devices for the positioning and the remote manipulation of free drops and bubbles have been used during the recent STS-94 Spacelab flight. One apparatus is designed to levitate or to position single or a small number of individual droplets in air, to induce drop shape oscillations, and to control the residual drop rotation. Implementation in low gravity has allowed the measurement of droplet dynamics and rotational stability in the absence of the overwhelming constraint imposed by the high-intensity ultrasonic field required for levitation on Earth. The second apparatus permits the trapping or positioning of gas bubbles in a water-filled resonant cell with a square cross section. Confirming earlier results from a previous flight experiment, it was found that gravity plays a determining part in the stable centering of bubbles larger than resonant size. Both devices operate at about 22.5 kHz and allow the monitoring of the drop or bubble motion through the detection of scattered light from a collimated diode laser beam illuminating the fluid particles. These investigations are low cost, they are built from commercially available components, and they are manually operated by crew members in the Middeck/Spacelab Glovebox facility. [Work funded by NASA.]

3:15

4pPA7. Single-bubble sonoluminescence in microgravity. Thomas J. Matula, Jarred E. Swalwell, Vassiloius Bezzerides, Paul Hilmo, Mike Chittick, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), David W. Kuhns (Univ. of Washington, Seattle, WA 98105), and Ronald A. Roy (Boston Univ., Boston, MA 02215)

A recent experiment involving single-bubble sonoluminescence aboard NASA's parabolic research aircraft will be described. Measurements of the intensity of the light emission were performed during periods of microgravity (near 0 g) and hypergravity (near 2 g). Gravitational effects on the luminescence and extinction pressure thresholds were examined. In addition, the light emission was monitored under constant drive conditions as the gravitational acceleration varied during the parabolas. Measurements show some variability in the sonoluminescence intensity for both thresholds. However, during otherwise constant conditions in which the drive pressure amplitude did not change, the intensity from stable single-bubble sonoluminescence was observed to increase during periods of microgravity. There was an initial increase in intensity that occurred simultaneous with the decrease in the gravitational acceleration, followed by a slow increase in intensity that appeared to level off near the end of the microgravity period. The longer time scales over which the intensity changed during the microgravity period may be indicative of gas diffusion occurring as the bubble attains a new equilibrium condition. [Research supported by NSF and NASA.]

3:30

4pPA8. Measurements of the transient response of single-bubble sonoluminescence subject to an abrupt change in the drive pressure amplitude. Thomas J. Matula and Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

A novel method utilizing abrupt drive pressure amplitude changes is described for exploring the dynamics of single-bubble sonoluminescence. Air bubbles that are initially below the luminescence threshold are subject to an abrupt increase in the drive pressure amplitude. The resulting turn-on time for light emission is qualitatively similar to nitrogen bubbles, except that over time scales of seconds, the light intensity of an air bubble increases [B. P. Barber *et al.*, ''Sensitivity of sonoluminescence to experimental parameters,'' Phys. Rev. Lett. **72**, 1380 (1994)], while the intensity from a nitrogen bubble remains low. However, when we begin with a stable sonoluminescing air bubble *above* the luminescence threshold, and then generate an abrupt decrease in the pressure amplitude below the threshold (no light regime), followed again by an abrupt increase back to its original (sonoluminescing) state, the resulting turn-on time for light emission appears qualitatively similar to noble gas bubbles: sudden, and relatively intense, unlike what occurs for air bubbles that were initially below the luminescence threshold. These measurements imply that sonoluminescence from an air bubble depends on the time the bubble spends in the sonoluminescing state. [Research supported by NSF.]

3:45

4pPA9. Could optical radiation pressure be used to move a sonoluminescing bubble to the pressure antinode of an acoustic standing wave? David B. Thiessen and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

It is generally recognized that in SBSL, the bubble is displaced slightly above the antinode of the acoustic standing wave which drives the bubble oscillations. This is because the acoustic radiation force must balance the average buoyancy of the bubble. For scientific purposes and to simulate SBSL in microgravity, it would be desirable to reposition the bubble at the antinode. One strategy to achieve that repositioning where the buoyancy force would be balanced entirely by the average optical radiation pressure of a downward propagating laser beam has been numerically investigated. Optical levitation of small bubbles has been previously demonstrated [B. T. Unger and P. L. Marston, J. Acoust. Soc. Am. **83**, 970–975 (1988)]; however, it is not trivially applied to SBSL; with cw illumination of a quiescent bubble, the laser power required increases as the cube of the bubble radius. To simulate the large-radius phase of SBSL bubble oscillations, a modified Rayleigh–Plesset equation was numerically integrated. The required power for cw and pulsed laser beams was estimated. The average power is significantly reduced if the beam is pulsed to coincide with the large-radius phase of the oscillations so as to maximize the relevant optical cross section.

4:00

4pPA10. Numerical research of the influence of external conditions on the sonoluminescence process. Vladimir N. Nogin, Nikolay G. Karlykhanov, Gennady V. Kovalenko, Vadim A. Simonenko (RFNC-VNIITF, Snezhinsk, Russia), and William C. Moss (Lawrence Livermore Natl. Lab., Livermore, CA)

Using a 1-D approximation, processes are numerically simulated that take place during oscillations of a single gas bubble located in the center of a spherical flask with liquid exposed to a periodical acoustic wave. Real equations of state for water, gases (air, argon), surface tension, and heat conduction (molecular and radiative) are taken into account in the simulation. The role of shock-wave processes at bubble compression is investigated. Because of the peculiarities in the behavior of the gas, the equation of state is shown to be the most important at densities greater than 1 g/cm^3, with the parameters of bubble compression being weakly affected by the water equation of state. Time dependencies of the bubble radius during one acoustic cycle are obtained for different ambient conditions: acoustic-wave amplitude and frequency, liquid temperature, and magnitude of surface tension. Comparison with experimental data is performed. [Work supported by CRDF.]

4:15

4pPA11. Scaling analysis for single-bubble cavitation. Vadim A. Simonenko, Nikolay G. Karlykhanov, Gennady V. Kovalenko, Vladimir N. Nogin (RFNC-VNIITF, Snezhinsk, Chelyabinsk Region, Russia), and William C. Moss (Lawrence Livermore Natl. Lab., Livermore, CA)

There is a new interest to study caviation phenomena for gas-filled bubble in liquid basing on recent experimental and theoretical results on single-bubble sonoluminescence. Scaling analysis is applied to these phenomena. It allows a role of main control parameters to be clarified which have an influence on energy transfer to central gas-filled cavity and on energy dissipation in gas. The main parameters are linear dimensions, initial equilibrium pressure, initial temperature, rate and amplitude of pressure change. Of principle importance are convergent symmetry parameters and time-shape for external pressure. The most important are phenomena which provide maximum energy transfer to a central bubble, and to internal gas layers in the bubble. Factors decreasing efficiency of energy transfer are shown. The main factors are (1) compressibility of liquid caused by additional energy dissipation in layers adjacent to bubble; (2) early origin of shock wave near gas–liquid boundary. Theoretical estimations are confirmed by simulations calibrated on single bubble sonoluminescence. Criteria are shown to optimize gas compression in bubble. There are proposed prospective experiments to study large-scale effects in systems with external size of few meters and equilibrium pressure of several kilobars. [Work supported by the CRDF.]

THURSDAY AFTERNOON, 4 DECEMBER 1997 CHAMBER ROOM, 2:00 TO 4:45 P.M.

Session 4pPP

Psychological and Physiological Acoustics: Psychophysical Beachcombing

Kourosh Saberi, Chair
Division of Biology, Caltech, Pasadena, California 91125

Contributed Papers

2:00

4pPP1. First and second pitch shift effects as alternative manifestations of a single phenomenon. Pantelis N. Vassilakis (MA Ethnomusicol., Dept. of Ethnomusicol., Program in Systematic Musicol., 2539 Schoenberg Hall, UCLA, Los Angeles, CA 90095, pantelis@ucla.edu)

The second pitch shift effect describes a drop in the perceived pitch of a complex stimulus when the frequency spacing among its components is increased. Based on existing experimental data describing the first pitch shift effect, the present paper demonstrates that the mathematical model introduced by de Boer [Doctoral dissertation, University of Amsterdam (1956)] and modified later by Smoorenburg [J. Acoust. Soc. Am. **48**, 1055–1060 (1970)], predicts without any further modification the second pitch shift effect as well. The Smoorenburg model also predicts that the perceived pitch of complex stimuli will not always drop when increasing the frequency spacing among components, but may rise depending on the structure of the stimuli. A perceptual experiment was conducted using nine complex stimuli. For eight of the stimuli, the model predicted a rise in pitch with increasing frequency spacing while for the ninth stimulus the opposite was predicted. The results of the experiment support the conclusion that the relationship between the direction of pitch motion (rise/drop) and the direction of changes in frequency spacing among components of complex stimuli (increasing/decreasing) can be predicted by the same model that explains the first pitch shift effect, making the second pitch shift effect an unnecessary concept.

2:15

4pPP2. Psychometric functions for two-transient dichotic stimuli. Agavni Petrosyan (Dept. of Psych., California State Univ., Los Angeles, CA 90032) and Kourosh Saberi (Caltech, Pasadena, CA 91125)

To better understand the parameters that limit interaural-delay sensitivity in a precedence-effect task, psychometric functions were measured for the detection of an interaural time difference (ITD) in the second transient of a two-transient binaural stimulus. The stimuli were broadband 1-ms Gaussian pulses centered on the cosine phase of a 6-kHz carrier. Twenty-point psychometric functions were obtained using a 2IFC method of constant stimuli, with each point based on a minimum of 500 trials. Four functions were obtained for each observer at interclick intervals (ICI) of 3, 6, and 12 ms, in addition to a single-transient control condition. Proportion of correct responses were transformed to d′ units and linearly fitted with $\ln(d') = \ln\beta + \upsilon \ln(\Delta \mathrm{ITD})$, where $\ln\beta$ and υ are the parameters of the fit. The slope parameter υ was independent of ICI and remained near unity, discounting nonlinear transformations of the stimulus scale [J. P. Egan, J. Acoust. Soc. Am. **38**, 1043–1049 (1965)]. However, $\ln\beta$ decreased with decreasing ICI, suggesting linear changes in either the variance or expected value of the perceived lateral position. Further experiments examined the contribution of each of these latter factors to changes in $\ln\beta$. Implications for causes that underlie the precedence effect are discussed.

2:30

4pPP3. The effects of nonconcurrent tonal and noise distractors upon auditory spatial resolution: The temporal, spatial, and spectral parameters underlying the ''disruption effect.'' Eric A. Erpenbeck, David R. Perrott, and Raymond Lim (Dept. of Psych., California State Univ., 5154 State University Dr., Los Angeles, CA 90032)

The minimum audible angle (MAA) thresholds obtained by Mills [A. W. Mills, J. Acoust. Soc. Am. **30**, 237–246 (1958)] involved two 1000-ms sinusoidal tone pulses presented sequentially with an interstimulus interval (ISI) of 1000 ms. In contrast to the generally good localization performance obtained by Mills, localization of two simultaneous events (concurrent localization or CMAA) results in relatively poor performance [D. R. Perrott, J. Acoust. Soc. Am. **76**, 1704–1712 (1984)]. The current set of experiments attempts to bridge the gap between these two extreme paradigms. While the target tones (14-ms, 1000-Hz sinusoids) were presented sequentially (an MAA task), an attempt was made to systematically examine the consequences of the addition of a nonoverlapping, auditory

event (tones or noise) during the interval between the target events. The results of this work indicate that localization performance is extremely sensitive to events that occur between the target events. The disruptive effect of a ''nontarget'' event during the ISI could even be observed when relatively long ISIs (814-ms) and very brief (14-ms) distractors events were used. Interesting interactions were encountered as a function of the spatial positioning of the nontarget (''distractor'') event. Implications of these results will be discussed.

2:45

4pPP4. Individual subjective preference of simulated sound fields by listeners for opera sound sources in relation to the subsequent reverberation time. Hiroyuki Sakai (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan), Hiroshi Setoguchi (Miyama Conceru, Aira, Kagoshima, 899-66 Japan), and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

The purpose of this study is to evaluate the subjective preference of a simulated sound field by listeners in changing subsequent reverberation time T_{sub} using vocal sources. A great deal of effort has been made in the study using music or speech [for example: Ando *et al.*, Acustica **50**, 134–141 (1982); Ando *et al.*, J. Acoust. Soc. Jpn. **39**, 89–95 (1983), in Japanese]. Subjective evaluation using vocal sources has never been examined, although vocal performance is often played in concert halls and opera houses. Subjective preference tests were conducted changing T_{sub}, which is one of the four objective parameters in relation to subjective preference of sound field. Individual differences of subjective preference as well as global preference are also considered.

4pPP5. Abstract withdrawn

3:00

4pPP6. Effects of changes of the spectral masking slope on sound quality and clarity of music sounds in the normal and impaired ear. Bernhard Laback (Acoust. Res. Lab., Austrian Acad. of Sci., Liebigg. 5, A-1010 Wien, Austria), Niek Versfeld (Free-Univ.-Hospital, 1007 MC Amsterdam, The Netherlands), and Werner Deutsch (Austrian Acad. of Sci., Liebigg. 5, A-1010 Wien, Austria)

Reduced frequency selectivity as occurs in the cochlear impaired ear has been shown to be a main factor for degraded speech perception [J. M. Festen and R. Plomp, J. Acoust. Soc. Am. **76**, 652–662 (1983)]. It is also assumed to have negative effects on music perception. In this paper the hypothesis is tested that the loudness relations between spectral peaks and nonmasked lower-level components, which depend on the individual frequency selectivity, are important for (a) subjective sound quality (listening comfort) and (b) clarity, as measured in terms of detectability of altered notes in music excerpts. In order to test this hypothesis a signal processing algorithm has been utilized which enhances or suppresses the lower-level spectral components of a complex signal according to a masking function. Processed signals have been presented to normal-hearing and cochlear-impaired subjects. Results revealed that subjects with reduced frequency selectivity—as measured with psychoacoustical tuning curves—tend to prefer music signals with enhanced lower-level components. In the clarity experiment, results obtained with hearing-impaired subjects show improved detectability with suppression of the lower-level components (increase of spectral contrasts). [Work supported by Austrian Academy of Sciences.]

3:15–3:30 Break

3:30

4pPP7. Evaluation of simulations of the effects of hearing loss produced by combinations of dynamic expansion and spectral smearing. Isaac J. Graf, David S. Lum, and Louis D. Braida (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

By processing sounds with combinations of multiband expansion (MBE) and spectral smearing (SS), listeners with normal hearing can experience the elevated thresholds, abnormal growth of loudness, and reduced frequency selectivity similar to those of listeners with sensorineural hearing impairments. The effects of simulations of flat 40-dB losses produced by such combinations on masking patterns and loudness summation, and of simulations of a more severe loss on speech perception in noise, were studied. Both MBE and SS reduce frequency selectivity as measured by masking patterns for tones produced by narrow-band noise, and both broaden forward-masked tuning curves for tones. However, for equally loud probe tones, MBE elevates the tips of psychoacoustic tuning curves much more than SS. As measured by loudness balances between narrow-band and wideband noises, MBE reduces loudness summation much more than SS. The effects of MBE and SS simulations of a severe hearing loss on consonant identification in speech-shaped noise depend on the SNR, with MBE more detrimental at high SNR and SS more detrimental at low SNR. MBE tends to degrade reception of cues for voicing; SS, cues for place of articulation; both degrade cues for manner of articulation. [Work supported by NIH.]

3:45

4pPP8. A psychophysically based model of consonant perception by multichannel cochlear implant users. Ted A. Meyer, Mario A. Svirsky, Michael B. Castor, Sassan Falsafi (Dept. of Otolaryngol., Indiana Univ. School of Medicine, 702 Barnhill Dr., Indianapolis, IN 46202), and Peter M. Simmons (Duke Univ., Durham, NC)

Despite advances in implant technology, cochlear implant (CI) users demonstrate a wide range in the ability to perceive speech in the absence of visual or contextual cues. Some progress has been made recently in the ability to explain perceptual performance with a CI. A quantitative, psychophysically based model of vowel perception by CI users of the SPEAK processing strategy generated confusion matrices that were remarkably similar to actual data [M. A. Svirsky and T. A. Meyer, Assoc. Res. Otolaryngol. Abs. **20**, 59 (1997)]. The three dimensions of the parameter space are the centers of gravity of stimulation in the cochlea in response to the first three vowel formants. The free parameter of the model is the subject's ability to scale pitch percepts associated with different electrodes. Although this single-parameter model was not successful in predicting consonant confusions, when the model was expanded to incorporate psychophysical estimates of gap detection and high- versus low-frequency intensity difference discrimination, the model successfully predicted performance on a medial consonant test [Meyer *et al.*, Conf. Implant. Aud. Prosthes. Abs. (1997)]. Relations between predicted and obtained performance on selected psychophysical tests will be discussed. [Work supported by NIH.]

4:00

4pPP9. Discrimination of single-formant stimuli by chinchillas (*Chinchilla villidera*). Lori L. Holt, Keith R. Kluender (Dept. of Psych., Univ. of Wisconsin, 1202 W. Johnson St., Madison, WI 53706), and Andrew J. Lotto (Loyola Univ., Chicago, IL 60625)

The present study assessed the behavioral capacity of six chinchillas (*Chinchilla villidera*) to discriminate complex acoustic differences, like those observed for the first formant ($F1$) of human speech, using operant methods. In a minimum-uncertainty task, chinchillas discriminated single-formant stimuli with fundamental frequencies of 100 or 200 Hz, 60 Hz ($Q6$) bandwidths, and center frequencies ranging from 2.5–7.0 Bark (approximately 253–709 Hz) in 0.25-Bark steps. Each trial began with the repeated presentation of a ''target'' stimulus. The chinchillas' task was to discriminate whether this repeating target changed in center frequency or remained constant throughout the trial. The chinchillas completed 20 experimental sessions for each of six conditions in which fundamental frequency and center frequency of the target stimulus were varied. Stimulus characteristics were designed to reveal perceptual consequences of interactions between individual harmonics and overall spectral shape of the single formant. Behavioral sensitivities were estimated by d′ and percent correct. Data will be discussed in the context of human psychophysical performance and observations of single-unit responses in the chinchilla cochlear nucleus from studies utilizing the same stimulus corpus. [Work supported by NSF.]

4:15

4pPP10. Auditory processing is related to reading ability. Athanassios Protopapas, Merav Ahissar, and Michael M. Merzenich (Sci. Learning Corp., 417 Montgomery St., Ste. 500, San Francisco, CA 94104)

Aspects of auditory processing related to temporal and spectral resolution were investigated in 50 adult subjects (aged 18-58) with varying reading abilities. The tests administered included tone detection, tone frequency discrimination, tone sequencing, interval discrimination, and gap detection. Tones were either long (250 ms) or short (20 ms) and ranged in frequency between 600 and 1400 Hz. Detection and frequency discrimination tasks were given in the clean and in a masking context, in which tones were followed by 300-ms bandpass noise. Thresholds were determined using an adaptive procedure. Backward masking interferences for short tones, as well as interval and frequency discrimination limens were substantially elevated for most poor readers, and were correlated with one another. Three-tone sequencing and short-tone frequency discrimination in a masking context were especially strongly correlated with single-word reading ability. Gap detection thresholds did not correlate with other tasks or with reading ability. In sum, spectral-temporal auditory processing resolution appears to be related to reading ability, possibly via an acoustically based deficit in phonetic development. These findings raise the question of whether nonspeech acoustic training can be used to improve adults' reading ability. [Work supported by Scientific Learning Corporation.]

4:30

4pPP11. On the apparent source width for music sources related to the IACC and the width of the interaural crosscorrelation function (W_{IACC}). Shin-ichi Sato and Yoichi Ando (Grad. School of Sci. and Technol., Kobe Univ., Rokkodai, Nada, Kobe, 657 Japan)

It has been shown that the ASW for 1/3 octave band noise can be described as functions of the IACC and the W_{IACC}, which is defined by the time interval between values of 10% below the IACC [S. Sato and Y. Ando, J. Acoust. Soc. Am. **100**, 2592 (1996)]. In this study, it is examined whether the ASW for music sources also can be calculated with the IACC and the W_{IACC}. In order to control of both the IACC and the W_{IACC}, the delay time of two reflections relative to the direct sound was changed in the range of 0–50 ms. Results of the scale value of ASW obtained by paired comparison are compared with calculated ASW.

THURSDAY AFTERNOON, 4 DECEMBER 1997 PRESIDIO ROOM, 1:00 TO 4:45 P.M.

Session 4pSC

Speech Communication: Phonetic Perception and Word Recognition

Lynne E. Bernstein, Cochair
Spoken Language Processing Laboratory, House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

Edward T. Auer, Jr., Cochair
Spoken Language Processing Laboratory, House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

Invited Papers

1:00

4pSC1. Chairs' introduction to special session on phonetic perception and word recognition. Edward T. Auer, Jr. and Lynne E. Bernstein (Spoken Lang. Processes Lab., House Ear Inst., 2100 West Third St., Los Angeles, CA 90057)

A classic statement of the speech perception problem is how to account for the perception of phonemes based on the acoustic stimulus. A classic statement of the word recognition problem is how to account for the access and selection of lexical representations, with little concern for phoneme/phonetic perception. Over the past two decades, research in spoken-word recognition has begun to explicitly model the stages of word recognition in terms of phonetic/phonemic perception and later stages involving lexical knowledge. These different models incorporate very specific proposals for the interaction of information from the bottom-up stimulus and the lexicon. They also raise issues concerning the appropriate model for phoneme perception, that is, for example, does it precede or follow the lexical processing stage? It would be convenient if theories/models of phoneme perception could ignore effects of the lexicon, and therefore, autonomously account for the ''front-end'' of word recognition. But can they, and be adequate theories/models of either phoneme perception or the front-end to word recognition? Alternatively, are there phoneme identification effects that are irrelevant to word recognition? In this session, researchers representing a variety of perspectives on the true relationship between lexical processes and phoneme/phonetic perception are brought together.

1:10

4pSC2. The beginnings of word recognition in infancy. Peter W. Jusczyk (Dept. of Psych. and Dept. of Cognit. Sci., Johns Hopkins Univ., Baltimore, MD 21218, jusczyk@jhu.edu)

At around 7.5 months, English-learning infants begin to show some ability to segment words from fluent speech. Moreover, there is evidence that infants at this age begin to engage in some long-term encoding and retention of information about the sound patterns of words that occur frequently in fluent speech. Infants' memory for these sound patterns apparently does not depend on their pairing with any concrete referents. Thus, at least in some instances, word learning may occur by first storing information about sound patterns, and then subsequently attaching a meaning to these. Additional findings from our laboratory suggest that these early representations of the sound patterns of words are rather detailed with respect to phonetic properties and may include information about such indexical properties as talker identity. [Work supported by NICHD and NIMH.]

1:35

4pSC3. Spoken-word recognition. Cynthia M. Connine (Dept. of Psych., Univ. of Binghamton, SUNY, Binghamton, NY 13902-6000)

Spoken-word recognition is an efficient and generally error-free process that occurs under a variety of speaking and listening conditions. The talk will focus on the mapping process between the speech signal and access of form and meaning. The nature of the representation that supports spoken-word recognition will be discussed with a focus on the consequence of ambiguity and mismatching information. Research has been conducted in the past few years suggesting that activation of lexical representations is accomplished via feature mapping. It is argued that this architecture permits lexical activation given incomplete or erroneous input. Phonological variation and some recent work concerning representation and processing of common variants will also be discussed.

2:00

4pSC4. A comparison of perceptual word similarity metrics. Paul Iverson, Edward T. Auer, Jr., and Lynne E. Bernstein (Spoken Lang. Processes Lab., House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

Contemporary theories of spoken-word recognition rely on the notion that the stimulus word is mapped against, or selected from, words in long-term memory in terms of its phonetic (form-based) attributes. A few metrics have been proposed to model the form-based similarity of words, including an abstract phonemic metric computed directly on the lexicon (i.e., Coltheart-n), and perceptual metrics derived from the results of phoneme identification experiments. The results of applying several different metrics to phoneme and word identification data (open-set and forced-choice tasks) will be discussed, and these metrics across stimulus conditions with a range of intelligibility levels and similarity structures (visual-only lipreading, audio-only conditions processed by a vocoder, and audiovisual conditions pairing vocoded audio with lipreading) will be compared. The our results suggest that graded perceptual metrics may be most useful for understanding the results of word identification experiments across a wide range of stimulus conditions. [Work supported by NIH DC00695.]

2:25–2:40 Break

2:40

4pSC5. The relation between phoneme reception and word perception. Guido F. Smoorenburg (Hear. Res. Lab., Dept. of Otolaryngol., Univ. Hospital Utrecht F.02.504, P.O. Box 85500, 3500 GA Utrecht, The Netherlands)

Word perception can be modeled as a two-stage speech recognition process involving, first, recognition of individual phonemes and, second, substitution of missing elements using contextual information. Within this conceptual framework, contextual information may be acoustic in nature, coarticulation, and it may be related to lexical context. Using nonsense CVC syllables, we have analyzed the effects of coarticulation; using meaningful CVC syllables, the additional effects of lexical context. Supplementary estimates of the effects of lexical context were obtained from an incomplete orthographic presentation of the meaningful CVC syllables. The acoustic materials were presented to normal-hearing subjects and to hearing-impaired subjects. The results show large differences in recognition probability of the individual phonemes at low speech-to-noise ratios and marked contextual effects, attributed to coarticulation, for nonsense CVC syllables presented in noise. In the Dutch language, meaningful CVC syllables appear to consist of 2.5 statistically independent elements (rather than 3), which is attributed to the lexical context effects. Near threshold, response behavior of the hearing impaired is heavily based on lexical information, which lowers this number of statistically independent elements to a value close to 1. [This research was performed in collaboration with Dr. A. J. Bosman and Dr. A. W. Bronkhorst.]

3:05

4pSC6. Lexical neighborhood and phoneme frequency effects in phonetic processing. James R. Sawusch (Psych. Dept., Park Hall, Univ. of Buffalo, SUNY, Buffalo, NY 14260, jsawusch@acsu.buffalo.edu)

The phonetic coding of speech reflects both a stimulus driven process and influences from the mental lexicon. There are at least three distinct forms of influence of the mental lexicon on phonetic processing: lexical status (is the item a word), lexical neighborhood (how many words is the item similar to), and phoneme frequency (how often does the phoneme occur in the language). Using natural speech series, the identification of an ambiguous phoneme is influenced by both the lexical neighborhood of the stimulus (syllable), and the probability of occurrence of the target phonemes in the language. Listeners consistently identify ambiguous phonemes with

the label that corresponds to the end of the series with a larger lexical neighborhood, and listeners are biased to report more probable phonemes. Neighborhood and phoneme frequency effects are also found in a lexical decision task for clear, unambiguous tokens. Together with other work on higher-level influences on phoneme perception, these results indicate that phoneme perception is the result of an interaction between acoustic–phonetic information and the lexical knowledge of the individual, regardless of whether the target item is a word or not. [Work supported by NIDCD Grant No. R01 DC00219 to SUNY at Buffalo.]

3:30

4pSC7. Transitional probability, not lexical knowledge, influences compensation. James M. McQueen (Max-Planck-Inst. for Psycholinguist., Wundtlaan 1, 6525 XD Nijmegen, The Netherlands) and Mark A. Pitt (Ohio State Univ., Columbus, OH 43210)

The perceptual system compensates for fricative–stop coarticulation. Ambiguous stops between /t/ and /k/ tend to be heard as /k/ after /s/-final words and as /t/ after /S/-final words, even when the word–final fricatives are replaced with an ambiguous phoneme [J. L. Elman and J. L. McClelland, J. Mem. Lang. **27**, 143–165 (1988)]. Biases in the transitional probabilities (TPs) of /s/ and /S/ in the words in that study mean that the effect following ambiguous fricatives could be due either to lexical involvement in prelexical compensation for coarticulation or to a sensitivity to TP at the prelexical level. In three experiments, listeners categorized nonword–final fricatives with TP biases, word–final fricatives where TPs were controlled, and word–initial stops following both fricative contexts. Categorization of ambiguous fricatives and stops was influenced by the TPs of the fricatives in the nonwords. But although there was a lexical bias in the categorization of the word–final ambiguous fricatives, there was no lexical influence on the categorization of the stops. These results suggest that the lexicon does not influence prelexical processing, and that the prelexical level is sensitive to TP. [Work supported by NIDCD, the Human Frontiers Science Project, and the Max Planck Society.]

3:55

4pSC8. Predicting the perception of words from the perception of phonemes. Terrance M. Nearey (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada)

The present paper explores the degree to which the perception of CVC syllables can be factored into the perception of its constituent phonemes. Categorization experiments with synthetic speech typically manifest such factorability to a remarkable degree (regardless of lexical status of the syllables.) Such factorability is also compatible with intelligibility studies of English CVCs [e.g., A. Boothroyd and S. Nittrouer, J. Acoust. Soc. Am. **84**, 101–114 (1988)]. These show that the identification of nonsense syllables can be predicted as the product of the probabilities of identification of its phoneme parts, while word identification is systematically higher. Simulation studies are reported here involving a ''factored'' perceptual model that consists of a set of phoneme-likelihood estimators whose outputs are modulated by prior probabilities related to lexical status. This model can approximate the patterns observed in human perception. Simulations were also run with nonfactorable models, where syllables and words involve information about unique stimulus properties that cannot be predicted from their constituent phonemes. Consistent with a conjecture of Allen [J. Allen, IEEE Trans. Speech Audio Process. **2**, 567–577 (1994)], such syllable template models do not produce behavior compatible with human performance. [Work supported by SSHRC.]

4:20–4:45

Discussion

THURSDAY AFTERNOON, 4 DECEMBER 1997 FRIARS/PADRE ROOM, 1:30 TO 3:15 P.M.

Session 4pSP

Signal Processing in Acoustics: Sound Quality

Patricia Davies, Chair
School of Mechanical Engineering, Purdue University, Ray W. Herrick Laboratories, West Lafayette, Indiana 47907

Invited Papers

1:30

4pSP1. Artificial neural network processing of time-varying loudness data to model annoyance of noise. Peter C. Laux (Prince Corp., One Prince Ctr., Holland, MI 49423) and Patricia Davies (Purdue Univ., West Lafayette, IN 47907)

In this talk, a procedure is presented for the development of an artificial neural network model of human annoyance to noise. This model development is based on using available sound quality metrics (loudness, sharpness, fluctuation strength, and roughness) that were calculated using post-process temporal filtering and numerical formulation operations on measured time-varying critical band loudness data obtained from a Bruel and Kjaer loudness analyzer. The artificial neural network (ANN) was initially developed to model Zwicker's equation for unbiased annoyance (UBA). Subsequent training of the artificial neural network was done by using a series of 700 noise signals that were evaluated for their annoyance by a pool of normal-hearing young adult U.S. born subjects in a parameter estimation subjective test procedure. The noise stimuli were varied in seven different properties, which included: peak loudness, modulation frequency, modulation envelope shape, modulation depth, frequency band of modulated noise, addition of continuous complex tone structures, and the relative level of complex tone set to the underlying noise. The response data were averaged across subjects and the UBA-ANN model was retrained using the average subjective responses. The results of this new ANN–annoyance model are compared to other metrics for sound measurement/rating.

2:00

4pSP2. A new aircraft interior noise simulator for psychoacoustic testing. Brenda M. Sullivan and Clemans A. Powell (Fluid Mech. and Acoust. Div., NASA Langley Res. Ctr., Hampton, VA 23681, b.m.sullivan@larc.nasa.gov)

NASA is conducting a research program in passenger response to aircraft interior noise to develop tools for use in design decisions for interior noise treatments. A new interior simulator has been built at NASA–Langley to be used in this program. The simulator is a shell fitted with interior trim and seats from 737/727 aircraft. It contains five listening stations, each having a pair of headphones for binaural signal presentation. Binaural recordings made in interiors of a number of aircraft were processed on a workstation into 50 sound stimuli. These were played back over the headphones and analyzed for repeatability within and between headphones. Initial results indicated that ranges averaged within headphones were 0.3 dB (A weighted), 0.3 phons and 0.7 dB (unweighted). When averaged between headphones, ranges were 1.0 dB (A weighted), 0.9 phons and 1.4 dB (unweighted). A first test in the simulator presented stimuli from propeller airplanes digitally modified to reduce the tonal components. Regressions between subjects' preference responses and measured metrics indicate that the subjective response correlates as well with the arithmetic mean of measurements from the left and right ears of the headphones as with the energy or pressure sums, and better than with measurements from the worse ear.

Contributed Papers

2:30

4pSP3. Perceived unpleasantness of natural sounds: Ratio-scale measurement and psychoacoustic analysis. Wolfgang Ellermeier, Markus Mader (Inst. für Psychol., Univ. Regensburg, D-93040 Regensburg, Germany), and Peter Daniel (Neutrik Cortex Instruments, Regensburg, Germany)

Paired comparisons of a heterogeneous set of ten natural sounds were collected from 60 listeners in order to determine: (1) if the sensation of unpleasantness is judged consistently across a wide range of acoustic stimuli; and (2) which sound features contribute to that sensation. The judgments conformed with the highly restrictive *BTL model* [R. D. Luce, *Individual Choice Behavior* (Wiley, New York, 1959)], thus justifying ratio-scale representation of perceived unpleasantness. The resulting scale values varied by a factor of 100 (diesel engine versus jackhammer). While they were not predicted by differences in A-weighted sound-pressure levels, a linear combination of the psychoacoustic attributes of loudness, roughness, and sharpness accounted for 98% of the variance in perceived unpleasantness.

2:45

4pSP4. Calibration of sound quality recording and playback systems. Poul Ladegaard (Bruel & Kjaer, Skodsborgvej 307, DK-2850 Naerum, Denmark)

In sound quality, a head and torso simulator (HATS) plays as important a role as the recording media. Used with playback of the signals through headphones, it allows the most accurate substitute for actually having the listener in the live situation. The human perception of sounds is very sensitive to changes in sound level, frequency response, and directional information. The value of a subjective sound quality evaluation is directly dependent on how well these parameters are maintained correctly. The paper describes the design methods and goals for a good HATS. Also the requirements for choice of headphones is discussed. In the complete recording/playback process, there is a need for some carefully selected frequency equalizations; these will be treated in some detail. Special attention will be given to the level calibration of the entire system. Here, the CIC check facility comes in as a very attractive and relevant tool. Finally, the summary will discuss the few limitations of the entire recording/playback system. Hints on how to overcome these will be given.

3:00

4pSP5. Voice selection for speech synthesis. Ann K. Syrdal, Alistair Conkie, Yannis Stylianou, Juergen Schroeter (AT&T Labs.—Res., 180 Park Ave., Florham Park, NJ 07932), Laurie F. Garrison, and Dawn L. Dutton (AT&T, Holmdel, NJ 07733)

A TTS voice quality experiment was conducted to select a speaker and to evaluate synthesis techniques. Small-scale TTS diphone inventories using six professional female speakers who were pre-selected in an audition were recorded. Two types of inventories were recorded for each speaker: a series of nonsense words and a series of English sentences. Using these 12 inventories, two synthesis methods were compared: PSOLA [Charpentier and Moulines, Eurospeech '89] and Harmonic Plus Noise (HNM) [Stylianou *et al.*, Eurospeech '97]. Synthetic prosody closely modeled naturally spoken versions of the target utterances. Three fully synthetic (TTS) and two hybrid (i.e., partly recorded from the human speaker and partly synthesized) sentences formed the experimental stimuli for subjective testing. For references, two MNRU versions of the naturally spoken sentences were used: (a) Q10 (resembling low-end commercial 16-kbps encoded speech) and (b) Q35 (resembling high-quality telephone speech). Forty-one subjects rated intelligibility [I], naturalness [N], and pleasantness [P] on five-point MOS scales. A total of 936 ratings were collected from each subject. Repeated measures of analyses of variance (ANOVAs) were performed on the data. There were significant main effects of speaker, synthesis method, and inventory, plus interactions. It was found that (1) the best speaker consistently outperformed the others on all three rating scales; for the optimal combination of parameters, TTS ratings ranged (across speakers) as follows: [I] 3.64–2.94, [N] 3.36–2.7, [P] 3.34–2.53. (2) HNM outperformed PSOLA (consistently 0.25 points higher for [I], [N], [P] scores), and (3) the diphone inventory extracted from sentences was preferred over that extracted from nonsense words (with a significantly smaller difference of 0.10 for HNM than 0.19 for PSOLA).

Session 4pUW

Underwater Acoustics: Signal Processing and Land-to-Sea Acoustics

Ahmad T. Abawi, Chair
Ocean and Atmospheric Science Division, NCCOSC RDTE Division, Code D881, San Diego, California 92152-5001

Contributed Papers

1:00

4pUW1. Matched Doppler processing with multiple receivers. Stephen J. Searle (CRC for Robust and Adaptive Systems, CSSIP and Univ. of Adelaide, Commun. Div., Knowledge Systems Bldg., P.O. Box 1500, Salisbury, SA 5108, Australia) and Douglas A. Gray (Univ. of Adelaide, Salisbury SA 5108, Australia)

Matched Doppler processing (MDP) is a new technique that estimates the range and speed of a near-field narrow-band sound source as it moves past an underwater acoustic sensor by coherently matching the received signal against replicas. Replica vectors are generated according to a simple propagation model, which accounts for the effect of the Doppler phenomena on the signal phase. The data to be matched are the vector of complex outputs of Fourier transforms of contiguous blocks of hydrophone data and the matching is by means of a cost or correlation function, which processes the signal phase coherently. The main contribution reported is the extension of MDP to an array of receivers for localization of a source's azimuth and heading as well as range and speed. Two methods are considered: The first method performs single receiver MDP on each sensor independently, then sums the cost values. The second method processes the signal phase coherently across space and also time. Extensive simulations compare the performance of both methods for an array of varying apertures and the number of receivers. Further simulation indicates that the second method affords sidelobe suppression, which improves performance when multiple signals are present. Application to tracking multiple sources is suggested and discussed.

1:15

4pUW2. On the use of signal autocorrelation matching in localization algorithms. Kevin B. Smith, Joachim Brune (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943), and Ching-Sang Chiu (Naval Postgrad. School, Monterey, CA 93943)

Many new ''robust'' localization algorithms rely on the insensitivity of a high-order processor. In this work, the robustness of a simple, Bartlett-type processor based on matching broadband signal autocorrelation functions is investigated. Specifically, the autocorrelation of the time-domain signal and the autocorrelation of the frequency-domain response will be matched. In order to highlight weaker aspects of these quantities, logarithms of these functions will also be used. Measures of robustness to be examined include the size of the localization footprint on the ambiguity surface and the peak-to-sidelobe levels in the presence of environmental mismatch. A full-wave PE model is used to produce broadband replicas. Model-generated synthetic signals with complete knowledge of the environment will provide base-line results. [This work is supported by the Naval Undersea Warfare Center, Code 2121.]

1:30

4pUW3. Channel-sensitive matched filter processor: Sensitivity and optimization study. Georgios Haralabus and Angela D'Amico (SACLANTCEN Undersea Res. Ctr., Viale S. Bartolomeo 400, 19138 La Spezia, Italy)

The detection performance of the channel-sensitive processor (CSP) [D. Alexandrou and G. Haralabus, SACLANTCEN Report No. SR-263 (1997)] has been tested in dense multipath conditions. It has been demonstrated that for a known propagation channel the CSP method outperforms the conventional matched filter technique. However, in an uncertain environment, the probability of detection of CSP degrades according to the degree of mismatch between the assumed and the actual channel characteristics. It has been found that the processor is more sensitive to geometric parameters (source range and depth) than to environmental parameters (sound velocity profile, sediment-subbottom interface, sediment depth). To overcome the performance degradation due to channel mismatch, the CSP method has been utilized in conjunction with two nonlinear optimization algorithms: the classical simulated annealing (SA) and the multilayer simulated annealing (MUSA) method. At the expense of processing time, it has been found that the optimization methods reduce the channel mismatch effect and considerably improve the detection performance of the CSP.

1:45

4pUW4. Inverse noise rejection. James H. Wilson (Neptune Sci., Inc., 4711 Viewridge Ave., Ste. 150, San Diego, CA 92123), James E. Donald, and Albert S. Nuttall (Naval Undersea Warfare Ctr., Newport, RI)

Inverse beamforming (IBF) has been developed over the last decade by the authors, and has proved highly successful in numerous at-sea tests. IBF is a unique signal processing method because the noise field properties are first determined by inverse techniques as a function of frequency, azimuth, magnitude, and the ''widths'' of these three dependent variables. The statistical properties of the noise field are determined, and a beamforming algorithm and six-dimensional tracker were developed to reject the noise and pass the signal (whose properties are well studied). IBF applies to deep water only, and this paper addresses the development of a similiar noise rejection algorithm in shallow water. The shallow-water problem is more complex due to the normal mode propagation characteristics at low frequencies, and the diversity of noise mechanisms in shallow water. Noise measurements are critically needed in shallow water to validate this approach.

2:00

4pUW5. Acoustic array analysis using multiple processors in parallel. Nils Paz and Ronald W. Townsen (CST Images, 3836 Browning St., San Diego, CA 92107, csti@ix.netcom.com)

Acoustic arrays are expensive to deploy and require a multidisciplinary team of scientists and engineers to develop, deploy, and monitor. This paper addresses one particular effort developed as a commercial software package, the Underwater Acoustic Analysis Tool (UWAA) by CST Images. UWAA is an acoustic array analysis tool for pre- and post-deployment modeling and analysis. Typically, the modeling of acoustic arrays and their deployment has been developed as separate and fractured modeling efforts. The array laydown modeling was separate from the acoustic beam formation and the propagation modeling. In addition to integrating all the components necessary to deploy an acoustic array, this effort describes the capability to compute the modeling in a parallel multithreaded environment using attached array processors. UWAA has integrated the modeling of the 3-D sensor positioning and sensor response, 3-D beam formation, 3-D acoustic modeling, and complex acoustic array interaction for beam formation using a high-performance multiprocessor system.

2:15

4pUW6. AWSUM EAP: An environmentally sensitive adaptive fluctuation-based processing algorithm that exploits amplitude and phase. Ronald A. Wagstaff and Jackson A. Mobbs (Naval Res. Lab., Stennis Space Center, MS 39529-5004, wagstaff@nrlssc.navy.mil)

Signals that propagate from submerged sources to deep receivers in the ocean generally interact less with the fluctuation generation mechanisms near the sea surface than do clutter signals and noise which originate near the surface. As a result, signals from submerged sources generally have smaller fluctuations than clutter signals and noise. The order dependence in the amplitudes of fluctuations has been used to devise an environmentally adaptive signal processing algorithm that provides preferential gains for signals having smaller fluctuation amplitudes than those of clutter and noise [R. A. Wagstaff and J. A. Mobbs, J. Acoust. Soc. Am. **101**, 3027 (1997)]. Gains include increases in signal-to-noise ratio, clutter suppression, and unalerted automatic detection. Similar gains can be achieved by exploiting the fluctuations in the phases of the acoustic pressures in the same manner that the order dependence was exploited. By replacing the order dependence with phase dependence, a similar algorithm has been devised that is sensitive to both amplitude and phase fluctuations and still adapts to the input data. The resulting algorithm, designated the AWSUM environmentally adaptive phase (EAP) is described, and results from measured data are presented. [Work supported by ONR and NRL.]

2:30

4pUW7. Array beamforming in long-range deep-water environments. Elena Yu. Gorodetskaya, Alexander I. Malekhanov, Alexander G. Sazontov, and Nadezhda K. Vdovicheva (Inst. of Appl. Phys., Russian Acad. of Sci., 46 Ul'yanov St., 603600 Nizhny Novgorod, Russia, almal@hydro.appl.sci-nnov.ru)

Realistic estimations of the deep-water acoustic coherence effects on the array gain for linear and quadratic beamformers, optimal ones included, were obtained for the North-West Pacific environments. An advanced technique was developed to calculate the signal coherence under the basic assumption that the long-range acoustic fluctuations are caused by internal waves in the summer channel or surface wind waves in the winter channel. Simulations were carried out both for horizontal and vertical arrays for the frequency of 250 Hz and ranges 500–1000 km. The following effects were shown to be the most essential points: (i) angular displacement and degradation of the plane-wave beampattern; (ii) large-array gain loss; (iii) coherence-induced gain ''gap'' between the optimal quadratic and linear beamformers; and (iv) gain dependence on the ambient modal noise. The optimal quadratic beamformer was shown to reduce the gain loss at a cost of increased processor complexity: The number of partial weight-sum channels is equal to the number of the largest signal eigenvalues. In some environments, a proper performance/complexity was realized using suboptimal beamformers which are, therefore, of particular interest for large-array applications. [Work supported by RFBR.]

2:45

4pUW8. Near-field polarization processing of ice fracturing events. Yuriy V. Dudko and Henrik Schmidt (Dept. of Ocean Eng., MIT, Cambridge, MA 02139)

As part of the ONR Sea-Ice Mechanics Initiative (SIMI) experiments in the spring of 1994, arrays of high-resolution seismic sensors were deployed in areas of high seismicity identified by real-time processing of acoustic emission events recorded by a wide, horizontal aperture hydrophone array. One of the areas most rich in ice events was located on a small ice floe ($\sim 100 \times 100$ m) 4 km Northeast of the main camp, where the seismic array consisting of five 3-component geophones in a 70-m aperture pentagon was deployed. One of the methods employed for analysis of the near-field data from this deployment was motion-product seismograms introduced by J. E. White [Geophysics **24**, 288–298 (1964)]. In this method each of the horizontal components of ice motion was multiplied by the vertical component, with or without phase shift, and after integration, the two resulting products identified a vector pointing to the source of the seismic waves. The ability of this method to separate different polarizations of the seismic waves and to determine the direction to the source was especially useful for geophone array data, because other analysis methods for such data occasionally failed due to the overlapping of waves generated by different ice events. Using the polarization processing method the development of ice fractures in the array near field was successfully tracked in the time and spatial domains. One result of the polarization analysis was that these fractures seemed to mostly generate vertically polarized shear (SV) waves. [Research supported by ONR.]

3:00–3:15 Break

3:15

4pUW9. Detection and tracking of land vehicle activity by offshore underwater acoustic arrays. Gerald L. D'Spain, Lewis P. Berger, William S. Hodgkiss, William A. Kuperman, LeRoy M. Dorman, and William A. Gaines (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Land-based vehicle activity can be tracked by underwater acoustic sensors located at least as far as 3.4 km offshore. This capability is demonstrated with data from two near-surf-zone experiments as part of the Marine Physical Laboratory's Adaptive Beach Monitoring program. In one case, four tracked vehicles spaced every 150 m traveled down the beach. Spectra estimated from a single hydrophone of a bottom line array 1.5 km offshore in 12-m water show a high-pass character compared to land geophone recordings because of the coupling characteristics between the land seismic and underwater acoustic fields. Adaptive plane-wave beamforming over the 30- to 70-Hz band shows, at times, beam-peak-to-background levels exceeding 35 dB. The acoustic tracks agree with those from visual logs and suggest that the four vehicles can be enumerated acoustically, particularly when combined with results from the 120- to 130-Hz band. In a second case, plane-wave beamforming results for two tracked vehicles traveling on the beach and recorded by a bottom array 3.4 km offshore in 20-m water show a beam-peak-to-background level approaching 20 dB. A videotape shows the evolution of 2-D wave-number spatial spectra during the vehicle transit. [Work supported by ONR, Code 32.]

3:30

4pUW10. Coupling of land-based signals into the nearshore underwater acoustic field. Gerald L. D'Spain, Lewis P. Berger, LeRoy M. Dorman, William S. Hodgkiss, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Small, controlled land detonations on and near the beach recorded by land geophones and by offshore seismoacoustic sensors are used to characterize the coupling between land seismic and underwater acoustic fields. These data were collected in the Marine Physical Laboratory's Adaptive Beach Monitoring program. The predominant arrival on land is an elastic surface wave. Although an ocean bottom interface wave can be seen clearly in seismoacoustic recordings just outside the surf zone, it attenuates to nearly background noise levels by the time it reaches a bottom hydrophone array 1.5 km offshore in 12-m water. Travel-time and frequency/wave-number analyses indicate that the hydrophone array's received energy is composed mostly of dispersive body waves propagating as the lowest water-borne mode, with frequency-dependent phase and group velocities around 2.0 and 1.8 km/s, respectively. The frequency at which this lowest mode cuts off is the corner frequency of the high-pass filter representing the propagation from land to the underwater acoustic field. These results and numerical modeling, guided by insights obtained from classic work on acoustic propagation in a fluid wedge, are used to explain offshore underwater recordings of land vehicle activity. [Work supported by ONR, Code 32.]

3:45

4pUW11. Sound radiation from motion-related ice processes in the central Arctic. Catherine Stamoulis (Dept. of Civil and Environ. Eng., MIT, Cambridge, MA 02139) and Ira Dyer (MIT, Cambridge, MA 02139)

The radiation characteristics of individual acoustic events from the central Arctic SIMI experiment data were studied, in the frequency range 10–350 Hz, with the purpose to identify the event generating ice mechanisms. Four event subpopulations and consequently four ice mechanisms were distinguished. The first subpopulation consisted of events attributed to floe unloading, a postfracture process, which was consistently detected in the ambient noise data following intervals of high ice fracture activity. The event radiation patterns were independent of azimuth and were modeled by a stationary vertical dipole. The second subpopulation consisted of events attributed to tensile fracture, the radiation pattern of which was modeled by a weighted superposition of longitudinal octopoles, modified by a Doppler factor to account for source motion. The last two subpopulations consisted of events attributed to shear fracture. A lateral octopole and a weighted combination of lateral and longitudinal octopoles, both modified by a Doppler factor, were used to describe the radiation characteristics of events in these subpopulations, respectively. The adequacy of the latter model indicated that events in the fourth subpopulation had been induced by shear fractures, which propagated through the formation of arrays of tensile cracks in their tips and edges.

4:00

4pUW12. Statistics of very deep ocean noise. H. M. Walkinshaw (Box 72, Peapack, NJ 07977)

Sea noise was measured continually over a period of 1 year on experimental installations 40 miles south of Bermuda. One-hour stretches of data were recorded daily on four hydrophones, two on the seafloor and two suspended off bottom in 2420 fm s. Analysis of spectrum levels at six frequencies from 0.1 to 1.5 kHz included computation of means, standard deviations, distributions, and higher-order statistical moments. Spectral shape, levels, and variability closely resembled sea noise measured at shallower receivers throughout the western North Atlantic. The suspended very deep hydrophones were 2 to 3 dB noisier than those on the bottom; other statistical properties of both suspended and bottomed receivers were alike. The mean monthly noise migrated through an annual cycle of frequency-dependent change, from 7 dB at 0.1 kHz to 12 dB at 1.5 kHz. This seasonal trend progressed smoothly from winter high to midsummer low, though not at a constant rate. Monthly probability distributions changed form correspondingly, becoming skewed for months with a rapidly shifting mean and tightly peaked for seasons of relative climatic stability. [Work supported by NavElecs.]

4:15

4pUW13. Acoustic image analysis by color mapping. Frank A. Boyle (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

Acoustic returns from different types of structures can appear very similar, making acoustic image classification difficult. Current research is aimed toward developing novel methods of generating acoustic images that convey relevant information for target identification and classificatin. A color-mapping method [Boyle and Chotiros, J. Acoust. Soc. Am. **99**(4), 2553(A) (1996)], which maps an image's spectral content into color has been applied to several examples of broadband acoustic data, with apparent success. Structural features that produce similar echoes in amplitude and duration often emerge with clear differences in color according to their spectral signature. Possibilities for using color to convey acoustic properties other than spectral content are also being explored. One such method maps the phase of a signal into colors, while another is based on wavelet filtering of signals. The presentation will include descriptions of color mapping and related methods, and examples of their application to acoustic data.

4:30

4pUW14. Uncertainty in fisheries echosounder calibrations. David A. Demer (Southwest Fisheries Sci. Ctr., P.O. Box 271, La Jolla, CA 92038) and Michael A. Soule (Sea Fisheries Res. Inst., Capetown, South Africa)

Calibration of echosounders for fish stock assessment are commonly performed using the standard sphere method. A complete calibration includes characterization of system sensitivity versus detection angle for measuring target strengths (TS) of individual scatterers and the equivalent beam angle for measuring backscattering strengths of insonified volumes. To explore the accuracy of the method, on-axis TS measurements were made of three standard spheres (23-mm copper and 33.0- and 38.1-mm tungsten carbide), under controlled conditions at a frequency of 120 kHz (bandwidth = 1.2 kHz). The TS measurements, derived by both integrated and peak intensities, were compared to their theoretical counterparts. Also, system sensitivities and split-beam detection angles were characterized versus target bearings. To investigate measurement precision, a commercial echosounder (Simrad EK500) was used to measure the TS of the spheres over 15-h periods with constant pulse duration (0.3 ms) and water temperature (19 °C). To characterize the temperature dependence of the calibrations, measurements of system gain and transducer admittance were made versus water temperature (0–17 °C). Finally, bootstrap simulations were used to estimate the combined uncertainty in 120-kHz echosounder calibrations using optimal standard spheres and over ranges of temperatures typically encountered in fisheries surveys.

4:45

4pUW15. ATOC signal enhancement using adaptive filtering. Gary E. J. Bold and Sze M. Tan (Dept. of Phys., Univ. of Auckland, Private Bag 92019, Auckland, New Zealand)

A moored, autonomous recording system deployed off the east coast of New Zealand in early 1996 acquired signals from the Pioneer seamount ATOC source. Unfortunately, the data were corrupted by a very strong, aliased interfering signal electrically coupled from a malfunctioning power supply which drifted in frequency by over 400 Hz, and also by a mechanically coupled vibration at the rotation frequency of the hard disk used to store the data. Pulse compression of the biphase modulated psuedorandom sequence used to encode the transmission spreads the energy in these signals across the desired signal's spectral passband, degrading the signal-to-noise ratio of the receptions. Since both unwanted components are unstable in frequency, classical digital filters are unable to eliminate them. However, an adaptive LMS filter has been used to track and virtually remove the interfering signals before pulse compression, resulting in signal-to-noise ratio gains of up to 6 dB.

Meeting of Accredited Standards Committee S1 on Acoustics

J. P. Seiler, Chair S1
U.S. Department of Labor, Cochran Mill Road, P.O. Box 18233, Building 038, Pittsburgh, Pennsylvania 15236

G. S. K. Wong, Vice Chair S1
Institute for National Measurement Standards, National Research Council, Ottawa, Ontario K1A 0R6, Canada

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
U. S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U. S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology, Building 233, Room A149, Gaithersburg, Maryland 20899

Accredited Standards Committee S1 on Acoustics. Working group chairs will report on their preparation of standards on methods of measurement and testing, and terminology, in physical acoustics, electroacoustics, sonics, ultrasonics, and underwater sound. Work in progress includes measurement of noise sources, noise dosimeters, integrating sound-level meters, and revision and extension of sound level meter specifications. Open discussion of committee reports is encouraged.

Scope of S1: Standards, specifications, methods of measurement and test and terminology in the field of physical acoustics including architectural acoustics, electroacoustics, sonics and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance and comfort.

THURSDAY AFTERNOON, 4 DECEMBER 1997 SHEFFIELD ROOM, 3:30 P.M. TO 4:45 P.M.

Meeting of Accredited Standards Committee S3 on Bioacoustics

L. S. Finegold, Chair S3
USAF Armstrong Laboratory, Noise Effects Branch AL/OEBN, 2610 Seventh Street, Wright–Patterson Airforce Base, Ohio 43433-7901

R. F. Burkard, Vice Chair S3
Hearing Research Laboratory, State University of New York at Buffalo, 215 Parker Hall, Buffalo, New York 14214

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
U. S. CERL, P.O. Box 9005, Champaign, Illinois 61826-9005

J. Erdreich, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
Ostergaard Acoustical Associates, 100 Executive Drive, West Orange, New Jersey 07052

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics and ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U.S. Technical Adviser (TA) for IEC/TC29, Electroacoutics
National Institute of Standards and Technology, Building 233, Room A149, Gathersburg, Maryland 20899

Accredited Standards Committee S3 on Bioacoustics. The current status of standards under preparation will be discussed. In addition to those topics of interest, including hearing conservation, noise, dosimeters, hearing aids, etc., consideration will be given to new standards which might be needed over the next few years. Open discussion of committee reports is encouraged.

Scope of S3: Standards, specifications, methods of measurement and test, and terminology in the fields of mechanical shock and physiological acoustics, including aspects of general acoustics, shock and vibrations which pertain to biological safety, tolerance, and comfort.