Stable dereverberation using microphone arrays for speaker verification

A. C. Surendran and J. L. Flanagan

Citation: The Journal of the Acoustical Society of America **96**, 3261 (1994); doi: 10.1121/1.411005
View online: https://doi.org/10.1121/1.411005
View Table of Contents: https://asa.scitation.org/toc/jas/96/5
Published by the Acoustical Society of America

4:30

2pSA11. The application of a biologically inspired controller to control sound transmission. James P. Carneal and Chris R. Fuller (Vib. and Acoust. Labs., Dept. of Mech. Eng., Virginia Polytech. Inst. and State Univ., Blacksburg, VA 24061)

A biologically inspired control approach for reducing sound transmission through a distributed elastic system has been theoretically and experimentally verified for narrow-band excitation. The control paradigm approximates natural biological systems for initiating movement, in that a low number of signals are sent from an advanced, centralized controller (analogous to the brain) and are then distributed by local rules and actions to multiple actuators (analogous to muscle fiber). A local learning rule that was developed from linear quadratic optimal control theory and solved *a priori* was implemented. The investigation considered a plate excited by normal plane wave, oblique plane wave, and reverberant acoustic fields. Radiated sound power was the defined cost function and therefore used as the controller error signal. Four control inputs in the form of piezoelectric actuators were mounted on the plate in a two-by-two array. Results indicate that increases in transmission loss of approximately 18 dB are attainable for off-resonance excitation. In general, comparisons of theoretical and experimental data show good agreement. This investigation has demonstrated that the biological control approach has the potential to control multimodal response in distributed elastic systems using an array of many actuators with a reduced order main controller. Thus significant reductions in control system computational complexity have been realized by this approach. [Work supported by NASA Langley.]

4:45

2pSA12. Experimental results from hybrid control with a sensoriactuator. Robert L. Clark and Jeffrey S. Vipperman (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708-0300)

Both transient and persistent disturbance rejection were demonstrated experimentally on a cantilevered beam configured with a piezoelectric sensoriactuator. The transient response of the system was suppressed through direct-rate feedback control, and adaptive feedforward control was utilized to minimize the response to a harmonic input disturbance. A time-averaged gradient descent algorithm was implemented to adapt a finite impulse response filter in the feedforward control approach. Experimental results demonstrate that rate-feedback control can be utilized to enhance the transient adaptation of the feedforward control algorithm. Furthermore, the sensoriactuator provides a convenient method of performing both sensing and actuation simultaneously in feedback and feedforward control of adaptive structures.

5:00

2pSA13. Determination of effective secondary source locations for active noise control. Jihe Yang (Automated Analysis Corp., 2805 S. Industrial, Ste. 100, Ann Arbor, MI 48104) and David K. Holger (Iowa State University, Ames, IA 50011)

A numerical method for determining effective secondary source locations for active control of interior sound fields has been investigated. The method uses intermediate results from an indirect boundary element simulation of a sound field to determine effective boundary locations for secondary sources. In the indirect boundary element method (IBEM), an interior sound field is simulated by replacing the physical boundaries with a fictitious source distribution that is determined from the geometry, the properties of the physical boundaries, and the primary source location(s). Locations of high fictitious source strength, as determined by the IBEM, are found to be particularly effective locations for secondary sources that are components in three dimensional active noise control systems. Numerical results for simple geometries are in agreement with previous experimental results [Elliott *et al.*, J. Sound Vib. **117**, 35–58 (1987)], and numerical predictions of active noise control using the proposed method for locating secondary sources resulted in sound pressure level reductions of more than 20 dB in reverberant and semi-reverberant spaces. The results obtained suggest that the method has significant potential for efficiently locating effective secondary sources for a variety of active noise control applications.

TUESDAY AFTERNOON, 29 NOVEMBER 1994 BALLROOM A, 1:30 TO 4:30 P.M.

Session 2pSP

Speech Communication and Engineering Acoustics: Microphone Arrays: Design and Analysis II

James L. Flanagan, Cochair
CAIP Center, Busch Campus, Rutgers University, Core Building 706, Piscataway, New Jersey 08855-1390

Harvey F. Silverman, Cochair
Laboratory for Engineering Man/Machine Systems, School of Engineering, Brown University, Providence, Rhode Island 02912

Qiguang Lin, Cochair
CAIP Center, Rutgers University, Core Building, Frelinghuysen Road, Piscataway, New Jersey 08855-1390

Contributed Papers

1:30

2pSP1. Microphone arrays for reducing reverberation and noise in speech communication. James G. Ryan (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

Microphone pickup of sound in typical rooms is impaired by the combined effects of reverberation and noise. This degrades speech intelligibility and quality particularly in applications where the microphone is located far away from the talker. Recent advances in microphone array technology suggest a potential solution to such problems. This paper gives an overview of current microphone array techniques and discusses the potential benefits for speech communication. Various criteria for measuring the performance of a microphone array are described. A flexible, experimental microphone array intended for research in speech communications is under construction and will be described.

1:45

2pSP2. Stable dereverberation using microphone arrays for speaker verification. A. C. Surendran and J. L. Flanagan (Ctr. for Comput. Aids for Indust. Productivity, P.O. Box 1390, Piscataway, NJ 08855-1390)

The impulse response of a reverberant environment, in general, is a nonminimum phase and cannot be inverted. But an exact inverse of the environment can be obtained by modeling the room as a multiple input–output (MINT) system [M. Miyoshi and Y. Kaneda ICASSP (1986)]. In this report, this model is applied to a microphone array and is used as a front-end processor for a speaker verification system. The G matrix is inverted using row action projection (RAP), an iterative approach to solving a system of linear equations. Starting from an initial guess, the solution is repeatedly projected onto each hyperplane of the equation system until it converges. The method is stable, robust to noise, and converges to the pseudo-inverse solution. In computer-simulated experiments, the signal-to-reverberant-noise ratio is found to improve with the number of microphones in the array. A speaker verification system using the array is evaluated at various signal-to-competing-noise ratios (SCNR). Results suggest that verification performance can be substantially elevated in adverse acoustic environments.

2:00

2pSP3. Binaural arrays for hearing enhancement. Michael V. Scanlon and Stephen M. Tenney (Army Res. Lab., AMSRL-SS-SH, 2800 Powder Mill Rd., Adelphi, MD 20783-1197)

Two hearing augmentation devices developed at the Army Research Laboratory can enhance normal listening abilities and restore hearing degraded by encapsulating headgear. Surrounding sounds are localized with a head-mounted binaural pinna attachment that recreates the head-related transfer function associated with the normal listening. The user's brain interprets the recreated stereo signals that enter the ear canals through intra-aural speakers, giving excellent restoration of omnidirectional hearing. A hand-held, ultra-directional array extends the user's listening range. The use of delay and sum beamforming in the array assures maximum directivity in the pointing direction. The binaural long-range hearing device has two linear endfire arrays of eight cardioid microphones each. The slightly offset directivity patterns of the two arrays create stereo outputs, so that the user can interpret differences in amplitude, phase, time-of-arrival, and frequency content of sounds in the forward area. These devices provide aural protection and an intra-aural input for communications, without removing the user from his acoustic environment. Both devices can be monitored remotely, and are ideally suited for detecting speech, personnel, equipment, or vehicles during military or law enforcement missions. Performance measurements of various array configurations will be shown.

2:15

2pSP4. A multisensor connectivist model for the preprocessing of the speech signals. Turker Kuyel and Elmer L. Hixson (Dept. of Elec. Eng., Univ. of Texas at Austin, Austin, TX 78712)

Due to the inherent redundancy of the speech data, the design of a redundancy reducing speech preprocessor is very important. Preprocessor design is also very important because it can greatly reduce the computational load on the later stages of speech processing. A special laboratory oriented method in speech data acquisition, which is called near-field spectral wave number estimation is implemented. In this method multiple microphones are used. The goal is to incorporate air flow velocity into speech feature vector. This extra feature is used in addition to the short time cepstrum of the sound data to make the final speech vectors. The speech vectors are then quantized into a determined number of categories using a self-organizing neural network. These quantized and extended vectors are then used for the modeling of higher speech constructs such as phonemes and words. The preprocessing scheme reduced the computational complexity considerably at the expense of slight reduction of the recognition accuracy.

2:30

2pSP5. Position-tolerant differential microphones for noisy environments. James E. West, G. W. Elko, D. R. Morgan, and R. A. Kubli (Acoust. Res. Dept., AT&T Bell Labs., 600 Mountain Ave., Murray Hill, NJ 07974)

Directional microphones are best noted for their noise reduction properties in communication systems. Close-talking differential microphones are particularly useful when the noise environment disturbs the ability to communicate without error, such as in public and cellular telephony, aircraft communications, etc. These differential microphones work best when they are placed within 1 cm from the lips of the talker where the sound field has a large gradient. For a plane-wave sound field the sensitivity rises proportional to ω^n, where n is the order of the difference. Users of differential microphones do not always correctly position the sensor at the proper distance from the mouth and therefore the sensitivity of the microphone may also rise proportional to ω^n especially at high frequencies. A method is described of correcting for this high-frequency gain without significantly degrading the noise canceling properties of first- and second-order differential microphones.

2:45

2pSP6. A new adaptive differential microphone array. Gary W. Elko and Anh-Tho Nguyen Pong (Acoust. Res. Dept., AT&T Bell Labs., 600 Mountain Ave., Murray Hill, NJ 07974)

An adaptive differential microphone has been implemented by combining two omnidirectional elements to form back-to-back cardioid directional microphones. By combining the weighted subtraction of these two outputs, any first-order array can be realized. If certain simple constraints are placed on the combination weighting, the null location can be constrained to defined angular regions. Three algorithms that control the constrained adaptation are presented and discussed for the array: the LMS algorithm, Newton's algorithm, and a time-varying least-squares Wiener filter. A real-time implementation utilizing an AT&T DSP32C digital signal processor is also described.

3:00–3:15 Break

3:15

2pSP7. An adaptive subband differential microphone. Juergen Cezanne and Anh-Tho Nguyen Pong (Acoust. Res. Dept., AT&T Bell Labs., 600 Mountain Ave., Murray Hill, NJ 07974)

In a previous talk, "A new adaptive differential microphone array" by Elko and Pong, a differential microphone has been introduced that adapts its directivity pattern to the particular acoustic environment to provide for a good signal-to-noise ratio. There, the selected pattern remains more or less constant with respect to frequency. In this talk an approach is described that contains one more degree of freedom. The spectrum of the signals is partitioned in uniform subbands and different directivity patterns are adaptively chosen in each subband. This allows to cancel multiple noise sources with nonoverlapping spectra. An LMS-based algorithm will be derived with focus on a low computational load and a short delay for the desired signal. Consequences on the speed of adaptation are discussed. Further, experimental results of a first implementation with 33 subbands on a PC-based DSP32C board will be presented. The measurements verify the ability of the algorithm to cancel multiple noise sources with disjoint spectra without distorting the desired signal.

3:30

2pSP8. Adaptive enhancement of microphone array signals. Carsten Sydow (Inst. for Electroacoust., Tech. Univ. of Darmstadt, Merckstr. 25, D-64283 Darmstadt, Germany)

The signal-to-noise ratio of a speech signal picked up by a microphone array can be improved by adaptive post processing. Enhancement techniques known from single microphone or dual microphone signal processing, like noise canceling and spectral subtraction can be extended to a multimicrophone array system. The noise canceling technique and derived structures try to model the room impulse response by an adaptive transversal filter. Thus the performance of these algorithms is limited by the ratio of filter length to reverberation time and by the capability to track the nonstationary impulse response. Reduction of the noise of approximately 8 dB can be achieved with acceptable filter length in a stationary environment, but precautions must be taken to avoid canceling of the desired speech signal. The spectral subtraction met' ͻd yields higher improvements