

AES

JOURNAL OF THE AUDIO ENGINEERING SOCIETY

AUDIO / ACOUSTICS / APPLICATIONS

Volume 50 Number 11

2002 November

In this issue...

Noisy Decay Parameters

**Improved Time-Frequency
Analysis**

Expanding Surround Sound

Synthesizing Surround Sound

Features...

**113th Convention Report,
Los Angeles**

**Bluetooth and
Wireless Networking**

**11th Regional Convention, Tokyo—
Call for Papers**





AUDIO ENGINEERING SOCIETY, INC.

INTERNATIONAL HEADQUARTERS

60 East 42nd Street, Room 2520, New York, NY 10165-2520, USA

Tel: +1 212 661 8528 Fax: +1 212 682 0477

E-mail: HQ@aes Internet: http://www.aes.org

ADMINISTRATION

Roger K. Furness Executive Director

Sandra J. Requa Executive Assistant to the Executive Director

OFFICERS 2002/2003

Kees A. Immink President

Ronald Streicher President-Elect

Garry Margolis Past President

Jim Anderson Vice President

Eastern Region, USA/Canada

James A. Kaiser Vice President

Central Region, USA/Canada

Bob Moses Vice President,

Western Region, USA/Canada

Søren Bech Vice President

Northern Region, Europe

Markus Erne

Vice President, Central Region, Europe

Daniel Zalay

Vice President, Southern Region, Europe

Mercedes Onorato Vice President,

Latin American Region

Neville Thiele

Vice President, International Region

Han Tendeloo Secretary

Marshall Buck Treasurer

GOVERNORS

Karl-Otto Bäder

Curtis Hoyt

Roy Pritts

Don Puluse

David Robinson

Annemarie Staepelaere

Roland Tan

Kunimara Tanaka

COMMITTEES

AWARDS

Roy Pritts Chair

CONFERENCE POLICY

Søren Bech Chair

CONVENTION POLICY & FINANCE

Marshall Buck Chair

EDUCATION

Don Puluse Chair

FUTURE DIRECTIONS

Kees A. Immink Chair

HISTORICAL

J. G. (Jay) McKnight Chair

Ted Sheldon Vice Chair

Donald J. Plunkett Chair Emeritus

LAWS & RESOLUTIONS

Ron Streicher Chair

MEMBERSHIP/ADMISSIONS

Francis Rumsey Chair

NOMINATIONS

Garry Margolis Chair

PUBLICATIONS POLICY

Richard H. Small Chair

REGIONS AND SECTIONS

Subir Pramanik Chair

STANDARDS

John P. Nunn Chair

TELLERS

Christopher V. Freitag Chair

WOMEN IN AUDIO

Kees A. Immink Chair

TECHNICAL COUNCIL

Wieslaw V. Woszczyk Chair

Jürgen Herre and

Robert Schulein Vice Chairs

TECHNICAL COMMITTEES

ACOUSTICS & SOUND

REINFORCEMENT

Mendel Kleiner Chair

Kurt Graffy Vice Chair

ARCHIVING, RESTORATION AND

DIGITAL LIBRARIES

David Ackerman Chair

AUDIO FOR

TELECOMMUNICATIONS

Bob Zurek Chair

Andrew Bright Vice Chair

CODING OF AUDIO SIGNALS

James Johnston and

Jürgen Herre Cochairs

AUTOMOTIVE AUDIO

Richard S. Stroud Chair

Tim Nind Vice Chair

HIGH-RESOLUTION AUDIO

Malcolm Hawksford Chair

Vicki R. Melchior and

Takeo Yamamoto Vice Chairs

LOUDSPEAKERS & HEADPHONES

David Clark Chair

Juha Backman Vice Chair

MICROPHONES & APPLICATIONS

David Josephson Chair

Wolfgang Niehoff Vice Chair

MULTICHANNEL & BINAURAL

AUDIO TECHNOLOGIES

Francis Rumsey Chair

Gunther Theile Vice Chair

NETWORK AUDIO SYSTEMS

Jeremy Cooperstock Chair

Robert Rowe and **Thomas**

Sporer Vice Chairs

AUDIO RECORDING & STORAGE

SYSTEMS

Derk Reefman Chair

PERCEPTION & SUBJECTIVE

EVALUATION OF AUDIO SIGNALS

Durand Begault Chair

Søren Bech and **Eiichi Miyasaka**

Vice Chairs

SIGNAL PROCESSING

Ronald Aarts Chair

James Johnston and **Christoph M.**

Musialik Vice Chairs

STUDIO PRACTICES & PRODUCTION

George Massenburg Chair

Alan Parsons, **David Smith** and

Mick Sawaguchi Vice Chairs

TRANSMISSION & BROADCASTING

Stephen Lyman Chair

Neville Thiele Vice Chair

STANDARDS COMMITTEE

John P. Nunn

Chair

Bruce Olson

Vice Chair, Western Hemisphere

Mark Yonge

Secretary, Standards Manager

Yoshizo Sohma

Vice Chair, International

SC-02 SUBCOMMITTEE ON DIGITAL AUDIO

John P. Nunn Chair

Robert A. Finger Vice Chair

Working Groups

SC-02-01 Digital Audio Measurement Techniques

Richard C. Cabot, I. Dennis, M. Keyhl

SC-02-02 Digital Input-Output Interfacing: Julian Dunn

Robert A. Finger, John Grant

SC-02-05 Synchronization: Robin Caine

SC-03 SUBCOMMITTEE ON THE PRESERVATION AND RESTORATION OF AUDIO RECORDING

Ted Sheldon Chair

Dietrich Schüller Vice Chair

Working Groups

SC-03-01 Analog Recording: J. G. McKnight

SC-03-02 Transfer Technologies: Lars Gaustad, Greg Faris

SC-03-04 Storage and Handling of Media: Ted Sheldon, Gerd Cyrener

SC-03-06 Digital Library and Archives Systems: William Storm

Joe Bean, Werner Deutsch

SC-03-12 Forensic Audio: Tom Owen, M. McDermott

Eddy Bogh Brixen

SC-04 SUBCOMMITTEE ON ACOUSTICS

Mendel Kleiner Chair

Mark Ureda Vice Chair

Working Groups

SC-04-01 Acoustics and Sound Source Modeling

Richard H. Campbell, Wolfgang Ahnert

SC-04-02 Characterization of Acoustical Materials

Peter D'Antonio, Trevor J. Cox

SC-04-03 Loudspeaker Modeling and Measurement

David Prince, Steve Hutt

SC-04-04 Microphone Measurement and Characterization

David Josephson, Jackie Green

SC-04-07 Listening Tests: David Clark, T. Nousaine

SC-05 SUBCOMMITTEE ON INTERCONNECTIONS

Ray Rayburn Chair

John Woodgate Vice Chair

Working Groups

SC-05-02 Audio Connectors

Ray Rayburn, Werner Bachmann

SC-05-03 Audio Connector Documentation

Dave Tosti-Lane, J. Chester

SC-05-05 Grounding and EMC Practices

Bruce Olson, Jim Brown

SC-06 SUBCOMMITTEE ON NETWORK AND FILE TRANSFER OF AUDIO

Robin Caine Chair

Steve Harris Vice Chair

Working Groups

SC-06-01 Audio-File Transfer and Exchange

Mark Yonge, Brooks Harris

SC-06-02 Audio Applications Using the High Performance Serial

Bus (IEEE: 1394): John Strawn, Bob Moses

SC-06-04 Internet Audio Delivery System

Karlheinz Brandenburg

SC-06-06 Audio Metadata

C. Chambers

AES REGIONAL OFFICES

Europe Conventions

Zevenbunderslaan 142/9, BE-1190 Brussels, Belgium, Tel: +32 2 345 7971, Fax: +32 2 345 3419, E-mail for convention information: euroconventions@aes.

Europe Services

B.P. 50, FR-94364 Bry Sur Marne Cedex, France, Tel: +33 1 4881 4632, Fax: +33 1 4706 0648, E-mail for membership and publication sales: euroservices@aes.

United Kingdom

British Section, Audio Engineering Society Ltd., P. O. Box 645, Slough, SL1 8BJ UK, Tel: +44 1628 663725, Fax: +44 1628 667002, E-mail: UK@aes.

Japan

AES Japan Section, 1-38-2 Yoyogi, Room 703, Shibuyaku-ku, Tokyo 151-0053, Japan, Tel: +81 3 5358 7320, Fax: +81 3 5358 7328, E-mail: aes_japan@aes.

AES REGIONS AND SECTIONS

Eastern Region, USA/Canada

Sections: Atlanta, Boston, District of Columbia, New York, Philadelphia, Toronto
Student Sections: American University, Berklee College of Music, Carnegie Mellon University, Duquesne University, Fredonia, Full Sail Real World Education, Hampton University, Institute of Audio Research, McGill University, Pennsylvania State University, University of Hartford, University of Massachusetts-Lowell, University of Miami, University of North Carolina at Asheville, William Patterson University, Worcester Polytechnic University

Central Region, USA/Canada

Sections: Central Indiana, Chicago, Detroit, Kansas City, Nashville, New Orleans, St. Louis, Upper Midwest, West Michigan
Student Sections: Ball State University, Belmont University, Columbia College, Michigan Technological University, Middle Tennessee State University, Music Tech College, SAE Nashville, Northeast Community College, Ohio University, Ridgewater College, Hutchinson Campus, Southwest Texas State University, University of Arkansas-Pine Bluff, University of Cincinnati, University of Illinois-Urbana-Champaign

Western Region, USA/Canada

Sections: Alberta, Colorado, Los Angeles, Pacific Northwest, Portland, San Diego, San Francisco, Utah, Vancouver
Student Sections: American River College, Brigham Young University, California State University-Chico, Citrus College, Cogswell Polytechnical College, Conservatory of Recording Arts and Sciences, Denver, Expression Center for New Media, Long Beach City College, San Diego State University, San Francisco State University, Cal Poly San Luis Obispo, Stanford University, The Art Institute of Seattle, University of Southern California, Vancouver

Northern Region, Europe

Sections: Belgian, British, Danish, Finnish, Moscow, Netherlands, Norwegian, St. Petersburg, Swedish
Student Sections: All-Russian State Institute of Cinematography, Danish, Netherlands, St. Petersburg, University of Lulea-Pitea

Central Region, Europe

Sections: Austrian, Belarus, Czech, Central German, North German, South German, Hungarian, Lithuanian, Polish, Slovakian Republic, Swiss, Ukrainian
Student Sections: Berlin, Czech Republic, Darmstadt, Detmold, Düsseldorf, Graz, Ilmenau, Technical University of Gdansk (Poland), Vienna, Wrocław University of Technology

Southern Region, Europe

Sections: Bosnia-Herzegovina, Bulgarian, Croatian, French, Greek, Israel, Italian, Portugal, Romanian, Slovenian, Spanish, Turkish, Yugoslavian
Student Sections: Croatian, Conservatoire de Paris, Italian, Louis-Lumière School

Latin American Region

Sections: Argentina, Brazil, Chile, Colombia (Medellin), Mexico, Uruguay, Venezuela
Student Sections: Taller de Arte Sonoro (Caracas)

International Region

Sections: Adelaide, Brisbane, Hong Kong, India, Japan, Korea, Malaysia, Melbourne, Philippines, Singapore, Sydney

PURPOSE: The Audio Engineering Society is organized for the purpose of: uniting persons performing professional services in the audio engineering field and its allied arts; collecting, collating, and disseminating scientific knowledge in the field of audio engineering and its allied arts; advancing such science in both theoretical and practical applications; and preparing, publishing, and distributing literature and periodicals relative to the foregoing purposes and policies.

MEMBERSHIP: Individuals who are interested in audio engineering may become members of the Society. Applications are considered by the Admissions Committee. Grades and annual dues are: Full member, \$80.00; Associate member, \$80.00; Student member, \$40.00. A membership application form may be obtained from headquarters. Sustaining memberships are available to persons, corporations, or organizations who wish to support the Society. A subscription to the *Journal* is included with all memberships.

AES Journal of the Audio Engineering Society

(ISSN 0004-7554), Volume 50, Number 11, 2002 November
Published monthly, except January/February and July/August when published bi-monthly, by the Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA, Telephone: +1 212 661 8528. Fax: +1 212 682 0477. Periodical postage paid at New York, New York, and at an additional mailing office. Postmaster: Send address corrections to Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520.

The Audio Engineering Society is not responsible for statements made by its contributors.

EDITORIAL STAFF

Daniel R. von Recklinghausen Editor

Patricia M. Macdonald Executive Editor

William T. McQuaide Managing Editor

Gerri M. Calamusa Senior Editor

Abbie J. Cohen Senior Editor

Mary Ellen Ilich Associate Editor

Patricia L. Sarch Art Director

Flávia Elzinga Advertising

Ingeborg M. Stochmal

Copy Editor

Barry A. Blesser

Consulting Technical Editor

Stephanie Paynes Writer

REVIEW BOARD

Ronald M. Aarts

James A. S. Angus

George L. Augspurger

Jeffrey Barish

Jerry Bauck

James W. Beauchamp

Søren Bech

Durand Begault

Barry A. Blesser

John S. Bradley

Robert Bristow-Johnson

John J. Bubbers

Marshall Buck

Mahlon D. Burkhard

Richard C. Cabot

Edward M. Cherry

Robert R. Cordell

Andrew Duncan

John M. Eargle

Louis D. Fielder

Edward J. Foster

Mark R. Gander

Earl R. Geddes

David Griesinger

Malcolm O. J. Hawksford

Jürgen Herre

Tomlinson Holman

Andrew Horner

James D. Johnston

Arie J. M. Kaizer

James M. Kates

D. B. Keele, Jr.

Mendel Kleiner

David L. Klepper

W. Marshall Leach, Jr.

Stanley P. Lipshitz

Robert C. Maher

Dan Mapes-Riordan

J. G. (Jay) McKnight

Guy W. McNally

D. J. Meares

Robert A. Moog

James A. Moorer

Dick Pierce

Martin Polon

D. Preis

Daniel Queen

Francis Rumsey

Kees A. Schouhamer

Immlink

Manfred R. Schroeder

Robert B. Schuelein

Richard H. Small

Julius O. Smith III

Gilbert Soulodre

Herman J. M. Steeneken

John Strawn

G. R. (Bob) Thurmond

Jiri Tichy

Floyd E. Toole

Emil L. Torick

John Vanderkooy

Daniel R. von

Recklinghausen

Rhonda Wilson

John M. Woodgate

Wieslaw V. Woszczyk

COPYRIGHT

Copyright © 2002 by the Audio Engineering Society, Inc. It is permitted to quote from this *Journal* with customary credit to the source.

COPIES

Individual readers are permitted to photocopy isolated articles for research or other noncommercial use. Permission to photocopy for internal or personal use of specific clients is granted by the Audio Engineering Society to libraries and other users registered with the Copyright Clearance Center (CCC), provided that the base fee of \$1.00 per copy plus \$0.50 per page is paid directly to CCC, 222 Rosewood Dr., Danvers, MA 01923, USA. 0004-7554/95. Photocopies of individual articles may be ordered from the AES Headquarters office at \$5.00 per article.

REPRINTS AND REPUBLICATION

Multiple reproduction or republication of any material in this *Journal* requires the permission of the Audio Engineering Society. Permission may also be required from the author(s). Send inquiries to AES Editorial office.

SUBSCRIPTIONS

The *Journal* is available by subscription. Annual rates are \$160.00 surface mail, \$205.00 air mail. For information, contact AES Headquarters.

BACK ISSUES

Selected back issues are available: From Vol. 1 (1953) through Vol. 12 (1964), \$10.00 per issue (members), \$15.00 (nonmembers); Vol. 13 (1965) to present, \$6.00 per issue (members), \$11.00 (nonmembers). For information, contact AES Headquarters office.

MICROFILM

Copies of Vol. 19, No. 1 (1971 January) to the present edition are available on microfilm from University Microfilms International, 300 North Zeeb Rd., Ann Arbor, MI 48106, USA.

ADVERTISING

Call the AES Editorial office or send e-mail to: JournalAdvertising@aes.

MANUSCRIPTS

For information on the presentation and processing of manuscripts, see Information for Authors.

CONTENT

PAPERS

Estimation of Modal Decay Parameters from Noisy Response MeasurementsMatti Karjalainen, Poju Antsalu, Aki Mäkivirta, Timo Peltonen, and Vesa Välimäki	867
<i>Estimating decay parameters, such as room reverberation or string decay of musical instruments, becomes more inaccurate as the noise level increases if that noise is not also included in the parametric model. A new method based on nonlinear optimization of a linear model with additive noise is demonstrated as being more accurate than the traditional approaches, especially at extreme noise conditions. Several practical examples illustrate the utility of the proposed method.</i>	
On the Use of Time–Frequency Reassignment in Additive Sound ModelingKelly Fitz and Lippold Haken	879
<i>The conventional short-time spectral analysis is unable to parameterize signals that have both narrow-band components combined with transients because of the nature of the time–frequency definition. Moreover, a single musical partial, which often has time-varying amplitude and frequency envelopes, does not appear as a single component in the classical linear analysis approach. By reassigning energy components, temporal smear is greatly reduced because unreliable data points can be removed from the representation. Several examples illustrate the ability to reconstruct abrupt onset square waves.</i>	
Scalable Multichannel Coding with HRFT Enhancement for DVD and Virtual Sound SystemsM. O. J. Hawksford	894
<i>Using the perceptual properties embedded in the head-related transfer functions of a normal listener, a standard five- or six-channel audio source can be transformed to feed additional loudspeakers, perhaps as many as 18. The proposed system enhances image stability by optimizing the correct ear signals at the listener’s location. The approach improves spatial resolution over a wider area while retaining backward compatibility with an unprocessed reproduction environment.</i>	
Two-to-Five Channel Sound ProcessingR. Irwan and Ronald M. Aarts	914
<i>In order to reproduce two-channel audio sources in a surround listening environment, the authors evaluated a proposed processing system based on decoding the principle directional component using correlation techniques. By parameterizing direction a stable center channel was created. Moreover, the synthesized ambiance for the sides does not degrade image stability.</i>	

STANDARDS AND INFORMATION DOCUMENTS

AES Standards Committee News	927
<i>New Internet facilities; preservation of recordings; storage and handling; audio connectors; grounding and shielding; file exchange</i>	

FEATURES

113th Convention Report, Los Angeles	934
Exhibitors	950
Program	954
Bluetooth and Wireless Networking—A Primer for Audio Engineers	979
11th Tokyo Regional Convention, Call for Papers	1001

DEPARTMENTS

News of the Sections	985	Membership Information	993
Sound Track	989	Advertiser Internet Directory	999
Upcoming Meetings	990	In Memoriam	1000
New Products and Developments	990	Sections Contacts Directory	1002
Available Literature	992	AES Conventions and Conferences	1008

Estimation of Modal Decay Parameters from Noisy Response Measurements*

MATTI KARJALAINEN,¹ AES Fellow, POJU ANTSALO,¹ AKI MÄKIVIRTA,² AES Member,
TIMO PELTONEN,³ AES Member, AND VESA VÄLIMÄKI,^{1,4} AES Member

¹*Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland*

²*Genelec Oy, Iisalmi, Finland*

³*Akukon Oy, Helsinki, Finland*

⁴*Pori School of Technology and Economics, Tampere University of Technology, Pori, Finland*

The estimation of modal decay parameters from noisy measurements of reverberant and resonating systems is a common problem in audio and acoustics, such as in room and concert hall measurements or musical instrument modeling. Reliable methods to estimate the initial response level, decay rate, and noise floor level from noisy measurement data are studied and compared. A new method, based on the nonlinear optimization of a model for exponential decay plus stationary noise floor, is presented. A comparison with traditional decay parameter estimation techniques using simulated measurement data shows that the proposed method outperforms in accuracy and robustness, especially in extreme SNR conditions. Three cases of practical applications of the method are demonstrated.

0 INTRODUCTION

Parametric analysis, modeling, and equalization (inverse modeling) of reverberant and resonating systems find many applications in audio and acoustics. These include room and concert hall acoustics, resonators in musical instruments, and resonant behavior in audio reproduction systems. Estimating the reverberation time or the modal decay rate are important measurement problems in room and concert hall acoustics [1], where signal-to-noise ratios (SNRs) of only 30–50 dB are common. The same problems can be found, for example, in the estimation of parameters in model-based sound synthesis of musical instruments, such as vibrating strings or body modes of string instruments [2]. Reliable methods to estimate parameters from noisy measurements are thus needed.

In an ideal case of modal behavior, after a possible initial transient, the decay is exponential until a steady-state noise floor is encountered. The parameters of primary interest to be estimated are

- Initial level of decay L_I
- Decay rate or reverberation time T_D
- Noise floor level L_N .

In a more complex case there can be two or more modal frequencies, whereby the decay is no longer simple, but shows additional fluctuation (beating) or a two-stage (or multiple-stage) decay behavior. In a diffuse field (room acoustics) the decay of a noiselike response is approximately exponential in rooms with compact geometry. The noise floor may also be nonstationary. In this paper we primarily discuss a simple mode (that is, a complex conjugate pole pair in the transfer function) or a dense set of modes with exponential reverberant decay, together with a stationary noise floor.

Methods presented in the literature and commonsense or ad hoc methods will first be reviewed. Techniques based on energy–time curve analysis of the signal envelope are known as methods where the noise floor can be found and estimated explicitly. Backward integration of energy, so-called Schroeder integration [3], [4], is often applied first to obtain a smoothed envelope for decay rate estimation.

The effect of the background noise floor is known to be problematic, and techniques have been developed to compensate the effect of envelope flattening when the noise floor in a measured response is reached, including limiting the period of integration [5], subtracting an estimated noise floor energy level from a response [6], or using two separate measurements to reduce the effect of noise [7]. The iterative method by Lundeby et al. [8] is of particular interest since it addresses the case of noisy data with care.

* Presented in the 110th Convention of the Audio Engineering Society, Amsterdam, The Netherlands, 2001 May 12–15; revised 2001 November 8 and 2002 September 6.

This technique, as most other methods, analyzes the initial level L_I , the decay time T_D , and the noise floor L_N separately, typically starting from a noise floor estimate. Iterative procedures are common in accurate estimation.

A different approach was taken by Xiang [9], where a parameterized signal-plus-noise model is fitted to Schroeder-integrated measurement data by searching for a least-squares (LS) optimal solution. In this study we have elaborated a similar method of nonlinear LS optimization further to make it applicable to a wide range of situations, showing good convergence properties. A specific parameter and/or a weighting function can be used to fine-tune the method further for specific problems. The technique is compared with the Lundebly et al. method by it applying to simulated cases of exponential decay plus a stationary noise floor where the exact parameters are known. The improved nonlinear optimization technique is found to outperform traditional methods in accuracy and robustness, particularly in difficult conditions with extreme SNRs.

Finally, the applicability of the improved method is demonstrated by three examples of real measurement data: 1) the reverberation time of a concert hall, 2) the low-frequency mode analysis of a room, and 3) the parametric analysis of guitar string behavior for model-based sound synthesis. Possibilities for further generalization of the technique to more complex problems, such as two-stage decay, are discussed briefly.

1 DEFINITION OF PROBLEM DOMAIN

A typical property of resonant acoustic systems is that their impulse response is a decaying function after a possible initial delay and the onset. In the simplest case the response of a single-mode resonator system is

$$h(t) = Ae^{-\tau(t-t_0)} \sin[\omega(t-t_0) + \phi] u(t-t_0) \quad (1)$$

where $u(t-t_0)$ is a step function with value 1 for $t \geq t_0$ and 0 elsewhere, A is the initial response level, t_0 the response latency, for example, due to the propagation delay of sound, τ the decay rate parameter, $\omega = 2\pi f$ the angular frequency, and ϕ the initial phase of the sinusoidal response. In practical measurements, when there are multiple modes in the system and noise (acoustic noise plus measurement system noise), a measured impulse response is of the form⁵

$$h(t) = \sum_{i=1}^N A_i e^{-\tau_i(t-t_0)} \sin[\omega_i(t-t_0) + \phi_i] A_n n(t) \quad (2)$$

where A_n is the rms value of background noise and $n(t)$ is the unity-level noise signal. Fig. 1 illustrates a single delayed mode response corrupted by additive noise.

The task of this study is defined as finding reliable estimates for the parameter set $\{A_i, \tau_i, t_0, \omega_i, \phi_i, A_n\}$, given a noisy measured impulse response of the form of Eq. (2).

⁵In a more general case the initial delays of signal components may differ and there can be simple nonresonant exponential terms, but these cases are of less importance here.

The main interest here is focused on systems of 1) separable single modes of type (1), including additive noise floor, or 2) the dense (diffuse, noiselike) set of modes resulting also in exponential decay similar to Fig. 1. In both cases the parameters of primary interest are A , τ , A_n , and t_0 .

Often the decay time T_D is of main interest, for example, in room acoustics where the reverberation time [10], [11] of 60-dB decay⁶ T_{60} is related to τ by

$$T_{60} = -\frac{1}{\tau} \ln(10^{-3}) \approx \frac{6.908}{\tau} \quad (3)$$

Modern measurement and analysis techniques of system responses are carried out by digital signal processing whereby the discrete-time formulation for modal decay (without initial delay) with sampling rate f_s and sample period $T_s = 1/f_s$ becomes

$$h(n) = Ae^{-\tau_d n} \sin(\Omega n + \phi) \quad (4)$$

with n being the sample index, $\tau_d = T_s \tau$, and $\Omega = 2\pi T_s f$.

2 DECAY PARAMETER ESTIMATION

In this section an overview of the known techniques for decay parameter estimation will be presented. Initial delay

⁶In practice, the reverberation time is often determined from the slope of a decay curve using only the first 25 or 35 dB of decay and extrapolating the result to 60 dB. For a recommended practice of reverberation time determination, see [1].

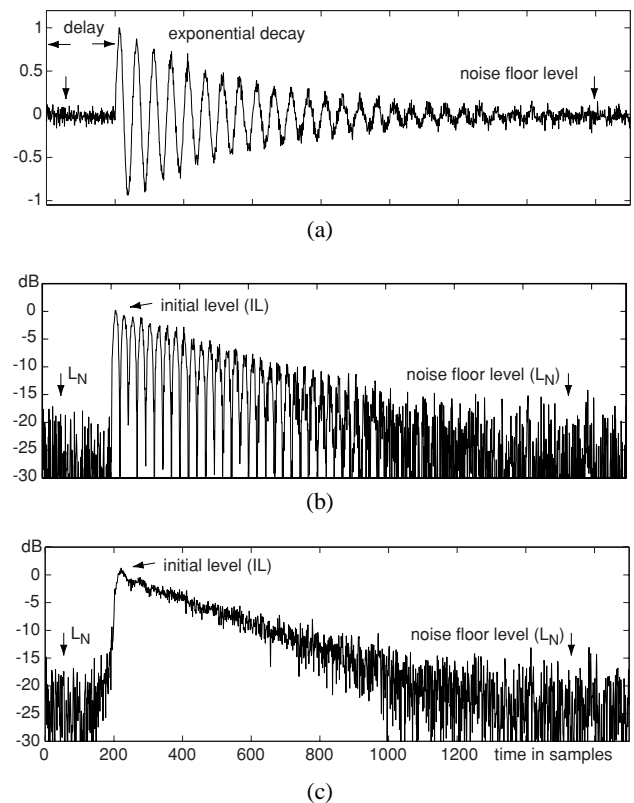


Fig. 1. (a) Single-mode impulse response (sinusoidal decay) with initial delay and additive measurement noise. (b) Absolute value of response on dB scale to illustrate decay envelope. (c) Hilbert envelope; otherwise same as (b).

and level estimation are first discussed briefly. The main problem, decay rate estimation, is the second topic. Methods to smooth the decay envelope from a measured impulse response are presented. Noise floor estimation, an important subproblem, is discussed next. Finally, techniques for combined noise floor and decay rate estimation are reviewed.

2.1 Initial Delay and Initial Level Estimation

In most cases the initial delay and the initial level parameters are relatively easy to estimate. The initial delay may be short, not needing any attention, or the initial bulk delay can be cut off easily up to the edge of response onset. Only when the onset is relatively irregular or the SNR is low, can the detection of onset time be difficult.

A simple technique to eliminate initial delay is to compute the minimum-phase component $h_{\text{mphase}}(t)$ of the measured response [12]. An impulse response can be decomposed as a sum of minimum-phase and excess-phase components, $h(t) = h_{\text{mphase}}(t) + h_{\text{ephase}}(t)$. Since the excess-phase component will have all-pass properties manifested as a delay, computation of the minimum-phase part will remove the initial delay.

The initial level in the beginning of the decay can be detected directly from the peak value of the onset. For improved robustness, however, it may be better to estimate it from the matched decay curve, particularly its value at the onset time.

In the case of a room impulse response, the onset corresponds to direct sound from the sound source. It may be of special interest for the computation of the source-to-receiver distance or in estimating the impulse response of the sound source itself by windowing the response prior to the first room reflection.

2.2 Decay Rate Estimation

Decay rate or time estimation is in practice based on fitting a straight line to the decay envelope, such as the energy–time curve, mapped on a logarithmic (dB) scale. Before the computerized age this was done graphically on paper. The advantage of manual inspection is that an expert can avoid data interpretation errors in pathological cases. However, in practice the automatic determination of the decay rate or time is highly desirable.

2.2.1 Straight-Line Fit to Log Envelope

Fitting a line in a logarithmic decay curve is a conceptually and computationally simple way of decay rate estimation. The decay envelope $y(t)$ can be computed simply as a dB-scaled energy–time curve,

$$y(t) = 10 \log_{10} [x^2(t)] \quad (5)$$

where $x(t)$ is the measured impulse response or a band-pass filtered part of it, such as an octave or one-third-octave band. It is common to apply techniques such as Schroeder integration and Hilbert envelope computation (to be described later) in order to smooth the decay curve before line fitting. Least-squares line fitting (linear regression) is done by finding the optimal decay rate k and the

initial level a ,

$$\min_{k, a} \int_{t_1}^{t_2} [y(t) - (kt + a)]^2 dt \quad (6)$$

using, for example, the Matlab function `polyfit` [13].

Practical problems with line fitting are related to the selection of the interval $[t_1, t_2]$ and cases where the decay of the measured response is inherently nonlinear. The first problem is avoided by excluding onset transients in the beginning and the noise floor at the end of the measurement interval. The second problem is related to such cases as two-stage decay (initial decay rate or early reverberation and late decay rate or reverberation) or beating (fluctuation) of the envelope because of two modes close in frequency [see Fig. 9(b)].

2.2.2 Nonlinear Regression (Xiang's Method)

Xiang [9] formulated a method where a measured and Schroeder-integrated energy–time curve is fitted to a parametric model of a linear decay plus a constant noise floor. Since the model is not linear in its parameters, nonlinear curve fitting (nonlinear regression) is needed. Mathematically, this is done by iterative means such as starting from a set of initial values for the model parameters and applying gradient descent to search for a least-squares optimum,

$$\min_{x_1, x_2, x_3} \int_{t_1}^{t_2} \left\{ y_{\text{sch}}(t) - [x_1 e^{-x_2 t} + x_3(L - t)] \right\}^2 dt \quad (7)$$

where $y_{\text{sch}}(t)$ is the Schroeder-integrated energy envelope, x_1 the initial level, x_2 the decay rate parameter, x_3 a noise floor related parameter, L the length of the response, and $[t_1, t_2]$ the time interval of nonlinear regression. Notice that the last term for the noise floor effect is a descending line, instead of a constant level, due to the backward integration of noise energy [9].

Nonlinear optimization is mathematically more complex than linear fitting, and care should be taken to guarantee convergence. Even when converging, the result may be only a local optimum, and generally the only way to know that a global optimum is found is to apply exhaustive search over possible value combinations of model parameters which, in a multiparameter case, is often computationally too expensive.

Nonlinear optimization techniques will be studied in more detail later in this paper by introducing generalizations to the method of Xiang and by comparing the performance of different techniques in decay parameter estimation.

2.2.3 AR and ARMA Modeling

For a single mode of Eq. (1) the response can be modeled as an impulse response of a resonating second-order all-pole or pole–zero filter. More generally, a combination of N modes can be modeled as a $2N$ -order filter. AR (autoregressive) and ARMA (autoregressive moving average) modeling [14] are ways to derive parameters for such models. In many technical applications the AR method is called linear prediction [15]. For example, the function `lpc` in Matlab [16] processes a signal frame

through autocorrelation coefficient computation and solving normal equations by Levinson recursion, resulting in the N th-order z -domain transfer function $1/(1 + \sum_{i=1}^N \alpha_i z^{-i})$. Poles are obtained by solving the roots of the denominator polynomial. Each modal resonance appears as a complex conjugate pole pair (z_i, z_i^*) in the complex z -plane with pole angle $\phi = \arg(z_i) = 2\pi f/f_s$ and pole radius $r = |z_i| = e^{-\tau/f_s}$, where f is the modal frequency, f_s the sampling rate, and τ the decay parameter of the mode in Eq. (1). ARMA modeling requires an iterative solution for a pole-zero filter.

Decay parameter analysis by AR and ARMA modeling is an important technique and attractive, for example, in cases where modes overlapping or very close to each other have to be modeled, which is often difficult by other means. For reverberation with high modal density the order of AR modeling may become too high for accurate modeling. Such accuracy is also not necessary for analyzing the overall decay rate (reverberation time) only. AR and ARMA modeling of modal behavior in acoustic systems are discussed in detail, for example, in [17].

2.2.4 Group Delay Analysis

A complementary method to AR modeling is to use the group delay, that is, the phase derivative $T_g(\omega) = -d\phi(\omega)/d\omega$, as an estimate of the decay time for separable modes of an impulse response. While AR modeling is sensitive to the power spectrum only, the group delay is based on phase properties only. For a minimum-phase single-mode response the group delay at the modal frequency is inversely proportional to the decay parameter, that is, $T_g = 1/\tau$. Group delay computation is somewhat critical due to the phase unwrapping needed, and the method can be sensitive to measurement noise.

2.3 Decay Envelope Smoothing Techniques

In the methods of linear or nonlinear curve fitting it is desirable to obtain a smooth decay envelope prior to the fitting operation. The following techniques are often used to improve the regularity of the decay ramp.

2.3.1 Hilbert Envelope Computation

In this method the signal $x(t)$ is first converted to an analytic signal $x_a(t)$ so that $x(t)$ is the real part of $x_a(t)$ and the Hilbert transform (90° phase shift) [12] is the imaginary part of $x_a(t)$. For a single sinusoid this results in an entirely smooth energy-time envelope. An example of a Hilbert envelope for a noisy modal response is shown in Fig. 1(c).

2.3.2 Schroeder Integration

A monotonic and smoothed decay curve can be produced by backward integration of the impulse response $h(t)$ over the measurement interval $[0, T]$ and converting it to a logarithmic scale,

$$L(t) = 10 \log_{10} \left[\frac{\int_t^T h^2(\tau) d\tau}{\int_0^T h^2(\tau) d\tau} \right] \text{ [dB]} . \tag{8}$$

This process is commonly known as Schroeder integration [3], [4]. Based on its superior smoothing properties it is used routinely in modern reverberation time measurements. A known problem with it is that if the background noise floor is included within the integration interval, the process produces a raised ramp that biases upward the late part of the decay. This is shown in Fig. 2 for the case of noisy single-mode decay [curve (a)] for full response integration [curve (d)].

The tail problem of Schroeder integration has been addressed by many authors (for example, in [18], [8], [5], [6]), and techniques to reduce slope biasing have been proposed. In order to apply these improvements, a good estimate of the noise floor level is needed first.

2.4 Noise Floor Level Estimation

The limited SNR inherent in practically all acoustical measurements, and especially measurements performed under field conditions, calls for attention concerning the upper time limit of decay curve fitting or Schroeder integration. Theoretically this limit is set to infinity, but in practical measurements it is naturally limited to the length of the measured impulse response data. In practice, measured impulse responses must be long enough to accommodate a large enough dynamic range or the whole system decay down to the background noise level.⁷

Thus the measured impulse response typically contains not only the decay curve under analysis, but also a steady level of background noise, which dominates at the end of

⁷This is needed to avoid time aliasing in MLS and other cyclic impulse response measurement methods.

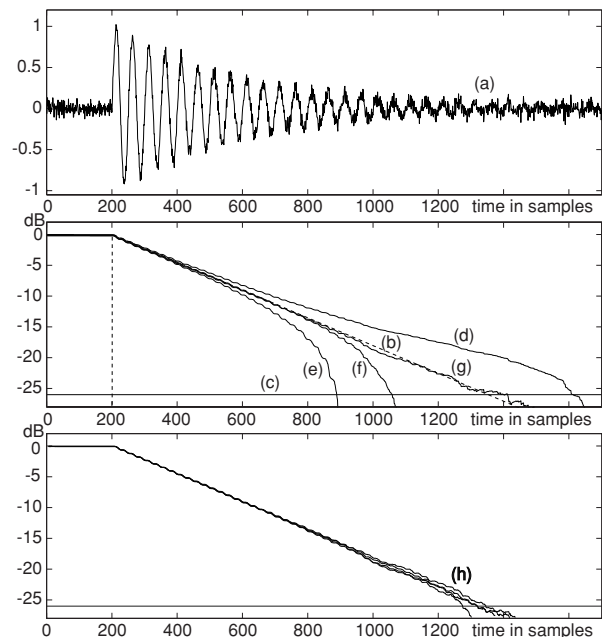


Fig. 2. Results of Schroeder integration applied to noisy decay of a mode. Curve (a) measured noisy response including initial delay; curve (b) true decay of noiseless mode (dashed straight line); curve (c) noise floor (-26 dB); curve (d) Schroeder integration of total measured interval; curve (e) integration over short interval (0, 900 ms); curve (f) integration over interval (0, 1100 ms), curve (g) integration after subtracting noise floor from energy-time curve; curve (h) a few decay curves integrated by Hirata's method.

the response. Fitting the decay line over this part of the envelope or Schroeder integrating this steady energy level along with the exponential decay curve causes an error both in the resulting decay rate (see Fig. 2) and in the time-windowed energies (energy parameters).

To avoid bias by noise, an analysis must be performed on the impulse response data to find the level of background noise and the point where the room decay meets the noise level. This way it is possible to effectively truncate the impulse response at the noise level, minimizing the noise energy mixed with the actual decay.

Determination of the noise floor level is difficult without using iterative techniques. The method by Lundeby et al., which will be outlined later, is a good example of iterative techniques integrated with decay rate estimation.

A simple way to obtain a reasonable estimate of a background noise floor is to average a selected part of the measured response tail or to fit a regression line to it [19]. The level is certainly overestimated if the noise floor is not reached, but this is not necessarily problematic, as opposed to underestimating it. Another technique is to look at the background level before the onset of the main response. This works if there is enough initial latency in the system response under study.

2.5 Decay Estimation with Noise Floor Reduction

In addition to determining the response starting point, it is thus essential to find an end point where the decay curve meets the background noise, and to truncate the noise from the end of the response. Fig. 2 illustrates the effect of limiting the Schroeder integration interval. If the interval is too short, as in curve (e), the curve is biased downward. Curve (f) shows a case where the bias due to noise is minimized by considering the decay only down to 10 dB above the noise floor.

There are no standardized exact methods for determining the limits for Schroeder integration and decay fitting or noise compensation techniques. The methods are discussed next.

2.5.1 Limited Integration or Decay Matching Interval

There are several recommendations for dealing with the noise floor and the point where the decay meets noise. For example, according to ISO 3382 [1] to determine room reverberation, the noise floor must be 10 dB below the lowest decay level used for the calculation of the decay slope. Morgan [5] recommends to truncate at the knee point and then measure the decay slope of the backward integrated response down to a level 5 dB above the noise floor.

Faiget et al. [19] propose a simple but systematic method for postprocessing noisy impulse responses. The latter part of a response is used for the estimation of the background noise level by means of a regression line. Another regression line is used for the decay, and the end of the useful response is determined at the crossing point of the decay and the background noise regression lines.

The decay parameter fitting interval ends 5 dB above the noise floor.

2.5.2 Lundeby's Method

Lundeby et al. [8] presented an algorithm for automatically determining the background noise level, the decay-noise truncation point, and the late decay slope of an impulse response. The steps of the algorithm are as follows.

1) The squared impulse response is averaged into local time intervals in the range of 10–50 ms to yield a smooth curve without losing short decays.

2) A first estimate for the background noise level is determined from a time segment containing the last 10% of the impulse response. This gives a reasonable statistical selection without a large systematic error if the decay continues to the end of the response.

3) The decay slope is estimated using linear regression between the time interval containing the response 0-dB peak and the first interval 5–10 dB above the background noise level.

4) A preliminary crosspoint is determined at the intersection of the decay slope and the background noise level.

5) A new time interval length is calculated according to the calculated slope, so that there are 3–10 intervals per 10 dB of decay.

6) The squared impulse is averaged into the new local time intervals.

7) The background noise level is determined again. The evaluated noise segment should start from a point corresponding to 5–10 dB of decay after the crosspoint, or a minimum of 10% of the total response length.

8) The late decay slope is estimated for a dynamic range of 10–20 dB, starting from a point 5–10 dB above the noise level.

9) A new crosspoint is found.

Steps 7–9 are iterated until the crosspoint is found to converge (maximum five iterations).

The response analysis may be further enhanced by estimating the amount of energy under the decay curve after the truncation point. The measured decay curve is artificially extended beyond the point of truncation by extrapolating the regression line on the late decay curve to infinity. The total compensation energy is formed as an ideal exponential decay process, the parameters of which are calculated from the late decay slope.

2.5.3 Subtraction of Noise Floor Level

Chu [18] proposed a subtraction method in which the mean square value of the background noise is subtracted from the original squared impulse response before the backward integration. Curve (g) in Fig. 2 illustrates this case. If the noise floor estimate is accurate and the noise is stationary, the resulting backward integrated curve is close to the ideal decay curve.

2.5.4 Hirata's Method

Hirata [7] has proposed a simple method for improving the signal-to-noise ratio by replacing the squared single impulse response $h^2(t)$ with the product of two impulse

responses measured separately at the same position,

$$\begin{aligned}
 \int_t^\infty h^2(t) dt &\Leftarrow \int_t^\infty [h_1(t) + n_1(t)][h_2(t) + n_2(t)] dt \\
 &= \int_t^\infty [h_1(t)h_2(t) + h_1(t)n_2(t) + h_2(t)n_1(t) + n_1(t)n_2(t)] dt \\
 &= \int_t^\infty \left\{ h_1(t)h_2(t) + n_1(t)n_2(t) \left[1 + \frac{h_1(t)}{n_1(t)} + \frac{h_2(t)}{n_2(t)} \right] \right\} dt \\
 &\approx \int_t^\infty h^2(t) dt + K(t) .
 \end{aligned} \tag{9}$$

The measured impulse responses consist of the decay terms $h_1(t)$, $h_2(t)$ and the noise terms $n_1(t)$, $n_2(t)$. The highly correlated decay terms $h_1(t)$ and $h_2(t)$ yield positive values corresponding to the squared response $h^2(t)$, whereas the mutually uncorrelated noise terms $n_1(t)$ and $n_2(t)$ are seen as a random fluctuation $K(t)$ superposed on the first term. Hirata’s method relies on the impulse response at large time values to be stationary. This condition is often not met in practical concert hall measurements.

Curves (h) in Fig. 2 illustrate a few decay curves obtained by backward integration with Hirata’s method. In this simulated case they correspond approximately to the case of curve (g), the noise floor subtraction technique.

2.5.5 Other Methods

Under adverse noise conditions, a direct determination of the T_{30} decay curve from the squared and time-averaged impulse response has been noted to be more robust than the backward integration method (Satoh et al. [20]).

3 NONLINEAR OPTIMIZATION OF A DECAY-PLUS-NOISE MODEL

The nonlinear regression (optimization) method proposed by Xiang [9] was briefly described earlier. In the present study we worked along similar ideas, using nonlinear optimization for improved robustness and accuracy. In the following we introduce the nonlinear decay-plus-noise model and its application in several cases.

Let us assume that in noiseless conditions the system under study results in a simple exponential decay of the response envelope, corrupted by additive stationary background noise. We will study two cases that fit into the same modeling category. In the first case there is a single mode (a complex conjugate pole pair in transfer function) that in the time domain corresponds to an exponential decay function,

$$h_m(t) = A_m e^{-\tau_m t} \sin(\omega_m t + \phi_m) . \tag{10}$$

Here A_m is the initial envelope amplitude of the decaying sinusoidal, τ_m is a coefficient that defines the decay rate, ω_m is the angular frequency of the mode, and ϕ_m is the initial phase of modal oscillation.

The second case that leads to a similar formulation is where we have a high density of modes (diffuse sound field) with exponential decay, resulting in an exponen-

tially decaying noise signal,

$$h_d(t) = A_d e^{-\tau_d t} n(t) \tag{11}$$

where A_d is the initial rms level of the response, τ_d is a decay rate parameter, and $n(t)$ is stationary Gaussian noise with an rms level of 1 (= 0 dB).

In both Eqs. (10) and (11) we assume that a practical measurement of the system impulse response is corrupted with additive stationary noise,

$$n_b(t) = A_n n(t) \tag{12}$$

where A_n is the rms level of the Gaussian measurement noise in the analysis bandwidth of interest, and it is assumed to be uncorrelated with the decaying system response. Statistically the rms envelope of the measured response is then

$$a(t) = \sqrt{h^2(t) + n_b^2(t)} = \sqrt{A^2 e^{-2\tau t} + A_n^2} . \tag{13}$$

This is a simple decay model that can be used for parametric analysis of noise-corrupted measurements. If the amplitude envelope of a specific measurement is $y(t)$, then an optimized least-squares (LS) error estimate for the parameters $\{A, \tau, A_n\}$ can be achieved by minimizing the following expression over a time span $[t_0, t_1]$ of interest:

$$\min_{A, \tau, A_n} \int_{t_0}^{t_1} [a(t) - y(t)]^2 dt \tag{14}$$

Since the model of Eq. (13) is nonlinear in the parameters $\{A, \tau, A_n\}$, nonlinear LS optimization is needed to search for the minimum LS error.

By numerical experimentation with real measurement data it is easy to observe that LS fitting of the model of Eq. (14) places emphasis on large magnitude values, whereby noise floors well below the signal starting level are estimated poorly. In order to improve the optimization, a generalized form of model fitting can be formulated as

$$\min_{A, \tau, A_n} \int_{t_0}^{t_1} [f(a(t), t) - f(y(t), t)]^2 dt \tag{15}$$

where $f(y, t)$ is a mapping with balanced weight for different envelope level values and time moments.

The choice of $f(y, t) = 20 \log_{10}[y(t)]$ results in fitting

on the dB scale. It turns out that low-level noise easily has a dominating role in this formulation. A better result in model fitting can be achieved by using a power law scaling $f(y, t) = y^s(t)$ with the exponent $s < 1$, which is a compromise between amplitude and logarithmic scaling. A value of $s = 0.5$ has been found to be a useful default value.⁸

A time-dependent part of mapping $f(y, t)$, if needed, can be separated as a temporal weighting function $w(t)$. A generalized form of the entire optimization is now to find

$$\min_{A, \tau, A_n} \int_{t_0}^{t_1} \left[w(t) a^s(t) - w(t) y^s(t) \right]^2 dt. \quad (16)$$

There is no clear physical motivation for the magnitude compression exponent s . A specific temporal weighting function $w(t)$ can be applied case by case, based on extra knowledge of the behavior of the system under study and goals of the analysis, such as focusing on the early decay time (early reverberation) of a room response.

The strengths of the nonlinear optimization method are apparent, especially under extreme SNR conditions where all three parameters $\{A, \tau, A_n\}$ are needed with greatest accuracy. This occurs both at very low SNR conditions where the signal is practically buried in background noise and at the other extreme where the noise floor is not reached within the measured impulse response, but an estimate of the noise level is nevertheless desired. A necessary assumption for the method to work in such cases is that the decay model is valid, implying an exponential decay and a stationary noise floor.

Experiments show that the model is useful for both single-mode decay and reverberant acoustic field decay models. Fig. 3 depicts three illustrative examples of decay model fitting to a single mode plus noise at an initial level of 0 dB and different noise floor levels. Because of simulated noisy responses it is easy to evaluate the estimation accuracy of each parameter. White curves show the estimated behavior of the decay-plus-noise model. In Fig. 3(a) the SNR is only 6 dB. Errors in the parameters in this case are a 0.5-dB underestimate of A , a 3.5% underestimate in the decay time related to parameter τ , and a 1.8-dB overestimate of the noise floor A_n . In Fig. 3(b) a similar case is shown with a moderate 30-dB SNR. Estimation errors of the parameters are +0.2 dB for A , -2.8% for decay time, and +1.2 dB for A_n . In the third case [Fig. 3(c)] the SNR is -60 dB so that the noise floor is barely reached within the analysis window. In this case the estimation errors are +0.002 dB for A , -0.07% for decay time, and -1.0 dB for A_n . This shows that the noise floor is estimated with high accuracy, even in this extreme case.

The nonlinear optimization used in this study is based on using the Matlab function `curvefit`,⁹ and the functions that implement the weighting by parameter s and the weighting function $w(t)$ can be found at <http://www.acoustics.hut.fi/software/decay>.

[acoustics.hut.fi/software/decay](http://www.acoustics.hut.fi/software/decay).

The optimization routines are found converging robustly in most cases, including such extreme cases as Fig. 3(a) and (c), and the initial values of the parameters for iteration are not critical. However, it is possible that in rare cases the optimization diverges and no (not even a local) optimum is found.¹⁰ It would be worth working out a dedicated optimization routine guaranteeing a result in minimal computation time.

Our experience in the nonlinear decay parameter fitting described here is that it still needs some extra information or top-level iteration for the very best results. It is advantageous to select the analysis frame so that the noise floor is reached neither too early nor too late. If the noise floor is reached in the very beginning of the frame, the decay may be missed. Not reaching the noise floor in the frame is a problem only if the estimate of this level is important. A rule for an optimal value of the scaling parameter s is to use $s \approx 1.0$ for very low SNRs such as in Fig. 3(a), and let it approach a value of 0.4–0.5 when the noise floor is low, as in Fig. 3(c) (see also Fig. 4).

4 COMPARISON OF DECAY PARAMETER ESTIMATION METHODS

The accuracy and robustness of the methods for decay parameter estimation can be evaluated by using synthetic

¹⁰The function `curvefit` also prints warnings of computational precision problems even when optimization results are excellent.

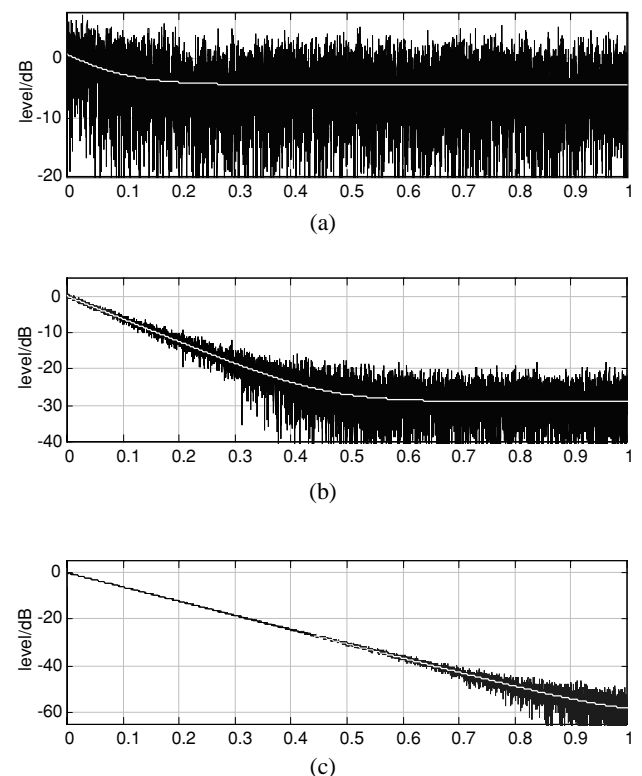


Fig. 3. Nonlinear optimization of decay-plus-noise model for three synthetic noisy responses with initial level of 0 dB and various noise levels. (a) -6 dB. (b) -30 dB. (c) -60 dB. Black curves—Hilbert envelopes of simulated responses; white curves—estimated decay behavior.

⁸Interestingly enough, this resembles the loudness scaling in auditory perception known from psychoacoustics [21].

⁹In new versions of Matlab, the function `curvefit` is recommended to be replaced by the function `lsqcurvefit`.

decay signals or envelope curves, computed for sets of the parameters $\{A, \tau, A_n\}$. By repeating the same for different methods, their relative performances can be compared. In this section we present results from a comparison of the proposed nonlinear optimization and the method of Lundeby et al. [8].

The accuracy of the two methods was analyzed in the following setting. A decaying sinusoid of 1 kHz with a 60-dB decay time (reverberation time) of 1 second was contaminated with white noise of Gaussian distribution and zero mean. The initial sinusoidal level to background noise ratio was varied from 0 to 80 dB in steps of 10 dB. Each method under study was applied to analyze the decay parameters, and the error to the “true” value was computed in dB for the initial and the noise floor levels and as a percentage of decay time.

Fig. 4 depicts the results of the evaluation for the nonlinear optimization proposed in this paper. The accuracy of the decay time estimation in Fig. 4(a) is excellent for SNRs above 30 dB and useful (below 10% typically) even for SNRs of 0–10 dB. The initial level is accurate within 0.1 dB for an SNR above 20 dB and about 1 dB for an SNR of 0 dB. The noise floor estimate is within approximately 1–2 dB up to an SNR of 60 dB and gives better than a guess up to 70–80 dB of SNR. (Notice that the SNR alone is not important here but rather whether or not the noise floor is reached in the analysis window.)

Fig. 5 plots the same information for decay parameter estimation using the method of Lundeby et al. without noise compensation, implemented by us in Matlab. Since this iterative technique is not developed for extreme SNRs, such as 0 dB, it cannot deal with these cases without extra tricks, and even then it may have severe problems. We used safety settings whereby we did not try to

obtain decay time values for SNRs below 20 dB, and low SNR parts of the decay parameter estimate curves are omitted.

For moderate SNRs the results of the method are fairly good and robust. The decay time shows a positive bias of a few percent, except for an SNR below 30 dB. The noise floor estimate is reliable in this case only up to about 50 dB SNR. Notice that the method is designed for practical reverberation time measurements rather than for this test case, where it could be tuned to perform better.

5 EXAMPLES OF DECAY PARAMETER ESTIMATION BY NONLINEAR OPTIMIZATION

In this section we present examples of applying the nonlinear estimation of a decay-plus-noise model to typical acoustic and audio applications, including reverberation time estimation, analysis and modeling of low-frequency modes of a room response, and decay rate analysis of plucked string vibration for model-based synthesis applications.

5.1 Reverberation Time Estimation

Estimating the reverberation time of a room or a hall is relatively easy if the decay curve behaves regularly and the noise floor is low enough. In practice the case is often quite different. Here we demonstrate the behavior of the nonlinear optimization method in an example where the measured impulse response includes an initial delay, an irregular initial part, and a relatively high measurement noise floor.

Fig. 6 depicts three different cases of fitting the decay-plus-noise model to this case of a control room with a short reverberation time. In Fig. 6(a) the fitting is applied

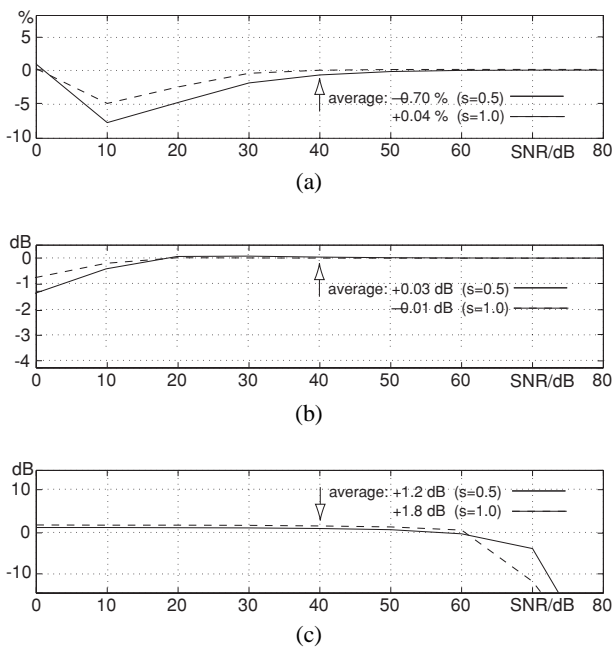


Fig. 4. Sine-plus-noise decay parameter estimation errors (average of 20 trials) for proposed nonlinear optimization method as a function of SNR. (a) Decay time estimation error in %. (b) Initial level estimation error in dB. (c) Noise floor estimation error in dB. — $s = 0.5$; --- $s = 1.0$.

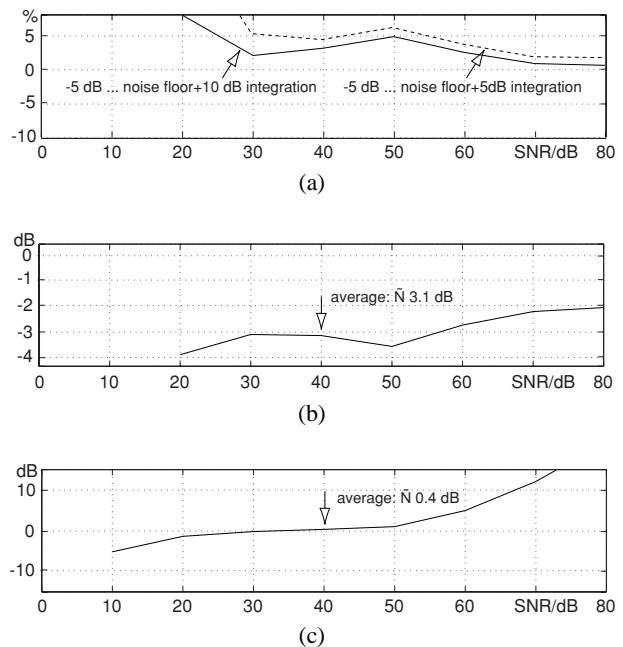


Fig. 5. Decay parameter estimation errors for Lundeby et al. method as a function of SNR. (a) Decay time estimation error in % with truncated Schroeder integration but without noise compensation. (b) Initial level estimation error in dB. (c) Noise floor estimation error in dB.

to the entire decay curve, including the initial delay, and the resulting model is clearly biased toward too long a reverberation time. In Fig. 6(b) the initial delay is excluded from model fitting, and the result is better. However, after the direct sound there is a period of only little energy during the first reflections prior to the range of dense reflections and diffuse response. If the reverberation time estimate is to describe the decay of this diffuse part, the case of Fig. 6(c), with a fitting starting from about 30 ms, yields the best match to reverberation decay,¹¹ and the approaching noise floor is also estimated well.

5.2 Modeling of Low-Frequency Room Modes

The next case deals with the modeling of the low-frequency modes of a room. Below a critical frequency (the so-called Schroeder frequency) the mode density is low and individual modes can be decomposed from the measured room impulse response. The task here was to find the most prominent modes and to analyze their modal parameters f_m and τ_m , the frequency and decay parameters, respectively. The case studied was a hard-walled, partially damped room with moderate reverberation time (≈ 1 second) at mid and high frequencies, but much longer decay times at the lowest modal frequencies. The following procedure was applied:

- A short-time Fourier analysis of the measured impulse response was computed to yield the time–frequency representation shown in Fig. 7 as a waterfall plot.
- At each frequency bin (1.3-Hz spacing is used) the dB-scaled energy–time decay trajectory was fitted to the decay-plus-noise model with the nonlinear optimization technique to obtain the optimal decay parameter τ .

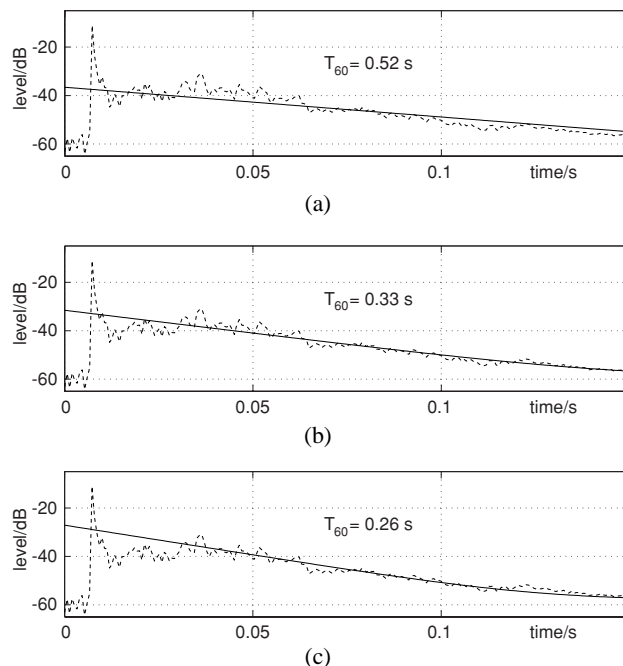


Fig. 6. Decay-plus-noise model fitting by nonlinear optimization to a room impulse response. (a) Fitting range includes initial delay, transient phase, and decay. (b) Fitting includes transient phase and decay. (c) Fitting includes only decay phase. Estimated T_{60} values are given.

- Based on decay parameter values and spectral levels, a rule was written to pick up the most prominent modal frequencies and the related decay parameter values.

In this context we are interested in how well the decay parameter estimation worked with noisy measurements. Application of the nonlinear optimization resulted in decay curve fits, some of which are illustrated in Fig. 8, by comparing the original decay and the decay-plus-noise model behavior. For all frequencies in the vicinity of a mode the model fits robustly and accurately.¹²

5.3 Analysis of Decay Rate of Plucked String Tones

A model-based synthesis of string tones can produce realistic guitar-like tones if the parameter values of the synthesis model are calibrated based on recordings [2]. The main properties of tones that need to be analyzed are their fundamental frequency and the decay time of all harmonic partials that are audible. While estimating the fundamental frequency is quite easy, measurement of the decay times of harmonics (= modes of the string) is complicated by the fact that they all have a different rate of decay and also the initial level can vary within a range of 20–30 dB. There may also be no information about the noise floor level for all harmonics.

One method used for measuring the decay times is based on the short-time Fourier analysis. A recorded single guitar tone is sliced into frames with a window function in the time domain. Each window function is then Fourier transformed with the fast Fourier transform using zero padding to increase the spectral resolution, and harmonic peaks are hunted from the magnitude spectrum using a peak-picking algorithm. The peak values from the consecutive frames are organized as tracks, which correspond to the temporal envelopes of the harmonics. Then it becomes possible to estimate the decay rate of each harmonic mode. In the following, we show how this works with the proposed decay parameter estimation algorithm. Finally the decay rate of each harmonic is converted into a corresponding target response, which is used for designing the magnitude response of a digital filter that controls the decay of harmonics in the synthesis model.

Fig. 9 plots three examples of modal decay analysis of guitar string harmonics (string 5, open string). Harmonic envelope trajectories were analyzed as described. The decay-plus-noise model was fitted in a time window that started from the maximum value position of the envelope curve. In Fig. 9(a) the second harmonic shows a highly regular decay after an initial transient of plucking, whereby decay fitting is almost perfect. Fig. 9(b), harmonic 24,

¹¹In this example, decay parameter analysis is applied to the entire frequency range of the impulse response. In practice it is computed as a function of frequency, that is, applied to octave or one-third-octave band decay curves.

¹²To obtain the best frequency resolution it may be desirable to replace short-time spectral analysis with energy-decay curves obtained by backward integration and the model of Eq. (16) with a corresponding formulation derived from Xiang's formula, Eq. (7).

depicts a strongly beating decay, where probably the horizontal and vertical polarizations have a frequency difference that after summation results in beating. Fig. 9(c), harmonic 54, shows a trajectory where the noise floor is reached within the analysis window. In all cases shown the nonlinear optimization works as perfectly as a simple decay model can do.

As can be concluded from Fig. 9(b), a string can exhibit more complicated behavior than simple exponential decay. Even more complex is the case of piano tones because there are two to three strings slightly off tune, and the envelope fluctuation can be more irregular. Two-stage decay is also common where the initial decay is faster than later decay [22].

In all such cases a more complex decay model is needed to achieve a good match with measured data. Such techniques are studied in [17].

6 SUMMARY AND CONCLUSIONS

An overview of modal decay analysis methods for noisy impulse response measurements of reverberant acoustic systems has been presented, and further improvements were introduced. The problem of decay time determination is important, for example, in room acoustics for characterizing the reverberation time. Another application where a similar problem is encountered is the estimation of string model parameters for model-based synthesis of

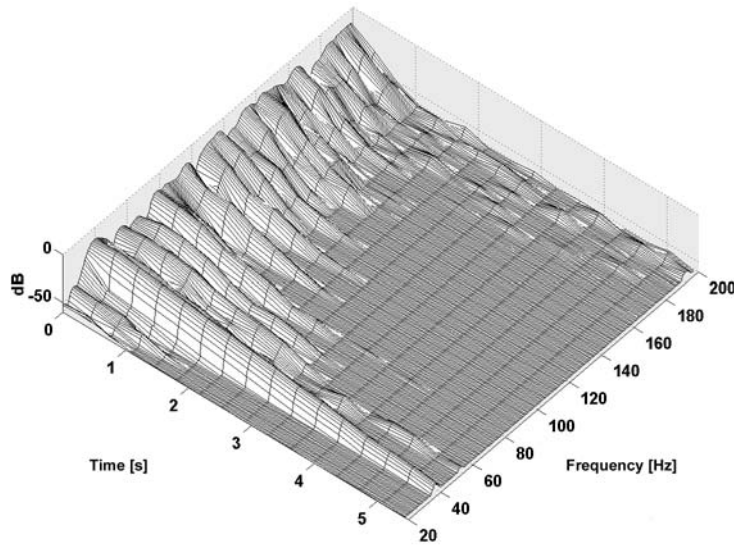


Fig. 7. Time–frequency plot of room response at low frequencies. Lowest modes (especially 40 Hz) show long decay times.

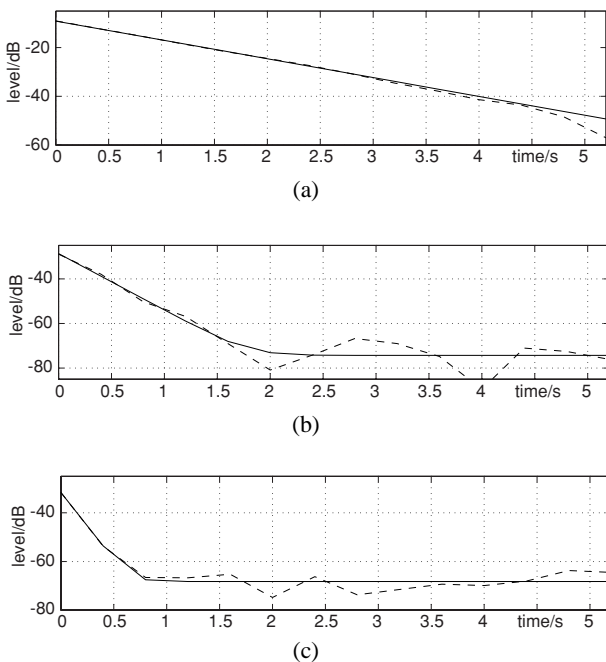


Fig. 8. Fitting of decay-plus-noise model to low-frequency modal data of room (see Fig. 7). (a) At 40-Hz. (b) At 104 Hz. (c) At 117 Hz (off-mode fast decay). --- measured; — modeled.

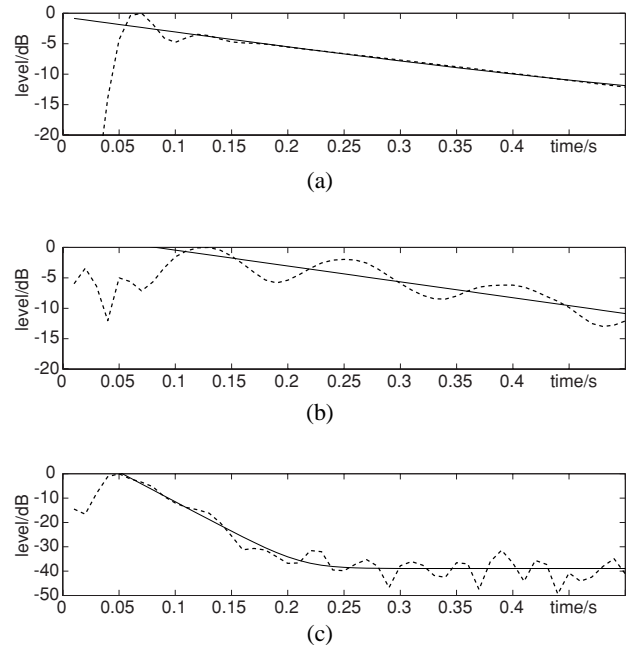


Fig. 9. Examples of modal decay matching for harmonic components of guitar string. (a) Regular decay after initial transient. (b) Strongly beating decay (double mode). (c) Fast decay that reaches noise floor. --- measured envelope; — optimized model fit.

plucked string instruments.

It is shown that the developed decay-plus-noise model yields highly accurate decay parameter estimates, outperforming traditional methods, especially under extreme SNR conditions.

There exist other methods, such as AR modeling, that show potential in specific applications. Challenges for further research are to make modal decay methods (with an increased number of parameters) able to analyze complex decay characteristics, such as double decay behavior and strongly fluctuating responses due to two or more modes very close in frequency.

A Matlab code for the nonlinear optimization of decay parameters, including data examples, can be found at <http://www.acoustics.hut.fi/software/decay>.

7 ACKNOWLEDGMENT

This study is part of the VÄRE technology program, project TAKU (Control of Closed Space Acoustics), funded by Tekes (National Technology Agency). The work of Vesa Välimäki has been financed by the Academy of Finland.

8 REFERENCES

[1] ISO 3382-1997, "Acoustics—Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters," International Standards Organization, Geneva, Switzerland (1997), 21 pp.

[2] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical Modeling of Plucked String Instruments with Application to Real-Time Sound Synthesis," *J. Audio Eng. Soc.*, vol. 44, pp. 331–353 (1996 May).

[3] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412 (1965).

[4] M. R. Schroeder, "Integrated-Impulse Method Measuring Sound Decay without Using Impulses," *J. Acoust. Soc. Am.*, vol. 66, pp. 497–500 (1979 Aug.).

[5] D. Morgan, "A Parametric Error Analysis of the Backward Integration Method for Reverberation Time Estimation," *J. Acoust. Soc. Am.*, vol. 101, pp. 2686–2693 (1997 May).

[6] L. Faiget, C. Legros, and R. Ruiz, "Optimization of the Impulse Response Length: Application to Noisy and

Highly Reverberant Rooms," *J. Audio Eng. Soc.*, vol. 46, pp. 741–750 (1998 Sept.).

[7] Y. Hirata, "A Method of Eliminating Noise in Power Responses," *J. Sound Vib.*, vol. 82, pp. 593–595 (1982).

[8] A. Lundeby, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of Measurements in Room Acoustics," *Acustica*, vol. 81, pp. 344–355 (1995).

[9] N. Xiang, "Evaluation of Reverberation Times Using a Nonlinear Regression Approach," *J. Acoust. Soc. Am.*, vol. 98, pp. 2112–2121 (1995 Oct.).

[10] W. C. Sabine, *Architectural Acoustics* (1900; reprinted by Dover, New York, 1964).

[11] L. Beranek, *Acoustic Measurements* (1949; reprinted by Acoustical Society of America, 1988).

[12] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1975), chap. 10, pp. 480–531.

[13] MathWorks, Inc., *MATLAB User's Guide* (2001).

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing*, vol. I, *Estimation Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[15] J. D. Markel and J. A. H. Gray, *Linear Prediction of Speech* (Springer, Berlin, 1976).

[16] MathWorks, Inc., *MATLAB Signal Processing Toolbox, User's Guide* (2001).

[17] M. Karjalainen, P. A. A. Esquef, P. Antsallo, A. Mäkivirta, and V. Välimäki, "Frequency-Zooming ARMA Modeling of Resonant and Reverberant Systems," to be published, vol. 50 (2002 Dec.).

[18] W. T. Chu, "Comparison of Reverberation Measurements Using Schroeder's Impulse Method and Decay-Curve Averaging Method," *J. Acoust. Soc. Am.*, vol. 63, pp. 1444–1450 (1978 May).

[19] L. Faiget, R. Ruiz, and C. Legros, "Estimation of Impulse Response Length to Compute Room Acoustical Criteria," *Acustica*, vol. 82, suppl. 1, p. S148 (1996 Sept.).

[20] F. Satoh, Y. Hidaka, and H. Tachibana, "Reverberation Time Directly Obtained from Squared Impulse Response Envelope," in *Proc. Int. Congr. on Acoustics*, vol. 4 (Seattle, WA, 1998 June), pp. 2755–2756.

[21] E. Zwicker and H. Fastl, *Psychoacoustics—Facts and Models* (Springer, Berlin, 1990).

[22] G. Weinreich, "Coupled Piano Strings," *J. Acoust. Soc. Am.*, vol. 62, pp. 1474–1484 (1977 Dec.).

THE AUTHORS



P. Antsallo



T. Peltonen

Poju Pietari Antsalo was born in Helsinki, Finland, in 1971. He studied electrical and communications engineering at the Helsinki University of Technology, and obtained a master's degree in acoustics and audio signal processing in 1999. He has carried out research on room acoustics and audio reproduction, particularly at low frequencies.



Timo Peltonen was born in Espoo, Finland, in 1972. He received an M.Sc. degree in electrical engineering from the Helsinki University of Technology in 2000. His master's thesis addressed the measurement and room acoustic analysis of multichannel impulse responses.

He is the author of IRMA, the Matlab-based measurement and analysis software. Currently he is working as an

acoustical consultant at Akukon Consulting Engineers Ltd., Helsinki, Finland.

Mr. Peltonen's interests include acoustical measurement techniques and signal processing methods for room acoustical analysis.

He is a member of the AES, a board member of the AES Finnish Section, and a member of the Acoustical Society of Finland. He is the author of several AES conference papers.

The biographies of Matti Karjalainen and Vesa Välimäki were published in the 2002 April issue of the *Journal*.

The biography of Aki Mäkivirta was published in the 2002 July/August issue of the *Journal*.

On the Use of Time–Frequency Reassignment in Additive Sound Modeling*

KELLY FITZ, *AES Member* AND LIPPOLD HAKEN, *AES Member*

Department of Electrical Engineering and Computer Science, Washington University, Pulman, WA 99164

A method of reassignment in sound modeling to produce a sharper, more robust additive representation is introduced. The reassigned bandwidth-enhanced additive model follows ridges in a time–frequency analysis to construct partials having both sinusoidal and noise characteristics. This model yields greater resolution in time and frequency than is possible using conventional additive techniques, and better preserves the temporal envelope of transient signals, even in modified reconstruction, without introducing new component types or cumbersome phase interpolation algorithms.

0 INTRODUCTION

The method of reassignment has been used to sharpen spectrograms in order to make them more readable [1], [2], to measure sinusoidality, and to ensure optimal window alignment in the analysis of musical signals [3]. We use time–frequency reassignment to improve our bandwidth-enhanced additive sound model. The bandwidth-enhanced additive representation is in some way similar to traditional sinusoidal models [4]–[6] in that a waveform is modeled as a collection of components, called partials, having time-varying amplitude and frequency envelopes. Our partials are not strictly sinusoidal, however. We employ a technique of bandwidth enhancement to combine sinusoidal energy and noise energy into a single partial having time-varying amplitude, frequency, and bandwidth parameters [7], [8].

Additive sound models applicable to polyphonic and nonharmonic sounds employ long analysis windows, which can compromise the time resolution and phase accuracy needed to preserve the temporal shape of transients. Various methods have been proposed for representing transient waveforms in additive sound models. Verma and Meng [9] introduce new component types specifically for modeling transients, but this method sacrifices the homogeneity of the model. A homogeneous model, that is, a model having a single component type, such as the breakpoint parameter envelopes in our reassigned bandwidth-enhanced additive model [10], is critical for many kinds of

manipulations [11], [12]. Peeters and Rodet [3] have developed a hybrid analysis/synthesis system that eschews high-level transient models and retains unabridged OLA (overlap–add) frame data at transient positions. This hybrid representation represents unmodified transients perfectly, but also sacrifices homogeneity. Quatieri et al. [13] propose a method for preserving the temporal envelope of short-duration complex acoustic signals using a homogeneous sinusoidal model, but it is inapplicable to sounds of longer duration, or sounds having multiple transient events.

We use the method of reassignment to improve the time and frequency estimates used to define our partial parameter envelopes, thereby enhancing the time–frequency resolution of our representation, and improving its phase accuracy. The combination of time–frequency reassignment and bandwidth enhancement yields a homogeneous model (that is, a model having a single component type) that is capable of representing at high fidelity a wide variety of sounds, including nonharmonic, polyphonic, impulsive, and noisy sounds. The reassigned bandwidth-enhanced sound model is robust under transformation, and the fidelity of the representation is preserved even under time dilation and other model-domain modifications. The homogeneity and robustness of the reassigned bandwidth-enhanced model make it particularly well suited for such manipulations as cross synthesis and sound morphing.

Reassigned bandwidth-enhanced modeling and rendering and many kinds of manipulations, including morphing, have been implemented in the open-source C++ class library Loris [14], and a stream-based, real-time implementation of bandwidth-enhanced synthesis is available in the Symbolic Sound Kyma environment [15].

* Manuscript received 2001 December 20; revised 2002 July 23 and 2002 September 11.

1 TIME-FREQUENCY REASSIGNMENT

The discrete short-time Fourier transform is often used as the basis for a time-frequency representation of time-varying signals, and is defined as a function of time index n and frequency index k as

$$X_n(k) = \sum_{l=-\infty}^{\infty} h(l-n)x(l) \exp\left[\frac{-j2\pi(l-n)k}{N}\right] \quad (1)$$

$$= \sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} h(l)x(n+l) \exp\left(\frac{-j2\pi lk}{N}\right) \quad (2)$$

where $h(n)$ is a sliding window function equal to 0 for $n < -(N-1)/2$ and $n > (N-1)/2$ (for N odd), so that $X_n(k)$ is the N -point discrete Fourier transform of a short-time waveform centered at time n .

Short-time Fourier transform data are sampled at a rate equal to the analysis hop size, so data in derivative time-frequency representations are reported on a regular temporal grid, corresponding to the centers of the short-time analysis windows. The sampling of these so-called frame-based representations can be made as dense as desired by an appropriate choice of hop size. However, temporal smearing due to long analysis windows needed to achieve high-frequency resolution cannot be relieved by denser sampling.

Though the short-time phase spectrum is known to contain important temporal information, typically only the short-time magnitude spectrum is considered in the time-frequency representation. The short-time phase spectrum is sometimes used to improve the frequency estimates in the time-frequency representation of quasi-harmonic sounds [16], but it is often omitted entirely, or used only in unmodified reconstruction, as in the basic sinusoidal model, described by McAulay and Quatieri [4].

The so-called method of reassignment computes sharpened time and frequency estimates for each spectral component from partial derivatives of the short-time phase spectrum. Instead of locating time-frequency components at the geometrical center of the analysis window (t_n, ω_k), as in traditional short-time spectral analysis, the components are reassigned to the center of gravity of their complex spectral energy distribution, computed from the short-time phase spectrum according to the principle of stationary phase [17, ch. 7.3]. This method was first developed in the context of the spectrogram and called the modified moving window method [18], but it has since been applied to a variety of time-frequency and time-scale transforms [1].

The principle of stationary phase states that the variation of the Fourier phase spectrum not attributable to periodic oscillation is slow with respect to frequency in certain spectral regions, and in surrounding regions the variation is relatively rapid. In Fourier reconstruction, positive and negative contributions to the waveform cancel in frequency regions of rapid phase variation. Only regions of slow phase variation (stationary phase) will contribute signifi-

cantly to the reconstruction, and the maximum contribution (center of gravity) occurs at the point where the phase is changing most slowly with respect to time and frequency.

In the vicinity of $t = \tau$ (that is, for an analysis window centered at time $t = \tau$), the point of maximum spectral energy contribution has time-frequency coordinates that satisfy the stationarity conditions

$$\frac{\partial}{\partial \omega} [\phi(\tau, \omega) + \omega(t - \tau)] = 0 \quad (3)$$

$$\frac{\partial}{\partial \tau} [\phi(\tau, \omega) + \omega(t - \tau)] = 0 \quad (4)$$

where $\phi(\tau, \omega)$ is the continuous short-time phase spectrum and $\omega(t - \tau)$ is the phase travel due to periodic oscillation [18]. The stationarity conditions are satisfied at the coordinates

$$\hat{t} = \tau - \frac{\partial \phi(\tau, \omega)}{\partial \omega} \quad (5)$$

$$\hat{\omega} = \frac{\partial \phi(\tau, \omega)}{\partial \tau} \quad (6)$$

representing group delay and instantaneous frequency, respectively.

Discretizing Eqs. (5) and (6) to compute the time and frequency coordinates numerically is difficult and unreliable, because the partial derivatives must be approximated. These formulas can be rewritten in the form of ratios of discrete Fourier transforms [1]. Time and frequency coordinates can be computed using two additional short-time Fourier transforms, one employing a time-weighted window function and one a frequency-weighted window function.

Since time estimates correspond to the temporal center of the short-time analysis window, the time-weighted window is computed by scaling the analysis window function by a time ramp from $-(N-1)/2$ to $(N-1)/2$ for a window of length N . The frequency-weighted window is computed by wrapping the Fourier transform of the analysis window to the frequency range $[-\pi, \pi]$, scaling the transform by a frequency ramp from $-(N-1)/2$ to $(N-1)/2$, and inverting the scaled transform to obtain a (real) frequency-scaled window. Using these weighted windows, the method of reassignment computes corrections to the time and frequency estimates in fractional sample units between $-(N-1)/2$ to $(N-1)/2$. The three analysis windows employed in reassigned short-time Fourier analysis are shown in Fig. 1.

The reassigned time $\hat{t}_{k,n}$ for the k th spectral component from the short-time analysis window centered at time n (in samples, assuming odd-length analysis windows) is [1]

$$\hat{t}_{k,n} = n - \Re \left[\frac{X_{t,n}(k) X_n^*(k)}{|X_n(k)|^2} \right] \quad (7)$$

where $X_{t,n}(k)$ denotes the short-time transform computed using the time-weighted window function and $\Re [\cdot]$ denotes the real part of the bracketed ratio.

The corrected frequency $\hat{\omega}_{k,n}(k)$ corresponding to the same component is [1]

$$\hat{\omega}_{k,n} = k + \Im \left[\frac{X_{f;n}(k) X_n^*(k)}{|X_n(k)|^2} \right] \quad (8)$$

where $X_{f;n}(k)$ denotes the short-time transform computed using the frequency-weighted window function and $\Im [\cdot]$ denotes the imaginary part of the bracketed ratio. Both $t_{k,n}$ and $\hat{\omega}_{k,n}$ have units of fractional samples.

Time and frequency shifts are preserved in the reassignment operation, and energy is conserved in the re-assigned time–frequency data. Moreover, chirps and impulses are perfectly localized in time and frequency in any reassigned time–frequency or time-scale representation [1]. Reassignment sacrifices the bilinearity of time–frequency transformations such as the squared magnitude of the short-time Fourier transform, since very data point in the representation is relocated by a process that is highly signal dependent. This is not an issue in our representation, since the bandwidth-enhanced additive model, like the basic sinusoidal model [4], retains data only at time–frequency ridges (peaks in the short-time magnitude spectra), and thus is not bilinear.

Note that since the short-time Fourier transform is invertible, and the original waveform can be exactly reconstructed from an adequately sampled short-time Fourier representation, all the information needed to precisely locate a spectral component within an analysis window is present in the short-time coefficients $X_n(k)$. Temporal information is encoded in the short-time phase

spectrum, which is very difficult to interpret. The method reassignment is a technique for extracting information from the phase spectrum.

2 REASSIGNED BANDWIDTH-ENHANCED ANALYSIS

The reassigned bandwidth-enhanced additive model [10] employs time–frequency reassignment to improve the time and frequency estimates used to define partial parameter envelopes, thereby improving the time–frequency resolution and the phase accuracy of the representation. Reassignment transforms our analysis from a frame-based analysis into a “true” time–frequency analysis. Whereas the discrete short-time Fourier transform defined by Eq. (2) orients data according to the analysis frame rate and the length of the transform, the time and frequency orientation of reassigned spectral data is solely a function of the data themselves.

The method of analysis we use in our research models a sampled audio waveform as a collection of bandwidth-enhanced partials having sinusoidal and noise-like characteristics. Other methods for capturing noise in additive sound models [5], [19] have represented noise energy in fixed frequency bands using more than one component type. By contrast, bandwidth-enhanced partials are defined by a trio of synchronized breakpoint envelopes specifying the time-varying amplitude, center frequency, and noise content for each component. Each partial is rendered by a bandwidth-enhanced oscillator, described by

$$y(n) = [A(n) + B(n)\zeta(n)] \cos[\theta(n)] \quad (9)$$

where $A(n)$ and $\beta(n)$ are the time-varying sinusoidal and noise amplitudes, respectively, and $\zeta(n)$ is a energy-normalized low-pass noise sequence, generated by exciting a low-pass filter with white noise and scaling the filter gain such that the noise sequence has the same total spectral energy as a full-amplitude sinusoid. The oscillator phase $\theta(n)$ is initialized to some starting value, obtained from the reassigned short-time phase spectrum, and updated according to the time-varying radian frequency $\omega(n)$ by

$$\theta(n) = \theta(n-1) + \omega(n), \quad n > 0 \quad (10)$$

The bandwidth-enhanced oscillator is depicted in Fig. 2.

We define the time-varying bandwidth coefficient $\kappa(n)$ as the fraction of total instantaneous partial energy that is attributable to noise. This bandwidth (or noisiness) coefficient assumes values between 0 for a pure sinusoid and 1 for a partial that is entirely narrow-band noise, and varies over time according to the noisiness of the partial. If we represent the total (sinusoidal and noise) instantaneous partial energy as $\tilde{A}^2(n)$, then the output of the bandwidth-enhanced oscillator is described by

$$y(n) = \tilde{A}(n) \left[\sqrt{1 - \kappa(n)} + \sqrt{2\kappa(n)} \zeta(n) \right] \cos[\theta(n)]. \quad (11)$$

The envelopes for the time-varying partial amplitudes and frequencies are constructed by identifying and following

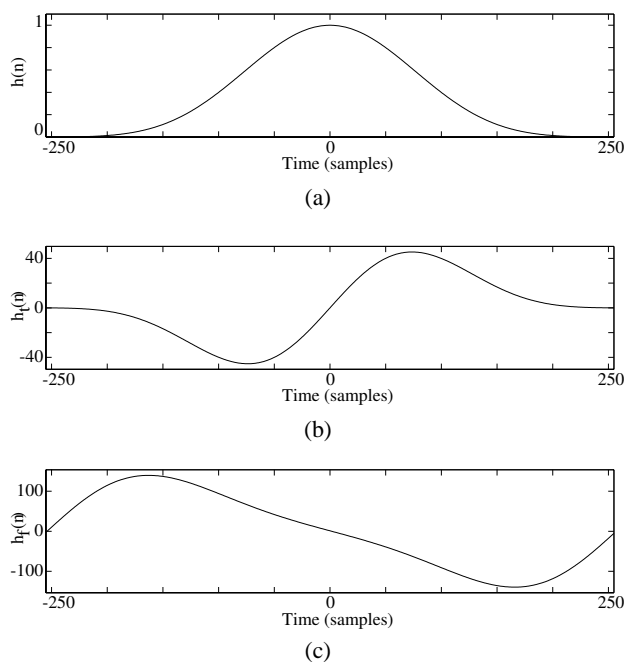


Fig. 1. Analysis windows employed in three short-time transforms used to compute reassigned times and frequencies. (a) Original window function $h(n)$ (a 501-point Kaiser window with shaping parameter 12.0 in this case). (b) Time-weighted window function $h_t(n) = nh(n)$. (c) Frequency-weighted window function $h_f(n)$.

the ridges on the time–frequency surface. The time-varying partial bandwidth coefficients are computed and assigned by a process of bandwidth association [7].

We use the method of reassignment to improve the time and frequency estimates for our partial parameter envelope breakpoints by computing reassigned times and frequencies that are not constrained to lie on the time–frequency grid defined by the short-time Fourier analysis parameters. Our algorithm shares with traditional sinusoidal methods the notion of temporally connected partial parameter estimates, but by contrast, our estimates are nonuniformly distributed in both time and frequency.

Short-time analysis windows normally overlap in both time and frequency, so time–frequency reassignment often yields time corrections greater than the length of the short-time hop size and frequency corrections greater than the width of a frequency bin. Large time corrections are common in analysis windows containing strong transients that are far from the temporal center of the window. Since we retain data only at time–frequency ridges, that is, at frequencies of spectral energy concentration, we generally observe large frequency corrections only in the presence of strong noise components, where phase stationarity is a weaker effect.

3 SHARPENING TRANSIENTS

Time–frequency representations based on traditional magnitude-only short-time Fourier analysis techniques (such as the spectrogram and the basic sinusoidal model [4]) fail to distinguish transient components from sustaining components. A strong transient waveform, as shown in Fig. 3(a), is represented by a collection of low-amplitude spectral components in early short-time analysis frames, that is, frames corresponding to analysis windows centered earlier than the time of the transient. A low-amplitude periodic waveform, as shown in Fig. 3(b), is also represented by a collection of low-amplitude spectral components. The information needed to distinguish these two critically different waveforms is encoded in the short-time phase spectrum, and is extracted by the method of reassignment.

Time–frequency reassignment allows us to preserve the temporal envelope shape without sacrificing the homogeneity of the bandwidth-enhanced additive model. Com-

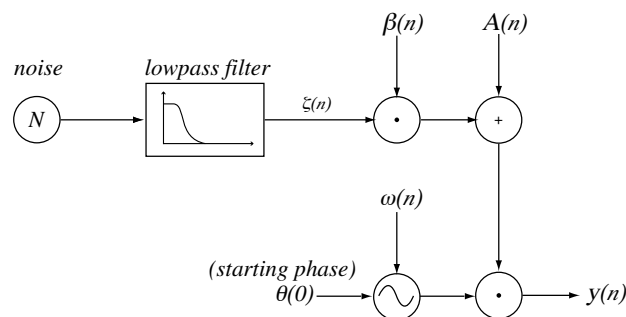


Fig. 2. Block diagram of bandwidth-enhanced oscillator. Time-varying sinusoidal and noise amplitudes are controlled by $A(n)$ and $\beta(n)$, respectively; time-varying center (sinusoidal) frequency is $\omega(n)$.

ponents extracted from early or late short-time analysis windows are relocated nearer to the times of transient events, yielding clusters of time–frequency data points, as depicted in Fig. 4. In this way, time reassignment greatly reduces the temporal smearing introduced through the use of long analysis windows. Moreover, since reassignment sharpens our frequency estimates, it is possible to achieve good frequency resolution with shorter (in time) analysis windows than would be possible with traditional methods. The use of shorter analysis windows further improves our time resolution and reduces temporal smearing.

The effect of time–frequency reassignment on the transient response can be demonstrated using a square wave that turns on abruptly, such as the waveform shown in Fig. 5. This waveform, while aurally uninteresting and uninformative, is useful for visualizing the performance of various analysis methods. Its abrupt onset makes temporal smearing obvious, its simple harmonic partial amplitude relationship makes it easy to predict the necessary data for a good time–frequency representation, and its simple waveshape makes phase errors and temporal distortion easy to identify. Note, however, that this waveform is pathological for Fourier-based additive models, and exaggerates all of these problems with such methods. We use it only for the comparison of various methods.

Fig. 6 shows two reconstructions of the onset of a square wave from time–frequency data obtained using overlapping 54-ms analysis windows, with temporal centers separated by 10 ms. This analysis window is long compared to the period of the square wave, but realistic for the case of a polyphonic sound (a sound having multiple simultaneous voices), in which the square wave is one voice. For clarity, only the square wave is presented in this example, and other simultaneous voices are omitted. The square wave

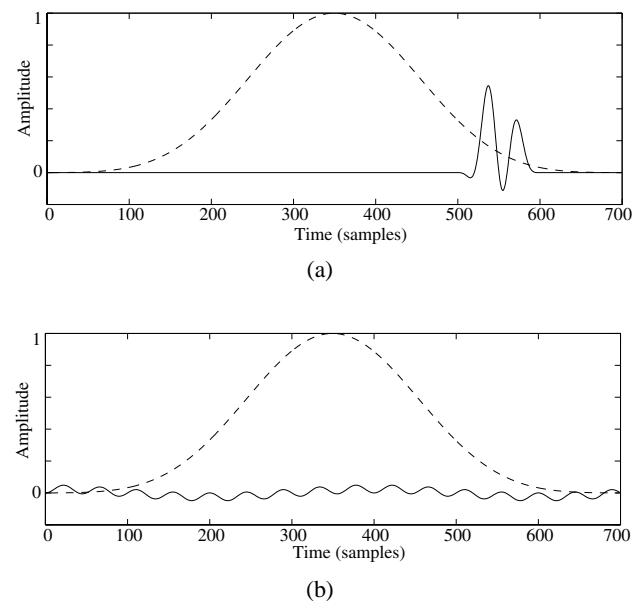


Fig. 3. Windowed short-time waveforms (dashed lines), not readily distinguished in basic sinusoidal model [4]. Both waveforms are represented by low-amplitude spectral components. (a) Strong transient yields off-center components, having large time corrections (positive in this case because transient is near right tail of window). (b) Sustained quasi-periodic waveform yields time corrections near zero.

has an abrupt onset. The silence before the onset is not shown. Only the first (lowest frequency) five harmonic partials were used in the reconstruction, and consequently the ringing due to Gibb's phenomenon is evident.

Fig. 6(a) is a reconstruction from traditional, nonreassigned time–frequency data. The reconstructed square wave amplitude rises very gradually and reaches full amplitude approximately 40 ms after the first nonzero sample. Clearly, the instantaneous turn-on has been smeared out by the long analysis window. Fig. 6(b) shows a reconstruction from reassigned time–frequency data. The transient response has been greatly improved by relocating components extracted from early analysis windows (like the one on the left in Fig. 5) to their spectral centers of gravity, closer to the observed turn-on time. The synthesized onset time has been reduced to approximately 10 ms. The corresponding time–frequency analysis data are shown in Fig. 7. The nonreassigned data are evenly distributed in time, so data from early windows (that is, windows centered before the onset time) smear the onset, whereas the reassigned data from early analysis windows are clumped near the correct onset time.

4 CROPPING

Off-center components are short-time spectral components having large time reassignments. Since they represent transient events that are far from the center of the analysis window, and are therefore poorly represented in the windowed short-time waveform, these off-center components introduce unreliable spectral parameter estimates that corrupt our representation, making the model data difficult to interpret and manipulated.

Fortunately large time corrections make off-center components easy to identify and remove from our model. By removing the unreliable data embodied by off-center components, we make our model cleaner and more robust. Moreover, thanks to the redundancy inherent in short-time analysis with overlapping analysis windows, we do not sacrifice information by removing the unreliable data points. The information represented poorly in off-center components is more reliably represented in well-centered components, extracted from analysis windows centered nearer the time of the transient event. Typically, data having time corrections

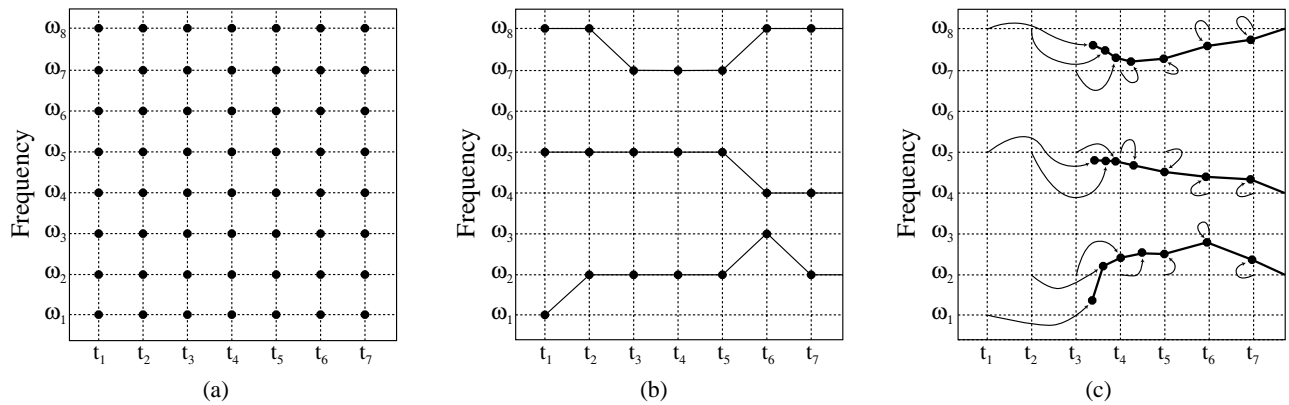


Fig. 4. Comparison of time–frequency data included in common representations. Only time–frequency orientation of data points is shown. (a) Short-time Fourier transform retains data at every time t_n and frequency ω_k . (b) Basic sinusoidal model [4] retains data at selected time and frequency samples. (c) Reassigned bandwidth-enhanced analysis data are distributed continuously in time and frequency, and retained only at time–frequency ridges. Arrows indicate mapping of short-time spectral samples onto time–frequency ridges due to method of reassignment.

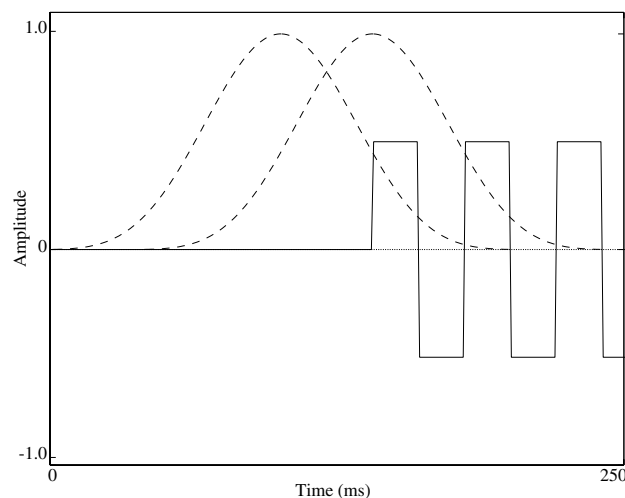


Fig. 5. Two long analysis windows superimposed at different times on square wave signal with abrupt turn-on. Short-time transform corresponding to earlier window generates unreliable parameter estimates and smears sharp onset of square wave.

greater than the time between consecutive analysis window centers are considered to be unreliable and are removed, or cropped.

Cropping partials to remove off-center components allows us to localize transient events reliably. Fig. 7(c) shows reassigned time–frequency data from the abrupt square wave onset with off-center components removed. The abrupt square wave onset synthesized from the cropped reassigned data, seen in Fig. 6(c), is much sharper than the uncropped reassigned reconstruction, because the taper of the analysis window makes even the time correc-

tion data unreliable in components that are very far off center.

Fig. 8 shows reassigned bandwidth-enhanced model data from the onset of a bowed cello tone before and after the removal of off-center components. In this case, components with time corrections greater than 10 ms (the time between consecutive analysis windows) were deemed to be too far off center to deliver reliable parameter estimates. As in Fig. 7(c), the unreliable data clustered at the time of the onset are removed, leaving a cleaner, more robust representation.

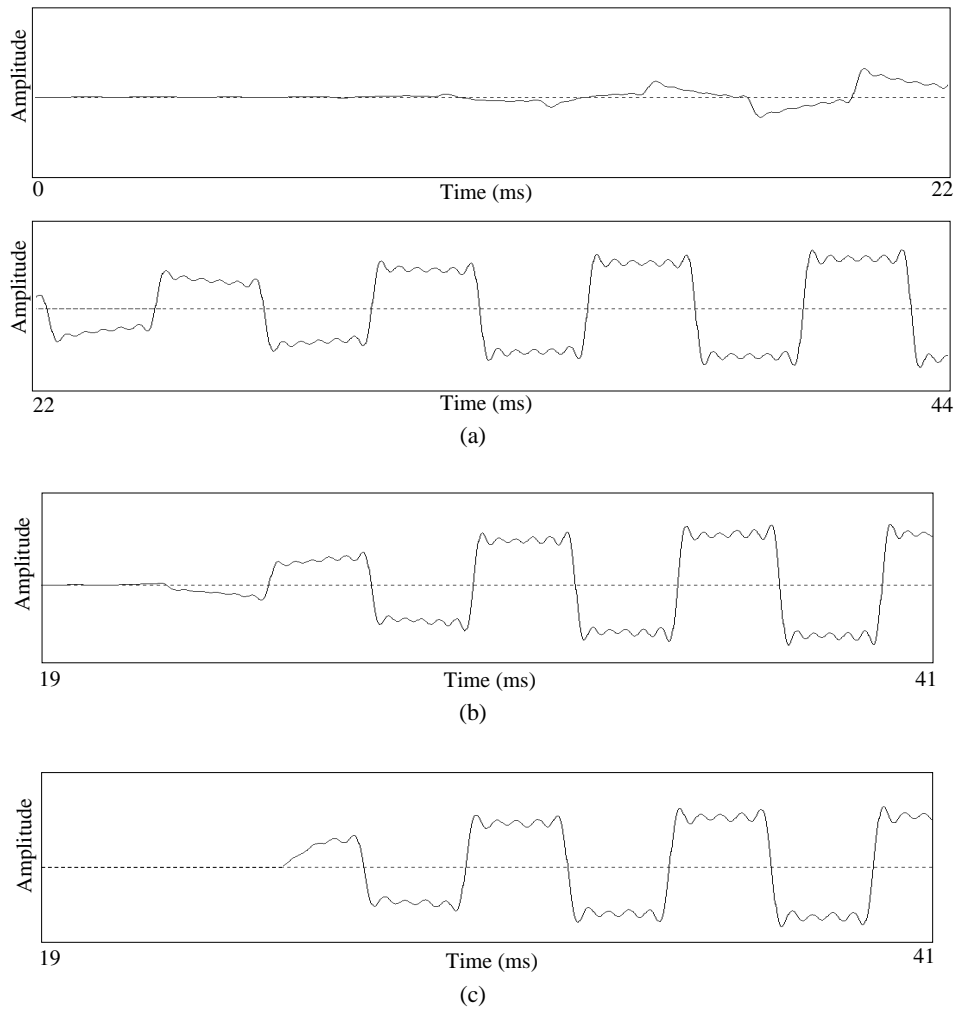


Fig. 6. Abrupt square wave onset reconstructed from five sinusoidal partials corresponding to first five harmonics. (a) Reconstruction from nonreassigned analysis data. (b) Reconstruction from reassigned analysis data. (c) Reconstruction from reassigned analysis data with unreliable partial parameter estimates removed, or cropped.

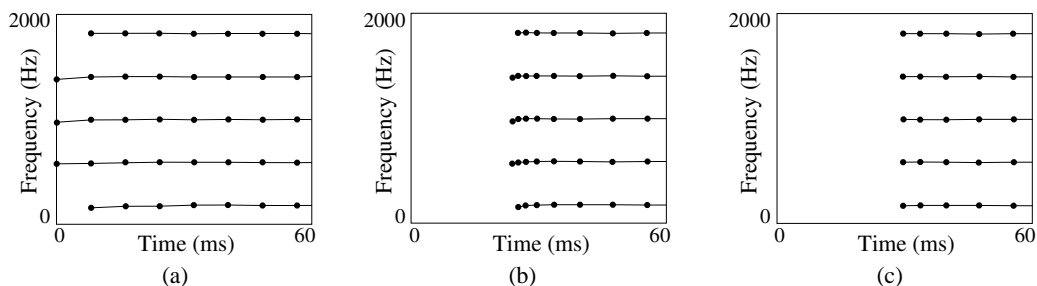


Fig. 7. Time–frequency analysis data points for abrupt square wave onset. (a) Traditional nonreassigned data are evenly distributed in time. (b) Reassigned data are clumped at onset time. (c) Reassigned analysis data after far off-center components have been removed, or cropped. Only time and frequency information is plotted; amplitude information is not displayed.

5 PHASE MAINTENANCE

Preserving phase is important for reproducing some classes of sounds, in particular transients and short-duration complex audio events having significant information in the temporal envelope [13]. The basic sinusoidal models proposed by McAulay and Quatieri [4] is phase correct, that is, it preserves phase at all times in unmodified reconstruction. In order to match short-time spectral frequency and phase estimates at frame boundaries, McAulay and Quatieri employ cubic interpolation of the instantaneous partial phase.

Cubic phase envelopes have many undesirable properties. They are difficult to manipulate and maintain under time- and frequency-scale transformation compared to linear frequency envelopes. However, in unmodified reconstruction, cubic interpolation prevents the propagation of phase errors introduced by unreliable parameter estimates, maintaining phase accuracy in transients, where the temporal envelope is important, and throughout the reconstructed waveform. The effect of phase errors in the unmodified reconstruction of a square wave is illustrated in Fig. 9. If not corrected using a technique such as cubic phase interpolation, partial parameter errors introduced by off-center components render the waveshape visually unrecognizable. Fig. 9(b) shows that cubic phase can be used to correct these errors in unmodified reconstruction.

It should be noted that, in this particular case, the phase errors appear dramatic, but do not affect the sound of the reconstructed steady-state waveforms appreciably. In many sounds, particularly transient sounds, preservation of the temporal envelope is critical [13], [9], but since they lack audible onset transients, the square waves in Fig. 9(a)–(c) sound identical. It should also be noted that cubic phase interpolation can be used to preserve phase accuracy, but does not reduce temporal smearing due to off-center components in long analysis windows.

It is not desirable to preserve phase at all times in modified reconstruction. Because frequency is the time derivative of phase, any change in the time or frequency scale of a partial must correspond to a change in the phase values at

the parameter envelope breakpoints. In general, preserving phase using the cubic phase method in the presence of modifications (or estimation errors) introduces wild frequency excursions [20]. Phase can be preserved at one time, however, and that time is typically chosen to be the onset of each partial, although any single time could be chosen. The partial phase at all other times is modified to reflect the new time–frequency characteristic of the modified partial.

Off-center components with unreliable parameter estimates introduce phase errors in modified reconstruction. If the phase is maintained at the partial onset, even the cubic interpolation scheme cannot prevent phase errors from propagating in modified syntheses. This effect is illustrated in Fig. 9(c), in which the square wave time–frequency data have been shifted in frequency by 10% and reconstructed using cubic phase curves modified to reflect the frequency shift.

By removing the off-center components at the onset of a partial, we not only remove the primary source of phase errors, we also improve the shape of the temporal envelope in the modified reconstruction of transients by preserving a more reliable phase estimate at a time closer to the time of the transient event. We can therefore maintain phase accuracy at critical parts of the audio waveform even under transformation, and even using linear frequency envelopes, which are much simpler to compute, interpret, edit, and maintain than cubic phase curves. Fig. 9(d) shows a square wave reconstruction from cropped reassigned time–frequency data, and Fig. 9(e) shows a frequency-shifted reconstruction, both using linear frequency interpolation. Removing components with large time corrections preserves phase in modified and unmodified reconstruction, and thus obviates cubic phase interpolation.

Moreover, since we do not rely on frequent cubic phase corrections to our frequency estimates to preserve the shape of the temporal envelope (which would otherwise be corrupted by errors introduced by unreliable data), we have found that we can obtain very good-quality reconstruction, even under modification, with regularly sampled partial parameter envelopes. That is, we can sample the frequency, amplitude, and bandwidth envelopes of our

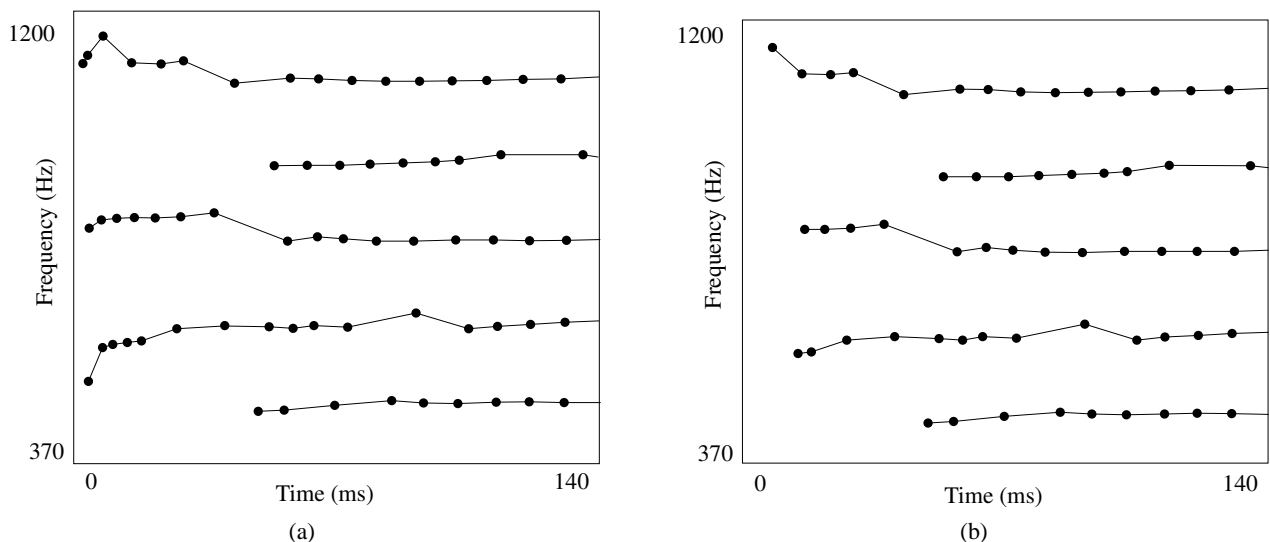


Fig. 8. Time–frequency coordinates of data from reassigned bandwidth-enhanced analysis. (a) Before cropping. (b) After cropping of off-center components clumped together at partial onsets. Source waveform is a bowed cello tone.

reassigned bandwidth-enhanced partials at regular intervals (of, for example, 10 ms) without sacrificing the fidelity of the model. We thereby achieve the data regularity of frame-based additive model data and the fidelity of reassigned spectral data. Resampling of the partial parameter envelopes is especially useful in real-time synthesis applications [11], [12].

6 BREAKING PARTIALS AT TRANSIENT EVENTS

Transients corresponding to the onset of all associated partials are preserved in our model by removing off-center components at the ends of partials. If transients always cor-

respond to the onset of associated partials, then that method will preserve the temporal envelope of multiple transient events. In fact, however, partials often span transients. Fig. 10 shows a partial that extends over transient boundaries in a representation of a bongo roll, a sequence of very short transient events. The approximate attack times are indicated by dashed vertical lines. In such cases it is not possible to preserve the phase at the locations of multiple transients, since under modification the phase can only be preserved at one time in the life of a partial.

Strong transients are identified by the large time corrections they introduce. By breaking partials at components having large time corrections, we cause all associated par-

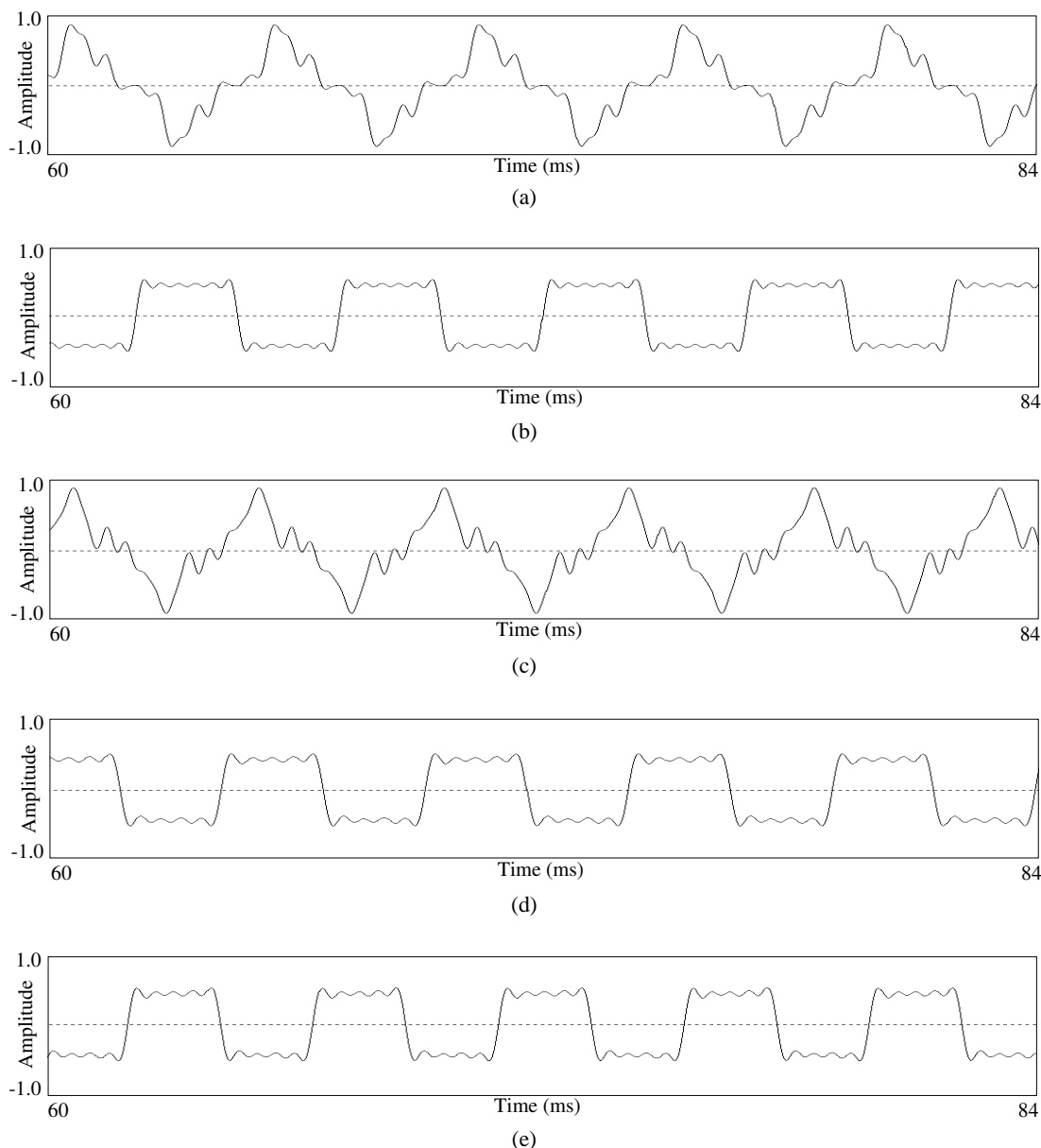


Fig. 9. Reconstruction of square wave having abrupt onset from five sinusoidal partials corresponding to first five harmonics. 24-ms plot spans slightly less than five periods of 200-Hz waveform. (a) Waveform reconstructed from nonreassigned analysis data using linear interpolation of partial frequencies. (b) Waveform reconstructed from nonreassigned analysis data using cubic phase interpolation, as proposed by McAulay and Quatieri [4]. (c) Waveform reconstructed from nonreassigned analysis data using cubic phase interpolation, with partial frequencies shifted by 10%. Notice that more periods of (distorted) waveform are spanned by 24-ms plot than by plots of unmodified reconstructions, due to frequency shift. (d) Waveform reconstructed from time-frequency reassigned analysis data using linear interpolation of partial frequencies, and having off-center components removed, or cropped. (e) Waveform reconstructed from reassigned analysis data using linear interpolation of partial frequencies and cropping of off-center components, with partial frequencies shifted by 10%. Notice that more periods of waveform are spanned by 24-ms plot than by plots of unmodified reconstructions, and that no distortion of waveform is evident.

tials to be born at the time of the transient, and thereby enhance our ability to maintain phase accuracy. In Fig. 11 the partial that spanned several transients in Fig. 10 has been broken at components having time corrections greater than the time between successive analysis window centers (about 1.3 ms in this case), allowing us to maintain the partial phases at each bongo strike. By breaking partials at the locations of transients, we can preserve the temporal envelope of multiple transient events, even under transformation.

Fig. 12(b) shows the waveform for two strikes in a bongo roll reconstructed from reassigned bandwidth-enhanced data.

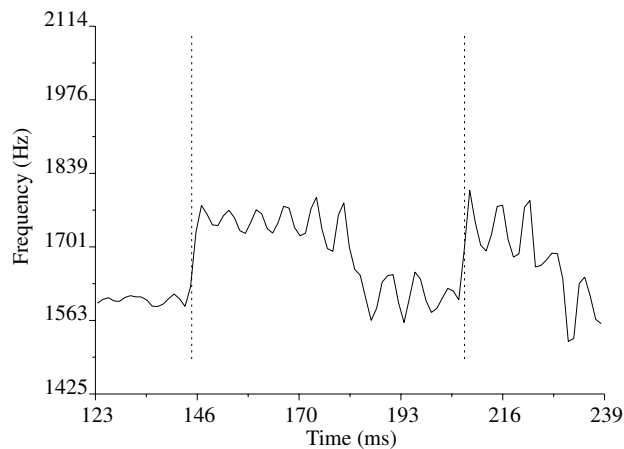


Fig. 10. Time–frequency plot of reassigned bandwidth-enhanced analysis data for one strike in a bongo roll. Dashed vertical lines show approximate locations of attack transients. Partial extends across transient boundaries. Only time–frequency coordinates of partial data are shown; partial amplitudes are not indicated.

The same two bongo strikes reconstructed from nonreassigned data are shown in Fig. 12(a). A comparison with the source waveform shown in Fig. 12(a) reveals that the reconstruction from reassigned data is better able to preserve the temporal envelope than the reconstruction from nonreassigned data and suffers less from temporal smearing.

7 REAL-TIME SYNTHESIS

Together with Kurt Hebel of Symbolic Sound Corporation we have implemented a real-time reassigned bandwidth-

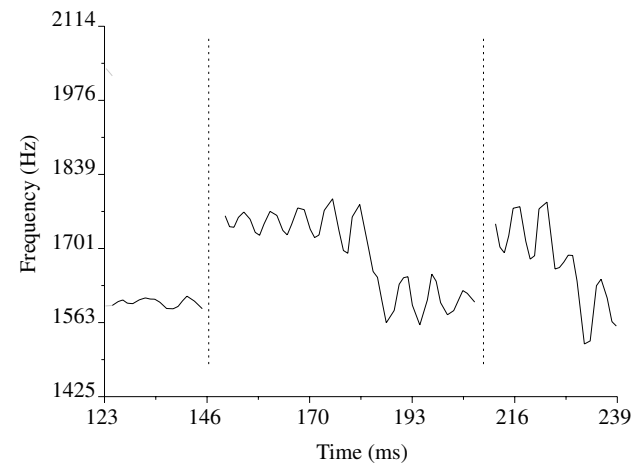
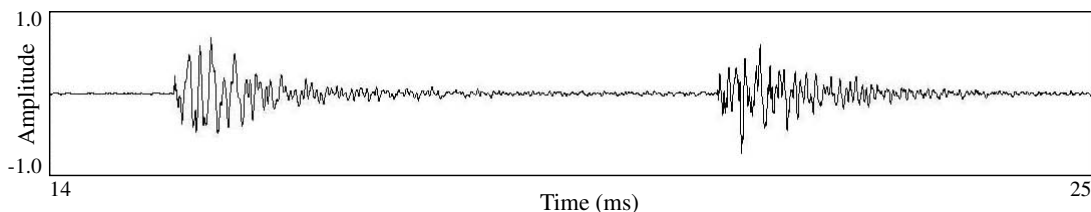
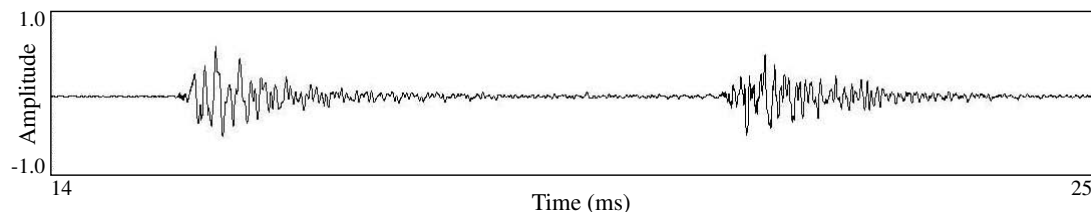


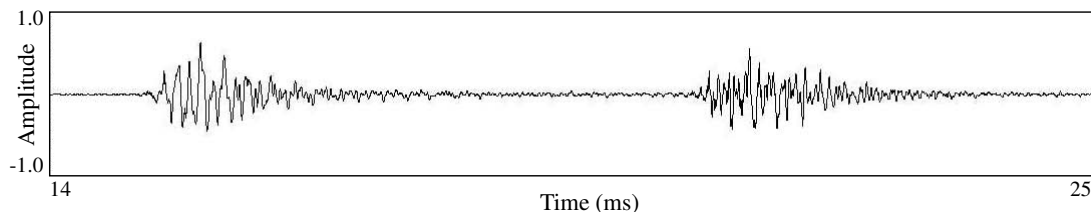
Fig. 11. Time–frequency plot of reassigned bandwidth-enhanced analysis data for one strike in a bongo roll with partials broken at components having large time corrections, and far off-center components removed. Dashed vertical lines show approximate locations of attack transients. Partial breaks at transient boundaries. Only time–frequency coordinates of partial data are shown; partial amplitudes are not indicated.



(a)



(b)



(c)

Fig. 12. Waveform plot for two strikes in a bongo roll. (a) Reconstructed from reassigned bandwidth-enhanced data. (b) Reconstructed from nonreassigned bandwidth-enhanced data. (c) Synthesized using cubic phase interpolation to maintain phase accuracy.

enhanced synthesizer using the Kyma Sound Design Workstation [15].

Many real-time synthesis systems allow the sound designer to manipulate streams of samples. In our real-time reassigned bandwidth-enhanced implementation, we work with streams of data that are not time-domain samples. Rather, our envelope parameter streams encode frequency, amplitude, and bandwidth envelope parameters for each bandwidth-enhanced partial [11], [12].

Much of the strength of systems that operate on sample streams is derived from the uniformity of the data. This homogeneity gives the sound designer great flexibility with a few general-purpose processing elements. In our encoding of envelope parameter streams, data homogeneity is also of prime importance. The envelope parameters for all the partials in a sound are encoded sequentially. Typically, the stream has a block size of 128 samples, which means the parameters for each partial are updated every 128 samples, or 2.9 ms at a 44.1-kHz sampling rate. Sample streams generally do not have block sizes associated with them, but this structure is necessary in our envelope parameter stream implementation. The envelope parameter stream encodes envelope information for a single partial at each sample time, and a block of samples provides updated envelope information for all the partials.

Envelope parameter streams are usually created by traversing a file containing frame-based data from an analysis of a source recording. Such a file can be derived from a reassigned bandwidth-enhanced analysis by resampling the envelopes at intervals of 128 samples at 44.1 kHz. The parameter streams may also be generated by real-time analysis, or by real-time algorithms, but that process is beyond the scope of this discussion. A parameter stream typically passes through several processing elements. These processing elements can combine multiple streams in a variety of ways, and can modify values within a stream. Finally a synthesis element computes an audio sample stream from the envelope parameter stream.

Our real-time synthesis element implements bandwidth-enhanced oscillators [8] with the sum

$$y(n) = \sum_{k=0}^{K-1} [A_k(n) + N_k(n)b(n)] \sin \theta_k(n) \quad (12)$$

$$\theta_k(n) = \theta_k(n-1) + 2^{F_k(n)} \quad (13)$$

where

- y = time-domain waveform for synthesized sound
- n = sample number
- k = partial number in sound
- K = total number of partials in sound (usually between 20 and 160)
- A_k = amplitude envelope of partial k
- N_k = noise envelope of partial k
- b = zero-mean noise modulator with bell-shaped spectrum
- F_k = \log_2 frequency envelope of partial k , radians per sample
- θ_k = running phase for k th partial.

Values for the envelopes A_k , N_k , and F_k are updated from the parameter stream every 128 samples. The synthesis element performs sample-level linear interpolation between updates, so that A_k , N_k , and F_k are piecewise linear envelopes with segments 128 samples in length [21]. The θ_k values are initialized at partial onsets (when A_k and N_k are zero) from the phase envelope in the partial's parameter stream.

Rather than using a separate model to represent noise in our sounds, we use the envelope N_k (in addition to the traditional A_k and F_k envelopes) and retain a homogeneous data stream. Quasi-harmonic sounds, even those with noisy attacks, have one partial per harmonic in our representation. The noise envelopes allow a sound designer to manipulate noiselike components of sound in an intuitive way, using a familiar set of controls. We have implemented a wide variety of real-time manipulations on envelope parameter streams, including frequency shifting, formant shifting, time dilation, cross synthesis, and sound morphing.

Our new MIDI controller, the Continuum Fingerboard, allows continuous control over each note in a performance. It resembles a traditional keyboard in that it is approximately the same size and is played with ten fingers [12]. Like keyboards supporting MIDI's polyphonic aftertouch, it continually measures each finger's pressure. The Continuum Fingerboard also resembles a fretless string instrument in that it has no discrete pitches; any pitch may be played, and smooth glissandi are possible. It tracks, in three dimensions (left to right, front to back, and downward pressure), the position for each finger pressing on the playing surface. These continuous three-dimensional outputs are a convenient source of control parameters for real-time manipulations on envelope parameter streams.

8 CONCLUSIONS

The reassigned bandwidth-enhanced additive sound model [10] combines bandwidth-enhanced analysis and synthesis techniques [7], [8] with the time-frequency reassignment technique described in this paper.

We found that the method of reassignment strengthens our bandwidth-enhanced additive sound model dramatically. Temporal smearing is greatly reduced because the time-frequency orientation of the model data is waveform dependent, rather than analysis dependent as in traditional short-time analysis methods. Moreover, time-frequency reassignment allows us to identify unreliable data points (having bad parameter estimates) and remove them from the representation. This not only sharpens the representation and makes it more robust, but it also allows us to maintain phase accuracy at transients, even under transformation, while avoiding the problems associated with cubic phase interpolation.

9 REFERENCES

- [1] F. Auger and P. Flandrin, "Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method," *IEEE Trans. Signal*

Process., vol. 43, pp. 1068–1089 (1995 May).

[2] F. Plante, G. Meyer, and W. A. Ainsworth, “Improvement of Speech Spectrogram Accuracy by the Method of Spectral Reassignment,” *IEEE Trans. Speech Audio Process.*, vol. 6, pp. 282–287 (1998 May).

[3] G. Peeters and X. Rode, “SINOLA: A New Analysis/Synthesis Method Using Spectrum Peak Shape Distortion, Phase and Reassigned Spectrum,” in *Proc. Int. Computer Music Conf.* (1999), pp. 153–156.

[4] R. J. McAulay and T. F. Quatieri, “Speech Analysis/Synthesis Based on a Sinusoidal Representation,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 744–754 (1986 Aug.).

[5] X. Serra and J. O. Smith, “Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic Plus Stochastic Decomposition,” *Computer Music J.*, vol. 14, no. 4, pp. 12–24 (1990).

[6] K. Fitz and L. Haken, “Sinusoidal Modeling and Manipulation Using Lemur,” *Computer Music J.*, vol. 20, no. 4, pp. 44–59 (1996).

[7] K. Fitz, L. Haken, and P. Christensen, “A New Algorithm for Bandwidth Association in Bandwidth-Enhanced Additive Sound Modeling,” in *Proc. Int. Computer Music Conf.* (2000).

[8] K. Fitz and L. Haken, “Bandwidth Enhanced Sinusoidal Modeling in Lemur,” in *Proc. Int. Computer Music Conf.* (1995), pp. 154–157.

[9] T. S. Verma and T. H. Y. Meng, “An Analysis/Synthesis Tool for Transient Signals,” in *Proc. 16th Int. Congr. on Acoustics/135th Mtg. of the Acoust. Soc. Am.* (1998 June), vol. 1, pp. 77–78.

[10] K. Fitz, L. Haken, and P. Christensen, “Transient Preservation under Transformation in an Additive Sound Model,” in *Proc. Int. Computer Music Conf.* (2000).

[11] L. Haken, K. Fitz, and P. Christensen, “Beyond Traditional Sampling Synthesis: Real-Time Timbre Morphing Using Additive Synthesis,” in *Sound of Music: Analysis, Synthesis, and Perception*, J. W. Beauchamp, Ed. (Springer, New York, to be published).

[12] L. Haken, E. Tellman, and P. Wolfe, “An Indiscrete Music Keyboard,” *Computer Music J.*, vol. 22, no. 1, pp. 30–48 (1998).

[13] T. F. Quatieri, R. B. Dunn, and T. E. Hanna, “Time-Scale Modification of Complex Acoustic Signals,” in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing* (IEEE, 1993), pp. I-213–I-216.

[14] K. Fitz and L. Haken, “The Loris C++ Class Library,” available at <http://www.cerlsoundgroup.org/Loris>.

[15] K. J. Hebel and C. Scaletti, “A Framework for the Design, Development, and Delivery of Real-Time Software-Based Sound Synthesis and Processing Algorithms,” presented at the 97th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 42, p. 1050 (1994 Dec.), preprint 3874.

[16] M. Dolson, “The Phase Vocoder: A Tutorial,” *Computer Music J.*, vol. 10, no. 4, pp. 14–27 (1986).

[17] A. Papoulis, *Systems and Transforms with Applications to Optics* (McGraw-Hill, New York, 1968), chap. 7.3, p. 234.

[18] K. Kodera, R. Gendrin, and C. de Villedary, “Analysis of Time-Varying Signals with Small *BT* Values,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-26, pp. 64–76 (1978 Feb.).

[19] D. W. Griffin and J. S. Lim, “Multiband Excitation Vocoder,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-36, p. 1223–1235 (1988 Aug.).

[20] Y. Ding and X. Qian, “Processing of Musical Tones Using a Combined Quadratic Polynomial-Phase Sinusoidal and Residual (QUASAR) Signal Model,” *J. Audio Eng. Soc.*, vol. 45, pp. 571–584 (1997 July/Aug.).

[21] L. Haken, “Computational Methods for Real-Time Fourier Synthesis,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-40, pp. 2327–2329 (1992 Sept.).

[22] A. Ricci, “SoundMaker 1.0.3,” MicroMat Computer Systems (1996–1997).

[23] F. Opolko and J. Wapnick, “McGill University Master Samples,” McGill University, Montreal, Que., Canada (1987).

[24] E. Tellman, cello tones recorded by P. Wolfe at Pogo Studios, Champaign, IL (1997 Jan.).

APPENDIX RESULTS

The reassigned bandwidth-enhanced additive model is implemented in the open source C++ class library Loris [14], and is the basis of the sound manipulation and morphing algorithms implemented therein.

We have attempted to use a wide variety of sounds in the experiments we conducted during the development of the reassigned bandwidth-enhanced additive sound model. The results from a few of those experiments are presented in this appendix. Data and waveform plots are not intended to constitute proof of the efficacy of our algorithms, or the utility of our representation. They are intended only to illustrate the features of some of the sounds used and generated in our experiments. The results of our work can only be judged by auditory evaluation, and to that end, these sounds and many others are available for audition at the Loris web site [14].

All sounds used in these experiments were sampled at 44.1 kHz (CD quality) so time–frequency analysis data are available at frequencies as high as 22.05 kHz. However, for clarity, only a limited frequency range is plotted in most cases. The spectrogram plots all have high gain so that low-amplitude high-frequency partials are visible. Consequently strong low-frequency partials are very often clipped, and appear to have unnaturally flat amplitude envelopes.

The waveform and spectrogram plots were produced using Ricci’s SoundMaker software application [22].

A.1 Flute Tone

A flute tone, played at pitch D4 (D above middle C), having a fundamental frequency of approximately 293 Hz and no vibrato, taken from the McGill University Master Samples compact discs [23, disc 2, track 1, index

3], is shown in the three-dimensional spectrogram plot in Fig. 13. This sound was modeled by reassigned bandwidth-enhanced analysis data produced using a 53-ms Kaiser analysis window with 90-dB sidelobe rejection. The partials were constrained to be separated by at least 250 Hz, slightly greater than 85% of the harmonic partial separation.

Breath noise is a significant component of this sound. This noise is visible between the strong harmonic components in the spectrogram plot, particularly at frequencies above 3 kHz. The breath noise is faithfully represented in the reassigned bandwidth-enhanced analysis data, and

reproduced in the reconstructions from those analysis data. A three-dimensional spectrogram plot of the reconstruction is shown in Fig. 14. The audible absence of the breath noise is apparent in the spectral plot for the sinusoidal reconstruction from non-bandwidth-enhanced analysis data, shown in Fig. 15.

A.2 Cello Tone

A cello tone, played at pitch D#3 (D sharp below middle C), having a fundamental frequency of approximately 156 Hz, played by Edwin Tellman and recorded by Patrick Wolfe [24] was modeled by reassigned bandwidth-

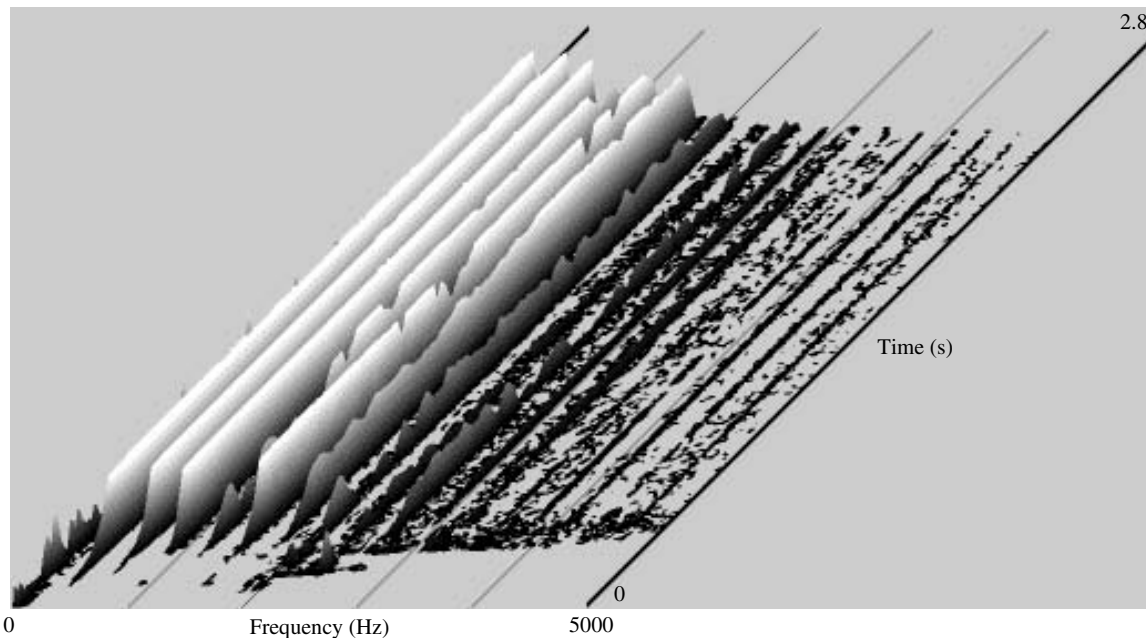


Fig. 13. Three-dimensional spectrogram plot for breathy flute tone, pitch D4 (D above middle C). Audible low-frequency noise and rumble from recording are visible. Strong low-frequency components are clipped and appear to have unnaturally flat amplitude envelopes due to high gain used to make low-amplitude high-frequency partials visible.

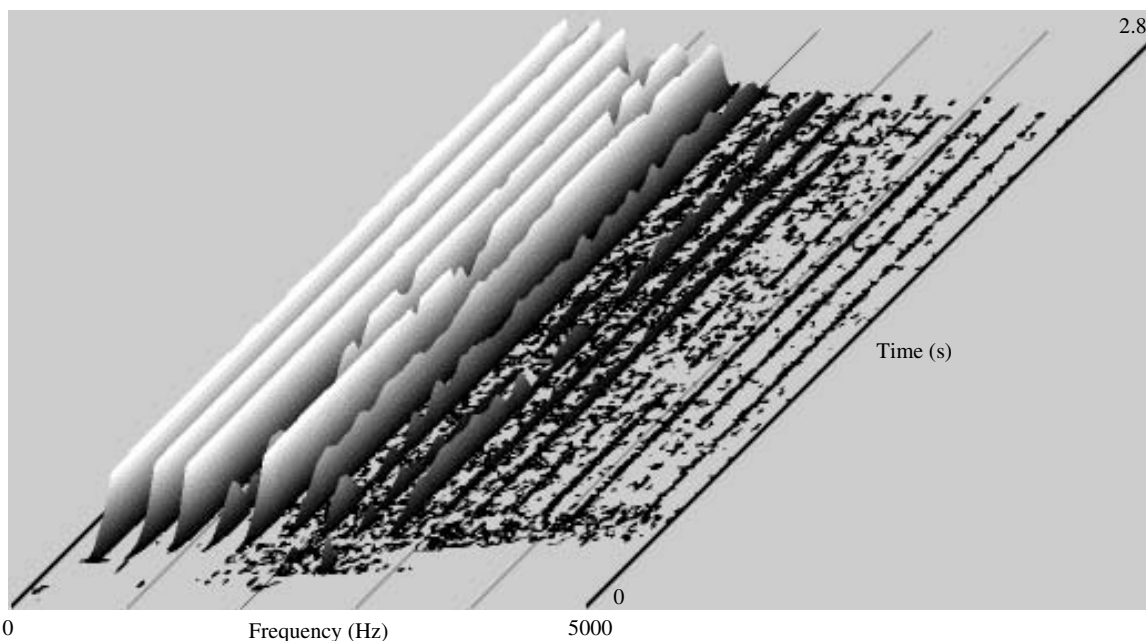


Fig. 14. Three-dimensional spectrogram plot for breathy flute tone, pitch D4 (D above middle C), reconstructed from reassigned bandwidth-enhanced analysis data.

enhanced analysis data produced using a 71-ms Kaiser analysis window with 80-dB sidelobe rejection. The partials were constrained to be separated by at least 135 Hz, slightly greater than 85% of the harmonic partial separation.

Bow noise is a strong component of the cello tone, especially in the attack portion. As with the flute tone, the noise is visible between the strong harmonic components in spectral plots, and was preserved in the reconstructions from reassigned bandwidth-enhanced analysis data and absent from sinusoidal (non-bandwidth-enhanced) reconstructions. Unlike the flute tone, the cello tone has an abrupt attack, which is smeared out in nonreassigned sinusoidal analyses (data from reassigned and nonreassigned cello analysis are plotted in Fig. 8), causing the reconstructed cello tone to have weak-sounding articulation. The characteristic “grunt” is much better preserved in reassigned model data.

A.3 Flutter-Tongued Flute Tone

A flutter-tongued flute tone, played at pitch E4 (E above middle C), having a fundamental frequency of approximately 330 Hz, taken from the McGill University Master Samples compact discs (23, disc 2, track 2, index 5), was represented by reassigned bandwidth-enhanced analysis data produced using a 17.8-ms Kaiser analysis window with 80-dB sidelobe rejection. The partials were constrained to be separated by at least 300 Hz, slightly greater than 90% of the harmonic partial separation. The flutter-tongue effect introduces a modulation with a period of approximately 35 ms, and gives the appearance of vertical stripes on the strong harmonic partials in the spectrogram shown in Fig. 16.

With careful choice of the window parameters, reconstruction from reassigned bandwidth-enhanced analysis data preserves the flutter-tongue effect, even under time dilation, and is difficult to distinguish from the original.

Fig. 17 shows how a poor choice of analysis window, a 71-ms Kaiser window in this case, can degrade the representation. The reconstructed tone plotted in Fig. 17 is recognizable, but lacks the flutter effect completely, which has been smeared by the window duration. In this case multiple transient events are spanned by a single analysis window, and the temporal center of gravity for that window lies somewhere between the transient events. Time–frequency reassignment allows us to identify multiple transient events in a single sound, but not within a single short-time analysis window.

A.4 Bongo Roll

Fig. 18 shows the waveform and spectrogram for an 18-strike bongo roll taken from the McGill University Master Samples compact discs [23, disc 3, track 11, index 31]. This sound was modeled by reassigned bandwidth-enhanced analysis data produced using a 10-ms Kaiser analysis window with 90-dB sidelobe rejection. The partials were constrained to be separated by at least 300 Hz.

The sharp attacks in this sound were preserved using reassigned analysis data, but smeared in nonreassigned reconstruction, as discussed in Section 6. The waveforms for two bongo strikes are shown in reassigned and nonreassigned reconstruction in Fig. 12(b) and (c). Inspection of the waveforms reveals that the attacks in the nonreassigned reconstruction are not as sharp as in the original or the reassigned reconstruction, a clearly audible difference.

Transient smearing is particularly apparent in time-dilated synthesis, where the nonreassigned reconstruction loses the percussive character of the bongo strikes. The reassigned data provide a much more robust representation of the attack transients, retaining the percussive character of the bongo roll under a variety of transformations, including time dilation.

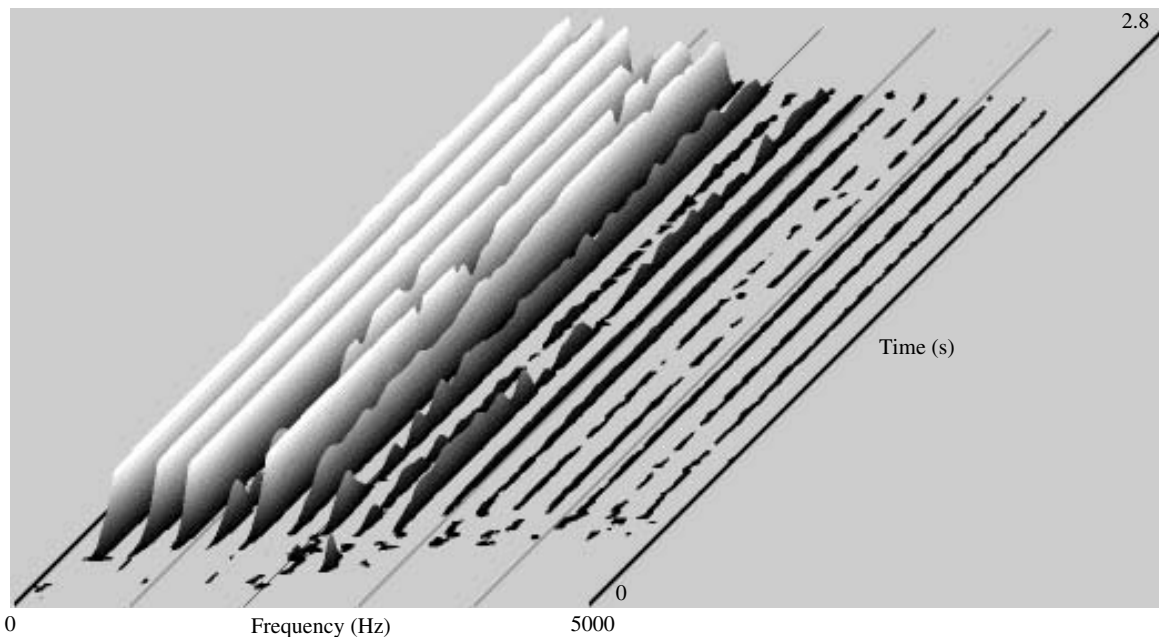


Fig. 15. Three-dimensional spectrogram plot for breathy flute tone, pitch D4 (D above middle C), reconstructed from reassigned non-bandwidth-enhanced analysis data.

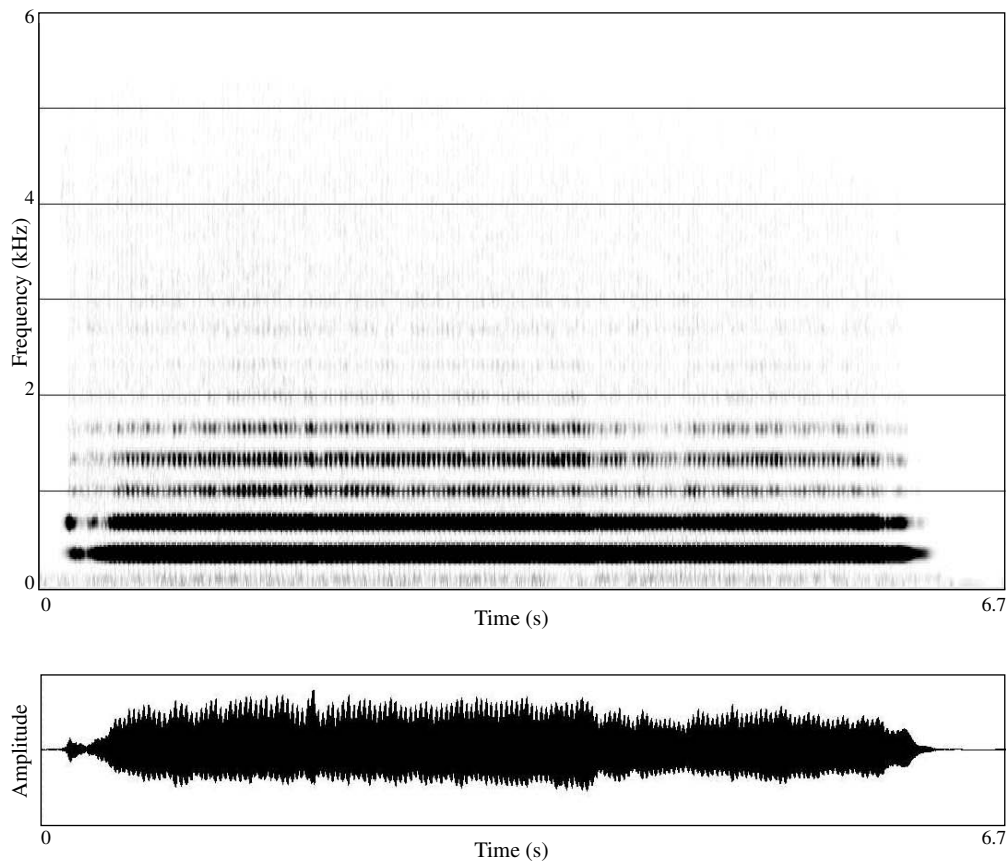


Fig. 16. Waveform and spectrogram plots for flutter-tongued flute tone, pitch E4 (E above middle C). Vertical stripes on strong harmonic partials indicate modulation due to flutter-tongue effect. Strong low-frequency components are clipped and appear to have unnaturally flat amplitude envelopes due to high gain used to make low-amplitude high-frequency partials visible.

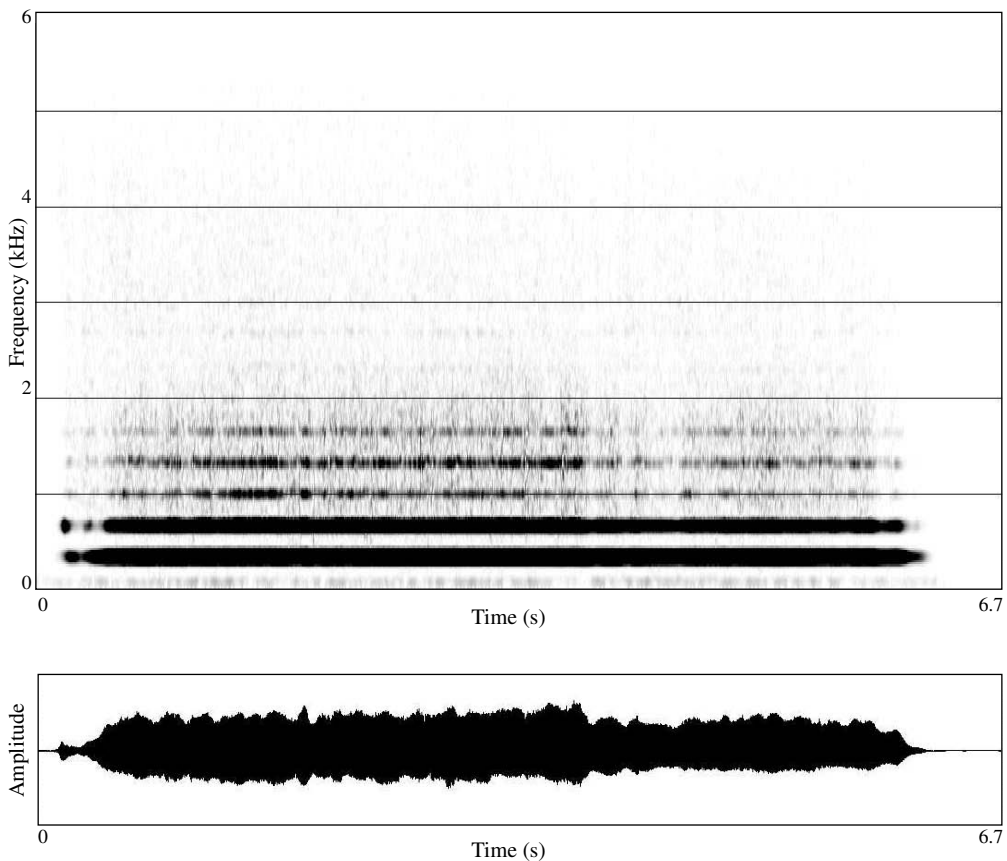


Fig. 17. Waveform and spectrogram plots for reconstruction of flutter-tongued flute tone plotted in Fig. 16, analyzed using long window, which smears out flutter effect.

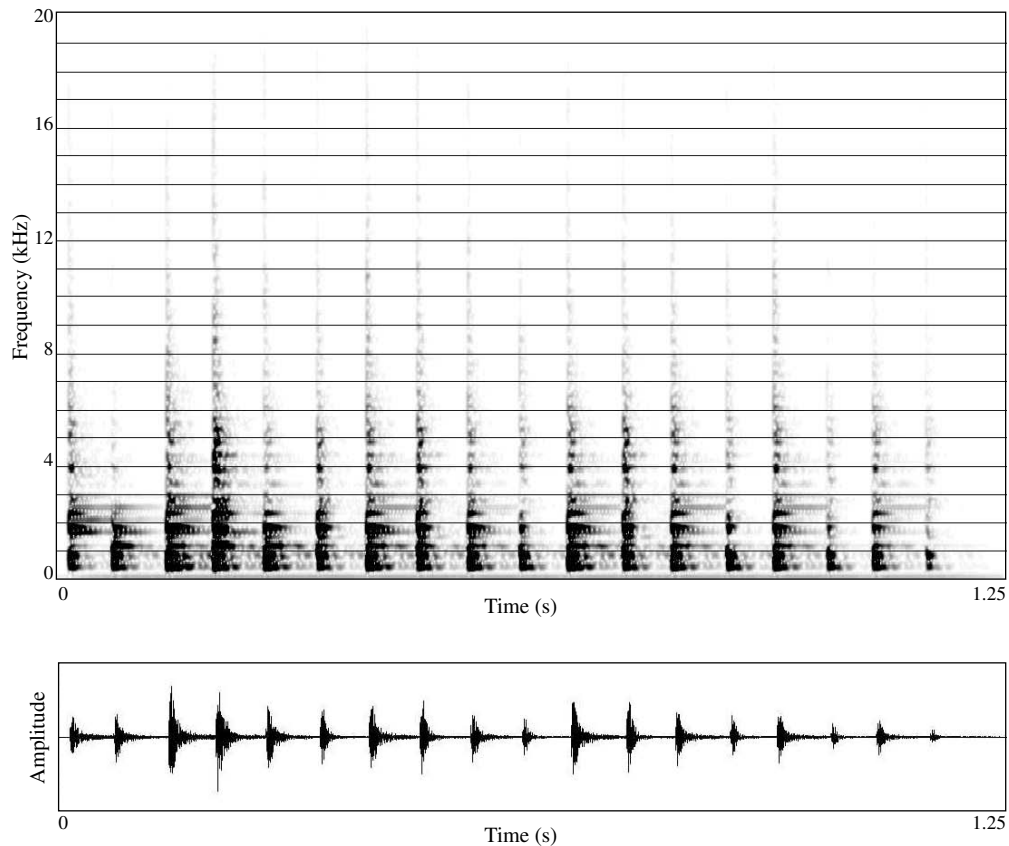


Fig. 18. Waveform and spectrogram plots for bongo roll.

THE AUTHORS



K. Fitz

Kelly Fitz received B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, in 1990, 1992, and 1999, respectively. There he studied digital signal processing as well as sound analysis and synthesis with Dr. James Beauchamp and sound design and electroacoustic music composition with Scott Wyatt using a variety of analog and digital systems in the experimental music studios.

Dr. Fitz is currently an assistant professor in the department of Electrical Engineering and Computer Science at the Washington State University.



Lippold Haken has an adjunct professorship in electrical and computer engineering at the University of



L. Haken

Illinois, and he is senior computer engineer at Prairie City Computing in Urbana, Illinois. He is leader of the CERL Sound Group, and together with his graduate students developed new software algorithms and signal processing hardware for computer music. He is inventor of the Continuum Fingerboard, a MIDI controller that allows continuous control over each note in a performance. He is a contributor of optimized real-time algorithms for the Symbolic Sound Corporation Kyma sound design workstation. He is also the author of a sophisticated music notation editor, Lime.

He is currently teaching a computer music survey course for seniors and graduate students in electrical and computer engineering.

Scalable Multichannel Coding with HRTF Enhancement for DVD and Virtual Sound Systems*

M. O. J. HAWKSFORD

Centre for Audio Research and Engineering, University of Essex, UK CO4 3SQ

A scalable and reverse compatible multichannel method of spatial audio using transaural coding designed for multiple-loudspeaker feeds is described with a focus on attaining optimum ear signals. A Fourier transform method for computing HRTF matrices is employed, including the generation of a subset of band-limited reproduction channels. Applications considered embrace multichannel audio, DVD, virtual reality, and telepresence.

0 INTRODUCTION

The purpose of this paper is to investigate how transaural processing can enhance conventional multichannel audio both by embedding perceptually relevant information and by improving image stability using additional loudspeakers integrated with supplementary digital processing and coding. The key objective is to achieve scalability in spatial performance while retaining full compatibility with conventional multichannel formats. This enables the system in its most basic form with unprocessed loudspeaker feeds to be used in a conventional multichannel installation. However, by appropriate signal processing additional loudspeaker feeds can be derived, together with the option of exploiting buried data to extract more signals in order to improve spatial resolution. The system is therefore hierarchical in terms of number of loudspeakers, channels, and ultimately spatial resolution, while in its simplest incarnation it remains fully compatible with the system configurations used with multichannel DVD-A and SACD replay equipment.

The multichannel capabilities of DVD¹ technology [1], [2] were designed to enhance stereo² sound reproduction by offering surround image and improved envelopment capabilities. Normally multichannel audio encoded onto DVD assumes the ITU standard of a five-loudspeaker configuration driven by five discrete wide-band “loudspeaker feeds.” However, a limitation of this system is the lack of a methodology to synthesize virtual images capable of three-dimensional audio (that is, a perception of direction, distance, and height together with acoustic envelopment)

rather than just “sound effects” often (although not exclusively) associated with surround sound in a home theater context. The ITU five-channel loudspeaker configuration can also be poor at side image localization, although this deficiency is closely allied to a sensitivity to room acoustics. Nevertheless, DVD formats still offer only six discrete channels, which if mapped directly into loudspeaker feeds remain deficient in terms of image precision, especially if height and depth information is to be encoded.

The techniques described in this paper support scalable spatial audio that can remain compatible with conventional multichannel systems. It is shown that in this class of system, under anechoic conditions, signal processing can be used to match theoretically the ear signals to either a real or an equivalent spatially synthesized sound source. Also, in order to improve image robustness, directional sound-field encoding is retained as exploited in conventional surround sound to match the image synthesized through transaural processing. It may be argued that as the number of channels is increased, there is convergence toward wavefront synthesis [3], where by default optimum ear-signal reconstruction is achieved. However, the proposed system is positioned well into the middle territory³ and is far removed from the array sizes required for broadband wavefront synthesis. Consequently, from the perspective of wavefront synthesis the transition frequency above which spatial aliasing occurs is located at a relatively low frequency, implying that for the proposed system the core concepts of wavefront synthesis do not apply

¹ Includes both DVD-A and SACD formats.

² Of Greek origin, meaning solid, stereo is applicable universally to multichannel audio.

³ A range of 5 to 32 loudspeakers is suggested.

over much of the audio band.

It is emphasized that an n -channel system does not necessarily imply n loudspeakers. Indeed, as is well known, it is possible for a two-loudspeaker system to reproduce virtual-sound sources [4], while using more than n loudspeakers can help create a more robust and stable illusion. Also, the mature technology of Ambisonics [5]–[7] is scalable and can accommodate both additional loudspeakers and information channels. However, here the encoding is hierarchical in terms of spatial spherical harmonics, although no attempt is made to reconstruct the ear signal directly at the listener. Consequently the approach taken in this paper differs in a number of fundamental aspects from that of Ambisonics, especially since there is no attempt to transform a sound field directly into a spherical harmonic set. Thus it remains for future work to establish the relative merits of these approaches although, because similar loudspeaker arrays are used, there is no fundamental compromise should the system be used either for Ambisonics or for conventional surround sound encoded audio.

The method of spatial audio described in this paper uses a conventional loudspeaker array to surround the listener and to reproduce a directional sound field. In addition, ear signals are simultaneously synthesized using head-related transfer functions (HRTFs) matched to the source image, where it is assumed in all cases that loudspeaker transfer functions have been equalized or otherwise taken into account. A number of examples illustrate the computational methods, which include pairwise transaural image synthesis⁴ reported in earlier work, where some preliminary experimental results were also discussed [8]–[10] to establish the efficacy of the method. This technique is especially well matched to multichannel multiloudspeaker installations, where transaural coding can be applied during encoding and recording while processing within the decoder located within the reproduction system can accommodate both additional loudspeakers and loudspeaker positions that differ from those assumed at the encoder. Consequently for an n -channel system it is straightforward to employ only n loudspeakers, although additional loudspeaker feeds can be derived, while still retaining correct ear signals, either by using matrix techniques or deterministically within the DVD-A format using additional embedded code. However, it is emphasized that in the simplest configuration, using only direct loudspeaker feeds and provided the loudspeakers are correctly located, there are no additional decoding requirements and the system remains fully compatible with all existing recordings.

Alternative technologies such as Ambisonics [11] have used fewer loudspeakers together with sophisticated matrix encoding. Also, there has been substantial research into perceptually based processing to reconstruct a three-dimensional environment using only two channels and two loudspeakers. More recently DOLBY EX⁵ has been introduced as a means of synthesizing a center rear channel using nonlinear Prologic⁶ processing applied to the rear two channels of a five-channel system. However, this technology is aimed principally at surround sound as conceived for cinema and home theater, with a bias toward sound effects and ambience creation. Nevertheless there

exists a grey area between cinema applications, music reproduction, gaming applications, and the synthesis of virtual acoustics, especially as at their core the same multichannel carriers can link all systems. It is therefore not unreasonable to anticipate some degree of convergence as similar theoretical models apply. Also, with conventional multichannel technology it is often the listening environment and the methods used to craft the audio signals that impose the greatest performance limitations.

Multichannel stereo on DVD allows for improved methods of spatial encoding that can transcend the common studio practice of using just pairwise amplitude panning with blending to mono. It is conjectured that by including perceptually motivated processing, three-dimensional “soundscapes” can be rendered rather than just peripheral surround sound. Complex HRTF data by default encapsulate all relevant spatial information [that is, interaural amplitude difference (IAD), interaural time difference (ITD), and directional spectral weighting] and form a generalized approach. However, to reduce signal coloration, a method of HRTF equalization is proposed with an emphasis on characterizing the interear difference signal computed in the lateral plane. The extension to height information in the equalized HRTFs is also discussed briefly.

In Section 5.2.1 a special case is presented for narrow subtended angle, two-channel stereo where it is shown that ear signals derived from a real acoustic source located on the arc of the loudspeakers can be closely synthesized using a mono source with amplitude-only panning. Critically in this example, the HRTFs are defined by the actual locations of the loudspeakers, and so are matched automatically to an individual listener’s HRTF characteristics. This is an important aspect of the proposal, which is directly extendable to the method of pairwise association where mismatch sensitivity between listener HRTFs and target HRTFs is reduced. This approach also encapsulates succinctly the principles of two-channel amplitude-only panning stereo while exposing inherent errors as the angle between the loudspeakers is increased.

To summarize, four core elements constitute the proposed scalable and reverse compatible spatial audio system:

- A vector component of the sound field is produced as a loudspeaker array surrounds the listener following conventional multichannel audio practice.
- Pairwise transaural techniques are used to code directional information and to create ear signals matched to the required source signal.
- Matrix processing can increase the number of loudspeakers used in the array while simultaneously preserving the ear signals resulting from transaural processing and nonoptimum loudspeaker placements.
- Embedded digital code⁷ [12] is used to create additional channels for enhanced resolution while remaining com-

⁴ Subject to a British Telecommunications patent application.

⁵ Dolby Laboratories, channel extension technology to AC-3 perceptual coding.

⁶ Registered trademark of Dolby Laboratories.

⁷ Applicable only to the DVD-A format.

patible with the basic system already enhanced by pairwise transaural processing.

1 HRTF NOTATION

A set of HRTFs is unique to an individual and describes a continuum of acoustic transfer functions linking every point in space to the listener's ears. HRTFs depend on the relative position of the source to the listener and are influenced by distance, reflection, and diffraction around the head, pinna, and torso. In this paper HRTFs were derived from measurements taken at BT Laboratories (BTL) utilizing an artificial head and small microphones mounted at the entrance of each ear canal. Measurements of head-related impulse responses (HRIRs) were performed in an anechoic chamber at 10° intervals using a maximum-length-sequence (MLS) excitation, and the corresponding HRTFs were computed using a time window and the Fourier transform.

To define the nomenclature used for various HRTF sub-functions, consider the arrangement shown in Fig. 1, where the listener's ears are labeled A (left) and B (right) when viewed from above. In a sound reproduction system all sound sources and loudspeakers have associated pairs of HRTFs, uniquely linking them to the listener, whereas in this paper these transfer functions are called the HRTF coordinates for each object given. In Fig. 1 the single sound source X has the HRTF coordinates $\{h_{xa}, h_{xb}\}$, while the three loudspeakers 1, 2, and n , with arbitrary positions, have the coordinates $\{h_a(1), h_b(1)\}$, $\{h_a(2), h_b(2)\}$, and $\{h_a(n), h_b(n)\}$. In specifying the loudspeaker HRTF coordinates, a left-right designation can be included when the loudspeaker array is known to be symmetrical about the centerline. Consequently $h_{la}(r)$ denotes the HRTF between the left-hand loudspeaker r and the left-hand ear A. However, for arrays having only three symmetrically positioned loudspeakers (left, center, and right) a simpler notation is used in Section 3, namely, $\{h_{la}$,

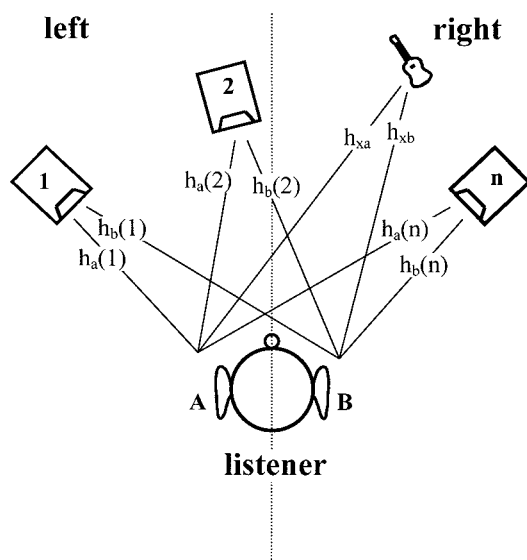


Fig. 1. Definition of HRTF notation for n loudspeakers positioned arbitrarily around the listener.

h_{lb} }, $\{h_{ca}, h_{cb}\}$, and $\{h_{ra}, h_{rb}\}$. It should be observed that in typical HRTF calculations, such as the evaluation of the positional transfer functions G_R and G_L used in transaural signal processing (see Section 3.1), ear-canal equalization need not be incorporated, provided the same set of HRTFs is used for both loudspeaker and image locations. Then the ear-canal transfer functions cancel, assuming they are not directionally encoded. For example, in Section 4 Eqs. (23a) and (23b) describe typical transaural processing to derive the positional transfer functions G_R and G_L , where any transfer function components common to all HRTFs cancel.

2 HRTF EQUALIZATION

The HRTFs used in transaural processing reveal frequency response variations that may contribute tonal coloration when sound is reproduced. In this section a strategy for equalization is studied that reduces the overall spectral variation, yet retains the key attributes deemed essential for localization. A simplified form of HRTF is also defined, which can prove useful in multiloudspeaker systems.

2.1 Methods of Equalization

When sound is reproduced over a conventional multi-loudspeaker array, where for example a signal is spatialized using pairwise amplitude panning, equalization as a function of direction is not normally employed. In such a system sound is perceived generally as uncolored, even though the ears, head, and torso impose direction-specific spectral weighting. However, although HRTFs used in transaural processing take account of both the source location and the loudspeakers, reducing frequency response variations can ameliorate tonal variance, which may become accentuated as the phantom image moves away from the loudspeaker locations and when the listener turns away from the optimum forward orientation. Also, systems are rarely optimally aligned and exhibit sensitivity to small head motions, both of which map into frequency response errors in the reconstructed ear signals. Consequently the aim is to introduce minimum spectral modifications commensurate with achieving spatialization.

It is proposed that for image localization within the lateral plane the relationship between the complex interaural difference signal and the signal components common to both ears is the critical factor. This conjecture is based on the premise that for lateral images, spectral components common to both ears relate closely to the source spectrum whereas the interaural spectrum is strongly influenced by source direction. Consequently it is argued that modification to the common spectrum causes principally tonal coloration, whereas the relationship between common spectrum and interaural spectrum is more critical to localization, even though spectral cues embedded in the source can induce an illusion of height. This approach may be extendable to include height localization, although it is recognized that additional spectral weighting of the monaural component can be required following, for example, the boosted-band experiments performed by Blauert [13].

2.1.1 Lateral-Plane HRTF Equalization

The proposed method of lateral-plane HRTF equalization first transforms each HRTF pair into sum and difference (M-S) coordinates and then performs equalization on the corresponding pair by dividing by the corresponding sum spectrum. It is proposed that all HRTFs in a set should be equalized using this technique in order to maintain relative group delay and, with appropriate weighting, relative level.

As defined in Section 1, let the HRTFs for a given source location X be h_{xa} and h_{xb} , and let the corresponding complex sum and difference transforms be $HSUM_x$ and $HDIFF_x$. Thus

$$HSUM_x = h_{xa} + h_{xb} \quad (1)$$

$$HDIFF_x = h_{xa} - h_{xb} \quad (2)$$

Four methods of HRTF equalization that match this objective are identified, where $\{h_{xea}, h_{xeb}\}$ are the resulting HRTFs after equalization.

Method 1: Equalization by the modulus of the complex sum spectrum,

$$h_{xea} = \frac{h_{xa}}{|h_{xa} + h_{xb}|} W_{nx} \quad (3a)$$

$$h_{xeb} = \frac{h_{xb}}{|h_{xa} + h_{xb}|} W_{nx} \quad (3b)$$

Method 2: Equalization by complex sum spectrum,

$$h_{xea} = \frac{h_{xa}}{h_{xa} + h_{xb}} W_{nx} \quad (4a)$$

$$h_{xeb} = \frac{h_{xb}}{h_{xa} + h_{xb}} W_{nx} \quad (4b)$$

Method 3: Equalization by the derived minimum-phase spectrum of the complex sum spectrum,

$$h_{xea} = \frac{h_{xa}}{\exp(\text{conj}(\text{hilbert}(\log(\text{abs}(h_{xa} + h_{xb}))))))} W_{nx} \quad (5a)$$

$$h_{xeb} = \frac{h_{xb}}{\exp(\text{conj}(\text{hilbert}(\log(\text{abs}(h_{xa} + h_{xb}))))))} W_{nx} \quad (5b)$$

Here W_{nx} are the normalization coefficients calculated to maintain the relative levels after equalization of all HRTF coordinates in the set. Each form of equalization delivers identical magnitude spectra in the HRTFs, although there are variations in the time-domain waveforms resulting from phase response differences. To illustrate these variations, consider an example HRTF pair corresponding to a nominal 30° off-axis image source. Fig. 2(a) shows the measured HRIRs, whereas Fig. 2(b)–(d) presents the impulse responses resulting from each form of equalization in order that both pre- and postering can be compared. Fig. 3 shows the corresponding amplitude spectra before

and after equalization, and Fig. 4 illustrates the sum and difference spectra, again before and after equalization.

In selecting a potential equalization strategy, it is a necessary condition that the relative time difference between HRTF pairs be maintained. Also, the time-domain waveforms should not accentuate or exhibit excessive pre- or postering, as this can produce unnatural sound coloration. Although each equalization method meets the principal objective, the technique of forging the denominator from the minimum phase of the sum spectrum yields results with minimum pre- or postering. In essence, the minimum-phase information common to both ear signals is removed, leaving mainly excess phase components to carry the essential time-delay information. Equalization using the complex sum spectrum (method 2) also yields results close to the requirements. However, inspection of Fig. 2(c) shows that the right-ear response, which in this case has the greater delay, exhibits pre- or postering extending back in time to the commencement of the left HRIR.

However, experience gained with equalization has revealed that certain image locations, particularly toward the center rear, can yield excessive ringing after equalization. Consequently a further equalization variant is proposed. This is similar to method 3, but it differs in the way the sum spectrum is computed and is defined as follows.

Method 4: Equalization by the derived minimum-phase sum of the moduli of each complex spectrum,

$$h_{xea} = \frac{h_{xa}}{\exp(\text{conj}(\text{hilbert}(\log(\text{abs}(h_{xa}) + \text{abs}(h_{xb}))))))} W_{nx} \quad (6a)$$

$$h_{xeb} = \frac{h_{xb}}{\exp(\text{conj}(\text{hilbert}(\log(\text{abs}(h_{xa}) + \text{abs}(h_{xb}))))))} W_{nx} \quad (6b)$$

In the denominator this algorithm factors out the interaural time difference between left and right signals, which otherwise map into artificial amplitude response variations in the complex sum spectrum. As such this procedure could be argued to be a better estimator of the common spectrum, as human auditory processing does not sum ear signals directly. Overall the effect on HRTFs is minor. Fig. 5 presents results that should be compared directly with those in Fig. 3.

Finally a further variant of equalization is where an average of all sum spectra is formed and the HRTFs are modified following procedures similar to those reported in this section but with particular emphasis on the minimum-phase and sum-of-moduli techniques. However, in this case, since all HRTFs are modified by a common equalization function in a way similar to ear-canal equalization, when positional transfer functions are calculated, their form is unchanged.

2.1.2 Equalization with the Addition of Height Cues

Research by Blauert [13] has shown that by introducing specific frequency-dependent characteristics into the HRTFs a sensation of height is achievable. However, the

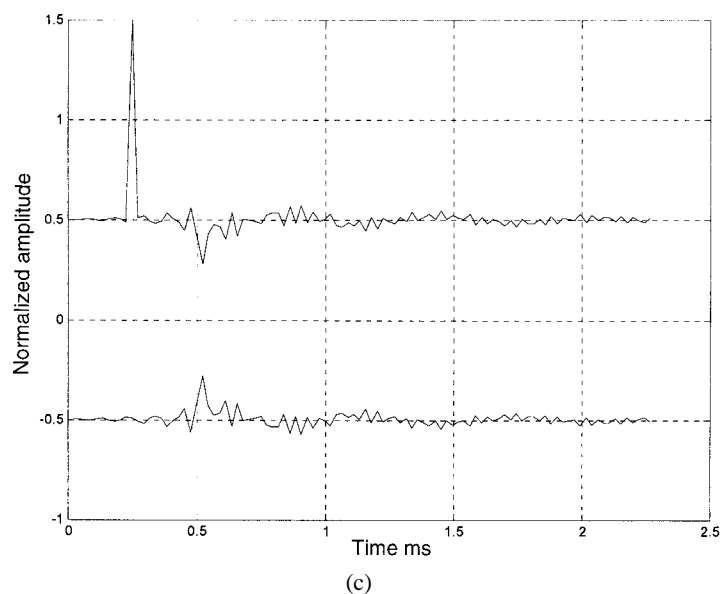
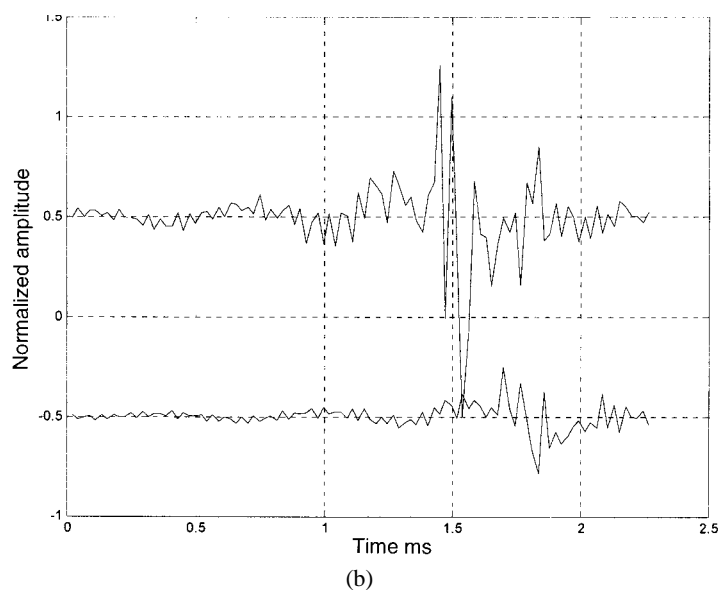
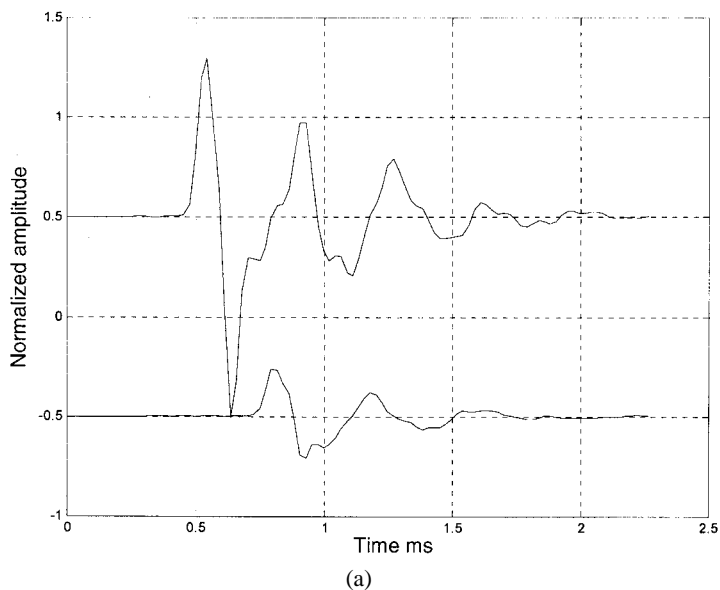


Fig. 2. Normalized time-domain left–right HRTFs at 30°. Top—left ear; bottom—right ear. (a) Measured. Observe relative time displacement revealing ITD and lack of pre-ringing in natural responses. (b) Method 1 equalized. Observe excessive pre-ringing that blurs commencement of the two HRIRs. (c) Method 2 equalized. HRIRs exhibit mirror images except for initial impulse. (d) Method 3 equalized. Pre-ringing reduced and initial ITD of HRIRs maintained.

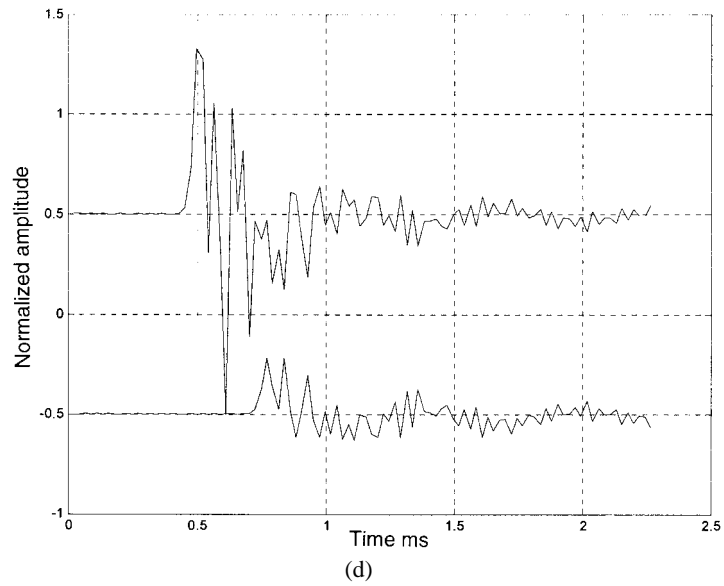


Fig. 2. Continued

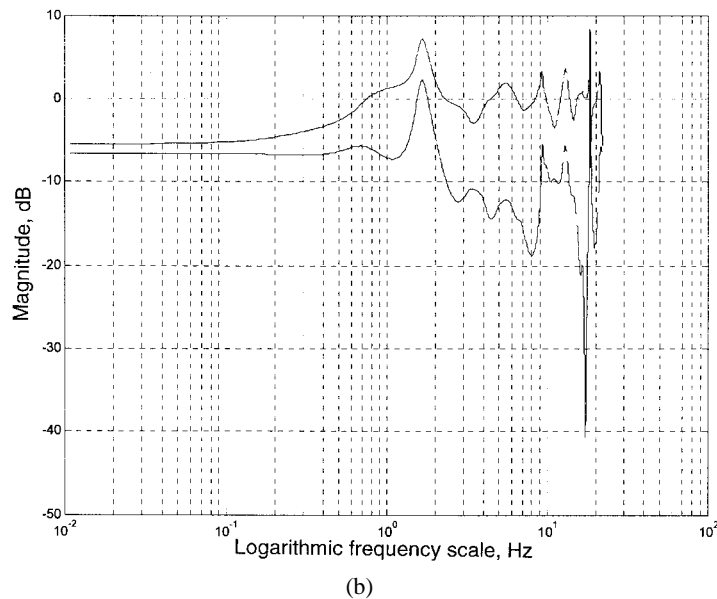
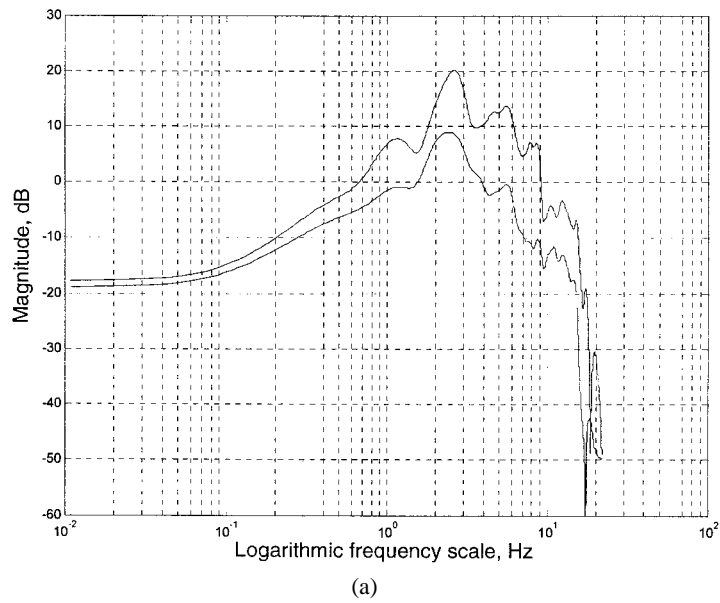
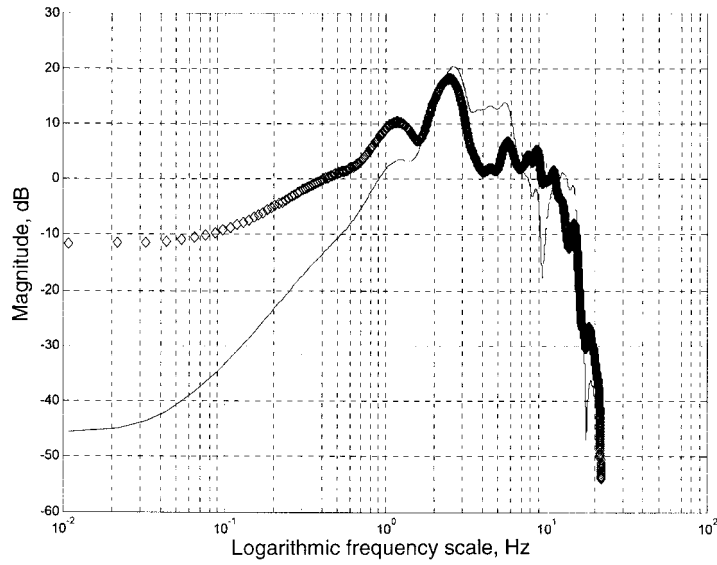
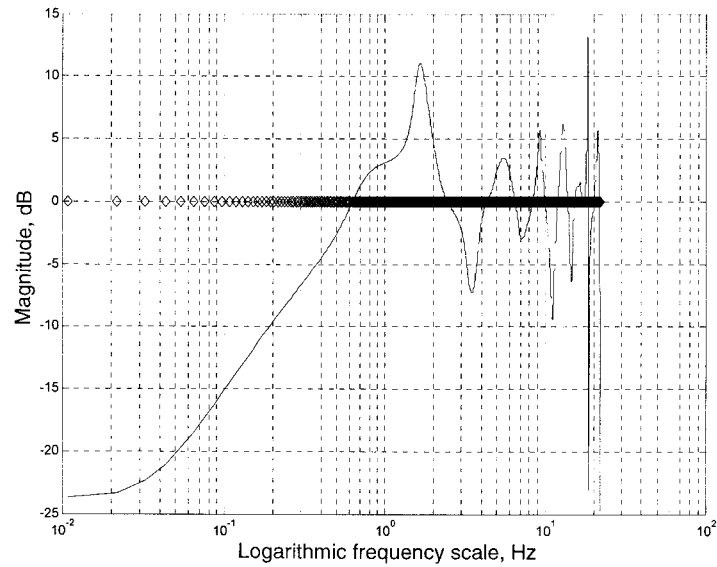


Fig. 3. Magnitude HRTF pair at 30°. Top—left ear; bottom—right ear. (a) Measured. (b). Equalized. Results are identical for methods 1, 2, and 3.



(a)



(b)

Fig. 4. Magnitude HRTF sum and difference spectra at 30°. Sum—"diamond" line; difference—continuous line. (a) Unequaled. (b) Equalized, applicable to methods 1, 2, and 3. (Note constant-level sum spectrum following equalization.)

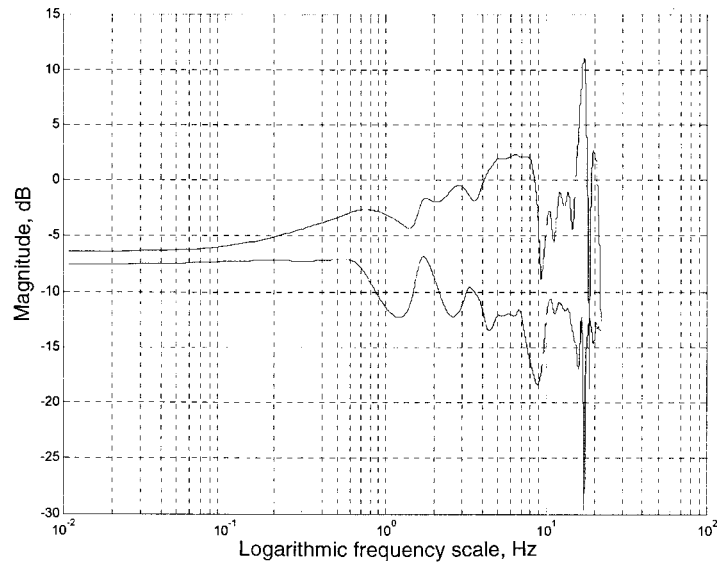


Fig. 5. Equalized HRTF pair at 30°, applicable to method 4 only. Top—left ear; bottom—right ear.

question arises as to whether this modification is compatible with the equalization strategies presented in Section 2.1.1. For example, the following questions need to be considered:

- Is it sufficient to measure the HRTF coordinates only at the required location above the lateral plane and then apply equalization, and will then sufficient information remain buried in the interaural difference signal with a unique characterization to discriminate against lateral images with equivalent interaural time differences?
- Does the absolute amplitude response variation, rather than just the difference response variation inherent in the HRTFs, represent a major factor in producing height cues?
- Are there secondary factors, such as ground reflections, which introduce additional cues to aid height localization? Effectively this would require at least two interfering sets of HRTFs to be summed.

A full investigation of these points relating to height is beyond the scope of the present study. However, if the ground reflection model were responsible, then the equalization methods could be applied individually to the direct source and to the ground reflection, with the results combined by taking the path difference into account.

2.2 Simplified HRTF Models

In applications where phantom images are positioned close to the locality of the loudspeakers, it may be sufficient to use a simple form of HRTF. This is particularly applicable with multiloudspeaker arrays where a vector component already forms a strong localization clue. Fig. 6 shows a source image at angle θ defined by h_{xa} and h_{xb} with respect to a human head of diameter d meters.

2.2.1 Simple HRTF Model 1

The model ignores head shadowing and assumes that only the interaural time difference is significant. Hence for

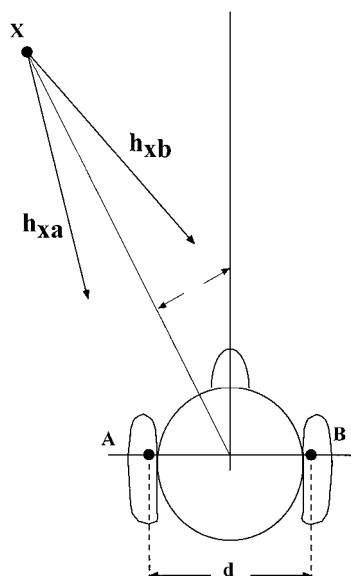


Fig. 6. Sound source X and simplified head model used to derive approximate HRTFs.

a source at angle θ from the forward position, the respective HRTF coordinates are approximately

$$h_{xa} = \exp \left\{ -j2\pi f \left[T - \frac{d}{2c} \sin(\theta) \right] \right\} \quad (7a)$$

$$h_{xb} = \exp \left\{ -j2\pi f \left[T + \frac{d}{2c} \sin(\theta) \right] \right\} \quad (7b)$$

where the velocity of sound is c m/s and T s is the time delay from the source to the center of the head. In this model the advantage of equalization method 3 is evident as no equalization need be applied.

2.2.2 Simple HRTF Model 2

In this second model both the front and the back waves are considered, where the back wave results from head defraction. In this representation a wave incident on ear A produces, by head defraction, a secondary signal at ear B. The head diffraction transfer function from ears A to B, $DH_{A-B}(r, \theta, \phi)$, is a function of the direction of the incident wave defined by the spherical coordinates $\{r, \theta, \phi\}$. A similar function linking ears B to A is defined, $DH_{B-A}(r, \theta, \phi)$. Hence for the source HRTF coordinates $\{h_{xa}, h_{xb}\}$,

$$h_{xa} = \exp \left\{ -j2\pi f \left[T - \frac{d}{2c} \sin(\theta) \right] \right\} + DH_{B-A}(r, \theta, \phi) \exp \left\{ -j2\pi f \left[T + \frac{d}{2c} \sin(\theta) \right] \right\} \quad (8a)$$

$$h_{xb} = \exp \left\{ -j2\pi f \left[T + \frac{d}{2c} \sin(\theta) \right] \right\} + DH_{A-B}(r, \theta, \phi) \exp \left\{ -j2\pi f \left[T - \frac{d}{2c} \sin(\theta) \right] \right\}. \quad (8b)$$

In a simple model the diffraction transfer functions could be represented as attenuation DH_k with a time delay of approximately the interaural time delay ΔT_{A-B} ,

$$DH_{A-B}(r, \theta, \phi) = DH_{B-A}(r, \theta, \phi) \approx DH_k \exp(-j2\pi f \Delta T_{A-B}). \quad (9)$$

3 MULTILOUSPEAKER ARRAYS IN TWO-CHANNEL STEREO

This section introduces variants to two-channel, two-loudspeaker transaural processing to demonstrate how a two-channel signal format can be mapped into n feeds to drive a multiloudspeaker array [14]. It is assumed that more than two loudspeakers are driven simultaneously by signals derived from a single-point sound source, while formal methods show that the correct ear signals can be retained. Besides supporting stand-alone applications, these transformations are relevant in the development of multichannel transaural stereo, as described in Section 5.

The outputs of an n -array of loudspeakers combine by acoustical superposition at the entrance to each ear canal. The principal condition for accurate sound localization is that these signals match the signals that would have been generated by a real sound source, both in the static case and in the case for small head rotations. Also, by using several loudspeakers placed to surround the listener, sound-field direction can make the system more tolerant to head motion. Consequently changes in ear signals with head motion match more closely those of a real image.

A static sound source, whatever its size and physical location, produces two ear signals that fully define the event, provided the relative head position to source is fixed. In practice it is possible to generate the correct ear signal from two or more noncoincident loudspeakers that can take any arbitrary position around the head. However, if the position of the head moves, then a change in the ear signals results, which no longer match the phantom image correctly, and a localization error is perceived. In a system of wavefront reconstruction this distortion is minimized, although the penalty is a large number of loudspeakers and channels. However, if a limited number n of loudspeakers is used (for example, $n = 12$), then although image position distortion still occurs, the effect is reduced as there is a robust directional component. Also, when a phantom image coincides with a loudspeaker position, the positional distortion as a function of head position tends to zero, although it is debatable whether this is a desirable situation as images from other locations are represented differently in terms of their radiated cones of sound.

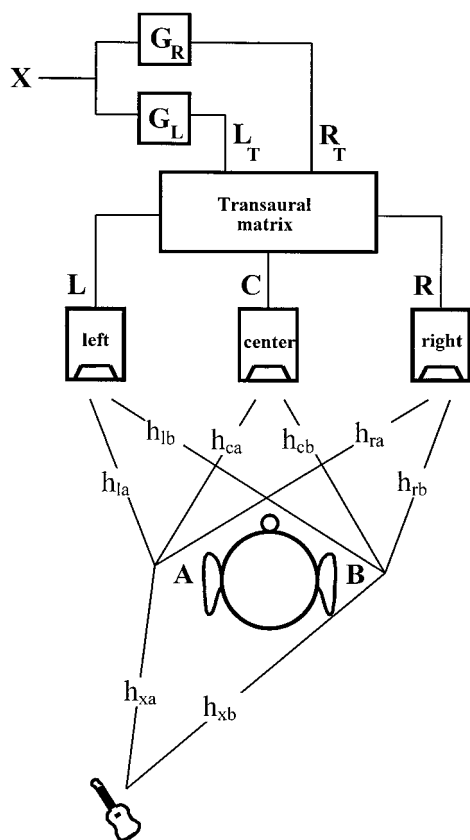


Fig. 7. Transaural processing using symmetrical three-loudspeaker array.

3.1 Three-Loudspeaker Transaural Processing

To illustrate how to accommodate more than two loudspeakers in an array while retaining the requirements for precise HRTF formulation, consider a three-loudspeaker array as illustrated in Fig. 7. In this system a mono source signal X is filtered by the positional transfer functions G_R and G_L to form L_T and R_T , which in turn form inputs to the matrix $[a]$. By way of example a Trifield⁸ matrix (after Gerzon [15]) is selected, which is defined as

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} 0.8850 & -0.1150 \\ 0.4511 & 0.4511 \\ -0.1150 & 0.8850 \end{bmatrix}. \tag{10}$$

By applying the coefficients defined in matrix $[a]$, the three loudspeaker signals L , C , and R (left, center, right) can be derived. However, to reproduce optimum localization, the system requires that the ear signals produced by the three loudspeakers match the ear signals that would be produced by the real source.

3.1.1 Analysis

The positional transfer function matrix $[G]$ converts the mono signal X to L_T and R_T as

$$\begin{bmatrix} L_T \\ R_T \end{bmatrix} = \begin{bmatrix} G_R \\ G_L \end{bmatrix} X. \tag{11}$$

Using matrix $[a]$, the loudspeaker feeds L , C , and R are then derived from $[G]X$,

$$\begin{bmatrix} L \\ C \\ R \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} G_L \\ G_R \end{bmatrix} X. \tag{12}$$

However, recalling the HRTFs as defined in Fig. 7, where h_{xa} and h_{xb} are the HRTF coordinates of source image X , then

$$\begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} X = \begin{bmatrix} h_{la} & h_{ca} & h_{ra} \\ h_{lb} & h_{cb} & h_{rb} \end{bmatrix} \begin{bmatrix} L \\ C \\ R \end{bmatrix} \tag{13}$$

where, substituting for L , C , and R ,

$$\begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} = \begin{bmatrix} h_{la} & h_{ca} & h_{ra} \\ h_{lb} & h_{cb} & h_{rb} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} G_L \\ G_R \end{bmatrix}. \tag{14}$$

The positional transfer functions G_R and G_L then follow by matrix inversion,

$$\begin{bmatrix} G_L \\ G_R \end{bmatrix} = \left\{ \begin{bmatrix} h_{la} & h_{ca} & h_{ra} \\ h_{lb} & h_{cb} & h_{rb} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right\}^{-1} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} \tag{15}$$

⁸Registered trademark describing a two-channel to three-loudspeaker mapping proposed by Gerzon [15].

from which the loudspeaker feeds L , C , and R are calculated,

$$\begin{bmatrix} L \\ C \\ R \end{bmatrix} = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} \right\} \left\{ \begin{bmatrix} h_{la} & h_{ca} & h_{ra} \\ h_{lb} & h_{cb} & h_{rb} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right\}^{-1} X. \quad (16)$$

In practice L , C , and R are calculated directly using matrix inversion. However, because the transfer functions can have several thousand elements, to avoid large-dimension matrices the solution can be decomposed as follows. Define

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} h_{la} & h_{ca} & h_{ra} \\ h_{lb} & h_{cb} & h_{rb} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \quad (17)$$

giving

$$\begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} G_L \\ G_R \end{bmatrix} \quad (18)$$

where, using matrix inversion, the positional transfer functions are

$$G_R = \frac{h_{xa} h_{21} - h_{xr} h_{11}}{h_{12} h_{21} - h_{11} h_{22}} \quad (19a)$$

$$G_L = \frac{h_{xa} h_{12} - h_{xb} h_{22}}{h_{12} h_{21} - h_{11} h_{22}} \quad (19b)$$

which enable L , C , and R to be calculated. Fig. 8 shows example transfer functions linking the system input to the three loudspeaker inputs (L , C , and R) located at -45° , 0° , and 45° , with a source location at 30° . Simulations confirm that the correct ear signals are produced as shown in Fig. 9, while Figs. 10 and 11 present the magnitudes of the positional transfer functions G_L and G_R and their differential phase response, respectively.

3.2 Three-Loudspeaker Matrix with Band-Limited Center Channel

This section extends multiloudspeaker transaural processing by considering a case where the center channel is band-limited by a low-pass filter with a transfer function $\lambda(f)$. For example, $\lambda(f)$ could constrain the center channel to operate only in the band where the ear and brain employ interaural time differences for localization. Alternatively the center channel may be used mainly for low-frequency reproduction.

The inclusion of $\lambda(f)$ yields effective HRTF center channel coordinates $\{h_{ca} * \lambda(f), h_{cb} * \lambda(f)\}$, where $*$ implies element-by-element vector multiplication. Hence, from the equations derived in Section 3.1, L , C , and R follow,

$$\begin{bmatrix} L \\ C \\ R \end{bmatrix} = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} \right\} \left\{ \begin{bmatrix} h_{la} & h_{ca} * \lambda(f) & h_{ra} \\ h_{lb} & h_{cb} * \lambda(f) & h_{rb} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right\}^{-1} \begin{bmatrix} 1 \\ \lambda(f) \\ 1 \end{bmatrix} X. \quad (20)$$

To illustrate this system with band-limited center channel, Fig. 12 shows again the system input to the loudspeaker transfer functions for the three loudspeakers located at -45° , 0° , and 45° , with a phantom source location at 30° . The low-pass filter $\lambda(f)$ in the center channel has a cutoff frequency of 100 Hz with an asymptotic attenuation slope of 40 dB per octave. The ear signals are formed correctly and are identical to those presented in Fig. 9.

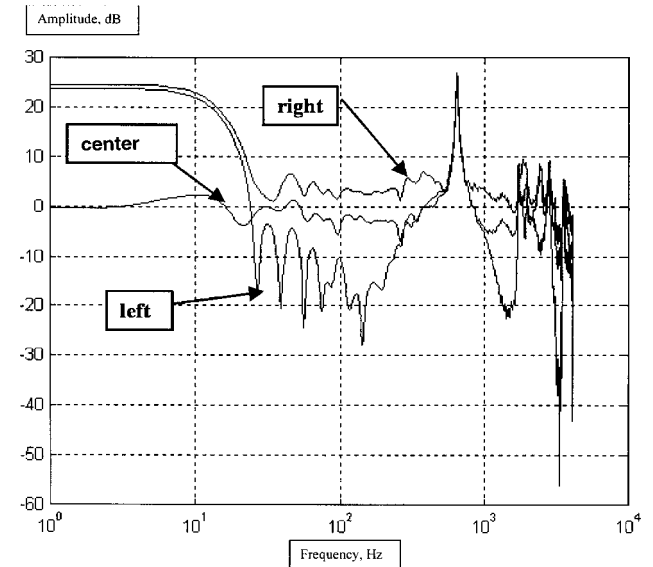


Fig. 8. Loudspeaker feed transfer functions for Trifield matrix linking input to three loudspeakers located at -45° , 0° , and 45° , image at 30° .

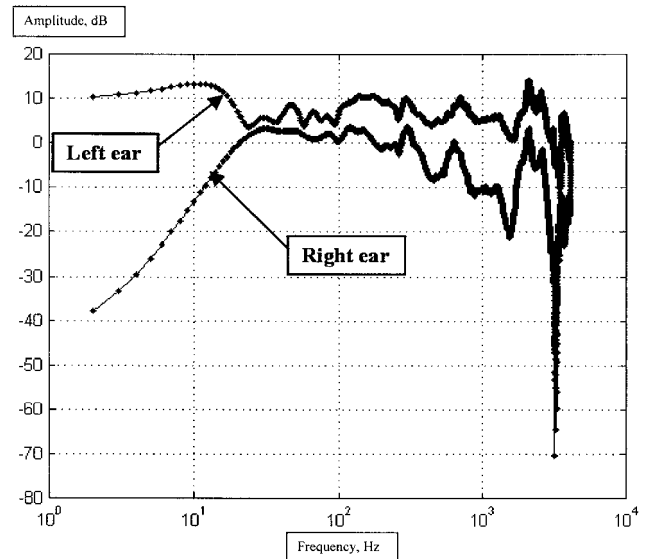


Fig. 9. Input-to-ear transfer functions for three-loudspeaker system with Trifield matrix, corresponding to functions shown in Fig. 8.

An attraction of this configuration is that the center channel can have a limited bandwidth while offering improvements in bass quality both in terms of power handling and by improving modal dispersion in the listening room. Alternatively, if a loss of spatial resolution at low frequency is permitted, then the center channel could function as a subwoofer with an upper response that extends only into the lower midband frequency range. The left- and right-hand loudspeakers would extend to high frequencies, although with restricted low-frequency performance.

3.3 *n*-Loudspeaker Array with Two-Channel Transaural Processing

The method of using more than two loudspeakers can be generalized to *n* loudspeakers while retaining only two information channels, where for example the loudspeakers surround the listener in a symmetrical array. The left- and right-hand loudspeakers in the array are fed by one of two information signals, and each loudspeaker has individual weighting defined by a coefficient matrix [*a*]

$$\begin{bmatrix} \text{LS}(1) \\ \text{LS}(2) \\ \vdots \\ \text{LS}(n) \end{bmatrix} = \begin{bmatrix} a_1 & a_1 \\ a_2 & a_2 \\ \vdots & \vdots \\ a_n & a_n \end{bmatrix} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} \begin{bmatrix} h_a(1) & h_a(2) & \dots & h_a(n) \\ h_b(1) & h_b(2) & \dots & h_b(n) \end{bmatrix} \begin{bmatrix} a_1 & a_1 \\ a_2 & a_2 \\ \vdots & \vdots \\ a_n & a_n \end{bmatrix}^{-1} X \quad (21)$$

Although this matrix equation cannot be solved in general as there are too many independent variables, solutions can be achieved when the matrix [*a*] is specified. For general two-channel stereophonic reproduction this system offers little advantage. However, in a telepresence and teleconference environment the coefficient matrix [*a*] may be transmitted for a given talker alongside the two information channels. A transaural reproduction system can then be conceived, where the coefficients are updated dynamically to enhance directional coding. This becomes particularly attractive where there are a number of talkers, as the

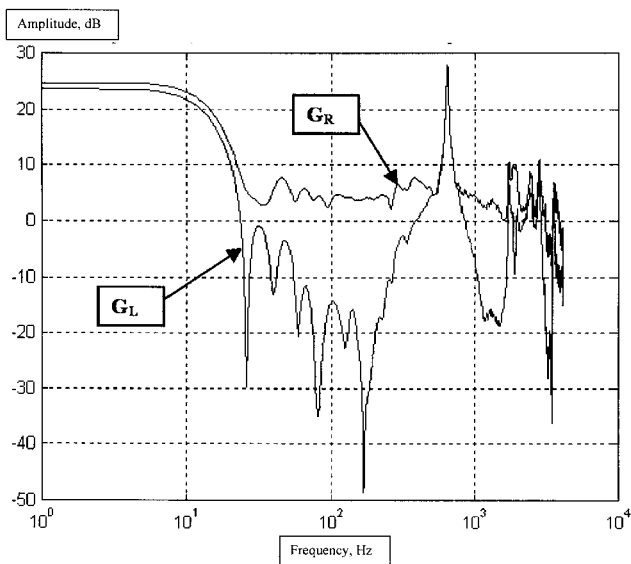


Fig. 10. Positional transfer functions G_L and G_R , corresponding functions shown in Fig. 8.

coefficient matrix could be adjusted dynamically to enhance localization.

4 MULTICHANNEL PARADIGM EXPLOITING PAIRWISE TRANSAURAL STEREO

This section reviews a spatial audio paradigm that links multichannel audio and transaural processing. The technique augments the directional clues inherent in multichannel stereo reproduction by embedding HRTF data such that the ear signals are matched more accurately to those produced by the source image. By default, such processing includes both frequency (interaural amplitude response) and time (interaural time response) information and therefore forms an elegant method of virtual image manipulation. Also, because HRTFs vary with both angular position and distance, sound sources can be synthesized and manipulated in three-dimensional space, together with

reflections spatialized using their HRTF coordinates, which further enhances this process. In a multichannel system this processing is performed during source coding and is therefore compatible with all DVD formats.

The proposal operates at six principal levels:

- Selecting a pair of loudspeakers whose subtended angle includes the position of the phantom image helps reinforce the sound direction and matches conventional mixing practice for localization in multichannel systems.
- Encoded amplitude differences in signals above about 2 kHz support localization using interaural amplitude differences.
- Encoded time differences in signals below about 2 kHz support localization using interaural time differences where an extended bass performance is desirable.
- The addition of transaural processing based on HRTF data enables the construction of ear signals that match the original event and aids localization.
- Closer spacing of loudspeakers in a multiloudspeaker array reduces sensitivity to the precise form of HRTF characteristics, thus making an averaged HRTF set more applicable to a wide range of listeners.
- The effect of moderate head motion, which is a desirable attribute for improving localization, is supported. For relatively small loudspeaker subtended angles the error in ear-signal reconstruction is reduced when the head is moved by a small angle such that the vector component reinforces localization.

As an example, consider a circular array of n loudspeakers, as shown in Fig. 13. In this system pairwise coding (PWC) selects the two closest loudspeakers such that an image X falls within the subtended angle at the listening position. Two-channel HRTF synthesis is then used to form the optimum ear signals. For example, if loudspeakers r and $r + 1$ are selected from the n -array of loudspeakers, then loudspeaker feeds $LS(r)$ and $LS(r + 1)$ are computed,

$$\begin{aligned} \begin{bmatrix} LS(r) \\ LS(r + 1) \end{bmatrix} &= \begin{bmatrix} h_a(r) & h_a(r + 1) \\ h_b(r) & h_b(r + 1) \end{bmatrix}^{-1} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} X \\ &= \begin{bmatrix} G_r \\ G_{(r+1)} \end{bmatrix} X. \end{aligned} \quad (22)$$

Matrix $[G]$ defines a set of positional transfer functions, which are effectively filters located between source and loudspeaker feed, where the HRTF notation was defined in Section 1 with reference to Fig. 1.

Solving for the positional transfer function matrix $[G]$,

$$G_r = \frac{h_{xa} h_b(r) - h_{xb} h_a(r)}{h_a(r + 1) h_b(r) - h_b(r + 1) h_a(r)} \quad (23a)$$

$$G_{r+1} = \frac{h_{xb} h_a(r + 1) - h_{xa} h_b(r + 1)}{h_a(r + 1) h_b(r) - h_b(r + 1) h_a(r)}. \quad (23b)$$

This result can be generalized for an n -array of loudspeakers as

$$\begin{bmatrix} LS(1) \\ LS(2) \\ \vdots \\ LS(n) \end{bmatrix} = \begin{bmatrix} a_1 & a_1 \\ a_2 & a_2 \\ \vdots & \vdots \\ a_n & a_n \end{bmatrix} \begin{bmatrix} h_{xa} \\ h_{xb} \end{bmatrix} \begin{bmatrix} h_a(1) & h_a(2) & \dots & h_a(n) \\ h_b(1) & h_b(2) & \dots & h_b(n) \end{bmatrix}^{-1} X = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \end{bmatrix} X. \quad (24)$$

If a sound source falls between loudspeakers r and $r + 1$, then $a_r = a_{r+1} = 1$; otherwise all remaining coefficients in matrix $[a]$ are set to zero. Consequently for a sound source to circumnavigate the head, the HRTF coordinates h_{xa} and h_{xb} must change dynamically, whereas as the source moves between loudspeaker pairs, the coefficient matrix is switched to redirect the sound.

A four-channel, four-loudspeaker PWC scheme is shown in Fig. 14. The positional transfer functions $\{G_1, G_2, G_3, G_4\}$ are calculated for each source location, which then filters the source signal X to form the loudspeaker feeds. Because transaural processing is performed at the encoder, a simple replay system is supported. Consequently complete compatibility with conventional multi-channel audio is retained.

5 INCREASED NUMBERS OF LOUDSPEAKERS

An increase in the number of loudspeakers can achieve more even sound distribution, distribute power handling, and possibly lower the sensitivity to room acoustics. This section considers methods by which the number of loudspeakers in an array can be increased.

In the basis system, where loudspeakers are linked directly to information channels located at coordinates compatible with the encoding HRTF coordinates, the loudspeakers are designated nodal loudspeakers. An n -

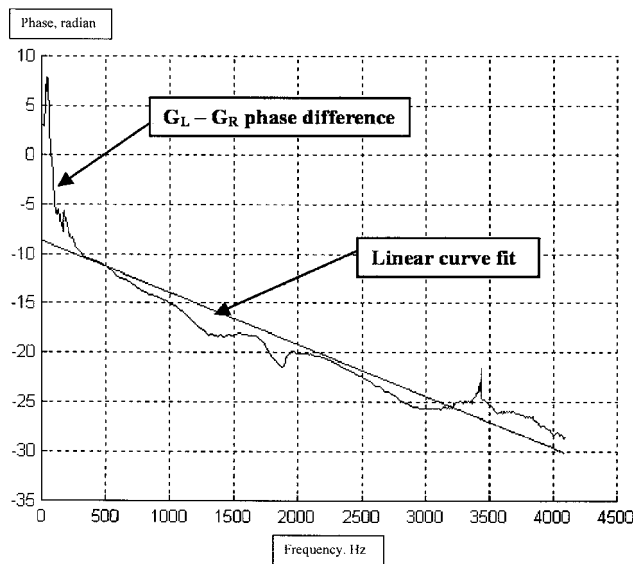


Fig. 11. Differential phase response between positional transfer functions G_L and G_R corresponding to functions shown in Fig 8.

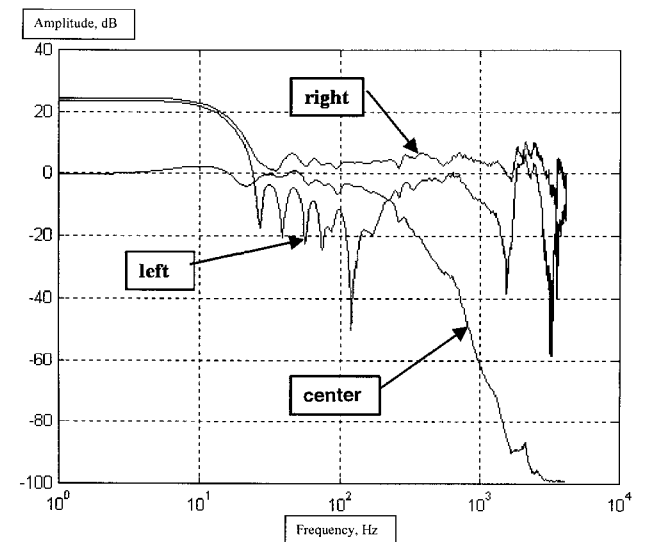


Fig. 12. Transfer functions linking input to three-loudspeaker feeds for Trifield matrix, but with center channel band-limited to 100 Hz, corresponding to functions shown in Fig. 8.

channel system therefore has n nodal loudspeakers, which constitute the basis array, with the corresponding drive signals, or primary signals, collectively forming the primary signal set. Loudspeakers in addition to the nodal loudspeakers are termed secondary loudspeakers.

5.1 Compensation for Inclusion of Secondary Loudspeakers

The objective is to derive additional signals within the decoder to drive secondary loudspeakers located between the nodal loudspeakers. However, the ear signals must be conserved and theoretically remain identical to the case where only the nodal loudspeakers are present. It is assumed here that all loudspeakers in the array have identical transfer functions and therefore do not affect the decoder process. If this is not the case, then they require

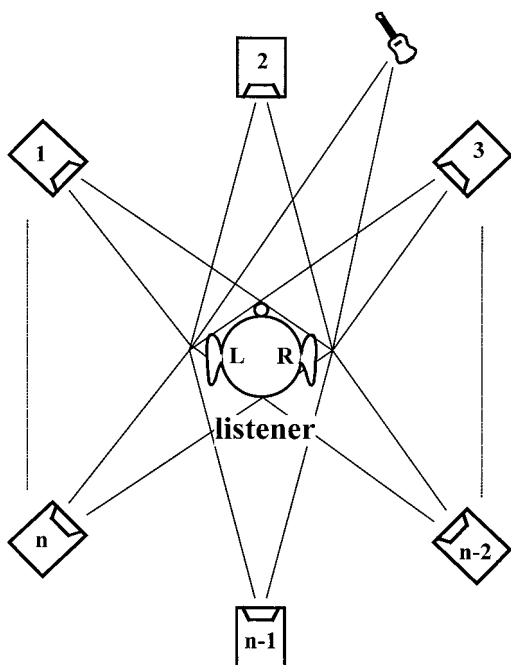


Fig. 13. n -loudspeaker array, suitable for transaural pairwise stereo.

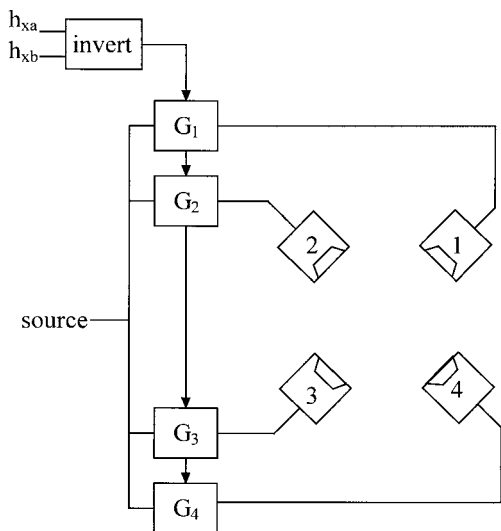


Fig. 14. Four-loudspeaker array with pairwise transaural synthesis.

individual correction. A sector of such an array is shown in Fig. 15. In this array nodal loudspeakers r and $r + 1$ have the respective HRTF coordinates $[h_a(r), h_b(r)]$ and $[h_a(r + 1), h_b(r + 1)]$ at the listener, whereas the single secondary loudspeaker p has the HRTF coordinates $[h_a(p), h_b(p)]$.

The synthesis of the drive signal for secondary loudspeaker p employs two weighting functions λ_{p1} and λ_{p2} applied to the respective primary signals LS_r and LS_{r+1} such that LS_p , the drive signal to the secondary loudspeaker p , is

$$LS_p = \lambda_{p1} LS_r + \lambda_{p2} LS_{r+1} \tag{25}$$

However, when the secondary loudspeaker enters the array, it is necessary to compensate the output from the two adjacent nodal loudspeakers in order that the ear signals remain unchanged. Ideally this needs to be achieved without knowledge of the encoding parameters. Otherwise the modified array cannot be used universally in a multi-channel audio system. A scheme capable of meeting this objective is shown in Fig. 16, where the compensation transfer functions γ_{p1} and γ_{p2} filter the signal LS_p to yield signals that are added to LS_r and LS_{r+1} .

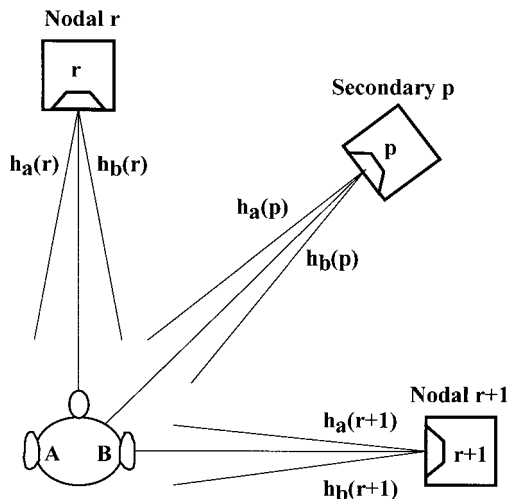


Fig. 15. Nodal loudspeakers with additional secondary loudspeaker.

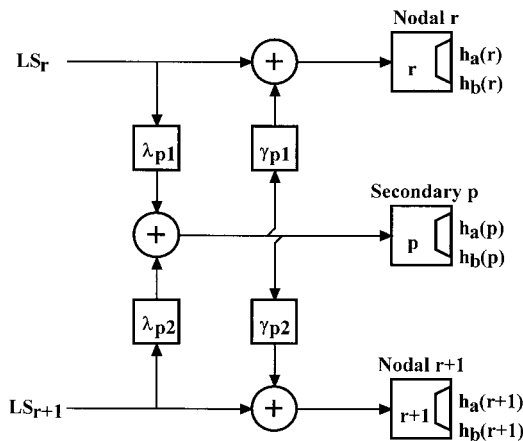


Fig. 16. Decoder processing to compensate for secondary loudspeaker p .

5.1.1 Analysis

Consider initially the case when the secondary loudspeaker p is absent from the array. The left- and right-ear signals e_a and e_b are expressed in terms of the HRTFs corresponding to the two adjacent nodal loudspeakers r and $r + 1$,

$$e_a = \text{LS}_r h_a(r) + \text{LS}_{r+1} h_a(r + 1) \quad (26a)$$

$$e_b = \text{LS}_r h_b(r) + \text{LS}_{r+1} h_b(r + 1). \quad (26b)$$

When the secondary loudspeaker p is introduced into the array, the modified ear signals e'_a and e'_b become

$$e'_a = \text{LS}_r h_a(r) + \text{LS}_{r+1} h_a(r + 1) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} \gamma_{p1} h_a(r) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} \gamma_{p2} h_a(r + 1) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} h_a(p) \quad (27a)$$

$$e'_b = \text{LS}_r h_b(r) + \text{LS}_{r+1} h_b(r + 1) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} \gamma_{p1} h_b(r) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} \gamma_{p2} h_b(r + 1) + \left\{ \text{LS}_r \lambda_{p1} + \text{LS}_{r+1} \lambda_{p2} \right\} h_b(p). \quad (27b)$$

Forcing $e'_a = e_a$ and $e'_b = e_b$, then

$$\begin{bmatrix} h_a(r) & h_a(r + 1) \\ h_b(r) & h_b(r + 1) \end{bmatrix} \begin{bmatrix} \gamma_{p1} \\ \gamma_{p2} \end{bmatrix} = - \begin{bmatrix} h_a(p) \\ h_b(p) \end{bmatrix} \quad (28)$$

from which the correction functions γ_{p1} and γ_{p2} follow,

$$\gamma_{p1} = - \left[\frac{h_a(p) h_b(r + 1) - h_b(p) h_a(r + 1)}{h_a(r) h_b(r + 1) - h_b(r) h_a(r + 1)} \right] \quad (29a)$$

$$\gamma_{p2} = - \left[\frac{h_b(p) h_a(r) - h_a(p) h_b(r)}{h_a(r) h_b(r + 1) - h_b(r) h_a(r + 1)} \right]. \quad (29b)$$

These equations reveal that the correction functions γ_{p1} and γ_{p2} depend only on the HRTFs corresponding to the loudspeaker array. Consequently they are purely a function of the replay system, and its loudspeaker layout thus can be computed within the replay decoder as part of the installation procedure. However, the weighting functions γ_{p1} and γ_{p2} can be selected independently, provided the system is stable.

5.2 Relationship of System Parameters to Loudspeaker HRTFs

Section 5.1 presented an analysis of decoder parameters where precise HRTF measurement data are known for each loudspeaker position. However, there are some interesting observations and simplifications that can be made, which are considered in this section.

5.2.1 HRTFs Derived Using Linear Interpolation

Consider a pair of nodal loudspeakers within an n -array PWC system where the proximity of loudspeakers r and $r + 1$ is such that the image source HRTFs can be approx-

imated by linear interpolation, such that

$$h_{xa} = \beta(m_r) \left\{ m_r h_a(r) + (1 - m_r) h_a(r + 1) \right\} \quad (30a)$$

$$h_{xb} = \beta(m_r) \left\{ m_r h_b(r) + (1 - m_r) h_b(r + 1) \right\} \quad (30b)$$

where m_r is the panning variable with a range of 1 to 0, which corresponds to an image pan from loudspeaker r to $r + 1$, and the function $\beta(m_r)$ is a moderator chosen to achieve a constant subjective loudness with variations in m_r . By substitution, the positional transfer functions G_r ,

and G_{r+1} then simplify to

$$G_r = m_r \beta(m_r) \quad (31a)$$

$$G_{r+1} = (1 - m_r) \beta(m_r). \quad (31b)$$

Consequently, for an image source that lies on a radial arc between the two nodal loudspeakers, simple amplitude panning yields the optimum panning algorithm. Effectively, HRTF coding information is derived directly from the loudspeaker locations and therefore is matched precisely to the listener. This assumes that intermediate HRTFs are derived by linear interpolation. It should also be noted that as the image source moves away from the radial arc containing the loudspeaker array, changes in HRTFs occur, making the positional transfer functions complex. Nevertheless, even when more exact HRTF data are available, there remains a strong desensitization to the exact form of the HRTFs when the positional transfer functions are calculated because of the relatively close proximity of loudspeakers in an n -array.

Consider next a secondary loudspeaker p that is added to the array, again assuming a linear interpolation model. It is assumed that the secondary loudspeaker is located midway along the same radial arc as the nodal loudspeakers and that its HRTFs $h_a(p)$ and $h_b(p)$ with respect to the listener can be determined by linear interpolation. Thus for the midpoint location,

$$h_a(p) = 0.5 h_a(r) + 0.5 h_a(r + 1) \quad (32a)$$

$$h_b(p) = 0.5 h_b(r) + 0.5 h_b(r + 1). \quad (32b)$$

From these data the compensation gamma functions γ_{p1}

and γ_{p2} follow,

$$\gamma_{p1} = \gamma_{p2} = 0.5 \tag{33}$$

revealing once more a simple form for this special case. The modified positional transfer functions GM_r , GM_{r+1} and GM_p for the respective nodal and secondary loudspeakers r , $r + 1$, and p are calculated as

$$GM_p = \lambda_{p1}G_r + \lambda_{p2}G_{r+1} \tag{34a}$$

$$GM_r = (1 + \gamma_{p1}\lambda_{p1})G_r + \gamma_{p1}\lambda_{p2}G_{r+1} \tag{34b}$$

$$GM_{r+1} = \gamma_{p2}\lambda_{p1}G_r + G_{r+1}(1 + \gamma_{p2}\lambda_{p2}) \tag{34c}$$

Assuming symmetry, let $\lambda_{p1} = \lambda_{p2} = 0.5$ and consider the following examples:

- 1) $G_r = 1$ and $G_{r+1} = 0$, yielding $GM_r = 0.75$, $GM_{r+1} = -0.25$, and $GM_p = 0.5$.
- 2) $G_r = 0.5$ and $G_{r+1} = 0.5$, yielding $GM_r = 0.25$, $GM_{r+1} = 0.25$, and $GM_p = 0.5$.

For an image located coincident with the secondary loudspeaker there is a 6-dB level difference between secondary and nodal loudspeaker input signals. Also signal processing is simple and uses only real coefficients in the matrices.

5.2.2 Compensation and Incorporation of Exact HRTF Data

It is instructive to compare three other options for incorporating secondary loudspeaker and image HRTF coordinates. In each case correction functions λ_{p1} and λ_{p2} are calculated to match the selected HRTFs and the corresponding transfer functions for system input-to-loudspeaker inputs evaluated for loudspeakers r , p , and $r + 1$ located at 20° , 30° , and 40° , respectively, with an image at 30° . The four cases are as follows.

Case 1:
Image Measured HRTF data
Secondary loudspeaker Measured HRTF data
Results See Fig. 17(a)

Case 2:
Image Measured HRTF data
Secondary loudspeaker HRTF derived by linear interpolation
Results See Fig. 17(b)

Case 3:
Image HRTF derived by linear interpolation
Secondary loudspeaker Measured HRTF data
Results See Fig. 17(c)

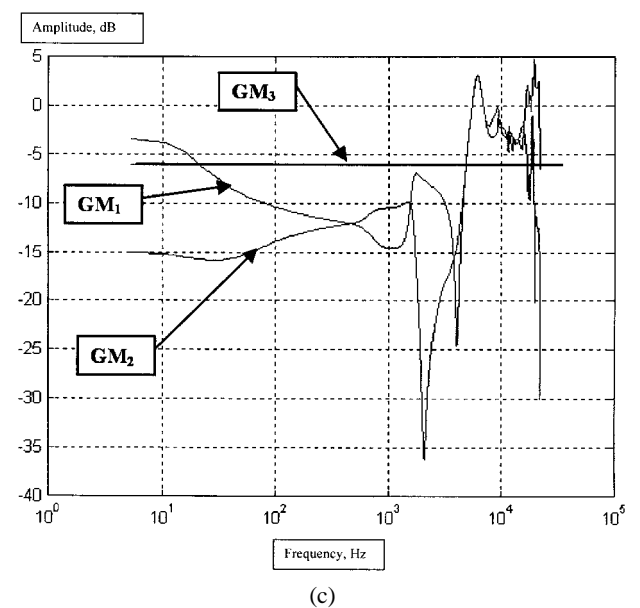
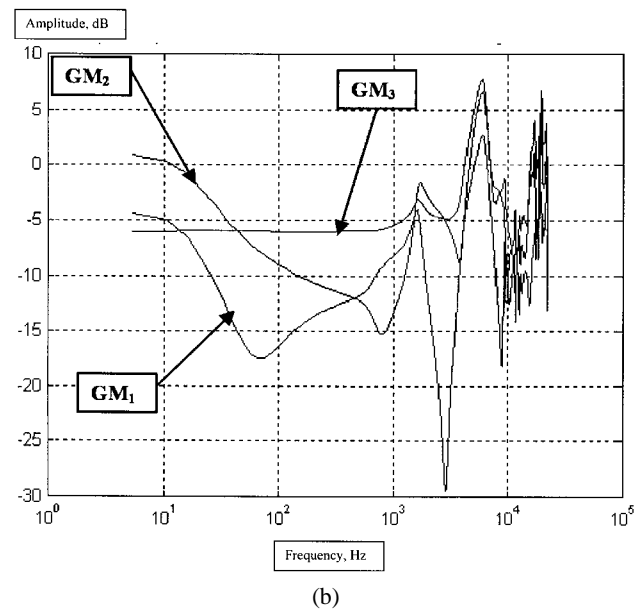
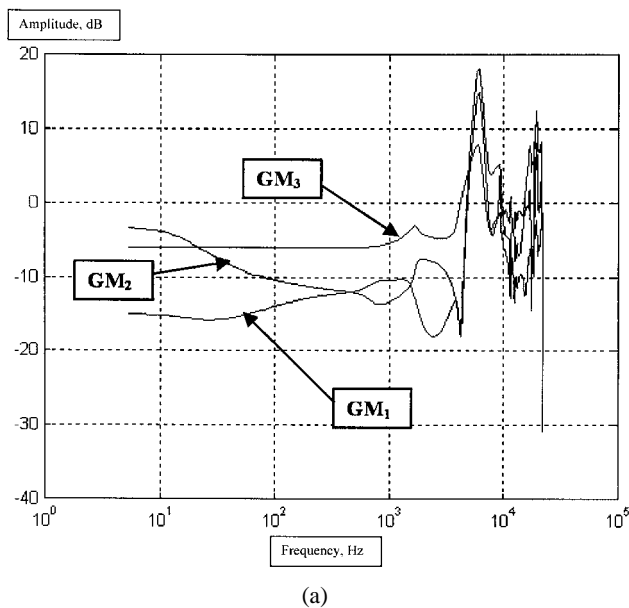


Fig. 17. Transfer functions linking input to three-loudspeaker feeds at 20° , 30° , and 40° , image at 20° . (a) Case 1. (b) Case 2. (c) Case 3. (d) Case 4.

Case 4:

Image	HRTF derived by linear interpolation
Secondary loudspeaker	HRTF derived by linear interpolation
Results	See Fig. 17(d) and discussion in Section 5.2.1

Fig. 17(c) shows that when the HRTF coordinates for the image are derived by linear interpolation, the center channel response is constant with frequency, even though the secondary loudspeaker HRTF coordinates are derived by measurement. Also, when both sets of coordinates are derived by linear interpolation, all three responses are constant, as demonstrated in Section 5.2.1. However, for cases 1 and 2 all responses vary with frequency, although case 2 maintains a greater high-frequency content in the center channel response, which is desirable for an image located coincident with the secondary loudspeaker.

5.2.3 Compensation for Encoder–Decoder Loudspeaker Displacement Error

The encoder assumes a nominal loudspeaker array location when determining the positional transfer functions. If the nodal loudspeakers are placed in the listening space in equivalent positions, then no additional processing is required at the decoder. However, in circumstances where the loudspeaker locations are displaced, positional compensation is required. It is important that positional compensation be independent of source coding and can be applied at the decoder without knowledge of the encoding algorithm. A pairwise positional correction scheme is shown in Fig. 18, where the compensation functions $GC_{11}(r)$, $GC_{12}(r)$, $GC_{11}(r + 1)$, and $GC_{12}(r + 2)$ are used to derive modified loudspeaker feeds. The form of this compensation is not unique, as correction signals can be applied to other loudspeakers in the array via appropriate filters. However, it is suggested that the loudspeaker

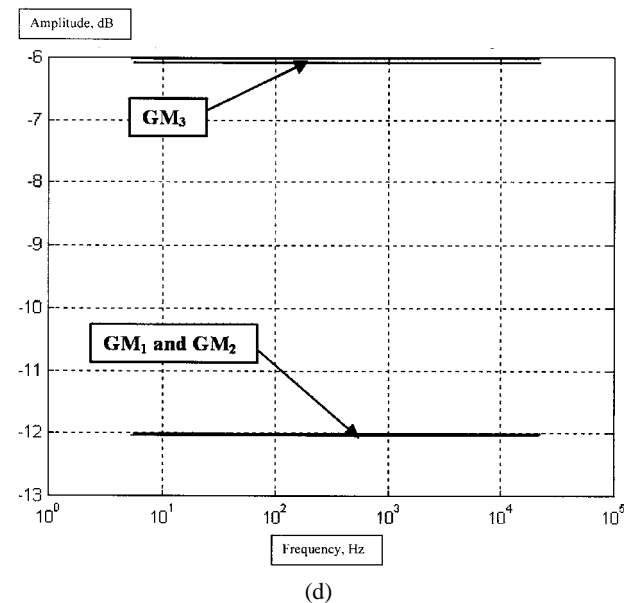


Fig. 17. Continued

selected to process the correction signal be the one closest to the loudspeaker that has been displaced. Hence by way of example, consider the scheme shown in Fig. 18. When the only active primary feed is LS_r , and equating the ear signals for both optimum and displaced loudspeaker locations, the positional compensation filters are as follows:

$$\begin{bmatrix} GC_{11}(r) \\ GC_{12}(r) \end{bmatrix} = \begin{bmatrix} h'_a(r) & h'_a(r + 1) \\ h'_b(r) & h'_b(r + 1) \end{bmatrix}^{-1} \begin{bmatrix} h_a(r) \\ h_b(r) \end{bmatrix}. \quad (35)$$

Similarly, when only LS_{r+1} is active, then

$$\begin{bmatrix} GC_{11}(r + 1) \\ GC_{12}(r + 1) \end{bmatrix} = \begin{bmatrix} h'_a(r + 1) & h'_a(r) \\ h'_b(r + 1) & h'_b(r) \end{bmatrix}^{-1} \begin{bmatrix} h_a(r + 1) \\ h_b(r + 1) \end{bmatrix}. \quad (36)$$

5.3 Standardization of HRTF Grid and Nodal Loudspeaker Locations

The techniques described in the preceding require knowledge of the HRTF coordinates for each nodal loudspeaker. Establishing a standardized encoding grid where each grid point is assigned nominal HRTF coordinates and where nodal loudspeakers are assigned nominal grid locations can satisfy this requirement. All encoder and decoder users then universally know this information. An example grid proposal, as shown in Fig. 19, is based on a 60° subtended angle for nodal loudspeakers with two additional

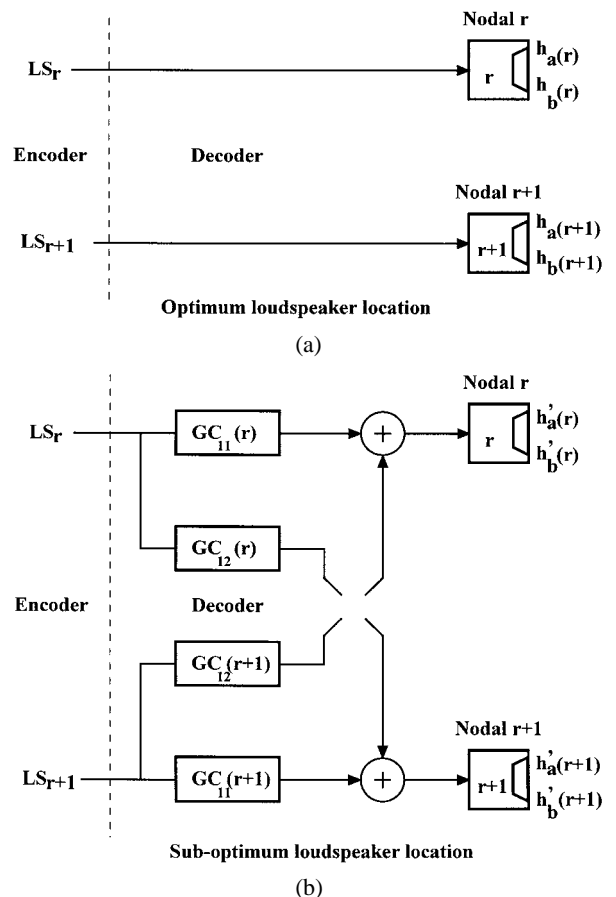


Fig. 18. Compensation process for nodal loudspeaker positional errors. (a) Optimum loudspeaker location. (b) Suboptimum loudspeaker location.

secondary loudspeakers. Three layers of nodes are suggested at radii of 1.5 m, 3 m, and 6 m. It is recognized that HRTFs are not unique, being listener specific, but when a multiloudspeaker array is formed using a set of HRTFs that are shared with image synthesis, errors are reduced.

HRTF coordinates that are noncoincident with the nodal points can be inferred by interpolation. For example, assume that an image X is located at the cylindrical coordinates $\{r, \theta\}$, where the four nearest nodes are $\{r_1, \theta_1\}$, $\{r_1, \theta_2\}$, $\{r_2, \theta_1\}$, and $\{r_2, \theta_2\}$. The interpolated HRTFs $h_a(r, \theta)$ and $h_b(r, \theta)$ are then

$$h_a(r, \theta) = m_r [m_\theta h_a(r_1, \theta_1) + (1 - m_\theta) h_a(r_1, \theta_2)] + (1 - m_r) [m_\theta h_a(r_2, \theta_1) + (1 - m_\theta) h_a(r_2, \theta_2)] \quad (37a)$$

$$h_b(r, \theta) = m_r [m_\theta h_b(r_1, \theta_1) + (1 - m_\theta) h_b(r_1, \theta_2)] + (1 - m_r) [m_\theta h_b(r_2, \theta_1) + (1 - m_\theta) h_b(r_2, \theta_2)] \quad (37b)$$

where m_θ and m_r are the angular and radial linear interpolation parameters defining the image X . For images that lie either within the inner radius or beyond the outer radius, angular interpolation is performed first, followed by an appropriate adjustment to the amplitude and time delays based on the radial distance from the head.

6 PERCEPTUALLY BASED CODING EXPLOITING EMBEDDED CODE IN PRIMARY SIGNALS TO ENHANCE SPATIAL RESOLUTION

The techniques described in Section 5 can be extended to a system with any number of nodal and secondary loudspeakers and thus can be matched to a wide variety of

multichannel system configurations. However, inevitably there is a limit to spatial resolution arising from the use of matrixing only, which imposes crosstalk between nodal and secondary loudspeaker feeds. Some advantage may be gained by using nonlinear decoding with dynamic parameterization, although for high-resolution music reproduction linear decoding should be retained.

Because DVD-A is capable of six channels at 24 bit 96

kHz, some of the lower bits in the LPCM stream can be sacrificed [12] while still retaining an exemplary dynamic range by using standard methods of psychoacoustically motivated noise shaping and equalization [16]. The least significant bits in the LPCM streams together with a randomization function can then be used to encode additional audio channels using perceptual coders such as AC-3,⁹ DTS,¹⁰ or MPEG.¹¹ For example, 4 bit per sample per LPCM channel

⁹Proprietary perceptual coding developed by Dolby Laboratories.
¹⁰Digital Theatre Systems.

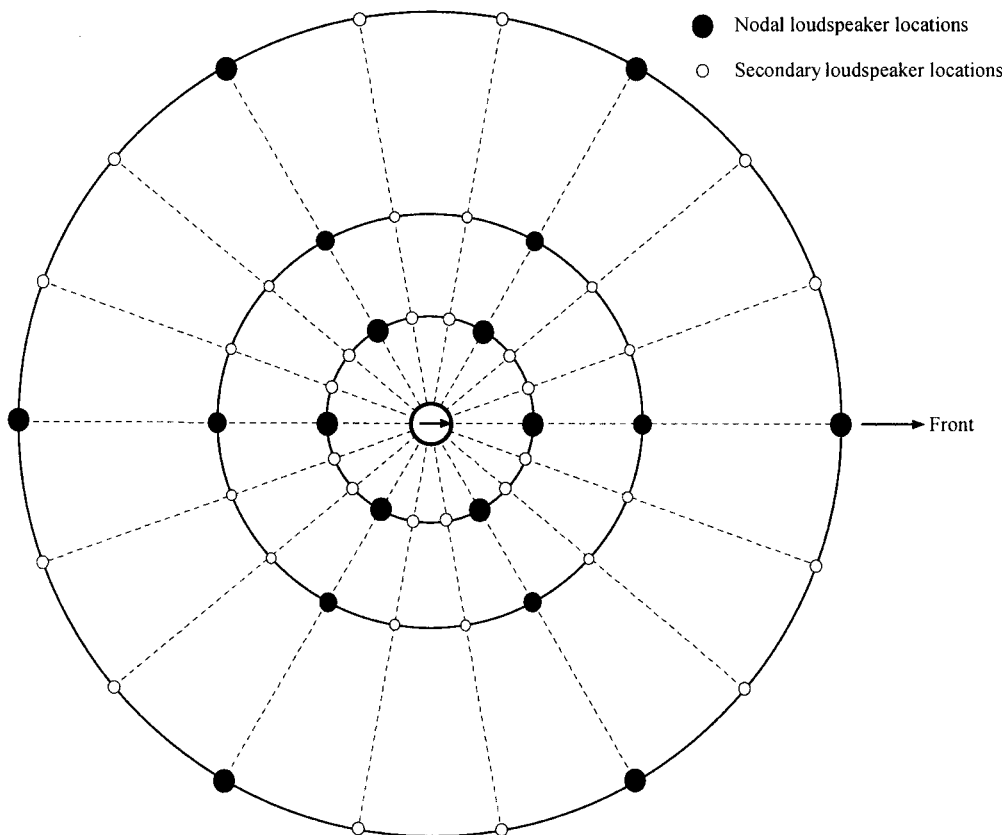


Fig. 19. Proposed 18-segment constellation map to define a standard set of HRTFs.

at 96 kHz yields a serial bit rate of 384 kbit/s.

The proposal retains the primary signals in high-resolution LPCM and uses the matrix methods described in Section 5 to estimate the secondary loudspeaker feeds. Spatially related difference signals are then calculated from the discrete secondary loudspeaker signals available at the encoder and the matrix-derived signals. Also, because of close spatial clustering of the additional channels there is a high degree of interchannel correlation with both the primary and the secondary loudspeaker signals, a factor that bodes well for accurate perceptual coding. Close clustering also implies that perceptual coding errors are not widely dispersed in space, yielding an improved masking performance. Given that DVD-A already supports six LPCM channels, it is suggested that an extra two encoded signals per primary signal is a realistic compromise, yielding a total of 18 channels, as proposed in the standardized HRTF constellation illustrated in Fig. 18. In the grand plan there would be n perceptual coders in operation, one per nodal loudspeaker feed. In such a scheme further gains are possible by integrating dynamic bit allocation across all coders as well as using a perceptual model designed specifically for multichannel stereo encoding. In difficult encoding situations dynamic spatial blending can be used to reduce the difference signals prior to perceptual encryption. Fig. 20 shows the basic encoder architecture where the error signals D_1 and D_2 are indicated. In Fig. 21 a decoder is shown where identical estimates are made of the secondary loudspeaker input signals, but with the addition of the difference signals to yield discrete loudspeaker feeds. Of course, if the embedded perceptually coded difference signals are not used, estimates can still be made for the secondary loudspeakers as shown in Fig. 16. Alternatively, for a basic scheme an array of nodal loudspeakers only can be used.

¹¹Perceptual based audio coding proposed by the Motion Picture Expert Group.

¹²Registered trademark of Company name, New Transducers plc, UK.

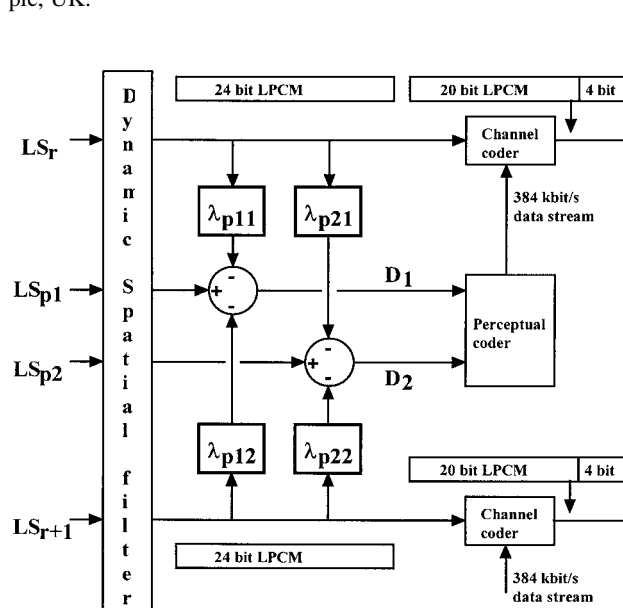


Fig. 20. High-level processor architecture for encoding discrete secondary loudspeaker feeds.

7 CONCLUSIONS

This paper has presented a method for multichannel audio that is fully compatible with DVD (DVD-A and SACD) multichannel formats, which have the capability of six high-resolution signals. The key to this technology is the exploitation of spatial coding using HRTF data to enhance the positional representation of sound sources. As such it forms a link between two-channel transaural techniques and conventional multichannel audio using many loudspeakers. This technique has already been demonstrated in telepresence and teleconferencing applications [9], [10] to be effective in representing spatial audio. However, the methods described here show how a particular loudspeaker array can be configured where issues of positional calibration were discussed for loudspeakers displaced from those locations assumed during coding.

The method is scalable and fully backward compatible. In a simple system there is no additional processing at the decoder where, for example, the outputs of a DVD player are routed directly to an array of loudspeakers. However, if additional loudspeakers are used, as might be envisaged with tiled walls of flat-panel NXT¹² loudspeakers (see, for example, Fig. 22), then formal methods exist, enabling the correct ear signals at the listening position to be maintained. Also for DVD-A, a method was suggested where perceptually coded information is embedded within the LPCM code to enable discrete loudspeaker signals to be derived. It was proposed that an upper limit of 18 channels should be accommodated, although full compatibility with systems down to the basic array is maintained.

An interesting observation for systems using a large number of closely spaced loudspeakers is that image positioning on the arc of the array can use simple linear amplitude panning applied between pairs of adjacent loud-

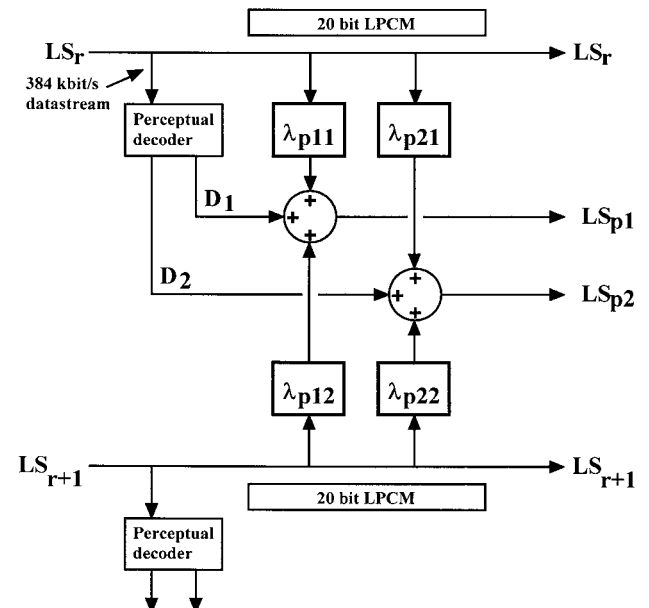


Fig. 21. High-level decoder architecture to derive discrete secondary loudspeaker feeds.

speakers. This approximation assumes that the image HRTF coordinates can be estimated to sufficient accuracy using linear interpolation between adjacent loudspeaker HRTF coordinates. This expedience effectively embeds HRTF data matched exactly to the listener simply because of the physical location of the loudspeakers. However, as the loudspeaker spacing increases, this approximation fails, requiring then the use of more accurate HRTF image coordinates together with transaural PWC, as described. This is particularly important where an image is located away from the arc of the loudspeaker array and where reflections are to be rendered to craft a more accurate virtual acoustic.

This work is also targeted at new communication formats for virtual reality, telepresence, and video conferencing [17], [18], where future research should investigate its application. Such schemes are not constrained by the normal paradigms of multichannel stereophonic reproduction, nor is compatibility necessarily sought. The approach is to form an optimum methodology for constructing phantom images and to consider coding paradigms appropriate for communication. For example, one possible communication format assigns a discrete channel to each phantom image. The channel then conveys the auditory signals together with the spatial coordinates updated at a rate compatible with motion tracking of the sound source. At the receiver, a processor carries a downloaded program with knowledge of the positional data and source acoustics from which the required reflections and reverberation are computed. These data would then be formatted to match the selected loudspeaker array. Such a scheme has great flexibility and can allow many mono sources to contribute to the final soundscape.

In conclusion, the techniques presented describe a

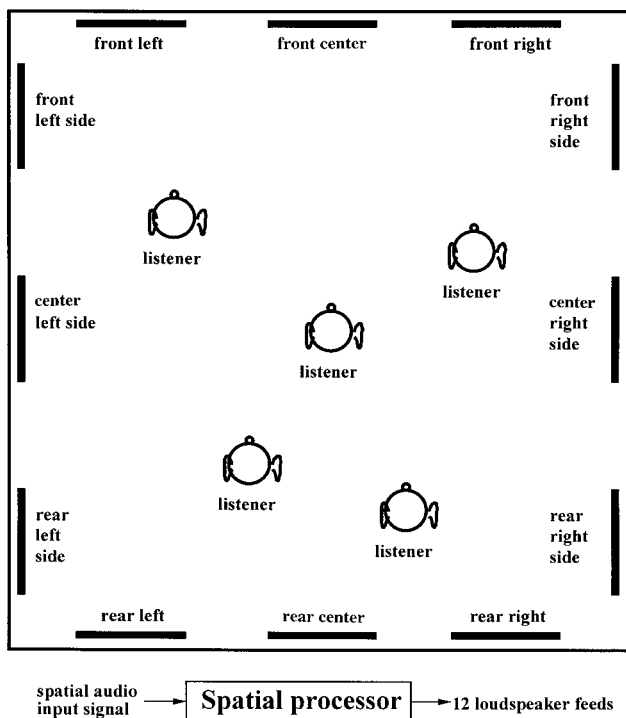


Fig. 22. Multichannel configuration using array of wall-mounted flat-panel diffuse loudspeakers.

means by which spatial resolution and image coding performance can transcend the six-channel limitation of the current DVD formats, yet without requiring additional storage capacity. Also, by basing signal processing on a perceptual model of hearing, it is revealed how sound images can be rendered and, in particular, how interaural amplitude differences and interaural time differences can be accommodated without seeking tradeoffs between time and amplitude clues. Essentially the work has presented a scalable and reverse compatible solution to multichannel audio that is particularly well matched to an LPCM format on DVD-A.

8 REFERENCES

[1] *HFN/RR*, “Digital Frontiers,” vol. 40, pp. 58–59, 106 (1995 Feb.).

[2] M. O. J. Hawksford, “High-Definition Digital Audio in 3-Dimensional Sound Reproduction,” presented at the 103rd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 45, p. 1016 (1997 Nov.), preprint 4560.

[3] A. J. Berkout, D. de Vries, and P. Vogel, “Acoustic Control by Wavefield Synthesis,” *J. Acoust. Soc. Am.*, vol. 93, pp. 2764–2778 (1993).

[4] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia* (AP Professional, 1994).

[5] M. A. Gerzon, “Periphery: With-Height Sound Reproduction,” *J. Audio Eng. Soc.*, vol. 21, pp. 2–10 (1973 Jan./Feb.).

[6] M. A. Gerzon, “Ambisonics in Multichannel Broadcasting and Video,” *J. Audio Eng. Soc.*, vol. 33, pp. 859–871 (1985 Nov.).

[7] M. A. Gerzon, “Hierarchical Transmission Systems for Multispeaker Stereo,” *J. Audio Eng. Soc.*, vol. 40, pp. 692–705 (1992 Sept.).

[8] K. C. K. Foo and M. O. J. Hawksford, “HRTF Sensitivity Analysis for Three-Dimensional Spatial Audio Using the Pairwise Loudspeaker Association Paradigm,” presented at the 103rd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 45, p. 1018 (1997 Nov.), preprint 4572.

[9] K. C. K. Foo, M. O. J. Hawksford, and M. P. Hollier, “Three-Dimensional Sound Localization with Multiple Loudspeakers Using a Pairwise Association Paradigm and Embedded HRTFs,” presented at the 104th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 46, p. 572 (1998 June), preprint 4745.

[10] K. C. K. Foo, M. O. J. Hawksford, and M. P. Hollier, “Pairwise Loudspeaker Paradigms for Multichannel Audio in Home Theatre and Virtual Reality,” presented at the 105th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 46, p. 1035 (1998 Nov.), preprint 4796.

[11] M. Gerzon, “Practical Periphery: The Reproduction of Full-Sphere Sound,” presented at the 65th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 28, p. 364 (1980 May), preprint 1571.

[12] M. A. Gerzon and P. G. Craven, “A High-Rate

Buried Data Channel for Audio CD,” *J. Audio Eng. Soc.*, vol. 43, pp. 3–22 (1995 Jan./Feb.).

[13] J. Blauert, *Spatial Hearing*, rev. ed. (MIT Press, Cambridge, MA, 1997).

[14] J. Bauck and D. H. Cooper, “Generalized Transaural Stereo and Applications,” *J. Audio Eng. Soc.*, vol. 44, pp. 683–705 (1996 Sept.).

[15] M. A. Gerzon, “Optimum Reproduction Matrices for Multispeaker Stereo,” *J. Audio Eng. Soc.*, vol. 40, pp. 571–589 (1992 July/Aug.).

[16] J. R. Stuart and R. J. Wilson, “Dynamic Range Enhancement Using Noise-Shaped Dither Applied to

Signals with and without Preemphasis,” presented at the 96th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 42, p. 400 (1994 May), preprint 3871.

[17] “Telepresence Theme,” *BT Technol. J.*, vol. 15 (1997 Oct.).

[18] D. M. Burraston, M. P. Hollier, and M. O. J. Hawksford, “Limitations of Dynamically Controlling the Listening Position in a 3-D Ambisonic Environment,” presented at the 102nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 45, p. 413 (1997 May), preprint 4460.

THE AUTHOR



Malcolm Hawksford received a B.Sc. degree with First Class Honors in 1968 and a Ph.D. degree in 1972, both from the University of Aston in Birmingham, UK. His Ph.D. research program was sponsored by a BBC Research Scholarship and investigated delta modulation and sigma-delta modulation (SDM, now known as bit-stream coding) for color television and produced a digital time-compression/time-multiplex technique for combining luminance and chrominance signals, a forerunner of the MAC/DMAC video system.

Dr. Hawksford is director of the Centre for Audio Research and Engineering and a professor in the Department of Electronic Systems Engineering at Essex University, where his research and teaching interests include audio engineering, electronic circuit design, and signal processing. His research encompasses both analog and digital systems with a strong emphasis on audio systems including loudspeaker technology. Since 1982, research into digital crossover networks and equalization for loudspeakers has resulted in an advanced digital and active loudspeaker system being designed at Essex

University. A first in 1986 was for a prototype system to be demonstrated at the Canon Research Centre in Tokyo, work sponsored by a research contract from Canon. Much of this work has appeared in the *JAES*, together with a substantial number of contributions at AES conventions.

His research has also encompassed oversampling and noise-shaping techniques applied to analog-to-digital and digital-to-analog conversion with a special emphasis on SDM. Other research has included the linearization of PWM encoders, diffuse loudspeaker technology, and three-dimensional spatial audio and telepresence including multichannel sound reproduction.

Dr. Hawksford is a recipient of the 1997/1998 AES Publications Award for his paper, “Digital Signal Processing Tools for Loudspeaker Evaluation and Discrete-Time Crossover.” He is a chartered engineer as well as a fellow of the AES, IEE, and IOA. He is currently chair of the AES Technical Committee on High-Resolution Audio and is a founder member of the Acoustic Renaissance for Audio (ARA). He is also a technical consultant for NXT, UK and LFD Audio, UK.

Two-to-Five Channel Sound Processing*

R. IRWAN AND RONALD M. AARTS, *AES Fellow*

Philips Research Laboratories, 5656 AA Eindhoven, The Netherlands

While stereo music reproduction was a dramatic advance over mono, recently a transition to multichannel audio has created a more involving experience for listeners. An algorithm to convert stereo to five-channel sound reproduction is presented. An effective sound distribution to the surround channels is achieved by using a cross-correlation technique, and a robust stereo image is obtained using principal component analysis. Informal listening tests comparing this scheme with other methods revealed that the proposed algorithm is preferred for both on and off the reference listening position (sweet spot).

0 INTRODUCTION

Since the introduction of the digital versatile disk (DVD) and the super audio CD (SACD), a revival of multichannel audio has appeared in sound systems for consumer use today. It is, however, desirable to maintain compatibility with the existing two-channel stereo recordings and/or broadcasting. Therefore the conversion of two-channel stereo to the multichannel format has been studied extensively over the decades, and a considerable number of publications exist [1]–[8]. Among these, Gerzon and Barton's is particularly notable, in which many schemes have been proposed (see [6] and references therein).

Although many authors have introduced multichannel sound systems with a large number of channels, we restrict ourselves to a home cinema setup for which it has been shown that five channels is sufficient for creating ambience effects [9]. Hence in this paper we focus on signal format conversion from two-channel stereo to the five-channel (two-to-five) sound processing algorithm.

The desired setup is shown in Fig. 1, in which the channels are labeled L (left), C (center), R (right), S_L (left surround), and S_R (right surround) according to convention. This setting is adopted from the ITU multichannel configuration [10], with three loudspeakers placed in front of the listener, and the other two at the back.

The front channels are used to provide a high degree of directional accuracy over a wide listening area for front-stage sounds, particularly dialogues, and the rear channels produce diffuse surround sounds, providing ambience and

environment affects. An additional loudspeaker (subwoofer) may be used to augment bass reproduction, which is often called 5.1 system, with .1 referring to the low-frequency enhancement (LFE) channel. In this paper, however, we do not use a subwoofer, since the system can easily be extended when necessary without affecting the algorithm.

The algorithm presented in this paper offers two improvements above the existing two-to-five channel sound systems. First a problem associated with channel crosstalk is reduced, and therefore sound localization is better. Listening tests have confirmed that good sound localization without the need to listen at the sweet spot gives more space to the listener to enjoy the program offered rather than restricting the listener to the sweet spot.

Second a better sound distribution to the surround channels is achieved by using a cross-correlation technique. Surround channels are crucial in creating the ambience effects, which is one of the main goals of multichannel audio. At the same time, the energy preservation criterion is an important constraint that has been used to design multichannel matrices [7]. The main reason for this is to maintain backward and forward stereo compatibility. Furthermore, the preservation criterion ensures that all signals present in the two-channel transmitted signals are produced at a correct power level, so that the balance between the different signal sounds in the recording is not disturbed.

This paper is organized as follows. In Section 1 a technique for deriving a robust center channel is outlined. A three-dimensional mapping to derive the surround channels is discussed in Section 2. The rest of the paper will discuss subjective assessments of some listening tests that have been performed in order to com-

*Partly presented at the AES 19th International Conference, Germany, 2001 June 21–24; revised 2001 October 5 and 2002 May 31.

pare the present method with other existing two-to-five channel sound systems. Concluding remarks are presented in Section 4.

1 CENTER LOUDSPEAKER

We consider the three-channel approach first. It is known that the sound quality of stereo sound reproduction can be improved by adding an additional loudspeaker between each adjacent pair of loudspeakers. For example, as proposed by Klipsch [1], an additional center loudspeaker C can be fed with the sum signal $\sqrt{2}(x_L + x_R)$, where x_L and x_R represent signals from left and right, respectively. The $\sqrt{2}$ factor was introduced to preserve the total energy from the three loudspeakers, assuming incoherent additive for left, center, and right sounds recorded by two widely spaced microphones. A major drawback of

this approach is that crosstalk with the left and right channels is inevitable, and therefore it will narrow the stereo image considerably.

We propose an algorithm to derive the center channel without these drawbacks, using principal component analysis (PCA) [11], which produces two vectors indicating the direction of both the dominant signal y and the remaining signal q , as shown in Fig. 2 by dashed lines. Note that these two directions are perpendicular to each other, creating a new coordinate system. These two signals are then used as basis signals in the matrix decoding, a point that is different from other existing two-to-five channel sound systems.

To derive the center channel's gain using the direction of a stereo image, we process the audio signal coming from a CD (sampling frequency $F_s = 44.1$ kHz) on a sample basis. Each sample of a stereo pair at a time index k

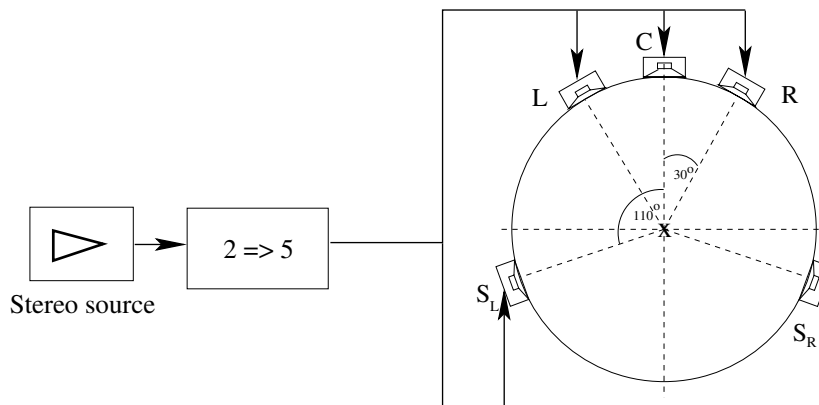


Fig. 1. ITU reference configuration [9]. ×—reference listening position (sweet spot). Left and right channels are placed at angles $\pm 30^\circ$ from C; two surround channels are placed at angles $\pm 110^\circ$ from C.

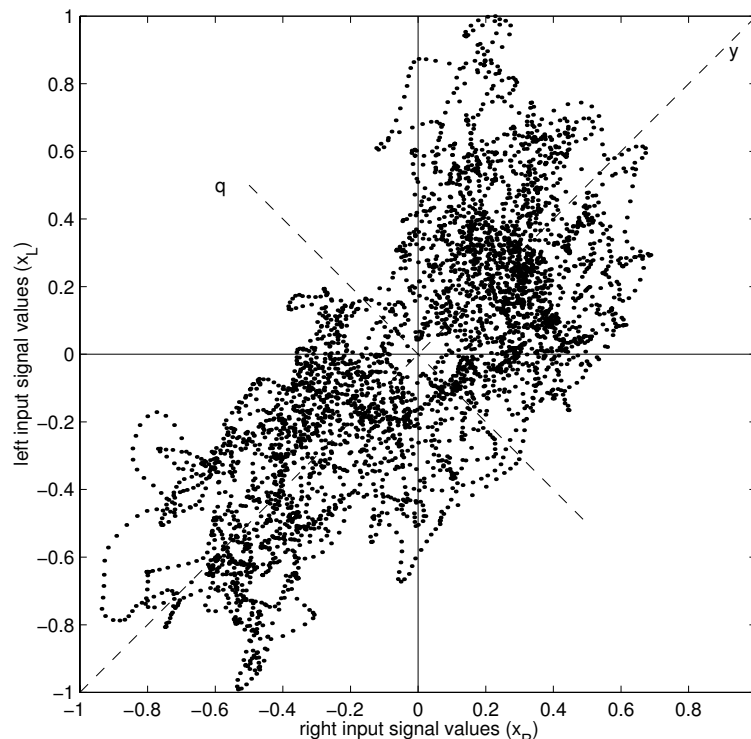


Fig. 2. Lissajous plot of stereo signal recorded from the fragment “The Great Pretender”: by Freddy Mercury. Dashed lines represent new coordinate system based on both dominant signal y and remaining signal q , forming the direction of a stereo image α .

can be expressed as

$$\mathbf{x}(k) = \begin{bmatrix} x_L(k) & x_R(k) \end{bmatrix}^T \quad (1)$$

where k is an integer.

Let us now define $y(k)$ to be a linear combination of the input signals,

$$y(k) = \mathbf{w}^T(k) \mathbf{x}(k) \quad (2)$$

where

$$\mathbf{w}(k) = \begin{bmatrix} w_L(k) & w_R(k) \end{bmatrix}^T \quad (3)$$

is a weight vector corresponding to the left and the right channels, respectively.

In order to find the optimum weighting vectors, we maximize the energy of Eq. (2) with respect to \mathbf{w} , that is,

$$\frac{\partial E \left[y(k)^2 \right]}{\partial \mathbf{w}} = 0 \quad (4)$$

where E denotes the expected value. Using a method presented by Haykin [12], we obtain by means of the steepest descent method

$$\begin{aligned} \mathbf{w}(k) &= \mathbf{w}(k-1) + \frac{1}{2} \mu \frac{\partial E \left[y(k-1)^2 \right]}{\partial \mathbf{w}} \\ &= \mathbf{w}(k-1) + \mu E \left[y(k-1) \mathbf{x}(k-1) \right] \end{aligned} \quad (5)$$

where μ is a step size. Since $E[y(k-1)]$ and $E[\mathbf{x}(k-1)]$ are both scalars, Eq. (5) can be estimated as

$$\mathbf{w}(k) = \mathbf{w}(k-1) + \mu y(k-1) \mathbf{x}(k-1). \quad (6)$$

Normalizing Eq. (6) such that $\|\mathbf{w}(k)\|_2 = 1$, gives the desired sample estimate of \mathbf{w} ,

$$\mathbf{w}(k) = \frac{\mathbf{w}(k-1) + \mu y(k-1) \mathbf{x}(k-1)}{\sqrt{\sum_{L,R} \left[\mathbf{w}(k-1) + \mu y(k-1) \mathbf{x}(k-1) \right]^2}}. \quad (7)$$

Assuming that the step size μ is small, Eq. (7) can be expanded as a power series in μ , yielding

$$\begin{aligned} \mathbf{w}(k) &= \mathbf{w}(k-1) + \mu y(k-1) \\ &\quad \times \left[\mathbf{x}(k-1) - \mathbf{w}(k-1) y(k-1) \right] \end{aligned} \quad (8)$$

which is a least-mean-square (LMS) algorithm with $y(k-1)$ as input. Writing out Eq. (8) for left and right channels, respectively, produces

$$\begin{aligned} w_L(k) &= w_L(k-1) + \mu y(k-1) \\ &\quad \times \left[x_L(k-1) - w_L(k-1) y(k-1) \right] \\ w_R(k) &= w_R(k-1) + \mu y(k-1) \\ &\quad \times \left[x_R(k-1) - w_R(k-1) y(k-1) \right]. \end{aligned} \quad (9)$$

Karhunen [13] has shown that the algorithm is stable if and only if

$$0 < \mu \mathbf{x}^T(k) \mathbf{x}(k) < 2 \quad (10)$$

or the step size must satisfy the following constraint:

$$0 < \mu < \frac{2}{\mathbf{x}^T(k) \mathbf{x}(k)} \quad (11)$$

and therefore it is input signal dependent.

The direction of a stereo image in terms of an angle, in radians, can easily be computed as

$$\alpha(k) = \arctan \left[\frac{w_L(k)}{w_R(k)} \right]. \quad (12)$$

Fig. 3 shows the values of α when it is calculated for a CD stereo music recording. Recalling Fig. 2 with the left channel corresponding to $\alpha = \pi/2$ and the right channel to $\alpha = 0$, we can see that α fluctuates around $\pi/4$, creating a phantom source almost equidistant between the left and right channels.

Fig. 4 shows the same response of the angle α , but now measured from a DVD movie fragment, where abrupt changes from one channel to the other are present. We intentionally take a shorter fragment in order to demonstrate that the algorithm is still able to detect abrupt changes in localizations within a short period of time.

Now we can represent a pair of stereo signals using a vector given by Eq. (3). This is a vector of unit length having the right channel gain in the horizontal axis, and the left channel gain in the vertical axis, as shown in Fig. 5(a). To map this stereo vector onto a three-channel vector, we double the angle α and produce a new mapping, as depicted in Fig. 5(b). We can then find the projections of the vector onto the LR axis and the C axis using sine and cosine rules,

$$\begin{aligned} c_{LR} &= w_R^2 - w_L^2 \\ c_C &= 2w_L w_R. \end{aligned} \quad (13)$$

It should be pointed out that the transformation illustrated in Fig. 5 works only for nonnegative α . This is because for negative α , multiplication by a factor of 2 results in the vector being in a lower quadrant, and therefore no gain can be derived for the center channel. To overcome this problem, extra information should be used, which is described in the next section.

2 SURROUND LOUDSPEAKERS

The surround channels are generally used to create ambience effects for music. For applications in the film industry the surround channels are used for sound effects. A common technique for ambience reconstruction is the use of delayed front channel information for the surround channels. Dolby Pro Logic, for instance, has delayed the surround sounds so as to arrive at the listeners' ears at least 10 ms later than the front sounds [7].

Environmental and ambience effects can be computed by considering left and right channel variations ($x_L - x_R$)

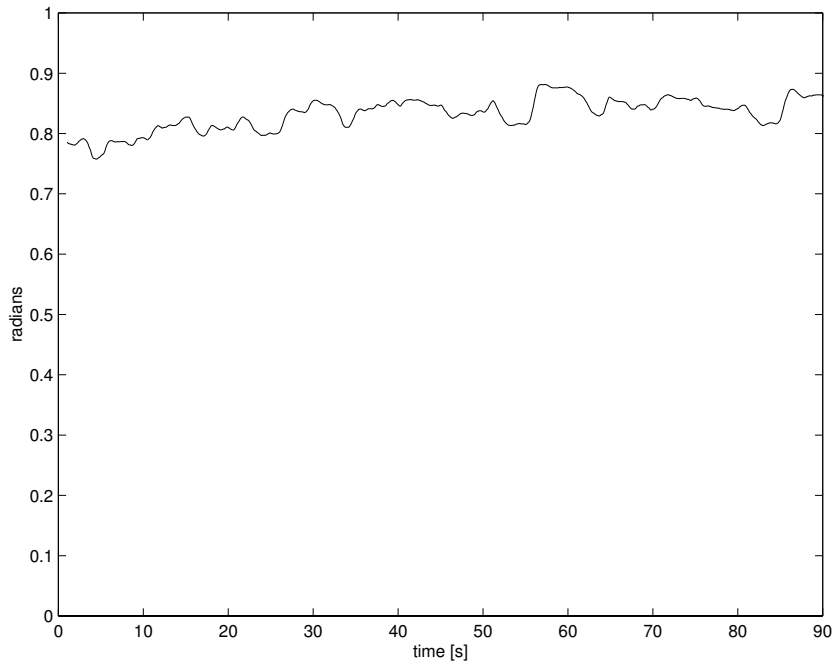


Fig. 3. Typical example of fluctuation of α , computed from a CD stereo music fragment with a stable phantom source.

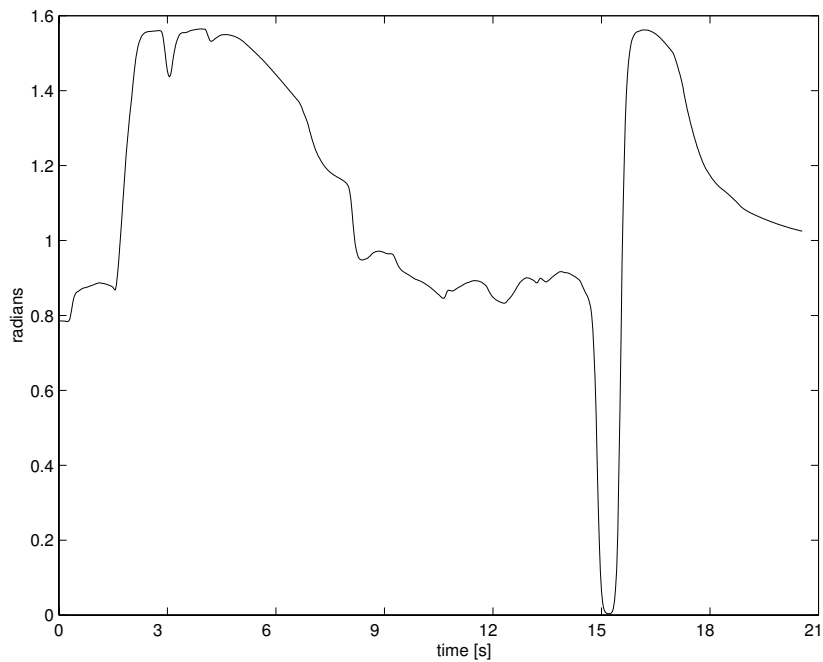


Fig. 4. Fluctuation of the direction α computed from a DVD fragment containing sounds of a car passing with high speed from one channel to the other. Total duration of fragment about 20 seconds.

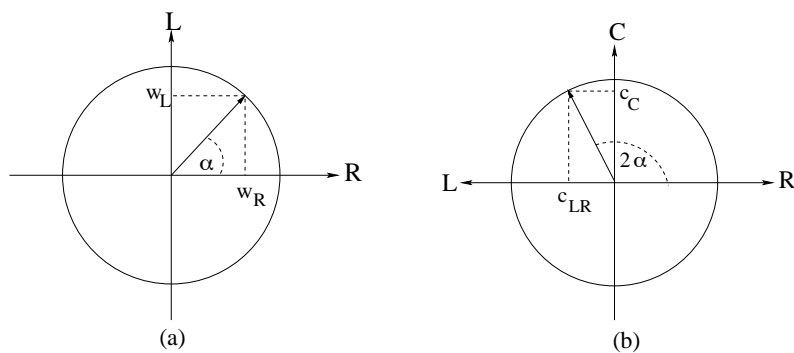


Fig. 5. (a) Direction vector plots of stereo signals. (b) Corresponding three-channel representation by doubling the angle α .

in the original signals. This variation is usually referred to as the antiphase components, the amount of which can be represented by the remaining signal q (see Fig. 2). However, it can be expected that when the amount of the dominant signal equals or almost equals that of the remaining signal, an ambiguity appears since there is no way of determining the direction vector uniquely. In this situation the distribution in Fig. 2 is no longer an ellipse but has a circlelike form ($|y| \approx |q|$), as illustrated in Fig. 6, causing α to be not well defined.

Obviously extra information is necessary when dealing with this sort of ambiguity. In this paper we propose to use a known technique to measure the amount of antiphase components, namely, the correlation coefficient, which is given in any text book on statistics as

$$\rho = \frac{\sum(x_L - \bar{x}_L)(x_R - \bar{x}_R)}{\sqrt{\sum(x_L - \bar{x}_L)^2 \sum(x_R - \bar{x}_R)^2}} \quad (14)$$

where \bar{x}_L and \bar{x}_R are the mean values of x_L and x_R , respectively.

Aarts et al. [14] have shown that Eq. (14) can be computed recursively by using only a few arithmetic operations,

$$\hat{\rho}(k) = \hat{\rho}(k - 1) + \gamma \left\{ 2x_L(k)x_R(k) - [x_L(k)^2 + x_R(k)^2] \hat{\rho}(k - 1) \right\} \quad (15)$$

where γ is the step size determining the time constant, and the caret (^) is used to denote that it is an estimate of the

true ρ . A summary of the mathematical derivations of Eq. (15) can be found in Appendix 3.

To give some ideas how this tracking algorithm works, we present two examples of measurements using Eq. (15), which are shown in Figs. 7 and 8. The measurements are performed within a time frame of 50 seconds, with the step size γ set to 10^{-3} at $F_s = 44.1$ kHz. Fig. 7 is a typical example of a modest stereo for which the correlation varies around 0.70, and it is thus neither too strong (mono sound) nor too weak (diffuse sound). On the other hand, Fig. 8 shows an example of an uncorrelated stereo signal with many antiphase components, for which α is difficult to detect (see Fig. 6).

It is worth mentioning here that there are three other variants of Eq. (15) which are evenly robust. For further information the reader is referred to [14].

Since

$$-1 \leq \rho \leq 1 \quad (16)$$

it is possible that the antiphase components exceed the dominant signal ($|y| < |q|$). In this case we treat the input signals as uncorrelated, and therefore

$$\rho_0 = \begin{cases} \rho, & 0 \leq \rho \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

It can be shown (see Appendix 1) that a relationship exists between this cross-correlation method and PCA described in the previous section.

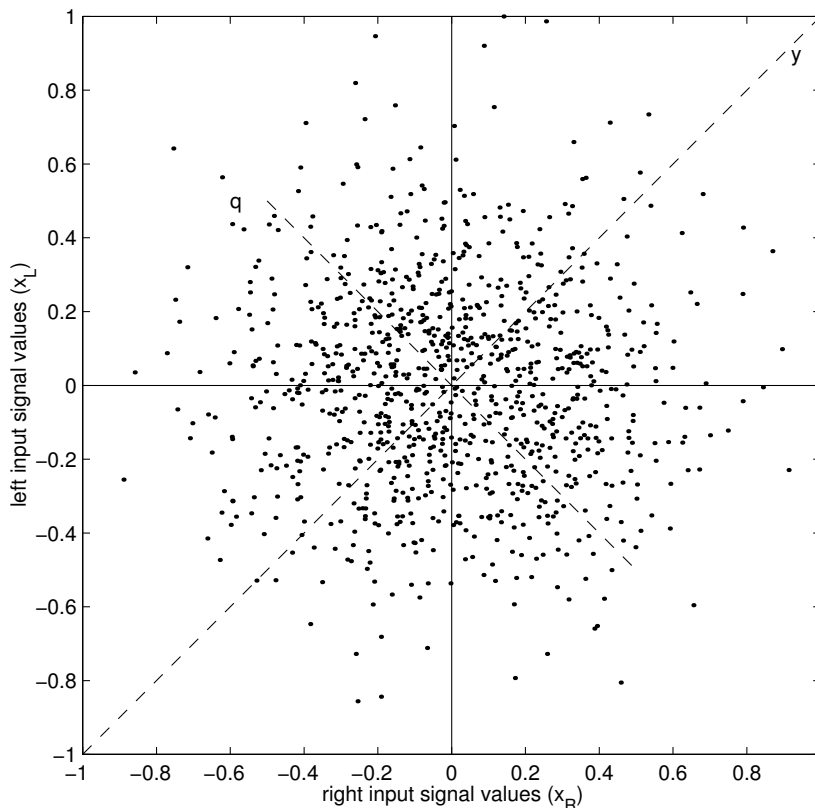


Fig. 6. Lissajous plot of the first 23-second stereo signal recorded from the fragment “Holiday” by Madonna, where the amount of a dominant signal is almost equal to that of a remaining signal, forming a circle-like distribution.

2.1 Three-Dimensional Mapping

To avoid ambiguity when the amount of the dominant signal approaches that of the remaining signals, the use of both the direction of the stereo image and the correlation coefficient is necessary. The latter is included in the mapping (see Fig. 5) by, for example, placing the surround channels in the vertical plane, as shown in Fig. 9.

The angle β can be defined to represent the actual surround information by means of the adaptive correlation coefficient, for example, by using the expression

$$\beta(k) = \arcsin[1 - \rho_0(k)] \tag{18}$$

and hence,

$$0 \leq \beta(k) \leq \frac{\pi}{2}. \tag{19}$$

Thus as the amount of the remaining signal increases (input signals become weakly correlated), the angle β also increases, which reduces the total distribution to the front channels. On the other hand, when the input signals are strongly correlated (quasi mono), β approaches zero, producing a larger contribution to the front channels. This principle satisfies the energy preservation criterion, which is discussed in more detail in Section 2.2.

Since the direction vector on the horizontal plane is now lifted by an angle β , recalculation of the projections is nec-

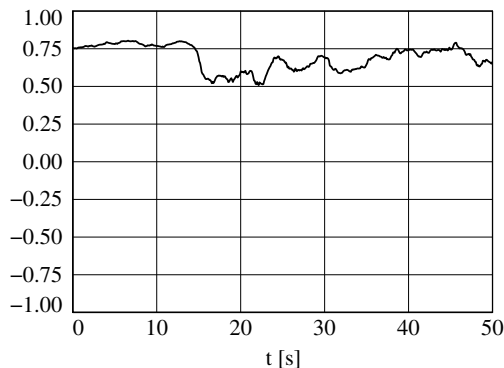


Fig. 7. Tracked cross-correlation coefficients obtained from the fragment “The Great Pretender” by Freddy Mercury. See Fig. 2 for the stereo image, which is a typical example of a modest stereo recording.

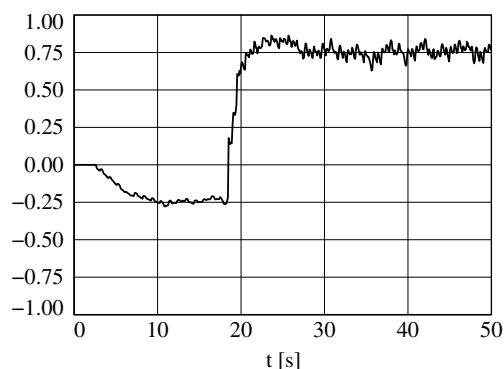


Fig. 8. Tracked cross-correlation coefficients obtained from the fragment “Holiday” by Madonna. See Fig. 6 for the corresponding Lissajous plot.

essary. Using straightforward trigonometry and keeping in mind that the vector is of unit length, we obtain

$$\begin{aligned} c'_{LR} &= c_{LR} \cos \beta \\ c'_C &= c_C \cos \beta \\ c_S &= \sin \beta. \end{aligned} \tag{20}$$

2.2 Matrixing

The system described so far reproduces four channel signals as L, C, R, and S from two input signals. Therefore we have a 4×2 reproduction matrix.

We now discuss the objective requirement on the energy preservation as emphasized in Section 2.1. A matrix preserves energy if and only if its columns are of unit length, and the columns are pairwise orthogonal. Since the product of any two orthogonal matrices is also orthogonal, back and forward compatibility between stereo and multi-channel can also be achieved.

Following this energy criterion, we design the matrix as follows:

$$\begin{bmatrix} u_L(k) \\ u_R(k) \\ u_C(k) \\ u_S(k) \end{bmatrix} = \begin{bmatrix} c_L(k) & gw_L(k) \\ c_R(k) & gw_R(k) \\ c_C(k) & 0 \\ 0 & c_S(k) \end{bmatrix} \begin{bmatrix} y(k) \\ q(k) \end{bmatrix}. \tag{21}$$

The components of the left-hand side of Eq. (21) denote the signals for the left, right, and center loudspeakers, and u_S denotes the mono surround signal. The basis signals are obtained by rotating the coordinate system of x_L and x_R ,

$$\begin{aligned} y(k) &= w_L(k)x_L(k) + w_R(k)x_R(k) \\ q(k) &= w_R(k)x_L(k) - w_L(k)x_R(k) \end{aligned} \tag{22}$$

and

$$\begin{aligned} c_L &= \begin{cases} -c_{LR}, & c_{LR} < 0 \\ 0, & \text{otherwise} \end{cases} \\ c_R &= \begin{cases} c_{LR}, & c_{LR} \geq 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{23}$$

and g is a gain to control the energy preservation.

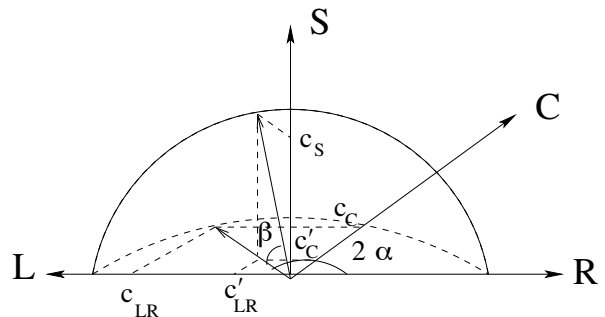


Fig. 9. Three-dimensional mapping showing front (horizontal plane) and surround channels (vertical plane). Parameter β determines the level of surround information with respect to the front channel sounds.

Since c_L and c_R can only produce one value depending on the condition in Eq. (23), the length of the first column of the matrix given in Eq. (21) is equal to $c_{LR}^2 + c_C^2$, which is unity. The second column of Eq. (21) contains mainly a projection of the vector onto the horizontal plane (see Fig. 9). The length of this column is equal to $g(w_L^2 + w_R^2) + c_S^2 = 1$. The two columns are thus of unit length and pairwise orthogonal if $g = \cos^2 \beta$, and therefore the matrix preserves the total energy.

Finally, the Lauridsen [15] decorrelator is used to obtain stereo surround because of its simplicity. This decorrelator can be viewed as two FIR comb filters (h_L and h_R) with two taps each for surround left and surround right. The impulse responses of these filters are illustrated in Fig. 10. A time delay of $\delta \approx 10$ ms (440 samples) is used between the taps, which is determined experimentally.

The choice of the time delay δ is a subtle compromise between the amount of widening and the sound diffuseness. The greater δ is, the more diffuse the sounds will be, and at some point it will lead to confusion.

Note that there are other decorrelator filters available, such as complementary comb filters, in which the “teeth” are distributed equally on a logarithmic frequency scale. Informal listening tests, however, revealed that the Lauridsen decorrelator is better appreciated when it is applied to the surround channels. Furthermore, its efficiency in the implementation also plays an important role in our application.

2.3 Analysis of Each Discrete Channel

The behavior of the proposed method can be analyzed in each channel and gives useful information for validating an implementation. In addition, such an analysis can be used to demonstrate the channel separation of our method. Such analyses are listed in Table 1.

First we feed the system with a sine wave in the left channel only and set the right channel to zero,

$$\begin{aligned} x_L(k) &= \sin(\omega k) \\ x_R(k) &= 0. \end{aligned} \tag{24}$$

In this situation we have $w_L = 1$ and $w_R = 0$. Therefore,

$$\begin{aligned} y(k) &= \sin(\omega k) \\ q(k) &= 0. \end{aligned} \tag{25}$$

Substituting Eqs. (24) and (25) into Eq. (21), and keeping in mind that $c_R = c_C = 0$, we obtain

$$\begin{aligned} u_L(k) &= c_L(k) \sin(\omega k) \\ u_R(k) &= 0 \\ u_C(k) &= 0 \\ u_S(k) &= 0. \end{aligned} \tag{26}$$

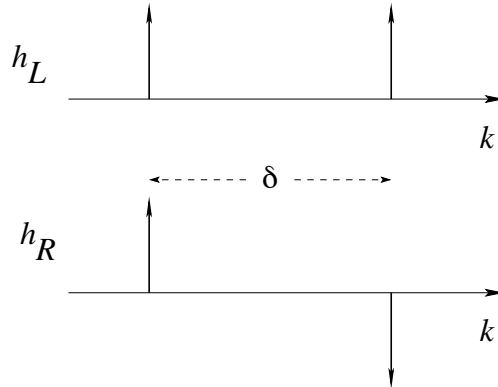


Fig. 10. Impulse response of left and right Lauridsen decorrelation filters. Time delay $\delta \approx 10$ ms (440 samples) is experimentally chosen to produce the most pleasant stereo sounds for the application concerned.

Table 1. Summary of extreme cases described in text.*

Input	u_L	u_R	u_C	u_S	w_L	w_R	ρ_0	β
$x_L = f, x_R = 0$	f	0	0	0	1	0	0	$\pi/2$
$x_L = 0, x_R = f$	0	f	0	0	0	1	0	$\pi/2$
$x_L = x_R = f$	0	0	κf	0	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{2}$	1	0
$x_L = f, x_R = -f$	κf	$-\kappa f$	0	κf	$\frac{1}{2}\sqrt{2}$	$-\frac{1}{2}\sqrt{2}$	0	$\pi/2$
$x_L, x_R \neq 0$, uncorrelated	0	0	0	κf	—	—	0	$\pi/2$

* A time signal f is used to represent any input signals fed into left, right, or a combination of left and right channels. κ represents a scalar due to mapping (Fig. 9). Note that when uncorrelated signals are fed into the system, w_L and w_R become undefined, meaning they can assume any value.

This is to be expected as any signal fed into one particular channel should be kept the same in the output.

Second, feeding the input signal into the right channel, some other combinations can be analyzed similarly. The outputs of these combinations are summarized in Table 1.

From the table it can be seen that fully correlated input signals will be reproduced in the center channel while all other channels are zero. This explains the strong sound localization that is achieved during the listening test, which is discussed in the next section.

Furthermore, when we feed the left and right channels with antiphase signals, no sound will be reproduced in the center channel, left and right outputs are in antiphase, and some sounds are going to the surround channels. This extreme case demonstrates how ambience effects are created when many antiphase signals are present in the original signals.

3 LISTENING TEST

In order to investigate the appreciation of the discussed conversion method, the system was tested together with a few other systems.

3.1 Method

The method of paired comparisons [16] was used to gather personal preference data. During each trial, subjects heard a music excerpt encoded by a certain method M_i , then the same excerpt encoded by M_j . The subject could repeat the pairs as often as desired, and finally had to indicate whether M_i was preferred above M_j or vice versa. There were four systems under test, so six pairs per repertoire per listening position were offered to the subjects. If M_i was preferred above M_j a 1 was placed in a preference matrix X at element x_{ij} , or otherwise at x_{ji} . This matrix was scaled with Thurstone's decision model [17] (see Appendix 2 for more details).

3.2 Technical Equipment and Repertoire

The subjects were either at the position advised by the ITU [10], referred to as the sweet spot, or 1 m aside of that spot, referred to as "off the sweet spot." The loudspeakers used were Philips DSS940 (digital) loudspeakers. The listening room was a rather dry listening room, which enabled a critical judgment for localization and crosstalk between the channels.

We compared four different systems: the system proposed in this paper (system 1), its variant, which puts more low-frequency to the surrounds (system 2), and two other commercially available systems (system 3 and 4, respectively).

Four different music excerpts were chosen. Three music fragments (Fish: "The Company," The Corrs: "What Can I Do," and Melanie C: "Never Be the Same Again") and a sound track from the movie picture "The Titanic" (Track #23 from the DVD) were used. In total a subject had to give 6×2 positions \times 4 fragments = 48 assessments.

The number of different tracks was somewhat limited. To derive more general conclusions more tracks would be necessary, such as used in [19]. However, our primary aim

was to focus on the theory, while more elaborate listening tests can still be done in the future.

3.3 Subjects

There were 17 subjects. Most were experienced listeners and all had no reported hearing loss.

3.4 Results

For each type of repertoire and each subject the scaled results were plotted. An example is given in Fig. 11. A high scale value means high appreciation. The value itself is of no importance. It is the value with respect to the others. The sum of the scale values equals zero.

The number on top of Figs. 11–13 denotes the coefficient of consistence ξ [16]. A value of $\xi = 1$ means fully consistent. In such a case there is never a violation of the triangular inequalities ($X_i > X_j > X_k > X_i$). $\xi = 0$ means fully inconsistent. This coefficient is important for various reasons. First it reveals how consistently a subject judges the stimuli. In this case subject SP appeared to be a very consistent judge. Second, if the subjects have different preferences with respect to each other, or for different repertoires, then summing their preference matrices will lower the ξ value.

For both positions, on the sweet spot and off the sweet spot, the results for all subjects and repertoires are scaled and plotted in Figs. 12 and 13, respectively.

It is clear that system 1—the system discussed in this paper—performs very well, both on the sweet spot and off the sweet spot. In Figs. 12 and 13 we see values of $\xi = 0.4$ and 0.6, respectively, revealing that not all the subjects act as one single fully consistent subject. However, it appeared that the individual subjects act rather consistently. Furthermore, the figures show that system 1 is rather robust in the face of a displacement from the

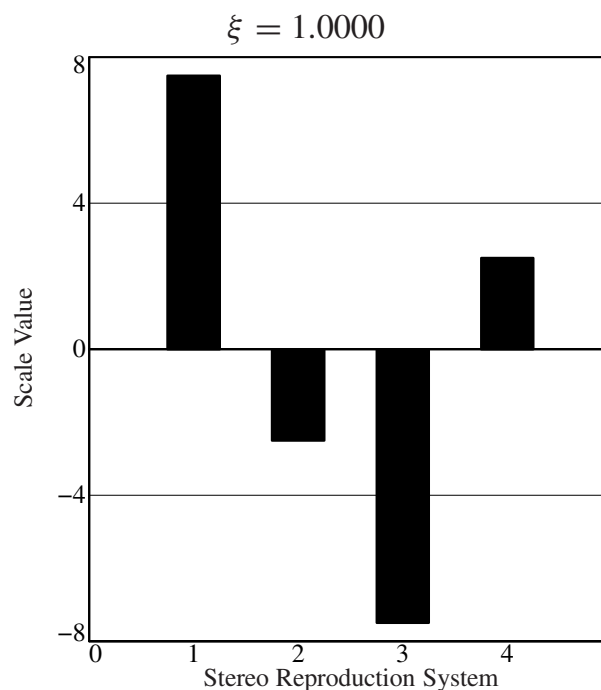


Fig. 11. Results of subject SP, four music excerpts, off the sweet spot. Scale values on vertical axis are arbitrary.

sweet spot, which is a desirable property.

Another analysis—using the Bradley–Terry model [16]—of the results of the listening test was performed [18]. This model was fitted by the maximum-likelihood method, and the goodness-of-fit was tested. It revealed that the four processing methods differed significantly from one another.

4 CONCLUSION

A new method to convert two-channel stereo to multi-channel sound has been presented. A three-dimensional representation has been used to produce each channel's gain, which is time varying. PCA proved to be a powerful tool to detect the direction of a stereo image, which is then used to derive the center channel's gain. Furthermore, a robust tracking algorithm for computing the cross correlation between left and right channels has been used to improve the sound quality of the surround channels.

A listening test comparing four different systems has been carried out using both music recordings and DVD movie tracks, and the results have been analyzed using the Thurstone scaling technique as well as the Bradley–Terry model. The preliminary listening test has shown that the proposed method is very good, both on and off the sweet spot. Moreover, it has been shown that the four processing methods differed significantly from one another.

5 ACKNOWLEDGMENT

The authors thank D. Roovers for his work in the initial phase of the project, and the statisticians T. J. J. Denteneer and J. Engel for interpreting the results of the listening tests.

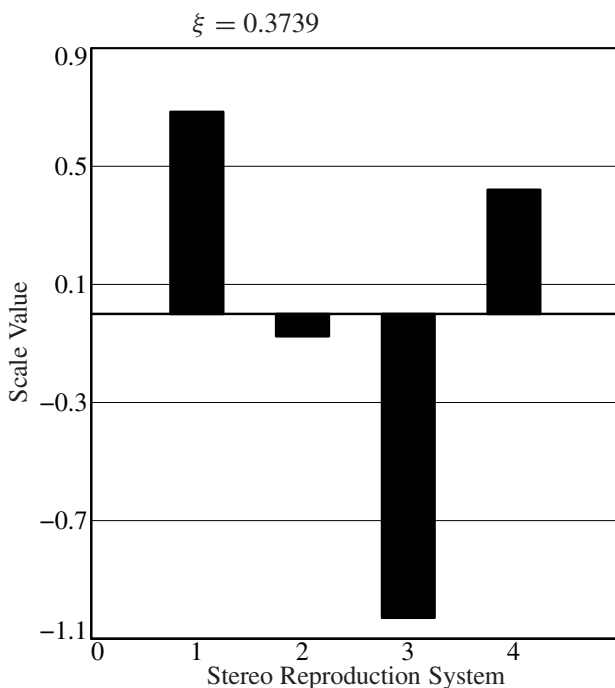


Fig. 12. Total scaling results of all 17 subjects, four music excerpts, on the sweet spot. Scale values on vertical axis are arbitrary.

6 REFERENCES

- [1] P. W. Klipsch, "Stereophonic Sound with Two Tracks, Three Channels by Means of a Phantom Circuit (2PH3)," *J. Audio Eng. Soc.*, vol. 6, p. 118 (1958).
- [2] P. Scheiber, "Four Channels and Compatibility," *J. Audio Eng. Soc.*, vol. 19, pp. 267–279 (1971 Apr.).
- [3] J. Eargle, "4–2–4 Matrix Systems: Standards, Practice, and Interchangeability," *J. Audio Eng. Soc.*, vol. 20, pp. 809–815 (1972 Dec.).
- [4] M. E. G. Willcocks, "Transformations of the Energy Sphere," *J. Audio Eng. Soc.*, vol. 31, pp. 29–36 (1983 Jan./Feb.).
- [5] S. Julstrum, "A High-Performance Surround Sound Process for Home Video," *J. Audio Eng. Soc.*, vol. 35, pp. 536–549 (1987 July/Aug.).
- [6] M. A. Gerzon and G. J. Barton, "Ambisonic Decoders for HDTV," presented at the 92nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc.*, vol. 40, p. 438 (1992 May), preprint 3345.
- [7] R. Dressler, "Pro Logic Surround Decoder, Principles of Operation," Dolby Laboratories, <http://www.dolby.com> (1997).
- [8] R. Irwan and R. M. Aarts, "A Method to Convert Stereo to Multichannel Sound," presented at the AES 19th International Conference, Germany (2001).
- [9] G. Theile, "HDTV Sound System: How Many Channels?" in *Proc. 10th AES Conf. on Images and Audio* (1991), pp. 147–162.
- [10] ITU-R BS.1116, "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems," International Telecommunications Union, Geneva, Switzerland (1994).
- [11] T. W. Lee, *Independent Component Analysis*:

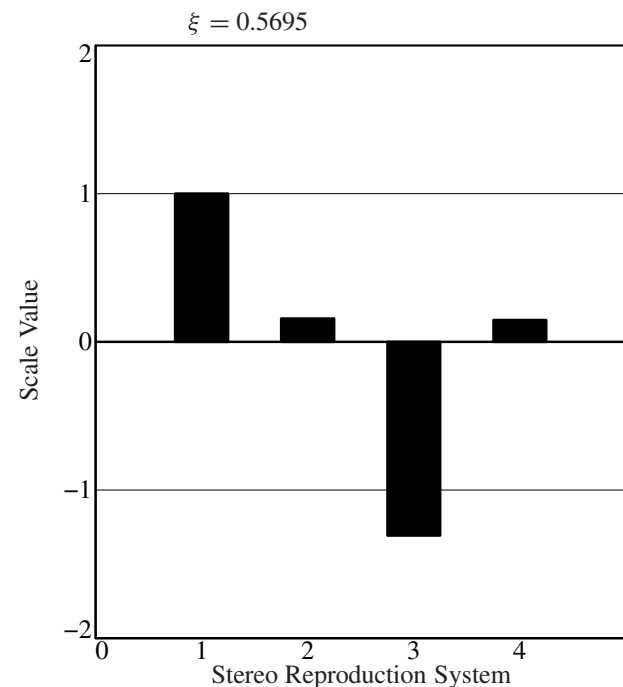


Fig. 13. Total scaling results of all 17 subjects, four music excerpts, off the sweet spot. Scale values on vertical axis are arbitrary.

Theory and Applications (Kluwer, Boston, MA, 1998).

[12] S. Haykin, *Neural Networks*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1999).

[13] J. Karhunen, "Stability of Oja's PCA Subspace Rule," *Neural Comput.*, vol. 6, pp. 739–747 (1994).

[14] R. M. Aarts, R. Irwan, and A. J. E. M. Janssen, "Efficient Tracking of the Cross-Correlation Coefficient," *IEEE Trans. Speech Audio Process.*, vol. 10, pp. 391–402 (2002 Sept.).

[15] H. Lauridsen, "Experiments Concerning Different Kinds of Room-Acoustics Recording" (in Danish), *Ingenioren*, vol. 47 (1954).

[16] H. A. David, *The Method of Paired Comparisons*, 2nd ed. (Griffin, London, 1988).

[17] W. S. Torgerson, *Theory and Methods of Scaling* (Wiley, New York, 1958).

[18] J. Engel, "Paired Comparisons on Audio: Modelling the System Differences," Intern. Rep. E1624-20, CQM (Centre for Quantitative Methods), Eindhoven, The Netherlands (2001 July).

[19] G. A. Soulodre, T. Grusec, M. Lavoie, and L. Thibault, "Subjective Evaluation of State-of-the-Art Two-Channel Audio Codecs," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 46, pp. 164–177 (1998 Mar.).

[20] R. M. Aarts, "A New Method for the Design of Crossover Filters," *J. Audio Eng. Soc.*, vol. 37, pp. 445–454 (1989 June).

[21] L. L. Thurstone, "A Law of Comparative Judgment," *Psychol. Rev.*, vol. 34, pp. 273–286 (1927).

[22] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

APPENDIX 1 RELATION BETWEEN CORRELATION COEFFICIENT ρ AND PCA

As was presented in Section 2, the cross-correlation technique is very useful in determining the surround channel distribution. In this appendix we discuss the relationship between the cross-correlation technique and the principal component analysis (PCA), also known as the Karhunen–Loève transformation. In general, PCA maximizes the rate of decrease in variance for each of its components. The solution lies in the eigenstructure of the covariance matrix C .

We may decompose C with the singular-value decomposition (SVD) [12] as

$$C = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} \quad (27)$$

where \mathbf{V} is an orthogonal (unitary) matrix of eigenvectors \mathbf{v}_j and $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues

$$\mathbf{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_j, \dots, \lambda_m] \quad (28)$$

arranged in decreasing order,

$$\lambda_1 > \lambda_2 > \dots > \lambda_j > \dots > \lambda_m \quad (29)$$

so that $\lambda_1 = \lambda_{\max}$.

It appears that the PCA and the SVD of C are basically

the same, just viewing the problem in different ways. From the theory of SVD it follows that the eigenvectors of the covariance matrix C define the unit vectors \mathbf{v}_j representing the principal directions along which the variance is maximized for each component. The associated eigenvalues define the total variance of the m elements,

$$\sum_{j=1}^m \sigma_j^2 = \sum_{j=1}^m \lambda_j. \quad (30)$$

In the present case C is constructed in the following way. Consider a segment of left and right audio samples $\mathbf{x}_L = [x_L(1) \dots x_L(B)]$ and $\mathbf{x}_R = [x_R(1) \dots x_R(B)]$; hence $m = 2$. With their correlation coefficients ρ the covariance matrix C can be written as

$$C = \begin{pmatrix} \sigma_{x_L}^2 & \rho\sigma_{x_L}\sigma_{x_R} \\ \rho\sigma_{x_L}\sigma_{x_R} & \sigma_{x_R}^2 \end{pmatrix}. \quad (31)$$

Now we see the relation between the correlation coefficient ρ on the left-hand side of Eq. (27) and the principal components on the right-hand side of Eq. (27). The latter we will develop more explicitly in the following.

Since C is positive definite, the eigenvalues of C are both real and positive and can be calculated as

$$\lambda_{1,2} = \frac{1}{2} (\sigma_{x_L}^2 + \sigma_{x_R}^2 \pm s) \quad (32)$$

$$s = \sqrt{(\sigma_{x_L}^2 - \sigma_{x_R}^2)^2 + (2\rho\sigma_{x_L}\sigma_{x_R})^2}. \quad (33)$$

The eigenvectors of C corresponding to the eigenvalues $\lambda_{1,2}$ are

$$\mathbf{v}_{1,2} = \gamma \left(\frac{\sigma_{x_L}^2 - \sigma_{x_R}^2 \pm s}{2\rho\sigma_{x_L}\sigma_{x_R}}, 1 \right) \quad (34)$$

where γ is such that $|\mathbf{v}_{1,2}| = 1$.

If we consider \mathbf{V} as a rotation matrix over the angle α , as used in Eq. (12), then we can derive

$$\rho = \frac{(\sigma_{x_L}^2 - \sigma_{x_R}^2) \tan(2\alpha)}{2\sigma_{x_L}\sigma_{x_R}}. \quad (35)$$

As special cases we consider $\rho = 0$. Then the left and right channels are uncorrelated and the eigenvectors are just $(1, 0)$, $(0, 1)$, which are coincident with the original left and right axes. As corresponding eigenvalues we have $\lambda_1 = \sigma_{x_L}^2$ and $\lambda_2 = \sigma_{x_R}^2$, which are the powers of the left and right channels, respectively. A similar case occurs if $\sigma_{x_L}^2 = \sigma_{x_R}^2 = \sigma^2$. Then $\lambda_{1,2} = \sigma^2$ and $\alpha = \pi/4$ and/or $\rho = 0$.

Since PCA can be seen as finding a decomposition of the covariance matrix C , the transformation into dominant and remaining signals using PCA described in Section 1 can also be carried out by computing \mathbf{v}_1 and \mathbf{v}_2 , respectively. It has, however, a limited practical usefulness since it requires more computation effort as opposed to an efficient way of using the LMS algorithm given in Eq. (6).

**APPENDIX 2
SCALING**

A2.1 Introduction

A problem encountered in many disciplines is how to measure and interpret the relationships between objects. A second problem is the lack, in general, of a mathematical relationship between the perceived response and the actual physical measure. With regard to this paper, how does the appreciation of our 2-to-5-channel system differ from others? How do we measure and what scale do we need? In the following we discuss some scales and techniques and give two examples.

A2.2 Scaling

The purpose of scaling is to quantify the qualitative relationships between objects by scaling data. Scaling procedures attempt to do this by using rules that assign numbers to qualities of things or events. There are two types of scaling, univariate scaling, which is explained hereafter, and multidimensional scaling (MDS), which is an extension of univariate scaling (see, for example, [20]). Univariate scaling is usually based on the law of comparative judgment [17], [21]. It is a set of equations relating the proportion of times any stimulus i is judged greater or is more highly appreciated relative to a given attribute (in our case the appreciation) than any other stimulus j . The set of equations is derived from the postulates presented in [17]. In brief, these postulates are as follows.

- 1) Each stimulus when presented to an observer gives rise to a discriminial process which has some value on the psychological continuum of interest (in our case the appreciation).
- 2) Because of momentary fluctuations in the organism, a given stimulus does not always excite the same discriminial process. This can be considered as noise in the process. It is postulated that the values of the discriminial process are such that the frequency distribution is normal on the psychological continuum.
- 3) The mean and the standard deviation of the distribution associated with a stimulus are taken as its scale value and discriminial dispersion, respectively.

Consider the theoretical distributions S_j and S_k of the discriminial process for any two stimuli j and k , respectively, as shown in Fig. 14(a). Let \bar{S}_j and \bar{S}_k correspond to the scale values of the two stimuli and σ_j and σ_k to their discriminial dispersion caused by noise.

Now we assume that the standard deviations of the distributions are all equal and constant (as in Fig. 14), and that the correlation between the pairs of discriminial processes is constant. This is called “condition C,” in Torgerson [17]. Since the distribution of the difference of the normal distributions is normal, we get

$$\bar{S}_k - \bar{S}_j = cx_{jk} \tag{36}$$

where c is a constant and x_{jk} is the transformed [see Eq. (39)] proportion of the number of times stimulus k is more highly appreciated than stimulus j . Eq. (36) is also known as Thurstone’s case V. The distribution of the discriminial

differences is plotted in Fig. 14(b). Eq. (36) is a set of $n(n - 1)$ equations with $n + 1$ unknowns, n scale values, and c . This can be solved with the least-square method. Setting $c = 1$ and the origin of the scale to the mean of the estimated scale values, that is,

$$1/n \sum_{j=1}^n s_j = 0 \tag{37}$$

we get

$$s_k = 1/n \sum_{j=1}^n x_{jk} . \tag{38}$$

Thus the least-square solution of the scale values can be obtained simply by averaging the columns of matrix X . However, the elements x_{jk} of X are not directly available. With paired comparisons we measure the proportion p_{kj} that stimulus k was judged greater than stimulus j . This proportion can be considered a probability that stimulus k was judged greater than stimulus j . This probability is equal to the shaded area in Fig. 14(b), or

$$x_{jk} = \text{erf}\left(p_{jk}\right) \tag{39}$$

where erf is the error function [22, §7, 26.2], which can easily be approximated (see, for example, [22, §26.2.23]). A problem may arise if $p_{jk} \approx \pm 1$ since $|x_{jk}|$ can be very large. In this case one can then replace x_{jk} by a large value.

It may be noted that this type of transformation is also known as Gaussian transform, where instead of the symbol x , z is used, known as the z score. Instead of using Eq. (39), other models are used, such as the Bradley–Terry model (see [16]). All forms of the law of comparative

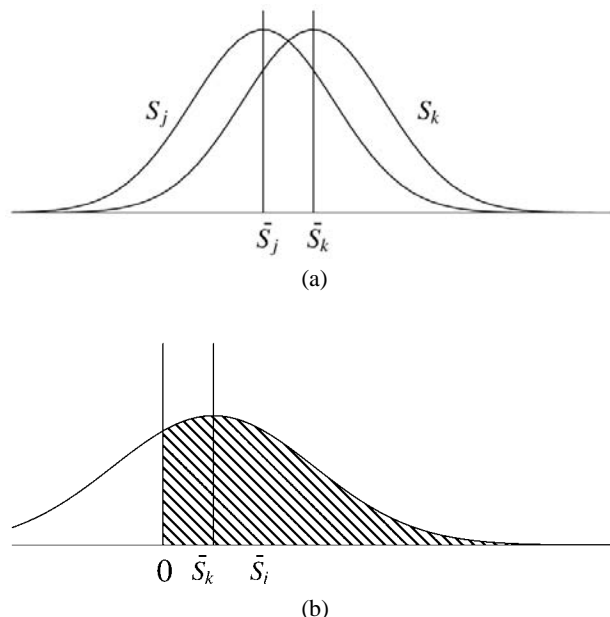


Fig. 14. (a) Probability distributions S_j and S_k of stimuli j and k on psychological continuum, with mean values \bar{S}_j and \bar{S}_k . (b) Probability distributions of difference of random variables. Shaded portion gives the proportion of times stimulus k was judged greater than stimulus j . $\bar{S}_k - \bar{S}_j$ is proportional to the difference in scale value for both stimuli.

judgment assume that each stimulus has been compared with the other stimuli a large number of times. The direct method of obtaining the values of p_{jk} is known as the method of paired comparisons (see, for example, [16]). As an example, the measured probabilities p_{jk} for a subject are listed in Table 2. The upper triangular is calculated as $p_{jk} = 1 - p_{kj}$. Using Eq. (39) the x_{jk} values are obtained. Using Eq. (38) the final scale values are determined and plotted in Fig. 11.

APPENDIX 3 EFFICIENT COMPUTATION OF CROSS-CORRELATION COEFFICIENT

In this appendix we summarize the mathematical derivations of the tracking cross-correlation coefficients as was fully reported in [14]. For the generalization we use x and y for two signals being correlated instead of x_L and x_R , which represent specifically stereo audio signals.

We show that ρ satisfies to a good approximation (when η is small) the recursion in Eq. (15) with γ given by

$$\gamma = \frac{c e^\eta}{2x_{\text{rms}}y_{\text{rms}}} \quad (40)$$

where $c = 1 - e^{-\eta}$, and the subscripts rms refer to the root mean-square values of x and y .

Using an exponential window we can redefine the correlation of x and y at time instant k as

$$\rho(k) = \frac{S_{xy}(k)}{[S_{xx}(k)S_{yy}(k)]^{1/2}} \quad (41)$$

for k an integer and where

$$\begin{aligned} S_{xy}(k) &= \sum_{l=0}^{\infty} c e^{-\eta l} x_{k-l} y_{k-l} \\ &= e^{-\eta} S_{xy}(k-1) + c x_k y_k \end{aligned} \quad (42)$$

and S_{xx} and S_{yy} are defined similarly. Hence,

$$\begin{aligned} \rho(k) &= \frac{S_{xy}(k-1) + c e^\eta x_k y_k}{\left\{ [S_{xx}(k-1) + c e^\eta x_k^2] [S_{yy}(k-1) + c e^\eta y_k^2] \right\}^{1/2}} \\ & \quad (43) \end{aligned}$$

Since we consider small values of η , we have that $c = 1 - e^{-\eta}$ is small as well. Expanding the right-hand side of Eq. (43) in powers of c and retaining only the constant and the linear term, we get, after some calculations,

$$\begin{aligned} \rho(k) &= \rho(k-1) + \frac{c e^\eta}{2[S_{xx}(k-1)S_{yy}(k-1)]^{1/2}} \\ & \quad \times \left\{ 2x_k y_k - \left[\frac{S_{yy}(k-1)}{S_{xx}(k-1)} \right]^{1/2} x_k^2 + \left[\frac{S_{xx}(k-1)}{S_{yy}(k-1)} \right]^{1/2} y_k^2 \right\} \rho(k-1) + O(c^2). \end{aligned} \quad (44)$$

Then, deleting the $O(c^2)$ term, we obtain the recursion in Eq. (15), with γ given by Eq. (40), when we identify

$$\begin{aligned} x_{\text{rms}}^2 &= S_{xx}(k) \\ y_{\text{rms}}^2 &= S_{yy}(k) \end{aligned} \quad (45)$$

for a sufficiently large k , and assuming that $x_{\text{rms}}^2 = y_{\text{rms}}^2$.

One may ask how to handle signals x and y that have nonzero, and actually time-varying, mean values. In those cases we still define $\rho(k)$ as in Eq. (42), however, with S_{xy} replaced by

$$S_{xy}(k) = \sum_{l=0}^{\infty} c e^{-\eta l} [x_{k-l} - \bar{x}(k)][y_{k-l} - \bar{y}(k)] \quad (46)$$

where

$$\bar{x}(k) = \sum_{l=0}^{\infty} c e^{-\eta l} x_{k-l} \quad (47)$$

$$\bar{y}(k) = \sum_{l=0}^{\infty} c e^{-\eta l} y_{k-l}$$

and S_{xx} and S_{yy} changed accordingly. It can then be shown that

$$\begin{aligned} \bar{x}(k) &= e^{-\eta} \bar{x}(k-1) + c x_k \\ \bar{y}(k) &= e^{-\eta} \bar{y}(k-1) + c y_k \end{aligned} \quad (48)$$

and

$$S_{xy}(k) = e^{-\eta} S_{xy}(k-1) + c p_k q_k \quad (49)$$

where $p_k = x_k - \bar{x}(k-1)$, $q_k = y_k - \bar{y}(k-1)$, while sim-

Table 2. Proportion p_{jk} of times that stimulus k was judged higher than stimulus j , obtained via paired comparisons.*

	—	—	—	—
↓ j	0.0	0.0	—	—
	0.5	1.0	1.0	—
	→ k			

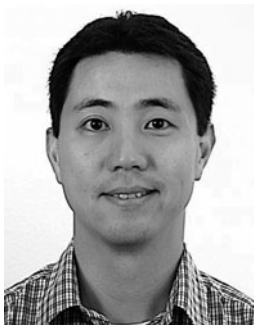
*For subject SP, averaged over the four fragments off the sweep spot.

ilar recursions hold for S_{xx} and S_{yy} . This then yields

$$\begin{aligned} \rho(k) &= \frac{S_{xy}(k-1) + cp_k q_k}{\left\{ \left[S_{xx}(k-1) + cp_k^2 \right] \left[S_{yy}(k-1) + cq_k^2 \right] \right\}^{1/2}} \\ &= \rho(k-1) + \frac{c}{2 \left[S_{xx}(k-1) S_{yy}(k-1) \right]^{1/2}} \\ &\quad \times \left\{ 2p_k q_k - \left\{ \left[\frac{S_{yy}(k-1)}{S_{xx}(k-1)} \right]^{1/2} p_k^2 + \left[\frac{S_{xx}(k-1)}{S_{yy}(k-1)} \right]^{1/2} q_k^2 \right\} \rho(k-1) \right\} + O(c^2). \end{aligned} \quad (50)$$

From this point onward, comparing with Eq. (44), one can proceed to apply many, if not all, of the developments presented in this appendix to this more general situation.

THE AUTHORS



R. Irwan

Roy Irwan received an M.Sc. degree in electrical engineering from the Delft University of Technology, Delft, The Netherlands in 1992, and a Ph.D. degree in electrical engineering from the University of Canterbury, Christchurch, New Zealand in 1999.

From 1993 to 1995 he was employed as a system engineer at NKF B.V. After obtaining his Ph.D. degree in 1999, he joined the Digital Signal Processing group at Philips Research Laboratories, Eindhoven, The Netherlands. Since 2002 he has been working with the State University Groningen, faculty of Medical Sciences.

Dr. Irwan has published a number of refereed papers in international journals and has more than 14 patent applications. His research interests include (medical) image and signal processing.



Ronald Aarts was born in 1956, in Amsterdam, The Netherlands. He received a degree in electrical engineering in 1977, and a Ph.D. degree from the Delft University



R. M. Aarts

of Technology in 1994.

In 1977 he joined the Optics group of Philips Research Laboratories, Eindhoven, The Netherlands, where he was involved in research into servos and signal processing for use in both Video Long Play players and Compact Disc players. In 1984 he joined the Acoustics group of the Philips Research Laboratories and was engaged in the development of CAD tools and signal processing for loudspeaker systems. In 1994 he became a member of the Digital Signal Processing group of the Philips Research Laboratories. There he has been engaged in the improvement of sound reproduction by exploiting DSP and psychoacoustical phenomena.

Dr. Aarts has published over one hundred technical papers and reports and is the holder of more than a dozen U.S. patents in his fields. He was a member of the organizing committee and chair for various conventions. He is a senior member of the IEEE, a fellow of the AES, and a member of the Dutch Acoustical Society and the Acoustical Society of America. He is past chair of the Dutch section of the AES.

AES STANDARDS COMMITTEE NEWS

Information regarding Standards Committee activities including meetings, structure, procedures, reports, and membership may be obtained via <http://www.aes.org/standards/>. For its published documents and reports, including this column, the AESSC is guided by International Electrotechnical Commission (IEC) style as described in the ISO-IEC Directives, Part 3. IEC style differs in some respects from the style of the AES as used elsewhere in this Journal. For current project schedules, see the project-status document on the Web site. AESSC document stages referenced are proposed task-group draft (PTD), proposed working-group draft (PWD), proposed call for comment (PCFC), and call for comment (CFC).

In Memoriam

Shortly before this *Journal* went to press, we received some sad news.

Daniel Queen, secretary of the AES Standards Committee 1980-2001, died on October 17 in Providence, Rhode Island, after a long illness. He had been living in Athens, Greece, for the past year, returning to the United States only two days before he died.

We are deeply saddened to hear this news. Dan's contribution to AES standards over many years has been immense and, personally, I don't know what I would have done without his help over the past year.

A fuller appreciation of Dan's life will appear in the next issue of this *Journal*.

MARK YONGE

The New AES Standards Internet Facilities, Mentioned in the September Issue, Are Now in Service

E-MAIL REFLECTORS

A total of 67 new e-mail reflectors—one for each of the currently active AES standards groups—have been in service since mid-September. Messages sent to a single group e-mail address are distributed to all members of that group to provide a fast and fluent medium for discussion of standards business. Additionally, the service operates more quickly and securely than before. Unsolicited e-mail and inappropriate content are filtered appropriately. All traffic is virus-checked.

WEB SITE

The new Web site, <http://www.aes.org/standards/>, was made public on 2002-10-02 and has two functions.

A public area provides general information, as previously, including information on standards projects and groups as well as the facility to download copies of published standards. The style and presentation of the site has been substantially updated to make access faster and more direct.

The AES Standards Web site will continue to be active as a responsive publication medium for news and public notices such as Calls for Comment. A Participate button provides general access to an on-line membership request form in addition to our established methods for membership request.

In addition to these public pages, there is a new Web area which provides services for AES Standards working groups.

A Log-In button allows members to access all of their registered groups, and their working documents, within a new secure Web area. Members will already have received their initial username and password. There is an automated option for members to request a password reminder and to change their username or password.

The new working area also allows members to check the status of their working groups and update their contact details directly.

Importantly, the new site allows us to exchange working documents more easily than before, replacing the FTP sites we have used previously.

Each group area has a link to a page showing all the current working documents for that group. The page is arranged so that subdirectories—for Task Groups, for example—appear at the top of the page. Documents for the main group are then shown. The list may be sorted by name, size or date; if you simply need to see the latest documents, you can simply sort by date to bring the newest to the top of the list, then click on the

documents you need to download them.

Please enjoy exploring this new facility. We now need to shut down the old e-mail reflectors and document FTP sites, which have served us so well over so many years.

MARK YONGE
AES STANDARDS SECRETARY

Report of the SC-03-02 Working Group on Transfer Technologies of the SC-03 Subcommittee on the Preservation and Restoration of Audio Recording meeting, held in conjunction with the AES 112th Convention in Munich, Germany, 2002-05-12

In the absence of the Chair and Vice Chairs, the meeting was convened by D. Schueller.

The agenda and report from the previous meeting were accepted as written.

Open projects

AES-X47 Minimum Set of Calibration Tones for Archival Transfer

D. Wickstrom published a working group draft in January with certain questions, none have been commented on by the group.

Current goal

A PWD is anticipated by 2002-09.

AES-X64 Test Methods and Materials for Archival Mechanical Media

The intent of the project is to re-press a metal master originating in the EMI factory, which complies to the IEC-98 standard for coarse groove recordings with bands of frequencies starting with a 1 kHz reference level signal and then in decreasing steps (18, 16, 14, 12, 10, 8, 6, 5, 4, 3, 2, 1 kHz; 700, 400, 200, 110, 50, 30 Hz).

It is further our intent to cut and press another test disk. The specifications for this second disc are:

—A sinusoidal sweep from 20 Hz to 20 kHz with constant velocity above 250 Hz. This would start with a 1 kHz tone and take 50 seconds for the sweep so allowing the use of the standard Bruel & Kjaer test equipment.

—A band of 1 kHz at a reference level: 8 cm/sec 20 mm light band width as used on Ortofon Test Record OR1001.

—A band of 1 kHz at a reference level: 16cm/sec post war standard (DIN 45533 1953).

—A repeat of the track 1 sweep. This would be cut at a smaller diameter so that the effect of playback tracing losses for a particular stylus/arm combination could be measured.

—The same programme would be pressed on both sides. The set of discs will be accompanied by a manual describing their use.

The working group will produce a PWD describing these test materials with the intent to publish a call for comment.

AES-X65 Rosetta Tone for Transfer of Historical Mechanical Media

The current goal is review of status. This has been on the agenda since 1997, and yet no progress has been made since G. Brock-Nannestad posted his paper in July 1998. Unless proposals on the contents of such a Rosetta Tone are put forward for discussion, the meeting felt that the project should be withdrawn.

AES-X90 Analog Transfer of Audio Program Material

There has been no communication on this project. Unless a dedicated working chair for Task Group SC-03-02-A can be nominated, the meeting recommends that the project should be withdrawn.

AES-X106 Styli Shape and Size for Transfer of Records

Brock-Nannestad and F. Lechleitner have said they will provide input for the PTD.

AES-X107 Compilation of Technical Archives for Mechanical Media

There has been no input on this project, but relevant liaison bodies have been contacted describing the project and requesting cooperation in putting together a reference document.

New projects

No new projects were received or proposed.

New business

There was no new business.

The next meeting will be held in conjunction with the AES 113th Convention in Los Angeles, CA, US.

Report of the SC-03-04 Working Group on Storage and Handling of Media of the SC-03 Subcommittee on the Preservation and Restoration of Audio Recording meeting, held in conjunction with the AES 112th Convention in Munich, Germany, 2001-11-30

The meeting was convened by T. Sheldon.

The agenda were approved with the addition of a note to discuss the future of the working group under the topic of AES-X80.

The report of the previous meeting was approved with two changes. 1) D. Schueller asked that his name be spelled correctly, with the umlaut. 2) In the notes for AES35-R, the second sentence should be changed to read, "D. Schueller indicated that he would like to see included in AES38-R a plan to investigate the light sensitivity issue in the next revision."

Open projects

AES22-R Review of AES22-1997 AES recommended practice for audio preservation and restoration—Storage

of polyester-based magnetic tape

Schueller will lead the review of this document.

AES28-R Review of AES28-1997 AES standard for audio preservation and restoration—Method for estimating life expectancy of compact discs (CD-ROM), based on effects of temperature and relative humidity

No one at the meeting felt able to conduct the review of this document. Sheldon agreed to refer it to the Joint Technical Commission (JTC) where the document was authored for consideration at the next meeting of SC-03-04.

AES35-R Review of AES35-2000 AES standard for audio preservation and restoration—Method for estimating life expectancy of magneto-optical (M-O) disks, based on effects of temperature and relative humidity

No action was taken.

AES38-R Review of AES38-2000 AES standard for audio preservation and restoration—Life expectancy of information stored in recordable compact disc systems—Method for estimating, based on effects of temperature and relative humidity

Schueller said that it was of vital importance to include a section on light sensitivity of these media as soon as possible. It was noted that J. M. Fontaine and D. Kunej were knowledgeable and could draft a new section. G. Cyrener agreed to investigate possible changes to include the issue of light sensitivity.

AES-X51 Procedures for the Storage of Optical Discs, Including Read Only, Write-once, and Re-writable

Sheldon noted that the draft document is a standard in the imaging area, but he has not brought it to the working group for consideration because of questions about the AES-X80 liaison relationship within the AES Standards Committee.

AES-X54 Magnetic Tape Care and Handling

Sheldon explained the status of this document within the JTC where it is being authored by AES members and others as a part of AES-X80. It is being submitted by JTC within ISO TC42 to form a second Draft International Standard, ISO 18933 DIS. Sheldon will commence the discussion of the text of this second DIS ballot document by July 2002. Further discussion will be conducted by e-mail.

AES-X55 Projection of the Life Expectancy of Magnetic Tape

The investigation to determine a suite of tests to determine the life expectancy of various formulations of magnetic tape continues. It was noted that the International Association of Sound & Audiovisual Archives (IASA), with UNESCO support, has initiated a project to encourage tape manufacturers to take an interest in preserving the 200 million hours of information stored on magnetic tape. One goal is to find ways to measure the life expectancy of magnetic tape. This project should be continued as these initiatives promise to develop new methods to determine life expectancies. However, it was noted that the tape manufacturers currently

are not thriving, so actions need to happen quickly if they are to participate. Also creating a sense of urgency, the signs of tape instability in the world's archives are growing. This subject also is being discussed at the JTC; the minutes of the JTC in Montreal on the subject of magnetic tape life expectancy testing were read. The search continues for a breakthrough that will open possibilities for success.

AES-X80 Liaison with ANSI/PIMA IT9-5

The current goal is a review of the liaison relationship. The liaison arrangement is now with International Imaging Industry Association (I3A) which was the creation of the merger of Photographic Image Manufacturers Association (PIMA) and two other imaging associations. The title of the liaison should be changed to "AES-X80 Liaison with I3A IT9-5."

Sheldon asked whether the AES-X80 liaison arrangement should be continued with I3A. In reality, virtually all research and preparation of standards documents published by AES comes from the JTC. The JTC has many AES members on its membership list, and many AES SC-03-04 members participate in the development of standards as a part of the JTC. AES SC-03-04 has not undertaken in recent years any work of its own apart from the work being conducted by the Joint Technical Commission. The question also should be asked, "Should the AES SC-03-04 working group continue?"

New projects

No project requests were received or introduced.

Sheldon noted, however, that at the AES-X80 Joint Technical Commission meeting April 19-20, 2002 the topic of the care and handling of transportable magnetic discs was raised and discussed. The JTC received a report from J. Lindner surveying the current state of these media. The JTC also is beginning consideration of the care and handling of optical discs. E. Zwanafelt agreed to prepare an outline for consideration at the next meeting.

New business

There was no new business.

The next meeting is scheduled to be held in conjunction with the AES 113th Convention in Los Angeles, CA, US.

Report of the SC-05-02 Working Group on Single-Programme Connections of the SC-05 Subcommittee on Interconnections meeting, held in conjunction with the AES 112th Convention in Munich, Germany, 2002-05-09

J. Brown convened the meeting in the absence of Working Group Chair R. Rayburn and Vice Chair W. Bachmann.

The agenda and the report from the previous meeting were approved as written.

Current development projects

AES-X11 Fiber-Optic Audio Connections: Connectors ➔

THE PROCEEDINGS OF THE AES 22ND INTERNATIONAL CONFERENCE

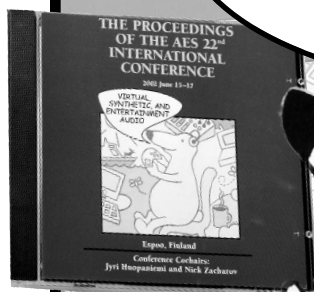
2002 June 15-17

Espoo, Finland

These 45 papers are devoted to virtual and augmented reality, sound synthesis, 3-D audio technologies, audio coding techniques, physical modeling, subjective and objective evaluation, and computational auditory scene analysis.

You can purchase the book and CD-ROM online at www.aes.org. For more information email publication_orders@aes or telephone +1 212 661 8528.

VIRTUAL,
SYNTHETIC, AND
ENTERTAINMENT
AUDIO



Also available on CD-ROM

and Cables Being Used and Considered for Audio

J. Woodgate reported on the meeting of Task Group SC-05-02-F held on 2002-05-09.

A proposed draft, AES32-TU, has been published as a Trial-Use Publication. The current goal is to monitor the trial use. Members present considered that a revision or amendment of the document is now necessary.

The draft specifies a high-quality connector, and there were suggestions that others should be included. It was suggested that the insistence on including only the SC connector had caused a loss of interest in the project. IEC SC 86C has standardized several new connectors and it is not known whether any of these are likely to be used for audio. R. Caine mentioned a connector designated as "ST1," recommended for use with MADI and FCDI, of interest to SC-02-02 as a connector already in use for AES3. A dual version has been adopted by an ISO committee. Caine was invited to submit a document on the subject, preferably with proposed texts for an amendment or a new document.

J. Gaunt considered that the type of connector was less important than the signal protocol. The chair suggested a set of tables showing which connectors, fiber types, and signal protocols worked together and which did not.

It appears unacceptable to let the existing AES32-tu draft go forward to publication as a full standard; it needs to be reviewed for further improvements. It was agreed that, at the present stage of technology, the document should be a report of what is actually being used, rather than a standard specifying what shall be used.

It was agreed that a user survey should be carried out to determine which connectors are actually in use. A form was designed and some responses were obtained during the Convention. Responses would be collected by the Standards Secretariat and sent to the Task Group. The closing date for the submission of responses was set at 31 July 2002. The draft was then reviewed in detail.

The following text was agreed: "This Report mainly deals with connectors for use with IR of wavelengths 1300 nm and 850 nm. 1300 nm is most widely used, while 850 nm is occasionally used for graded-index applications."

It is possible that more terms should be defined, if really necessary. The texts in the Glossary need not be as formal as that of a definition.

Additional connectors, ST1 and MTRJ should be mentioned, together with any others found to be in significant use according to the results of the survey. More information should be submitted on the ST1 (Caine) and the MTRJ (Gaunt) connectors.

It was stated that IEC MT 61806 looks to AES for input on professional applications of optical fiber.

AES-X40 Compatibility of Tip-Ring-Sleeve Connectors Conforming to Different Standards

AES-R3 has been issued as a project report. Current goal is review of this document with a target date of 2005-10. No further action is required.

AES-X105 Modified XLR-3 Connector for Digital Microphones

The work done by members of SC-04-04D on a digital microphone connector will be used by M. Natter to develop mechanical specifications for the connector. It is our hope to have that material prepared for inclusion in the IEC XLR standard as an amendment during the next maintenance cycle (roughly three years). The Secretary of IEC SC 48B has been notified. The digital connector should include the option of a capacitor for the concentric connection of the shield to the shell. The current goal is a standard. The target date is 2003-05.

AES-X113 Universal Female Phone Jack

The project intent is to develop a specification for a jack that will take both international IEC 60603-11 jacks and B-gauge plugs. The original project initiation form or a new one will be produced as soon as possible. A. Eckhart has previously indicated that such a connector exists and volunteered to document it.

AES-X123 XL Connectors to Improve Electromagnetic Compatibility

The new intent is a Performance Standard for both male and female connectors that includes performance limits, defines test fixtures and methods, and defines an objective for each generic type. The goal is a PTD. The target date is 2003-10.

At least six connector types were discussed, as set out in the draft X13 document. Two are cable-mounted types, both male and female, intended for the termination of both microphone and line level circuits. These connectors should contain a concentric capacitor to terminate the shield to the shell at radio frequencies. Two connector types, male and female, are intended for use within equipment, and according to the recommendations of SC-05-05, should offer greatly improved contact between the shells of mating connectors, greatly improved contact between pin 1, the shell, and the outside of the chassis (or shielding enclosure). It was expected that small radio frequency bypass capacitors would be fitted between pins 2 and 3 and the shell. Two connector types, male and female, are intended for use on wiring panels external to equipment.

M. Natter reported on research work on several pre-production prototypes of a male cable-mount connector incorporating a capacitor of concentric construction. B. Whitlock reported, via e-mail, on preliminary measurements of concentric capacitors. Measurement techniques, including suitable test fixtures, were discussed. The test jig should allow determination of the impedance of the capacitor over a wide frequency range by the measurement of the voltage divider ratio when a generator of known impedance drives the capacitor.

Woodgate described a test jig he has constructed that connects the center conductor of a coaxial connector from the generator to the shield connection of the concentric capacitor. The shell of the coaxial connector is connected nearly concentrically to the shell of the connector in which the concentric connector is mounted. The cable shield also connects to pin 1 through a suitable ferrite bead. The center

conductor of a coaxial feed to an rf voltmeter (spectrum analyzer, network analyzer, receiver) connects to pin 1 of a mating connector, and the shell of that coaxial connector is connected to the shell of the connector holding the capacitor being tested. The effectiveness of the test jig could be determined by replacing the capacitor with a copper disc that short circuits the generator but makes contact with pin 1 in the same manner as it would with the capacitor, and by measuring at the same point as before. Under the short circuit condition a very high value of attenuation should be measured over the frequency range of interest. Brown noted the need for measurements that test for detection of radio frequency energy by a differential input stage connected when a signal is injected in the same manner as in the Woodgate tests.

Natter and Woodgate will continue work on one or more prototypes of connectors that include a concentric capacitor and measurements of their performance.

Commonly used panel-mount connectors with mounting flanges that contact the outside of the chassis or panel were seen as nearly ideal for contact between the shell and the panel. Some redesign is needed to reduce the impedance (principally the inductance) of the connection between pin 1 and the chassis by shortening the signal path. This may be accomplished by connecting pin 1 to the shell within the connector.

Male and female connectors were described to meet the requirement of insulating XL connector shells from a mounting panel external to equipment. These connectors are expected to use a capacitor between pin 1 and the panel and another between the shell and the panel. Because of the way SC-05-05 expects these connectors to be used, these capacitors can be conventional types having good high frequency properties. It was noted that the capacitor between the shell and the chassis can be subjected to considerable stress from ESD, and should be of a type suitably rated for that condition.

AES-X130 Category-6 Data Connector in an XL Connector Shell

The intent of the standard to be developed needs clarification. J. Woodgate will put a summary of IEC SC 48B documents on CAT6 connectors on the reflector. The current goal is a PTD. The target date is 2003-05.

New projects

No project requests were received or introduced.

New business

There was a discussion of the changes in scope for SC-05-02 and SC-05-03 expressed in the minutes of the November 2001 meeting. It was felt that, although accurately reported in those minutes, the wording should be revised slightly for clarity.

Secretariat note: At the subsequent meeting of Subcommittee SC-05, the proposed clarifications were adopted with minor amendments so the the scopes now read:

“The scope of the SC-05-02 Working Group on Audio →

Connections shall include, within the bounds of the scope of SC-05, new usage, description, and contact designation for connectors for audio and ancillary functions.

“The scope of the SC-05-03 Working Group on Audio Connectors shall include, within the bounds of the scope of SC-05, documentation of established connector usages for audio and ancillary functions.”

The next meeting is scheduled to be held in conjunction with the AES 113th Convention in Los Angeles, CA, US.

Report of the SC-05-05, Working Group on Grounding and EMC Practices of the SC-05 Subcommittee on Interconnections, held in conjunction with the AES 112th Convention in Munich, Germany, 2002-05-10

The meeting was convened by Chair B. Olson.

The agenda order was revised by the chair to have the AES-X13 item appear at the end of Current Projects. The revised agenda was approved. The report of the meeting in November 2001 in New York City was accepted as written.

Current projects

AES-X27 Test Methods for Measuring Electromagnetic Interference

This document is intended to be an Engineering Report outlining useful procedures for measuring the electromagnetic interference created by real-world conditions.

A report called “Informal immunity testing of small microphone pre-amplifiers and mix consoles” was presented by J. Brown. This showed the general direction that we will need to go to produce a set of guidelines allowing users to determine the susceptibility of equipment to electromagnetic interference.

R. Caine pointed out that this testing could be much more stringent than the legal requirements and is intended to show the ability of equipment that is most suitable for the highest quality professional audio systems.

It is hoped that manufacturers will be encouraged to produce a low-cost test generator that will allow suitable testing of equipment using a variety of test fixtures.

Alternatively, the document will describe procedures that can be used with radios transmitters or cellular telephones as the interference source. These radios and cellular telephones would need to be used in their normal operation by licensed operators as required by relevant statute.

The following observations were made from initial experience of testing with radios and cellular telephones:

- 1) As frequency increased, the immunity improved.
- 2) As the distance from the device increased past a wavelength of the RF carrier, the cable attenuation reduced the level of the interference and therefore the immunity was also increased.
- 3) The headphone output stages of some mixers were often the most susceptible to interference.

This report on informal immunity testing was NOT directed only at AES-X27, but rather at demonstrating the

general level of immunity of some typical low cost products to a real-world interference source. Inclusion of such a test into AES-X27 was discussed and considered useful by those members present, but the need for cautions regarding the avoidance of interference to radio communications and conformance with national regulations was emphasized.

Caine pointed out that a coupling device would be useful for the radiated immunity tests.

J. Woodgate pointed out that this project originally started as a set of tests for low-frequency interference. He also pointed out that it should reference IEC 61004 as the primary test method.

Olson proposed a foreword for AES-X27 that clearly presents why these tests are different from either EN 55103-2 or IEC 61004. AES-X27 should address testing for pin 1 problems at both audio and radio frequencies, testing of excessive input and output bandwidth, and general RF immunity testing.

Brown offered to lead the writing effort for AES-X27.

The expectation is for a PWD for the Standards Project Report by 2002-09.

AES-X35 Installation Wiring Practices

Olson proposed an outline to guide further development of the document.

I. Cable Grouping

II. Grounding

- A. Mesh
- B. Star
- C. Hybrid
- D. Wrong

III. Safety

- A. Fire and Smoke

IV. Shielding

V. Cable Types

- A. Conduit
- B. Plenum requirements

VI. Wiring Practices

- A. Permanent installations
- B. Temporary installations

VII. Connectors

Woodgate pointed out the need for Normative References to the relevant safety standards and documents. This will be a substantial amount of work.

The chair will find help to merge the existing X35 document into the new outline.

AES-X112 Insulating Cable-Mount XL Connectors

A new, clearer title was proposed: “XLR free connectors with non-conducting shells.”

The intent is to create an Information Document on applications of connectors for facilities. This should include an explanation of stopgap measures using nonconductive covers to prevent inadvertent ground connections to the shielding contact. This work should appear as part of the X35 document. Woodgate offered to write a section for X35 regarding this.

AES-X125 Input Filtering for Electromagnetic Compatibility

Waldron indicated that K. Armstrong would write an information document for this project.

Brown proposed that the scope be expanded to include output and control port filtering. This need not be specific but should list the appropriate issues to be considered. The WG decided that just input and output filtering would be included at this stage.

AES-X13 Guidelines for Grounding

The scope of the document has been greatly simplified. Specific wording was discussed to remove performance requirements and clarify definitions. The meeting agreed on all the wording changes and decided that the PWD was ready to be progressed to a PCFC. The revised document will be posted to the reflector before being forwarded to the Secretariat to be formatted as a PCFC.

M. Natter presented preliminary results of measurements made on a connector prototype that incorporates some ideas for coaxial connection through a quasi-discoid capacitor to the shell of an XLR connector. The results look promising but need further investigation and testing.

Much discussion ensued about the need for protection from surges across the capacitor. More testing and measurement is required to show whether this is a problem. R. Cabot pointed out that, for the value of capacitance under consideration, it did not appear that a damaging voltage would appear across the capacitor in any configuration.

New projects

There were no new projects.

New business

There was no new business.

The next meeting will be held in conjunction with the AES 113th Convention in Los Angeles, CA, US.

Report of the SC-06-01 Working Group on Audio-File Transfer and Exchange of the SC-06 Subcommittee on Network and File Transfer of Audio meeting, held in conjunction with the AES 112th Convention in Munich, Germany, 2002-05-09

The meeting was convened by Chair M. Yonge.

The agenda was approved with the addition of a report on the progress of AES46-xxxx. The report of the previous meeting at 111th Convention in New York was approved as written.

Open projects

AES46-xxxx Radio Traffic Data Extension to Broadcast Wave Files

A Call for Comment was issued on 2002-03-07. M. Gerhardt had sent a comment requesting clarification of the order of RIFF chunks within a Broadcast Wave File (BWF) which had been satisfactorily addressed by an editorial change. There had been no other comments so far within the comment period which was due to close on 2002-06-08.

AES31-1-R AES Standard for Network and File Transfer of Audio, Part 1: Disk Format

No action was taken in this maintenance project.

AES31-3-R AES Standard for Network and File Transfer of Audio, Part 3: Simple Project Interchange

It was observed that the sampling frequencies supported within AES31-3-1999 did not include some of the higher frequencies anticipated in a revision of AES5, specifically 96 kHz and 192 kHz. These frequencies are important for record companies who wish to archive music recording, for instance. An amendment was proposed to include a multiplier parameter in the header chunk that would indicate multiples of the sampling frequencies already defined.

U. Henry is preparing a proposal document for Edit Automation.

The question of supporting video in AES31 was raised. There was a general feeling that this would be too complex for this body to handle, although there was no objection to working with some other body to assist the development of a complementary standard for simple video project interchange.

AES-X66 File Format for Transferring Digital Audio Data Between Systems of Different Type and Manufacture

There was a discussion of the filename extension to be used with Broadcast Wave Files. Some had assumed that “.bwf” would be used. It was pointed out that this was not part of the EBU specification and that the “.wav” extension was intended.

There was some concern that the usage of the Unique Material Identifier (UMID) was under renewed discussion within SMPTE with the potential to impact its use in AES-X66. More information will be sought.

The issue of higher sampling frequencies will be discussed with the EBU to avoid the risk of divergent specifications.

The document for AES-X66 still needs to be re-drafted in IEC style; the secretariat will arrange for this to happen.

AES-X68 A Format for Passing Edited Digital Audio Between Systems of Different Type and Manufacture that Is Based on Object Oriented Computer Techniques No action was taken.

AES-X71 Liaison with SMPTE Registration Authority No action was reported. It was felt that the AES should investigate registering a SMPTE label for the AES31-1 data format.

AES-X128 Liaison with AAF Association No action was taken.

New projects

No new projects were received or proposed.

New business

There was no new business.

The next meeting is scheduled to be held in conjunction with the AES 113th Convention in Los Angeles, CA, US.

AES 113th Convention

2002 October 5–8

Los Angeles Convention Center
Los Angeles, California, USA





To no one's surprise, the sun was shining brightly throughout the days of the AES 113th Convention held October 5–8 in Los Angeles, illuminating the positive outlook of the audio industry after a difficult year. The combination of a large number of exhibitors and delegates together with a wide range of new products and excellent technical events indicates an industry in good shape growing stronger.

OPENING CEREMONIES

The opening of the convention commenced with a welcome from Roger Furness, executive director of the AES, who was glad to see that so many people had come to attend the convention in spite of the sunny weather outside. He introduced AES President Garry Margolis, who said that it was an honor and a pleasure to be the president of the Society during a challenging and rewarding year. He proclaimed the AES to be a forward-looking society and hinted that exciting new developments are in the pipeline.

Floyd Toole, 113th Convention chair, discussed the convention theme: "Science in the Service of Art." He praised the audio industry as a marvelous enterprise and explained that this was shown perfectly by the diversity of activities within the convention: from the events hosted by the Historical Committee to the innovative products and software in the exhibition. But he warned that without art the business would die, and he saluted the artists who would be performing in the Songwriters Showcase throughout the convention.

Next the presentation of awards was overseen by David Robinson, the Awards Committee chair. Citations were awarded to Elmar Leal for many years of outstanding work in South America, culminating in the formation of the Latin America Region, and Roland Tan for his work in Southeast Asia and the formation of the Singapore Section. Board of Governors Awards were presented to Roy Pritts and Ron Streicher for cochairing the 109th Convention in Los Angeles in September 2000. AES Fellowships were awarded to Durand Begault for his contributions to our understanding of spatial hearing and its applications, Gilbert Soulodre for his significant contributions to procedures for subjective testing of audio systems, and Carson Taylor in recognition of lifelong contributions to the art and science of music recording techniques.

The keynote speech was given by Leonardo Chiariglione of the Telecom Italia Lab. He offered further examination of the convention theme by discussing the effect of new distribution technology on the industry and the art. He compared the repercussions of sharing MP3 files on the Internet with the changes brought about by the advent of radio and recording technologies. He outlined the tension currently created by the differing opinions of the consumers, who usually want free access to any content in a form that can be manipulated and transferred to any device, and the distributors, who want to retain control of the publishing mechanism and need a financial return on content. Encouraging both sides in the dispute to take a step back from confrontation, he proposed a solution based on open access to protected content. This solution is ➤

OPENING CEREMONIES AND AWARDS



Leonardo Chiariglione, 113th keynote speaker



Roger Furness, AES executive director



Garry Margolis, AES president



Floyd Toole, 113th Convention chair



David Robinson, Awards Committee chair, announces award recipients.



Garry Margolis presents Citation Awards to Elmar Leal, left photo, and Roland Tan.

based on interoperable technology making readily accessible content easy to distribute and requiring that the consumer obtain the right to play the content. He expanded on this by telling of his dream that any person has the opportunity to create art, distribute it, and gain just reward for it. He believes that the technology to achieve this is available and it is time to make technology and content friends of mankind. He concluded by saying that “unless people get value from their art, they can’t make it. All artists should have means to produce their art and receive remuneration.”

Toole concluded the opening ceremony by thanking his

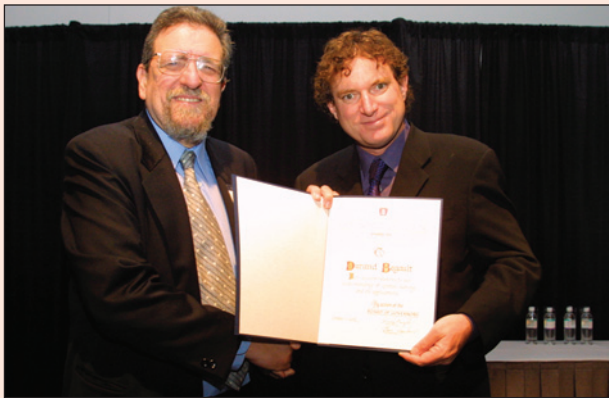
convention committee for their dedicated efforts that made the convention possible, and asked the audience to go out and enjoy the results of that hard work.

EXHIBITION

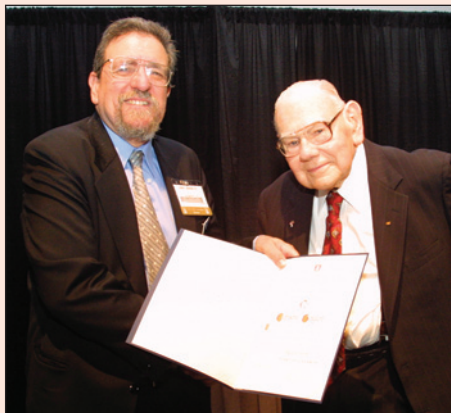
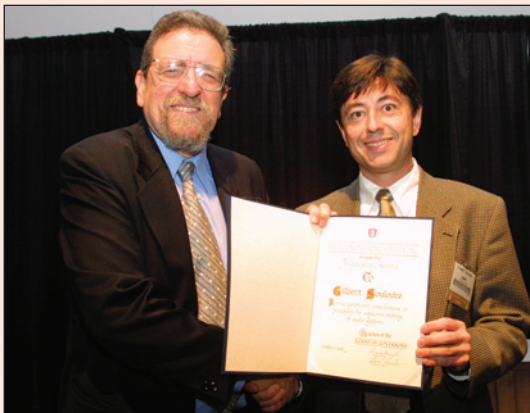
There was a very positive atmosphere on the exhibition floor throughout the convention, with the quality and quantity of the attendees beating the expectations of exhibitors. Of prime importance was the fact that the audience included a large number of the industry’s key decisionmakers. The format of the exhibition also allowed for the attendees to talk at length



Roy Pritts, left photo, and Ron Streicher receiving Board of Governors Awards.



Fellowship Awards recipients: counterclockwise from left, Durand Begault, Gilbert Souldre, and Carson Taylor.



line array products. JBL introduced three new VerTec models aimed at medium and small sound reinforcement applications, designed to give high output power and quality within a lightweight package. Meyer Sound also exhibited compact line array units, with their MID and M2D products. Electro-Voice used the convention to launch its XLC range of compact line-array loudspeakers designed to meet the challenges of difficult acoustical environments. SLS Loudspeakers also introduced its new compact line array product, the RLA/2 that includes dual 8-inch woofers and a high-frequency ribbon driver.

Renkus-Heinz demonstrated an interesting combination of technologies for the live-sound market with the incorporation of the Ethernet-based CobraNet audio distribution system into their ST-STX loudspeaker range. They claim that this allows improved signal quality, minimal data loss, and precision remote control over long distances, together with networking and interoperability with a range of other products incorporating CobraNet. Continuing this theme, Neutrik showed a range of professional EtherCon Ethernet RJ45 connectors built within a standard XLR casing in order to better cope with the stresses of the professional audio environment.

At the other end of the live sound signal chain, Audix demonstrated its new D6 kick drum microphone that was developed from detailed analysis of the sound of a large number of drums. Audio-Technica launched a range of live-sound instrument microphones, including the

with the product manufacturers and give valuable feedback.

Numerous exciting new products were displayed at the convention. A small number of these are summarized here, and a full list of exhibitors begins on page 950 of this issue.

One focus of the convention was on live sound, with two workshops on line arrays, a roundtable of live sound engineers discussing the latest trends, techniques, and tools in modern sound reinforcement, and a large number of new products from the major manufacturers. In addition to the workshops on line array loudspeakers, Meyer Sound, JBL, SLS, McCauley Sound, and Electro-Voice all showed their

AE5100 and the AE3000, which are large diaphragm condensers, and the AE2500, which combines condenser and dynamic elements within one microphone.

A large number of significant new software releases were announced at the convention. Steinberg launched version 2.0 of Nuendo, featuring a reengineered 32-bit floating-point mixer, enhanced routing options (including surround sound up to 10.2 channels) and new networking capabilities over TCP/IP. Digidesign announced version 6.0 of ProTools for Mac OS X, including a cleaner user interface, improved MIDI facilities with support for the OS X CoreMIDI en- ➤

gine, and the ability to take advantage of dual-processor Macintosh G4 computers. SADiE unveiled its Series 5 products, based on a newly designed hardware platform. This includes four new products that encompass authoring and editing for either PCM or DSD technologies, with the option of up to 8 channels.

Plugzilla is a unique product that combines the traditionally separate realms of computer-based and outboard processing. This is a stand-alone 2U unit which runs VST software plug-ins. It contains two independent machines that will run up to eight plug-ins simultaneously and has enough processing power for up to 16 channels of reverberation. The parameters of the plug-ins are accessed via the front-panel rotary controllers, MIDI, USB, and assignable foot switches.

The Virtual Mixer from the Virtual Mixing Company was another unique product. Acting as a front end to a MIDI-controlled system, it allows tracks or MIDI instruments to be mixed in a visual environment where the instruments are positioned in space according to their pitch, level, and panned position. The mix can then be manipulated visually using a mouse or a touch screen.

Continuing the popular trend for multichannel surround sound production, a number of innovative new products were released at the convention. Fostex showed a 6-channel location recorder that records directly onto DVD-RAM disks and includes microphone preamplifiers and time-code features. Lexicon demonstrated updates to the 960L processor, including an automation option and LOGIC7 up-mixing algorithms to create 5-channel surround sound from 2-channel stereo. TASCAM exhibited a surround sound monitor controller, the DS-M7.1, which is designed to add multiloudspeaker monitoring control to consoles with limited output busses.

The exhibition floor also contained a large number of small



Speaking at AES annual business meeting: clockwise from above left, Ron Streicher, secretary; Garry Margolis, president; Marshall Buck, treasurer; and Chris Freitag, Board of Tellers chair.



stands exhibiting specialist products. These included manufacturers of outboard recording equipment, such as compressors and microphone preamplifiers. Anthony De-Maria Labs showed a number of tube compressors that are

based on classic designs. Universal Audio unveiled a new channel strip, the 6176, which combines a microphone preamplifier with a

compressor in a single unit, providing vintage character with modern specifications.

Finally, a number of microphone products were shown at the convention, though not all were new designs. Most notable among these was the Telefunken Ela-M 251, a reissue of the classic tube microphone that is hand built to original specifications using the same methods used to make the original version 40 years ago. With the mixture of cutting-edge and classic, the exhibition contained something for every audio professional.

PAPERS SESSIONS

Papers cochairs Eric Benjamin and John Strawn organized sessions at the 113th Convention that covered the advanced audio research being performed across a wide range of topics. Based on the number of papers presented in each area, it appears that the themes of psychoacoustics, low bit-rate coding, signal processing, and the design and measurement of transducers are those currently of greatest research interest.

A number of papers were related to the measurement of loudspeaker drivers. One such paper by Siegfried Linkwitz investigated the loudspeaker parameters that are of greatest importance. He argued that volume displacement, intermodulation distortion, stored energy, and off-axis frequency response all need to be tightly controlled, and that the importance of phase linearity and cabinet diffraction are



Author Siegfried Linkwitz presenting paper.



Multichannel demo room at 113th allowed presenters to demonstrate new theories with 5.1-channel samples.

Your resource for advanced audio and voice solutions



Stereo

Mono

SRS Labs is a recognized leader in developing audio solutions for any application. Its diverse portfolio of proprietary technologies includes mono and stereo enhancement, voice processing, multichannel audio, headphones, and speaker design. • With over seventy patents, established platform partnerships with analog and digital implementations, and hardware or software solutions, **SRS Labs is the perfect partner for companies reliant upon audio performance.**

Multichannel

Technologies

- Circle Surround II
- FOCUS
- SRS 3D
- SRS Headphone
- TruBass
- TruSurround XT
- VIP
- WOW

Product Applications

- Home Theater/Entertainment
- Wireless + Portable
- Telecom + Voice
- Gaming
- Internet + Broadcast

World Wide Partners

Aiwa, AKM, Analog Devices, Broadcom, Cirrus Logic, ESS, Fujitsu, Funai, Hitachi, Hughes Network Systems, Kenwood, Marantz, Microsoft, Mitsubishi, Motorola, NJRC, Olympus, Philips, Pioneer, RCA, Samsung, Sanyo, Sherwood, Sony, STMicroelectronics, Texas Instruments, Toshiba



Van Webster (left), facilities chair, and Floyd Toole, convention chair



Marshall Buck (center), workshops chair, flanked by workshops assistants David Sheirman (left) and Bejan Amini



Papers cochairs John Strawn (left) and Eric Benjamin



Ron Streicher (right), special events cochair



From left, Theresa Leonard, education events chair, Elliot Scheiner, renowned producer, and Scott Cannon, education events assistant



From left, historical events cochairs Dale Manquen and Wes Dooley, and Eric Gaskell who assisted with *When Vinyl Ruled* exhibit.



From left, Mel Lambert, special event cochair and technical tours cochair, Lisa Roy, Platinum Series chair, and Roger Furness, AES executive director

sometimes exaggerated. Other papers in this field were on topics such as techniques for measuring the amplitude response of loudspeaker systems in domestic environments by Allan Devantier and methods for interpreting nonlinearity measurements of transducers by Alexander Voishvillo, Alex Terekhov, Gene Czerwinski, and Sergei Alexandrov.

The sessions on signal processing included papers on filter morphing by Rob Clark, Emmanuel Ifeachor, and Glenn Rogers, sound synthesis by genetic programming by Ricardo Garcia, and noise shaping in test-signal generation by Stanley Lipshitz, John Vanderkooy, and Edward Semyonov. There was also a paper by Frank Siebenhaar, Christian Neubauer, Robert Bäuml, and Jürgen Herre on audio watermarking based on a Scalar Costa Scheme, which allows the inclusion of higher data rates within the watermark for em-



Floyd Toole (left) and Bob Lee, signs chair



Shelley Herman (left), chair of *The Jazz Singer*, and Ken Lopez, volunteers coordinator



Han Tendeloo, program coordinator



Peter Chaikin (right) technical tours cochair with Rich Tozzoli



Claudia Koal, music chair, with Gordon van Ekstrom



Ken Lopez (center), volunteers coordinator, with 113th volunteers



Marla Egan (right), producer for Platinum Series, with Lisa Roy

We Thank...

Floyd Toole
chair

Eric Benjamin John Strawn
papers cochairs

Marshall Buck
workshops chair

Bejan Amini David Scheirman
workshops assistants

Mel Lambert Ron Streicher
special events cochairs

Lisa Roy
Platinum Series chair

Theresa Leonard
education events chair

Scott Cannon
education events assistant

Wes Dooley Dale Manquen
historical events cochairs

Van Webster
facilities chair

Peter Chaikin Mel Lambert
technical tours cochairs

Claudia Koal
music chair

Shelley Herman
chair, *The Jazz Singer*

Bob Lee
signs chair

Han Tendeloo
program coordinator

Ken Lopez
volunteers coordinator

bedding such things as song lyrics or pictures.

Multichannel sound and spatial audio appeared frequently within the sessions mentioned above, as well as in a separate session on multichannel sound. Listener envelopment in surround sound was investigated in a paper by Gilbert Soulo-dre, Michel Lavoie, and Scott Norcross, based on applying measurements developed for concert hall acoustics. Other papers in the session on multichannel sound included an investigation of the interchannel interference from multiple loudspeakers at the listening position and its relationship to microphone technique by Geoff Martin, the effect of early reflections on localization by Jason Corey and Wieslaw Woszczyk, an approach for synthesizing surround sound from mono and two-channel stereo by Ching-Shun Lin and Chris Kyriakakis, and an investigation of loudspeaker ar-

rangements for creating a diffuse sound field by Koichiro Hiyama, Setsu Komiyama, and Kimio Hamasaki.

Further papers sessions included *Room Acoustics and Sound Reinforcement*, *High Resolution Audio*, *Recording and Reproduction of Audio*, and *Audio Networking and Automotive Audio*. Individual papers presented at the convention can be purchased at www.aes.org/publications/preprints, and a CD-ROM with all the 113th Convention papers is also available. A full list of paper titles and abstracts begins on page 954 of this issue.

WORKSHOPS

The workshops, organized by Marshall Buck with the assistance of Bejan Amini and David Scheirman, provided an interesting mix of the current issues facing the industry. ➤



Jim West (second from left), who gave Richard C. Heyser Memorial Lecture at 113th, receives citation from Wieslaw Woszczyk (left), Technical Council chair, and vice chairs, Robert Schulein and Jürgen Herre (right).



Quantity and quality of attendees exceeded exhibitors' expectations.



Vintage equipment displayed at *When Vinyl Ruled* exhibit.

the summary that it is most important for engineers to trust their ears and adapt the available techniques to the particular situation. John Eargle then gave a historical view of stereo and surround microphone techniques, before Doug Botnick, Michael Bishop, Richard King, and Mick Sawaguchi discussed how they apply the available techniques to their own productions.

There was a standing-room-only crowd for *Mixing and Mastering in Multichannel Surround*, chaired by Michael Bishop. The first presentation was by Bob Ludwig, who explained the challenges of mastering in surround sound, including advice on the use of the LFE channel. John Eargle described the way in which he simultaneously mixes for surround and stereo, referring to the problems of producing stereo from the surround mix by downmixing. Frank Filipetti discussed the problems of creating a multichannel product from a stereo mix, and encouraged record companies to specify how multichannel mixes were created. Finally, the opportunities provided by multichannel surround sound were explained by George Massenburg and Elliot Scheiner. Massenburg described it as a "sandbox" as he talked about why he enjoyed surround mixing, and Scheiner explained that, "there are no rules, but you have to maintain the integrity of the music when you redo it in surround."



Stereo and Surround Microphone Techniques was one of 15 well-attended workshops.

In addition to these two workshops on multichannel surround sound, there was also *The Application of Multichannel Sound Formats* ➔

In addition to the usual lively discussion sessions, this convention continued the recent trend of an increasing number of tutorial workshops featuring recognized experts in their respective fields.

Surround sound and multichannel audio continued to be a popular topic at this convention. The first workshop was *Stereo and Surround Microphone Techniques*, chaired by Geoff Martin. Despite the early start of this workshop, a large crowd attended to hear the panel discuss microphone techniques that are applicable to a wide range of genres. The session opened with an informative presentation by Martin on the basic theories behind simple microphone arrays, with



Attendees were able to sample advanced automotive audio systems.



THAT 300 Series

Low-Noise Matched Transistor Arrays

- 4 Matched NPNs (Model 300)
4 Matched PNPs (Model 320)
2 Matched NPNs and PNPs (Model 340)
- Monolithic Construction
- Low Noise:
0.75 nV / Hz^{1/2} (PNP) 0.80 nV / Hz^{1/2} (NPN)
- High Speed:
 $f_T = 325$ MHz (PNP) $f_T = 350$ MHz (NPN)
- Excellent Matching:
500 μ V typ. between devices of same gender
- Dielectrically Isolated for low crosstalk and high DC isolation
- 36V V_{CEO} Minimum

THAT Corporation

Making Good Sound Better®

45 Sumner Street, Milford MA 01757-1656 USA

Tel: +1 508 478 9200 Fax: +1 508 478 0990 Web: www.thatcorp.com

113TH SPECIAL EVENTS

1. Platinum Producers Panel 1: from left, Ben Grosse, Tal Herzberg, Rob Cavallo, moderator Mitch Gallagher, Ron Fair, and Mike Elizondo.

2. Platinum Producers Panel 2: from left, Larry Levine, Michael Bradford, moderator Howard Massey, Phil Ramone, Patrick Leonard, and Bob Ezrin.

3. Road Warriors Panel: from left, Kirk Kelsey, David Morgan, Greg Dean, moderators Steve Harvey and Clive Young, and Paul Gallo.

4. GRAMMY Recording Soundtable: AES President Garry Margolis welcoming audience and panelists, from left, Ken Jordan, George Massenburg, Jack Joseph Puig, moderator Howard Massey, Elliot Scheiner, Brian Garten, and Al Schmitt.

5. An Afternoon with...Tom Holman: Tom Holman (left) and moderator George Petersen.



in *Vehicles*, chaired by Richard Stroud, and *Coding of Spatial Audio*, chaired by Christof Faller. The surround sound workshops were enhanced by the availability of a multi-channel reproduction system in a small room which enabled the presenters to demonstrate their recordings in a more acoustically suitable environment. Other demonstrations relating to the workshops program included some automobiles with multichannel audio systems.

As part of the promotion of new areas of audio technology within the AES, the 113th Convention included the workshop *Game Audio*, with presentations by experts from a range of sections of the computer game industry. Topics covered the use of surround sound in video games, issues in programming audio for games, and the audio capabilities of the Xbox. Also included was a discussion of

how to enter the game audio industry.

Loudspeaker technology received significant attention, with *Large Signal Loudspeaker Parameter Measurement*, chaired by John Stewart, and the two-part *Loudspeaker Line Arrays*, chaired by Jim Brown and John Murray.

The other workshops covered a wide range of topics: *New Media for Music*, chaired by Dave Davis; *What Audio Engineers Should Know About Human Sound Perception*, presented by Durand Begault and William Martens; *AES 42-2001*, chaired by Stephan Peus; *Studio Production and Practices*, chaired by George Massenburg and David Smith; *Perceptual Issues Related to Cascaded Audio Codecs*, chaired by Thomas Sporer; *Protecting Your Hearing Against Loss*, chaired by Dilys Jones and Sigfrid Soli; and *Recent Developments in MPEG-4 Audio*, chaired by Jürgen Herre.

HEYSER LECTURE

The Richard C. Heyser Memorial Lecture was presented at this convention by James E. West, coinventor of the modern electret microphone. His lecture focused on its development and the unique applications of this technology to a wide range of technical areas. He started by reviewing the history of the technology, from the first experiments with electrets in the 15th century to the early electret microphones used by the Japanese in World War II. He went on to explain that the problem with these electrets was that they were wax-based, which caused their characteristics to be dependent on the atmospheric conditions and limited their lifespan to about six months.

The main achievement of West and his colleagues at Bell Labs was the discovery of stable charge storage in thin polymers and using them in microphone technology. West explained that whereas condenser microphones might be the best choice for some applications, electret microphones can be used in a wider range of situations, such as compact arrays.

West demonstrated the use of electret microphones together with signal processing for use in telecommunications. He showed an audio and video clip of the use of four microphones to obtain a second-order directional response to reduce background noise. This was followed by a demonstration of the use of signal processing to adapt for the distance between the mouth of the talker and the microphone array. In this case, the variations in the level and the frequency response caused by varying the talker-microphone distance were compensated based on analysis of the audio signals reaching each of the microphones in the array. West also demonstrated how this technology was adapted for special applications, such as communications in motor racing to overcome the high ambient noise levels in race cars.

In conclusion, West described the use of electret microphones in arrays of up to three dimensions. This enables their use in a wide range of applications, from reducing background noise for speech recognition, through steered directivity for amplifying audience questions, to recording and reproducing 3D sound.

In addition to the lecture, the Technical Council was busy throughout the convention with committee meetings that covered a wide range of topics. The AES Technical Committees review trends in audio technology and practice and provide expertise to the Society, such as suggesting topics for workshops and conferences, publishing technical documents, and collating written papers and audio demonstrations on important topics. More information on the Technical Council and the Committees can be found at www.aes.org/technical.



Among many artists who performed at Songwriters Showcase: top to bottom, Severin Browne and James Coberly Smith, Deni Bonet, Vinx, and Tim Janis.



113TH MEETINGS

1. Technical Council: from left, Jürgen Herre, vice chair, Wieslaw Woszczyk, chair, Bob Schulein, vice chair.
2. Publications Policy Committee: from left, Daniel von Recklinghausen, Emil Torick, and Garry Margolis.
3. Standards Committee: Mark Yonge (left) and John Nunn.
4. Historical Committee: Jay McKnight and Dale Manquen.
5. Regions and Sections Committee: Subir Pramanik.



SPECIAL EVENTS

The program of special events at the convention included the 14th annual GRAMMY Recording Soundtable, two Platinum Producers sessions, and a Road Warriors Panel. Those who attended the Recording Soundtable heard Phil Ramone, Ken Jordan, George Massenburg, Jack Joseph Puig, Elliot Scheiner, and Al Schmitt discuss how they cope with the challenges of the recording industry. In the first Platinum Producers session Rob Cavallo, Mike Elizondo, Ron Fair, Ben Grosse, and Tal Herzberg, with moderator Mitch Gallagher, discussed trends in engineering and producing, such as how the traditionally separate roles of producer, engineer, and studio technician are starting to blend into one.

In the second Platinum Producer session Howard Massey moderated a discussion with panelists Michael Bradford, Bob Ezrin, Patrick Leonard, Larry Levine, and Phil Ramone on the past, present, and future of the recording industry. Continuing the theme of the keynote speech at the opening ceremony, they discussed the difficulties of creating and distributing new and possibly commercially risky music. They noted that less risks are being taken with new artists in the current business climate, meaning that the songs produced are becoming more and more similar. They briefly touched on the possibility that new methods of distribution could help this situation, but might also lead to the problem of not getting an adequate financial return.

On Tuesday Paul Gallo introduced the Road Warriors Panel of Kirk Kelsey, David Morgan, Greg Dean, and moderators Steve Harvey and Clive Young, who discussed the latest trends, techniques, and tools in the live-sound industry.

To commemorate the 75th anniversary of the first successful talking picture, on Sunday evening the convention was host to an old-time radio recreation of *The Jazz Singer*. The show featured Richard Halpern in the starring role made famous by Al Jolson, along with a supporting cast that included AES President Garry Margolis. Convention Chair Floyd Toole commented that this event exemplified the convention theme, "Science in the Service of Art," as the production was art at its highest.

Throughout the convention attendees had the opportunity to take advantage of free hearing screening cosponsored by the AES and the House Ear Institute in response to a growing interest in hearing conservation and to heighten awareness of the need for hearing protection and the safe management of sound.

The Historical Committee again put on a number of demonstrations and talks, as well as an exhibition of production technology during the age of vinyl, organized by Wes Dooley and Dale Manquen. The presentations included Carson Taylor discussing location recording from the 1950s to the 1970s and a review of 12 landmark microphones by Jim Webb. Other presentations covered topics such as vinyl disk mastering, record manufacturing, the development of powered monitors, and the development of mixing console technology.

The Art part of the convention theme was on display at the Songwriters Showcase held in the main entrance area of the convention center. Attendees had the chance to listen to a wide range of live music performed throughout the four days of the convention. The music continued each evening at mixer events, which allowed the convention attendees to meet in a relaxed social atmosphere.

Other special events at the convention were: "An Afternoon with...Tom Holman," in which moderator George Petersen interviewed Holman about his career in audio; the SPARS Business Panel; and there were sessions on technical areas that are currently of interest such as Audio Post Production in 24p HDTV and The Virtual Studio: DAW-Based Recording.

EDUCATION EVENTS

Numerous education events also took place, arranged by Theresa Leonard and Don Puluse with the assistance of Scott Cannon and members of the student delegate assembly. The program gave students the opportunity to present their own work, including paper and poster sessions for technical work and a recording competition for practical work. For the first time a number of the educational events were included in the main convention workshop program,



113th Education events: counterclockwise from top right: Education Fair, Mentoring Panel, students anxiously await start of Recording Competition, and Stereo and Surround Microphone Techniques workshop.



dak Theater, the Hollywood Bowl, and Platinum Live. A final event included a tour of the newly completed Cathedral of Our Lady of the Angels. This event was held in combination with the American Institute of Organ Builders, and included lectures on the new 105-rank, 6,019-pipe organ. This visit concluded with a recital by AES organist-in-residence Graham Blyth. He performed works by Soler, Franck, Messiaen, Bach, and Viern on the new organ to an enthusiastic audience.

AES COMMITTEE MEETINGS

The various committee meetings that help to keep the Society running were held in the days before, during, and after the four days of the convention.

The meetings of the Standards Committee commenced under the administration of Standards Manager Mark Yonge and Standards Committee Chair John Nunn two days before the convention began. As part of the mission of the committee to develop standards relating to current practice and technology in professional audio, there were discussions on a wide range of topics, including digital audio, forensics, listening tests, file transfer, and measurement. The convention also saw the launch of the new website and email facilities for the Standards Committee to help develop the electronic communication of the working groups. For more information see www.aes.org/standards.

Other Society meetings were held throughout the convention. Following the convention the AES Board of Governors (see next page) met to discuss the issues facing the Society and to consider aspects of policy and planning.

Attendees left the convention with insights, possibilities, ideas, and inspiration in their chosen field and about the industry in general. Now they can look forward to the next European convention, the AES 114th to be held March 22–25, 2003 in Amsterdam. For more information on this and other Society activities, visit the AES website at www.aes.org.

these being the workshops on multichannel microphone techniques and mixing and mastering in surround sound mentioned above. It turned out that these events were also popular with nonstudents, resulting in some of the most highly attended events of the convention.

There was also a strong focus on mentoring. This enabled students to meet and talk to a number of respected audio engineers from a wide range of backgrounds, both in a group situation and in one-to-one sessions. The Education Forum hosted a discussion about the issues facing education in audio engineering, and enabled those involved in education to meet and share opinions. At the Education Fair attendees were able to learn more about the course offerings in audio at numerous colleges and universities.

TECHNICAL TOURS

Peter Chaikin and Mel Lambert organized a fascinating range of technical tours to complement the other convention activities. There were tours of a number of the leading recording studios in the Los Angeles area, including Sunset Sound Studios, Glenwood Place Studios, Capitol Recording Studios, and Cello Recording Studios. Tours were also organized to a number of performance venues including the Ko-



Board of Governors Meets

Meeting on October 9, members of the AES Board of Governors gather from around the world to hear reports from AES officials and standing committees:

- 1** Kunimara Tanaka, incoming governor
- 2** Daniel von Recklinghausen, editor; Marina Bosi, governor; Roger Furness, executive director
- 3** Theresa Leonard, governor and 24th International Conference chair; John Nunn, Standards Committee chair
- 4** Annemarie Staepelaere, governor; Roland Tan, governor; Markus Erne, Europe Central Region vice president
- 5** Jim Anderson, USA/Canada Eastern Region vice president; Curtis Hoyt, incoming governor
- 6** Marshall Buck, treasurer, Convention Policy Committee chair, and Finance Committee chair; Karl-Otto Bäder, governor
- 7** Mercedes Onorato, Latin American Region vice president; Daniel Zalay, Europe Southern Region vice president
- 8** Juergen Wahl, governor; Neville Thiele, International Region vice president



9 Scott Cannon, student representative, Roy Pritts, past president and Nominations Committee chair



10 Han Tendeloo, incoming secretary; Wieslaw Woszczyk, Technical Council chair

11 Jay Fouts, legal counsel; Richard Small, Publications Policy Committee chair



12 Jay McKnight, Historical Committee chair; Subir Pramanik, Regions and Sections Committee chair

13 Bob Sherwood, financial advisor; Nick Zacharov, governor

14 Bob Moses, USA/Canada Western Region vice president; James Kaiser, USA/Canada Central Region vice president

15 Søren Bech, Europe Northern Region vice president and Conference Policy Committee chair



16 Garry Margolis, president, Future Directions Committee chair, and Women in Audio Committee acting chair; Ron Streicher, secretary and incoming president elect

17 Don Puluse, Education Committee chair and incoming governor; David Robinson, governor and Awards Committee Chair



113th Exhibitors

Los Angeles • California • 2002 October 5–8



A.D.A.M. Audio GmbH
Berlin, Germany
A Designs
West Hills, CA, USA

Aardvark
Ann Arbor, MI, USA
Ableton
Arcadia, CA, USA

***ACO Pacific, Inc.**
Belmont, CA, USA

Acoustic Systems
Austin, TX, USA

Acoustical Solutions, Inc.*
Richmond, VA, USA

Adamson Systems Engineering

Ajax, Ontario, Canada

ADK Microphones
Ridgefield, WA, USA

Advanced Audio
North Hollywood, CA, USA

AEA
Pasadena, CA, USA

AEQ, S. A.
Leganes, Madrid, Spain

Akai Musical Instrument Corporation
Ft. Worth, TX, USA

***AKG Acoustics, US**
Nashville, TN, USA

***AKM Semiconductor, Inc.**
San Jose, CA, USA

Alcorn McBride, Inc.
Orlando, FL, USA

Alesis Distribution LLC
Los Angeles, CA, USA

Allen & Heath USA
Agoura Hills, CA, USA

AlterMedia
Burbank, CA, USA

Amek
Nashville, TN, USA

Ametron Audio/Video
Hollywood, CA, USA

***AMS Neve PLC**
Burnley, Lancashire, UK

Analog Devices, Inc.
Norwood, MA, USA

Anthony DeMaria Labs
New Paltz, NY, USA

Apex N.V.
Genk, Belgium

Aphex Systems
Sun Valley, CA, USA

Apogee Electronics, Inc.
Santa Monica, CA, USA

Applied Microphone Technology
Berkeley Heights, NJ, USA

ART-Applied Research & Technology
Rochester, NY, USA

APRS
Totnes, UK

Arboretum Systems, Inc.
Pacifica, CA, USA

Arkaos
Arcadia, CA, USA

***ATC Loudspeaker Technology**
Stroud, Gloucester, UK

ATC / Transamerica Audio Group, Inc.
Las Vegas, NV, USA

ATI - Audio Technologies, Inc.
Horsham, PA, USA

The ATI Group
Jessup, MD, USA

ATR Service Co.
York, PA, USA

Audio Accessories, Inc.
Marlow, NH, USA

Audio Developments
Portland, ME, USA

Audio Engineering Associates
Pasadena, CA, USA

***Audio Precision**
Beaverton, OR, USA

Audio Specialties
Herndon, VA, USA

***Audio-Technica U.S., Inc.**
Stow, OH, USA

Audio Technology Magazine
Dee Why, NSW, Australia

Audix Corporation
Wilsonville, OR, USA

Auralex Acoustics
Indianapolis, IN, USA

Avalon Design, Inc.
San Clemente, CA, USA

Bader LLC
Danville, KY, USA

Baltic Latvian Universal Electronics "B.L.U.E."
Westlake Village, CA, USA

Belden Electronics Division
Richmond, IN, USA

Benchmark Media Systems, Inc.
Syracuse, NY, USA

Berklee College of Music
Boston, MA, USA

BIAS (Berkley Integrated Audio Software)
Petaluma, CA, USA

BitHeadz, Inc.
East Greenwich, RI, USA

Blue Sky / Group One
Farmingdale, NY, USA

Boss
Los Angeles, CA, USA

Brainstorm Electronics, Inc.
North Hollywood, CA, USA

Brauner / Transamerica Audio
Las Vegas, NV, USA

Brill Electronics
Oakland, CA, USA

Bruel & Kjaer North America
Norcross, GA, USA

Bryston Ltd.
Arcadia, CA, USA

***BSS Audio**
Nashville, TN, USA

Cable Factory
Burnaby, B.C., Canada

***Cadac Electronics Ltd.**
Luton, Bedfordshire, UK

Cakewalk
Boston, MA, USA

Caldwell Bennett Inc.
Oriskany, NY, USA

***Calrec Audio Ltd.**
Hebden Bridge, West Yorkshire, UK

CB Electronics
Charvil, Berks., UK

***Cedar Audio Limited**
Fullbourn, Cambridge, UK

Celestion / Group One Ltd.
Farmingdale, NY, USA

Cherry Lane Magazines, LLC - Home Recording Magazine

New York, NY, USA

Chevin Research
Old Lyme, CT, USA

Church Production Magazine
Carey, NC, USA

Cirrus Logic Inc.
Austin, TX, USA

Cliff Electronic Components, Inc.
Benicia, CA, USA

Club Systems
Port Washington, NY, USA

Coding Technologies
Nürnberg, Germany

Coles
Portland, ME, USA

Coles / AEA
Pasadena, CA, USA

* Sustaining Member of the Audio Engineering Society

***Community Professional Loudspeakers**
Chester, PA, USA

Compuplus Inc. / Mac Hollywood
Hollywood, CA, USA

Convention TV @ AES
Port Washington, NY, USA

Cooper Sound Systems
San Luis Obispo, CA, USA

Countryman Associates, Inc.
Redwood City, CA, USA

Crane Song Ltd.
Superior, WI, USA

Creative Network Design
Alameda, CA, USA

Crest Audio
Fair Lawn, NJ, USA

Crown International
Elkhart, IN, USA

Cycling '74
San Francisco, CA, USA

***D.A.S. Audio of America, Inc.**
Valencia, Spain

D.W. Fearn
Westchester, PA, USA

D2Audio Corporation
Austin, TX, USA

DACS Ltd.
Portland, ME, USA

Dan Dugan Sound Design
San Francisco, CA, USA

Data Linked Products
Orange, CA, USA

dbx Pro
Sandy, UT, USA

***dCS Ltd.**
Portland, ME, USA

Demeter Tube Amplification
Van Nuys, CA, USA

Denon Electronics
Pine Brook, NJ, USA

Design FX Audio
Burbank, CA, USA

DiGiCo / Soundtracs
Epsom, Surrey, UK

***Digidesign**
Daly City, CA, USA

Digidesign Development Partners
Daly City, CA, USA

***Digigram**
Arlington, VA, USA

Digital Music Technologies
Burbank, CA, USA

Disc Makers
Pennsauken, NJ, USA

DJ Times
Port Washington, NY, USA

DK Audio A/S
Herlev, Denmark

***Dolby Laboratories, Inc.**
San Francisco, CA, USA

Doremi Labs, Inc.
Burbank, CA, USA

Dorrrough Electronics
Woodland Hills, CA, USA

DPA Microphones / TGI North America Inc.
Kitchener, Ontario, Canada

Drawmer (USA) / Transamerica Audio Group, Inc.
Las Vegas, NV, USA

Digital Theater Systems, Inc.
Agoura Hills, CA, USA

Dupco
Salem, NH, USA

Dynaudio Acoustics
Westlake Village, CA, USA

Earth Works Audio Products
Milford, NH, USA

***Eastern Acoustic Works, Inc.**
Whitinsville, MA, USA

ESI (EGO SYS Inc.)
San Jose, CA, USA

Electro-Voice
Burnsville, MN, USA

Electronic Musician
Emeryville, CA, USA

Elsevier Science
New York, NY, USA

Embracing Sound Experience
Stockholm, Sweden

Equi=Tech Corporation
Selma, OR, USA

ETA Systems
Twinsburg, OH, USA

Euphonix, Inc.
Palo Alto, CA, USA

Eventide, Inc.
Little Ferry, NJ, USA

Expression Center for New Media
Emeryville, CA, USA

Fairlight USA
Hollywood, CA, USA

FAR Fundamental Acoustic Research
Ougree, Belgium

***Fostex America**
Norwalk, CA, USA

***Fraunhofer Institut Fuer Integrierte Schaltungen**
Erlangen, Germany

FriendChip America
West Hollywood, CA, USA

Furman Sound
Petaluma, CA, USA

Gefen Systems
Woodland Hills, CA, USA

Genelec
Natick, MA, USA

Genex Audio
Santa Monica, CA, USA

Geoffrey Daking & Co., Inc.
Treasure Island, FL, USA

Gepco International, Inc.
Gee Plains, IL, USA

Gibson Labs
Redondo Beach, CA, USA

GML, LLC
Franklin, TN, USA

Gold Line
West Redding, CT, USA

Gordon Instruments
Nashville, TN, USA

Gotham Audio Cables
Regensdorf, Switzerland

Grace Design
Boulder, CO, USA

Great River Electronics Inc.
Inver Grove Heights, MN, USA

Groove Tubes
San Fernando, CA, USA

Group One Ltd.
Farmingdale, NY, USA

H.E.A.R. — Hearing Education & Awareness for Rockers
San Francisco, CA, USA

Hacousto International / Sonic Systems, Inc.
Stratford, CT, USA

Harrison by GLW, Inc.
Laverne, TN, USA

Hear Technologies
Huntsville, AL, USA

***HHB Communications USA**
Simi Valley, CA, USA

Hosa Technology, Inc.
Buena Park, CA, USA

House Ear Research Institute
Los Angeles, CA, USA

Ibiquity Digital Corp.
Warren, NJ, USA

ILIO Entertainments
Malibu, CA, USA

Imas Publishing / Audio Media / Pro Audio Review
Falls Church, VA, USA

Independent Audio
Portland, ME, USA

Industrial Acoustics Co.
Bronx, NY, USA

***Innova-SON**
Old Lyme, CT, USA

Inter-M / Algorithmix (European Office, Germany)
Karisruhe, Germany

International DJ Expo / The Club Show
Port Washington, NY, USA

IZ Technology Corp.
Burnaby, B. C., Canada

***JBL Professional**
Northridge, CA, USA

***Jensen Transformers**
Van Nuys, CA, USA

JLCooper Electronics
El Segundo, CA, USA

The John Hardy Company
Evanston, IL, USA

Josephson Engineering
Santa Cruz, CA, USA

JRF Magnetic Sciences
Greendell, NJ, USA

***Klark-Teknik**
Burnsville, MN, USA



Klippel GmbH
Dresden, Germany

KRK Systems, LLC
Hollywood, FL, USA

Kurzweil Music Systems
Lakewood, WA, USA

***L-Acoustics US**
Oxnard, CA, USA

Lavry Engineering (formerly dB Technologies)
Seattle, WA, USA

Lectrosonics, Inc.
Rio Rancho, NM, USA

Level Control Systems
Sierra Madre, CA, USA

Lexicon, Inc.
Bedford, MA, USA

Line 6
Agoura Hills, CA, USA

Linn Products Ltd.
Glasgow Scotland, UK

Listen, Inc.
Boston, MA, USA

Little Labs
Los Angeles, CA, USA

Live Sound! International
San Francisco, CA, USA

Logitek Electronic Systems, Inc.
Houston, TX, USA

Lundahl Transformers AB
Norrtaälje, Sweden

Lynx Studio Technology, Inc.
Newport Beach, CA, USA

M.-Audio Inc.
Arcadia, CA, USA

MAGMA
San Diego, CA, USA

Magtrax Surround
Portland, ME, USA

Manley Laboratories, Inc.
Chino, CA, USA

Marian
West Hollywood, CA, USA

Marquette Audio Labs / Mercury
Hayward, CA, USA

Marshall Electronics, Inc. (Mogami Cables)
El Segundo, CA, USA

Marshall Electronics, Inc. (MXL Microphones)
El Segundo, CA, USA

***Martin Audio**
Kitchener, Ontario, Canada

Martinsound, Inc.
Alhambra, CA, USA

Master Recording Supply
Newport Beach, CA, USA

MBHO — Haun Microphones Germany
Brooklyn, NY, USA

MC2 Audio Ltd./Group One
Farmingdale, NY, USA

McCauley Sound, Inc.
Puyallup, WA, USA

Mercury Recording Equipment Co.
Hayward, CA, USA

Merging Technologies
Northbrook, IL, USA

Metric Halo Distribution, Inc.
Castle Point, NY, USA

Meyer Sound Laboratories, Inc.
Berkeley, CA, USA

Microboards Technology, LLC
Chanhassen, MN, USA

Microtech Gefell GmbH
Thuringia, Germany

Midas
Burnsville, MN, USA

Midiman
Arcadia, CA, USA

Millennia Media, Inc.
Placerville, CA, USA

Miller & Kreisel Sound Corporation
Chatsworth, CA, USA

Mix
Emeryville, CA, USA

Mixed Logic Studio Electronics
Brook Park, OH, USA

Motorola, Inc.
Austin, TX, USA

mSoft Inc.
Woodland Hills, CA, USA

The Museum of Sound Recording
New York, NY, USA

Music And More (MAM)
West Hollywood, CA, USA

Music Connection Magazine
Studio City, CA, USA

Music Maker Publications / Recording Magazine
Boulder, CO, USA

Mytek Digital
New York, NY, USA

Nagra USA, Inc.
Nashville, TN, USA

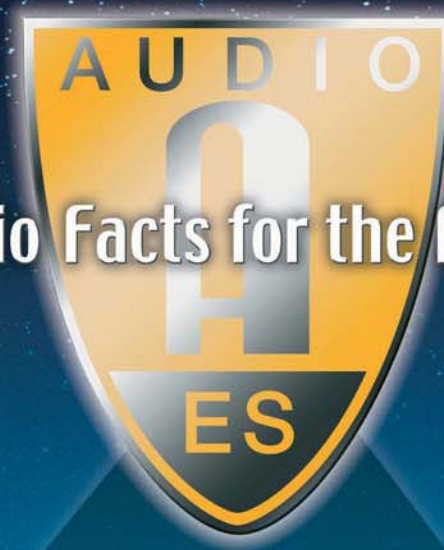


A Sound Investment

114th CONVENTION

OF THE

AUDIO ENGINEERING SOCIETY



Audio Facts for the Future

Amsterdam
The Netherlands

RAI Congress Center

22-25 March 2003

www.aes.org

114th@aes.org

National Academy of Recording Arts and Science

Santa Monica, CA, USA
Native Instruments GmbH
 Berlin, Germany

***Georg Neumann**
 Old Lyme, CT, USA

***Neutrik USA, Inc.**
 Lakewood, NJ, USA

Neutrik Test Instrument (NTI)
 St. Laurent, QC, Canada

***New Transducers Ltd. (NXT)**
 Huntingdon, Cambs., UK

Nexo USA
 San Rafael, CA, USA

NHT Pro
 Benicia, CA, USA

Noren Products, Inc.
 Menlo Park, CA, USA

Norris-Whitney Communications
 St. Catharines, Ontario, Canada

OKM Microphones
 Portland, ME, USA

Opticom
 Erlangen, Germany

Otari Corporation
 Canoga Park, CA, USA

Pacific Radio Electronics
 Los Angeles, CA, USA

Pearl Lab
 Portland, ME, USA

Peavey Electronics Corp.
 Meridian, MS, USA

Pelonis Sound and Acoustics, Inc.
 Gaviota, CA, USA

Pendulum Audio, Inc.
 Berkeley Heights, NJ, USA

Penn Fabrication, Inc.
 Oxnard, CA, USA

Penny and Giles Control Inc.
 Schaumburg, IL, USA

Philips Electronics — Super Audio CD
 Eindhoven, The Netherlands

Phoenix Audio International (UK)
 Stevenage, Herts., UK

Pilchner Schoustal Inc.
 Toronto, Ontario, Canada

Plugzilla
 New York, NY, USA

Plus24
 West Hollywood, CA, USA

PMC Monitors
 Welwyn Garden City, Herts., UK

PMI Audio Group
 Torrance, CA, USA

Post Magazine (AdvanStar Communications)
 New York, NY, USA

Powerphysics
 Newport Beach, CA, USA

Precision Laboratories
 Altamonta Springs, FL, USA

PreSonus Audio Electronics
 Baton Rouge, LA, USA

Prime LED
 San Francisco, CA, USA

***Primedia Business**
 New York, NY, USA

Primera Technology Inc.
 Plymouth, MN, USA

Princeton Digital
 Princeton, NJ, USA

Prism Media Products, Inc.
 Rockaway, NJ, USA

Professional Audio Design, Inc.
 Rockland, MA, USA

Propellerhead
 Arcadia, CA, USA

QSC Audio Products, Inc.
 Costa Mesa, CA, USA

RADIAN Audio Engineering, Inc.
 Orange, CA, USA

Radikal Technologies
 Teaneck, NJ, USA

RDL Radio Design Labs
 Carpinteria, CA, USA

***Rane Corporation**
 Mukilteo, WA, USA

RCS Enterprises
 White Plains, NY, USA

Renkus-Heinz, Inc.
 Foothill Ranch, CA, USA

Resolution (S2 Publications Ltd.)
 London, UK

RODE Microphones
 Torrance, CA, USA

Rohde & Schwarz GmbH & Co. KG
 Munich, Germany

Roland Corporation
 Los Angeles, CA, USA

Rolls Corporation
 Murray, UT, USA

Royer Labs
 Burbank, CA, USA

RTS
 Burnsville, MN, USA

Sabra-som Ltd. (K-IV Enterprises)
 Mahwah, NJ, USA

***SADIE, Inc.**
 Nashville, TN, USA

SAE Institute of Technology
 North Miami Beach, FL, USA

Sanken Microphones
 W. Hollywood, CA, USA

Schoeps/Redding Audio, Inc.
 Newtown, CT, USA

SE Electronics
 Rohnert Park, CA, USA

SEK'D
 North Hollywood, CA, USA

Seltron Components Ltd.
 Leadgate, Co Durham, Suffolk, UK

***Sennheiser Electronics Corp.**
 Old Lyme, CT, USA

The Serafine Collection
 Venice, CA, USA

***Shure Incorporated**
 Evanston, IL, USA

SLS Loudspeakers
 Springfield, MO, USA

Solid State Logic
 Begbroke, Oxford, UK

SommerCable America
 West Hollywood, CA, USA

Sonic Studio LLC
 Plymouth, MN, USA

Sonifex
 Portland, ME, USA

Sonomic
 New York, NY, USA

***Sony Electronics, Inc.**
 Park Ridge, NJ, USA

Sony Super Audio CD
 New York, NY, USA

***Sound Devices, LLC**
 Reedsburg, WI, USA

Sound Ideas
 Richmond Hill, Ontario, Canada

***Sound on Sound Magazine**
 Bar Hill, Cambridge, UK

Sound & Video Contractor Magazine
 Emeryville, CA, USA

***Soundcraft**
 Nashville, TN, USA

Soundelux — Hollywood Edge
 Hollywood, CA, USA

Soundelux Microphones
 Hollywood, CA, USA

Soundfield Research / Transamerica Audio Group
 Las Vegas, NV, USA

SPARS
 Memphis, TN, USA

Speck Electronics
 Fallbrook, CA, USA

Spectrasonics
 Burbank, CA, USA

SPL Electronics GmbH
 Niederkruechten, Germany

***SRS Labs, Inc.**
 Santa Ana, CA, USA

Steinberg N. America
 Chatsworth, CA, USA

StorCase Technology
 Fountain Valley, CA, USA

***STUDER**
 Toronto, Ontario, Canada

Studio Network Solutions
 St. Louis, MO, USA

Studio Technologies, Inc.
 Skokie, IL, USA

Summit Audio, Inc.
 Watsonville, CA, USA

Sunrise E. & E. Inc.
 Diamond Bar, CA, USA

Swissonic America
 West Hollywood, CA, USA

Switchcraft, Inc.
 Chicago, IL, USA

Symetrix, Inc.
 Lynwood, WA, USA

Syntrillium Software Corp.
 Scottsdale, AZ, USA

TAD — Technical Audio Devices / Pioneer
 Long Beach, CA, USA

Tamura Corporation
 Tokyo, Japan

***Tannoy North America Inc.**
 Kitchener, Ontario, Canada

***TASCAM**
 Montebello, CA, USA

Taytrix, Inc.
 Jersey City, NJ, USA

TC Electronic Inc.
 Westlake Village, CA, USA

TC Helicon
 Westlake Village, CA, USA

TC Works
 Westlake Village, CA, USA

TECA Spa
 Bresua, Italy

Telefunken North America, LLC
 South Windsor, CT, USA

Telex Communications, Inc.
 Burnsville, MN, USA

TerraSonde
 Boulder, CO, USA

Sound & Communication Magazine
 Port Washington, NY, USA

Testa Communications
 Port Washington, NY, USA

Texas Instruments
 Dallas, TX, USA

***THAT Corporation**
 Milford, MA, USA

The Music & Sound Retailer
 Port Washington, NY, USA

THE — Taylor Hohendahl Engineering
 S. Woodstock, CT, USA

THX Ltd.
 San Rafael, CA, USA

Tonian Laboratories
 Burbank, CA, USA

ToteVision
 Seattle, WA, USA

Tour Guide / Session Guide Magazines
 Nashville, TN, USA

Triam Electronics/Quested Monitoring Systems/Shep Associates
 Waunakee, WI, USA

Trident Audio
 Hook Green, Meopham, Kent, UK

Tube Tech
 Westlake Village, CA, USA

Ultrason AG
 Walnut Creek, CA, USA

Under Cover
 New York, NY, USA



***United Entertainment Media**
 New York, NY, USA

Universal Audio
 Santa Cruz, CA, USA

***VidiPax**
 New York, NY, USA

Vienna Symphonic Library GmbH
 Wien, Austria

Vintage Studio Rentals
 North Hollywood, CA, USA

Vintech Audio
 Seffner, FL, USA

Virtual Mixing Company
 Daly City, CA, USA

Voyager Sound Inc.
 Weston, MA, USA

Wacom Technology
 Vancouver, WA, USA

Wave Distribution
 Ringwood, NJ, USA

Wave:Space, Inc.
 Foothill Ranch, CA, USA

Wave Arts, Inc.
 Arlington, MA, USA

Waves, Inc.
 Knoxville, TN, USA

***Wenger Corporation**
 Owatonna, MN, USA

West Penn Wire / CDT
 Washington, PA, USA

Westlake Audio
 Newbury Park, CA, USA

Whirlwind
 Rochester, NY, USA

WhisperRoom, Inc.
 Morristown, TN, USA

Wireworks Corporation
 Hillside, NJ, USA

Wohler Technologies, Inc.
 S. San Francisco, CA, USA

World Link Digital
 Burbank, CA, USA

X-Vision Audio US Ltd.
 Youngstown, OH, USA

XTA Electronics / Group One Ltd.
 Farmingdale, NY, USA

***Yamaha Corporation of America**
 Buena Park, CA, USA

Yamaha mLAN Licensing Office
 Buena Park, CA, USA

Z-Systems, Inc.
 Gainesville, FL, USA

Zaxcom Audio
 Midland Park, NJ, USA

AES 113th Convention Program

2002 October 5–8

Los Angeles Convention Center

Los Angeles, California, USA

**Session A Saturday, October 5 9:00 am–11:30 am
Room 404AB**

TRANSDUCERS, PART 1

Chair: **Marshall Buck**, Gibson Labs, Redondo Beach, CA, USA; Psychotechnology Inc., Los Angeles, CA, USA

9:00 am

A-1 Which Loudspeaker Parameters Are Important to Create the Illusion of a Live Performance in the Living Room?—*Siegfried Linkwitz*, Linkwitz Lab, Corte Madera, CA, USA

The preference in loudspeaker product design is for a small size, while preserving maximum low-frequency extension and output volume. If the goal is to create a realistic reproduction of a live event, then certain speaker parameters must be adequately controlled, such as volume displacement, intermodulation distortion, stored energy, and off-axis frequency response. Components must be carefully selected for low distortion performance. Parameters like phase linearity and cabinet diffraction are sometimes overrated. Multichannel loudspeaker setups require propagation delay correction and bass management if not all of the loudspeakers cover the full frequency range. These issues are reviewed at the advent of high resolution surround sound. The new technology can only fulfill its promise and expand into more than a niche market if capable loudspeakers are widely available.

Convention Paper 5637

9:30 am

A-2 Characterizing the Amplitude Response of Loudspeaker Systems—*Allan Devantier*, Harman International Industries, Inc., Northridge, CA, USA

The amplitude response of a loudspeaker system is characterized by a series of spatially averaged measurements. The proposed approach recognizes that the listener hears three acoustical events in a typical domestic environment: the direct sound, the early arrivals, and the reverberant sound field. A survey of fifteen domestic multichannel installations was used to determine the typical angle of the direct sound and the early arrivals. The reflected sound that arrives at the listener after encountering only one room boundary is used to approximate the

early arrivals, and the total sound power is used to approximate the reverberant sound field. Two unique directivity indices are also defined, and the in-room response of the loudspeaker is predicted from anechoic data.

Convention Paper 5638

10:00 am

A-3 Graphing, Interpretation, and Comparison of Results of Loudspeaker Nonlinearity Measurement—*Alexander Voishvillo, Alex Terekhov, Gene Czerwinski, Sergei Alexandrov*, Cerwin-Vega Inc., Simi Valley, CA, USA

Harmonic distortion and THD do not convey sufficient information about nonlinearity in loudspeakers and horn drivers. Multitone stimulus and Gaussian noise produce more informative nonlinear response. Reaction to Gaussian noise can be transformed into a coherence or incoherence function. They provide information about nonlinearity in the form of easy-to-grasp frequency-dependent curves. Alternatively, the results of multitone measurement are difficult to interpret, compare, and overlay. A new method of depicting the results of multitone measurements has been developed. The distortion products are averaged in a moving frequency window. The result of the measurement is a single, continuous, frequency-dependent curve that takes into account the level of distortion products and their density. The curves can be easily overlaid and compared. Future development of a new method may lead to correlation between the level of distortion curves and the audibility of nonlinear distortion.

Convention Paper 5639

10:30 am

A-4 The Effects of Voice-Coil Axial Rest Position on Amplitude Modulation Distortion in Loudspeakers—*Ryan J. Mihelich*, Harman/Becker Automotive Systems, Martinsville, IN, USA

The magnetic field in the air gap of a conventional loudspeaker motor is often an asymmetric nonlinear function of axial position. Placement of the voice-coil into this asymmetrical field yields an asymmetric nonlinear force-factor, Bl , which is a primary cause of amplitude modulation distortion in loudspeakers. Adjustment of the rest position of the voice-coil in this field can alter the nature of this modulation distortion. Common practice is to nominally set the voice-coil at the geometric center of

the gap or at the position generating maximum BL. A time-domain nonlinear simulator has been used to investigate effects of voice-coil placement in an asymmetric flux field on amplitude modulation distortion.

Convention Paper 5640

11:00 am

A-5 Nonlinearity in Horn Drivers—Where the Distortion Comes From?—*Alexander Voishvillo*, Cerwin-Vega Inc., Simi Valley, CA, USA

Compared to other components of professional sound systems (omitting free propagation distortion), horn drivers have the worst nonlinear distortion. Some of the driver's distortions can be mitigated by proper mechanical measures. However, distortions caused by nonlinear air compression and propagation are inherent to any horn driver. In this paper the comparison of nonlinear distortions caused by different sources is carried through measurements and modeling. The new dynamic model of the compression driver is based on the system of nonlinear differential and algebraic equations. Complex impedance of an arbitrary horn is considered by turning the impedance into a system of differential equations describing the pressure and velocity at the horn's throat. The comparison is carried out using harmonic distortion and the reaction to multitone stimulus.

Convention Paper 5641

Session B Saturday, October 5 9:00 am–11:00 am Room 406AB

SIGNAL PROCESSING, PART 1

Chair: **Rob Maher**, Montana State University, Bozeman, MT, USA

9:00 am

B-1 A Talker-Tracking Microphone Array for Teleconferencing Systems—*Kazunori Kobayashi*, *Ken'ichi Furuya*, *Akitoshi Kataoka*, NTT, Musashino-shi, Tokyo, Japan

We propose a beamforming method that is applicable to near sound fields where a filter-and-sum microphone array maintains better quality for the target sound than the conventional delay-and-sum array. We also describe a real-time implementation that includes steering of the beam to detected talker locations. With the use of a microphone array, our system also cuts levels of noise to achieve high-quality sound acquisition. Furthermore, it allows the talker to be in any position. Computer simulation and experiments show that our method is effective in teleconferencing systems.

Convention Paper 5642

9:30 am

B-2 An Alternative Model for Sound Signals Encountered in Reverberant Environments; Robust Maximum Likelihood Localization and Parameter Estimation Based on a Sub-Gaussian Model—*Panayiotis G. Georgiou*, *Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

In this paper we investigate an alternative to the Gaussian density for modeling signals encountered in audio environments. The observation that sound signals are impulsive in

nature, combined with the reverberation effects commonly encountered in audio, motivates the use of the sub-Gaussian density. The new sub-Gaussian statistical model and the separable solution of its maximum likelihood estimator are derived. These are used in an array scenario to demonstrate with both simulations and two different microphone arrays the achievable performance gains. The simulations exhibit the robustness of the sub-Gaussian-based method while the real world experiments reveal a significant performance gain, supporting the claim that the sub-Gaussian model is better suited for sound signals.

Convention Paper 5643

10:00 am

B-3 Imperceptible Echo for Robust Audio Watermarking—*Hyeon-O Oh*¹, *Jong Won Seok*², *Jin Woo Hong*², *Dae-Hee Youn*¹

¹Yonsei University, Seoul, Korea

²Electronics & Telecommunications Research Institute (ETRI), Daejeon, Korea

In echo watermarking, the effort to improve robustness often conflicts with the requirement of imperceptibility. There have been inherent trade-offs in general audio watermarking techniques. In this paper we challenge the development of imperceptible but detectable echo kernels being directly embedded into the high-quality audio signal. Mathematical and perceptual characteristics of echo kernels are analyzed in a frequency domain. Finally, we can obtain a greater flat frequency response in perceptually significant bands by combining closely located positive and negative echoes. The proposed echo makes it possible to improve the robustness of an echo watermark without breaking the imperceptibility.

(Paper not presented at convention, but Convention Paper 5644 is available.)

10:30 am

B-4 New High Data Rate Audio Watermarking Based on SCS (Scalar Costa Scheme)—*Frank Siebenhaar*¹, *Christian Neubauer*¹, *Robert Baum*², *Jürgen Herre*¹

¹Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

²Friedrich-Alexander University, Erlangen, Germany

Presently, distribution of audio material is no longer limited to physical media. Instead, distribution via the Internet is of increasing importance. In order to attach additional information to the audio content, either for forensic or digital rights management purposes or for annotation purposes, watermarking is a promising technique since it is independent of the audio format and transmission technology.

State-of-the-art spread spectrum watermarking systems can offer high robustness against unintentional and intentional signal modifications. However, their data rate is typically comparatively low, often below 100 bit/s. This paper describes the adaptation of a new watermarking scheme called Scalar Costa Scheme (SCS), which is based on dithered quantization of audio signals. In order to fulfill the demands of high quality audio signal processing, modifications of the basic SCS, such as the introduction of a psychoacoustic model and new algorithms to determine quantization intervals, are required. Simulation figures and results of a sample implementation, which show the potential of this new watermarking scheme, are presented in this paper along with a short theoretical introduction to the SCS watermarking scheme.

Convention Paper 5645



**Workshop 1 Saturday, October 5 9:00 am–11:00 am
Room 408A**

STEREO AND SURROUND MICROPHONE TECHNIQUES (TUTORIAL)

Chair: **Geoff Martin**, McGill University, Montreal,
Quebec, Canada, and Bang & Olufsen a/s,
Struer, Denmark

Presenters: *Michael Bishop*, Telarc International
Doug Botnick, Independent Engineer/Producer
John Eargle, JME/JBL Professional
Richard King, Sony Music Studios
Mick Sawaguchi, NHK Broadcasting Center,
Programme Production Operations, Sound Division

This tutorial workshop includes a panel of leading industry professionals from the areas of pop, classical and film music, as well as radio drama. Issues to be discussed include the characteristics of various microphone configurations as well as upward-and-downward compatibility considerations. This workshop, which has been organized by the AES Education Committee, will be of benefit to audio engineers of all backgrounds, including students.

11:10 am

Technical Committee Meeting on Microphones

Special Event WHEN VINYL RULED!

Saturday, October 5	12:00 noon–6:00 pm
Sunday, October 6	12:00 noon–6:00 pm
Monday, October 7	10:00 am–6:00 pm
Tuesday, October 8	10:00 am–4:00 pm

Room 308A

Organizers: **Wes Dooley**, Audio Engineering
Associates, Pasadena, CA, USA
Dale Manquen, Consultant, Thousand
Oaks, CA, USA

Vinyl records ushered in an age of consumer high-fidelity. Magnetic tape recording was the technology that made practical the production of long-playing records. Assembling a component high-fidelity system became a widespread hobby for many of us, and for decades fueled the development of loudspeakers, electronics, and microphones. Listening to recorded music became a part of many people's daily life.

Our historical exhibit will offer attendees an overview of production technology during the age of vinyl and spotlight its relevance to current music production.

The analog audio presentation will take the visitor from Ampex's first machine, the 30 in/s, one-quarter-inch, full-track recordings to the leading edge technology of Michael Spitz's contemporary 30 in/s, one-inch, two-track mastering decks. Capitol Records will have the spotlight on Saturday afternoon when Carson Taylor talks about and plays examples of Tower and location recordings of the fifties, sixties, and seventies. Other key highlights will include Jim Webb's presentation of "12 Landmark Microphones" that made history; Kevin Gray's and Stan Ricker's presentation of vinyl disc mastering and record manufacturing; Paul McManus on the development and history of powered monitors from the fifties; and Ken Hirsch and David Gordon on the development of mixing console technology.

Saturday, Sunday, Monday, and Tuesday come up to Demo Room Row and hear good sounds, vintage and contemporary.

Be sure to check the daily schedule outside Room 308A for exact presentation times.

Special Event FREE HEARING SCREENINGS CO-SPONSORED BY THE AES AND HOUSE EAR INSTITUTE

Saturday, October 5	12:00 noon–6:00 pm
Sunday, October 6	10:00 am–6:00 pm
Monday, October 7	10:00 am–6:00 pm
Tuesday, October 8	10:00 am–4:00 pm

Booth 1881

Attendees are invited to take advantage of a free hearing screening co-sponsored by the AES and House Ear Institute. Four people can be screened simultaneously in the mobile audiological screening unit located on the exhibit floor. A daily sign-up sheet at the unit will allow individuals to reserve a screening time for that day. This hearing screening service has been developed in response to a growing interest in hearing conservation and to heighten awareness of the need for hearing protection and the safe management of sound.

Special Event SONGWRITERS SHOWCASE: "IT'S ALL ABOUT THE SONG"

October 5–October 7	1:00 pm–8:00 pm
October 8	12:00 noon–3:00 pm

South Hall Lobby

Produced by: **Claudia Koal**

The Audio Engineering Society will present the Songwriters Showcase, "It's All About the Song." This event marks the third time the AES will incorporate live original music into its convention schedule. The Songwriters Showcase is a means of acknowledging the song as a key element in the motivation for technological advancements in audio. The roster of performers includes international composers and songwriters. The event features a wide spectrum of musical genres. Members of performing rights organizations ASCAP, BMI, and SESAC will perform original works at daily performances. A complete schedule of performers will be posted in the registration, press, and exhibition areas throughout the four-day convention. Enjoy special performances during the AES Mixers each evening.

Education Event STUDENT DELEGATE ASSEMBLY 1

Saturday, October 5, 1:00 pm–2:30 pm
Room 402A

Chair: **Scott Cannon**, Stanford University Student
Section, Stanford, CA, USA

Vice Chair: **Dell Harris**, Hampton University Student
Section, Hampton, VA, USA

All students and educators are invited to participate in a discussion of opportunities in the audio field and issues of importance to audio education. A descriptive overview of conference events for students will also be given. This opening meeting of the Student Delegate Assembly will introduce the candidates for the coming year's election for chair and vice chair of the North and Latin America Regions. Each AES regional vice president may present two candidates for the election to be held at the closing meeting of the SDA. Election results and

Recording Competition and Poster Awards will be given at the Student Delegate Assembly 2 on Tuesday, October 8, at 10:00 am, in Room 402A.

**Session C Saturday, October 5 2:00 pm–6:00 pm
Room 404AB**

TRANSDUCERS, PART 2

Chair: **Steve Hutt**, Harman/Becker Automotive Systems, Martinsville, IN, USA

2:00 pm

C-1 The Bidirectional Microphone: A Forgotten Patriarch—Ron Streicher¹, Wes Dooley²

¹Pacific Audio-Visual Enterprises, Pasadena, CA, USA

²Audio Engineering Associates, Pasadena, CA, USA

Despite being one of the progenitors of all modern microphones and recording techniques, the bidirectional pattern is still not very well understood. Its proper and effective use remains somewhat of a mystery to many recording and sound reinforcement engineers. In this paper the bidirectional microphone is examined from historical, technical, and operational perspectives. We review how it developed and exists as a fundamental element of almost all other single-order microphone patterns. In the course of describing how this unique pattern responds to sound waves arriving from different angles of incidence, we show that it very often can be successfully employed where other more commonly-used microphones cannot.

Convention Paper 5646

2:30 pm

C-2 Gaussian Mixture Model-Based Methods for Virtual Microphone Signal Synthesis—

Athanasios Mouchtaris, Shrikanth S. Narayanan, Chris Kyriakakis, University of Southern California, Los Angeles, CA, USA

Multichannel audio can immerse a group of listeners in a seamless aural environment. However, several issues must be addressed such as the excessive transmission requirements of multichannel audio, as well as the fact that to date only a handful of music recordings have been made with multiple channels. Previously, we proposed a system capable of synthesizing the multiple channels of a virtual multichannel recording from a smaller set of reference recordings. In this paper these methods are extended to provide a more general coverage of the problem. The emphasis here is on time-varying filtering techniques that can be used to enhance particular instruments in the recording, which is desired in order to simulate virtual microphones in several locations close to and around the sound source.

Convention Paper 5647

3:00 pm

C-3 Driver Directivity Control by Sound Redistribution—Jan Abildgaard Pedersen, Gert Munch, Bang & Olufsen a/s, Struer, Denmark

The directivity of a single loudspeaker driver is controlled by adding an acoustic reflector to an ordinary driver. The driver radiates upward and the sound is redis-

tributed by being reflected off the acoustic reflector. The shape of the acoustic reflector is nontrivial and yields an interesting and useful directivity both in the vertical and horizontal plane. Two-dimensional FEM simulations and 3-D BEM simulations are compared to free field measurements performed on a loudspeaker using the acoustic reflector. The resulting directivity is related to results of previously reported psychoacoustic experiments.

Convention Paper 5648

3:30 pm

C-4 Pressure Response of Line Sources—Mark S. Ureda, JBL Professional, Northridge, CA, USA

The on-axis pressure response of a vertical line source is known to decrease at 3 dB per doubling of distance in the near field and at 6 dB in the far field. This paper shows that the conventional mathematics used to achieve this result understates the distance at which the -3 dB to -6 dB transition occurs. An examination of the pressure field of a line source reveals that the near field extends to a greater distance at positions laterally displaced from the centerline, normal to the source. The paper introduces the endpoint convention for the pressure response and compares the on-axis response of straight and hybrid line sources.

Convention Paper 5649

4:00 pm

C-5 High-Frequency Components for High-Output Articulated Line Arrays—Doug Button, JBL Professional, Northridge, CA, USA

The narrow vertical pattern achieved by line arrays has prompted much interest in the method for many forms of sound reinforcement in recent years. The live sound segment of the audio community has used horns and compression drivers for sound reinforcement for several decades. To adopt a line array philosophy to meet the demands of high level sound reinforcement, requires an approach that allows for the creation of a line source from the output of compression drivers. Additionally, it is desired that the line array take on different vertical patterns dependent upon use. This requires the solution to allow for the array to be articulated. Outlined in this paper is a waveguide/compression driver combination that is compact and simple in approach and highly suited for articulated arrays.

Convention Paper 5650

4:30 pm

C-6 High-Efficiency Direct-Radiator Loudspeaker Systems—John Vanderkooy¹, Paul M. Boers²

¹University of Waterloo, Waterloo, Ontario, Canada

²Philips Research Labs, Eindhoven, The Netherlands

Direct-radiator loudspeakers become more efficient as the total magnetic flux is increased, but the accompanying equalization and amplifier modify the gains thus made. We study the combination of an efficient high-BI driver with several amplifier types, including a highly efficient class-D amplifier. Comparison is made of a typical simulated driver, excited with a few different amplifier types, using various audio signals. The comparison is quite striking as the BI value of the driver increases, significantly favoring the class-D amplifier.

Convention Paper 5651

5:00 pm

C-7 Audio Application of the Parametric Array—Implementation through a Numerical Model—Wontak Kim¹, Victor W. Sparrow²

¹Bose Corporation, Framingham, MA, USA

²Pennsylvania State University, University Park, PA, USA

Implementing the parametric array for audio applications is examined through numerical modeling and analytical approximation. The analytical solution of the nonlinear wave equation is used to provide guidelines on the design parameters of the parametric array. The solution relates the source size, input pressure level, and the carrier frequency to the audible signal response including the output level, beam width, and length of the interaction region. A time domain finite difference code that accurately solves the KZK nonlinear parabolic wave equation is used to predict the response of the parametric array. The accuracy of the numerical model is established by a simple parametric array experiment. In considering the implementation issues for audio applications of the parametric array, the emphasis is given to the poor frequency response and the harmonic distortion. Signal processing techniques to improve the frequency response and the harmonic distortion are suggested and implemented through the numerical simulation.

Convention Paper 5652

5:30 pm

C-8 Implementation of Straight-Line and Flat-Panel Constant Beamwidth Transducer (CBT) Loudspeaker Arrays Using Signal Delays—D. B. (Don) Keele, Jr., Harman/Becker Automotive Systems, Martinsville, IN, USA

Conventional CBT arrays require a driver configuration that conforms to either a spherical cap-curved surface or a circular arc. CBT arrays can also be implemented in flat-panel or straight-line array configurations using signal delays and Legendre function shading of the driver amplitudes. Conventional CBT arrays do not require any signal processing except for simple frequency-independent shifts in loudspeaker level. However, the signal processing for the delay-derived CBT configurations, although more complex, is still frequency independent. This is in contrast with conventional constant-beamwidth flat-panel and straight-line designs which require strongly frequency-dependent signal processing. Additionally, the power response roll-off of the delay-derived CBT arrays is one-half the roll-off rate of the conventional designs, i.e., 3- or 6-dB/octave (line or flat) for the CBT array versus 6- or 12-dB/octave for the conventional designs. Delay-derived straight-line CBT arrays also provide superior horizontal off-axis response because they do not exhibit the ± 90 degree right-left off-axis sound pressure buildup or bulge as compared to conventional circular-arc CBT arrays. In comparison to conventional CBT arrays, the two main disadvantages of delay-derived straight-line or flat-panel CBT arrays are 1) the more complicated processing required, which includes multiple power amplifiers and delay elements; and 2) the widening of the polar response at extreme off-axis angles particularly for arrays that provide wide coverage with beamwidths greater than 60 degrees. This paper illustrates its findings using numerical simulation and modeling.

Convention Paper 5653

**Session D Saturday, October 5 2:00 pm–4:30 pm
Room 406AB**

SIGNAL PROCESSING, PART 2

Chair: **Robert Bristow-Johnson**, Wave Mechanics, Burlington, VT, USA

2:00 pm

D-1 Automatic Design of Sound Synthesis Techniques by Means of Genetic Programming—Ricardo A. Garcia, Chaoticom, Hampton Falls, NH, USA

Design of sound synthesis techniques (SST) is a difficult problem. It is usually assumed that it requires human ingenuity to find a suitable solution. Many of the SSTs commonly used are the fruit of experimentation and long refinement processes. An automated approach for design of SSTs is proposed. The problem is stated as a search in the multidimensional SST space. It uses genetic programming (GP) to suggest valid functional forms and standard optimization techniques to fit their internal parameters. A psychoacoustic auditory model is used to compute the perceptual distance between the target and test sounds. The developed AGeSS (automatic generator of sound synthesizers) system is introduced, and a simple example of the evolved SSTs is shown.

Convention Paper 5654

2:30 pm

D-2 A Consumer-Adjustable Dynamic Range Control System—Keith A. McMillen, Octiv, Inc., Berkeley, CA, USA

Advances in technology have afforded listeners an available dynamic range in excess of 120 dB. While impressive in proper concert halls and listening rooms, large dynamic ranges are not always realistic for all environments and musical styles. This paper describes a practical multiband dynamics processor software object that can reside in low cost consumer products and allow the user to adjust dynamic range to fit his or her taste and listening environment.

Convention Paper 5655

3:00 pm

D-3 A Simple, Efficient Algorithm for Reduction of Hiss Amplification Under High Dynamics Compression—Guillermo Garcia, CCRMA, Stanford, CA, USA

We present a very simple, effective, and computationally efficient algorithm to reduce the typical hiss amplification (or breathing) artifact of dynamics compressors working under high compression ratios. The algorithm works in the time domain, is very easy to implement, has a very low computational cost, and requires little program memory, therefore being of special interest for consumer-audio applications.

(Paper not presented at convention, but Convention Paper 5656 is available.)

3:30 pm

D-4 Adaptive Predistortion Filter for Linearization of Digital PWM Power Amplifier Using Neural Networks—Minki Yang¹, Jong-Hoon Oh²

¹Pulsus Technologies, Inc., Seoul, Korea

²Pohang University of Science and Technology, Pohang, Korea

The paper presents a method to compensate for nonlinear distortion of digital pulse width modulation (PWM) power amplifiers by prefiltering the input signals using artificial neural networks. We first construct a model of the digital amplifier using artificial neural networks. Using this model, the artificial neural network model of a predistortion filter is trained such that the combined sys-

tem, the digital amplifier, and predistortion filter produces an output that is linearly proportional to the input. The simulation results show that the artificial neural networks of the predistortion filter can effectively correct the nonlinear distortion of the digital amplifier.
Convention Paper 5657

4:00 pm

D-5 Objective Measures of the Quality of Speech Transmission in a Real Mobile Network—Measuring, Estimate, and Prediction Method—

Dragana Sagovnovic, Telekom Srbija, Belgrade, Yugoslavia

Contemporary means of communication (e.g., mobile telephony) have brought new limitations that telecom operators have to take into consideration. One of them is the fact that the types of deterioration of speech quality, perceived in mobile telephony, are different from the degradations noted in fixed telephony. This paper discusses, during the process of estimating the quality of speech transmission, the comparative (objective) methods for a formed database of degradations of a real mobile communications system. The consideration of the results of the objective-method tests is based on the development of a new objective method of speech quality. The results gathered during the comparison tests have been displayed and interpreted for different types of Serbian vowels: front vowels, /e/; mid vowels, /a/; and back vowels, /u/.

(Paper not presented at convention, but Convention Paper 5658 is available.)

4:40 pm

Technical Committee Meeting on Signal Processing

**Workshop 2 Saturday, October 5 2:00 pm–5:00 pm
Room 408A**

NEW MEDIA FOR MUSIC—ADAPTIVE RESPONSE TO TECHNOLOGY

Chair: **Dave Davis**, QCA Mastering, Cincinnati, OH, USA

Panelists: *Bob Ludwig*, Gateway Mastering and DVD, Portland, ME
Bill McQuay, Radio Expeditions: National Public Radio and the National Geographic Society, Baltimore, MD
Bobby Owsinski, Surround Associates, Studio City, CA
Mike Sokol, JMS Productions & Fits and Starts Productions, Hagerstown, MD

In 2002 the audio industry is at a crossroad. The CD is seemingly vulnerable to piracy by file trading and easy replication. Studios and engineers are seeing their markets shrink and rates decline. New formats are available to consumers, but there are no budgets for creating content. While technology gets the blame, any solution requires new technologies to succeed. We will examine product-oriented solutions that are available today and how to create new kinds of products that expand both artistic expression and markets.

This workshop focuses on using existing technology and adapting existing models to expand the music industry with new products. It is not an esoteric romp through dot.bomb schemes or a complicated strategy requiring consumers to replace their stereo equipment and record collections. The current crisis requires us to adapt and evolve incrementally, so this workshop is about identifying successful models and

applying them to what we collectively do. Many of the technologies and models that can save us have already matured and found their way into the homes of music fans. 2002 is our last, best hope to restore the health and vitality to the music industry through our products, rather than legislation or litigation.

Dave Davis (QCA Mastering/UltraInteractive) will introduce the qualities of new media and how they can be used to add value to music products. He will explain why new media strategies are essential to our industry's adaptation to current cultural and technological challenges and demonstrate a number of existing products that provide fans with a richer experience. Finally, he will present some works that were conceived for rich media environments and require a DVD or computer to experience; for many new artists these technologies not only address economic challenges, but broaden their vision and expand their palette of tools.

Bob Ludwig (Gateway Mastering and DVD) will discuss the role of the mastering studio regarding SACD, DVD-V, DVD-A and enhanced CD. He will discuss how Gateway uses its ftp site for projects and various pro and consumer file transfer technologies (Liquid Audio, DigiProNet, Rocket etc) and explain how an integrated facility can support an artist's vision from soup to nuts in one the building.

Bill McQuay (Radio Expeditions: National Public Radio and the National Geographic Society) is going to discuss the National Geographic *Expeditions* series as a forward looking model to support rich, challenging production for programming in a radio format. The Radio Expeditions feature runs once a month on NPR's flagship news magazine Morning Edition. It lives on in that form for fans on NPR's web site, where it can be streamed at will. The material was expanded into a lecture/presentation series with surround sound for a live venue. A DVD is in the works, presenting an even richer version of the program. These additional modes of presentation collectively support the use of audio in a deeper, richer way than is traditionally possible on broadcast radio.

Bobby Owsinski (Surround AssociatesA) is deeply involved in the development and exploration of surround mixing and will address the role of the creative engineer in production for multiple mix targets, as well as what studios and tracking engineers can do to facilitate surround work downstream.

Mike Sokol (JMS Productions & Fits and Starts Productions) will address the opportunities created by surround for audio professionals. Mike will be discussing a specific project in radio drama that was staged live in surround, and is working to broadcast it to home theaters via Digital cable using Dolby Digital. This represents both a reborn market, and the beginnings of surround-driven audio-only content in a digital broadcast environment. Mike will also discuss other alternative 5.1 audio markets, such as Powerpoint 5.1, as well as ways to do large music concerts with 5.1 surround elements. All of these markets offer additional production and revenue streams for studios trying to justify 5.1 production gear.

5:10 pm

Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

**Workshop 3 Saturday, October 5 2:00 pm–5:00 pm
Room 408B**

WHAT AUDIO ENGINEERS SHOULD KNOW ABOUT HUMAN SOUND PERCEPTION

Presenters: *Durand Begault*, NASA Ames Research Center, Mountain View, CA, USA
William Martens, University of Aizu, Aizuwakamatsu-shi, Japan

Audio engineering professionals are intimately familiar with the art of producing acoustical phenomena to achieve specific psychoacoustic goals for applications such as music, multimedia, speech reinforcement, etc. However, they may be less fa- ➔

miliar with the underlying mechanisms of hearing and perception that influence their everyday decisions.

This tutorial will review fundamental concepts of human sound perception, first from a monaural hearing perspective and then from a spatial hearing perspective. Topics include pitch, loudness, timbre, sensitivity to phase; temporal resolution and temporal integration; masking; spatial perception; perceptual effects of rooms; and differences between headphone and loudspeaker reproduction. Questions and comments from the audience will be encouraged.

5:10 pm

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Education Event

MENTORING PANEL: STEPS TO THE FUTURE—EFFECTIVELY USING MENTORS TO HELP BUILD A CAREER

Saturday, October 5, 2:30 pm–4:00 pm
Room 402A

Chair: **Theresa Leonard**, The Banff Centre, Banff, Alberta, Canada

Panelists: *Jim Anderson*, Independent Recording Engineer, AES Regional VP
Edwin Dolinsky, Electronic Arts
Bill Dooley, SPARS, The Village Recorder
Lynn Fuston, Skywalker Sound
Richard King, Sony Music
David Moulton, Moulton Laboratories
Julie Perez, NBC TV, Music Mixer, “Late Night with Conan O’Brien”

Thriving in the audio industry takes more than technical aptitude—it helps to have guidance from industry professionals. Learn about the benefits of mentor relationships and how to develop your own network of industry connections via mentoring. Students have a chance to learn firsthand from industry veterans by attending a mentoring panel and then signing up for one-on-one time with one of the distinguished participants.

Special Event

14TH ANNUAL GRAMMY® RECORDING SOUNDTABLE: 21ST CENTURY REALITIES

Saturday, October 5, 4:00 pm–6:00 pm
Room 411 Theater

Moderator: **Howard Massey**, On The Right Wavelength Consulting

Panelists: *Brian Garten*
Ken Jordan
George Massenbourg
Jack Joseph Puig
Elliot Scheiner
Al Schmitt

The National Academy of Recording Arts & Sciences, Inc. will present the 14th Annual GRAMMY® Recording SoundTable, moderated by Howard Massey. It's the 21st century: Recording budgets are shrinking, technologies are advancing, and public taste in music is changing. How does today's producer cope with these realities? Do 24 bits truly matter when most people are listening to music in MP3 format? Is there still room for analog in today's digital world? And, with entire hit records being made on laptop computers, is the professional recording studio becoming an endangered species? Don't miss the lively discussion as our panelists attack these topics head on.

Brian Garten is engineer for the Neptunes, and the man behind the console for two of 2002's biggest hits: Nelly's "Hot in Here" and "Dilemma," featuring Kelly of Destiny's Child. He has also recorded by Beyonce Knowles, Britney Spears, Mary J. Blige, Alicia Keys, Usher, No Doubt, Busta Rhymes, and Justin Timberlake.

Beatmaster Ken Jordan of the alternative-electronic duo The Crystal Method brings a variety of influences to their dance-based productions, merging hip-hop, rock, and breakbeats into a unique blend that have made them an integral part of club culture. Their debut LP, *Vegas*, was released to critical acclaim in 1997, and their latest album, *Tweekend*, was released in 2001. The two albums have sold over a million and a half copies. For the very latest on this duo, check out www.thecrystal-method.com.

George Massenbourg is known as an engineer's engineer. His pristine work with artists like Little Feat, Toto, Linda Ronstadt, Lyle Lovett, Emmylou Harris, Carly Simon, Mary Chapin Carpenter, and Aaron Neville constantly sets new standards in audio excellence, and he is a technical pioneer as well. Recipient of a Technical GRAMMY® Award for his development of innovative recording tools such as the parametric equalizer, Massenbourg is actively involved in surround sound mixing and the standardization of archiving materials.

Jack Joseph Puig is one of the most sought-after engineer/producers crafting music today. He owns an astonishing collection of vintage and modern recording equipment that he uses to create a unique engineering and production sound that merges the styles of the last four decades. You can hear his work on releases by Sheryl Crow, Vanessa Carlton, John Mayer, No Doubt, Counting Crows, Green Day, Goo Goo Dolls, Hole, Shelby Lynne, Weezer, Stone Temple Pilots, Tonic, Stevie Nicks, and Robbie Williams.

Engineer extraordinaire Elliot Scheiner has been nominated for more than a dozen GRAMMY® Awards and has won five of them, including the 2000 GRAMMY® Award for Best Engineered Album, Non-Classical. He has worked with a wide variety of multiplatinum artists, including Faith Hill, Beck, Steely Dan, Fleetwood Mac, The Eagles, Van Morrison, The Doobie Brothers, America, John Fogerty and Jimmy Buffett. He recently completed a surround sound remix of the classic Queen album *A Night at the Opera*, which won the 2002 DVD award for Best Audio.

There are few people truly deserving of the term legendary, but Al Schmitt is one of them. Not only has he won an astonishing nine GRAMMY® Awards, which include the 2001 Best Engineered Album, Non-Classical Award for Diana Krall's *The Look of Love*, he has also won two Latin GRAMMY® Awards. He has produced, engineered, and/or mixed more than 150 gold and platinum records for a diverse range of artists, including Frank Sinatra, Jefferson Airplane, Steely Dan, Barbra Streisand, Henry Mancini, and Duane Eddy.

Moderator Howard Massey is a noted industry consultant and veteran audio journalist. He has written for *Home Recording*, *Surround Professional*, *EQ*, *Musician*, *Billboard*, and many other publications, and is the author of eleven books, including *Behind the Glass*, a collection of interviews with the world's top record producers. Massey has also worked extensively as an audio engineer, producer, songwriter, and touring/session musician.

Special Event AES MIXER

October 5–October 7
South Hall Lobby

6:00 pm–8:00 pm

A series of informal Mixers will be held each evening in the South Lobby entrance to the Los Angeles Convention Center, to enable convention attendees to meet in a social atmosphere after the day's activities and catch up with friends and colleagues from the industry. There will be music featuring performances from the “Songwriter's Showcase,” a cash bar, and snacks.

SIGNAL PROCESSING, PART 3

Chair: **Jayant Datta**, Discrete Labs, Syracuse, NY, USA

9:00 am

E-1 Computationally Efficient Inversion of Mixed Phase Plants with IIR Filters—*Timoleon Papadopoulos, Philip A. Nelson*, University of Southampton, Southampton, UK

Inverse filtering in a single or in multiple channels arises as a problem in a number of applications in the areas of communications, active control, sound reproduction, and virtual acoustic imaging. In the single-channel case, when the plant $C(z^{-1})$ sought to be inverted has zeros outside the unit circle in the z -plane, an approximation to the inverse $1/C(z^{-1})$ can be realized with an FIR filter if an appropriate amount of modeling delay is introduced to the system. But the closer the zeros of $C(z^{-1})$ are to the unit circle (either inside or outside it), the longer the FIR inverse has to be, typically several tens of times longer than the plant. An off-line implementation utilizing a variant of the backward-in-time filtering technique usually associated with zero-phase FIR filtering is presented. This forms the basis on which a single-channel mixed phase plant can be inverted with an IIR filter of order roughly double than that of $C(z^{-1})$, thus decimating the processing time required for the inverse filtering computation.
Convention Paper 5659

9:30 am

E-2 Optimal Filter Partition for Efficient Convolution with Short Input/Output Delay—*Guillermo García*, CCRMA, Stanford, CA, USA

A new algorithm to find an optimal filter partition for efficient long convolution with low input/output delay is presented. For a specified input/output delay and filter length, our algorithm finds the nonuniform filter partition that minimizes computational cost of the convolution. We perform a detailed cost analysis of different block convolution schemes and show that our optimal-partition finder algorithm allows for significant performance improvement. Furthermore, when several long convolutions are computed in parallel and their outputs are mixed down (as is the case in multiple-source 3-D audio rendering), the algorithm finds an optimal partition (common to all channels) that allows for further performance optimization.
Convention Paper 5660

10:00 am

E-3 Filter Morphing—Topologies, Signals, and Sampling Rates—*Rob Clark*^{1,2}, *Emmanuel Ifeakor*², *Glenn Rogers*¹

¹Allen & Heath Limited, Penryn, Cornwall, UK
²University of Plymouth, Plymouth, UK

Digital filter morphing techniques exist to reduce audible transient distortion during filter frequency response change. However, such distortions are heavily dependent on signal content, frequency response settings, filter topology, interpolation scheme, and sampling rates. This paper presents an investigation into these issues, implementing various filter topologies using different

input stimuli and filter state change scenarios. The paper identifies the mechanisms causing these distortions, specifying worst case filter state change scenarios. The effects of existing interpolator schemes, finite word length, and system sampling rates on signal distortion are presented. The paper provides an understanding of filter state change, critical in the design of filter morphing algorithms.
Convention Paper 5661

10:30 am

E-4 Evaluation of Inverse Filtering Techniques for Room/Speaker Equalization—*Scott G. Norcross, Gilbert A. Soulodre, Michel C. Lavoie*, Communications Research Centre, Ottawa, Ontario, Canada

Inverse filtering has been proposed for numerous applications in audio and telecommunications, such as speaker equalization, virtual source creation, and room deconvolution. When an impulse response (IR) is at nonminimum phase, its corresponding inverse can produce artifacts that become distinctly audible. These artifacts produced by the inverse filtering can actually degrade the signal rather than improve it. The severity of these artifacts is affected by the characteristics of the filter and the method (time or frequency domain) used to compute its inverse. In this paper objective and subjective tests were conducted to investigate and highlight the potential limitations associated with several different inverse-filtering techniques. The subjective tests were conducted in compliance with the ITU-R MUSHRA method.
Convention Paper 5662

11:00 am

E-5 Using Subband Filters to Reduce the Complexity of Real-Time Signal Processing—*J. Michael Peterson, Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

Several high quality audio applications require the use of long finite impulse response (FIR) filters to model the acoustical properties of a room. Various structures for subband filtering are examined. These structures have the ability to divide long FIR filters into smaller FIR filters that are easier to use. Two structures will be discussed to process the signals in a real-time manner, time-convolution of spectrograms, and generalized filterbanks. Also filter estimation will be discussed.
Convention Paper 5663

11:30 am

E-6 Noise Shaping in Digital Test-Signal Generation—*Stanley P. Lipshitz*¹, *John Vanderkooy*¹, *Edward V. Semyonov*²

¹University of Waterloo, Waterloo, Ontario, Canada
²Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia

In an earlier paper we put forth an idea to use noise-shaping techniques in the generation of digital test signals. The previous paper proposed using noise shaping around an undithered quantizer to generate sinusoidal digital test signals with spectra having error nulls at the harmonics of the signal frequency, thus making digital distortion measurements of very great dynamic range possible. We extend this idea in this present paper in a number of ways. We show a) that dither is necessary in order to suppress spurious artifacts caused by the nonlinearity of an undithered noise shaper; b) that wider and deeper nulls at the harmonic frequencies can be achieved

by using higher-order noise-shaper designs; c) that IIR filter designs can moderate the increased noise power that accompanies an increased FIR filter order; and d) some other novel uses of noise shaping in digital signal generation.

Convention Paper 5664

12:10 pm

Technical Committee Meeting on Multichannel and Binaural Audio Technologies

**Session F Sunday, October 6 9:00 am–12:00 noon
Room 404ABI**

ROOM ACOUSTICS AND SOUND REINFORCEMENT

Chair: **Eric Benjamin**, Dolby Laboratories, San Francisco, CA, USA

9:00 am

F-1 Factors Affecting Accuracy of Loudspeaker Measurements for Computational Predictions— *Roger Schwenke*, Meyer Sound Laboratories, Berkeley, CA, USA

The delay of a signal from the input terminals of a loudspeaker amplifier to the output terminals of a microphone can be represented as two parts: one from the electrical input to acoustical transmission and an acoustical propagation delay from some point on the loudspeaker to the microphone. For computational models of mixtures of loudspeakers to be correct, these delays must be measured accurately. It will be shown that temperature differences as small as 1 degree Celsius between measurements of two models of loudspeakers can cause significant differences in the predicted sound field. Though sound speed is much less sensitive to changes in humidity, the difference between assuming a typical humidity and assuming zero humidity (which is the norm) can be significant.

Convention Paper 5665

9:30 am

F-2 Systems for Stereophonic Sound Reinforcement: Performance Criteria, Design Techniques, and Practical Examples— *Jim Brown*, Audio Systems Group, Inc., Chicago, IL, USA

Although stereo systems for large rooms were pioneered in a well documented work at Bell Labs in the 1930s, most modern practitioners appear to be ignorant of the most important aspect of that work as applied to modern sound reinforcement. This paper draws on the author's experience of over twenty years with both portable and permanent systems using two and three front-referenced channels. Design criteria and examples are presented to illustrate both good and bad design practices.

Convention Paper 5666

10:00 am

F-3 Cable Impedance and Digital Audio— *Stephen H. Lampen*, David A. DeSmidt, Belden Electronics Division, San Francisco, CA, USA

One of the key differences between cable designed for analog signals and cable designed for digital signals is the impedance of the cable. Why is impedance important for digital but not for analog? What effect do impedance

variations or mismatching have on digital signals? Can you use Category 5e or Category 6 computer cable to run digital audio? Can you use coaxial cable to carry digital audio? This paper addresses all these questions and also outlines the limitations of digital cable designs.

Convention Paper 5667

10:30 am

F-4 Limitations of Current Sound System Intelligibility Verification Techniques— *Peter Mapp*, Peter Mapp Associates, Colchester, Essex, UK

The role of emergency sound and voice alarm systems in life safety management has never been so important. However, to be effective, it is essential that such systems are adequately intelligible. Verification of system intelligibility is therefore assuming ever-greater importance. While a number of verification techniques are available, each is not without its drawbacks. The paper reviews the available methods and, using the results of new research, highlights areas of weakness of the current techniques.

Convention Paper 5668

11:00 am

F-5 Robustness of Multiple Listener Equalization with Magnitude Response Averaging— *Sunil Bharitkar*, Philip Hilmes, Chris Kyriakakis, University of Southern California, Los Angeles, CA, USA

Traditionally, room response equalization is performed to improve sound quality at a given listener. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through spatial averaging of magnitude responses at locations of interest. However, the performance of averaging-based equalization, at the listeners, may be affected when listener positions change. In this paper, we present a statistical approach to map variations in listener positions to a performance metric of equalization for magnitude response averaging. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on the distance of microphones from a source and the frequencies in the sound.

Convention Paper 5669

11:30 am

F-6 Coax and Digital Audio— *Stephen H. Lampen*, David A. DeSmidt, Belden Electronics Division, San Francisco, CA, USA

Coaxial cables have been used to run digital audio signals for many years, and have been added to the AES specifications (AES3-id). How is coax different from twisted pairs? What are the distance limitations? What trade-offs are made going from digital twisted pairs to digital coax? These questions are all answered in this paper including a discussion of baluns, which are used to convert from one format to the other.

Convention Paper 5670

12:10 pm

Technical Committee Meeting on Acoustics and Sound Reinforcement

**Workshop 4 Sunday, October 6 9:00 am–12:00 noon
Room 408B**

LARGE SIGNAL LOUDSPEAKER PARAMETER MEASUREMENT

Chair: **John Stewart**, Harman/Becker Automotive Systems, Martinsville, IN, USA

Panelists: *Doug Button*, JBL Professional, Northridge, CA, USA
David Clark, DLC Design, Wixom, MI, USA
Earl Geddes, GedLee Associates, Northville, MI, USA
Steve Hutt, Harman/Becker Automotive Systems, Martinsville, IN, USA
Wolfgang Klippel, Klippel, GmbH, Dresden, Germany
Steve Temme, Listen, Inc., Boston, MA, USA
Alexander Voishvillo, Cerwin-Vega Inc., Simi Valley, CA, USA

This workshop will present measurement methods used to characterize loudspeaker performance at large signal levels. A brief discussion of each technique—how it evolved, what benefit it brings, etc.—will be followed by a demonstration of the technique when possible. The discussion will include: high-power impedance measurement, X_{\max} , model parameter nonlinearities, flux modulation, harmonic distortion, maximum SPL, and power compression.

12:10 pm

Technical Committee Meeting on Audio for Telecommunication

**Workshop 5 Sunday, October 6 9:00 am–12:00 noon
Room 408A**

MIXING AND MASTERING IN MULTICHANNEL SURROUND (TUTORIAL)

Chair: **Michael Bishop**, Telarc International Corporation, Cleveland, OH, USA

Panelists: *John Eargle*, JME/JBL Professional
Frank Filipetti, Independent Engineer/Producer
Bob Ludwig, Gateway Mastering & DVD
George Massenburg, GML, LLC
Elliott Scheiner, Independent Engineer/Producer

This tutorial workshop includes a panel of leading industry professionals who will discuss various formats for mixing and mastering in surround. Topics to be covered include:

- Mixing and mastering for DTS/Dolby Digital Surround vs. discrete multichannel surround
- Mixing for 5.1 vs. 6.0, 7.1 or 10.2
- Multichannel mix recording formats in most common use and compatibility with mastering needs
- Mixing and mastering monitor loudspeaker setup/formats and compatibility with common home theater setups in the home
- Essential tools for mixing and mastering in multichannel surround
- Mix approaches and preferences: ambient surround vs. direct surround
- Challenges of mastering for DVD-A
- Challenges of mastering for SACD

Audio and music examples will be played. This workshop, which has been sponsored by the AES Education Committee, will be of benefit to audio engineers of all backgrounds, including students.

Regrettably, Dennis Purcel was scheduled to be a panelist in this workshop but has since passed away.

12:10 pm

Technical Committee Meeting on Transmission and Broadcasting

Special Event

PLATINUM PRODUCERS PANEL 1 PRODUCER, ENGINEER, STUDIO TECHNICIAN— BLURRING OF ROLES

Sunday, October 6, 12:30 pm–2:30 pm
Room 403A

Moderator: **Mitch Gallagher**, EQ Magazine

Panelists: *Rob Cavallo*
Mike Elizondo
Ron Fair
Ben Grosse
Taz Herzberg

It is no secret that the demands of creating audio and music in today's marketplace require a broad range of skills and flexibility. It is increasingly common for one person to handle what once were multiple dedicated jobs: producing, engineering, performing . . . and to do them all at the same time. We've gathered five of today's hottest producers and engineers; join as they discuss the tricks, techniques, mindset, and technologies that allow them to cover multiple roles at once—and how you can apply this information to your own situation.

Mitch Gallagher, the Editor in Chief of *EQ* magazine, began working in music professionally over 20 years ago. His background includes a degree in music, as well as graduate studies in electronic music composition and classical guitar. He is an author, journalist, teacher, touring and studio musician, recording engineer, project studio/multimedia production company owner, and award-winning composer. His first book, *Make Music Now!*, was recently released by Backbeat Books.

Mike Elizondo has shared production credit with Dr. Dre on the Rolling Stones "I Miss You" for the *Austin Powers* soundtrack. His songs and musicianship can be found on Eve's Grammy winning "Let Me Blow Ya Mind," Mary J. Blige's "Family Affair," Eminem's "My Dad's Gone Crazy." Elizondo a production company "Tone Down Productions" with its first release, *Rebel through Columbia* due in early 2003.

Ron Fair is a veteran record man who is a rare combination of producer, A&R executive, musician, arranger, and engineer. He signed and produced multi-Grammy winner Christina Aguilera and also signed Lit. He is currently President of A&M Records and produced the vocal performances of Christina Aguilera, Mya, Pink, and Lil' Kim.

Ben Grosse is a producer/engineer/mixer whose credits include Vertical Horizon, Six Pence None The Richer, Fuel, Filter, Guster, Ben Folds, Live, BT, Megadeth, 3rd Eye Blind, Red Hot Chili Peppers, Madonna and Crystal Method. He owns The Mix Room, Burbank, CA, which features two SSL-equipped rooms.

Taz Herzberg is a Los Angeles-based programmer and engineer. His credits include Counting Crows, Vanessa Carlton, Christina Aguilera, and the Grammy-awarded "Lady Marmalade" remake.

**Workshop 7 Sunday, October 6 1:00 pm–3:00 pm
Room 406AB**

AES42-2001: AES STANDARD FOR ACOUSTIC-DIGITAL INTERFACE FOR MICROPHONES

Chair: **Stephan Peus**, Georg Neumann GmbH, Berlin, Germany

Panelists: *Manfred Bleichwehl, Steven Harris, Juergen Wahl, Sennheiser Electronic Corporation, Van Nuys, CA, USA*

This workshop introduces the new *AES42-2001 Standard for Acoustic-Digital Interface for Microphones*. Among its most interesting potentials are the many options for controlling remotely a wide variety of microphone features, which are not possible with traditional analog microphones. Microphones that are built according to this new standard will not only have an internal analog/digital converter but also extended digital signal processing capabilities (DSP). It can incorporate, for example, the equivalent of a high-gain, high-precision traditional microphone preamplifier, inside the digital microphone. Presenters will demonstrate some of these features and advantages and show potential applications using an operational microphone system, designed according to the new AES42-2001 Standard.

**Session G Sunday, October 6 2:00 pm–6:00 pm
Room 408A**

MULTICHANNEL SOUND

Chair: **Durand Begault**, NASA Ames Research Center, Mountain View, CA, USA

2:00 pm

G-1 Interchannel Interference at the Listening Position in a Five-Channel Loudspeaker Configuration—*Geoff Martin*, McGill University, Montreal, Quebec, Canada, and *Bang & Olufsen a/s, Struer, Denmark*

It is commonly-accepted thinking that the use of a five-channel surround sound reproduction system increases the size of the listening area over that for two-channel stereophonic systems. In actual fact, for many types of program material, the area of this so-called sweet spot is reduced due to interference between the channels at the listener's ears. This effect is described and analyzed through theoretical evaluation and psychoacoustic listening tests.
Convention Paper 5677

2:30 pm

G-2 Breaking and Making the Rules: Sound Design in 6.1—*Ambrose Field*, University of York, York, UK

It is often colloquially said that there are no rules for the spatialization of multichannel content. This paper seeks to identify creative ways in which the multichannel systems of today and tomorrow can be harnessed to provide aesthetically convincing surround environments. The suitability of production techniques based on Ambisonic methods is re-evaluated for this task, and the discussion focuses on the role of multichannel systems in sound design for presentation in the movie theater.
Convention Paper 5672

3:00 pm

G-3 Localization of Lateral Phantom Images in a 5-Channel System With and Without Simulated Early Reflections—*Jason Corey, Wieslaw Woszczyk*, McGill University, Montreal, Quebec, Canada

Phantom images that rely on interchannel level differences can be produced easily for two-channel stereo. Yet one of the most difficult challenges in production for a five-channel environment is the creation of stable phantom images

to the side of the listening position. The addition of simulated early reflection patterns from all five loudspeakers influences the localization of lateral phantom sources. Listening tests were conducted to compare participants' abilities to localize lateral sources under three conditions: power-panned sources alone, sources with simulated early reflection patterns, and simulated early reflection patterns alone (without direct sound). Our results compare localization error for the three conditions at different locations and suggest that early reflection patterns alone can be sufficient for source localization.
Convention Paper 5673

3:30 pm

G-4 The Minimum Number of Loudspeakers and its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field—*Koichiro Hiyama, Setsu Komiyama, Kimio Hamasaki*, NHK, Setagaya, Tokyo, Japan

It is important to find out how many loudspeakers are necessary for multichannel sound systems to reproduce the spatial impression of a diffuse sound field. This paper discusses the issue in the case where loudspeakers are placed symmetrically along a concentric circle around the listener and at the same height as the listener's ear. It becomes clear that the spatial impression of the diffuse sound field can be obtained by only two symmetrical pairs of loudspeakers (that is, four loudspeakers in all). In this arrangement, one pair of loudspeakers should be placed in the frontal area around the listener at an angle of about 60 degrees, and the other pair should be in the rear area with an angle of 120 degrees to 180 degrees.
Convention Paper 5674

4:00 pm

G-5 A Fuzzy Cerebellar Model Approach for Synthesizing Multichannel Recordings—*Ching-Shun Lin, Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

New high capacity optical discs and high bandwidth networks provide the capability for delivering multichannel audio. Although there are many one- and two-channel recordings in existence, only a handful of multichannel recordings exist. In this paper we propose a neural network approach that can synthesize microphone signals with the correct acoustical characteristics of specific venues that have been characterized in advance. These signals can be used to generate a multichannel recording with the acoustical characteristics of the original venue. The complex semi-cepstrum technique is employed to extract features from musical signals recorded in a venue and these signals are sent into the fuzzy cerebellar model articulation controller (FCMAC) for training.
Convention Paper 5675

4:30 pm

G-6 Investigation of Listener Envelopment in Multichannel Surround Systems—*Gilbert A. Soulodre, Michel C. Lavoie, Scott G. Norcross*, Communications Research Centre, Ottawa, Ontario, Canada

It is now well understood that listener envelopment (LEV) is an essential component of good concert hall acoustics. An objective measure based on the late-lateral energy has been shown to perform well at predicting LEV. One goal of multichannel surround systems is to improve the re-creation of the concert hall experience in a home listening environment. By varying the amount of

late-lateral energy, such systems should allow the perception of LEV to be enhanced and controlled. In this paper the loudspeaker/listening room interactions are shown to limit the range of acoustical conditions that can be re-created. A series of formal subjective tests were conducted to determine if objective measures of late-lateral energy are suitable for predicting LEV in multichannel surround systems.

Convention Paper 5676

5:00 pm

G-7 The Significance of Interchannel Correlation, Phase, and Amplitude Differences on Multichannel Microphone Techniques—*Geoff*

Martin, McGill University, Montreal, Quebec, Canada, and Bang & Olufsen a/s, Struer, Denmark

There is a measurable interference between correlated signals produced by multiple loudspeakers in a standard five-channel loudspeaker configuration, resulting in an audible comb filter effect. This is due to small individual differences in distances between the ears of the listener and the various loudspeakers. Although this effect is caused by the dimensions and characteristics of the monitoring environment, it can be minimized in the recording process, particularly through the relative placement of microphones and choice of their directional characteristics. In order to analyze this effect, the correlation of microphone signals and their amplitude differences in a recording environment are evaluated using theoretical models. This procedure is applied to coincident and spaced pairs of transducers for direct and reverberant sounds.

Convention Paper 5671

5:30 pm

G-8 A Multichannel Surround Audio Format Applying 3-D Sound Technology—*Myung-Soo*

Kang, Kyoo-Nyun Kim, University of Ulsan, Ulsan, Korea

As systems employ multichannel audio more and more, it is necessary to consider the surround sound capability in the process of sound recording and playing. This paper presents a structured audio format and its application to design a more efficient surround sound system. Three-dimensional sound technology is used for localization of the sound source. We define the reusable sound object to clarify the audio format. Sound object is a unit of recorded sound samples that can be changed by various effect properties. Filter and 3-D properties are applied to change the sound objects in each track.

(Paper not presented at convention, but Convention Paper 5678 is available.)

Session H Sunday, October 6 2:00 pm–4:30 pm
Room 404AB

LOW BIT-RATE CODING, PART 1

Chair: **Gary Brown**, Tensilica, Santa Clara, CA, USA

2:00 pm

H-1 Scalable Lossless Audio Coding Based on MPEG-4 BSAC—*Doh-Hyung Kim, Jung-Hoe Kim,*

Sang-Wook Kim, Samsung Advanced Institute of Technology, Suwon, Korea

In this paper a new hybrid type of scalable lossless audio coding scheme based on MPEG-4 BSAC (bit sliced arithmetic coding) is proposed. This method introduces two residual error signals, lossy coding error signal and prediction error signal, and utilizes the rice coding as a lossless coding tool. These kinds of processes enable an increase in the compression ratio. As a result of experiment, average total file size can be reduced about 50 to 60 percent of the original size. Consequently, a slight modification of the conventional MPEG-4 general audio coding scheme can give a scalable lossless audio coding functionality between lossy and lossless bitstream.

Convention Paper 5679

2:30 pm

H-2 Lossless Audio Coding Using Adaptive Multichannel Prediction—*Tilman Liebchen,*

Technical University of Berlin, Berlin, Germany

Lossless audio coding enables the compression of digital audio data without any loss in quality due to a perfect reconstruction of the original signal. The compression is achieved by means of decorrelation methods such as linear prediction. However, since audio signals usually consist of at least two channels, which are often highly correlated with each other, it is worthwhile to make use of interchannel correlations as well. This paper shows how conventional (mono) prediction can be extended to stereo and multichannel prediction in order to improve compression efficiency. Results for stereo and multichannel recordings are given.

Convention Paper 5680

3:00 pm

H-3 Design of an Error-Resilient Layered Audio Codec—*Dai Yang, Hongmei Ai, Chris Kyriakakis, C.-C. Jay Kuo,*

University of Southern California, Los Angeles, CA, USA

Current high quality audio coding techniques mainly focus on coding efficiency, which makes them extremely sensitive to channel noise, especially in high error rate wireless channels. In this paper we propose an error-resilient layered audio codec (ERLAC) which provides functionalities of both fine-grain scalability and error-resiliency. A progressive quantization, a dynamic segmentation scheme, a frequency interleaving technique, and an unequal error protection scheme are adopted in the proposed algorithm to construct the final error robust layered audio bitstream. The performance of the proposed algorithm is tested under different error patterns of WCDMA channels with several test audio materials. Our experimental results show that the proposed approach achieves excellent error resilience at a regular user bit rate of 64 kb/s.

Convention Paper 5681

3:30 pm

H-4 Application of a Concatenated Coding System with Convolutional Codes and Reed-Solomon Codes to MPEG Advanced Audio Coding—*Dong Yan Huang¹, Say Wei Foo², Weisi Lin³, Ju-Nia Al Lee⁴*

¹Institute of Microelectronics, Singapore, Singapore

²Nanyang Technological University, Singapore, Singapore

³Laboratories of Information Technology, Singapore, Singapore

⁴National University of Singapore, Singapore

Reliable delivery of audio bitstream is vital to ensure the acceptable audio quality perceived by 3G network customers. In other words, an audio coding scheme that is employed must be fairly robust over the error-prone channels. Various error-resilience techniques can be utilized for the purpose. Due to the fact that some parts of the audio bitstream are less sensitive to transmission errors than others, the unequal error protection (UEP) is used to reduce the redundancy introduced by error resilience requirements. The current UEP scheme with convolutional codes and multistage interleaving has an unfortunate tendency to generate burst errors at the decoder output as the noise level is increased. A concatenated system combining Reed-Solomon codes with convolutional codes in the UEP scheme is investigated for MPEG advanced audio coding (AAC). Under severe channel conditions with random bit error rates of up to 5×10^{-2} , the proposed scheme achieved more than 50 percent improvement in residual bit error rate over the original scheme at a bit rate of 64 kb/s and sampling frequency of 48 kHz. Under burst error conditions with burst error length of up to 4 ms, the proposed scheme achieved more than 65 percent improvement in bit error rate over the original scheme. The average percentage overhead incurred by using the concatenated system is about 3.5 percent of the original UEP scheme. Further improvements are made by decreasing the puncturing rate of convolutional codes. However, this can only be adopted when high protection is needed in extremely noisy conditions (e.g., channel BER significantly exceeds 1.00×10^{-2}) since it incurs increased overheads.

Convention Paper 5682

4:00 pm

H-5 A Simple Method for Reproducing High Frequency Components at Low Bit-Rate Audio Coding—*Jeongil Seo, Daeyoung Jang, Jinwoo Hong, Kyeoungok Kang*, Electronics & Telecommunications Research Institute (ETRI), Yuseong-Gu, Deajon, Korea

In this paper we describe a simple method for reproducing high frequency components at low bit-rate audio coding. To compress an audio signal at low bit rates (below 16-kb/s per channel) we can use a lower sampling frequency (below 16 kHz) or high performance audio coding technology. When an audio signal is sampled at a low frequency and coded at a low bit rate, high frequency components and reverberant sound are lost because of quantization noise between pitch pulses. In a short-term period, the harmonic characteristic of audio signals is stationary, so the replication of high-frequency bands with low-frequency bands can extend the frequency range of resulting sound and enhance the sound quality. In addition, for reducing the number of bands to be reproduced we adapted this algorithm at the Bark scale domain. For compatibility with a conventional audio decoder, the additional bitstream is added at the end of each frame, which is generated by a conventional audio coder. We adapted this proposed algorithm to MPEG-2 AAC and increased the quality of audio in comparison with the conventional MPEG-2 AAC coded audio at the same rate. The computational cost of the proposed algorithm is similar to or a little more than a conventional MPEG-2 AAC decoder.

Convention Paper 5683

4:40 pm

Technical Committee Meeting on Coding of Audio Signals

**Workshop 6 Sunday, October 6 2:00 pm–5:00 pm
Room 408B**

GAME AUDIO

Chair: **Aaron Marks**, On Your Mark Music Productions, CA, USA

Speakers: *Keith Arem*, PCB Productions (CD-ROM audio)
Jack Buser, Dolby Labs, San Francisco, CA, USA (Use of Surround Sound in Video Games)
Brian Schmidt, Microsoft (XBox audio)
Martin Wilde, Game Audio Programmer (Game audio programming issues)

This game audio workshop is designed to answer one of the most often asked questions: What does it take to make great game audio? Not only will we provide the answers, but we also will come armed with many specific audio examples designed to showcase the specific talents, techniques, and technologies used to help make the games business the billion dollar industry that it is today.

5:10 pm

Technical Committee Meeting on Audio for Games

Education Event

ONE-ON-ONE MENTORING SESSION 1

Sunday, October 6, 2:30 pm–4:30 pm
Room 402A

Students are invited to sign up for an individual meeting with audio mentors from the industry. The sign-up sheet will be located near the student center of the convention, and all students are invited to participate in this exciting and rewarding opportunity for focused discussion.

Education Event

EDUCATION FAIR

Sunday, October 6, 2:30 pm–4:30 pm
Room 403B

Institutions offering studies in audio—from short courses to graduate degrees—will be represented in a tabletop session. Information on each school's respective programs will be made available through the display of literature and academic guidance sessions with representatives. There is no charge for schools to participate and admission is free and open to everyone.

Education Event

STUDENT POSTER SESSION

Sunday, October 6, 2:30 pm–4:30 pm
Room 403B

The event will display the scholarly/research/creative works from AES student members. Since many institutions are engaged in both research and applied sciences of audio, this session will provide an opportunity to display and discuss these accomplishments with professionals, educators, and other students.

Special Event
AN AFTERNOON WITH . . . TOM HOLMAN
Sunday, October 6, 2:30 pm–5:00 pm
Room 411 Theater

Moderator: **George Petersen**, Mix Magazine,
Emeryville, CA, USA

Tom Holman, whose distinguished career in audio, video, and film spans more than three decades, will participate in an interactive discussion with the audience. President of the TMH Corporation, and Professor of Film Sound at the University of Southern California School of Cinema-Television, Tom Holman has the rare ability to push the state of the art and challenge entire industries to achieve new quality standards. He creates new markets and redefines exiting markets by introducing new quality standards and designing products that perform to them.

Workshop 8 Sunday, October 6 3:00 pm–5:40 pm
Room 406AB

**PROTECTING YOUR HEARING AGAINST LOSS:
ASSESSMENT OF FUNCTIONAL HEARING ABILITY
IN NOISE**

Co-chairs: **Dilys Jones, Sigfrid Soli**, House Ear
Institute, Los Angeles, CA, USA

These experts in the field of hearing will discuss the methods in which hearing can be protected from loss and will assess functional hearing ability in noise. The workshop will feature a tutorial on the anatomy of the ear, a discussion of the factors that contribute to hearing loss, and hearing evaluations.

Sigfrid Soli will also discuss: audiometric testing; use and interpretation of an audiogram; extended high-frequency testing; and how to protect and conserve one's hearing by using the Occupational Safety and Health Administration's (OSHA) guidelines, hearing protection devices, in-ear monitors, and safe behavioral practices.

Education Event
EDUCATION FORUM

Sunday, October 6, 4:30 pm–6:00 pm
Room 402A

Co-hosts: **Theresa Leonard**, The Banff Centre,
Banff, Alberta, Canada
Don Puluse

This event is a meeting of the AES Education Committee, teachers, authors, students, and members interested in the issues of primary and continuing education of the audio industry. It is an opportunity to discuss the programs of the Education Committee and to provide input for future projects of this committee.

Special Event
THE JAZZ SINGER

Sunday, October 6, 6:30 pm–7:30 pm
Room 403A

An old time radio re-creation of the Lux Radio Theatre production of "The Jazz Singer," the first successful talking picture starring Al Jolson, will be staged on the 75th anniversary of the film's original debut. Several veteran radio actors have been assembled to commemorate both the film and the original 1947 broadcast. This radio re-creation will feature Richard Halpern in the starring role, with Herb Ellis directing.

Session I Monday, October 7 9:00 am–12:00 noon
Room 404AB

LOW BIT-RATE CODING, PART 2

Chair: **Marina Bosi**, MPEG LA, Denver, Colorado

9:00 am

**I-1 Perceptually-Based Joint-Program Audio
Coding—Christof Faller¹, Raziel Haimi-Cohen²,
Peter Kroon¹, Joseph Rothweiler¹**

¹Agere Systems, Murray Hill, NJ, USA

²Lucent Technologies, Murray Hill, NY, USA

Some digital audio broadcasting systems, such as Satellite Digital Audio Radio Services (SDARS), transmit many audio programs over the same transmission channel. Instead of splitting up the channel into fixed bit-rate subchannels, each carrying one audio program, one can dynamically distribute the channel capacity among the audio programs. We describe an algorithm which implements this concept taking into account statistics of the bit rate variation of audio coders and perception. The result is a dynamic distribution of the channel capacity among the coders depending on the perceptual entropy of the individual programs. This solution provides improved audio quality compared with fixed bit-rate subchannels for the same total transmission capacity. The proposed scheme is non-iterative and has a low computational complexity.
Convention Paper 5684

9:30 am

**I-2 Incorporation of Inharmonicity Effects into
Auditory Masking Models—Hossein Najaf-Zadeh,
Hassan Lahdili, Louis Thibault, Communications
Research Centre, Ottawa, Ontario, Canada**

In this paper the effect of inharmonic structure of audio maskers on the produced masking pattern is addressed. In most auditory models, the tonal structure of the masker is analyzed to determine the masking threshold. Based on psychoacoustic data, masking thresholds caused by tonal signals are lower compared to those produced by noise-like maskers. However, the relationship between spectral components has not been considered. It has been found that for two different multitonal maskers with the same power, the one with a harmonic structure produces a lower masking threshold. This paper proposes a modification to the MPEG psychoacoustic model 2 in order to take into account the inharmonic structure of the input signal. Informal listening tests have shown that the bit rate required for transparent coding of inharmonic (multitonal) audio material can be reduced by 10 percent if the new modified psychoacoustic model 2 is used in the MPEG 1 Layer II encoder.
Convention Paper 5685

10:00 am

**I-3 Binaural Cue Coding Applied to Audio
Compression with Flexible Rendering—Christof
Faller, Frank Baumgarte, Agere Systems, Murray Hill,
NJ, USA**

In this paper we describe an efficient scheme for compression and flexible spatial rendering of audio signals. The method is based on binaural cue coding (BCC) which was recently introduced for efficient compression of multichannel audio signals. The encoder input consists of separate signals without directional spatial cues, such as separate sound source signals, i.e., several monophon-

ic signals. The signal transmitted to the decoder consists of the mono sum-signal of all input signals plus a low bit rate (e.g., 2 kb/s) set of BCC parameters. The mono signal can be encoded with any conventional audio or speech coder. Using the BCC parameters and the mono signal, the BCC synthesizer can flexibly render a spatial image by determining the perceived direction of the audio content of each of the encoder input signals. We provide the results of an audio quality assessment using headphones, which is a more critical scenario than loudspeaker playback.

Convention Paper 5686

10:30 am

I-4 Two-Pass Encoding of Audio Material Using MP3 Compression—*Martin Weishart¹, Ralf Göbel², Jürgen Herre¹*

¹Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

²Fachhochschule Koblenz, Koblenz, Germany

Perceptual audio coding has become a widely-used technique for economic transmission and storage of high-quality audio signals. Audio compression schemes, as known from MPEG-1, 2, and 4 allow encoding with either a constant or a variable bit rate over time. While many applications demand a constant bit rate due to channel characteristics, the use of variable bit-rate encoding becomes increasingly attractive, e.g., for Internet audio and portable audio players. Using such an approach can lead to significant improvements in audio quality compared to traditional constant bit-rate encoding, but the consumed average bit rate will generally depend on the compressed audio material. This paper presents investigations into two-pass encoding which combines the flexibility of variable bit rate encoding and a predictable target bit consumption.

Convention Paper 5687

11:00 am

I-5 Technical Aspects of Digital Rights Management Systems—*Christian Neubauer¹, Karlheinz Brandenburg², Frank Siebenhaar¹*

¹Fraunhofer Institute for Integrated Circuits IIS-A, Erlangen, Germany

²Fraunhofer Arbeitsgruppe Elektronische Medientechnologie AEMT and Ilmenau University, Ilmenau, Germany

In today's multimedia world digital content is easily available to and widely used by end consumers. On the one hand high quality, the ability to be copied without loss of quality, and the existence of portable players make digital content, in particular digital music, very attractive to consumers. On the other hand the music industry is facing increasing revenue loss due to illegal copying. To cope with this problem so-called Digital Rights Management (DRM) systems have been developed in order to control the usage of content. However, currently no vendor and no DRM system is widely accepted by the market. This is due to the incompatibility of different systems, the lack of open standards, and other reasons. This paper analyzes the current situation of DRM systems, derives requirements for DRM systems, and presents technological building blocks to meet these requirements. Finally, an alternative approach for a DRM system is presented that better respects the rights of the consumers.

Convention Paper 5688

11:30 am

I-6 Configurable Microprocessor Implementation of Low Bit-Rate Audio Decoding—*Gary S. Brown, Tensilica, Inc., Santa Clara, CA, USA*

Using a configurable microprocessor to implement low bit-rate audio applications by tailoring the instruction set reduces algorithm complexity and implementation cost. As an example, this paper describes a Dolby digital (AC-3) decoder implementation that uses a commercially-available configurable microprocessor to achieve 32-bit floating-point precision while minimizing the required processor clock rate and die size. This paper focuses on how the audio quality and features of the reference decoder algorithm dictate the customization of the microprocessor. This paper shows examples of audio specific extensions to the processor's instruction set to create a family of AC-3 decoder implementations that meet multiple performance and cost points. How this approach benefits other audio applications is also discussed.

Convention Paper 5689

12:10 pm

Technical Committee Meeting on Network Audio Systems

**Session J Monday, October 7 9:00 am–11:00 am
Room 406AB**

HIGH-RESOLUTION AUDIO

Chair: **Rhonda Wilson**, Meridian Audio Limited, Huntingdon, Cambridgeshire, UK

9:00 am

J-1 DSD Processing Modules—Architecture and Application—*Nathan Bentall, Gary Cook, Peter Eastty, Eamon Hughes, Michael Page, Christopher Sleight, Mike Smith, Sony Broadcast & Professional Research Labs, Oxford, UK*

As demand continues to grow for production equipment targeting the high resolution, multichannel capabilities of SACD (super audio CD), there is increasing interest in adding DSD capability to both new and existing systems. The prospect of researching and implementing the necessary algorithms from scratch can be daunting. The high data rates, coupled with the asymmetric multipliers often required by the algorithms, make conventional von-Neumann-type DSP platforms (where most developers traditionally have their DSP expertise) seem sub-optimal. Based on custom DSD audio processing engines and packaged in the very compact SODIMM form factor, the modules described in this paper can add high quality, real time, low latency DSD audio processing functionality to a system with a minimum of development time.

Convention Paper 5690

9:30 am

J-2 Multichannel Audio Connection for Direct Stream Digital—*Michael Page, Nathan Bentall, Gary Cook, Peter Eastty, Eamon Hughes, Christopher Sleight, Sony Broadcast & Professional Research Labs, Oxford, UK*

The development of large-scale multitrack production equipment for direct stream digital (DSD) audio will be

substantially aided by the availability of a convenient multichannel interface. DSD is the high-resolution one-bit audio coding system for the super audio CD consumer disc format. Existing multichannel audio interconnection and networking protocols are not easily able to support the high-frequency sample clock required by DSD. A new interconnection type is introduced to provide reliable, low-latency, full-duplex transfer of 24 channels of DSD, using a single conventional office networking cable. The interconnection transfers audio data using Ethernet physical layer technology, while conveying a DSD sample clock over the same cable.

Convention Paper 5691

10:00 am

J-3 The Effect of Idle Tone Structure on Effective Dither in Delta-Sigma Modulation Systems: Part 2—*James A. S. Angus*, University of Salford, Salford, UK

This paper clarifies some of the confusion that has arisen over the efficacy of dither in PCM and sigma-delta modulation (SDM) systems. It presents a means of analyzing in-band, idle tone structure using chaos theory and describes a fair means of comparison between PCM and SDM. It presents results that show that dither can be effective in SDM systems.

Convention Paper 5692

10:30 am

J-4 DC Analysis of High Order Sigma-Delta Modulators—*Derk Reefman, Erwin Janssen*, Philips Research, Eindhoven, The Netherlands

A new method for the DC analysis of a sigma-delta modulator (SDM) is presented. The model used for the description of an SDM is adopted from Candy's model for a first order SDM. However, where Candy's model is exact for a first order SDM, it fails to be so in a higher order case. In our model, we deal with this by the introduction of stochastic behavior of the SDM and obtain the probability density distribution function of some variables which determine many of the characteristics of the SDM in the time domain. Comparison with simulation results shows that the assumption of stochastic behavior is rather good for SDM orders greater than 3, which display significant noise shaping. For lower orders (or less aggressive noise shaping) the approximation is less good. As an aside, the new model of sigma-delta modulation also clarifies why the time-quantized dither approach presented by Hawksford is much better compared to standard quantizer dithering.

Convention Paper 5693

11:10 am

Technical Committee Meeting on High-Resolution Audio

Workshop 9 Monday, October 7 9:00 am–12:00 noon Room 408A

STUDIO PRODUCTION AND PRACTICES

Chair: **George Massenburg**, GML, North Hollywood, CA, USA
with **David Smith**, Sony Music Studios, New York, NY, USA

This workshop will be an open discussion led by esteemed recording engineers George Massenburg and David Smith on de-

livery standards in digital and analog audio in the modern studio.

Mr. Massenburg and Mr. Smith, who are respectively the Chair and a member of the AES Technical Committee on Studio Practices, will detail their proven techniques for establishing high quality, archive-ready audio, using today's state of the art professional tools. They will concentrate on delivery formats, media requirements, and proper documentation, and will cover the myriad of formats embedded in current digital audio workstations. In addition, they will share their practical secrets for everyday success in the studio environment.

12:10 pm

Technical Committee Meeting on Studio Practices and Production

Workshop 10 Monday, October 7 9:00 am–12:00 noon Room 408B

LOUDSPEAKER LINE ARRAYS, PART 1: THEORY AND HISTORY

Chair: **Jim Brown**, Audio Systems Group, Inc., Chicago, IL, USA

Panelists: *Wolfgang Ahnert*, Acoustic Design Ahnert, Berlin, Germany
Stefan Feistel, Software Design Ahnert, Berlin, Germany
Kurt Graffy, Arup Acoustics, Los Angeles, CA, USA
David Guinness, Eastern Acoustic Works, Inc., Whitinsville, MA, USA
Don Keele, Harman/Becker Automotive Systems, Martinsville, IN, USA
Mike Klasco, Menlo Scientific, El Sobrante, CA, USA

This tutorial workshop is targeted at sound reinforcement professionals. The session includes development of basic concepts underlying line arrays, both simple and complex, an historical survey of commercial line array products, and a variety of line array design techniques. Methods of predicting line array performance are described, both at the conceptual and practical level. This workshop is a highly useful precursor to Workshop 12, Loudspeaker Line Arrays, Part 2: Practice and Applications, at 1:30 pm this afternoon.

Presentations:

Historical Review of Line Array Products, by *Mike Klasco*

Basic Line Array Concepts," *Don Keele*

Electronically-Steered Arrays, by *David Guinness*

Measurement and Modeling of Multi-Element Loudspeakers and Modeling of Waveguide Elements, by *David Guinness*

Line Array System Design Considerations, by *Wolfgang Ahnert*

3-D Modeling of Line Arrays in Software, by *Stefan Feistel*

Sort-of Line Arrays? by *Jim Brown*

12:10 pm

Technical Committee Meeting on Loudspeakers

Education Event

ONE-ON-ONE MENTORING SESSION 2
Monday, October 7, 10:00 am–12:00 noon
Room 402A

Students are invited to sign up for an individual meeting with ➔

audio mentors from the industry. The sign-up sheet will be located near the student center of the convention, and all students are invited to participate in this exciting and rewarding opportunity for focused discussion.

Special Event

PLATINUM PRODUCERS PANEL 2: PAST, PRESENT, AND FUTURE OF RECORDING

Monday, October 7, 12:30 pm–2:30 pm
Room 403A

Moderator: **Howard Massey**, On The Right Wavelength Consulting

Panelists: *Michael Bradford*
Bob Ezrin
Patrick Leonard
Larry Levine
Phil Ramone

How has the art of music recording changed over the past four decades? What is the state-of-the-art today? Where is recording going? Join us for a lively, fascinating discussion as four of the world's top record producers talk about how far we have come and make predictions for the future.

Multi-instrumentalist Michael Bradford (a.k.a. Chunky Style) is one of today's hottest producers. Best known for his co-productions of Kid Rock's controversial *The History of Rock* and Uncle Kracker's hit single *Follow Me*, Bradford has also worked with Run-DMC, Terence Trent D'Arby, and Gregg Alexander, and has collaborated with renowned arranger Paul Buckmaster on a number of movie soundtracks.

Bob Ezrin has produced multiplatinum albums and live events for some of the world's leading artists, including Pink Floyd (*The Wall*), Alice Cooper, KISS, Rod Stewart, Peter Gabriel, Lou Reed, Roger Daltrey, and Nine Inch Nails. He has also established several successful production and publishing companies, a recording and mastering studio, and an independent label. He is also active in community service, and currently serves as Co-Chairman of the Board of Directors of The Los Angeles Mentoring Partnership, and is VP of the Board of Directors of the Mr. Holland's Opus Foundation. He is currently producing the new Jane's Addiction album.

Keyboardist Patrick Leonard produced Elton John's critically acclaimed 2001 album *Songs from the West Coast* and recently completed the new album for Duncan Sheik. He has also produced Madonna, Jewel, Bryan Ferry, Roger Waters, and Rod Stewart, and has enjoyed fame as both a solo artist and as half of the nineties rock duo Toy Matinee.

Larry Levine is responsible for some of the greatest classic rock recordings of all time. As Phil Spector's longtime engineer, he helped craft the wall of sound that made records by The Ronettes and other Spector artists sonically unique. Levine also manned the board for numerous Beach Boys albums, including the legendary *Pet Sounds*, as well as Ike and Tina Turner's famed *River Deep, Mountain High*. He has also recorded Herb Alpert, The Righteous Brothers, Billy Eckstine, and The Ramones.

Phil Ramone is one of the most revered names in the music business. A former child prodigy and classically trained violinist, the Pope of Pop has produced an astonishing list of platinum-selling artists, including Billy Joel, Frank Sinatra, Paul Simon, Ray Charles, Elton John, Bono, Quincy Jones, Madonna, Tony Bennett, Carly Simon, Gloria Estefan, Luciano Pavarotti, Natalie Cole, B.B. King, Paul McCartney, Sinéad O'Connor, George Michael, James Taylor, and Jon Secada. He has also produced numerous theater, television, and film soundtracks and is currently working on album projects with Rod Stewart and Liza Minnelli.

Moderator Howard Massey is a noted industry consultant and veteran audio journalist. He has written for *Home Recording*,

Surround Professional, *EQ*, *Musician*, *Billboard*, and many other publications, and is the author of 11 books, including *Behind the Glass*, a collection of interviews with the world's top record producers. Massey has also worked extensively as an audio engineer, producer, songwriter, and touring/session musician.

Education Event

STUDENT RECORDING COMPETITION

Monday, October 7, 1:00 pm–6:00 pm
Room 309

Hosts: **Theresa Leonard**, The Banff Centre,
Banff, Alberta, Canada
Don Puluse

Finalists selected by an elite panel of judges will give brief descriptions and play recordings in the Classical and Jazz/Pop categories. One submission per category per student per student section or school is allowed. Meritorious awards will be presented at the closing Student Delegate Assembly meeting on Tuesday, October 8, at 10:00 am in Room 402AB.

1:00 pm–2:00 pm Classical Category¹
2:00 pm–3:00 pm Surround Classical Category¹
3:00 pm–4:00 pm Jazz/Folk Category²
4:00 pm–5:00 pm Pop/Rock Category³
5:00 pm–6:00 pm Surround Non-Classical Category³

¹**Judges:** *Michael Bishop, John Eargle, Richard King*

²**Judges:** *Jim Anderson, Lynn Fuston, Bill Schnee*

³**Judges:** *Frank Filipetti, Bob Ludwig, Elliot Scheiner*

Workshop 12 Monday, October 7 1:30 pm–5:30 pm
Room 408B

LOUDSPEAKER LINE ARRAYS, PART 2: PRACTICE AND APPLICATION

Chair: **John Murray**, Live Sound International Magazine!, San Francisco, CA, USA

Panelists: *François Deffarges*, Nexo, San Rafael, CA, USA
Mark Engebretson, JBL Professional, Northridge, CA, USA
Christian Heil, L'Acoustics, Oxnard, CA, USA
Tom McCauley, McCauley Sound, Inc., Puyallup, WA, USA
Jeff Rocha, Eastern Acoustic Works, Whitinsville, MA, USA
Evert Start, Duran Audio, Zaltbommel, The Netherlands

This workshop focuses on commercially available modular line array systems for performance audio use. Presenters who represent engineering departments of several competitive manufacturers and are actively working in the development of loudspeaker line arrays and for manufacturers of such systems will be featured.

Presentations:

Fresnel Zone Analysis, Wave Sculpturing Technology (WST), and HF DOSC Wave Guide, by *Christian Heil*

Arithmetic Spiral Arrays, MF Diffraction-Slot Band-Pass Section, and High-Efficiency Driver Design, by *Mark Engebretson*

Divergence Shading, MF Parabolic Separator, and MF & LF Band-Pass Technology, by *Jeff Rocha*

Hyperboloid Reflective Wave Source, MF Phase Plug, and Hypercardioid Subbass, by *François Deffarges*

Adaptive Density Inverse Flat Lens, MF Driver Design, and Inter-Cell Summation Aperture, by *Tom McCauley*.

Logarithmic Driver Spacing, Bessel FIR Filtering, and Steerable Lobes/2-Lobe Arrays, by *Evert Start*

**Session K Monday, October 7 2:00 pm–4:00 pm
Room 404AB**

RECORDING AND REPRODUCTION OF AUDIO

Chair: **Bob Moses**, Island Digital Media Group, Vashon, WA, USA

2:00 pm

K-1 Power Supply Regulation in Audio Power Amplifiers—*Eric Mendenhall*, Gibson Labs, Redondo Beach, CA, USA

Audio power amplifiers have typically been supplied power by the simplest possible means, usually an off-line supply with no line or load regulation, most commonly based on a line frequency transformer. Even modern amplifiers utilizing switchmode power supplies are usually designed without line or load regulation. The exception has been made for high-end audiophile amplifiers. The pros and cons of a regulated power supply are investigated.
Convention Paper 5694

2:30 pm

K-2 Audio Power Amplifier Output Stage Protection—*Eric Mendenhall*, Gibson Labs, Redondo Beach, CA, USA

This paper reviews a progression of circuits used for protecting bipolar power transistors in the output stages of audio power amplifiers. Design oriented methods of determining the protection locus are shown in a mathematical and graphical procedure. The circuits are then expanded from their standard configurations to allow for transient excursion beyond steady state limits, and thermally dependent protection limits, to better match the protection limits to the actual output stage capability. This allows the protection scheme to prevent output stage failure in the least restrictive way. A new method is shown for achieving a junction temperature estimation system without the use of a multiplier.
Convention Paper 5695

3:00 pm

K-3 Archiving Audio—*Jim Wheeler*, Tape Restoration & Forensics Company, Oceano, CA, USA

As hundreds of millions of tapes age, they begin to deteriorate. This paper describes how to recover these unplayable tapes as well as how to store them properly. This paper will also cover all of the issues of archiving audio, including high-capacity and inexpensive hard disk drives, as well as equipment obsolescence and new media.
Convention Paper 5696

3:30 pm

K-4 Finding a Recording Audio Education Program that Suits Your Career Choice—*Laurel Cash-Jones*, Burbank, CA, USA

This paper discusses the past, present, and future of recording audio education. It describes how the job mar-

ket and educational requirements have changed, and takes a look at how to plan for a successful career. It also provides valuable information on getting and keeping a job in today's fast-paced world of professional audio.
Convention Paper 5697

4:10 pm

Technical Committee Meeting on Audio Recording and Storage Systems

**Session L Monday, October 7 2:00 pm–4:30 pm
Room 406AB**

AUDIO NETWORKING AND AUTOMOTIVE AUDIO

Chair: **John Strawn**, S Systems, Larkspur, CA, USA

2:00 pm

L-1 Studio Exploring Using Universal Plug and Play—*Rob Laubscher*¹, *Richard Foss*²

¹Seattle, WA, USA

²Rhodes University, Grahamstown, South Africa

This paper explores the use of universal plug and play (UPnP) as a studio control technology. The architecture of a possible studio control technology is introduced. The elements of this studio control architecture are related to the architecture of UPnP. A sample implementation demonstrates the key aspects of using UPnP as a studio control technology.
Convention Paper 5698

2:30 pm

L-2 mLAN: Current Status and Future Directions—*Jun-chi Fujimori*¹, *Richard Foss*²

¹Yamaha Corporation, Hamamatsu, Japan

²Rhodes University, Grahamstown, South Africa

“mLAN” describes a network that allows for the transmission and receipt of audio and music control data by audio devices. IEEE 1394 was chosen as the specification upon which to implement mLAN. mLAN has built on IEEE 1394 and related standards, introducing formats, structures, and procedures that enable the deployment of IEEE 1394 within a music studio context. This paper discusses these standards, their implementations, and provides pointers to the future evolution of mLAN.
Convention Paper 5699

3:00 pm

L-3 Mutually-Immersive Audio Telepresence—*Norman P. Jouppi*¹, *Michael J. Pan*²

¹HP Labs, Palo Alto, CA, USA

²UCLA, Los Angeles, CA, USA

Mutually-immersive audio telepresence attempts to create for the user the audio perception of being in a remote location, as well as simultaneously to create the perception for people in the remote location that the user of the system is present there. The system provides bidirectional multichannel audio with relatively good fidelity, isolation from local sounds, and a reduction of local reflections. The system includes software for reducing unwanted feedback and joystick control of the audio environment. We are investigating the use of this telepresence system as a substitute for business travel. Initial

user response has been positive, and future improvements in networking quality of service (QoS) will improve the interactivity of the system.

Convention Paper 5700

3:30 pm

L-4 Assessment of the Sound Field in a Car—

Chulmin Choi¹, Lae-Hoon Kim¹, Sejin Doo², Yangki Oh³, Koeng-Mo Sung¹

¹Seoul National University, Seoul, Korea

²Dong-Ah Broadcasting College, Ansong, Korea

³Mokpo National University, Mokpo, Korea

This paper describes the measurement and assessment method for the sound field in a car cabin, which is so small that conventional room acoustical parameters cannot be employed directly. First, we measured the sound field using a multichannel microphone system and calculated some room acoustical parameters for judging their validity in the car cabin. As a result, we concluded that many of the conventional parameters do not have useful meaning in a car and its audio system. By analyzing the impulse responses from many cars, we developed some parameters for a more profound assessment of the sound field in a car.

Convention Paper 5701

4:00 pm

L-5 Measurement of the Speech Intelligibility Inside Cars—*Angelo Farina, Fabio Bozzoli, University of Parma, Parma, Italy*

The paper describes a measurement system developed for assessing speech intelligibility inside car compartments. This relates directly to the understandability of the voices being reproduced through the radio receiving system (news, traffic information, etc.), but in the future it will also be used for assessing the direct speech communications between the passengers and the performance of hands-free communication devices. The system is based on the use of two head-and-torso simulators, one equipped with an artificial mouth, the second equipped with binaural microphones. Only the second is used when the sound is being reproduced through the car's sound system. The MLS-based version of the STI method is used for performing the measurements, taking into account the effect of the background noise and the electro-acoustic propagation inside the compartment.

Convention Paper 5702

4:40 pm

Technical Committee Meeting on Automotive Audio

Workshop 11 **Monday, October 7** **2:00 pm–5:00 pm**
Room 408A

RECENT DEVELOPMENTS IN MPEG-4 AUDIO

Chair: **Jürgen Herre**, Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

Panelists: *Martin Dietz*, Coding Technologies, Nürnberg, Germany
Ralf Geiger, Fraunhofer Institute for Integrated Circuits IIS-A, Erlangen, Germany
Leon van de Kerkhof, Philips Digital Systems Laboratory, Eindhoven, The Netherlands

For more than a decade MPEG audio standards have been defin-

ing the state of the art in perceptual coding of audio signals. Several phases of standardization (MPEG-1, MPEG-2, MPEG-4) have been pushing borders beyond everyone's initial expectations. This workshop reports on three new extensions to the existing MPEG-4 audio standard which are currently under development. The work on compatible bandwidth extension of audio signals provides an additional performance boost for AAC coding at very low bit rates, permitting considerable signal quality at bit rates around 24 kbit/s per channel. A second extension augments the existing MPEG parametric audio coder with modes for representing high-quality audio signals. Finally, methods for lossless audio coding are under consideration, extending current MPEG-4 audio coders toward perfect representation at word lengths and sampling rates typically associated with high-definition audio. The workshop will provide background and demonstrations on these technologies.

Special Event

SPARS BUSINESS PANEL: WHAT I'VE HAD TO DO TO SURVIVE SINCE 9/11, ENRON, "DESKTOP" AUDIO, AND STIFFER COMPETITION

Monday, October 7, 3:30 pm–5:30 pm

Room 403A

Moderator: **David Amlen**, SPARS President, Sound on Sound (NYC)

Panelists: *Russ Berger*, Russ Berger Design Group, Addison, TX, USA
Steve Davis, Crawford Audio, Atlanta, GA, USA
Kevin Mills, Larrabee Studios, Los Angeles, CA, USA
Joe Montarello, Capitol Region Insurance
David Porter, Annex Digital, San Francisco, CA, USA

The Society of Professional Audio Recording Services (SPARS) is a 23-year-old professional organization dedicated to sharing practical, hands-on business information about audio facility ownership, management, and operations. This event, which is co-hosted by the AES, features an elite panel of studio owners and managers who will explore strategies for adapting your business to changing times. Panelists will illustrate, with stories and solutions, innovative ways they have handled unique business challenges faced in successfully running today's audio production facility.

Special Event

OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE

Monday, October 7, 6:00 pm–8:00 pm

Room 403A

Lecturer: **James E. West**

The Heyser Series is an endowment for lectures by eminent individuals with outstanding reputations in audio engineering and its related fields. The series is featured twice annually at both the United States and European AES conventions. Established in May 1999, The Richard C. Heyser Memorial Lecture honors the memory of Richard Heyser, a scientist at the Jet Propulsion Laboratory, who was awarded nine patents in audio and communication techniques and was widely known for his ability to clearly present new and complex technical ideas. Mr. Heyser was also an AES governor and AES Silver Medal recipient.

The distinguished Richard C. Heyser lecturer is Professor James West of Avaya Labs and Johns Hopkins University, co-

inventor of the electret condenser microphone. His lecture is titled, "Modern Electret Microphones and Their Applications."

It is well known that condenser microphones are the transducer of choice when accuracy, stability, frequency characteristics, dynamic range, and phase are important. But conventional condenser microphones require critical and costly construction as well as the need for a high dc bias for linearity. These disadvantages ruled out advanced practical microphone designs such as multi-element arrays and the use of linear microphones in telephony. The combination of our discovery of stable charge storage in thin polymers and the need for improved linearity in communications encouraged the development of modern electret microphones in the early 1960s at Bell Labs.

Others had suggested the use of electrets in transducers (electrical analog of a permanent magnet) in the late 1920s, but these and subsequent efforts all suffered from the insufficient stability of wax electrets under normal environmental conditions. Water molecules from atmospheric humidity were the main depolarizing factor in Carnauba and other wax electrets. The first broad application of wax electret microphones was discovered in captured World War II Japanese field equipment. Because of the decay of the polarization of the electret, these microphones had a lifetime of about six months.

Modern polymer electret transducers can be constructed in various sizes and shapes mainly because of the simplicity of the transducer. All that is needed in the mechanical system is a thin (25 microns) charged polymer, a small (irregular) air gap and a back plate. An impedance converter is necessary and is provided by a FET transistor to better match conventional electronic equipment. Applications of electret microphones range from very small hearing aid microphones (a few square millimeters) to very large single element units (20 cm diameter) for underwater and airborne reception of very low frequencies. Because the frequency and phase response of electret microphones are relatively constant from unit to unit, multiple element two-dimensional arrays can be constructed. We have constructed a two-dimensional, 400-element array for Arnold Auditorium at Bell Labs with electret elements that are available for under \$1.00 each.

Telephone bandwidth and frequency characteristics have remained constant for the past 30 years while entertainment has brought high fidelity including surround sound into most homes throughout the world. People are accustomed to good quality sound and expect it in communication systems. The Internet Protocol (IP) offers the needed bandwidth to improve audio quality for telephony, but it will require broadband microphones and loudspeakers to provide customers with voice presence and clarity. Directional microphones for both handheld and hands-free modes are necessary to improve signal-to-noise ratios and to enable automatic speech recognition. Arrays with dynamic beam forming properties are required especially for conference rooms. Signal processing has made possible stereo acoustic echo cancellers and many other signal enhancements that improve audio quality. Dr. West will discuss some of the current work on broadband communications at Avaya Labs.

James E. West is a Bell Laboratories Fellow and former member of the Acoustics and Speech Research Department at Lucent Technologies, specializing in electroacoustics, physical acoustics, and architectural acoustics. He is now research scientist in the Multimedia Technologies Research Lab of Avaya. His pioneering research on charge storage and transport in polymers led to the development of electret transducers for sound recording and voice communication. Almost 90 percent of all microphones built today are based on the principles first published by West in the early 1960s. This simple but rugged transducer is the heart of most new telephones manufactured by Lucent and other producers of communication equipment. He is the recipient of the Callinan Award (1970), sponsored by the Electrochemical Society of America, the Senior Award

(1970), sponsored by the IEEE Group on Acoustics, the Lewis Howard Latimer Light Switch and Socket Award (1989), sponsored by the National Patent Law Association, the George R. Stibitz Trophy, sponsored by the Third Annual AT&T Patent Award (1993), New Jersey Inventor of the Year for 1995, The Acoustical Society of America's Silver Medal in Engineering Acoustics (1995), an honorary Doctor of Science degree from New Jersey Institute of Technology (1997), the Golden Torch Award (1998) sponsored by the National Society of Black Engineers, the Industrial Research Institutes 1998 Achievement Award, and The Ronald H. Brown American Innovator Award (1999). The Acoustical Society of America has chosen one of his early papers on Electret Microphones as a Benchmark publication.

The lecture will be followed by a reception hosted by the Technical Council.

Special Event TOUR OF CATHEDRAL OF OUR LADY OF THE ANGELS AND ORGAN RECITAL

Tuesday, October 8, 8:30 am–3:00 pm
Cathedral of Our Lady of the Angels
555 W. Temple Street
Los Angeles

Join with the American Institute of Organbuilders for a tour and series of lectures in the newly-completed Cathedral of Our Lady of the Angels. The prime focus of the tour will be "Opus 75"—the Cathedral's new 105 rank, 6019 pipe organ by Dobson Pipe Organ Builders (Lake City, Iowa). The tour also will feature a recital by the AES' "resident organist" Graham Blyth.

Included in the series of lectures will be "The Cathedral Organ: The Truth Revealed," by Manuel Rosales, consultant to the Cathedral for the organ project; "Design Adventures: Collaboration with Architects," by Lynn Dobson, President, Dobson Pipe Organ Builders; "How to Hold Up 55 Tons of Organ," by Jon Thieszen, a member of the organ design team; "Entering the Tonal Unknown," by John Panning and Samuel Soria, Cathedral Organists (this will be a lecture and demonstration of the features of "Opus 75").

At noon, lunch will be provided in the Archdiocesan Conference Center, accompanied by a lecture, "Organ Structure: Preparing for the Big One!" John Seest will offer a look into the support of the Dobson Organ in the massive new concrete Cathedral and the attention given to the high earthquake design loads that make this a unique organ structure. Following the lunch and lecture, Graham Blyth will perform a recital on the Dobson Organ.

To conclude the tour, Dennis Paoletti and John Prohs, acoustical consultants for the Cathedral, will discuss some of the unique design features and challenges they faced in realizing this massive and complex project.

Cost for the tour, including lunch, is \$45 per person. Reservations are required and must be made at the Special Events Desk no later than 4:00 pm Sunday, October 6th. Convention attendees wanting to attend just the organ recital by Graham Blyth, but not the entire tour, are encouraged to check the poster in the Convention Center lobby for the starting time and additional details.

Attendance at the recital is free.

**Session M Tuesday, October 8 9:00 am–12:00 noon
Room 404AB**

PSYCHOACOUSTICS, PART 1

Chair: **Christopher Struck**, Dolby Laboratories, San Francisco, CA, USA

9:00 am

M-1 Comparison of Objective Measures of Loudness Using Audio Program Material—*Eric Benjamin*, Dolby Laboratories, San Francisco, CA, USA

Level measurements do not necessarily correlate well with subjective loudness. Several methods are described for making objective measurements that are designed to correlate better with actual loudness. Some of these measures are: A-weighting and B-weighting, the methods of Stevens and Zwicker as described in ISO 532A and B, and the method described by Moore and Glasberg. All of these measures are intended to describe (measure) the loudness of sounds with constant spectra. How well do these measures work with typical audio signals distributed via broadcast or recording?
Convention Paper 5703

9:30 am

M-2 Descriptive Analysis and Ideal Point Modelling of Speech Quality in Mobile Communication—*Ville-Veikko Mattila*, Nokia Research Center, Tampere, Finland

Descriptive analysis of processed speech quality was carried out by semantic differentiation, and external preference mapping was used to relate the attributes to overall quality judgments. Clean and noisy speech samples from different speakers were processed by various processing chains representing mobile communications, resulting in a total of 170 samples. The perceptual characteristics of the test samples were described by 18 screened subjects, and the final descriptive language with 21 attributes were developed in panel discussions. The scaled attributes were mapped to overall quality evaluations collected from 30 screened subjects by partial least square regression. The Phase II ideal point modeling was used to predict the quality with an average error of about 6 percent and to study the linearity of the attributes.
Convention Paper 5704

10:00 am

M-3 Relating Multilingual Semantic Scales to a Common Timbre Space—*William L. Martens*, *Charith N. W. Giragama*, University of Aizu, Aizu-Wakamatsu, Fukushima-ken, Japan

A single, common perceptual space for a small set of processed guitar timbres was derived for two groups of listeners, one group comprising native speakers of the Japanese language, the other group comprising native speakers of Sinhala, a language of Sri Lanka. Subsets of these groups made ratings on 13 bipolar adjective scales for the same set of sounds, each of the two groups using anchoring adjectives taken from their native language. The adjectives were those freely chosen most often in a preliminary elicitation. The results showed that the Japanese and Sinhalese semantic scales related differently to the dimensions of their shared timbre space that was derived using MDS analysis of the combined dissimilarity ratings of listeners from both groups.
Convention Paper 5705

10:30 am

M-4 Design and Evaluation of Binaural Cue Coding Schemes—*Frank Baumgarte*, *Christof Faller*, Agere Systems, Murray Hill, NJ, USA

Binaural cue coding (BCC) offers a compact parametric representation of auditory spatial information such as lo-

calization cues inherent in multichannel audio signals. BCC allows reconstruction of the spatial image given a mono signal and spatial cues that require a very low rate of a few kbit/s. This paper reviews relevant auditory perception phenomena exploited by BCC. The BCC core processing scheme design is discussed from a psychoacoustic point of view. This approach leads to a BCC implementation based on binaural perception models. The audio quality of this implementation is compared with low-complexity FFT-based BCC schemes presented earlier. Furthermore, spatial equalization schemes are introduced to optimize the auditory spatial image of loudspeaker or headphone presentation.
Convention Paper 5706

11:00 am

M-5 Evaluating Influences of a Central Automotive Loudspeaker on Perceived Spatial Attributes Using a Graphical Assessment Language—*Natanya Ford*¹, *Francis Rumsey*¹, *Tim Nind*²

¹University of Surrey, Guildford, Surrey, UK

²Harman/Becker Automotive Systems, Martinsville, IN, USA

An investigation is described which further develops a graphical assessment language (GAL) for subjectively evaluating spatial attributes of audio reproductions. Two groups of listeners, those with previous experience of using a GAL and listeners new to graphical elicitation, were involved in the study which considered the influence of a central automotive loudspeaker on listeners' perception of ensemble width, instrument focus, and image skew. Listeners represented these attributes from both driver's and passenger's seats using their own graphical descriptors. Source material for the study consisted of simple instrumental and vocal sources chosen for their spectral and temporal characteristics. Sources ranged from a sustained cello melody to percussion and speech extracts. When analyzed using conventional statistical methods, responses highlighted differences in listeners' perceptions of width, focus, and skew for the various experimental conditions.
Convention Paper 5707

11:30 am

M-6 Multidimensional Perceptual Calibration for Distortion Effects Processing Software—*Marui Atsushi*, *William L. Martens*, University of Aizu, Aizuwakamatsu-shi, Fukushima-ken, Japan

Controlled nonlinear distortion effects processing produces a wide range of musically useful outputs, especially in the production of popular guitar sounds. But systematic control of distortion effects has been difficult to attain, due to the complex interaction of input gain, drive level, and tone controls. Rather than attempting to calibrate the output of commercial effects processing hardware, which typically employs proprietary distortion algorithms, a real-time software-based distortion effects processor was implemented and tested. Three distortion effect types were modeled using both waveshaping and a second order filter to provide more complete control over the parameters typically manipulated in controlling effects for electric guitars. The motivation was to relate perceptual differences between effects processing outputs and the mathematical functions describing the nonlinear waveshaping producing variation in distortion. Perceptual calibration entailed listening sessions where listeners adjusted the tone of each of nine test outputs, and then made both

pairwise dissimilarity ratings and attribute ratings for those nine stimuli. The results provide a basis for an effects-processing interface that is perceptually-calibrated for system users.

Convention Paper 5708

**Workshop 13 Tuesday, October 8 9:00 am–12:00 noon
Room 408B**

**PERCEPTUAL ISSUES RELATED TO CASCADED
AUDIO CODECS**

Chair: **Thomas Sporer**, Fraunhofer Institute for Integrated Circuits, Ilmenau, Germany

Panelists: *Louis Fielder*, Dolby Laboratories, Inc., San Francisco, USA
Jürgen Herre, Fraunhofer IIS, Erlangen, Germany
Representatives from IRT, Munich, Germany;
BBC, UK

Digital processing seemed to solve the problem of reduced audio quality due to copying forever, but today several stages of the transmission chain include perceptual audio coding schemes such as MPEG-2 Layer-2 and Layer-3, MPEG-4 AAC, AC-3, and ATRAC. This mixture of different coding formats leads to significant accumulation of coder distortions in each encoding step (so-called tandem coding) and associated serious loss of subjective audio quality. Additional problems can be caused by processing such as equalization and compression.

In this workshop typical usage scenarios will be presented and discussed. Results from listening tests conducted by ITU and EBU over the last few years will be summarized. Different solutions to reduce tandem artifacts and to increase robustness will be presented.

**Workshop 14 Tuesday, October 8 9:00 am–12:00 noon
Room 408A**

**THE APPLICATION OF MULTICHANNEL SOUND FOR-
MATS IN VEHICLES**

Chair: **Richard Stroud**, Stroud Audio, Inc., Kokomo, IN, USA

Panelists: *David Clark*
Neal House
Mark Ziemba

Multichannel audio has made its way into the car. Many more applications of this technology will doubtless appear in new vehicles. Workshop presenters will discuss vehicular multichannel audio installation design goals, system development, available source material, recording considerations, and evaluation methodologies. Demonstration vehicles for multichannel listening will be present.

**Special Event
AUDIO POST PRODUCTION IN 24p HDTV
AND RELATED FORMATS**

Tuesday, October 8, 10:00 am–11:30 am
Room 403A

Moderator: **Dennis Weinreich**, Videosonics Cinema Sound, London, UK

Panelists: *Colin Broad*, CB Electronics
Doug Ford, Skywalker Sound
Robert Predovich, Soundmaster Group
Scott Wood, Digidesign/Avid Technologies

The challenge of new technologies has significant impact on our professional workflow. As cutting-edge sound people we must be able to understand how best to utilize these technologies to benefit most from them—both technically and creatively. We need to understand what the technology is attempting to address and how we should use it to our best advantage. Some of us will try to find where it fits into our existing work flow. Others will see new creative opportunities. Many advances require complete rethinking of how we approach our job in order to benefit from them.

HD 24p is one such advance that will have significant impact on how we do our jobs. Currently, the lighter and seemingly less complex production tools cause filmmakers to think that HD will give a better look and sound to production with less effort. In time, after the advantages are weighed against the disadvantages, we will probably find ourselves working in ways we currently cannot foresee.

Our panel will discuss the present position on working with HD and 24p, with particular emphasis on Audio Post Production. We will look at issues from the production floor to sound editorial to the dubbing stage. The panel will be made up of contributors from Post facilities, manufacturers' representatives, and professionals in the field who can share their HD 24p experiences and enlighten us about this new and interesting technology.

This event is being presented by the APRS (UK) and co-hosted by the AES.

**Education Event
STUDENT DELEGATE ASSEMBLY 2**

Tuesday, October 8, 10:00 am–11:30 pm
Room 402A

Chair: **Scott Cannon**, Stanford University Student Section, Stanford, CA, USA

Vice Chair: **Dell Harris**, Hampton University Student Section, Hampton, VA, USA

At this meeting the SDA will elect new officers. One vote will be cast by the designated representative from each recognized AES student section in the North and Latin America Regions. Judges' comments and awards will be presented for the Recording Competitions and the Student Poster Session. Plans for future student activities at local, regional, and international levels will be summarized.

**Special Event
ROAD WARRIORS PANEL: TRUE STORIES FROM
THE FRONT LINES OF SOUND REINFORCEMENT**

Tuesday, October 8, 12:30 pm – 2:00 pm
Room 403A

Introduction: **Paul Gallo**, founder of Pro Sound News, New York, NY, USA

Moderators: **Steve Harvey**, **Clive Young**, Pro Sound News, New York, NY, USA

Participants: *Greg Dean*
Kirk Kelsey
David Morgan

Introduced by Paul Gallo, publishing veteran and long-time friend to the live sound industry, this freewheeling panel of ➡

touring professionals will cover the latest trends, techniques, and tools that shape modern sound reinforcement. Moderators Steve Harvey and Clive Young will guide the all-star panel across subjects ranging from gear to gossip, in what promises to be an entertaining and educational 90 minutes, with the engineers on the business side of the microphone, saying something besides “testing” and “check” for a change!

Special Event ORGAN RECITAL

Tuesday, October 8, 1:00 pm–2:00 pm
Cathedral of Our Lady of the Angels
555 W. Temple Street
Los Angeles

Organist: **Graham Blyth**

Visit the newly completed Dobson Organ at the new Cathedral of Our Lady of the Angels in downtown Los Angeles. The specially organized event will feature a lunchtime performance by Graham Blyth, the AES’ resident organ recitalist. His program will include: Antonio Soler’s “The Emperor’s Fanfare,” Cesar Franck’s “Chorale No. 3 in A Minor,” Olivier Messiaen’s “Dieu Parmi Nous,” plus works by J. S. Bach and Louis Vierne. Mr. Blyth’s performance will be preceded by a series of presentations from the design and installation team for the Dobson Organ, and will be followed by a brief lecture from Dennis Paoletti and John Prohs, the Cathedral’s acoustical consultants.

Graham Blyth received his early musical training as a Junior Exhibitioner at Trinity College of Music in London, England. Subsequently at Bristol University, he took up conducting, performing Bach’s St. Matthew Passion before he was 21. He holds diplomas in Organ Performance from the Royal College of Organists, The Royal College of Music, and the Trinity College of Music. In the late 1980s he renewed his studies with Sulemita Aronowsky for piano and with Robert Munns for organ.

Mr. Blyth made his international debut with an organ recital at St. Thomas Church, New York, in 1993, and since then has played in San Francisco (Grace Cathedral), Los Angeles, Amsterdam, Copenhagen, Munich, and Paris (Madeleine Church). He gives numerous concerts each year, principally as an organist and a pianist, but also as a conductor and a harpsichord player.

Mr. Blyth is founder and technical director of Soundcraft. He divides his time between his main career as a designer of professional audio equipment and organ-related activities. He has lived in Wantage, Oxfordshire, U.K., since 1984, where he is currently artistic director of the Wantage Chamber Concerts and director of the Wantage Festival of Arts. He is also founder and conductor of the Challow Chamber Singers & Players. He is involved with Musicom Ltd., a British company at the leading edge of the pipe organ control system and digital pipe synthesis design. He also acts as tonal consultant to the Saville Organ Company and is recognized as one of the leading voicers of digital pipe modeling systems.

**Session N Tuesday, October 8 1:00 pm–4:00 pm
Room 404AB**

PSYCHOACOUSTICS, PART 2

Chair: **Nick Zacharov**, Nokia Research Center,
Tampere, Finland

1:00 pm

N-1 Industry Evaluation of In-Band On-Channel Digital Audio Broadcast Systems—David Wilson,

Consumer Electronics Association, Arlington, VA,
USA

The National Radio Systems Committee’s testing and evaluation program for in-band on-channel digital audio broadcast systems is described. The results of laboratory and field tests performed during 2001 on iBiquity Digital Corporation’s AM-band and FM-band IBOC DAB systems are reported. The conclusions drawn from the laboratory and field test results are also reported, and implications for the future are discussed.

Convention Paper 5709

1:30 pm

N-2 Comparisons of De Facto and MPEG Standard Audio Codecs in Sound Quality—Eunmi L. Oh,

JungHoe Kim, Samsung Advanced Institute of
Technology, Suwon, Korea

The current paper is concerned with assessing the sound quality of various audio codecs including ubiquitous de facto standards. Formal listening tests were conducted based on the ITU-R Recommendation BS.1116 in order to provide an objective measure of sound quality. Codecs tested included de facto standards that were commercially and noncommercially available and the MPEG general audio. In addition, our recently updated codec was tested. Test items consisted of usual MPEG test sequences and other sensitive sound excerpts at the bit rate of 64- and 96-kb/s stereo. Experimental results show that the sound quality of our newest codec out paces that of most of other codecs.

Convention Paper 5710

2:00 pm

N-3 Evaluating Digital Audio Artifacts with PEAQ—

Eric Benjamin, Dolby Laboratories, San Francisco, CA,
USA

Portions of the digital audio chain have been incrementally improved by development, such that objective specifications indicate a very high level of performance. Subjective reviews of these components often claim to observe substantial differences between products. This investigation uses the tool PEAQ (perceptual evaluation of audio quality) to measure the audio degradation caused by analog to digital converters, digital to analog converters, and sample rate conversion, and also to measure the minute incremental changes of codec audio quality that accompany very small changes in data rate.

Convention Paper 5711

2:30 pm

N-4 The Use of Head-and-Torso Models for Improved Spatial Sound Synthesis—Ralph Algazi,

Richard O. Duda, Dennis M. Thompson, University of
California, Davis, CA, USA

This paper concerns the use of a simple head-and-torso model to correct deficiencies in the low-frequency behavior of experimentally measured head-related transfer functions (HRTFs). This so-called snowman model consists of a spherical head located above a spherical torso. In addition to providing improved low-frequency response for music reproduction, the model provides the major low-frequency localization cues, including cues for low-elevation as well as high-elevation sources. The model HRTF and the measured HRTF can be easily combined by using the phase response of the model at all frequencies and by cross-fading between the dB magni-

tude responses of the model and the measurements. For efficient implementation, the exact snowman HRTF is approximated by two time delays and two first-order IIR filters. Because the poles are independent of the location of the virtual source, this supports a simple real-time implementation that allows for arbitrarily rapid head and source motion.

Convention Paper 5712

3:00 pm

N-5 Three-Dimensional Headphone Sound Reproduction Based on Active Noise Cancellation—*Daniël Schobben, Ronald Aarts*, Philips Research Laboratories, Eindhoven, The Netherlands

Headphone signal processing systems that are commercially available today are not optimized for the individual listener. This results in large localization errors for most listeners. In this paper a system is introduced that requires a one-time calibration procedure, which can be carried out conveniently by the listener. This system consists of conventional headphones into which small microphones have been mounted. An active noise cancellation method is used to achieve a sound reproduction via headphones, which is as close as possible to a reference loudspeaker setup. The active noise cancellation system is based on adaptive filters that are implemented in the frequency domain.

Convention Paper 5713

3:30 pm

N-6 On the Design of Canonical Sound Localization Environments—*Eric J. Angel, Ralph Algazi, Richard O. Duda*, University of California, Davis, CA, USA

This paper addresses the design of virtual auditory spaces that optimize the localization of sound sources under engineering constraints. Such a design incorporates some critical cues commonly provided by rooms and by head motion. Different designs are evaluated by psychoacoustics tests with several subjects. Localization accuracy is measured by the azimuth and elevation errors and the front/back confusion rate. We present a methodology and results for some simple canonical environments that optimize the localization of sounds.

Convention Paper 5714

Workshop 15
Tuesday, October 8
Room 408A

1:00 pm–4:00 pm

CODING OF SPATIAL AUDIO—YESTERDAY, TODAY, TOMORROW

Chair: **Christof Faller**, Agere Systems, Murray Hill, NJ, USA



113th Convention Papers and CD-ROM

Convention Papers of many of the presentations given at the 113th Convention and a CD-ROM containing the 113th Convention Papers are available from the Audio Engineering Society. Prices follow:

113th CONVENTION PAPERS (single copy)

Member: \$ 5. (USA) or £3.50 (UK)

Nonmember: \$ 10. (USA) or £7.00 (UK)

113th COMPLETE SET (78 papers)

Member: \$100. (USA) or £ 66. (UK)

Nonmember: \$150. (USA) or £ 99. (UK)

113th CD-ROM (78 papers)

Member: \$100. (USA) or £ 66. (UK)

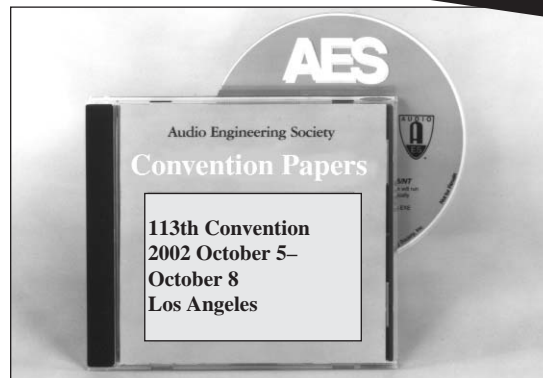
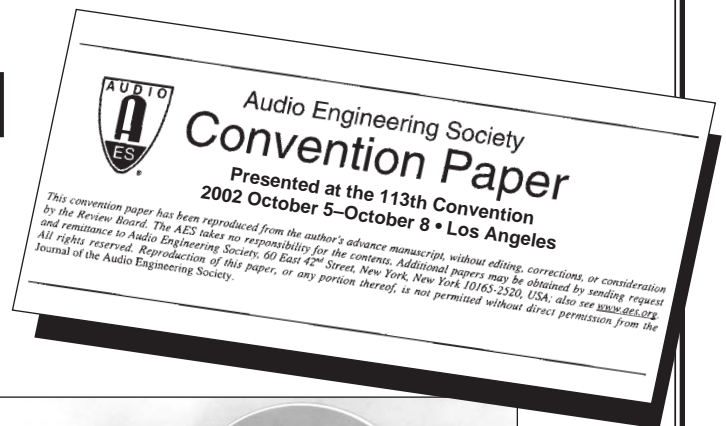
Nonmember: \$150. (USA) or £ 99. (UK)

113th CONVENTION PAPER SET AND CD-ROM (78 papers, each)

Member: \$150. (USA) or £ 99. (UK)

Nonmember: \$225. (USA) or £149. (UK)

A complete list of the 113th Convention Papers with a convenient order form is inserted in this issue. In addition, audio cassettes of the papers, workshops, and opening ceremonies are available from Mobiltape, and may be purchased by using the enclosed order form.



Panelists: *Frank Baumgarte*, Agere Systems, Murray Hill, NJ, USA
Mark Davis, Dolby, San Francisco, CA, USA
Martin Dietz, Coding Technologies, Nürnberg, Germany
Gerald Schuller, *Thomas Sporer*, Fraunhofer Arbeitsgruppe Elektronische Medientechnologie AEMT; and Ilmenau University, Ilmenau, Germany

Low bit-rate audio coding has become ubiquitous in many of today's audio systems, most of which are able to handle stereo or multichannel audio signals. A closer examination of issues related to the compression of spatial audio reveals a number of complex perceptual and coding issues that need to be considered in order to achieve optimum coder performance. While a wealth of different approaches have evolved over the recent decade, there is no single technique serving all purposes equally well. This workshop will provide a review of the principles and practical approaches for coding of spatial audio, and a discussion of the different dimensions of the tradeoff bit rate vs. spatial quality, and characterizes commonly used coders w.r.t. these aspects.

Introduction, by *Christof Fallner*

Spatial Perception, by *Thomas Sporer*

This talk will review spatial perception that is relevant for reproduction and coding of spatial audio.

Historical Development of Spatial Audio Reproduction and Transmission, by *Mark Davis*

The historical development of stereophonic and multichannel audio reproduction will be presented. It includes a description of early experiments for multiloudspeaker and two-loudspeaker stereophony, and the early techniques for pseudo stereophony, etc. For the historical development of spatial audio transmission, sum/difference transmission for FM radio stereophony will be described, and new techniques based on perceptual audio coding will be briefly mentioned.

Coding of Stereophonic Signals, by *Gerald Schuller*

This talk will focus on traditional ways of encoding of stereo audio signals such as separate coding (fails for certain signals), L/R coding and S/D coding and BMLD consideration, redundancy reduction/irrelevancy reduction, intensity stereo, and matrices as used in popular coders.

Parametric Stereo for Very Low Bit-Rate Stereo Coding, by *Martin Dietz*

This talk will describe the idea of parametric stereo coding: a mono signal, downmixed from the stereo source and compressed by powerful modern coding algorithms, is converted back into stereophonic sound by means of a coded parametric description of the spatial properties of the original signal.

Future Directions for Coding of Spatial Audio, by *Frank Baumgarte*

The technique and philosophy behind binaural cue coding (BCC) and other potential forward-looking technologies for coding of spatial audio will be described.

Education Event STUDENT PAPER SESSION

Tuesday, October 8, 1:00 pm–3:00 pm
Room 406AB

The paper session provides opportunity for students to present their research work on all subjects covered by the AES. Professionals from the audio industry will be invited to discuss the students' projects.

Special Event THE VIRTUAL STUDIO: DAW-BASED RECORDING

Tuesday, October 8, 2:30 pm–5:30 pm
Room 403A

Moderator: **Frank Wells**, Pro Sound News, New York, NY, USA

Panelists: *David Channing*, Independent Engineer
Lynn Fuston, 3D Audio
Jim Kaiser, Mastermix
Nathaniel Kunkel, Independent Engineer
Ralf Schluezen, TC Works
Rich Tozzoli, Independent Engineer

Everyday, more of the recording and post production studio infrastructure collapses into the frame of a computer. This two-part event will focus on technical concerns unique to DAW-based recording, as well as present power user application techniques.

Bits and Bytes and Bottlenecks

Inside a DAW, music exists as data, but a special kind of data that is sensitive to issues like word length, processor resolution, headroom and delay compensation. These aren't the sexy parts of the feature set, to be sure, but attention to the details can noticeably and measurably improve the final product. Our presentations will focus on some of the core issues that require attention.

Beyond the Basics

While a DAW can function as a simple recorder or editor, with plug-ins and other DSP tricks at their disposal, a DAW can also manipulate audio in sophisticated ways that were hitherto impossible. Power users will present advanced, cross platform techniques for the DAW user.

Moderator Frank Wells came to the recording industry after years of work as a radio broadcast engineer. Following nearly a decade as Chief of Technical Services for Masterfonics, Nashville, he defected to the world of trade journalism, first as the founding editor of *Audio Media, USA* and currently editor of *Pro Sound News* and executive editor of *Surround Professional*. He is also the current chair of the AES Nashville section.

David Channing has engineered and/or edited two Elton John records, the new Duncan Sheik record, Jewel's "Spirit" and projects with Enrique Iglesias, Melissa Etheridge, Shawn Colvin, and many more.

Lynn Fuston is owner/chief engineer of 3D Audio, with engineering credits that include Amy Grant, Mark O'Connor, DC Talk, Lee Greenwood, and the Newsboys. He has also produced a series of microphone and microphone pre-amplifier comparison CDs, and will soon release an A/D comparison disc.

Jim Kaiser is VP Central Region for the AES, and Director of Technology at Mastermix, Nashville. His background includes session engineering and technical consulting, including full facility technical design.

Nathaniel Kunkel is known for a broad body of recording work that includes CSNY, Little Feat, Linda Ronstadt and Lyle Lovett, including extensive work in 5.1 for DVD-A and Super Audio CD (the Graham Nash project, *Songs for Survivors*, and the 5.1 remix of James Taylor's *J.T.*).

Ralf Schluezen is CEO of TC Works, the software division of TC Electronic, based in Hamburg, Germany. He has extensive experience in plug-in design, including a stint with Steinberg before joining TC.

Rich Tozzoli is a New York-based engineer whose long list of credits includes extensive surround work with the likes of David Bowie, Foghat and Average White Band, along with numerous live classical recordings. He is also a contributing editor with *Surround Professional*.

BLUETOOTH

and wireless networking

a primer for audio engineers

Wireless networking, using radio frequency links instead of wires, is becoming increasingly popular as a means of connecting digital devices and computers. It provides greater flexibility and mobility to the user and does away with those tangled strands of wires. And wireless networks can be integrated with wired systems, thereby acting as extensions of existing networks. Audio information is increasingly transferred over networks instead of dedicated digital interfaces or analog cabling, making it important that audio engineers have an appreciation of the issues involved.

Bluetooth is a technology for wireless personal area networks (WPANs) that has been widely publicized recently, but it is only one of a number of op-

tions for wireless data communication. This article puts Bluetooth into context and describes wireless data communications in relation to audio systems. This field is changing fast like most aspects of computing, making it hard for standardization to keep pace with technology development. A number of wireless networking standards have been or are being standardized by the IEEE, but alternative solutions also exist that overlap with or differ from IEEE standards. As a rule it seems that IEEE has attempted where possible to adopt commercial technology in response to calls for suitable solutions.

WIRELESS PERSONAL AREA NETWORKS

WPANs connect desktop digital devices, usually over short dis-

tances within a so-called personal operating space of 10 m enveloping the person. WPANs are intended mainly for indoor and fixed outdoor applications and operate at data rates of up to about 1 Mbit/s. They are a useful next step from infrared communications. Bluetooth is the prime example of a WPAN. Home RF is potentially similar, but it sits somewhere between a PAN and a LAN (local area network). Fig. 1 shows an example of the interrelationship between PANs, LANs, and WANs. WPANs are broadly covered in IEEE standards work by the 802.15 groups. The primary concerns when designing successful WPAN technology are power consumption, simplicity, and product size. ➔

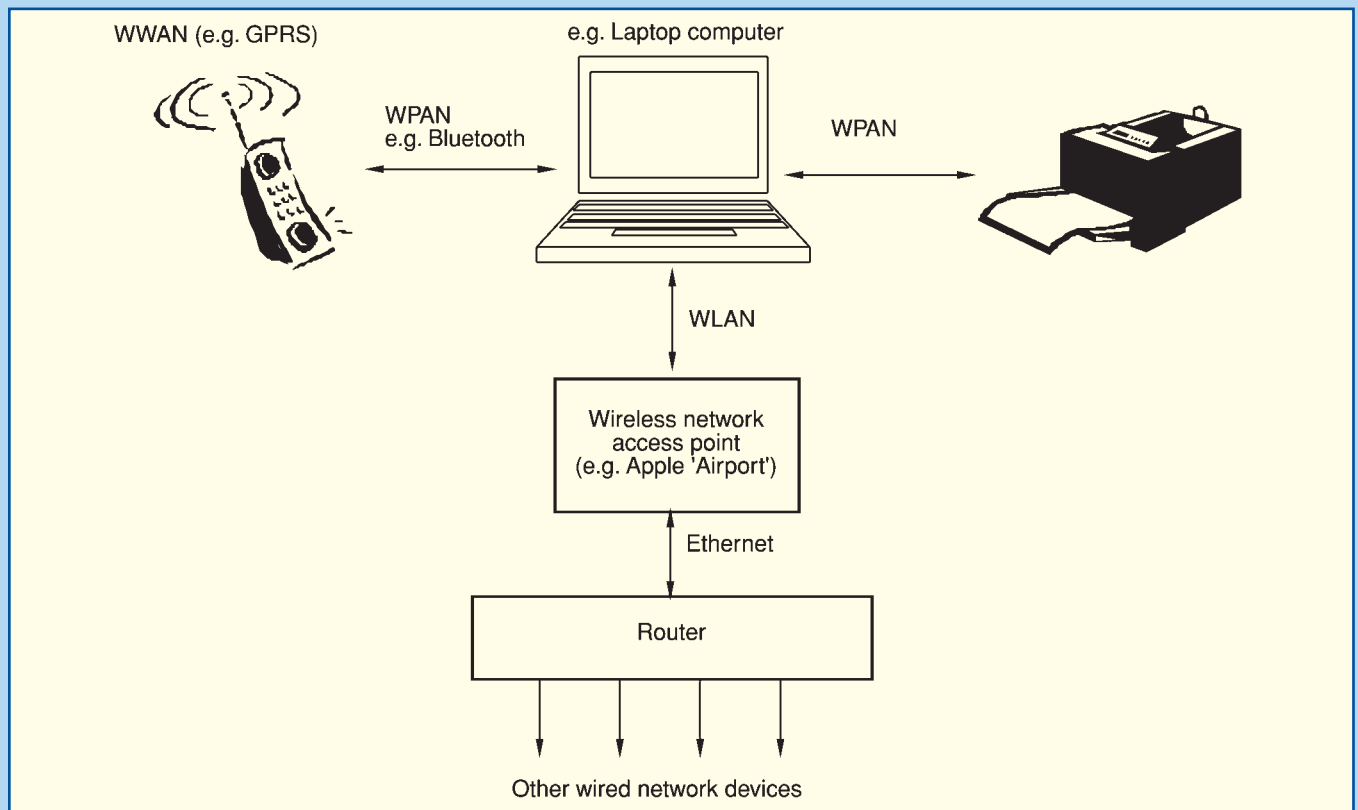


Fig. 1. Example of WPAN, WLAN, and WWAN interacting.

WIRELESS LOCAL AREA NETWORKS

WLANs are used over larger distances, up to 500 m, in areas where people may move around with digital devices but usually remain stationary while using the devices. They typically operate around hot-spot access points, for example in offices, homes, and schools. Data rates are typically higher than WPANs, between 2 and about 50 Mbit/s. WLANs are covered in IEEE standards activity by the 802.11 groups. Typically above this data rate are WMANs (wireless metropolitan area networks) that interconnect LANs over large regions.

WIRELESS WIDE AREA NETWORKS

WWANs have more to do with global telecommunications systems, for example mobile phone operators using GPRS (General Packet Radio Service) equipment, and are mainly designed for applications in which people move around while using devices. They are not featured in this article as they have little to do with potential audio applications in consumer and professional environments.

General radio frequency issues

Wireless networks typically use low-power, spread-spectrum, and frequency-hopping techniques to ensure robust communications in the face of interference and fading. They typically use coding techniques to make the

data signal look like noise or very short-term impulsive interference to a narrow-band receiver in the radio frequency (RF) domain, as shown in Fig. 2. Most current systems operate in a license-free region of the RF spectrum between 2.4 and 5 GHz, relying principally on the low-power and interference-avoiding mechanisms of systems to operate satisfactorily in this free-for-all region.

The 2.4-GHz band is becoming increasingly crowded owing to Bluetooth devices, wireless networks, and telephones, so there can be significant RF interference in this band. The 5-GHz band is somewhat quieter in the United States, but signals have a much shorter range at this frequency and are more easily obstructed by walls, doors, and other objects. Communication therefore tends to be over line-of-sight paths or short distances. Some chipsets are appearing that will operate in both bands for IEEE 802.11 standards. The issue of interference is covered in greater detail later in this article.

EXAMPLES OF WPAN TECHNOLOGY

Bluetooth operates in the 2.4-GHz band at data rates up to 1 Mbit/s over distances of up to 10 m or 100 m depending on the implementation. A relatively simple binary FM modulation method is used because Bluetooth has to be implementable in basic devices. Communication between devices can be either point-to-point or point-to-

multipoint. In the latter case one device acts as master, synchronizing, controlling, and sharing the channel with as many as seven slaves in a so-called piconet of up to eight devices. Frequency hopping between 23 or 79 RF channels is employed to increase robustness, and data packets are each transmitted on a different hop frequency. The IEEE 802.15.1 WPAN standard is based on Bluetooth.

The primary modes for data communication are ACL (asynchronous connectionless) links and SCO (synchronous connection-oriented) links. The latter are limited to a data rate of 64 kbit/s each, and up to three concurrent SCO links are allowed per Bluetooth channel. They are set up between a master and a single Bluetooth device at a time for time-critical purposes such as voice audio. SCO links are established as real-time links between two points for the duration of the connection. Packets are never retransmitted, owing to their real-time nature. In the time remaining within each transmission slot the master can set up an asynchronous communication with one or more slaves. ACL communications operate rather like conventional packet-switched networks, with the packet header being used to indicate the destination on the piconet.

It is possible to operate either one asynchronous and three synchronous voice channels (bidirectional) or one channel that handles both asynchronous data and synchronous voice

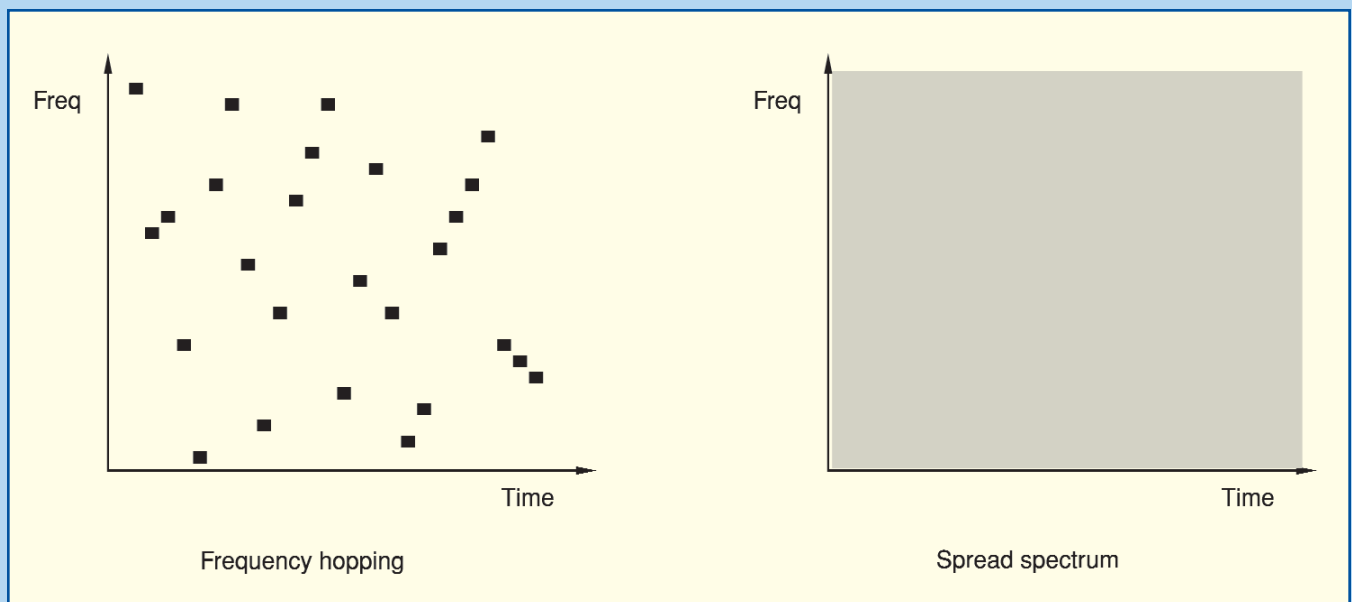


Fig. 2. Wireless networks use low-power, spread-spectrum, and frequency-hopping techniques for robust communications.

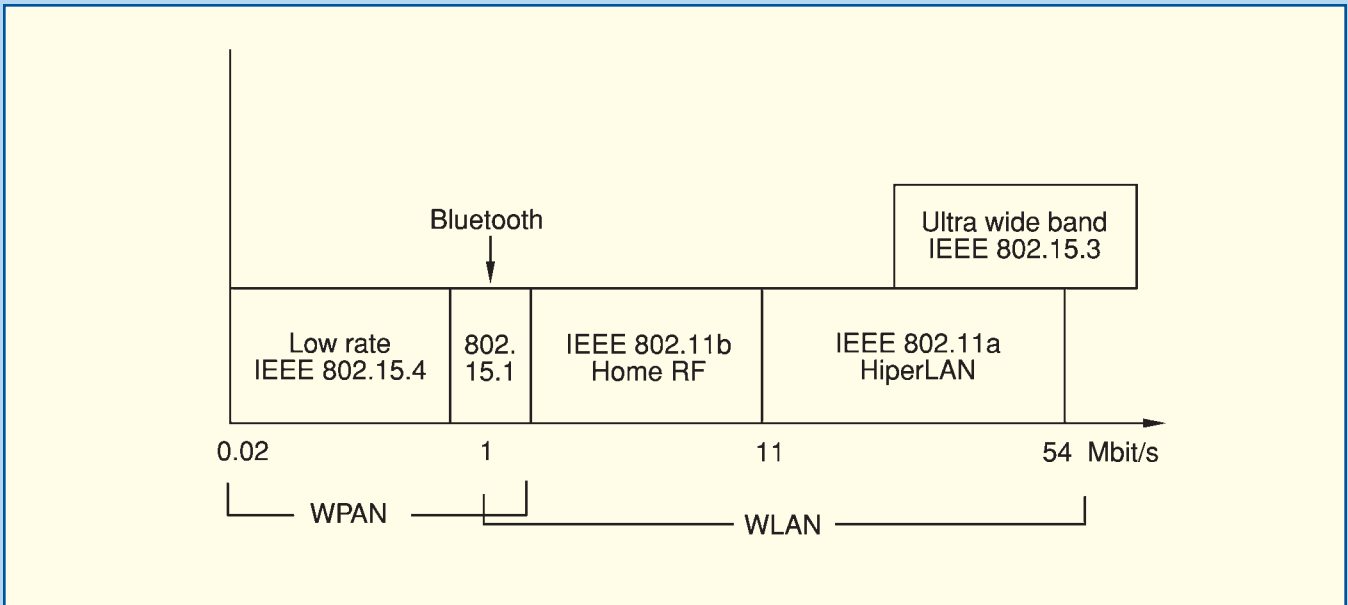


Fig. 3. Data rates of common wireless technologies.

(this is essentially made up of single SCO packets with two parts). The maximum asynchronous data rate is 723.2 kbit/s in one direction, with a return rate of 57.6 kbit/s, or 433.3 kbit/s symmetrically.

In addition to the Bluetooth Core Specification that describes the principles of communication, the Bluetooth SIG (special interest group) publishes a number of profile documents that describe specific approaches to data transfer for particular applications. The documents of primary interest to audio engineers are the Generic Audio/Video Distribution Profile (GAVDP) and the Advanced Audio Distribution Profile (A2DP). These are discussed later in this article.

The long-term future of Bluetooth is not certain, but the same can be said of virtually any data communications technology today. For example, Apple and a number of mobile-telephone manufacturers have significant investment in Bluetooth, and this may help to ensure its future.

Many commentators appear enthusiastic about the market potential for the technology. It is a relatively low-rate system but is intended for low-cost, low-power implementation in highly portable devices. Some commentators cite the relative advantages of broadband technology such as UWB (ultra-wide band) that may ultimately become more useful in high-rate multimedia environments. Some concern has also been expressed about the

interoperability of Bluetooth devices, as peer-to-peer networking interoperability is apparently not mandatory for device certification. Success may therefore come down to a question of which Bluetooth devices actually work together and at what level of sophistication.

High-rate and ultrawide band systems

The IEEE 802.15.3 group is studying high-rate applications above 20 Mbit/s. These could ultimately be up to 100 times faster than Bluetooth. UWB is also being developed for applications requiring very high data rates. UWB uses very wide band transmissions at extremely low power, requiring receivers and transmitters to be tightly synchronized. The U.S. Federal Communications Commission (FCC) has approved UWB within certain low-power limits, but other countries have not yet done the same. IEEE 802.15 Study Group 3a is studying alternative physical layer options for high-rate WPANs.

Low-rate, low-power consumption systems

IEEE 802.15.4 is devising a lower rate standard for simple devices, such as joysticks, that do not need high data rates but need small size and long battery life. Data rates of 20, 40, and 250 kbit/s are targeted for use within RF bands at 2.4 GHz, 915, and 868 MHz. This clearly lies below the typical

Bluetooth application range of data rates, whereas high rate and UWB lie above the Bluetooth range, as shown in Fig. 3.

EXAMPLES OF WLAN TECHNOLOGY

Broadband WLAN technology is developing so fast that the regulatory and commercial situation changes almost weekly. However, the following is an attempt to summarize the current situation as clearly as possible.

IEEE 802.11

IEEE 802.11a and IEEE 802.11b are wireless LAN standards and the basis of wireless Ethernet or WiFi. WiFi is a mark for compatible products awarded by the Wireless Ethernet Compatibility Alliance (WECA). Apple uses this technology in its AirPort products, as do many of the consumer wireless network systems currently marketed in computer stores.

The two standards define different physical layers, where 802.11b transmits at 2.4 GHz and 802.11a transmits at 5 GHz. The standards do not interoperate directly with each other, although bridges between them are possible and chipsets are appearing that operate in either mode. 802.11 deals with the physical and MAC (medium access control) layers of networking protocol, but the remaining layers are virtually identical to Ethernet, which is IEEE 802.3. In network terminology layers are different levels in the

USEFUL WEB SITES

Bluetooth specifications:
www.bluetooth.com/dev/specifications.asp

HiperLAN:
www.hiperlan2.com

HomeRF:
www.homerf.org

IEEE 802.11:
grouper.ieee.org/groups/802/11/

IEEE 802.15:
grouper.ieee.org/groups/802/15/

Ultra wide band:
www.uwb.org

Wireless networking links:
www.wireless-nets.com/links.htm

dynamic frequency selection (DFS) and transmission power control (TPC) have been required for conformity with European requirements, and these are now virtually completed for 802.11a (this project was originally coded 802.11h). However, as discussed in the next section, full ETSI approval of wireless networking in the 5-GHz band does appear to be feasible, and this band is apparently being cleared in many countries for broadband data communications.

The range of 802.11 activities is wide and fast moving. See <http://grouper.ieee.org/groups/802/11/> for the latest information.

HiperLAN

HiperLAN operates in the 5-GHz band up to 54 Mbit/s and requires approximately 330 MHz of bandwidth. These rates and frequencies are the same as 802.11a, and at the physical layer the two are almost identical. HiperLAN has been developed by ETSI and the Broadband Radio Access Network (BRAN), whose founding members are Tenovis (Bosch), Dell, Ericsson, Nokia, Telia, and Texas Instruments. Typically, the operating ranges are 30 m indoors and 150 m outdoors. In Europe specific bands from 5.15 to 5.35 GHz and 5.470 to 5.725 GHz seem to have been approved already for dedicated use by HiperLAN. In the U.S. the 5.15–5.35-GHz and 5.725–5.825-GHz bands are unlicensed and usable.

The main difference between HiperLAN and 802.11a is at the media access (MAC) level. Whereas 802.11a is an extension of wireless Ethernet, primarily based on contention mechanisms and packet switching, HiperLAN is claimed to offer connection-oriented communication and is compatible with other circuit-switched network standards such as ATM. As a result it offers quality-of-service (QoS) guarantees and therefore will be useful for real-time streaming applications.

HomeRF

HomeRF is yet another competing technology in the 2.4-GHz band. It offers a data rate of 10 Mbit/s and combines cordless telephone links (up to eight concurrently), wireless networking, and data streaming for home entertainment products.

somewhat arcane hierarchy of data communications, with the application at the top and the physical medium at the bottom. As with cell phones, users can roam between wireless access points (APs), using the beacon of the AP as a judge of signal strength. Collision avoidance and carrier detection are used for medium access, rather like Ethernet, and authentication or full encryption can be used for security purposes.

Both FHSS (frequency hopping spread spectrum) and DSSS (direct sequence spread spectrum) approaches are specified for 1- and 2-Mbit/s rates, but only DSSS is allowed for 11-Mbit/s communications. 5.5 Mbit/s is also allowed and even higher bit-rate standards were approved in 1999. In fact it appears that most developers actually went for DSSS even at the lower rates, for future compatibility with the higher rates. Generally, the higher the radio frequency (and therefore the greater the potential for high data rates), the shorter the range that may

be covered. For example the 802.11a standard can operate at 54 Mbit/s but only within a 50-meter range. Alternatively, 802.11b operates at a maximum speed of 11 Mbit/s but can cover 300 meters outdoors (100 meters indoors). As with most packet-switched RF networks, contention, interference, and overheads can reduce the realistic data rate closer to half the maximum.

Contrary to what may have been anticipated, 802.11b is already in wide use but 802.11a products are only just beginning to appear. In the UK an internal agreement is being formulated to allow 802.11a to operate in a limited fashion in the region from 5.15 to 5.25 GHz. This allows four access points as opposed to the eight possible if the system occupies the full band from 5.15 to 5.35 GHz. Full approval from the European Telecommunications Standards Institute (ETSI) is not yet finalized because of military and government use of the 5-GHz band. But other countries are striking individual agreements rather like the UK. Dy-

Up to eight simultaneous real-time streaming sessions are possible, either two-way, multicast, or receive only. Priority is given to real-time voice communications over streaming applications as the default condition. Asynchronous packets take last place in the queue. The IEEE describes HomeRF as a sort of trimmed-down 802.11.

RF INTERFERENCE ISSUES

Interference is a particular problem in the crowded 2.4-GHz band, which is license free in most parts of the world. Industrial, scientific, and medical applications (ISM) use it, and microwave ovens are the primary culprits when it comes to interference.

In order to limit the effects of interference, frequency hopping spread spectrum (FHSS) techniques are used by Bluetooth and HomeRF. Hopping reduces channel efficiency in favor of robustness and is simpler to implement than the DSSS techniques described below. Such techniques usually employ at least 75 frequencies with a maximum dwell time per frequency of 400 ms. This ensures that communication does not spend too long on each frequency and the signal looks like random impulsive interference to a narrow-band receiver. Both transmitter and receiver follow the same pseudorandom pattern of hops, and adaptive hopping can avoid frequencies that are known to be blocked. Typical FHSS products occupy about 1 MHz of bandwidth. They achieve higher power within each frequency band at any one time and may therefore result in a better instantaneous RF signal-to-noise ratio than other spread-spectrum techniques.

The Bluetooth hopping rate is quite fast (1,600 times per second) compared with some other systems. Per second, in any one band, it is said that one is more likely to encounter an 802.11b transmission than a Bluetooth signal. There is some concern over the collocation of Bluetooth and 802.11b devices, as Bluetooth transmitters have been shown to reduce the performance of the other network considerably if placed close to a receiver. Bluetooth has limited error handling and retry mechanisms beyond the frequency-hopping physical layer. So it is not particularly sophisticated for robust wire-

less LANs, for which it was not designed, intended instead as a simple and cheap short-distance cable-replacement technology.

The direct sequence spread spectrum (DSSS) approach, as used by 802.11b for example, uses more bandwidth than strictly required by the data rate. Data are coded onto chips (simply a pattern of data bits) that have redundant portions spread over the frequencies in the band. The data are exclusively ordered within an 11-bit Barker code (a pseudorandom sequence), the bits of which make up the chips, and there are usually an integral number of chips per bit. These are modulated onto the carrier using differential phase-shift keying (DPSK), so a narrowband receiver perceives this signal as low-level noise. Even if some parts are lost during transmission owing to interference, error correction can be used to recover the data. The so-called spreading ratio is related to the degree of redundancy employed (number of chips per bit); the spreading ratio of most wireless LANs is around eight. This ratio makes the use of the band reasonably efficient at the expense of higher robustness. Sometimes the band is split between more than one network; a maximum of three networks is possible in typical current implementations. Typical current DSSS products occupy 20 to 22 MHz of bandwidth no matter what the data rate. The power within any one band is relatively low, making the RF signal potentially less robust than with frequency-hopping techniques, but performance in practice depends on the coding scheme and spreading ratio employed.

BLUETOOTH AUDIO

The basic audio functionality in the Bluetooth Core Specification is really only suitable for telecommunications applications. The bandwidth is typically limited to 4 kHz for telephony, and the sampling rate is correspondingly 8 kHz. Relevant standards are ITU-T P.71, G.711, 712, and 714. The 64-kbit/s voice channels use μ law or A-law logarithmic PCM coding or CVSDM (continuous variable slope delta modulation). CVSDM is said to be preferable with respect to quality and robustness to errors. Audio error cor-

rection depends on which packet type carries the audio transmission: HV3 (High-quality Voice 3) packets have no forward error correction (FEC) and contain 3.75 ms of audio at 64 kbit/s, whereas HV1 and 2 packets have some error correction and contain shorter durations of audio, 1.25 and 2.5 ms respectively.

Bluetooth audio streaming using ACL packets

New draft profiles have been created for audio streaming using asynchronous (ACL) packets that can occupy the full remaining bit rate of Bluetooth (721 kbit/s, after the voice quality streams are taken into account). These use real-time protocol (RTP) for streaming. RTP was originally developed for managing streaming connections on asynchronous packet-switched networks such as parts of the Internet.

Either point-to-point or point-to-multipoint (such as one transmitter to multiple Bluetooth loudspeakers) are allowed. QoS is not guaranteed with ACL connections although RTP does provide for some real-time requirements provided that buffering is used at the receiver. Better QoS provision for streaming has been requested by the A/V working group of Bluetooth SIG in the next revision of the Bluetooth data-link specification. Although Bluetooth supports isochronous communication (the type required for clock-dependent processes) for streaming applications through the use of higher-level L2CAP (logical link control and adaptation protocol) connections, this is always only on a best-effort basis. The only truly synchronous reserved slots at the baseband level are for SCO packets, basically speech audio.

AVCTP is the Audio/Video Control Transport Protocol that can be used for conveying messages intended for controlling Bluetooth A/V devices. The AVDTP (Audio/Video Distribution Transport Protocol) uses L2CAP packets to transfer audio with connections established between a transmitter and a receiver. This protocol provides a mechanism for reporting QoS, optimizing the use of bandwidth, minimizing delays, and attaching time-stamp data to packets for synchronization ➔

purposes. The AVDTP document provides some advice in an appendix relating to the synchronization of devices either to each other or to a separate network clock. It also briefly mentions an approach to measuring timing jitter. Broadly, the approach relies on the timing mechanisms inherent in RTP and RTCP (real-time control protocol).

The A2DP (Advanced Audio Distribution Profile) describes a configuration of layers in the Bluetooth stack and higher application layers that can be used for the conveyance of audio. It was authored by the audio/video working group consisting of members from Sony, Toshiba, Nokia, Philips, Ericsson, and Matsushita.

Considering the bit rate, there is really no way that uncompressed high-quality audio can be carried, requiring that some form of low bit-rate coding be employed. Specified in A2DP is a mandatory audio codec, which is a low-complexity, subband codec (SBC), whereas other codec types (MPEG and ATRAC in the current version) are listed as optional. The SBC codec was developed for Bluetooth but based upon an earlier Philips system described by de Bont et al.¹ NonA2DP codec types can be accommodated, although the rubric is rather confusing in relation to this eventuality. This is supposed to be achieved either by upgrading the codec to optional status within A2DP (this requires the manufacturer to submit clear definitions of certain required characteristics) or by transcoding the audio data to SBC if the receiver does not support the decoding of the data type; this way interoperability is maintained as much as possible.

The profile document does not define anything in relation to nonA2DP codecs except that the vendor is supposed to use a Bluetooth Assigned Number to identify itself, and its parameters must be signalled within the standard packet headers. Audio may be encrypted for content protection or not, but this is application dependent. The mandatory subband codec should use

at least one of 44.1- or 48-kHz sampling frequencies, and other lower rates can be specified. The encoder should be able to handle at least mono and one stereo mode (such as dual channel, stereo, joint stereo) and the receiver should be able to decode all of these. Similar requirements exist for MPEG 1, Layer I, II or III, for MPEG 2 and 4-AAC, and for ATRAC codecs. MPEG codecs are also allowed variable bit-rate (VBR) encoding in the profile.

Quality and Robustness of Bluetooth Audio Streaming

For adequate audio quality the A2DP profile requires that the audio data rate be sufficiently lower than the available link data rate to allow for the retransmissions that will avoid packet loss. The margin allowed for this obviously depends on the robustness expected of the application in the face of interference or long distances. The profile limits the SBC bit rate to a maximum of 320 kbit/s for mono and 512 kbit/s for stereo. The maximum available bit rate is 721 kbit/s. The data overhead when one transmitter is communicating multiple ACL streams to different receivers can lower the overall bandwidth available for audio, hence the number of channels is tightly limited even with data compression. An alternative is to use broadcast mode in which the full data stream for all audio channels is broadcast to all receivers, requiring them to separate the channels themselves (they therefore need to have a means of channel identification). The maximum number of audio channels in this mode is seven, that is the maximum number of connections to a master device.

There are some problems with point-to-multipoint connection for audio using ACL and RTP because different retransmission rates will apply on each connection and possibly affect inter-channel synchronization. The result of this is timing differences between the audio channels or phase distortion. This

can be ameliorated using adequate buffering and resynchronization. Broadcasting gets around this problem, but packet loss can be encountered and not recovered because there is no dynamic retransmission method for broadcast mode. However, there is a fixed retransmission option for broadcast mode, which appears to act rather like a form of permanent redundancy, whereby packets are always retransmitted at the expense of a reduction in available bit rate on the channel. Floros et al.² found broadcast mode with no retransmissions only just acceptable for audio streaming owing to packet loss on the wireless link. One retransmission reduced audio data loss to 2 to 3% compared with 7 to 8%, but reduced the effective bit rate from 551 to 325 kbit/s. Because the losses were in compressed audio data, the resulting uncompressed audio was badly affected. They claim that retrieval mechanisms in the Bluetooth lower levels and application layer could be adapted to minimize the effects of packet loss on audio quality, but have some reservations about the use of the approach for synchronous compressed multichannel audio. They concluded that one could easily transmit stereo audio at 256 kbit/s per channel plus control information, within the bandwidth available, using two separate ACL links. The best data rate was obtained with DH5 packets, getting close to the upper limit of 721 kbit/s, but all DH packets have no forward error correction and so are more prone to data loss on the link. DM packets have forward error correction (FEC) and a correspondingly lower overall data rate.

Whether Bluetooth stands the test of time remains to be seen, but wireless technology for audio transmission will undoubtedly continue to expand. See the box on p. 982 for useful web sites with more information.

FURTHER READING

Specification of the Bluetooth System, version 1.1 (2001).

Bluetooth Audio/Video Working Group (2002), Advanced Audio Distribution Profile 0.95b.

Bluetooth Audio/Video Working Group (2002), Generic Audio/Video Distribution Profile 0.95b.

¹F. de Bont, M. Groenewegen, and W. Oomen, "A High Quality Audio-Coding System at 128 kb/s," presented at the 98th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 43, p. 388 (1995 May), preprint 3937.

²A. Floros, M. Koutroubas, N. Alexander-Tatlas, and J. Mourjopoulos, "A Study of Wireless Compressed Digital Audio Transmission," presented at the 112th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 50, p. 498 (2002 June), preprint 5516.

SECTIONS

We appreciate the assistance of the section secretaries in providing the information for the following reports.



Colorado Section addressed voice, stage and sound amplification. Panelists (left to right): Bob Mahoney, Patti Peterson, Bob Burnham and Peter Russel.

Vocal Amplification

The **Colorado Section** held its first meeting of the season on September 25, at the King Center of the Auraria Higher Education Center campus in Denver. Some 24 guests attended a panel discussion on voice, stage and sound amplification that covered a broad range of interests, including audience expectations for vocal performance in coming years.

Patti Peterson, professor of vocal studies at the University of Colorado, began the session by describing the aspects of vocal pedagogy that are common to both amplified and unamplified performance practices. She also talked about how vocal technique and tone production vary, depending on whether or not sound amplification is employed.

Peter Russell, president and general director of Opera Colorado, which in its early days was one of the first companies to practice amplified opera-in-the-round, spoke forcefully in favor of proscenium style presentations of unamplified opera. Russell characterized amplified vocal performance as appropriate for musical theater, but reiterated his position that conventionally produced opera must be unamplified to be true to the art form. Russell also said that since productions developed for in-the-round presentations have lit-

tle value to other (more traditional) companies, co-production with other opera companies for such performances is not economically feasible.

Bob Burnham then spoke to the group from a sound technician's point of view on the importance of establishing a warm and supportive rapport with the performing artists. Drawing on his 33 years of experience as a recording and sound reinforcement engineer, Burnham described several methods of miking and reinforcement and fielded a number of very pertinent questions from the audience.

Japan Tours

With 53 members and guests the **Japan Section** toured the new sound research facility at NHK Science & Technical Research Labs in Setagaya, Tokyo, on June 26.

Kimio Hamasaki of NHK welcomed visitors and talked briefly about NHK's objectives in undertaking the renovation, and about those sound research projects that are being undertaken now by the company. Members were able to visit the laboratories, talk with the heads, and take a look at some of the company's latest products. These include an extended dynamic range silicon diaphragm condenser microphone; a tiny moving coil microphone capable of picking up a

sound of insect chewing leaves, which was developed for the purpose of supporting microscopic scenes taken by a high resolution camera; a highly directional stereo gun microphone; and a sound image control system that uses loudspeaker arrays for a virtual studio.

Hamasaki concluded the tour with a demonstration of a 23-channel sound system developed for an experimental 4000-line HD-TV system, which impressed the group.

The section met again on August 22, for the annual review and to evaluate the fiscal year of 2001. During the evening, all the section's activities throughout the year, as well as the settlement of accounts were reported and approved. Other business included the election of new officers and the announcement of activity and budget projections for 2002.

Vic Goh

Audio Tech Seminar

The **Singapore Section** presented an Audio Technology Seminar in conjunction with Broadcast Asia 2002, Singapore's premier broadcast exhibition and conference on June 17, 2002. Entitled, "Future Trends for Digital Audio: Principles, the Practice & the Possibilities," the seminar drew 48 professionals from various areas of the audio industry and included delegates from 12 countries, including the USA, Japan and China.

The full-day seminar featured seven speakers, who covered topics relating to digital TV, digital audio, and loudspeaker systems. The first lecture focused on the work being done in the area of compression schemes and digital television. Mary Ann Seidler, international director of sales for Telos Systems/Omnia USA, talked →



Singapore discusses future trends for digital audio. *photo/Singapore Exhibition Services*

about some of the new schemes now available for audio compression. In particular, she focused on the advantages of MPEG4, such as its more efficient use of bandwidth.

Henk Mensinga, sales manager of TC Electronic of Denmark, then spoke about audio for DTV, covering the requirements of broadcasting for digital TV. Mensinga discussed the assortment of digital TV formats that are now being adopted by governments in different countries.

After a 15-minute break for refreshments, the discussion moved on to the implementation of digital audio. John Wigglesworth, vice president & director of Creative Services for Four Media Company Asia of Singapore, and Danial Shimiaei, chief digital systems engineer of Todd-Soundelux USA, talked about how international program producers are currently choosing to complete their programs in high-definition video with Dolby 5.1 soundtracks. Even though many of the final delivery formats for TV are still in analog and capable of supporting only Dolby Surround 2.0, the practice of exceeding current requirements permits a degree of future proofing that will ultimately prolong the shelf-life of these products.

Wigglesworth then talked about digital audio from a management point of view. He deciphered budget breakdowns and compared major Hollywood movie budgets with typical Asian budgets. Shimiaei covered the topic from an operational angle. He talked about the responsibilities involved in the numerous mixing

stages of major TV shows, including "The Sopranos," "Alias" and "Star Trek: Next Generation." Local delegates were able to get a glimpse into the workings of the companies that produce so many of the programs that make up many of the prime time programming in the region.

After a lunch break, two more seminars on digital audio focused on the developments in network audio. Yves Ansade, head of Product Development of Digigram, France, gave a presentation on using Ethernet cables for sound distribution that included an in-depth explanation of applications possible in a format that permits 24-bit/64-channel sound. Ansade also discussed EtherSound, a proprietary card system developed by Digigram that utilizes Ethernet cables for sound distribution.

The last presentation focused on a dual-loudspeaker project involving mLAN. Kunihiko Maeda, chief planner of the R&D Planning Section for Otari Inc System Engineering Division, USA, and Junichi Fujimori, project leader of Yamaha Corporation, Japan, described the efforts made by their respective companies to bring mLAN to fruition for sound applications. According to Maeda and Fujimori, using the FireWire/IEEE 1394 format single-cable connections within studio and stage environments is now possible. This eliminates cable clutter as well as permitting network connections that allow other settings and parameters to be transmitted via the single FireWire cable.

Lively Q&A sessions followed each

of the three pairs of presentations and AES mementos were given to the speakers by Robert Soo, section chair. The section is already preparing for its 2003 seminar to be held in conjunction with Broadcast Asia 2003.

Kenneth J. Delbridge

Seattle Symphony Center

Twenty-eight members and guests of the **Pacific Northwest Section** held a business meeting at Soundbridge, the Seattle Symphony's Music Discovery Center, an educational outreach section of the Benaroya Hall Complex on June 25.

Aurika Hays, section chair, opened the meeting with some remarks about the success of the preceding year and thanked members for all their hard work in helping to bring Soundbridge to life.

Ron Hyder and Rick Chinn (section committeemen) gave details of continuing improvements to the sound reinforcement system in Benaroya Hall S. Mark Taper Auditorium. The designers purchased a JBL Vertec line array system for left/right arrays and redid the built-in central cluster system, which is normally intended for basic use such as announcements, with JBL components to match the Vertec sound. They also installed a distributed loudspeaker array under the balcony, and purchased a Soundcraft Series 4 board. Unfortunately, the Seattle Symphony was in closed rehearsal at the time, so the group could not see the items firsthand. However, attendees were able to view the rehearsal on the Soundbridge plasma TV.

Bryan Stratton, manager of Soundbridge, then talked about how Soundbridge strives to be a hub for the Seattle Symphony's education and community programs. The program offers high-tech, hands-on interactive exhibits such as "Be a Virtual Conductor;" real instruments visitors can try; and a "Music Bar" featuring hundreds of musical pieces that can be called up for listening by visitors. The organization also offers programs and classes to educate people of all ages about music.

Soundbridge was originally slated to



Bryan Stratton (in suit) explains a kiosk to Pacific Northwest members. Aurika Hays and Gary Louie (inset) look on.

photos/Rick Smargiassi

occupy a small restaurant space. Stratton walked the group through the exhibits and demonstrated the Music Bar, which consisted of a row of PCs connected to a server. Each station has touchscreen controls, which allow the user to access information about hundreds of musical pieces, composers and styles, as well as the Seattle Symphony recordings.

There were also several video kiosks on myriad musical topics. One explained the forms and styles of various musical genres; another showed how music is used in films and cartoons. There was a "Meet the Conductor" kiosk, that featured a video of symphony conductor Gerard Schwarz and offered the participant a view of the orchestra from the conductor's standpoint, and several kiosks on various instruments and musicians. One could even try playing the instruments on display.

After experiencing the array of delightful exhibits, attendees reassembled for the announcement of election results and a presentation of door prizes. For the list of new officers see the December issue of the Journal. Door prizes included a Silk Road poster won by Seth St. John, and two tickets to the Seattle Symphony Bugs Bunny on Broadway concert, won by Gary Louie. Members then discussed several ideas for next season's meeting topics. Film scoring, IMAX theaters, acoustic analyzers, grounding and shielding were some being considered.

Gary Louie

Waring Collection

Fourteen members of the **Penn State Student Section** held the first meeting of the semester on September 23, at the Fred Waring Collection of the Patee Library. Waring, who is probably best known for the Waring blender, hailed from the local area of Tyrone and donated his collection of musical archives and memorabilia to Penn State after his death.

Curator Peter Keifer, who was a sound person and road manager for Waring, began with a short history of Waring's career as a choral conductor and showman. The Waring Collection houses more than 10 000 recordings, sheet music, scrapbooks, photographs, original cartoon artwork, correspondence, and other items such as costumes and the renown Waring blender. Keifer talked about the many changes in recording media and microphone techniques as reflected in Waring's archives and souvenirs from his Pennsylvania shows. Early shows were recorded using only one or two microphones for the soloists. Later, Keifer explained, the shows incorporated more microphones for the band.

The group gathered in the archive area, where Keifer transfers recordings from old formats such as aluminum records and wire recordings to digital formats like DAT and CD. He demonstrated the wire recorder as well as various types of records that require different size styli.

Keifer concluded the session by

talking a little about the Association for Recorded Sound Collections (ARSC), a group dedicated to the preservation of historical recordings and their importance in cultural heritage. For more information on this group, visit the Web site at: www.arsc-audio.org.

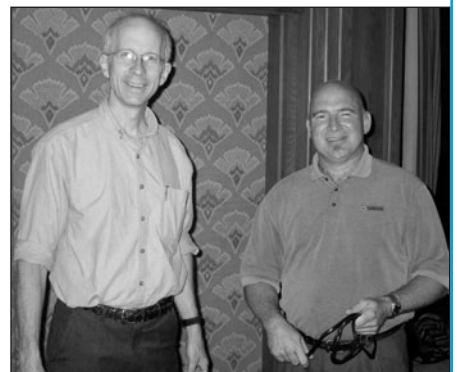
Alexandra Loubeau

IEEE 1394 in SF

The September meeting of the **San Francisco Section** at Dolby Labs, featured John Strawn of S Systems and Mike Overlin of Yamaha. They spoke to 65 audio professionals about IEEE 1394. The IEEE 1394 bus can be used to distribute audio and MIDI signals to practically any audio device including personal computers, keyboards, mixers, samplers and signal processors. IEEE 1394 provides configurable, high-speed (400 Mbps) data transfer and can greatly reduce the complexity of cables commonly found in audio systems.

Strawn talked in general about how audio is distributed over IEEE 1394, which was originally developed to debug back planes and has since evolved into a set of standards for a high-speed serial bus. Apple Computer then developed Firewire, which in turn, became the prototype for the IEEE 1394 standard.

According to Strawn, it is important to remember that IEEE 1394 is a bus, not a network, and is quite versatile in carrying SMPTE code, MIDI, 1-bit, floating point or linear audio. Data transfer on IEEE 1394/Firewire can be isochronous (real-time), or asyn- ➔



John Strawn (left) and Mike Overlin tell San Francisco members about IEEE 1394.

Tobias Lentz tells Central Germany members about audio-virtual reality.



chronous (not real-time).

Overlin described a family of products that incorporate IEEE 1394/-Firewire into a practical system called mLAN by Yamaha and its licensees. The core of Yamaha's mLAN is a series of chipsets, which can be embedded in audio devices to allow connection to the IEEE 1394 bus.

Overlin gave an example of mLAN applied to sound reinforcement at a sports stadium in Florida. In this situation, amplifier power totaling 350 000 W drives 250 loudspeakers. Two rooms contain power amplifiers, one 70 meters from the audio control room, the other 400 meters away. Interconnections were done using mLAN over glass fiber optics to avoid a wiring maze.

Strawn and Overlin concluded with a discussion of some of the major advancements in interconnections, an important and often overlooked part of the audio world.

Paul Howard

British Look at Pro Studios

British Section members gathered June 19, at the Department of Sound Recording at Surrey University to participate in a one-day course organized to give members an opportunity to gain hands-on experience in professional studio operations. Entitled Studio Appreciation Day, the day's events offered a practical introduction to five principal areas of studio operation: surround sound mixing on a Sony Oxford Console; multitrack pop mixdown on Neve Series V; music

and dialog editing on the SADiE Audio Workstation; audio postproduction on Sonic Solutions; and live classical or jazz recording straight to stereo.

The range of activities covered the principal areas of work of professional studios and production facilities, as well as a wide variety of equipment, media and musical styles. During the course of the day members made and edited recordings under the guidance of experts. By using the equipment first hand, participants gained a deeper appreciation of the decisions and compromises involved in the design. The course organizer, Ken Blair of BMP Recording, assisted by other AES members as well as staff and students of the Tonmeister course, were on hand to guide participants through the studio process.

Virtual Acoustics

The **Central Germany Section** held a meeting June 24, at the Technical University of Aachen on virtual acoustics and the simulation of sound fields and multichannel reproduction. Present were 31 members and a large group of students, who were interested in finding out more about the work of the AES and its student sections all over the world.

The University of Aachen, under the leadership of Professor Vorländer, head of the Institute for Technical Acoustics, has for many years been a center of acoustics and electroacoustics. The school first became famous with Professor Kuttruff and his extended investigations in room

acoustics and electroacoustics.

The meeting was full of interesting reports, information and demonstrations. Tobias Lentz reported on his investigations in the field of audio-virtual reality. Simulation, in many cases, is much cheaper and faster than reality; consider flight simulators and simulations of a production line in the automobile industry. Lentz talked about how such models can be applied to the creation of the ideal acoustics for new buildings and/or rooms. For new buildings room acoustical conditions demonstrated in virtual reality enable decisions to be faster and easier. The reproduction of sound fields is based on the fast calculation of sound transmission and the use — in this case — of only two channels and the reproduction with loudspeakers instead of headphones.

Professor Völker then took over with a report on a listening test that compared the sound of 5.1 multichannel stereo reproduction in an Institute for Acoustics and Building Physics control room with a normal living room. The living room contained furniture and had the acoustical features of a normal room, i.e., no specially tailored design.

Alexander Bob described the test results, which showed that a normal living room with additional sound reflections placed on the ceiling walls and floor was not as acoustically bad as expected. Some listeners actually preferred these acoustic conditions to the normal non-reverberant IAB control room, where the reproduction was reportedly more precise and sharply directed.

EMPIRE STATE ANTENNA

The **Empire State Building** has reaffirmed its pledge to provide the power, space and infrastructure necessary to permanently accommodate the television and FM broadcasters in the New York metropolitan area who previously transmitted from the World Trade Center. Adequate power and electricity has already been added to the Empire State Building antenna to allow transmission at full power and full range in cooperation with the Metropolitan Television Alliance.

Since the destruction of the World Trade Center, all television broadcasters have been transmitting from the Empire State Building. In the wake of the attacks, the building accommodated broadcasters by providing access to its antenna at no cost from September 11 through October 2001. The building immediately installed a temporary electrical AC power supply circuit of 2000 amperes, which it replaced in March 2002 by a permanent AC power supply circuit of 9000 amperes that permits analog broadcast at full range and power. Plans to install an additional 16 700 amperes of electricity will permit full analog and digital transmission by all broadcasters by mid-2003. If the proposed 2000-ft. antenna is built at an alternative site, the Empire State Building will remain the primary antenna in the three to four years required to complete such a structure, and then could serve as a permanent secondary antenna.

MUSIC AWARDS FOR 2004

The **University of Louisville School of Music** has announced that it is now accepting applications for the University of Louisville Grawemeyer Award for Music Composition 2004. The University will offer an international prize in recognition of outstanding achieve-

ment by a living composer in a musical genre: choral, orchestral, chamber, electronic, song-cycle, dance, opera, musical theater, extended solo work, etc. The \$200,000 award will be granted to a composer for a work premiered during the five-year period from January 1, 1998 through December 31, 2002. For details regarding the submission of scores and the rules and procedures for the selection of the winning work, contact: Grawemeyer Music Award Committee, School of Music, University of Louisville, Louisville, Kentucky 40292, USA. The committee must receive all materials plus the application form by January 27, 2003.

DVB-T PROJECT

Broadcast specialist Rohde & Schwarz has been awarded an exclusive contract to supply the required transmitter systems needed to convert the Berlin metropolitan area from terrestrial analog to **digital signal transmission**.

In the first phase, the Munich-based company will supply 13 new transmitters and modify the existing analog high-power transmitter so it can be used for digital transmission. The DVB-T coverage of households is scheduled for completion by August 2003. The existing conventional analog terrestrial TV transmissions will be fully replaced. The digitization of terrestrial frequencies will allow more TV program channels to be transmitted on fewer frequencies and provide viewers with additional services such as electronic program guides.

EDISON EXHIBIT AT VIRTUAL MUSEUM

The **IEEE Virtual Museum (VM)** has launched its newest exhibit: Thomas Edison: A Lifetime of Invention. The exhibit, funded by the Charles Edison

Fund, explores the different stages of Edison's career, from his entrepreneurial youth to the disappointments of his later years.

Most are familiar with the works that marked the apex of Edison's career: the incandescent bulb, a lighting system, and the phonograph. These achievements earned Edison the moniker: "The Wizard of Menlo Park," and made him famous.

Lesser known, but also compelling, are the research and inventions that bracket Edison's most prolific years. Edison's first patented device, an automatic vote-recording machine (1865), worked well, but did not meet with commercial success. From that, Edison learned to create things that people needed and would buy. For the next fifty years, he followed that dictum and, in addition to his most famous works, introduced such things as the universal stock ticker, the carbon button transmitter and the storage battery. Along the way, he also helped launch the modern electric utility industry, founded many companies (one of which became General Electric) and created the precursor to the modern research laboratory.

Unfortunately, as he grew older, Edison's ability to identify potentially commercial products declined. His later years were marked by ephemeral success in the concrete and movie businesses and complete failure in iron ore production. Nonetheless, by the time of his death in 1931, he had close to 1100 patents, a number which remains unsurpassed by any inventor.

The exhibit does a thorough job of explaining to a pre-college audience how different technologies worked, how they were developed, and the impact they had on the people who used them. To find out more about Edison, his inventions and his era, visit: www.ieee.org/museum.

Upcoming Meetings

2003 March 22-25: *114th AES Convention*, RAI Conference and Exhibition Centre, Amsterdam, the Netherlands. See p. 1008 for more information.

•

2003 April 28-May 2: *145th Meeting of the Acoustical Society of America*, Nashville, TN, USA. For information call: 516-576-2360, fax: 516-576-2377.

•

2003 May 23-25: *AES 23rd International Conference*, "Signal Processing in Audio Recording and Reproduction," Copenhagen, Denmark. Marienlyst Hotel, Helsingor, Copenhagen. For details see p. 1008.

•

2003 June 26-28: *AES 24th International Conference*, "Multichannel Audio: The New Reality," The Banff Centre, Banff, Alberta, Canada. For more information see: www.banffcentre.ca.

•

2003 July 7-10: *Tenth International Congress on Sound and Vibration*, Stockholm, Sweden.

•

2003 October 10-13: *AES 115th AES Convention*, Jacob K. Javits Convention Center, New York, NY, USA. See p.1008 for details.

•

2003 October 20-23: *NAB Europe Radio Conference*, Prague, Czech Republic. Contact Mark Rebholz (202) 429-3191 or website: www.nab.org.

NEW PRODUCTS

AND

DEVELOPMENTS

Product information is provided as a service to our readers. Contact manufacturers directly for additional information and please refer to the Journal of the Audio Engineering Society.



CD AND DVD DUPLICATOR is an advanced, stand-alone, full-featured system appropriate for video data or professional audio applications. The industrial ES-DVD8 features four user control buttons; two-line backlit LCD display; dedicated read-only DVD-ROM drive; supports DVD/CD video, data and professional audio formats; a burn-on-the-fly copy option; power on self-diagnostics; internal hard drive for multiple image storage and firmware upgrade via CD-ROM. Echo Star Systems, 1223 Osborne Road, Downingtown, PA 19335, USA; tel. +1 610 518 2860; fax +1 267 200 0462; e-mail Joe@echostarsystems.com (Joe Alfonsi); Web site www.echostarsystems.com.

AES SUSTAINING MEMBER
COMPACT PRODUCTION MIXING CONSOLE is an entry-level desk ideally suited for repertory theater, corporate events, industrials, and rental purposes. The console is highly specified with high quality microphone input and full four-band para-

metric EQ with the ability to switch the EQ pre the insert point, as featured on the J-Type live production console. The standard S-Type configuration has a 25-way frame with 16 mono inputs, 8 x group outputs, auxiliary outputs, matrix outputs and DC master faders, plus an oscillator/communications module. However, the configuration is flexible and modules can be slotted into any position. The desk is available in three frame sizes: 17-way, 25-way and 33-way, and may be linked via optional bus connectors. Cadac Electronics Plc, One New Street, Luton, Bedfordshire, LU1 5DX, UK; tel. +44 1582 404202; fax +44 1582 412799; e-mail info@cadac-sound.com; Web site www.cadac-sound.com.



AES SUSTAINING MEMBER
LIGHTWEIGHT LOUDSPEAKER for small PA systems is operable between 70-Hz and 16-kHz. The MVP25 houses a single 12-in low frequency driver and proprietary LM20™ non-metallic, ferro fluid-cooled compression driver coupled to a 90-degree

by 40-degree horn. With a maximum output of 116-dB SPL, the enclosure's sensitivity is rated at 94-dB SPL, while power handling is 150-W RMS, 375-W program. The unit stands 20.4-in high and measures 14.5-in wide by 13.8-in deep. Ruggedly constructed using an internally braced MDF board, the MVP25 is covered with black carpeting and has recessed handles. Community Professional Loudspeakers, 333 East Fifth Street, Chester, PA 19013, USA; tel. +1 610 876 3400 or 800 523 4934 (toll-free); fax +1 610 874 0190; Web site www.loudspeakers.net.



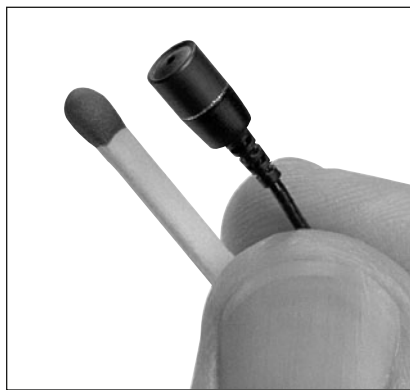
CLASSIC TUBE COMPRESSOR/LIMITER offers analog stereo tube compression and 24-bit/96-k digital conversion via a DC1 24/96-output module. The TS2 allows users to dial in the desired amount of tube warmth via a variable tube drive control and A/B the results by means of a separate switch located on the front panel. The soft-knee compressor performs well in both mixing and tracking applications and has a stereo (link) mode for warming up digital mixdowns, and a dual-mono mode for tracking. A compress control makes best use of threshold and ratio settings and attack and release controls are independent. The TS2 also includes a fixed threshold (+16) limiter to prevent overload of the digital output and provide maximum output without distortion. This product is distributed by the Transamerica Audio Group, www.transaudiogroup.com. Drawmer, Charlotte Street Business Centre, Charlotte Street, Wakefield, West Yorkshire WF1 1UH; tel. +44 1924 378669; fax +44 1924 290460; e-mail dramer.com; Web site www.drawmer.com.

AES SUSTAINING MEMBER

SUBWOOFER LINE ARRAY ELEMENT is an 18-in subwoofer that can either be flown in the array as a VT4889 full-range unit, suspended as a dedicated subwoofer line array, or ground-stacked as desired. The new

VT4880 weighs 59.9 kilograms (132 lbs), delivers a maximum peak output of 138 dB and has an input power rating of 4800 W. The subwoofer has two JBL 2258H 18-in dual voice coil, Direct Cooled™ cone transducers with Neodymium Differential Drive®. The rigid enclosure has DuraFlex™ finish and weatherized component transducers. JBL Professional, 8500 Balboa Boulevard, Northridge, CA 91329, USA; tel. +1 818 894 8850; fax +1 818 894 3479; Web site www.jbl-pro.com/pressroom.

SOUND CALIBRATOR conforms to IEC 60942, Class 1 requirements and is suitable for calibrating high-precision sound pressure level meters. Nominal frequency rating is 1 kHz and sound pressure level is 94 dB. The unit can accept both 1-in and 1/2-in microphones and runs on two alkaline batteries for over 30 hours of continuous use. The small, handheld NC-74 automatically compensates for atmospheric pressure fluctuations, so that manual compensation is not required. Rion Co., Ltd., 20-41, Higashimotomachi 3-chome; Kokubunji, Tokyo 185-8533, Japan; tel. +81 42 359 7888; fax +81 42 359 7442; e-mail info@rion; Website www.rion.co.jp, distributed by Scantek, Inc., 7060-L Oakland Mills Road, Columbia, MD 21046, USA; tel. +1 410 290 7726; fax +1 410 290 9167; e-mail: info@scantekinc; Web site www.scantekinc.com.



AES SUSTAINING MEMBER

SUB-MINIATURE CLIP-ON MICROPHONE is practically invisible and built to exacting tolerances. The MKE Platinum features optimized treble response for improved headroom

at the wireless body-pack transmitter. Bass response is optimized through capsule ventilation design and features a smooth THD curve. The microphone will handle sound pressure levels up to 142 dB. Hermetically sealed capsule and improved embossed umbrella diaphragm protects the microphone against moisture and ensures a longer working life. Low-capacitance, ultra-thin flexible cable allows for unobtrusive attachment of the microphone. Sennheiser Electronic Corporation, 1 Enterprise Drive, Old Lyme, CT 06371, USA; tel. +1 860 434 9190; fax +1 860 434 1759; Web site www.sennheiserusa.com.



SOFTWARE UPGRADE to the SymNet modular audio routing and digital signal processing system is now available. The latest version supports two new hardware add-ons to the SymNet line: BreakIn12 and BreakOut 12, both of which extend the input and output capacity of the system. In addition, two new subclasses of modules called Programmable Filters have been added to the Filters and EQ class in the software's module view toolkit. They include mono and stereo programmable 6 dB/octave sloped low pass and high pass filters, and band pass filters, which afford greater control than standard filters by adding controls based on filter type, resonance, and slope. In addition, the stereo and mono sound pressure level (SPL) modules now include SPL computers with gap sensing technology. These new modules are located under the Dynamics class in the software's module view toolkit and control the volume of a signal as measured by a sensing microphone during gaps in the program material. Symetrix, Inc., 14926 35th Avenue, West Lynnwood, WA 98037, USA; tel. +1 425 787 3222; fax +1 425 787 3211; Web site www.symetrixaudio.com.

AVAILABLE

LITERATURE

The opinions expressed are those of the individual reviewers and are not necessarily endorsed by the Editors of the Journal.

SIGNAL PROCESSING NOISE by Vyacheslav P. Tuzlukov, CRC Press, Boca Raton, FL, 652 pages, 2002, \$129.95, ISBN: 0-8493-1025-3.

The author presents much of his own research in this book on processing of signals with additive and multiplicative noise. This noise can significantly limit the potential of complex signal processing systems, especially when those systems use signals with complex phase structure.

The book sets forth a generalized approach to signal processing in the presence of multiplicative and additive noise that represents an advance in signal processing and detection theory. This approach extends the limits of the noise immunity set by classical and modern signal processing theories. Systems constructed on this basis achieve better detection performance than that of systems currently in use. Addressing a fundamental problem in complex signal processing systems, this book offers a theoretical development for raising noise immunity in various applications, but mainly on signal detection.

Readers who are not experts in advanced DSP will find this a difficult book. This is addressed by the author himself when he states: "To better understand the fundamental statements and concepts..., the reader should consult my two earlier books." The novice DSP engineer should acquire the necessary background to fully benefit from this text.

This work, while containing almost 700 pages, is definitely too concise to be a first text on this topic, but it is a good supplementary text for knowledgeable readers who are involved in the topic. A second difficulty with it is that the author refers sometimes in one sentence to 45 journal papers, of

which many are in Russian. Furthermore, there are quite a number of misspellings of names like Hancel and Crammer for Hankel and Cramér, respectively; and errors in the bibliographic data in the references. Few audio engineers will fully benefit from the theories explained in this book.

Ronald M. Aarts
The Netherlands

IN BRIEF AND OF INTEREST...

The Life and Works of Alan Dower Blumlein: The Inventor of Stereo by Robert Charles Alexander (Focal Press), 2000, explores the fascinating career of the man known as "the inventor of stereo."

Blumlein was born in 1903 and demonstrated an early proclivity for academics and engineering. He attended Imperial College in London and in 1924, after graduating, joined International Western Electric, launching his career as an engineer. Over the course of the next several years Blumlein worked on various projects related to telephony and telegraphy. When he joined the Gramophone Company, he helped devise one of his key inventions, a method of recording 2-channel audio in a single phonograph groove. Many of these features would later be used in the stereo records introduced in the late 1950s.

By the early 1930s, Blumlein was working with the British firm EMI toward establishing electronic television standards, and he helped in the creation of a new high-definition (for the time) TV system. With the coming of World War II, however, Blumlein and his colleagues at EMI were drawn toward war work. In 1942, during airborne experiments on a new type of radar set, Blumlein was killed with sev-

eral others when their airplane crashed.

In this 398-page biography, Alexander carefully documents the life of this inventor, placing his numerous contributions into the broader context of research and development in the British communications industry. The book includes illustrations, references and tables. Price is: \$29.99 (paperback). Focal Press, 225 Wildwood Avenue, Woburn, MA 01801, USA; tel: 800-366-2665, fax: 781-904-2620, Internet: www.focalpress.com.

Noise Reduction in Speech Applications by Gillian M. Davis (CRC Press), 2002, provides a comprehensive introduction to modern techniques for removing or reducing background noise from a range of speech-related applications.

Edited by Gillian Davis, the book begins with a tutorial, which provides the background material necessary to understanding system issues, digital algorithms, and their implementation. The following chapters are written by international experts and offer a practical systems approach to help readers decide whether or not digital noise reduction will solve specific problems in their own systems. A closing section explores a variety of applications and demonstrates the satisfying results that can be achieved with various noise reduction techniques.

Other features of the book include: a section on how to select the most appropriate technique for a given situation; background on digital signal processing techniques and noise in digital speech communication systems; an examination of single-channel speech enhancers, microphone arrays and echo canceller noise reduction algorithms. The price: \$129.95. CRC Press, or www.crcpress.com.

In Memoriam



John G. Humble
1922 – 2002

John Granville Humble was born December 31, 1922, in Lincoln, Nebraska. As a boy, he loved to take things apart to see how they worked, and then put them back together, trying carefully not to have any left over parts. He often used his little brother to determine if certain “high voltage” areas of his disassembled radios were fully discharged and safe to touch.

John was always, intrigued with all communications media. He loved newspapers, and was knowledgeable about the history of all of the famous publications of the time. In the eighth grade, he enlisted the help of several of his friends and published his own newspaper, *The News Recorder*, which he sold every Friday for two cents a copy. It wasn't long before he and the other boys associated with the paper had completed a wired communication network to each of their houses spanning several city blocks.

John always excelled in his studies and was recommended by his chemistry and physics teachers to apply for the Midland Radio School Scholarship. Midland, located in Missouri, was one of the most prestigious radio training schools at that time. John took their advice and was accepted after graduating from high school in 1941. Midland Radio School was founded by Arthur B. Church and was affiliated with KMBC in Kansas City, MO. Many radio personalities began their careers at KMBC including John

Cameron Swayze, Ted Malone, Walter Cronkite and Caroline Ellis. After completing his studies at Midland, he worked for a radio station in Waycross, GA, for about a year. He then returned to Kansas City and worked for Radio Station KMBC as a radio engineer.

In the early 1960s John came to California to work as the Northwest sales representative for Altec Lansing Corporation. It was at Altec that he began to develop his well regarded reputation in the audio industry. As the years passed, he moved in the direction of the theater sound product division. In theater sound, the Altec name was held in high regard because most of the solid engineering practices were adopted from the early days at Western Electric.

Because John never took any endeavor lightly, he quickly became an authoritative figure and a bit of a historian for the theater sound division. He represented quite a lot more than that of salesman. He used his well trained background in engineering and radio, and adapted it well to audio. Many distinguished engineers at Altec, like Bill Hayes, remember having technical conversations with him as often as three times a week.

John was never pretentious; he never came across as a “know-it-all.” He did, however, know quite a bit, and for questions a customer might have that he didn't know, he was always sure to find the answer. In this way,

he would learn little extra pearls of audio wisdom and history along the way. During these years, and many to follow, he earned the respect of many leaders in the audio industry,

As the theater business started to decline, Altec began to focus more on engineered sound. John saw this as an opportunity to begin his own business. He became an independent sales representative for a number of products. Shortly after taking on a new, unknown line of commercial sound equipment manufactured in Japan by the TOA Electric Company, he moved to Southern California.

After talking to many people since John's death, the one word most often mentioned was integrity. His honesty and integrity along with his sincere love of the business was his key to success. He was always on the side of the customer; and everyone enjoyed seeing him show up at their office.

During the 29,063 days he spent here on earth (he died July 27),s John made many friends and took an interest in young people starting out in the sound industry, mentoring them to successful careers. He was also involved with neighborhood charity projects of his own and, living up to his name, never mentioned them.

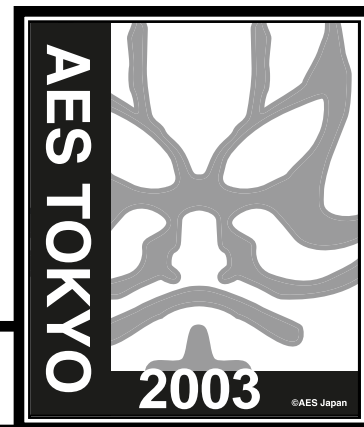
Those who have worked with John will never hear the familiar words “John Humble” in the receiver again. A dear friend will be sorely missed.

Tim Clark, John Murray
and Thomas P. Leach

AUDIO ENGINEERING SOCIETY

CALL for PAPERS

AES 11th Regional Convention, 2003 Tokyo, Japan



Dates: 2003 July 7–9

Location: Science Museum, Chiyoda, Tokyo, Japan

The 11th Regional Convention committee invites you to submit technical papers for presentation at the 2003 July meeting in Tokyo, Japan. **By 2003 February 28** a proposed title, 60- to 120-word abstract (English), and 500- to 750-word précis of the paper should be submitted via Email to the papers chair (see below). The précis should describe the work performed, methods employed, conclusion(s), and significance of the paper. Title, abstract, and précis should follow the guidelines in Information for Authors at www.aes.org/journal/con_infoauth.html. Authors without Email and Internet access should contact the papers chair for hardcopy forms and instructions. Acceptance of papers will be determined by a review committee based on an assessment of the précis. A preprint manuscript will be a condition for acceptance of the paper for presentation at the convention. Abstracts of accepted papers will be published in the convention program. Please choose your wording carefully.

PROPOSED TOPICS FOR PAPERS

High-Resolution Audio
Signal Processing for Audio
Architectural Acoustics and Room Acoustics
Audio Networking
Recording and Reproduction
Audio Games and Entertainment
Sound Reinforcement
Digital Audio Equipment and Media
Automotive Audio
Instrumentation and Measurement
Transmission

Transducers
Psychoacoustics, Perception, Listening Tests
Multichannel Sound
Program Production
Audio Coding
Audio Watermarking
Virtual Reality Audio
Digital Broadcasting
Audio Restoration and Enhancement
Analysis and Synthesis of Sound

CONVENTION CHAIRS

Convention chair: Kimio Hamasaki
NHK Science & Technical Research
Laboratories
Telephone: +81 3 5494 3208
Fax: +81 3 5494 3219
Email: hamasaki.k-ku@nhk

Vice chair: Hiroaki Suzuki
Victor Company of Japan, Limited (JVC)
Telephone: +81 45 450 1779
Fax: +81 45 450 1639
Email: hr-suzuki@skt.jvc-victor

SUBMISSION SCHEDULE

Please submit via Email proposed title, abstract, and précis to papers chair no later than **2003 February 28**. If you have any questions contact **Papers Chair**:

Shinji Koyano

Email: shinji_koyano@post.pioneer
Pioneer Corporation, Corporate R & D Lab.
6-1-1, Fujimi, Tsurugashima-shi
Saitama 350-2288, Japan
Telephone: +81 49 279 2627
Fax: +81 49 279 1513

Authors whose contributions have been accepted for presentation will receive additional instructions for submission of their manuscripts.

Proposal deadline:
2003 February 28

Acceptance emailed:
2003 March 17

Preprint deadline:
2003 May 10

SECTIONS CONTACTS

DIRECTORY

The following is the latest information we have available for our sections contacts. If you wish to change the listing for your section, please mail, fax or e-mail the new information to: Mary Ellen Ilich, AES Publications Office, Audio Engineering Society, Inc., 60 East 42nd Street, Suite 2520, New York, NY 10165-2520, USA. Telephone +1 212 661 8528. Fax +1 212 661 7829.

Updated information that is received by the first of the month will be published in the next month's *Journal*. Please help us to keep this information accurate and timely.

EASTERN REGION, USA/CANADA

Vice President:

Jim Anderson
12 Garfield Place
Brooklyn, NY 11215
Tel. +1 718 369 7633
Fax +1 718 669 7631
E-mail

UNITED STATES OF AMERICA

CONNECTICUT

**University of Hartford
Section (Student)**
Howard A. Canistraro
Faculty Advisor
AES Student Section
University of Hartford
Ward College of Technology
200 Bloomfield Ave.
West Hartford, CT 06117-1599
Tel. +1 860 768 5358
Fax +1 860 768 5074
E-mail

FLORIDA

**Full Sail Real World
Education Section (Student)**
Bill Smith, Faculty Advisor
AES Student Section
Full Sail Real World Education
3300 University Blvd., Suite 160
Winter Park, FL 327922
Tel. +1 800 679 0100
E-mail

**University of Miami Section
(Student)**
Ken Pohlmann, Faculty Advisor
AES Student Section
University of Miami
School of Music
PO Box 248165
Coral Gables, FL 33124-7610
Tel. +1 305 284 6252
Fax +1 305 284 4448
E-mail

GEORGIA

Atlanta Section
Robert Mason
2712 Leslie Dr.
Atlanta, GA 30345
Home Tel. +1 770 908 1833
E-mail

MASSACHUSETTS

**Berklee College of Music
Section (Student)**

Eric Reuter, Faculty Advisor
Berklee College of Music
Audio Engineering Society
% Student Activities
1140 Boylston St., Box 82
Boston, MA 02215
Tel. +1 617 747 8251
Fax +1 617 747 2179
E-mail

Boston Section

J. Nelson Chadderdon
c/o Oceanwave Consulting, Inc.
21 Old Town Rd.
Beverly, MA 01915
Tel. +1 978 232 9535 x201
Fax +1 978 232 9537
E-mail

University of Massachusetts -Lowell Section (Student)

John Shirley, Faculty Advisor
AES Student Chapter
University of Massachusetts-Lowell
Dept. of Music
35 Wilder St., Ste. 3
Lowell, MA 01854-3083
Tel. +1 978 934 3886
Fax +1 978 934 3034
E-mail

Worcester Polytechnic Institute Section (Student)

William Michalson
Faculty Advisor
AES Student Section
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609
Tel. +1 508 831 5766
E-mail

NEW JERSEY

**William Paterson University
Section (Student)**
David Kerzner, Faculty Advisor
AES Student Section
William Paterson University
300 Pompton Rd.
Wayne, NJ 07470-2103
Tel. +1 973 720 3198
Fax +1 973 720 2217
E-mail

NEW YORK

Fredonia Section (Student)
Bernd Gottinger, Faculty Advisor
AES Student Section
SUNY-Fredonia
1146 Mason Hall
Fredonia, NY 14063
Tel. +1 716 673 4634
Fax +1 716 673 3154
E-mail

Institute of Audio Research Section (Student)

Noel Smith, Faculty Advisor
AES Student Section
Institute of Audio Research
64 University Pl.
New York, NY 10003
Tel. +1 212 677 7580
Fax +1 212 677 6549
E-mail

New York Section

Robbin L. Gheesling
Broadness, LLC
265 Madison Ave., Second Floor
New York, NY 10016
Tel. +1 212 818 1313
Fax +1 212 818 1330
E-mail

NORTH CAROLINA

University of North Carolina at Asheville Section (Student)

Wayne J. Kirby
Faculty Advisor
AES Student Section
University of North Carolina at
Asheville
Dept. of Music
One University Heights
Asheville, NC 28804
Tel. +1 828 251 6487
Fax +1 828 253 4573
E-mail

PENNSYLVANIA

Carnegie Mellon University Section (Student)

Thomas Sullivan
Faculty Advisor
AES Student Section
Carnegie Mellon University
University Center Box 122
Pittsburg, PA 15213
Tel. +1 412 268 3351
E-mail

Duquesne University Section (Student)

Francisco Rodriguez
Faculty Advisor
AES Student Section
Duquesne University
School of Music
600 Forbes Ave.
Pittsburgh, PA 15282
Tel. +1 412 434 1630
Fax +1 412 396 5479
E-mail

Pennsylvania State University Section (Student)

Brian Tuttle
AES Penn State Student Chapter
Graduate Program in Acoustics

217 Applied Science Bldg.
University Park, PA 16802
Home Tel. +1 814 863 8282
Fax +1 814 865 3119
E-mail

Philadelphia Section

Rebecca Mercuri
P.O. Box 1166.
Philadelphia, PA 19105
Tel. +1 609 895 1375
E-mail

VIRGINIA

Hampton University Section (Student)

Bob Ransom, Faculty Advisor
AES Student Section
Hampton University
Dept. of Music
Hampton, VA 23668
Office Tel. +1 757 727 5658,
+1 757 727 5404
Home Tel. +1 757 826 0092
Fax +1 757 727 5084
E-mail

WASHINGTON, DC

American University Section (Student)

Benjamin Tomassetti
Faculty Advisor
AES Student Section
American University
Physics Dept.
4400 Massachusetts Ave., N.W.
Washington, DC 20016
Tel. +1 202 885 2746
Fax +1 202 885 2723
E-mail

District of Columbia Section

John W. Reiser
DC AES Section Secretary
P.O. Box 169
Mt. Vernon, VA 22121-0169
Tel. +1 703 780 4824
Fax +1 703 780 4214
E-mail

CANADA

McGill University Section (Student)

John Klepko, Faculty Advisor
AES Student Section
McGill University
Sound Recording Studios
Strathcona Music Bldg.
555 Sherbrooke St. W.
Montreal, Quebec H3A 1E3
Canada
Tel. +1 514 398 4535 ext. 0454
E-mail

SECTIONS CONTACTS

DIRECTORY

Toronto Section

Anne Reynolds
606-50 Cosburn Ave.
Toronto, Ontario M4K 2G8
Canada
Tel. +1 416 957 6204
Fax +1 416 364 1310
E-mail

CENTRAL REGION, USA/CANADA

Vice President:

Jim Kaiser
Master Mix
1921 Division St.
Nashville, TN 37203
Tel. +1 615 321 5970
Fax +1 615 321 0764
E-mail

UNITED STATES OF AMERICA

ARKANSAS

**University of Arkansas at
Pine Bluff Section (Student)**
Robert Elliott, Faculty Advisor
AES Student Section
Music Dept. Univ. of Arkansas
at Pine Bluff
1200 N. University Drive
Pine Bluff, AR 71601
Tel. +1 870 575 8916
Fax +1 870 543 8108
E-mail

INDIANA

**Ball State University Section
(Student)**
Michael Pounds, Faculty Advisor
AES Student Section
Ball State University
MET Studios
2520 W. Bethel
Muncie, IN 47306
Tel. +1 765 285 5537
Fax +1 765 285 8768
E-mail

Central Indiana Section

James Latta
Sound Around
6349 Warren Ln.
Brownsburg, IN 46112
Office Tel. +1 317 852 8379
Fax +1 317 858 8105
E-mail

ILLINOIS

Chicago Section

Robert Zurek
Motorola
2001 N. Division St.
Harvard, IL 60033
Tel. +1 815 884 1361
Fax +1 815 884 2519
E-mail

Columbia College Section (Student)

Dominique J. Chéenne
Faculty Advisor
AES Student Section

676 N. LaSalle, Ste. 300
Chicago, IL 60610
Tel. +1 312 344 7802
Fax +1 312 482 9083

University of Illinois at Urbana-Champaign Section (Student)

David S. Petruccio Jr.
AES Student Section
University of Illinois, Urbana-
Champaign
Urbana, IL 61801
Tel. +1 217 621 7586
E-mail

KANSAS

Kansas City Section

Jim Mitchell
Custom Distribution Limited
12301 Riggs Rd.
Overland Park, KS 66209
Tel. +1 913 661 0131
Fax +1 913 663 5662

LOUISIANA

New Orleans Section

Joseph Doherty
6015 Annunciation St.
New Orleans, LA 70118
Tel. +1 504 891 4424
Fax +1 504 891 6075

MICHIGAN

Detroit Section

Tom Conlin
DaimlerChrysler
E-mail

Michigan Technological University Section (Student)

Andre LaRouche
AES Student Section
Michigan Technological
University
Electrical Engineering Dept.
1400 Townsend Dr.
Houghton, MI 49931
Home Tel. +1 906 847 9324
E-mail

West Michigan Section

Carl Hordyk
Calvin College
3201 Burton S.E.
Grand Rapids, MI 49546
Tel. +1 616 957 6279
Fax +1 616 957 6469
E-mail

MINNESOTA

Music Tech College Section (Student)

Michael McKern
Faculty Advisor
AES Student Section
Music Tech College
19 Exchange Street East
Saint Paul, MN 55101
Tel. +1 651 291 0177
Fax +1 651 291 0366
E-mail

Ridgewater College, Hutchinson Campus Section (Student)

Dave Igl, Faculty Advisor
AES Student Section
Ridgewater College, Hutchinson
Campus
2 Century Ave. S.E.
Hutchinson, MN 55350
E-mail

Upper Midwest Section

Greg Reiersen
Rare Form Mastering
4624 34th Avenue South
Minneapolis, MN 55406
Tel. +1 612 327 8750
E-mail

MISSOURI

St. Louis Section

John Nolan, Jr.
693 Green Forest Dr.
Fenton, MO 63026
Tel./Fax +1 636 343 4765
E-mail

NEBRASKA

Northeast Community College Section (Student)

Anthony D. Beardslee
Faculty Advisor
AES Student Section
Northeast Community College
P.O. Box 469
Norfolk, NE 68702
Tel. +1 402 644 0581
Fax +1 402 644 0650
E-mail

OHIO

Ohio University Section (Student)

Erin M. Dawes
AES Student Section
Ohio University
RTVC Bldg.
9 S. College St.
Athens, OH 45701-2979
Home Tel. +1 740 597 6608
E-mail

University of Cincinnati Section (Student)

Thomas A. Haines
Faculty Advisor
AES Student Section
University of Cincinnati
College-Conservatory of Music
M.L. 0003
Cincinnati, OH 45221
Tel. +1 513 556 9497
Fax +1 513 556 0202

TENNESSEE

Belmont University Section (Student)

Wesley Bulla, Faculty Advisor
AES Student Section
Belmont University
Nashville, TN 37212

Middle Tennessee State University Section (Student)

Phil Shullo, Faculty Advisor
AES Student Section
Middle Tennessee State University
301 E. Main St., Box 21
Murfreesboro, TN 37132
Tel. +1 615 898 2553
E-mail

Nashville Section

Tom Edwards
MTV Networks
330 Commerce St.
Nashville, TN 37201
Tel. +1 615 335 8520
Fax +1 615 335 8608
E-mail

SAE Nashville Section (Student)

Larry Sterling, Faculty Advisor
AES Student Section
7 Music Circle N.
Nashville, TN 37203
Tel. +1 615 244 5848
Fax +1 615 244 3192
E-mail

TEXAS

Southwest Texas State University Section (Student)

Mark C. Erickson
Faculty Advisor
AES Student Section
Southwest Texas State
University
224 N. Guadalupe St.
San Marcos, TX 78666
Tel. +1 512 245 8451
Fax +1 512 396 1169
E-mail

WESTERN REGION, USA/CANADA

Vice President:

Bob Moses
Island Digital Media Group,
LLC
26510 Vashon Highway S.W.
Vashon, WA 98070
Tel. +1 206 463 6667
Fax +1 810 454 5349
E-mail

UNITED STATES OF AMERICA

ARIZONA

Conservatory of The Recording Arts and Sciences Section (Student)

Glen O'Hara, Faculty Advisor
AES Student Section
Conservatory of The Recording
Arts and Sciences
2300 E. Broadway Rd.
Tempe, AZ 85282
Tel. +1 480 858 9400, 800 562
6383 (toll-free)
Fax +1 480 829 1332
E-mail

CALIFORNIA

American River College Section (Student)
Eric Chun, Faculty Advisor
AES Student Section
American River College Chapter
4700 College Oak Dr.
Sacramento, CA 95841
Tel. +1 916 484 8420
E-mail

Cal Poly San Luis Obispo State University Section (Student)
Jerome R. Breitenbach
Faculty Advisor
AES Student Section
California Polytechnic State University
Dept. of Electrical Engineering
San Luis Obispo, CA 93407
Tel. +1 805 756 5710
Fax +1 805 756 1458
E-mail

California State University -Chico Section (Student)
Keith Seppanen, Faculty Advisor
AES Student Section
California State University-Chico
400 W. 1st St.
Chico, CA 95929-0805
Tel. +1 530 898 5500
E-mail

Citrus College Section (Student)
Gary Mraz, Faculty Advisor
AES Student Section
Citrus College
Recording Arts
1000 W. Foothill Blvd.
Glendora, CA 91741-1899
Fax +1 626 852 8063

Cogswells Polytechnical College Section (Student)
Tim Duncan, Faculty Sponsor
AES Student Section
Cogswell Polytechnical College
Music Engineering Technology
1175 Bordeaux Dr.
Sunnyvale, CA 94089
Tel. +1 408 541 0100, ext. 130
Fax +1 408 747 0764
E-mail

Expression Center for New Media Section (Student)
Scott Theakston, Faculty Advisor
AES Student Section
Ex'pression Center for New Media
6601 Shellmount St.
Emeryville, CA 94608
Tel. +1 510 654 2934
Fax +1 510 658 3414
E-mail

Long Beach City College Section (Student)
Nancy Allen, Faculty Advisor
AES Student Section
Long Beach City College
4901 E. Carson St.
Long Beach, CA 90808

Tel. +1 562 938 4312
Fax +1 562 938 4409
E-mail

Los Angeles Section
Andrew Turner
1733 Lucile Ave., #8
Los Angeles, CA 90026
Tel. +1 323 661 0390
E-mail

San Diego Section
J. Russell Lemon
2031 Ladera Ct.
Carlsbad, CA 92009-8521
Home Tel. +1 760 753 2949
E-mail

San Diego State University Section (Student)
John Kennedy, Faculty Advisor
AES Student Section
San Diego State University
Electrical & Computer Engineering Dept.
5500 Campanile Dr.
San Diego, CA 92182-1309
Tel. +1 619 594 1053
Fax +1 619 594 2654
E-mail

San Francisco Section
Bill Orner
1513 Meadow Lane
Mountain View, Ca 94040
Tel. +1 650 903 0301
Fax +1 650 903 0409
E-mail

San Francisco State University Section (Student)
John Barsotti, Faculty Advisor
AES Student Section
San Francisco State University
Broadcast and Electronic Communication Arts Dept.
1600 Holloway Ave.
San Francisco, CA 94132
Tel. +1 415 338 1507
E-mail

Stanford University Section (Student)
Jay Kadis, Faculty Advisor
Stanford AES Student Section
Stanford University
CCRMA/Dept. of Music
Stanford, CA 94305-8180
Tel. +1 650 723 4971
Fax +1 650 723 8468
E-mail

University of Southern California Section (Student)
Kenneth Lopez
Faculty Advisor
AES Student Section
University of Southern California
840 W. 34th St.
Los Angeles, CA 90089-0851
Tel. +1 213 740 3224
Fax +1 213 740 3217
E-mail

COLORADO

Colorado Section
Robert F. Mahoney
Robert F. Mahoney & Associates
310 Balsam Ave.
Boulder, CO 80304
Tel. +1 303 443 2213
Fax +1 303 443 6989
E-mail

Denver Section (Student)
Roy Pritts, Faculty Advisor
AES Student Section
University of Colorado at Denver
Dept. of Professional Studies
Campus Box 162
P.O. Box 173364
Denver, CO 80217-3364
Tel. +1 303 556 2795
Fax +1 303 556 2335
E-mail

OREGON

Portland Section
Tony Dal Molin
Audio Precision, Inc.
5750 S.W. Arctic Dr.
Portland, OR 97005
Tel. +1 503 627 0832
Fax +1 503 641 8906
E-mail

UTAH

Brigham Young University Section (Student)
Jim Anglesey,
Faculty Advisor
BYU-AES Student Section
School of Music
Brigham Young University
Provo, UT 84602
Tel. +1 801 378 1299
Fax +1 801 378 5973 (Music Office)
E-mail

Utah Section
Deward Timothy
%o Poll Sound
4026 S. Main
Salt Lake City, UT 84107
Tel. +1 801 261 2500
Fax +1 801 262 7379

WASHINGTON

Pacific Northwest Section
Gary Louie
University of Washington
School of Music
PO Box 353450
Seattle, WA 98195
Office Tel. +1 206 543 1218
Fax +1 206 685 9499
E-mail

The Art Institute of Seattle Section (Student)
David G. Christensen
Faculty Advisor
AES Student Section
The Art Institute of Seattle

2232 Elliott Ave.
Seattle, WA 98121-1622
Tel. +1 206 239 2397
E-mail
.....

CANADA

Alberta Section
Frank Lockwood
AES Alberta Section
Suite 404
815 - 50 Avenue S.W.
Calgary, Alberta T2S 1H8
Canada
Home Tel. +1 403 703 5277
Fax +1 403 762 6665
E-mail

Vancouver Section
Peter L. Janis
C-Tec
#114, 1585 Broadway
Port Coquitlam, B.C. V3C 2M7
Canada
Tel. +1 604 942 1001
Fax +1 604 942 1010
E-mail

Vancouver Student Section
Gregg Gorrie, Faculty Advisor
AES Greater Vancouver Student Section
Centre for Digital Imaging and Sound
3264 Beta Ave.
Burnaby, B.C. V5G 4K4, Canada
Tel. +1 604 298 5400
E-mail
.....

NORTHERN REGION, EUROPE

Vice President:
Søren Bech
Bang & Olufsen a/s
CoreTech
Peter Bangs Vej 15
DK-7600 Struer, Denmark
Tel. +45 96 84 49 62
Fax +45 97 85 59 50
E-mail
northern

BELGIUM

Belgian Section
Hermann A. O. Wilms
AES Europe Region Office
Zevenbunderslaan 142, #9
BE-1190 Vorst-Brussels, Belgium
Tel. +32 2 345 7971
Fax +32 2 345 3419

DENMARK

Danish Section
Knud Bank Christensen
Skovvej 2
DK-8550 Ryomgård, Denmark
Tel. +45 87 42 71 46
Fax +45 87 42 70 10
E-mail

Danish Student Section

Torben Poulsen
Faculty Advisor
AES Student Section
Technical University of Denmark
Ørsted-DTU, Acoustic
Technology
DTU - Building 352
DK-2800 Kgs. Lyngby, Denmark
Tel. +45 45 25 39 40
Fax +45 45 88 05 77
E-mail
...

FINLAND

Finnish Section

Kalle Koivuniemi
Nokia Research Center
P.O. Box 100
FI-33721 Tampere, Finland
Tel. +358 7180 35452
Fax +358 7180 35897
E-mail

NETHERLANDS

Netherlands Section

Rinus Boone
Voorweg 105A
NL-2715 NG Zoetermeer
Netherlands
Tel. +31 15 278 14 71, +31 62
127 36 51
Fax +31 79 352 10 08
E-mail

Netherlands Student Section

Dirk Fischer
AES Student Section
Groenewegje 143a
Den Haag, Netherlands
Home Tel. +31 70 3885958
E-mail
netherlands

NORWAY

Norwegian Section

Jan Erik Jensen
Nøklesvingen 74
NO-0689 Oslo, Norway
Office Tel. +47 22 24 07 52
Home Tel. +47 22 26 36 13
Fax +47 22 24 28 06
E-mail

RUSSIA

**All-Russian State Institute of
Cinematography Section
(Student)**

Leonid Sheetov, Faculty Sponsor
AES Student Section
All-Russian State Institute of
Cinematography (VGIK)
W. Pieck St. 3
RU-129226 Moscow, Russia
Tel. +7 095 181 3868
Fax +7 095 187 7174
E-mail

Moscow Section

Michael Lannie
Research Institute for

Television and Radio
Acoustic Laboratory
12-79 Chernomorsky bulvar
RU-113452 Moscow, Russia
Tel. +7 095 2502161, +7 095
1929011
Fax +7 095 9430006
E-mail

St. Petersburg Section

Irina A. Aldoshina
St. Petersburg University of
Telecommunications
Gangutskaya St. 16, #31
RU-191187 St. Petersburg

Russia
Tel. +7 812 272 4405
Fax +7 812 316 1559
E-mail

St. Petersburg Student Section

Natalia V. Tyurina
Faculty Advisor
Prosvescheniya pr., 41, 185
RU-194291 St. Petersburg, Russia
Tel. +7 812 595 1730
Fax +7 812 316 1559
E-mail
st_peter

SWEDEN

Swedish Section

Mikael Olsson
Audio Data Lab
Katarinavägen 22
SE-116 45 Stockholm, Sweden
Tel. +46 8 30 29 98
Fax +46 8 641 67 91
E-mail

**University of Luleå-Piteå
Section (Student)**

Lars Hallberg, Faculty Sponsor
AES Student Section
University of Luleå-Piteå
School of Music
Box 744
S-94134 Piteå, Sweden
Tel. +46 911 726 27
Fax +46 911 727 10
E-mail

UNITED KINGDOM

British Section

Heather Lane
Audio Engineering Society
P.O. Box 645
Slough GB-SL1 8BJ
United Kingdom
Tel. +44 1628 663725
Fax +44 1628 667002
E-mail

**CENTRAL REGION,
EUROPE**

Vice President:
Markus Erne
Scopein Research
Sonnmattweg 6
CH-5000 Aarau, Switzerland

Tel. +41 62 825 09 19
Fax +41 62 825 09 15
E-mail

AUSTRIA

Austrian Section

Franz Lechleitner
Lainergasse 7-19/2/1
AT-1230 Vienna, Austria
Office Tel. +43 1 4277 29602
Fax +43 1 4277 9296
E-mail

Graz Section (Student)

Robert Höldrich
Faculty Sponsor
Institut für Elektronische Musik
und Akustik
Inffeldgasse 10
AT-8010 Graz, Austria
Tel. +43 316 389 3172
Fax +43 316 389 3171
E-mail

Vienna Section (Student)

Jürg Jecklin, Faculty Sponsor
Vienna Student Section
Universität für Musik und
Darstellende Kunst Wien
Institut für Elektroakustik und
Experimentelle Musik
Rienösslgasse 12
AT-1040 Vienna, Austria
Tel. +43 1 587 3478
Fax +43 1 587 3478 20
E-mail

CZECH REPUBLIC

Czech Section

Jiri Ocenasek
Dejvicka 36
CZ-160 00 Prague 6
Czech Republic
Home Tel. +420 2 24324556
E-mail

**Czech Republic Student
Section**

Libor Husník, Faculty Advisor
AES Student Section
Czech Technical University at
Prague
Technická 2,
CZ-116 27 Prague 6
Czech Republic
Tel. +420 2 2435 2115
E-mail

GERMANY

Berlin Section (Student)

Bernhard Güttler
Zionskirchstrasse 14
DE-10119 Berlin, Germany
Tel. +49 30 4404 72 19
Fax +49 30 4405 39 03
E-mail

Central German Section

Ernst-Joachim Völker
Institut für Akustik und
Bauphysik
Kiesweg 22-24
DE-61440 Oberursel, Germany
Tel. +49 6171 75031

Fax +49 6171 85483
E-mail

Darmstadt Section (Student)

G. M. Sessler, Faculty Sponsor
AES Student Section
Technical University of
Darmstadt
Institut für Übertragungstechnik
Merkstr. 25
DE-64283 Darmstadt, Germany
Tel. +49 6151 162869
E-mail
darmstadt

Detmold Section (Student)

Andreas Meyer, Faculty Sponsor
AES Student Section
c/o Erich Thienhaus Institut
Tonmeisterausbildung
Hochschule für Musik
Detmold
Neustadt 22, DE-32756
Detmold, Germany
Tel/Fax +49 5231 975639
E-mail

**Düsseldorf Section
(Student)**

Ludwig Kugler
AES Student Section
Bilker Allee 126
DE-40217 Düsseldorf, Germany
Tel. +49 211 3 36 80 38
E-mail
duesseldorf

**Ilmenau Section
(Student)**

Karlheinz Brandenburg
Faculty Sponsor
AES Student Section
Institut für Medientechnik
PF 10 05 65
DE-98684 Ilmenau, Germany
Tel. +49 3677 69 2676
Fax +49 3677 69 1255
E-mail
ilmenau

North German Section

Reinhard O. Sahr
Eickhopskamp 3
DE-30938 Burgwedel, Germany
Tel. +49 5139 4978
Fax +49 5139 5977
E-mail

South German Section

Gerhard E. Picklapp
Landshuter Allee 162
DE-80637 Munich, Germany
Tel. +49 89 15 16 17
Fax +49 89 157 10 31
E-mail

HUNGARY

Hungarian Section

István Matók
Rona u. 102. II. 10
HU-1149 Budapest, Hungary
Home Tel. +36 1 383 24 81
Fax +36 1 383 24 81
E-mail

LITHUANIA

Lithuanian Section
Vytautas J. Stauskis
Vilnius Gediminas Technical
University
Sauletekio al. 11
LT-2040 Vilnius, Lithuania
Tel. +370 2 700 492
Fax +370 2 700 498
E-mail

POLAND

Polish Section
Jan A. Adamczyk
University of Mining and
Metallurgy
Dept. of Mechanics and
Vibroacoustics
al. Mickiewicza 30
PL-30 059 Cracow, Poland
Tel. +48 12 617 30 55
Fax +48 12 633 23 14
E-mail

**Technical University of Gdansk
Section (Student)**

Pawel Zwan
AES Student Section
Technical University of Gdansk
Sound Engineering Dept.
ul. Narutowicza 11/12
PL-80 952 Gdansk, Poland
Home Tel. +48 58 347 23 98
Office Tel. +4858 3471301
Fax +48 58 3471114
E-mail

**Wroclaw University of
Technology Section (Student)**

Andrzej B. Dobrucki
Faculty Sponsor
AES Student Section
Institute of Telecommunications
and Acoustics
Wroclaw University of
Technology
Wybrzeze Wyspianskiego 27
PL-503 70 Wroclaw, Poland
Tel. +48 71 320 30 68
Fax +48 71 320 31 89
E-mail
wroclaw

REPUBLIC OF BELARUS

Belarus Section
Valery Shalatonin
Belarusian State University of
Informatics and
Radioelectronics
vul. Petrusya Brouki 6
BY-220027 Minsk
Republic of Belarus
Tel. +375 17 239 80 95
Fax +375 17 231 09 14
E-mail

SLOVAK REPUBLIC

Slovakian Republic Section
Richard Varkonda
Centron Slovakia Ltd.
Podhaj 107
SK-841 03 Bratislava
Slovak Republic

Tel. +421 7 6478 0767
Fax. +421 7 6478 0042
E-mail
slovakian

SWITZERLAND

Swiss Section
Joel Godel
AES Swiss Section
Sonnmattweg 6
CH-5000 Aarau
Switzerland
E-mail

UKRAINE

Ukrainian Section
Valentin Abakumov
National Technical University
of Ukraine
Kiev Politechnical Institute
Politechnical St. 16
Kiev UA-56, Ukraine
Tel./Fax +38 044 2366093

**SOUTHERN REGION,
EUROPE**

Vice President:
Daniel Zalay
Conservatoire de Paris
Dept. Son
FR-75019 Paris, France
Office Tel. +33 1 40 40 46 14
Fax +33 1 40 40 47 68
E-mail
southern

BOSNIA-HERZEGOVINA

Bosnia-Herzegovina Section
Jozo Talajic
Bulevar Mese Selimovica 12
BA-71000 Sarajevo
Bosnia-Herzegovina
Tel. +387 33 455 160
Fax +387 33 455 163
E-mail
bosnia

BULGARIA

Bulgarian Section
Konstantin D. Kounov
Bulgarian National Radio
Technical Dept.
4 Dragan Tzankov Blvd.
BG-1040 Sofia, Bulgaria
Tel. +359 2 65 93 37, +359 2
9336 6 01
Fax +359 2 963 1003
E-mail
bulgaria

CROATIA

Croatian Section
Silvije Stamac
Hrvatski Radio
Prisavlje 3
HR-10000 Zagreb, Croatia
Tel. +385 1 634 28 81
Fax +385 1 611 58 29
E-mail

Croatian Student Section

Hrvoje Domitrovic
Faculty Advisor
AES Student Section
Faculty of Electrical
Engineering and Computing
Dept. of Electroacoustics (X. Fl.)
Unska 3
HR-10000 Zagreb, Croatia
Tel. +385 1 6129 640
Fax +385 1 6129 852
E-mail
croatian

FRANCE

**Conservatoire de Paris
Section (Student)**

Alessandra Galleron
36, Ave. Parmentier
FR-75011 Paris, France
Tel. +33 1 43 38 15 94

French Section

Michael Williams
Ile du Moulin
62 bis Quai de l'Artois
FR-94170 Le Perreux sur
Marne, France
Tel. +33 1 48 81 46 32
Fax +33 1 47 06 06 48
E-mail

**Louis Lumière Section
(Student)**

Alexandra Carr-Brown
AES Student Section
Ecole Nationale Supérieure
Louis Lumière
7, allée du Promontoire, BP 22
FR-93161 Noisy Le Grand
Cedex, France
Tel. +33 6 18 57 84 41
E-mail
louis

GREECE

Greek Section
Vassilis Tsakiris
Crystal Audio
Aiantos 3a Vrillissia
GR 15235 Athens, Greece
Tel. +30 2 10 6134767
Fax +30 2 10 6137010
E-mail

ISRAEL

Israel Section
Ben Bernfeld Jr.
H. M. Acustica Ltd.
20G/5 Mashabim St..
IL-45201 Hod Hasharon, Israel
Tel./Fax +972 9 7444099
E-mail

ITALY

Italian Section
Carlo Perretta
% AES Italian Section
Piazza Cantore 10
IT-20134 Milan, Italy
Tel. +39 338 9108768
Fax +39 02 58440640
E-mail

Italian Student Section

Franco Grossi, Faculty Advisor
AES Student Section
Viale San Daniele 29
IT-33100 Udine, Italy
Tel. +39 0432227527
E-mail
italian

PORTUGAL

Portugal Section

Rui Miguel Avelans Coelho
R. Paulo Renato 1, 2A
PT-2745-147 Linda-a-Velha
Portugal
Tel. +351 214145827
E-mail
portugal

ROMANIA

Romanian Section

Marcia Taiachin
Radio Romania
60-62 Gr. Berthelot St.
RO-79756 Bucharest, Romania
Tel. +40 1 303 12 07
Fax +40 1 222 69 19

SLOVENIA

Slovenian Section

Tone Seliskar
RTV Slovenija
Kolodvorska 2
SI-1550 Ljubljana, Slovenia
Tel. +386 61 175 2708
Fax +386 61 175 2710
E-mail
slovenian

SPAIN

Spanish Section

Juan Recio Morillas
Spanish Section
C/Florencia 14 3oD
ES-28850 Torrejon de Ardoz
(Madrid), Spain
Tel. +34 91 540 14 03
E-mail
spanish

TURKEY

Turkish Section

Sorgun Akkor
STD
Gazeteciler Sitesi, Yazarlar
Sok. 19/6
Esentepe 80300
Istanbul, Turkey
Tel. +90 212 2889825
Fax +90 212 2889831
E-mail

YUGOSLAVIA

Yugoslavian Section

Tomislav Stanojevic
Sava centre
M. Popovica 9
YU-11070 Belgrade, Yugoslavia
Tel. +381 11 311 1368

SECTIONS CONTACTS**DIRECTORY**

Fax +381 11 605 578
E-mail
yugo

LATIN AMERICAN REGION*Vice President:*

Mercedes Onorato
Talcahuano 141
Buenos Aires, Argentina
Tel./Fax +5411 4 375 0116
E-mail
latin

ARGENTINA**Argentina Section**

Hernan Ranucci
Talcahuano 141
Buenos Aires, Argentina
Tel./Fax +5411 4 375 0116
E-mail

BRAZIL**Brazil Section**

Rosalfonso Bortoni
Rua Doutor Jesuino Maciel,
1584/22
Campo Belo
São Paulo, SP, Brazil 04615-
004
Tel.+55 11 5533-3970
Fax +55 21 2421 0112
E-mail

CHILE**Chile Section**

Andres Pablo Schmidt Ilic
Tonmusik
Hernan Cortez 2768
Ñuñoa, Santiago de Chile
Tel: +56 2 3792064
Fax: +56 2 2513283
E-mail

COLOMBIA**Colombia Section**

Tony Penarredonda Caraballo
Carrera 51 #13-223
Medellin, Colombia
Tel. +57 4 265 7000
Fax +57 4 265 2772
E-mail

MEXICO**Mexican Section**

Javier Posada
Div. Del Norte #1008
Col. Del Valle
Mexico, D.F. MX-03100
Mexico
Tel. +52 5 669 48 79
Fax +52 5 543 60 37
E-mail
mexican

URUGUAY**Uruguay Section**

Rafael Abal
Sondor S.A.
Calle Rio Branco 1530
C.P. UY-11100 Montevideo
Uruguay
Tel. +598 2 901 26 70,
+598 2 90253 88
Fax +598 2 902 52 72
E-mail

VENEZUELA**Taller de Arte Sonoro,
Caracas Section (Student)**

Carmen Bell-Smythe de Leal
Faculty Advisor
AES Student Section
Taller de Arte Sonoro
Ave. Rio de Janeiro
Qta. Tres Pinos
Chuao, VE-1061 Caracas
Venezuela
Tel. +58 14 9292552
Tel./Fax +58 2 9937296
E-mail

Venezuela Section

Elmar Leal
Ave. Rio de Janeiro
Qta. Tres Pinos
Chuao, VE-1061 Caracas
Venezuela
Tel. +58 14 9292552
Tel./Fax +58 2 9937296
E-mail

INTERNATIONAL REGION*Vice President:*

Neville Thiele
10 Wycombe St.
Epping, NSW AU-2121,
Australia
Tel. +61 2 9876 2407
Fax +61 2 9876 2749
E-mail

AUSTRALIA**Adelaide Section**

David Murphy
Krix Loudspeakers
14 Chapman Rd.
Hackham AU-5163
South Australia
Tel. +618 8 8384 3433
Fax +618 8 8384 3419
E-mail

Brisbane Section

David Ringrose
AES Brisbane Section
P.O. Box 642
Roma St. Post Office
Brisbane, Qld. AU-4003, Australia
Office Tel. +61 7 3364 6510
E-mail

Melbourne Section

Graham J. Haynes
P.O. Box 5266

Wantirna South, Victoria

AU-3152, Australia
Tel. +61 3 9887 3765
Fax +61 3 9887 1688
E-mail
mel

Sydney Section

Howard Jones
AES Sydney Section
P.O. Box 766
Crows Nest, NSW AU-2065
Australia
Tel. +61 2 9417 3200
Fax +61 2 9417 3714
E-mail

HONG KONG**Hong Kong Section**

Henry Ma Chi Fai
HKAPA, School of Film and
Television
1 Gloucester Rd.
Wanchai, Hong Kong
Tel. +852 2584 8824
Fax +852 2588 1303
E-mail

INDIA**India Section**

Avinash Oak
Western Outdoor Media
Technologies Ltd.
16, Mumbai Samachar Marg
Mumbai 400023, India
Tel. +91 22 204 6181
Fax +91 22 660 8144
E-mail

JAPAN**Japan Section**

Katsuya (Vic) Goh
2-15-4 Tenjin-cho, Fujisawa-shi
Kanagawa-ken 252-0814, Japan
Tel. +81 466 81 0681
Fax +81 466 81 0698
E-mail

KOREA**Korea Section**

Seong-Hoon Kang
Taejeon Health Science College
Dept. of Broadcasting
Technology
77-3 Gayang-dong Dong-gu
Taejeon, Korea
Tel. +82 42 630 5990
Fax +82 42 628 1423
E-mail

MALAYSIA**Malaysia Section**

C. K. Ng
King Musical Industries
Sdn Bhd
Lot 5, Jalan 13/2
MY-46200 Kuala Lumpur
Malaysia
Tel. +603 7956 1668
Fax +603 7955 4926
E-mail

PHILIPPINES**Philippines Section**

Dario (Dar) J. Quintos
125 Regalia Park Tower
P. Tuazon Blvd., Cubao
Quezon City, Philippines
Tel./Fax +63 2 4211790, +63 2
4211784
E-mail

SINGAPORE**Singapore Section**

Cedric M. M. Tio
Apt. Block 237
Bishan Street 22, # 02-174
Singapore 570237
Republic of Singapore
Tel. +65 6887 4382
Fax +65 6887 7481
E-mail

**STUDENT DELEGATE
ASSEMBLY****NORTH/SOUTH
AMERICA REGIONS***Chair:*

Dell Harris
Hampton University Section
(AES)
63 Litchfield Close
Hampton, VA 23669
Tel +1 757 265 1033
E-mail

Vice Chair:

Scott Cannon
Stanford University Section
(AES)
P.O. Box 15259
Stanford, CA 94309
Tel. +1 650 346 4556
Fax +1 650 723 8468
E-mail

**EUROPE/INTERNATIONAL
REGIONS***Chair:*

Blaise Chabanis
Conservatoire de Paris
(CNSMDP) Student Section
(AES)
14, rue de la Faisanderie
FR-77200 Torcy, France
Tel. +336 62 15 29 97
E-mail

Vice Chair:

Werner de Bruijn
The Netherlands Student
Section (AES)
Korvezeestraat 541
NL-2628 CZ Delft
The Netherlands
Home Tel. +31 15 2622995
Office Tel. +31 15 2782021
E-mail
werner

AES CONVENTIONS AND CON

The latest details on the following events are posted on the AES Website: <http://www.aes.org>



113th Convention
Los Angeles, California, USA
Date: 2002 October 5–8
Location: Los Angeles Convention Center, Los Angeles, California, USA

Convention chair:
Floyd Toole
Harman International
8500 Balboa Blvd.
Northridge, CA 91329, USA
Telephone: +1 818 895 5761
Fax: +1 818 893 7139
Email
Papers cochair:
John Strawn
S Systems, Inc.

Telephone: +1 415 927 8856
Email
Papers cochair:
Eric Benjamin
Dolby Laboratories, Inc.
Telephone: +1 415 558 0236
Email
Exhibit information:
Chris Plunkett
Telephone: +1 212 661 8528



114th Convention
Amsterdam, The Netherlands
Date: 2003 March 22–25
Location: RAI Conference and Exhibition Centre, Amsterdam, The Netherlands

Convention chair:
Peter A. Swarte
P.A.S. Electro-Acoustics
Graaf Adolfstraat 85
5616 BV Eindhoven
The Netherlands
Telephone: +31 40 255 0889
Email
Papers chair: Ronald M. Aarts
Vice Chair: Erik Larsen

DSP-Acoustics & Sound
Reproduction
Philips Research Labs, WY81
Prof. Hostlaan 4
5656 AA Eindhoven, The Netherlands
Telephone: +31 40 274 3149
Fax: +31 40 274 3230
Email



23rd International Conference
Copenhagen, Denmark
“Signal Processing in Audio Recording and Reproduction”
Date: 2003 May 23–25
Location: Marienlyst Hotel, Helsingør, Copenhagen, Denmark

Conference chair:
Per Rubak
Aalborg University
Fredrik Bajers Vej 7 A3-216
DK-9220 Aalborg Ø
Denmark
Telephone: +45 9635 8682
Email

Papers cochair:
Jan Abildgaard Pedersen
Bang & Olufsen A/S
Peter Bangs Vej 15
P.O. box 40,
DK-7600 Struer
Phone: +45 9684 1122
Email
Papers cochair:
Lars Gottfried Johansen



24th International Conference
Banff, Canada
“Multichannel Audio: The New Reality”
Date: 2003 June 26–28
Location: The Banff Centre, Banff, Alberta, Canada

Conference chair:
Theresa Leonard
The Banff Centre
Banff, Canada
Email
Conference vice chair:
John Sorensen
The Banff Centre
Banff, Canada
Email

Papers chair:
Geoff Martin
Email



11th Regional Convention
Tokyo, Japan
Date: 2003 July 7–9
Location: Science Museum, Chiyoda, Tokyo, Japan

Convention chair:
Kimio Hamasaki
NHK Science & Technical
Research Laboratories
Telephone: +81 3 5494 3208

Fax: +81 3 5494 3219
Email
Convention vice chair:
Hiroaki Suzuki
Victor Company of Japan (JVC)
Telephone: +81 45 450 1779
Email
victor.co.jp
Papers chair: Shinji Koyano
Pioneer Corporation

Telephone: +81 49 279 2627
Fax: +81 49 279 1513
Email:
pioneer.co.jp
Workshops chair:
Toru Kamekawa
Tokyo National University of Fine
Art & Music
Telephone: +81 3 297 73 8663
Fax: +81 297 73 8670
Email



115th Convention
New York, NY, USA
Date: 2003 October 10–13
Location: Jacob K. Javits Convention Center, New York, New York, USA

REFERENCES

INFORMATION FOR AUTHORS

Email

Call for papers: Vol. 50, No. 1/2, p. 100 (2002 January/February)

Call for workshops participants: Vol. 50, No. 3, p. 206 (2002 March)

Convention preview: Vol. 50, No. 7/8, pp. 606–628 (2002 July/August)

Convention report: This issue, pp. 934–978 (2002 November)

Exhibit information:

Thierry Bergmans

Telephone: +32 2 345 7971

Fax: +32 2 345 3419

Email

Call for papers: Vol. 50, No. 6, p. 535 (2002 June)

Aalborg University

Niels Jernes Vej 14, 4

DK-9220 Aalborg Ø

Phone: +45 9635 9828

Email

Call for papers: Vol. 50, No. 9, p. 737 (2002 September)

Call for contributions: Vol. 50, No. 10, pp. 851–852 (2002 October)

Exhibit chair: Tadahiko Nakaoki

Pioneer Business Systems Division

Telephone: +81 3 3763 9445

Fax : +81 3 3763 3138

Email

Section contact: Vic Goh

Email

Call for papers: This issue, pp. 1001 (2002 November)

Presentation

Manuscripts submitted should be typewritten on one side of ISO size A4 (210 x 297 mm) or 216-mm x 280-mm (8.5-inch x 11-inch) paper with 40-mm (1.5-inch) margins. All copies including abstract, text, references, figure captions, and tables should be double-spaced. Pages should be numbered consecutively. Authors should submit an original plus two copies of text and illustrations.

Review

Manuscripts are reviewed anonymously by members of the review board. After the reviewers' analysis and recommendation to the editors, the author is advised of either acceptance or rejection. On the basis of the reviewers' comments, the editor may request that the author make certain revisions which will allow the paper to be accepted for publication.

Content

Technical articles should be informative and well organized. They should cite original work or review previous work, giving proper credit. Results of actual experiments or research should be included. The *Journal* cannot accept unsubstantiated or commercial statements.

Organization

An informative and self-contained abstract of about 60 words must be provided. The manuscript should develop the main point, beginning with an introduction and ending with a summary or conclusion. Illustrations must have informative captions and must be referred to in the text.

References should be cited numerically in brackets in order of appearance in the text. Footnotes should be avoided, when possible, by making parenthetical remarks in the text.

Mathematical symbols, abbreviations, acronyms, etc., which may not be familiar to readers must be spelled out or defined the first time they are cited in the text.

Subheads are appropriate and should be inserted where necessary. Paragraph division numbers should be of the form 0 (only for introduction), 1, 1.1, 1.1.1, 2, 2.1, 2.1.1, etc.

References should be typed on a manuscript page at the end of the text in order of appearance. References to periodicals should include the authors' names, title of article, periodical title, volume, page numbers, year and month of publication. Book references should contain the names of the authors, title of book, edition (if other than first), name and location of publisher, publication year, and page numbers. References to AES convention preprints should be replaced with *Journal* publication citations if the preprint has been published.

Illustrations

Figure captions should be typed on a separate sheet following the references. Captions should be concise. All figures

should be labeled with author's name and figure number.

Photographs should be black and white prints without a halftone screen, preferably 200 mm x 250 mm (8 inch by 10 inch).

Line drawings (graphs or sketches) can be original drawings on white paper, or high-quality photographic reproductions.

The size of illustrations when printed in the *Journal* is usually 82 mm (3.25 inches) wide, although 170 mm (6.75 inches) wide can be used if required. Letters on original illustrations (before reduction) must be large enough so that the smallest letters are at least 1.5 mm (1/16 inch) high when the illustrations are reduced to one of the above widths. If possible, letters on all original illustrations should be the same size.

Units and Symbols

Metric units according to the System of International Units (SI) should be used. For more details, see G. F. Montgomery, "Metric Review," *JAES*, Vol. 32, No. 11, pp. 890–893 (1984 Nov.) and J. G. McKnight, "Quantities, Units, Letter Symbols, and Abbreviations," *JAES*, Vol. 24, No. 1, pp. 40, 42, 44 (1976 Jan./Feb.). Following are some frequently used SI units and their symbols, some non-SI units that may be used with SI units (Δ) and some non-SI units that are deprecated (\square)

Unit Name	Unit Symbol
ampere	A
bit or bits	spell out
bytes	spell out
decibel	dB
degree (plane angle) (Δ)	$^{\circ}$
farad	F
gauss (\square)	Gs
gram	g
henry	H
hertz	Hz
hour (Δ)	h
inch (\square)	in
joule	J
kelvin	K
kilohertz	kHz
kilohm	k Ω
liter (Δ)	l, L
megahertz	MHz
meter	m
microfarad	μ F
micrometer	μ m
microsecond	μ s
milliampere	mA
millihenry	mH
millimeter	mm
millivolt	mV
minute (time) (Δ)	min
minute (plane angle) (Δ)	'
nanosecond	ns
oersted (\square)	Oe
ohm	Ω
pascal	Pa
picofarad	pF
second (time)	s
second (plane angle) (Δ)	"
siemens	S
tesla	T
volt	V
watt	W
weber	Wb

AES

sustaining member organizations

The Audio Engineering Society recognizes with gratitude the financial support given by its sustaining members, which enables the work of the Society to be extended. Addresses and brief descriptions of the business activities of the sustaining members appear in the October issue of the Journal.

The Society invites applications for sustaining membership. Information may be obtained from the Chair, Sustaining Memberships Committee, Audio Engineering Society, 60 East 42nd St., Room 2520, New York, New York 10165-2520, USA, tel: 212-661-8528. Fax: 212-682-0477.

ACO Pacific, Inc.
Air Studios Ltd.
AKG Acoustics GmbH
AKM Semiconductor, Inc.
Amber Technology Limited
AMS Neve plc
ATC Loudspeaker Technology Ltd.
Audio Limited
Audiomatica S.r.l.
Audio Media/IMAS Publishing Ltd.
Audio Precision, Inc.
AudioScience, Inc.
Audio-Technica U.S., Inc.
AudioTrack Corporation
Autograph Sound Recording Ltd.
B & W Loudspeakers Limited
BMP Recording
British Broadcasting Corporation
BSS Audio
Cadac Electronics PLC
Calrec Audio
Canford Audio plc
CEDAR Audio Ltd.
Celestion International Limited
Cerwin-Vega, Incorporated
ClearOne Communications Corp.
Community Professional Loudspeakers, Inc.
Crystal Audio Products/Cirrus Logic Inc.
D.A.S. Audio, S.A.
D.A.T. Ltd.
dCS Ltd.
Deltron Emcon Limited
Digidesign
Digigram
Digital Audio Disc Corporation
Dolby Laboratories, Inc.
DRA Laboratories
DTS, Inc.
DYNACORD, EVI Audio GmbH
Eastern Acoustic Works, Inc.
Eminence Speaker LLC

Event Electronics, LLC
Ferrotec (USA) Corporation
Focusrite Audio Engineering Ltd.
Fostex America, a division of Foster Electric U.S.A., Inc.
Fraunhofer IIS-A
FreeSystems Private Limited
FTG Sandar TeleCast AS
Harman Becker
HHB Communications Ltd.
Innova SON
Innovative Electronic Designs (IED), Inc.
International Federation of the Phonographic Industry
JBL Professional
Jensen Transformers Inc.
Kawamura Electrical Laboratory
KEF Audio (UK) Limited
Kenwood U.S.A. Corporation
Klark Teknik Group (UK) Plc
Klipsch L.L.C.
Laboratories for Information
L-Acoustics US
Leitch Technology Corporation
Lindos Electronics
Magnetic Reference Laboratory (MRL) Inc.
Martin Audio Ltd.
Meridian Audio Limited
Metropolis Group
Middle Atlantic Products Inc.
Mosses & Mitchell
M2 Gauss Corp.
Music Plaza Pte. Ltd.
Georg Neumann GmbH
Neutrik AG
NVision
NXT (New Transducers Ltd.)
1 Limited
Ontario Institute of Audio Recording Technology
Outline snc
PRIMEDIA Business Magazines & Media Inc.
Prism Sound

Pro-Bel Limited
Pro-Sound News
Radio Free Asia
Rane Corporation
Recording Connection
Rocket Network
Royal National Institute for the Blind
RTI Tech Pte. Ltd.
Rycote Microphone Windshields Ltd.
SADiE
Sanctuary Studios Ltd.
Sekaku Electron Ind. Co., Ltd.
Sennheiser Electronic Corporation
Shure Inc.
Snell & Wilcox Ltd.
Solid State Logic, Ltd.
Sony Broadcast & Professional Europe
Sound Devices LLC
Sound On Sound Ltd.
Soundcraft Electronics Ltd.
Soundtracs plc
Sowter Audio Transformers
SRS Labs, Inc.
Stage Accompany
Sterling Sound, Inc.
Studer North America Inc.
Studer Professional Audio AG
Tannoy Limited
TASCAM
THAT Corporation
TOA Electronics, Inc.
Touchtunes Music Corp.
United Entertainment Media, Inc.
Uniton AG
University of Derby
University of Salford
University of Surrey, Dept. of Sound Recording
VidiPax
Wenger Corporation
J. M. Woodgate and Associates
Yamaha Research and Development