

# canadian acoustics

# acoustique canadienne

Journal of the Canadian Acoustical Association - Journal de l'Association Canadienne d'Acoustique

SEPTEMBER 2005

Volume 33 -- Number 3

SEPTEMBRE 2005

Volume 33 -- Numéro 3

EDITORIAL / EDITORIAL

3

PROCEEDINGS OF THE ACOUSTICS WEEK IN CANADA 2005/ ACTES DE LA SEMAINE CANADIENNE D'ACOUSTIQUE 2005

Table of Contents / Table des matières

4

Conference Calendar	9
Underwater Acoustics	14
Acoustical Instrumentation and Applications	18
Physical Acoustics/Ultrasound	26
Engineering Acoustics and Noise Control	30
Hearing Aids	38
Vibration Control	42
Hearing Science and Hearing Conservation	44
Speech Sciences I	48
Signal Processing and Numerical Methods I	54
Signal Processing and Numerical Methods II	60
Architectural Acoustics	68
Physiological and Bioacoustics	72
Engineering Acoustics II	76
Speech Sciences II	78
Biomedical Ultrasound I	84
Speech Sciences III	94
Biomedical Ultrasound II	100
Abstracts for Presentations without Summary Papers	110

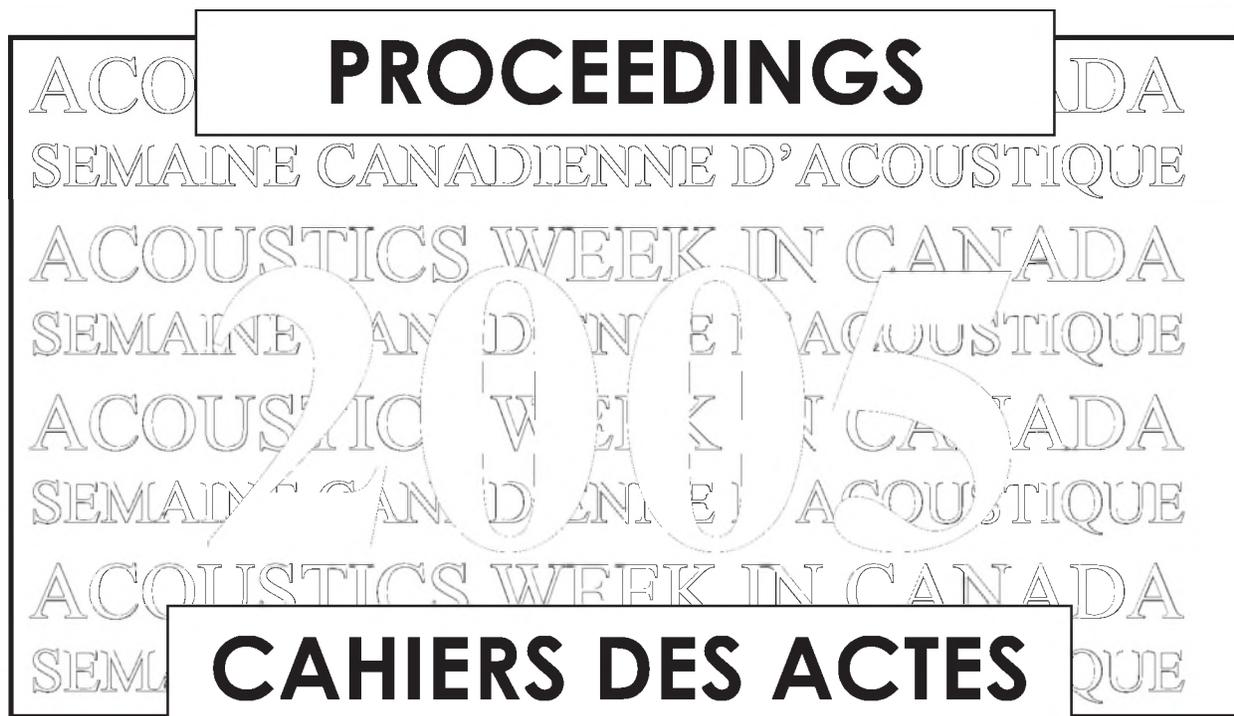
Other Features / Autres Rubriques

News / Informations

116

Book Reviews

120



# canadian acoustics

THE CANADIAN ACOUSTICAL ASSOCIATION  
P.O. BOX 1351, STATION "F"  
TORONTO, ONTARIO M4Y 2V9

CANADIAN ACOUSTICS publishes refereed articles and news items on all aspects of acoustics and vibration. Articles reporting new research or applications, as well as review or tutorial papers and shorter technical notes are welcomed, in English or in French. Submissions should be sent directly to the Editor-in-Chief. Complete instructions to authors concerning the required camera-ready copy are presented at the end of this issue.

CANADIAN ACOUSTICS is published four times a year - in March, June, September and December. The deadline for submission of material is the first day of the month preceeding the issue month. Copyright on articles is held by the author(s), who should be contacted regarding reproduction. Annual subscription: \$20 (student); \$60 (individual, institution); \$250 (sustaining - see back cover). Back issues (when available) may be obtained from the CAA Secretary - price \$10 including postage. Advertisement prices: \$500 (centre spread); \$250 (full page); \$150 (half page); \$100 (quarter page). Contact the Associate Editor (advertising) to place advertisements. Canadian Publication Mail Product Sales Agreement No. 0557188.

# acoustique canadienne

L'ASSOCIATION CANADIENNE D'ACOUSTIQUE  
C.P. 1351, SUCCURSALE "F"  
TORONTO, ONTARIO M4Y 2V9

ACOUSTIQUE CANADIENNE publie des articles arbitrés et des informations sur tous les domaines de l'acoustique et des vibrations. On invite les auteurs à soumettre des manuscrits, rédigés en français ou en anglais, concernant des travaux inédits, des états de question ou des notes techniques. Les soumissions doivent être envoyées au rédacteur en chef. Les instructions pour la présentation des textes sont exposées à la fin de cette publication.

ACOUSTIQUE CANADIENNE est publiée quatre fois par année - en mars, juin, septembre et décembre. La date de tombée pour la soumission de matériel est fixée au premier jour du mois précédant la publication d'un numéro donné. Les droits d'auteur d'un article appartiennent à (aux) auteur(s). Toute demande de reproduction doit leur être acheminée. Abonnement annuel: \$20 (étudiant); \$60 (individuel, société); \$150 (soutien - voir la couverture arrière). D'anciens numéros (non-épuisés) peuvent être obtenus du Secrétaire de l'ACA - prix: \$10 (affranchissement inclus). Prix d'annonces publicitaires: \$500 (page double); \$250 (page pleine); \$150 (demi page); \$100 (quart de page). Contacter le rédacteur associé (publicité) afin de placer des annonces. Société canadienne des postes - Envois de publications canadiennes - Numéro de convention 0557188.

---

## EDITOR-IN-CHIEF / RÉDACTEUR EN CHEF

**Ramani Ramakrishnan**  
Department of Architectural Science  
Ryerson University  
350 Victoria Street  
Toronto, Ontario M5B 2K3  
Tel: (416) 979-5000; Ext: 6508  
Fax: (416) 979-5353  
E-mail: rramakri@ryerson.ca

## EDITOR / RÉDACTEUR

**Chantai Laroche**  
Programme d'audiologie et d'orthophonie  
École des sciences de la réadaptation  
Université d'Ottawa  
451, chemin Smyth, pièce 3062  
Ottawa, Ontario K1H 8M5  
Tél: (613) 562-5800 # 3066; Fax: (613) 562-5428  
E-mail: claroche@uottawa.ca

## ASSOCIATE EDITORS / REDACTEURS ASSOCIES

### Advertising / Publicité

**Karen Fraser**  
City of Brampton  
2 Wellington Street West, 5th floor  
Brampton, Ontario L6Y 4R2  
Tel: (905) 874-2489  
Fax: (905) 874-2599  
E-mail: karen.fraser@city.brampton.on.ca

### News / Informations

**Steven Bilawchuk**  
aci Acoustical Consultants Inc.  
Suite 107, 9920-63rd Avenue  
Edmonton, Alberta T6E 0G9  
Tel: (780) 414-6373  
Fax: (780) 414-6376  
E-mail: stevenb@aciacoustical.com



**Acoustics Week in Canada  
Lamplighter Inn and Conference Centre  
591 Wellington Road, London, Ontario  
October 11 - 14, 2005**

**ORGANIZING COMMITTEE**

**M. Cheesman, Convenor  
National Centre for Audiology, School of Communication Sciences and Disorders  
University of Western Ontario**

**V. Parsa, Technical Program Chair  
National Centre for Audiology, School of Communication Sciences and Disorders  
and Department of Electrical and Computer Engineering  
University of Western Ontario**

**S. Sinclair Moodie, Exhibits Coordinator  
National Centre for Audiology  
University of Western Ontario**

**D. Grainger, Audiovisual Coordinator  
National Centre for Audiology  
University of Western Ontario**

**EXHIBITORS**

**H. L. Blachford Ltd.  
Briel & Kjaer  
Dalimar Instruments  
Durisol  
Eckel Industries of Canada Ltd.  
Novel Dynamics Test Inc.  
Scantek Inc.  
West Caldwell Calibrations Laboratories Inc.**

**Many thanks to Christian Giguère and Steve Beaulac for assistance with French translations.**

# SOUND SOLUTIONS FROM



## *Integrated Solutions from World Leaders*

- Precision Measurement Microphones
- Intensity Probes
- Outdoor Microphones
- Hydrophones
- Ear Simulation Devices
- Speech Simulation Devices
- Calibrators
- Array Microphones
- Sound Quality
- Sound Intensity
- Sound Power
- Room Acoustics
- Noise Monitoring
- Dynamic Signal Analyzers
- Multi Channel Dynamic Analyzer/Recorders
- Electro Dynamic Shaker Systems
- Advanced Sound & Vibration Level Meters
- Doppler Laser Optical Transducers (Laser Vibrometers)



**Ottawa**  
 200-440 Laurier Ave. West, K1R 7X6  
 613-598-0026 fax: 616-598-0019  
 info@noveldynamics.com



**Toronto**  
 RR#2 13652 4th Line. Acton, Ont. L7J 2L8  
 519-853-4495 fax: 519-853-3366  
 metelka@aztec-net.com

## EDITORIAL / EDITORIAL

The 2005 annual conference of the Canadian Acoustical Association will be held at the Lamplighter Inn and Conference Centre in London, Ontario, Canada on October 11-14, 2005. We have received a wide variety of submissions, thanks to your involvement in your Association. Vijay Parsa and the rest of the Organizing Committee look forward to joining you in a lively discussion of acoustical issues, from speech communication to underwater acoustics.

We have three plenary lectures on the schedule. Richard Seewald, Canada Research Chair in Infant Hearing, will speak about his work on hearing loss and hearing aids in infants. Brock Fenton, Professor and Chair of the Biology Department at the University of Western Ontario, will discuss echolocation and bat behavior. Tim Kelsall, of Hatch Associates Ltd., will update us on the acoustics-related activities of the Canadian Standards Association, following which Alberto Behar will lead us in a discussion on how CAA members can become more involved.

Many thanks to those who have assisted in making this conference a success so far: our exhibitors and sponsors, special session chairs, paper contributors, organizing committee, and many student volunteers.

Take a look at the preliminary program on the following pages and be sure to join us in London 2005.

M. Cheesman

La conférence de 2005 annuelle de l'Association canadienne d'acoustique se tiendra à Lamplighter Inn et le Centre de Conference à London, Ontario, Canada le 11 au 14 octobre 2005. Nous avons reçu une grande variété de soumissions, grâce à votre participation dans votre Association. Vijay Parsa et le reste du Comité d'organisation attendent avec intérêt de vous joindre dans une discussion animée des issues acoustiques, de la communication de la parole à l'acoustique sous-marine.

Nous avons trois conférences plénières sur le programme. Richard Seewald, Chaire de recherche du Canada sur l'audition chez l'enfant, parlera au sujet de son travail sur les déficiences auditives chez les jeunes enfants et des améliorations rendues possibles grâce aux appareils auditifs. Brock Fenton, Professeur et la Chaire du Département de Biologie à l'Université de Western Ontario, discutera l'écholocation et le comportement des chauve-souris. Tim Kelsall, de Hatch Associates Ltd., nous mettra à jour sur les activités acoustique relatant sur l'Association Canadienne de Normes, suivi par Alberto Behar qui nous mènera dans une discussion sur la façon dont les membres de CAA peuvent devenir plus impliqués.

Beaucoup de mercis à ceux qui ont aidé à faire à cette conférence un succès jusqu'ici: nos exposants et commanditaires, chaires spéciales de session, contributeurs de papier, comité d'organisation, et beaucoup de volontaires d'étudiant.

Jetez un coup d'oeil au programme préliminaire aux pages suivantes et soyez sûr de nous joindre à London 2005."

M. Cheesman

### WHAT'S NEW ??

Promotions  
Deaths  
New jobs  
Moves

Retirements  
Degrees awarded  
Distinctions  
Other news

Do you have any news that you would like to share with Canadian Acoustics readers? If so, send it to:

**Steven Bilawchuk**, aci Acoustical Consultants Inc., Edmonton, Alberta, Email: [stevenb@aciacoustical.com](mailto:stevenb@aciacoustical.com)

### QUOI DE NEUF ?

Promotions  
Décès  
Offre d'emploi  
Déménagements

Retraites  
Obtention de diplômes  
Distinctions  
Autres nouvelles

Avez-vous des nouvelles que vous aimeriez partager

## TABLE OF CONTENTS/TABLES DES MATIÈRES

Organising Committee, Sponsors and Exhibitors	1
Editorial/Éditorial	3
Table of Contents/Tables des matières	4
Conference Calendar	9
<b>Underwater Acoustics</b>	
Data Covariance Estimation for Geoacoustic Inversion Stan E. Dosso	14
Array Element Localization Accuracy and Survey Design Stan E. Dosso and Gordon R. Ebbeson	16
<b>Acoustical Instrumentation and Applications</b>	
Monitoring Fluidized-Bed Drying of Pharmaceutical Granules using Acoustic Sensors Derek Daniher, Lauren Briens	18
Development Of An Acoustic Method For The Measurement Of Mixing And Drying In A Vibrated Fluidized Bed Garret Book, Lauren Briens, Cedric Briens	20
Flow Regime Detection in Pneumatic Transport of Particulates Using Non-Intrusive Acoustic Probes Katherine Albion, Lauren Briens, Garret Book, Franco Berruti & Cedric Briens	22
Detection Of Conditions Leading To Breakage Of Pharmaceutical Tablets In Pneumatic Transport Katherine Albion, Lauren Briens*, Cedric Briens, and Franco Berruti	24
<b>Physical Acoustics/Ultrasound</b>	
Single Bubble Sonoluminescence: Effects of Signal Amplitude Modulation Borko Djurkovic, Igor Mastikhin, Dennis Tokaryk	26
Using Ultrasonic and Vision Sensors within Extended Kalman Filter for Robot Navigation Zhenhe Chen, Ranga Rodrigo, Vijay Parsa and Jagath Samarabandu	28
<b>Engineering Acoustics and Noise Control</b>	
Kalman Filtering of Acoustic Emission Signal Generated by Laser-Material Interactions Dmitry S. Solodovnikov, Evgueni V. Bordatchev	30
Collaborative Noise Control – A Case Study Ramani Ramakrishnan and Werner Richarz	32
Use of Psychoacoustic Metrics for the Analysis of Next Generation Computer Cooling Fan Noise Colin Novak, Helen Ule, Robert Gaspar, and Robert Wiley	34
Effect of next Generation Computer Cooling Fan Speed on Acoustic Noise Emissions Helen Ule, Colin Novak and Robert Gaspar	36

## **Hearing Aids**

Subband Adaptive Modeling of Digital Hearing Aids 38  
Michael Wirtzfeld, Vijay Parsa

An Evaluation of the Impact and Feasibility of the ANSI S3.22 2003 Standard for Testing Hearing Instruments 40  
Andreas Seelisch and Susan Scollie

## **Vibration Control**

Tuned Mass Dampers For Soldier Field Stadium Grandstand Vibration 42  
Mal P. Sacks, Ramin Behboudi, and John C. Swallow

## **Hearing Science and Hearing Conservation**

Normative Threshold Levels for a Calibrated, Computer-Assisted Version of the Ling Six-Sound Test 44  
Julianne J. Tenhaaf and Susan D. Scollie

Persistence of Auditory Streaming Preserves release from masking 46  
Kachina Allen, Simon Carlile, David Alaisand Karen Froud

## **Speech Sciences I**

Toward Better Automatic Speech Recognition 48  
D. O'Shaughnessy, W. Wang, W. Zhu, V. Barreaud, T. Nagarajan, and R. Muralishankar

Output-based Sound Quality Evaluation Using Statistical Model Method 50  
Guo Chen, Vijay Parsa

An Exploratory study onto the role of phonemic categories in speech perception 52  
Yun Zheng, Christian Giguère

## **Signal Processing and Numerical Methods I**

Measuring the Dynamic Performance of Video Conferencing Systems 54  
David I. Havelock, and David Green

Mistakes, Errors And Uncertainties 56  
Alberto Behar

Time-frequency Signal Decompositions for Audio and Speech Processing 58  
Karthikeyan Umapathy and Sridhar Krishnan

## **Signal Processing and Numerical Methods II**

Hearing Aid Modelling with Adaptive Nonlinear Filters 60  
Ranga Rodrigo, Zhenhe Chen, Vijay Parsa and Jagath Samarabandhu

Advanced Signalling Processing Technologies for Intelligent Hearing Aids 64  
Henry Luo and Horst Arndt

Wavelet Packets-based Speech Enhancement for Hearing Aids Application 66  
Jiming Yang and Sridhar Krishnan

## **Architectural Acoustics**

Vertical flanking sound transmission via the wall-floor junction in wood framed construction 68  
J. David Quirt, Trevor R.T. Nightingale, Robin E. Halliwell

Predicting Transmitted Levels and Audible Effects for Meeting Room Speech Security 70  
J.S. Bradley, M. Apfel and B.N. Gover

## **Physiological and Bioacoustics**

Individual and Roost Identification Using the Echolocation Calls of Wild Big Brown Bats, *Eptesicus fuscus* 72  
Elisabeth van Stam

Measurement and Modelling of the Response of the Cat Eardrum to Large Static Pressures 74  
Hanif M. Ladak, W. Robert J. Funnell, Willem F. Decraemer and Joris J. J. Dirckx

## **Engineering Acoustics II**

Evaluation of Pavement Noise Using Statistical Techniques 76  
Yuen-Ting Fiona Leung, Susan Tighe, Scott Penton, and Peter VanDelden

## **Speech Sciences II**

Cross-Linguistic Influences On Infant Babbling 78  
Karen Mattock, Susan Rvachew, And Linda Polka

Relationship Between Measures Of Intelligibility And Phonetic Accuracy In Children With And Without Cleft Palate 80  
Carrie L. Gotzke and Megan M. Hodge

Acoustic And Articulatory Space Before And After Lateral Tongue Resections In Oral Cancer Patients 82  
Orchid Rastadmehr, Tim Bressmann, Chiang-Le Heng, Jonathan C. Irish

## **Biomedical Ultrasound I**

Finite-Element Modelling of Acoustic Wave Scattering from Fluid, Rigid and Elastic Objects 84  
Omar Falou, J. Carl Kumaradas, and Michael C. Kolios

First-Order Speckle Statistics Of Ultrasound Breast Images Synthesized From A Computational Anatomy Model 86  
Yi-Ting Shen and James C. Lacefield

Development of Prototype Phased Array Ultrasound Systems for Hyperthermia and Targeted Drug Delivery 88  
Don Chorman and Robert J. McGough

Mode Scanning For Ultrasound Phased Arrays with Six Planes of Symmetry 90  
Duo Chen, and Robert J. McGough

Computer Modeling of a Hybrid RF/US Phased Array System for Hyperthermia Cancer Treatments in the Intact Breast 92  
Liyong Wu, Xiaozheng Zeng, and Robert J. McGough

## **Speech Sciences III**

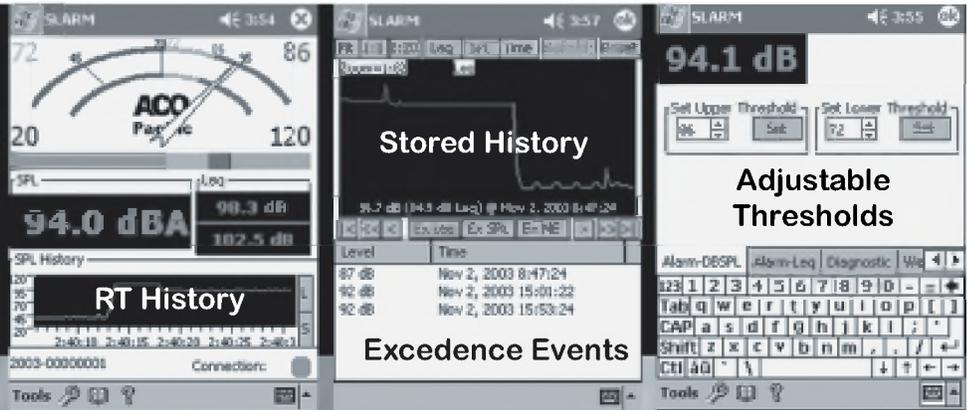
Effects of Multi-Talker Background Noise on the Intensity of Spoken Sentences in Parkinson's Disease 94  
Scott Adams, Olga Haralabous, Allyson Dykstra, Kayla Abrams, and Mandar Jog

The Emotional Prosody on Nasalance Scores Tim Bressman and Jayanthi Sasisekaran	96
Lingual Protrusion and Elevation in Lingual Dystonia: A Hypothesis Allyson D. Dykstra, Scott G. Adams, and Mandar Jog	98
<b>Biomedical Ultrasound II</b>	
The Effect of Packing Order on Ultrasound Backscatter from Cells at Different Volume Fractions Ralph E. Baddour and Michael C. Kolios	100
The Fast Nearfield Method Applied to Axisymmetric Radiators James F. Kelly, Xiaozheng Zeng, and Robert J. McGough	102
Multi-planar Angular Spectrum Approach Applied to Pressure Field Calculations of Spherically Focused Pistons Xiaozheng Zeng, James F. Kelly, and Robert J. McGough	104
A Time-Space Decomposition Method For Fast Calculations Of Transient Pressure Fields Generated By Ultrasound Phased Arrays James F. Kelly and Robert J. McGough	106
<b>Abstracts for Presentations without Summary Papers</b>	110

# Noise Pollution

## The SLARM™ Solution

**PDA & Laptop  
Displays  
Wired  
Wireless**



The **SLARM™** developed in response to increased emphasis on hearing conservation and comfort in the community and workplace incorporates **ACOustAlert™** and **ACOustAlarm™** technology. Making the **SLARM™** a powerful and versatile sound monitoring/alarm system.

### Typical Applications Include:

#### Community

- ◆ Amphitheaters
- ◆ Outdoor Events
- ◆ Nightclubs/Discos
- ◆ Churches
- ◆ Classrooms

#### Industrial

- ◆ Machine/Plant Noise
- ◆ Fault Detection
- ◆ Marshalling Yards
- ◆ Construction Sites
- ◆ Product Testing

### FEATURES

- ✓ **Wired and Wireless (opt)**
- ✓ **USB, Serial, and LAN(opt) Connectivity**
- ✓ **Remote Displays and Programming**
- ✓ **SPL, Leq, Thresholds, Alert and Alarm**
- ✓ **Filters (A,C,Z), Thresholds, Calibration**
- ✓ **Multiple Profiles (opt)**
- ✓ **100 dB Display Range:**
- ✓ **20-120 dBSPL and 40-140 dBSPL**
- ✓ **Real-time Clock/Calendar**
- ✓ **Internal Storage: 10+days @1/sec**
- ✓ **Remote Storage of 1/8 second events**
- ✓ **7052S Type 1.5™ Titanium Measurement Mic**



2604 Read Ave., Belmont, CA 94002 Tel: 650-595-8588 FAX: 650-591-2891  
[www.acopacific.com](http://www.acopacific.com) [acopac@acopacific.com](mailto:acopac@acopacific.com)

**ACOustics Begins With ACO™**

**CONFERENCE PROGRAM****TUESDAY, OCTOBER 11, 2005**

16:00 – 20:00	CAA Board of Directors Meeting
18:00 - 22:00	CSA Standards Meeting – open to all interested persons
18:00 – 20:00	Opening registration

**WEDNESDAY, OCTOBER 12, 2005**

08:00 – 15:30	CONFERENCE REGISTRATION		
08:45 – 09:00	CONFERENCE OPENING		
09:00 – 10:00	PLENARY SPEAKER: Richard Seewald		
09:00 – 17:00	EXHIBITS OPEN		
10:00 – 10:20		COFFEE BREAK	
10:20 – 12:00			
UNDERWATER ACOUSTICS		ACOUSTICAL INSTRUMENTATION & APPLICATIONS	
Data Covariance Estimation for Geoacoustic Inversion	S. E. Dosso	Understanding instrument calibration	R. Peppin, M.Buzduga
Acoustic Array Element Localization Accuracy and Survey Design	S. E. Dosso, G.Ebbeson	Monitoring fluidized-bed drying of pharmaceutical granules using acoustic sensors	D. Daniher, L. Briens
A smart hydrophone for underwater acoustics	N. Somayajula, B. Yan, S.E. Prasad, R. Blacow	Development of an acoustic method for the measurement of mixing and drying in a vibrated fluidized bed	G. Book, L. Briens, C. Briens
PHYSICAL ACOUSTICS/ULTRASOUND		Flow regime detection in pneumatic transport of particulates using non-intrusive acoustic probes	K. Albion, L. Briens, G. Book, F. Berruti, C. Briens
MRI measurements of acoustically cavitated fluid in a standing wave	I. Mastikhin, B. Newling		
Single bubble sonoluminescence: Effects of signal amplitude modulation	B. Djurkovic, I.V.Mastikhin, D. Tokaryk	Detection of conditions leading to breakage of Pharmaceutical tablets in pneumatic transport	K. Albion, L. Briens, C. Briens, F. Berruti
Using ultrasonic and vision sensors within extended Kalman filter for robot navigation	Z. Chen, R. Rodrigo, V. Parsa, J. Samarabandu		
12:20 – 13:20		LUNCH	

13:20 – 15:20			
ENGINEERING ACOUSTICS & NOISE CONTROL - I		HEARING AIDS	
Kalman filtering of acoustic emission signal generated by the laser-material interactions	D. S. Solodovnikova, E. V. Bordatchev	Subband adaptive modelling of hearing aids	M. Wirtzfeld, V. Parsa
Collaborative Noise Control - A Case Study	R. Ramakrishnan, W. Richarz	The effects of changes to the settings used for AGC tests	A. Seelisch, S. Scollie
Use of psychoacoustic metrics for the analysis of next generation computer cooling fan noise	C. Novak, H. Ule, R. Gaspar, R. Wiley	Using active control to reduce the occlusion effect in hearing aids	Jim Ryan
Investigation of fan-tip/heat-sink clearance on fan noise	H. Ule, C. Novak, R. Gaspar, R. Wiley	Combined Impact of a Rapid Feedback Manager and Telephone Use on Acoustic Insertion Gain	D.Hayes, D. Eddins, C. Secor
A Tribute to Herbert S. Ribner	W. Richarz		
15:20 – 15:40 COFFEE BREAK			
15:40 – 17:40			
VIBRATION CONTROL		HEARING SCIENCE & HEARING CONSERVATION	
Tuned Mass Dampers for Soldier Field Stadium Grandstand Vibration	M. P. Sacks, R. Behboudi, J. C. Swallow	Normative threshold levels for a calibrated, computer-assisted version of the Ling Six-Sound test.	Tenhaaf, J. J., Scollie, S.D
Transverse vibrations of centrifugally stiffened tapered beams using Mathematica	D. Jackson, S. O. Oyadiji	Noise exposures caused by respirators in neonatal intensive care	A. Mayrand, , S. D. Scollie, H. Roukema, M. Cheesman
Flexural sensing using piezoelectric	M.R. Kermani, M. Moallem, R. V. Patel	Sound sense: Saving our students' hearing	G. Hannan
		Persistence of auditory streaming preserves release from masking	K. Allen, S. Carlile, D. Alais, K. Froud
		Physiologic indicators of auditory dysfunction in children with suspected auditory processing disorders	P. Allen, C. Allan, N. Chelehmazadeh
19:00 – 21:00 TOUR OF NATIONAL CENTRE FOR AUDIOLOGY AND UWO ACOUSTIC FACILITES			

## THURSDAY, OCTOBER 13, 2005

08:00 – 15:30				CONFERENCE REGISTRATION			
08:00 – 13:00				EXHIBITS OPEN			
09:00 – 10:00				PLENARY SPEAKER: Brock Fenton			
10:00 – 10:20				COFFEE BREAK			
10:20 – 12:00							
SPEECH SCIENCES - I				SIGNAL PROCESSING & NUMERICAL METHODS – I			
Toward better automatic speech recognition		D. O'Shaughnessy, W. Wang, W. Zhu, V. Barreaud, T. Nagarajan. R.		Videoconferencing dynamic performance metrics		D.I. Havelock, D. Green	
Speech quality estimation using the Bayesian model		G. Chen, V. Parsa		Mistakes, errors and uncertainties		A. Behar	
Auditory perceptual evaluation voice acceptability and listener comfort in non-normal vocal signals.		P. C. Doyle, M. Durdle, P. G. Beaudin, T. L. Eadie		System identification with adaptive lattice filters for speech data		W. Shi, V. Parsa, J. Samarabandu	
An experimental study on the role of phonetic categories in speech perception		Y. Zheng, C. Giguère		Time-frequency signal decompositions for audio and speech processing		K. Umopathy, S. Krishnan	
				Acoustic ray tracing for 3D environment simulation		L. Meng, D. Gerhard	
12:00 – 13:00 LUNCH							
13:00 – 14:40							
SIGNAL PROCESSING & NUMERICAL METHODS - II				ARCHITECTURAL ACOUSTICS			
Hearing-aid modeling with adaptive nonlinear filters		R. Rodrigo, Z. Chen, V. Parsa, J. Samarabandu		Vertical flanking sound transmission via the wall-floor junction in wood framed construction		D. Quirt, T. Nightingale, R. Halliwell	
Advanced signaling processing technologies for intelligent hearing aids		H. Luo, H. Arndt		Predicting transmitted levels and audible effects for meeting room speech security		J.S. Bradley, B.N. Gover	
Wavelet packets-based speech enhancement for hearing aids application		J. Yang, S. Krishnan					
A modal analysis of an Indian Gong		P. G. Shrivage, S.Parmeswaran, K. V.deSa					
14:40 – 15:00 COFFEE BREAK							

15:00 – 16:20			
PHYSIOLOGICAL AND BIOACOUSTICS		ENGINEERING ACOUSTICS – II	
Recognizing individual wild Big brown bats ( <i>Eptesicus fuscus</i> ) using their echolocation calls	E. van Stam	Evaluation of pavement noise using statistical techniques	Y-T. F. Leung, P. VanDelden, S. Tighe, S. Penton
Characteristics of chimney swift in-flight vocalizations	J. Bouchard	Evaluating the accuracy of determining hourly road traffic noise levels using typical versus actual road traffic distributions	P. VanDelden, J. Wrobel, M. Li, C. Palis, S. Penton
Distress calls in Neotropical bats	J. J. Nagel	Validation of the STEAM rail traffic noise prediction model – initial results	P. VanDelden, C. Palis, D. Chin-Quee, S. Penton
Measurement and modelling of the response of the cat	H. M. Ladak, W. R. J. Funnell, W. F. Decraemer, J. J. J. Dirckx	Environmental considerations for noise barriers	J. Tsang, C. Vatcher, S. Penton
17:00 – 18:30 CAA/ACA ANNUAL GENERAL MEETING			
18:45 – 19:15 BANQUET RECEPTION/COCKTAILS			
19:15 - CONFERENCE BANQUET AND AWARDS CEREMONY (except STUDENT PRESENTATION AWARDS)			

## FRIDAY, OCTOBER 14 2005

08:00 – 13:30 CONFERENCE REGISTRATION			
09:00 – 10:00 PLENARY SPEAKER & ROUNDTABLE DISCUSSION: Tim Kelsall and Alberto Behar Topic: Activities in Canadian Acoustical Standards			
10:00 – 10:20 COFFEE BREAK			
10:20 – 12:00			
SPEECH SCIENCES - II		BIOMEDICAL ULTRASOUND - I	
Cross-linguistic influences on infant babbling	K. Mattock, S. Rvachew, L. Polka	Finite element modeling of acoustic wave scattering from fluid, rigid and elastic objects	O. Falou, M. C. Kolios
Relationship between measures of intelligibility and phonetic accuracy in children with and without cleft palate	C. L. Gotzke, M. M. Hodge	First-order speckle statistics of ultrasound breast images synthesized from a computational anatomy model	Y-T. Shen, J. C. Lacefield

Exploring the mathematical relationship between acoustic and visual speech for facial animation	M. Craig, P. van Lieshout, W. Wong	Development of prototype ultrasound phased array systems for hyperthermia and targeted drug delivery	D. A. Chorman, R. J. McGough
Perceptual correlates of compensation and adaptation to biteblock perturbation in people who stutter	A. K. Namasivayam, P. H. H. M. van Lieshout	Mode scanning for ultrasound phased arrays with six planes of symmetry	D. Chen, R. J. McGough
Acoustic and articulatory space before and after lateral tongue resections in oral cancer patients	O. Rastadmehr, T. Bressmann, C.-L. Heng, J. C. Irish	Computer modeling of a hybrid rf/us phased array system for hyperthermia cancer treatments in the intact breast	L. Wu, X. Zeng, R. J. McGough
12:00 – 13:00		LUNCH	
13:00 – 14:20			
SPEECH SCIENCES - III		BIOMEDICAL ULTRASOUND - II	
Effects of multi-talker background noise on the intensity of spoken sentences in Parkinson's Disease	S. G. Adams, O. Haralabous, A. D. Dykstra, K. Abrams, M. Jog	The effect of packing order on ultrasound backscatter from cells at different volume fractions	R. E. Baddour
The impact of emotional prosody on nasalance scores	T. Bressmann, J. Sasisekaran	The Fast Nearfield Method Applied to Axisymmetric Radiators	J. F. Kelly, X. Zeng, R. J. McGough
Abnormal lingual protrusion and elevation in lingual dystonia: A hypothesis	A. D. Dykstra, S. G. Adams, M. Jog	Multi-planar angular spectrum approach applied to pressure field calculations of spherically focused pistons	X. Zeng, J. F. Kelly, R. J. McGough
Reconstructing 3D tongue movement trajectories from multiplanar paced ultrasound scans	H. Flowers, T. Bressmann, B. Carmichael, C.-L. Heng	A time-space decomposition method for fast calculations of transient pressure fields generated by ultrasound phased arrays	J. F. Kelly, R. J. McGough
Acoustic and psychophysical relationships and the classification of dysphonic voice	P. C. Doyle, T. L. Eadie		
14:30 – 14:50		PRESENTATION OF STUDENT PAPER AWARDS	

# DATA COVARIANCE ESTIMATION FOR GEOACOUSTIC INVERSION

Stan E. Dosso

School of Earth and Ocean Sciences, University of Victoria, Victoria, B.C., Canada V8W 3P6, sdosso@uvic.ca

## 1. INTRODUCTION

Estimating seabed geoacoustic parameters from ocean acoustic fields measured at an array of sensors represents a challenging but important nonlinear inverse problem. In a Bayesian approach, the posterior probability density (PPD) of the unknown geoacoustic model parameters is formulated in terms of the data information (represented by an appropriate likelihood function) and independent prior information [1, 2]. PPD properties such maximum *a posteriori* (MAP) model estimates (i.e., the most probable parameters) and marginal probability distributions for each parameter can be computed numerically using global optimization [3] and Markov-chain Monte Carlo integration methods [1, 2], respectively.

The likelihood function represents the conditional data uncertainty distribution interpreted as a function of the model parameters for the (fixed) measured data. Hence, specifying data uncertainties is an important component of Bayesian inversion. Data uncertainties must include both measurement error (e.g., errors due to instrumentation and ambient noise) and theory errors (due to the simplified seabed parameterization and idealized treatment of the forward problem). Theory errors, in particular, are difficult to estimate independently, and in most cases physically reasonable assumptions are required about the form of the uncertainty distribution. To date, data errors have been assumed to be Gaussian distributed and spatially uncorrelated (i.e., represented by a diagonal covariance matrix). However, the assumption of uncorrelated errors is often not valid due to theory errors. Neglecting significant error correlations represents the data as more informative than they actually are, and leads to under-estimating parameter uncertainties. In this paper, data-error correlations are estimated to form a full covariance matrix which is explicitly incorporated into the inversion procedure.

## 2. THEORY

Matched-field geoacoustic inversion is based on estimating a model  $\mathbf{m}$  of seabed geoacoustic parameters by matching complex (frequency-domain) acoustic pressure fields  $\mathbf{d}_f$  measured at an  $N$ -sensor array at  $f=1, F$  frequencies. Assuming the data errors are complex Gaussian distributed random variables uncorrelated from frequency to frequency but spatially correlated with covariance matrix  $\mathbf{C}_f$  at the  $f$ -th frequency, the likelihood function is given by

$$L(\mathbf{m}) \propto \prod_{f=1}^F \frac{1}{|\mathbf{C}_f|} \exp[-\mathbf{r}_f(\mathbf{m})^T \mathbf{C}_f^{-1} \mathbf{r}_f(\mathbf{m})], \quad (1)$$

where  $\mathbf{r}_f(\mathbf{m}) = \mathbf{d}_f - \mathbf{d}_f(\mathbf{m})$  are data residuals (difference between measured and modelled data) and  $^T$  represents conjugate transpose. The MAP model  $\hat{\mathbf{m}}$  is computed by minimizing a mismatch function consisting of the negative log-likelihood plus a term representing prior information.

Under the usual assumption of spatially uncorrelated errors, the covariance matrix is diagonal,  $\mathbf{C}_f = \sigma_f^2 \mathbf{I}$ . In this case, the MAP estimate for the standard deviation is  $\hat{\sigma}_f^2 = |\mathbf{r}_f(\hat{\mathbf{m}})|^2 / N$ . However, if significant error correlations exist, the diagonal approximation is inadequate. Assuming the data residuals represent an ergodic random process, a non-parametric estimate of the full covariance matrix is given by

$$\hat{C}_{ij} = \sum_{k=1}^{N-|i-j|} [r_k(\hat{\mathbf{m}}) - \bar{r}(\hat{\mathbf{m}})]^* [r_{k+|i-j|}(\hat{\mathbf{m}}) - \bar{r}(\hat{\mathbf{m}})] / N, \quad (2)$$

where  $\bar{r}$  represents the residual mean and the subscript  $f$  is suppressed for clarity. Covariance elements that are located far off the main diagonal represent error correlations between widely spaced data points. These are expected to be small and are often poorly estimated due to the small number of samples in the average. Hence, it is generally beneficial to damp off-diagonal terms, e.g., by applying a cosine damping function. Since the MAP model is required to estimate the covariance but the covariance is itself required to estimate the MAP model (in the log-likelihood function), the above procedure must be applied iteratively.

The validity of the covariance estimate can be examined *a posteriori* by considering standardized residuals

$$\mathbf{w}_f(\hat{\mathbf{m}}) = [\hat{\mathbf{C}}_f^{-1/2}]^T \mathbf{r}_f(\hat{\mathbf{m}}), \quad (3)$$

where  $\hat{\mathbf{C}}_f^{-1/2}$  represents the inverse of the Cholesky decomposition (square root) of the covariance matrix. If the covariance estimate is valid,  $\mathbf{w}_f$  should represent an uncorrelated random process: this can be examined qualitatively by plotting the autocorrelation of  $\mathbf{w}_f$ , with a

narrow central peak indicating uncorrelated residuals. Quantitative statistical tests can also be applied to the standardized residuals (e.g., one-tailed runs test).

Finally, matched-field methods are typically employed without knowledge of the (complex) source spectrum. In this case, a maximum-likelihood estimate for the source strength can be derived [1, 2], leading to the substitution

$$\mathbf{d}_f(\mathbf{m}) \rightarrow \frac{\mathbf{d}_f(\mathbf{m})^T \mathbf{d}_f}{|\mathbf{d}_f(\mathbf{m})|^2} \mathbf{d}_f(\mathbf{m}) \quad (4)$$

in the above equations.

### 3. INVERSION EXAMPLE

Bayesian geoacoustic inversion with full-covariance estimation is illustrated here for ocean acoustic data measured in the Mediterranean Sea off the west coast of Italy near Elba Island [3]. Linear frequency-modulated signals (300–800 Hz) were transmitted from a ship-towed source at approximately 10-m depth and recorded at a 48-element vertical line array that extended from 26–120 m depth in water 132-m deep. The inversion here is for a source transmission at a range of approximately 4 km. The seabed was parameterized as a two-layer model with an upper sediment layer of thickness  $h$  and sound speed  $c_1$ , density  $\rho_1$  and attenuation  $\alpha_1$  overlaying a semi-infinite basement with parameters  $c_2, \rho_2$  and  $\alpha_2$ . Small corrections to the source range and depth and water depth were also included in the inversion, but are not discussed here.

Fig. 1 shows examples of the (complex) covariance matrices estimated from the acoustic data, and indicates that the spatial correlation scale decreases with frequency. Fig. 2 shows marginal PPDs for the geoacoustic parameters computed using both full covariance-matrix estimates and variance-only estimates. The uncertainty distributions based on variance-only estimates are overly optimistic for all geoacoustic parameters compared to the full-covariance

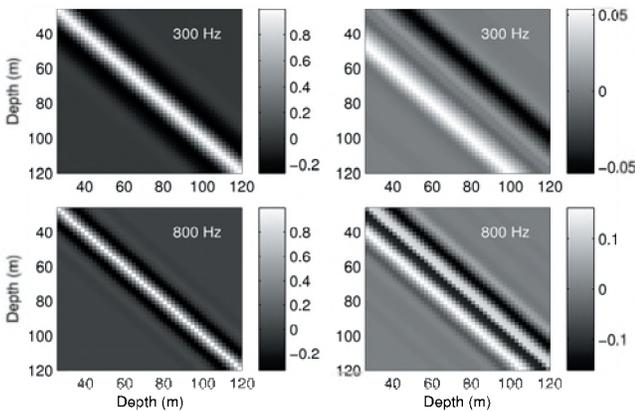


Fig. 1. Covariance matrices at 300 and 800 Hz (real part on left, imaginary part on right), normalized by amplitude of real part.

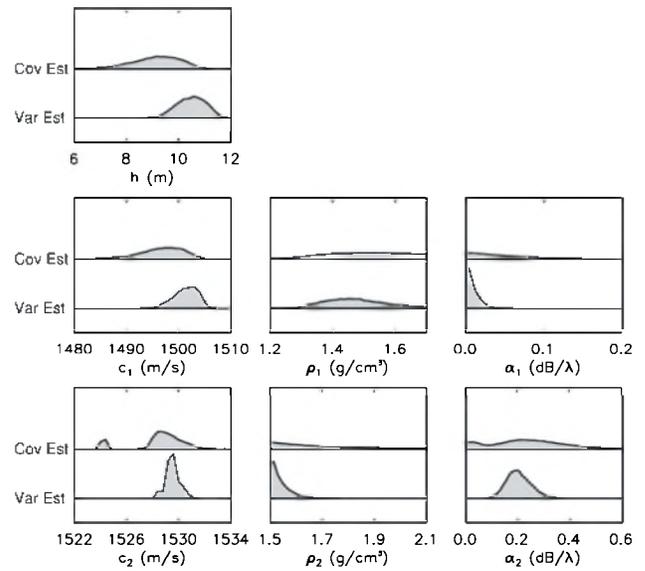


Fig. 2. Marginal probability distributions for geoacoustic parameters computed using full covariance matrix estimates (Cov Est) and variance-only estimates (Var Est).

estimates, in some cases indicating unrealistic parameter sensitivity (e.g.,  $\alpha_1, \alpha_2, \rho_2$ ). Finally, Fig. 3 shows that incorporating full covariance matrices in the inversion leads to essentially uncorrelated standardized residuals, indicating that the error correlations have been accounted for correctly in the inversion.

### REFERENCES

- [1] S.E. Dosso (2002). Quantifying uncertainty in geoacoustic inversion, I: A fast Gibbs sampler approach. *J. Acoust. Soc. Am.*, **111**, 129–142.
- [2] S.E. Dosso and P.L. Nielsen (2002). Quantifying uncertainty in geoacoustic inversion, II: Application to broadband shallow-water data. *J. Acoust. Soc. Am.*, **111**, 143–159.
- [3] S.E. Dosso, M.J. Wilmut and A.-L.S. Lapinski (2001). An adaptive hybrid algorithm for geoacoustic inversion. *IEEE J. Oceanic Eng.*, **26**, 324–336.

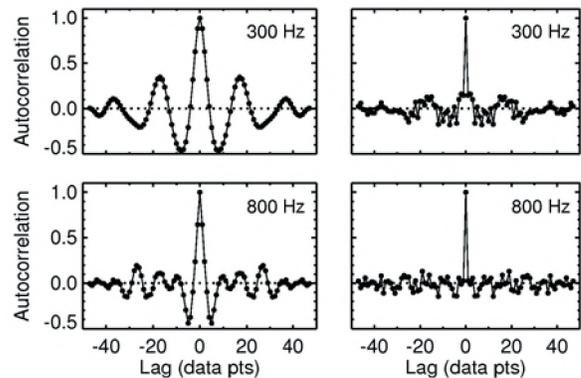


Fig. 3. Autocorrelation of standardized residuals (real part) for variance only estimates (left) and full covariance estimates (right).

# ARRAY ELEMENT LOCALIZATION ACCURACY AND SURVEY DESIGN

Stan E. Dosso<sup>1</sup> and Gordon R. Ebbeson<sup>2</sup>

<sup>1</sup>School of Earth and Ocean Sciences, University of Victoria, Victoria BC, Canada, V8W 3P6, sdosso@uvic.ca

<sup>2</sup>Defence Research and Development Canada–Atlantic, Dartmouth NS, Canada, B2Y 3Z7

## 1. INTRODUCTION

Advanced array processing methods in underwater acoustics require knowledge of the locations of individual elements in a sensor array. However, sufficiently accurate sensor locations are often not known after array deployment, and array element localization (AEL) surveys are required. AEL is based on inverting acoustic arrival-time measurements from a series of controlled sources to the sensors to be localized [1, 2]. AEL is usually based on direct acoustic paths, but can also include surface and/or seabed reflected paths to provide more information. Synchronized AEL surveys, in which the source transmission times are known, are more complicated logistically than non-synchronized surveys, but provide more informative data. AEL accuracy also depends on a number of other factors including arrival time uncertainty, number of sources, source configuration, source-position uncertainties, and water-column sound-speed and depth uncertainties.

This paper quantifies sensor-localization accuracy in terms of an analytic result for the posterior uncertainties of a Bayesian formulation of AEL inversion which takes all of the above factors into account. This provides a rigorous and general measure of the accuracy that can be achieved in AEL applications without resorting to computationally-intensive Monte Carlo simulations. The approach can be applied to study the relative importance of the various factors influencing AEL accuracy, and to guide in planning efficient and effective AEL surveys.

## 2. THEORY

### 2.1 Inverse Theory

Let  $\mathbf{d}$  and  $\mathbf{m}$  be vectors of measured data and unknown model parameters, respectively, with the elements of each considered to be random variables. According to Bayes' rule, the posterior probability density (PPD)  $P(\mathbf{m}|\mathbf{d})$  is then proportional to the product of the likelihood function (quantifying data information) and the prior probability distribution (expressing independent model information). Assuming that the errors (uncertainties) for the measured data and prior parameter estimates  $\hat{\mathbf{m}}$  are Gaussian-distributed random variables with covariance matrices  $\mathbf{C}_d$  and  $\mathbf{C}_m$  (typically diagonal matrices with variances on the main diagonal), the PPD may be expressed

$$P(\mathbf{m}|\mathbf{d}) \propto \exp\left\{-\frac{1}{2}[(\mathbf{d} - \mathbf{d}(\mathbf{m}))^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{d}(\mathbf{m})) + (\hat{\mathbf{m}} - \mathbf{m})^T \mathbf{C}_m^{-1} (\hat{\mathbf{m}} - \mathbf{m})]\right\}, \quad (1)$$

where  $\mathbf{d}(\mathbf{m})$  represents data predicted for model  $\mathbf{m}$ . For a linear problem with  $\mathbf{d}(\mathbf{m}) = \mathbf{A}\mathbf{m}$ , the maximum *a posteriori* (MAP) solution (i.e., the most probable parameter set) is found by setting the PPD derivative to zero, leading to

$$\mathbf{m} = [\mathbf{A}^T \mathbf{C}_d^{-1} \mathbf{A} + \mathbf{C}_m^{-1}]^{-1} \mathbf{A}^T \mathbf{C}_d^{-1} [\mathbf{d} - \mathbf{A}\hat{\mathbf{m}}]. \quad (2)$$

Further, the PPD is a multi-dimensional Gaussian distribution with expected values given by the MAP parameters and posterior covariance matrix

$$\mathbf{C}_m = [\mathbf{A}^T \mathbf{C}_d^{-1} \mathbf{A} + \mathbf{C}_m^{-1}]^{-1}. \quad (3)$$

In particular, the standard deviation for parameter  $m_i$  is given by the square-root of the  $i$ th diagonal element of  $\mathbf{C}_m$ .

### 2.2 AEL Inversion

A general formulation of AEL inversion includes as unknown parameters not only the positions ( $x, y, z$ ) of the sensors to be localized, but also source locations and water-column parameters to properly account for the effect of uncertainties in these quantities. Source transmission instants can be treated as either unknown, known, or known to within a common timing offset to account for system synchronization error. Treating the above quantities as unknown parameters leads to an under-determined inverse problem. However, incorporating prior estimates with uncertainties, as outlined above, regularizes the inversion and provides a stable, well-determined solution and quantitative posterior uncertainty estimates. AEL is based on inverting the acoustic ray-tracing equations

$$t = \int_{\text{ray path}} \frac{dz}{c(z)[1 - p^2 c^2(z)]^{1/2}}, \quad (4)$$

$$r = \int_{\text{ray path}} \frac{pc(z)dz}{[1 - p^2 c^2(z)]^{1/2}}. \quad (5)$$

In Eq. (5), the arrival time data  $t$  are equal to the transmission time  $t_0$  plus the travel time along the ray path through sound-speed profile  $c(z)$  (possibly including surface

and/or bottom reflections). Ray parameters  $p = \cos(\theta)/c(z)$  ( $\theta$  is grazing angle) for eigenrays connecting source and receiver are determined by searching for values which produce the correct range  $r$  via Eq. (6). An efficient search uses Newton's method to refine an initial approximation based on straight-line propagation, with boundary reflections incorporated using the method of images [1].

The ray equations are functionally nonlinear, but can be linearized using a truncated Taylor-series expansion about an arbitrary starting model  $\mathbf{m}_0$ , leading to a MAP estimate

$$\mathbf{m} = [\mathbf{J}^T \mathbf{C}_d^{-1} \mathbf{J} + \mathbf{C}_m^{-1}]^{-1} \mathbf{J}^T \mathbf{C}_d^{-1} [\mathbf{t} - \mathbf{t}(\mathbf{m}_0) + \mathbf{J}(\mathbf{m}_0 - \hat{\mathbf{m}})], \quad (6)$$

where  $\mathbf{J}$  is the Jacobian matrix of partial derivatives with elements  $J_{ik} = \partial t_i(\mathbf{m}_0) / \partial m_k$ . Due to the linearization, the inversion must be repeated iteratively to convergence [1, 2]. Parameter uncertainties are given by Eq. (4) with  $\mathbf{J}$  (evaluated at the final model) substituted for  $\mathbf{A}$  and  $\mathbf{t}$  for  $\mathbf{d}$ . The required integrals and derivatives can be derived analytically for a piecewise-linear sound-speed profile [1].

### 3. AEL EXAMPLE

As a synthetic AEL example, Fig. 1(a) shows a plan view of a 30-element horizontal sensor array (at 75-m depth) together with 8 acoustic sources (30-m depth). Fig. 1(b) shows the sound-speed profile, and Fig. 1(c) shows ray paths with up to one bottom reflection for the maximum range of 800 m. In this example, arrival time uncertainties are 0.5 ms and prior information consists of source locations known to within 10 m in  $x$  and  $y$  and 5 m in  $z$ , and sensor locations known to within 50 m in  $x$  and  $y$  and 10 m in  $z$ .

Fig. 2 shows that posterior sensor-location uncertainties for direct-path AEL inversion are smaller by up to  $\sim 1.5$  m for known source-transmission timing over unknown timing. Results for timing known to within a constant offset are similar to known timing. The unknown timing results degrade (relative to known timing) for cases involving fewer sources or inferior source geometries (not shown).

Finally, Fig. 3 shows mean sensor-location uncertainties computed using different acoustic arrivals (d—direct, s—surface-reflected, b—bottom-reflected, all—all paths in Fig. 1c) and various water-depth uncertainties. Sensor depth is improved most by including reflected arrivals (since these have different vertical angles but follow the same  $x$ - $y$  paths as direct arrivals); however, the benefits of bottom-reflected paths diminish with increasing water-depth uncertainty.

### REFERENCES

- [1] S.E. Dosso & N.E. Collison (2001). Regularized inversion for towed-array shape estimation. *Inverse Problems in Underwater Acoustics*, Eds: M.I. Taroudakis & G.N. Makrakis (Springer-Verlag, New York), 77–103.
- [2] S.E. Dosso, N.E. Collison, G.J. Heard & R.I. Verrall (2004). Experimental validation of regularized array element localization. *J. Acoust. Soc. Am.*, **115**, 2129–2137.

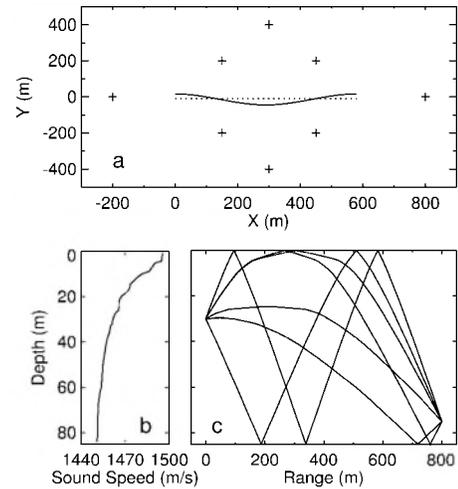


Fig. 1. AEL example: (a) plan view of array (solid line) compared to straight (dotted) line and source positions (crosses); (b) sound-speed profile; and (c) ray paths with up to one bottom reflection.

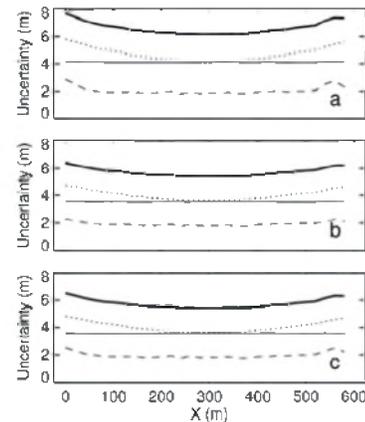


Fig. 2. Posterior sensor-location uncertainties in  $x$ ,  $y$ ,  $z$  and  $R = [x^2 + y^2 + z^2]^{1/2}$  (solid, dotted, dashed and heavy solid lines, respectively) when transmission times are (a) unknown, (b) known, and (c) known to within a common offset.

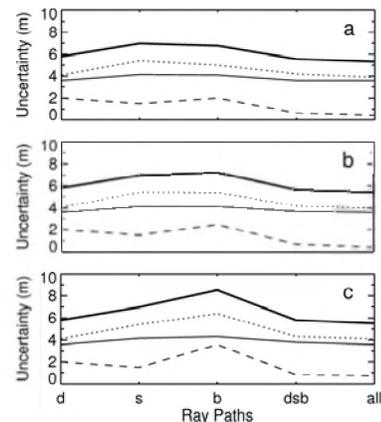


Fig. 3. Mean sensor-location uncertainties in  $x$ ,  $y$ ,  $z$  and  $R$  (solid, dotted, dashed and heavy solid lines, respectively) for indicated ray paths when the water-depth uncertainty is 1, 2 and 5 m in (a), (b) and (c), respectively.

# MONITORING FLUIDIZED-BED DRYING OF PHARMACEUTICAL GRANULES USING ACOUSTIC SENSORS

Derek Daniher, Lauren Briens

Western Fluidization Group, Faculty of Engineering, The University of Western Ontario,  
London, Ontario, Canada, N6A 5B9 [lbriens@eng.uwo.ca](mailto:lbriens@eng.uwo.ca)

## 1. INTRODUCTION

In the manufacturing of pharmaceutical tablets, one of the steps is the granulation of the excipient, binder and active drug. Often, a two-step process is required to accomplish this. First, agglomeration is performed in a high-shear mixer. Second, the wet granules are dried. The result is a dry mixture with a larger size distribution and better flowability than the starting materials. Fluidized-bed drying is often the drying method of choice due to its rapid mixing and uniform temperature distribution. The particles are suspended in warm up-flowing air, maximizing the exposed solid surface area. Care must be taken to avoid over-drying, which may cause the granules to become brittle, and cause product loss due to attrition.

Typically, fluidized-bed drying is monitored by removing samples from the bed at certain time intervals and measuring the moisture content of the samples. Sampling is often inaccurate due to the difficulties in collecting representative samples. Also, the moisture analysis is performed off-line, and takes several minutes to complete. Another technique is to monitor the temperature of the bed or of the air at the outlet. While the surfaces of the granules remain moist, the temperature of the granules will not exceed the wet bulb temperature of the drying gas. When the surface of the granule becomes dry, the granule temperature will approach the dry bulb temperature of the gas. This increase in bed temperature is commonly used to detect the drying end-point. Monitoring temperature only provides information about the moisture content of the granules. Often, attrition becomes a problem well before the desired moisture content is reached. This cannot be detected by monitoring temperature alone.

Acoustic monitoring has been shown to be a useful monitoring tool in high-shear wet granulation (Daniher et al, 2005), pneumatic transport (Albion et al., 2005) and rotary drying (Smith, 2004). The application to fluidized-bed drying monitoring is based on the fact that the sound produced by the collisions of dry granules is different from sound produced by the collisions of moist granules. These collisions are dependant on the quality of the fluidization hydrodynamics. The sounds from a well fluidized bed will differ from that of a non-fluidized bed. If these sounds can

be recorded and transformed into a useful voltage signal, it may provide a non-intrusive, passive method of monitoring the drying process. This objective of this research is to show the potential of acoustics for monitoring the drying of pharmaceutical granules.

## 2. METHOD

The fluidized-bed drying apparatus was a conical column with the dimensions given in Fig. 1. The distributor was stainless steel 100  $\mu\text{m}$  mesh. Four air outlets were located on the top of the column. These outlets were fitted with filter cloth to prevent dust from escaping. A port fitted with a sampling thief was located at the base of the stainless steel column. Sound data was obtained using two PCB Piezotronics model 130D10 electret microphones and 130P10 preamplifiers. The positions of the sensors on the granulator equipment are indicated in Fig. 1. Microphone 1 was attached flush to the exterior of column. Microphone 2 was centered in one of the air outlets, on the exterior of the filter cloth. The data from both sensors was acquired using a 16-bit National Instruments DAQCard-6036E. The samples were recorded at a sampling rate of 40,000 Hz.

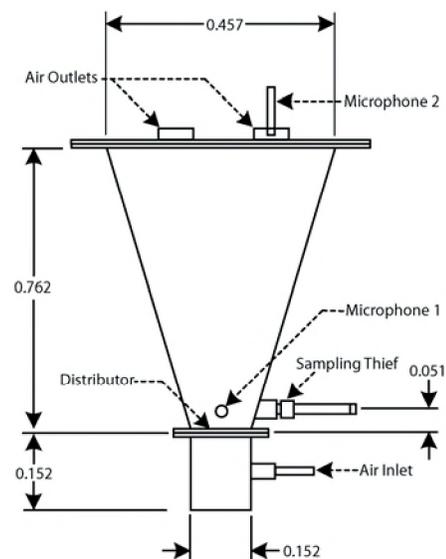


Fig. 1. The fluidized-bed dryer dimensions in meters.

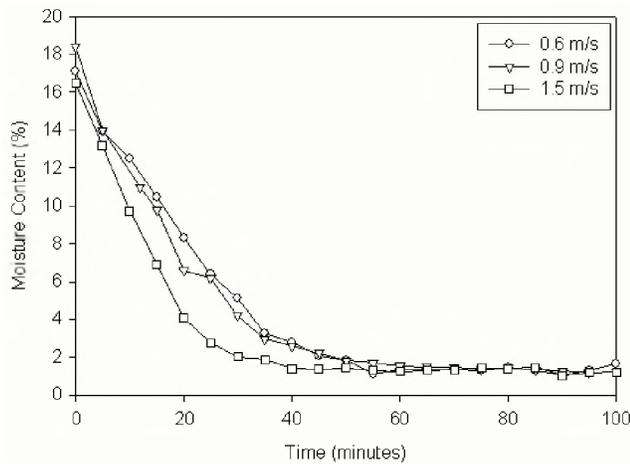


Fig. 2. Moisture content of granule samples taken during the drying process.

Placebo granules consisting of lactose monohydrate (87wt% dry basis), corn starch (10wt%) and polyvinylpyrrolidone (3wt%) were granulated with 18wt% de-ionized water. 2.1 kg of wet granules were dried in the fluidized bed dryer over 100 minutes. The inlet air temperature was 21 °C and the superficial air velocities at the distributor were 0.6, 0.9 and 1.5 m/s. Samples were taken every 5 minutes and analyzed for moisture content using a Mettler Toledo HG63 halogen moisture analyzer. The microphone signals were recorded for the entire drying process and analyzed offline in 10 second segments.

### 3. RESULTS

Fig. 2 is a plot of granule moisture content as a function of time during the drying process. The change in superficial air velocity between 0.6 and 0.9 m/s did not affect the drying profile. There was an increase in the drying rate at 1.5 m/s. The profile of the mean frequency of the microphone 1 signal (Fig. 3) showed a peak occurring at about 40 minutes for air velocities from 0.6 to 0.9 m/s

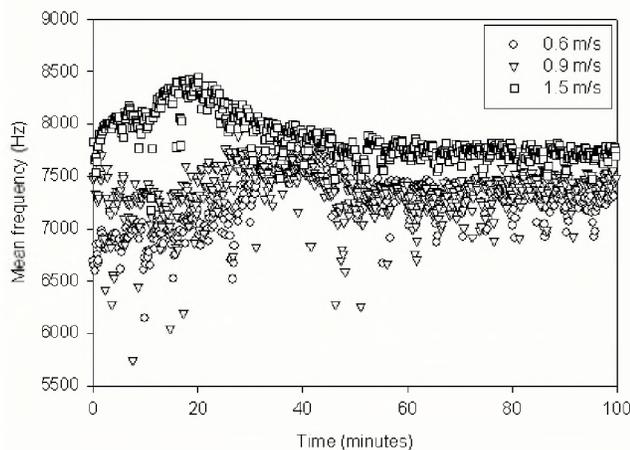


Fig. 3. Mean frequency of the microphone 1 signal (base of cone).

and 20 minutes for 1.5 m/s. These times approximately corresponded to a bed moisture content of 3 to 4%. The mean frequency of the microphone 2 signal (Fig. 4) showed a sharp decrease occurring at approximately the same times.

### 4. CONCLUSIONS

The mean frequency profiles detect the 3 to 4% moisture range rather than the equilibrium moisture content of 1.5%. This indicates that rather than moisture content, the signal may be monitoring a range of other factors such as granule hardness or size distribution. As the granules approach 3 to 4% moisture, attrition may become significant. The resulting increase in fines would build up on the surface of the filter cloth and could cause the sudden drop in mean frequency. Since the detection of the onset of attrition cannot be detected by monitoring temperature, having a method to do this is advantageous. Acoustic monitoring is non-intrusive, in-expensive, and requires no modification to existing fluidized-bed drying equipment if a microphone is located externally.

### REFERENCES

- Albion, K., Briens, L., Briens, C. & Berruti, F. (2005). Flow regime determination in horizontal pneumatic transport of fine powders using non-intrusive acoustic probes. *Powder Technology*, (submitted).
- Daniher, D., Briens, L. & Tallevi, A. (2005). End-point detection in high-shear granulation using sound and vibration signal analysis. *Powder Technology*, (submitted).
- Smith, R.J. (2004). Monitoring of a rotary dryer using acoustic emissions, Master's Thesis. The University of Western Ontario.

### ACKNOWLEDGEMENTS

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarship Program (OGS) for their financial support.

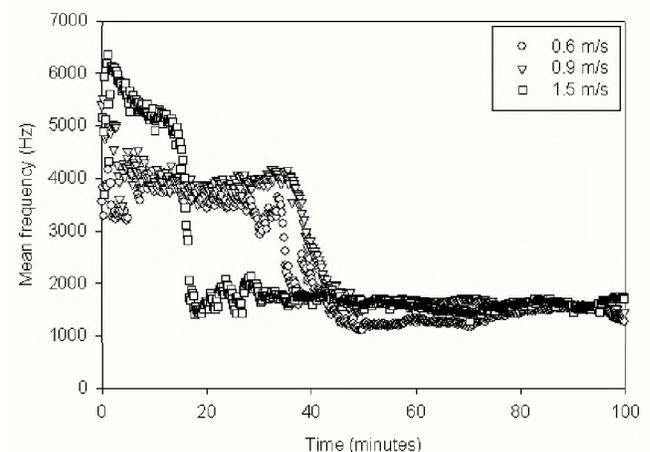


Fig. 4. Mean frequency of the microphone 2 signal (air outlet).

# DEVELOPMENT OF AN ACOUSTIC METHOD FOR THE MEASUREMENT OF MIXING AND DRYING IN A VIBRATED FLUIDIZED BED

Garret Book, Lauren Briens, Cedric Briens

Western Fluidization Group, Faculty of Engineering  
The University of Western Ontario, London, Ontario, Canada, N6A 5B9

## 1. INTRODUCTION

Gas-solid fluidized beds can be used for physical processes such as mixing, drying and granulation. Fine powders, however, are difficult to fluidize well with gas flow. Vibration can be used to improve the fluidization.

The objective of this research was to develop a non-invasive acoustic method to detect fluidization regimes and measure mixing and drying in a vibrated fluidized bed.

### 1.1 Fluidization

At low gas velocities, the bed of powder is fixed. As the gas velocity is increased above the minimum fluidization velocity ( $U_{mf}$ ), the bed becomes fluidized. Increasing the gas velocity further, past the minimum bubbling velocity ( $U_{mb}$ ) the bed becomes well mixed due to the action of the bubbles. Applied vibration reduces the required gas velocity for fluidization and bubbling. Pressure profiles and bed height methods for determining  $U_{mf}$  and  $U_{mb}$  are not practical for industrial applications.

The fluidizing gas dries the wet particles as it flows through the bed. For drying, the bed must be operated above  $U_{mb}$  to provide the agitation, mixing and good heat transfer required for effective drying. Monitoring is required to ensure that the solids reach the appropriate moisture.

### 1.2 Acoustics

Acoustic sensors are inexpensive, can withstand a wide range of process conditions, and can provide reliable, fast, on-line and non-intrusive monitoring. Passive acoustics detect the sound generated by the process itself. In vibrated fluidized beds, acoustic emissions are caused by particles colliding with each other, the bed walls, the motion and eruption of bubbles and the vibration source [1].

### 1.3 Signal Analysis

The standard deviation measures the dispersion of the values within a given signal. The information entropy of a signal is a measure of the randomness. The power spectrum decomposes a signal into frequency bands that allow dominant frequencies to be identified.

## 2. METHOD

A Plexiglas fluidized bed (I.D. of 0.113 m and height of 0.274 m) was mounted on a vibration source and vibrated at 60 Hz with variable amplitudes. The powder was ceramic microspheres of 33  $\mu\text{m}$ . This powder was non-porous and analogous to glass beads. A microphone was attached in the baghouse at the top of the column directed down towards the bed. Each acoustic measurement was recorded at a frequency of 40000 Hz.

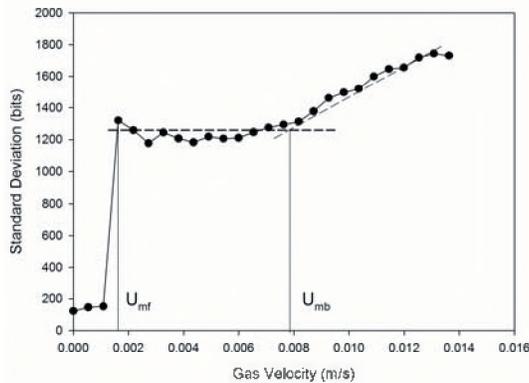
To determine the minimum bubbling velocity and the minimum fluidization velocity, the bed was initially very well fluidized. Acoustic emissions were recorded, the gas velocity was incrementally lowered, the bed was allowed to reach steady state, and measurements repeated. Measurements were recorded at vibration amplitudes of 0, 0.01, 0.026 and 0.05 mm.

To evaluate drying, solids were removed from the bed, water was mixed in with the solids and then the particles were replaced and the bed was re-fluidized. Measurements were recorded at gas velocities of 0.008 and 0.010 m/s and vibration amplitudes of 0 and 0.05 mm.

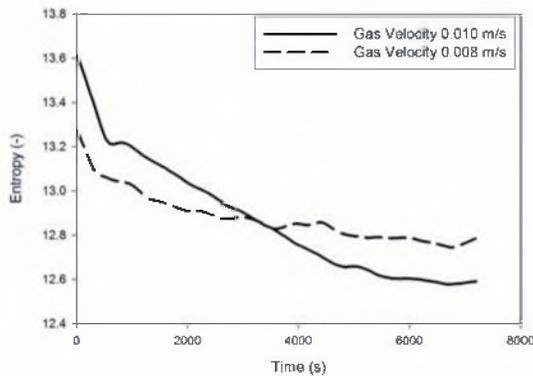
## 3. RESULTS & DISCUSSION

As shown in Figure 1, the standard deviation of the acoustic emissions for the vibrated bed identified both the minimum fluidization and the minimum bubbling velocities. The standard deviation indicates the amplitude of the acoustic emissions. Below  $U_{mf}$ , the bed was fixed: the acoustic emissions were low as there was no particle movement and the bed did not easily transmit emissions from the vibration source. The standard deviation of the acoustic emissions increased significantly above  $U_{mf}$  due to particle movement and the fluid properties of the bed allowing better transmission of emissions from the vibration source. Above  $U_{mb}$ , the standard deviation again increased due to the emissions from bubble motion and eruption at the bed surface and the decreased bed density for emission transmission. The power at the 60 Hz frequency of the vibration source showed similar profiles and the standard deviation was much lower in the non-vibrated bed indicating that observed variations in acoustic emissions are

primarily due to the changes in transmission properties of the bed as it changes fluidization regime.



**Fig. 1.** Standard deviation of acoustic emissions at a vibration amplitude of 0.026 mm.



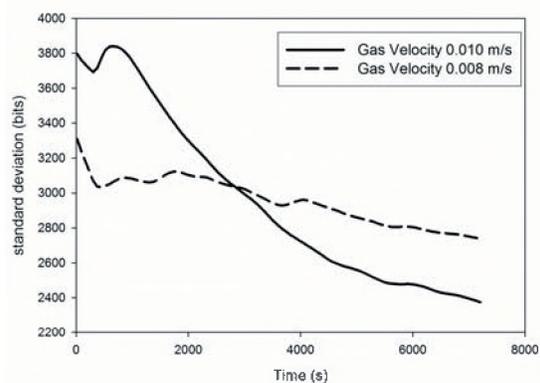
**Fig. 2.** Evolution of the information entropy of the acoustic emissions as the bed dries.

Figure 2 shows that the drying process can be monitored with the acoustic signals. It shows how the information entropy of the microphone signal decreased gradually as drying progressed. At a superficial gas velocity of 0.010 m/s, the information entropy reached an asymptotic value at a time of about 6000 s (100 minutes). Drying at a superficial gas velocity of 0.08 m/s was not as effective, since the information entropy had not yet reached its asymptotic value after 7000 s of drying.

Figure 3 shows that similar results can be obtained from the standard deviation of the acoustic signals. The information entropy, however, seems more sensitive to low moistures.

Since the initial moisture was 100 ppm, the non-invasive acoustic signals therefore allow for the detection of

moistures of a few tens of ppms. Very few on-line methods can achieve such sensitivity.



**Fig. 3.** Evolution of the standard deviation of the acoustic emissions as the bed dries.

## 4. CONCLUSIONS

A non-invasive acoustic method was developed to identify fluidization regimes and to monitor drying within a vibrated fluidized bed.

The minimum fluidization and bubbling velocities can be reliably determined from the standard deviation of a microphone signal.

The drying of fluidized solids can be monitored by using either the information entropy or the standard deviation of a microphone signal. This acoustic method can detect moistures of a few tens of ppms.

## REFERENCES

- [1] Belchamber, R. (2000) Acoustic Emission Monitoring – A New Tool for Particle Technology. *Pharmaceutical Technology Europe*, **12**, 26-29.

## ACKNOWLEDGEMENTS

The authors would like to thank the National Sciences and Research Council of Canada for their financial support of this research.

# FLOW REGIME DETECTION IN PNEUMATIC TRANSPORT OF PARTICULATES USING NON-INTRUSIVE ACOUSTIC PROBES

Katherine Albion, Lauren Briens, Garret Book, Franco Berruti & Cedric Briens

Western Fluidization Group, Faculty of Engineering  
The University of Western Ontario, London, Ontario, N6A 5B9

## 1. INTRODUCTION

Pneumatic transport is used to convey solids in many industrial plants. Changes in particle properties or operating conditions may cause particles to settle out and deposit on the bottom of horizontal sections of the transport line. Such deposits result in unsteady solids flow, pressure surges and eventually, blockage of the line. Early detection of deposits would allow for quick corrective action.

The objective of this paper is to present an early detection method for solids deposits at the bottom of the transport line using only non-invasive sensors. External microphones were selected as non-invasive sensors and their signals were processed with advanced filters and signal analysis methods.

### 1.1 Pneumatic Transport and Flow Regimes

In dilute phase horizontal pneumatic transport, particles do not move forward in a straight line, parallel with the pipe wall. Gravity pulls the particles toward the bottom of the pipe, but are ideally kept in suspension by turbulent gas eddies. As the gas flowrate is reduced or the solids flowrate increased, particles eventually deposit on the bottom of the pipe [1].

Flow patterns in horizontal transport vary depending on the solids concentration and superficial gas velocity. Different flow regimes occur at different transport conditions. At high superficial gas velocities and low solids fluxes, there is fully suspended dilute phase flow; increasing the solids flux or decreasing the superficial gas velocity results in dilute phase flow with a higher concentration of solids transported along the bottom of the pipe, and later as solids deposits as dunes or a stationary bed [2].

Flow patterns in an inclined line are similar. At high gas velocities and low solids fluxes, there is dilute phase flow. Increasing the solids flux or decreasing the superficial gas velocity eventually results in solids deposition on the bottom of the pipe [3].

### 1.2 Acoustics

Acoustic sensors are inexpensive and can withstand a wide range of process conditions. They can provide reliable, on-line and non-intrusive monitoring.

Passive acoustics detect the acoustic emissions generated by the process. In processes involving the movement of solid particles, acoustic emissions are caused by particles colliding with each other, vessel walls or other objects [4].

### 1.3 Signal Analysis – Wavelet Residual and V Statistic

The wavelet residual is a method used for detecting outliers in the signal [5]. Wavelets can be used to denoise signals while preserving sharp, rapid variations in the signal. However, in the present work the rejected “noise” actually contains valuable information about the process. The V Statistic is used to detect cyclic, non-periodic behaviour.

## 2. METHOD

The pneumatic transport loop consisted of a 0.1 m inside diameter, stainless steel pipe. The gas-solids mixture flows through a 5.3 m vertical transport line, a 5.5 m horizontal pipeline or 5 m inclined pipeline, with a return line into a cyclone. Each acoustic measurement was recorded at a frequency of 40 000 Hz. Acoustic sensors were located on the bottom or side of the pipe, at measurement locations of 0.05, 0.20, 0.35, 0.90, 1.05, 1.20, 1.35, 1.50, and 1.65 m from the elbow in the horizontal line for glass beads or PVC powder, or at 0.50, 1.05, 2.50 and 3.3 m from the elbow in the inclined line for polyethylene pellets. An acrylic section of pipe in each section allowed for visual observations of the flow regimes. The flow regimes could easily be determined from high speed video files.

## 3. RESULTS & DISCUSSION

The V Statistic at 0.000425 s was calculated from the wavelet residual signal. At this subperiod length, there was a significant difference in the V Statistic which allowed for regime identification. Three flow regimes were identified: conveying over settled solids, dilute phase conveying and a transition region between the phases. Figure 1 shows that the narrow transition region could be detected from acoustic signals. This transition region occurred at  $164 \pm 4 \text{ s}^{-0.5}$ .

This method was applied to PVC powder in the same system, and resulted in a similar transition region of  $169.5 \pm 5.5 \text{ s}^{-0.5}$ .

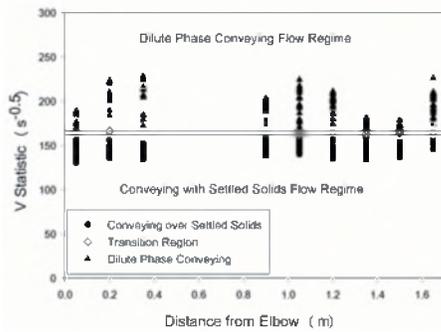


Fig. 1. V Statistic values for glass beads using wavelet residual of acoustic signals for all distances from the elbow, identifying two main flow regimes and narrow transition region.

Figure 2 is a flow regime map that was developed from the acoustic signals using the criteria shown in Figure 1. Each curve on the map represents the average value of the transition region, as it occurs relative to all solids fluxes and superficial gas velocities, and corresponds to a given distance from the elbow. The area located above each line represents the conveying with settled solids flow regime, whereas the area located below each curve represents dilute phase conveying. At 0.05 m from the elbow, the settled solids regime predominated. At distances further from the

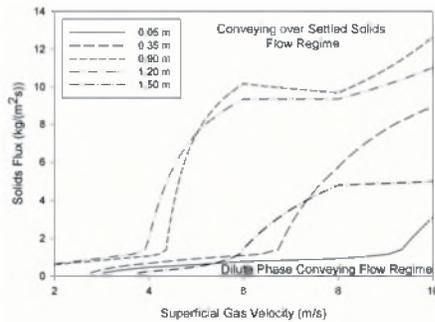


Fig. 2. Flow regime map for glass beads at various distances from the elbow.

elbow, 0.35 m and 0.90 m, the position of the line changed, and there were more conditions where dilute phase flow occurred. The majority of dilute phase flow occurred at 0.90 m and 1.20 m, where the shapes of the curves were nearly identical. Further from the elbow, at 1.50 m, there was less dilute phase flow, and the shape of the curve was similar to the curves closer to the elbow. This variation in the flows indicated where dunes existed rather than locations of a stationary bed. At 0.90 m and 1.20 m, there were breaks in the stationary bed, forming distinct dunes, existing at these locations and conditions, whereas at locations further from the elbow, a settled bed was detected.

This flow regime method was applied to an inclined line transporting polyethylene pellets. Figure 3 shows the V Statistic values for the wavelet residual acoustic signal at 0.000425 s. Two regimes were identified, conveying over settled solids and dilute phase flow. There was no detectable transition region: minute changes in either gas

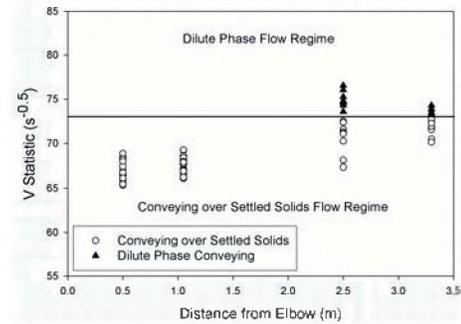


Fig. 3. V Statistic values for polyethylene pellets from the wavelet residual acoustic signal for all distances from the elbow, identifying the two flow regimes.

velocity or solids flowrate were sufficient to switch from one regime to another. The boundary between the two regimes occurred at  $73 \text{ s}^{-0.5}$ .

## 4. CONCLUSIONS

Flow regimes in pneumatic transport can be detected by applying a wavelet filter to the raw acoustic signal and calculating the V Statistic at 0.000425 s on the residual signal.

There was a narrow transition region between the dilute phase regime and conveying with settled solids regime for horizontal transport of glass beads and PVC powders. There was a sharp boundary between the two flow regimes for inclined transport of polyethylene pellets.

This acoustic monitoring method is useful for process control. It allows for easy, rapid and non-intrusive on-line monitoring of flow regimes in pneumatic transport lines. It can help maintain the pneumatic transport line at conditions that maximize product quality and system efficiency.

## REFERENCES

- [1] L.S. Fan, C. Zhu, Principles of Gas-Solids Flows: Cambridge Series in Chemical Engineering, Cambridge University Press, United Kingdom, 1998.
- [2] G.E. Klinzing, R. Marcus, F. Rizk, L.S. Leung, Pneumatic Conveying of Solids, 2nd Edition, Chapman and Hall, New York, 1997.
- [3] J.G.E. Klinzing, N.D. Rohatgi, C.A. Myler, S. Dhodapkar, A. Zaltash, M.P. Mathur, Pneumatic transport of solids in an inclined geometry, Can. J. Chem. Eng. 67 (1989) 237 - 244.
- [4] J.W.R. Boyd, J. Varley, The uses of passive measurement of acoustic emissions from chemical engineering processes, Chem. Eng. Sci. 56 (2001) 1749 - 1767.
- [5] J. Trygg, N. Kettaneh-Wold, L. Wallback, 2D wavelet analysis and compression of on-line industrial process data, J. Chemometrics 15, (2001) 299 - 319.

## ACKNOWLEDGEMENTS

The authors would like to thank the National Sciences and Engineering Research Council of Canada for their financial support of this research.

# DETECTION OF CONDITIONS LEADING TO BREAKAGE OF PHARMACEUTICAL TABLETS IN PNEUMATIC TRANSPORT

Katherine Albion, Lauren Briens\*, Cedric Briens, and Franco Berruti

Department of Chemical and Biochemical Engineering, Faculty of Engineering  
The University of Western Ontario, London, Ontario, Canada, N6A 5B9, lbriens@eng.uwo.ca

## 1. INTRODUCTION

Pharmaceutical tablets are a popular form of drug delivery due to their convenient and safe means of drug administration, ease of handling, and mass production using quality-controlled procedures for consistent high quality. In pneumatic conveying, the size distribution and appearance of tablets can change significantly and the product may no longer meet the required product specifications. Tablets must meet stringent specifications and broken tablets cannot be passed on to the consumer.

The objective of this research was to develop a detection method for particle breakage in a pneumatic transport line. External microphones were selected as the non-invasive sensors and their signals were analyzed with advanced methods.

### 1.1 Pneumatic Transport and Particle Breakage

Parameters which influence particle breakage and chipping in pneumatic conveying include air velocity, solids loading ratio, bend configuration and particle properties [1]. Particles may undergo different attrition mechanisms, such as erosion, which produces dust with a slight change in the original particle size, chipping, or breakage of the particle into two or more particles of nearly identical sizes [2]. Literature suggests that there exists a threshold velocity, below which particle breakage does not occur [3]. Therefore, to avoid particle breakage it is essential that the velocity remain below this critical value.

### 1.2 Acoustics

Acoustic sensors are inexpensive and can withstand a wide range of process conditions. They provide reliable, on-line and non-intrusive monitoring. Passive acoustics detect the acoustic emissions generated by the process itself. In processes involving the movement of solid particles, acoustic emissions are caused by particles colliding with each other, vessel walls or other objects [4].

### 1.3 Signal Analysis – Kurtosis and Multiple Regression

Kurtosis is used to describe a distribution, and is a measure of the relative peakedness of the distribution. It

can thus be used to detect peaks in a signal. Multiple regression is used to determine the relationship between independent variables and a dependent variable, and identifies the main contributing variables that predict this relationship based on a linear or power law fit.

## 2. METHOD

Two different pneumatic transport systems were examined in this study. For most of the experiments, the pneumatic transport loop consisted of a 0.1 m inside diameter, PVC re-enforced hose and, for some experiments, this hose was replaced with a 0.1 m inside diameter steel pipe. The gas-tablet mixture flows through a 5.0 m vertical upward transport line, a 4.0 m horizontal pipeline and a downward inclined line to a padded collection container. Three different elbow configurations were used: long radius PVC hose, 90° steel elbow and 90° steel elbow lined with a thin foam pad. Three superficial gas velocities were examined: 15, 17 and 19 m/s as well as four tablet feedrates of 15, 60, 105 and 185 tablets/s. Acoustic sensors were located on the side of the pipe, at measurement locations of 2.30 and 3.20 m in the vertical section from the first elbow, on the elbow at the top of the vertical section and at 1.80 and 2.20 m in the horizontal section from the elbow. Each acoustic measurement was recorded at a frequency of 40 000 Hz. A notch filter was used to filter out 60 Hz electrical noise before analysis.

## 3. RESULTS & DISCUSSION

Figure 1 shows peaks existed in the raw signal, which corresponded to collisions between tablets and between tablets and the pipe wall. Kurtosis of the signal was calculated; the amplitude of peaks in the raw signal corresponded to the magnitude of the peaks of kurtosis. The number of peaks in the signal was calculated based on the number of peaks above a set threshold value. This indicated the number of collisions occurring during the transport through the system.

In industrial applications, a high breakage rate of tablets is unacceptable. In this study, an acceptable proportion of broken tablets was chosen to be 0.3%. A breakage index of 1 was assigned to trials with acceptable proportions of

broken tablets and an index of 2 for trials with unacceptable proportions. The proportion of broken tablets was determined by collecting all the conveyed tablets and sorting them manually.

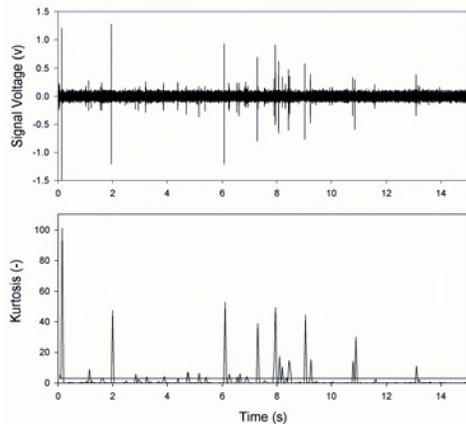


Fig. 1. Raw acoustic signal and kurtosis of signal at 2.2 m in horizontal line

Multiple regression was used to correlate the observed breakage index to the collision rate determined from the microphone signals. Figure 2 shows that there is a perfect agreement between the observed breakage index and the index obtained with the signals from the elbow and the location 1.8 m downstream of the elbow in the horizontal line. Microphone signals can, therefore, be used to monitor the breakage rate.

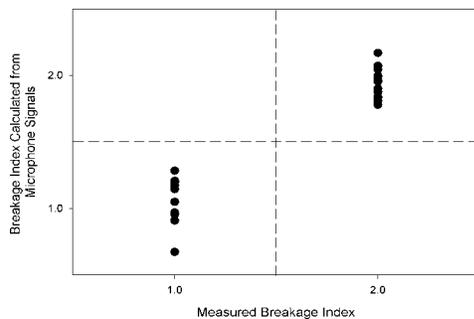


Fig. 2. Using acoustic signals to determine whether the proportion of broken tablets is acceptable (1) or unacceptable (2).

PVC hose and steel pipe were investigated to determine the effect of pipe material on tablet breakage. It was determined that the breakage rate in the steel pipe was less than in the PVC hose at high superficial gas velocities and tablet feedrates. Further investigations into elbow material and shape led to the conclusion that the higher breakage rate in the PVC hose was not due to its wall material but to the long radius of the elbow which caused the particles to significantly decelerate and affected the hydrodynamics of the horizontal section.

A boundary map which defines acceptable operating conditions was developed. Figure 3 shows that a maximum

acceptable tablet feedrate of 220 tablet/s could be achieved with a conveying gas velocity of approximately 15.25 m/s in the steel pipe. With the PVC hose, the maximum acceptable tablet feedrate was lower at 160 tablet/s at a superficial gas velocity of 14.75 m/s.

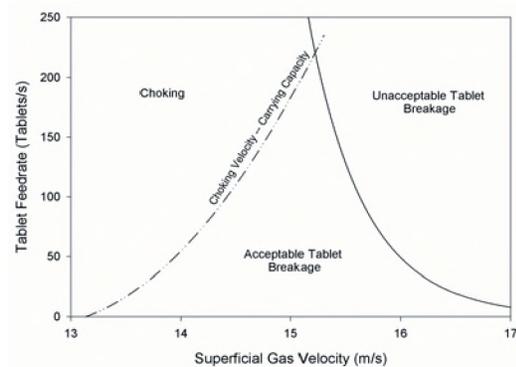


Fig. 3. Boundary conditions of acceptable and unacceptable tablet breakage based on transport conditions in steel pipe.

#### 4. CONCLUSIONS

Unacceptable breakage rates of acetaminophen tablets in pneumatic transport could be detected with non-invasive microphones located at the elbow and horizontal sections of the pneumatic transport line.

A boundary map was developed to determine operating conditions. There was an optimum conveying gas velocity at which the tablet throughput could be maximized with acceptable breakage rates, and where tablet conveying cannot occur due to choking in the vertical section of pipe.

A steel pipe provided much lower breakage rates than a reinforced PVC hose at high superficial gas velocities and tablet feedrates. The gradual elbow associated with the PVC hose induced hydrodynamic conditions in the downstream horizontal section that promoted tablet breakage.

#### REFERENCES

- [1] H. Kalman, Attrition control by pneumatic conveying, *Powder Tech.* 104 (1999) 214-220.
- [2] H. Kalman, Attrition of powders and granules at various bends during pneumatic conveying, *Powder Tech.* 112 (2000) 244-250.
- [3] A.D. Salman, M.J. Hounslow, A. Verba, Particle fragmentation in dilute phase pneumatic conveying, *Powder Tech.* 126 (2002) 109-115.
- [4] J.W.R. Boyd, J. Varley, The uses of passive measurement of acoustic emissions from chemical engineering processes, *Chem. Eng. Sci.* 56 (2001) 1749 - 1767.

#### ACKNOWLEDGEMENTS

The authors would like to thank the National Sciences and Engineering Research Council of Canada for their financial support of this research, and Accucaps Industries Limited for the acetaminophen tablets.

# SINGLE BUBBLE SONOLUMINESCENCE: EFFECTS OF SIGNAL AMPLITUDE MODULATION

Borko Djurkovic, Igor Mastikhin, Dennis Tokaryk

Dept. of Physics, University of New Brunswick, PO Box 4400, Fredericton, NB, Canada, E3B 5A3

## 1. Introduction

### 1.1 Single Bubble Sonoluminescence

When bubbles in a fluid are exposed to high frequency sound pressure, they emit light. This phenomenon is called sonoluminescence. Single bubble sonoluminescence is the study of a single light-emitting bubble.

To generate SBSL, a sinusoidal ultrasound signal is applied to a water-filled flask with a standing wave formation. Bjerknes force drives a small gas bubble (ambient radius of the bubble is around 5  $\mu\text{m}$  [1]) towards a pressure antinode [2]. This force arises due to a pressure gradient across the bubble. Reacting to the pressure variations of the acoustic transducer, the bubble undergoes nonlinear radial oscillations. During the negative phase of the pressure swing, the bubble expands to a maximum radius of about 45  $\mu\text{m}$  [1]. Next, during the positive phase, the bubble compresses, undergoing a violent supersonic collapse with bubble wall velocity in excess of 1.4 km/s [3]. This implosion is so rapid that almost no heat is able to escape the bubble. Pressure of about 1000 atm and temperatures greater than 1000 K [1] are formed inside the bubble, and the result is an emission of light.

### 1.2 Modulated Single Bubble Sonoluminescence

We studied SBSL for the case of driving frequency modulated by lower frequency with an offset. There exists a minimum critical pressure at which stable SL is observed. With modulated SBSL, the drive horn oscillates the acoustic pressure above and below this threshold. Our goal was to introduce the bubble into a dynamic environment, and observe the effects of amplitude modulation on the properties of the emitted light. We used this information to infer how bubble dynamics responds to the modulation.

The driving signal equation becomes:

$$f(t) = \sin(2\pi\nu_0 t)[a + b \sin(2\pi\nu t)] \quad (1)$$

where  $\nu_0$  is the driving frequency,  $\nu$  is the modulation frequency and  $a$  and  $b$  are the lowest and highest pressures over the modulation period. The modulation strength was defined as the difference of highest and lowest pressures over the modulation period:

$$\frac{a-b}{a+b} 100$$

## 2. Experiment

The experimental setup includes a rectangular cell with dimensions 2 1/4" x 2 1/4" x 5". An oscillating voltage is applied across a piezoelectric ceramic transducer to produce ultrasound. The water-filled cell was driven in (1,1,3) mode which corresponds to resonant frequency of about 27 kHz. The drive pressure amplitude of about 1 atm was used. The bubbles are seeded by passing a brief current through a loop of NiCr wire which boils the surrounding fluid.

Modulation frequencies of 1 - 1000 Hz were used. Modulation strengths of 20%, 50%, and 80% were used in the experiments. Glycerol was added to water to produce mixtures of various viscosities. The measurements were performed for pure water and 8%, 16%, and 24% water-glycerol mixtures.

A photomultiplier tube (PMT) was used to collect the emitted light. A Tektronix TDS3000B oscilloscope was used to monitor and record the PMT signal.

To study the bubble motion, we illuminated the bubble using an external light source, and recorded its displacement using a microscope.

## 3. Results and Discussion

The measured SBSL signal appeared as a train of flashes<sup>a</sup> for modulation frequencies below 250 Hz, and as a continuous modulated signal for higher frequencies. From this we can infer that SBSL is possible and stable in the presence of amplitude modulation. For the case of small modulation frequency, the bubble goes through a cycle that repeats itself with modulation period. When the drive pressure is high enough, the bubble emits light. Consequently, as the drive pressure falls, the bubble stops emitting light.

There exists a balance between the Bjerknes, buoyancy, and viscous forces that determines the equilibrium position of the bubble. As the amplitude of the sound field decreases, the equilibrium position will shift, and the bubble will move upward away from the antinode due to buoyancy force. During the next half of the modulation period, the increasing pressure will push the bubble back towards the antinode, and the bubble will begin light emission once

---

<sup>a</sup> Here a flash refers to a fraction of modulation period during which the bubble emits light. This, of course, consists of many hundreds of individual bubble flashes.

again. For high modulation frequencies, the bubble would not move very far away from the antinode since the modulation period is shorter. This suggests that SL intensity is position dependent since for high modulation frequencies, the bubble does not stop emitting light, but rather experiences continuous light emission with intensity oscillating with the modulation frequency.

We looked at the relationship between the flash length to modulation period ratio and the modulation frequency. We found that the ratio increases linearly as modulation frequency increases up to a frequency of about 150 Hz. At higher frequencies, the ratio of flash width to period remains constant. The bubble position responds to modulation frequency which is reflected in the increasing ratio. However, at high modulation frequencies, the changes become too rapid and the bubble is unable to respond to them. Instead, perhaps, a new equilibrium position is reached.

We found that the flash length to modulation period ratio increases as viscosity increases. Increased viscosity results in increased viscous force, which means that the bubble will move slower, and hence less away from the antinode. Increased viscosity also leads to a more symmetric collapse, and slows convection around the bubble which reduces bubble cooling.

We also examined the ratio of the time it takes for the bubble to reach the maximum light intensity, called the rise time, to the time it takes for the intensity to fall back down, called the decay time. The decay time to rise time ratio remains constant as modulation frequency is varied. As it can be seen in Figure 1, the decay time to rise time ratio decreases as modulation strength decreases and as viscosity increases. For normal SBSL, one can expect this ratio to be one. With weaker modulation, the bubble is disturbed less, which means it experiences conditions more similar to normal SBSL. This is indeed the observed behaviour.

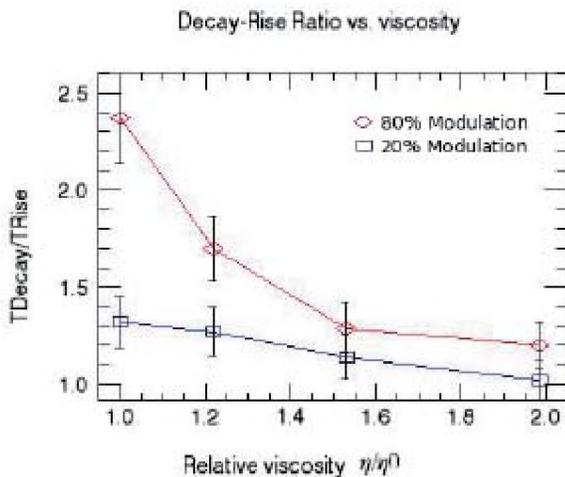


Figure 1: Decay time/Rise time ratio vs. viscosity.

As viscosity increases, bubble collapse becomes more symmetric and bubble position remains closer to the antinode, which means that bubble will again approach normal SBSL conditions.

The bubble motion responds to the pressure swings of the sound field caused by the amplitude modulation. As explained earlier, the bubble moves away from the antinode as the amplitude decreases and towards the antinode as the amplitude increases. Using side illumination and a microscope, we measured the bubble displacement (shown in Figure 2). The displacement decreases as modulation frequency increases and as modulation strength decreases.

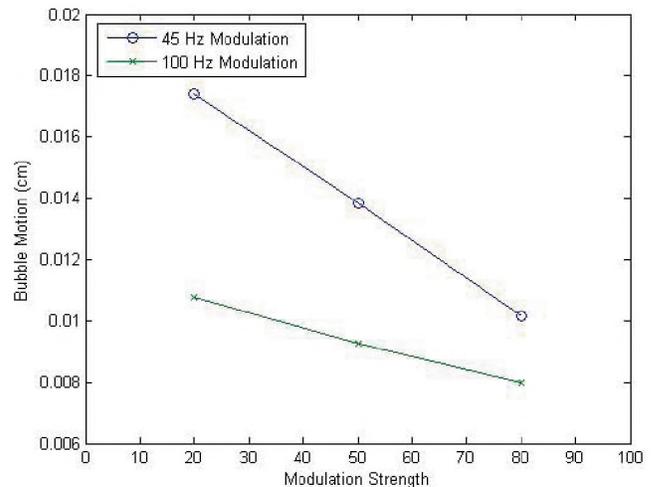


Figure 2: Bubble Motion

Spectroscopy measurements of the modulated SBSL revealed a maximum in the spectra for low modulation frequencies. This is a remarkable result, since spectra of normal SBSL show no maxima, growing steadily towards UV region. As a possible explanation, bubble cooling can take place due to contamination of the interior between flashes and subsequent longer rise times.

## REFERENCES

- [1] B. P. Barber, R. A. Hiller, R. Lofstedt, S. J. Putterman, K. R. Weninger, "Defining the unknowns of sonoluminescence", Physics Reports 281, 65-143 (1997).
- [2] T. J. Matula, S. M. Cordry, R. A. Roy, and L. A. Crum, "Bjerknes force and bubble levitation under single-bubble sonoluminescence conditions", J. Acoust. Soc. Am. 102, 1522-1527 (1997).
- [3] K. R. Weninger, B. P. Barber, and S. J. Putterman, "Pulsed Mie Scattering Measurements of the Collapse of a Sonoluminescing Bubble", Phys. Rev. Lett. 78, 1799-1802 (1997)

## ACKNOWLEDGEMENTS

We wish to thank Ben Newling for his assistance. This research was supported by NSERC Undergraduate Research Award (B.D.) and UNB Research Award (I.M.).

# USING ULTRASONIC AND VISION SENSORS WITHIN EXTENDED KALMAN FILTER FOR ROBOT NAVIGATION

Zhenhe Chen, Ranga Rodrigo, Vijay Parsa and Jagath Samarabandu

Dept of ECE, the University of Western Ontario, London, ON, N6A 5B9, Canada

{zchen56, brodrigo, jagath}@uwo.ca, parsa@nca.uwo.ca

## 1. INTRODUCTION

One of the recent and consistently interesting topics in robotics research community is the simultaneous localization and map-building (SLAM) problem. It requires an autonomous mobile vehicle starting in an unknown environment. Then the vehicle incrementally builds an environment map and simultaneously localizes its pose within this map.

Wolter and his colleagues point out that any approach to master the SLAM problem can be decomposed into two aspects: handling of map features (extraction from sensor data and matching against the (partially) existing map) and handling of uncertainty (Wolter et al. 2004). Therefore, robotic research community keeps tackling the SLAM problem in these two aspects. As for the first aspect, some sensing systems can successfully interpret natural features for mapping, such as ultrasound, computer vision, laser and their fusion. As for the second aspect, some probabilistic methods are proven to minimize the SLAM uncertainties (Thrun 2002).

In this study, we address the problem of robot SLAM within extended Kalman filter (EKF) framework, which falls into the second aspect. Two types of sensing systems, ultrasound and computer vision, which handle the first aspect, are mounted separately on the robot to perceive the environment so as to simultaneously localize the robot and build the landmark map.

## 2. METHODS

### 2.1 Extended Kalman Filter

The Kalman filter (KF) is a linear, discrete-time, finite dimensional system endowed with a recursive structure that makes a digital computer well suited for its implementation (Hayin 2002). While state vector of the KF is suitable only for a linear model, EKF can deal with the nonlinear models in which the majority of the real world applications lie. The conversion from KF to EKF is through a linearization procedure based on first order Taylor approximations (Hayin 2002).

The key reasons for using EKF in SLAM are due to the facts as follows: firstly, EKF directly provides a real-time solution to the navigation problem and to on-going estimation of the uncertainty acquired from vehicle motion and landmark observations; secondly, a number of methods and experience have been developed in aerospace, subsea, and other navigation applications that robotics research community can study. Such wide application is due to the fact that EKF is

an optimal minimum mean square error (MMSE) estimate method and its covariance matrix is proven to converge strongly (Dissanayake et al. 2001).

A classic two-stage EKF algorithm is as follows:

**Prediction Process** (Chen & Samarabandu 2005):

$$\begin{aligned}\hat{\mathbf{x}}(k|k-1) &= \mathbf{f}(\hat{\mathbf{x}}(k-1|k-1), \mathbf{u}(k)), \\ \mathbf{P}(k|k-1) &= \nabla \mathbf{F}_x \mathbf{P}(k-1|k-1) \nabla \mathbf{F}_x^T + \nabla \mathbf{F}_u \mathbf{Q} \nabla \mathbf{F}_u^T, \\ \mathbf{z}(k|k-1) &= \mathbf{h}(\hat{\mathbf{x}}(k|k-1)).\end{aligned}\quad (1)$$

**Update Stage:**

$$\begin{aligned}\hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}\nu(k) \\ \mathbf{P}(k|k) &= \mathbf{P}(k|k-1) - \mathbf{K}\mathbf{S}\mathbf{K}^T,\end{aligned}\quad (2)$$

where

$$\begin{aligned}\nu(k) &= \mathbf{z}(k) - \mathbf{z}(k|k-1) \\ \mathbf{K} &= \mathbf{P}(k|k-1) \nabla \mathbf{H}_x^T \mathbf{S}^{-1} \\ \mathbf{S} &= \nabla \mathbf{H}_x \mathbf{P}(k|k-1) \nabla \mathbf{H}_x^T + \mathbf{R}.\end{aligned}\quad (3)$$

Interested readers can refer (Chen & Samarabandu 2005) for more details of this probabilistic framework.

### 2.2 Ultrasonic Sensing System

Ultrasonic sensors provide a cheap and reliable means for robot localization and environmental sensing when the physical principles and limitations of their operation are well understood.

In this study, beam pattern of an ultrasonic range finder is modeled in a 2-D plane, which will be further discussed in Section 3. Considering the speed of sound is much faster than that of the wheeled robot, the robot movement is omitted between the time interval when transmitter fires and receiver receives the echo. Kleeman *et al.* develop algorithms to localize and classify the features (Kleeman & Kuc 1995), which indicates that sonar is one of the good tools for robotic feature detection. After detecting the feature range of the landmark as well as bearing (from an other sensor, e.g. computer vision or compass) are fed to EKF in Equations. 2 and 3 in order to localize the robot.

### 2.3 Multiple View Geometry

For the purpose of study and comparison, multiple view geometry (MVG) technique is utilized within EKF framework to solve the SLAM problem. Singular value decomposition (SVD) based factorization is applied to 2-D snapshot pictures to reconstruct the landmark in 3-D world coordinates. Map is augmented by the adding reconstructed landmarks.

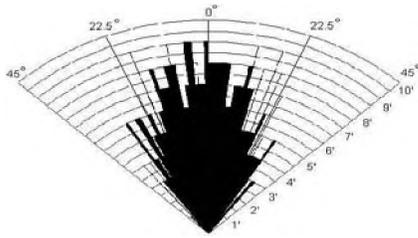


Fig. 1. A Devantech SRF04 sonar range finder beam pattern

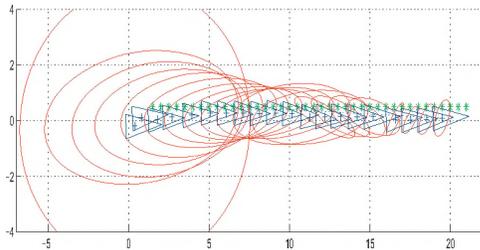


Fig. 2. The trajectory of a mobile robot with ultrasonic sensor for localization

The camera position is estimated using a direct reconstruction technique. This reconstruction outcome is optimized by EKF. Robot navigation uncertainty is reduced (Chen & Samarabandu 2005).

### 3. RESULTS

Simulations are performed in both sensing systems. For the ultrasonic system, a Devantech SRF04 sonar range finder beam pattern is simulated, which is illustrated in Figure 1. It is apparent that the angular scope of the ultrasonic sensor is narrow. Thus a quite large number of ultrasonic sensors is required to mount on the robot and make sure that the robot can “observe” all 360° of its environment. In the simulation, observation Gaussian noise with  $\sigma_z = 2.8$  is added to the ultrasound detection. In Figure 2, the triangles represent the robot trajectory. The asterisks represent features the robot detected, and the ellipses are the uncertainties of robot locations. This shows that the uncertainties decrease as the robot moves. The final average error can be as small as 0.22m. This indicates that the ultrasonic sensors are reliable for robot navigation. However, its short range detection property (less than 10 ft) restricts it to local measurement. Additionally, a non-scanning ultrasonic sensor can only build a 2-D map.

Computer vision technique is suitable for high resolution and long range measurements. MVG integrated with EKF framework is also implemented with observation noise  $\sigma_z = 0.002$  (Figure 3). Simulation results show that EKF can improve the localization accuracies, and recursively build the environment map for robot navigation (Table I).

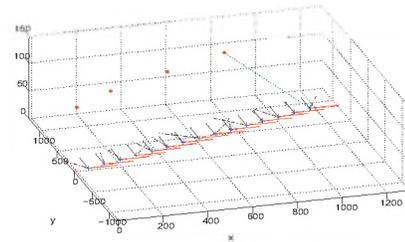


Fig. 3. SLAM by MVG-EKF

TABLE I  
AVERAGE ERRORS WHEN  $\sigma_z = 0.002$ .

	MVG-EKF	MVG	Map
Smooth trajectory	1.7403	4.3289	4.6907
Sharp trajectory	5.8833	14.0874	12.1794

### 4. DISCUSSION

For robotic SLAM problem, ultrasonic and computer vision sensors have their own advantages. For example, ultrasonic sensors are cheap and reliable for short distance detection. Computer vision does well in high resolution and long range measurements, and provides rich research results in feature detection and recognition. Of course, each has their own drawbacks. For ultrasound, its short distance, narrow angular scope of detection and 2-D properties limit its application. For computer vision, it is sensitive to observation noise.

In the SLAM research community, ultrasound based techniques are well established. On the other hand, MVG has a promising future in 3-D SLAM, as the algorithm can be easily applied to single or multiple camera sensing system. If applied to monocular vision-based system, it is benefited from redundant information and avoids unnecessary calibration. MVG-EKF based technique can reduce the robot localization estimation errors compared to using MVG solely.

Future work will include different sensor fusion, e.g. computer vision with ultrasound sensor. Other works will focus on real-time implementation of MVG-EKF algorithm.

### REFERENCES

- Chen, Z. & Samarabandu, J. (2005), Using multiple view geometry within extended kalman filter framework for simultaneous localization and map-building, in 'IEEE International Conference on Mechatronics and Automation', Niagara Falls, Canada.
- Dissanayake, M. W. M. G., Newman, P. & Clark, S. (2001), 'A solution to the simultaneous localization and map building (SLAM) problem', *IEEE Transactions on Robotics and Automation* **17**(3), 229-241.
- Hayin, S. (2002), *Adaptive Filter Theory, 4th Edition*, Prentice Hall, Upper Saddle River, New Jersey.
- Kleeman, L. & Kuc, R. (1995), 'Mobile robot sonar for target localization and classification', *International Journal of Robotics Research* **14**(4).
- Thrun, S. (2002), *Robotic Mapping: A Survey*, In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millennium*, Morgan Kaufmann. URL: <http://robots.stanford.edu/papers.html>
- Wolter, D., Latecki, L., Lakamper, R. & Sun, X. (2004), 'Shape-based robot mapping', *Advances in Artificial Intelligence, Proceedings Lecture Notes in Computer Science* (3238), 439-452.

# KALMAN FILTERING OF ACOUSTIC EMISSION SIGNAL GENERATED BY LASER-MATERIAL INTERACTIONS

Dmitry S. Solodovnikov<sup>1,2</sup>, Evgueni V. Bordatchev<sup>2,1</sup>

<sup>1</sup>Dept. of Mechanical Engineering, The University of Western Ontario, 1151 Richmond St., London, ON, Canada, N6A 5B8

<sup>2</sup>NRC Integrated Manufacturing Technologies Institute, 800 Collip Circle, London, ON, Canada, N6G 4X8

## 1. INTRODUCTION

Laser material removal process (LMRP) requires a proper selection and optimization of a large number of interdependent process parameters related to laser, optics, workpiece material, and the motion system. All of these parameters significantly influence accuracy, precision and surface quality of the laser machined parts. Therefore, on-line process monitoring and control is required based on effective signal processing algorithms and reliable, physically observed information on actual process parameters and the state of the laser-material interactions.

This work was conducted to study an applicability of the Kalman filtering for on-line monitoring of the laser material removal process. The study is based on an experimental investigation of the informational properties of the acoustic emission (AE) signal generated by the surface acoustic waves in the laser-material interaction zone, selection of the signature that reliably characterizes the LMRP, and comparative analysis of the state variables affected by a working distance, which is one of the most critical process parameters.

## 2. EXPERIMENTAL SET-UP AND PROCEDURE

Figure 1 shows the schematic of the experimental set-up and procedure. The laser-material removal experiments were carried out on a brass foil with a thickness of 152  $\mu\text{m}$  using Q-switched diode pumped solid state Nd:YVO<sub>4</sub> laser system with a pulse width of 20 ns and wavelength of 1064 nm. During the experiments, working distance (WD) was varied from 14.710 mm to 15.385 mm. Several trenches were machined by moving a sample with a feed rate of 2 mm/s and by applying laser pulses with a frequency of the 10 kHz and pulse energy of 0.055  $\mu\text{J}$ . The surface acoustic waves were measured during LMRP experiments using a high-fidelity acoustic emission transducer with a frequency bandwidth of up to 500 kHz, which was directly placed in contact with the top surface of the laser ablated sample 30 mm away from the LMRP zone. The AE signals were recorded by an oscilloscope for a total duration of 130 ms with a sampling period of 1  $\mu\text{s}$ .

## 3. LMRP'S DYNAMIC CHARACTERISTIC

During the preliminary signal analysis, it was observed that each measured AE signal corresponding to an

unchanged working distance had a noticeable periodic structure with a period of applied laser pulses. In addition, all the periodic structures had consistent signal signatures with small variations for the LMRP with unchanged process parameters. This signature was selected for further analysis as an LMRP's dynamic characteristic, which characterizes an LMRP's state through the AE response on applied laser pulse. Figure 2 shows an LMRP's dynamic characteristic,  $y(t)$ , experimentally obtained for a working distance of 15.01 mm as an averaged signature of 1000 synchronized periodic structures with a duration of 100  $\mu\text{s}$  between applied 10 kHz laser pulses.

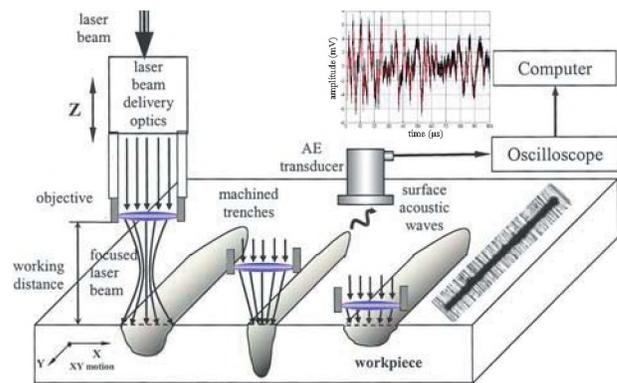


Figure 1. Schematic of the experimental set-up and procedure.

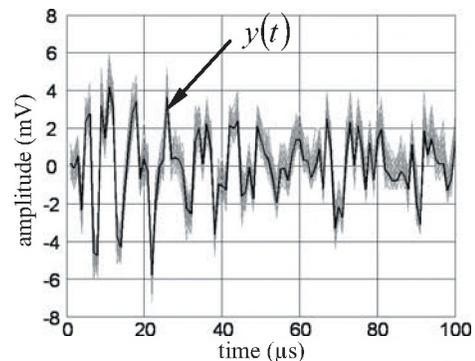


Figure 2. Experimentally obtained LMRP's dynamic characteristic.

The LMRP was described in the general form of a linear, discrete time, time-invariant state space model:

$$\begin{aligned} \hat{\mathbf{x}}(n+1) &= \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}u(n) + \mathbf{D}w(n) \\ y(n) &= \mathbf{C}\hat{\mathbf{x}}(n) + v(n) \end{aligned} \quad (1)$$

where  $\hat{\mathbf{x}}(n)$  is the LMRP's state vector,  $u(n)$  is the control input vector associated with an applied laser pulse;  $y(n)$  is the measured output vector, which is an LMRP's dynamic

characteristic;  $w(n)$  and  $v(n)$  are the process and measurement noises, respectively;  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  are the system state, input, measurement and noise matrices, respectively. Eq. (1) was considered as an informational model of LMRP in the form of a transfer function between applied laser pulse and measured AE signal and it was used to calculate  $\hat{y}(t)$  as an estimation of  $y(t)$ . The optimal system order of 14 and elements of  $\mathbf{A}_{14 \times 14}$ ,  $\mathbf{B}_{14 \times 1}$ ,  $\mathbf{C}_{1 \times 14}$ ,  $\mathbf{D}_{14 \times 1}$ , matrices were identified using the deterministic-stochastic realizations algorithm [2] applied to the experimentally obtained LMRP's dynamic characteristic. The measurement noise was estimated based on the measured AE signal before the actual laser-material interactions. The process noise was estimated using the deviations between the measured AE signal and  $y(t)$ .

#### 4. KALMAN FILTERING

The matrices,  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$ , were used in the standard Kalman filtering algorithm [1] for linear time-invariant stochastic systems:

$$\begin{aligned} \tilde{\mathbf{x}}(n+1) &= \mathbf{A}\tilde{\mathbf{x}}(n) + \mathbf{B}u(n) + \mathbf{H}r(n) \\ r(n) &= y(n) - \tilde{y}(n) \\ \tilde{y}(n) &= \mathbf{C}\tilde{\mathbf{x}}(n) \end{aligned} \quad (2)$$

where  $\mathbf{H}$  is the Kalman Filter gain matrix, and  $r(n)$  is the residual vector which represents the quality of the Kalman filtering as a difference between measured and predicted LMRP's dynamic characteristics.

Figure 3 shows a comparison between the measured, estimated, and predicted LMRP's dynamic characteristics obtained for a working distance of 15.01 mm. The quality of the identification and simulation was estimated from the correlation coefficient and the mean-square error with respect to  $y(t)$ . The results are shown in Table 1.

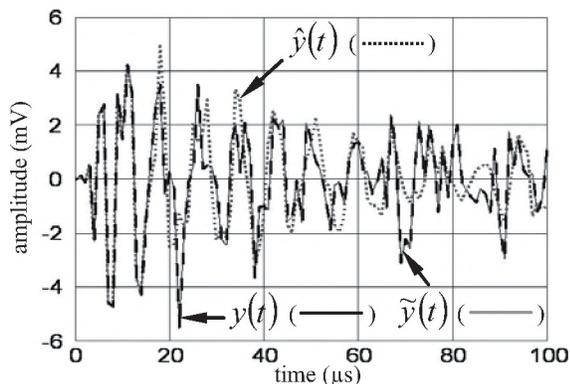


Figure 3. Comparison of measured, estimated and predicted LMRP's dynamic characteristics.

The LMRP's dynamic characteristics were obtained experimentally and estimated using Eq. (1), the Kalman filters were determined and the corresponding state vectors were studied in details for three different working distances,

{14.76, 15.01, 15.26} mm. One of the 14 state variables considered was found to be affected considerably more than others by the working distance. Figure 4 shows the comparison of these vector signatures, which had variances of {3.82, 1.64, 4.72}  $\text{mV}^2$  respectively.

Table 1. Quality of identification and simulation

	$\hat{y}(t)$	$\tilde{y}(t)$
correlation coefficient, dimensionless	0.714	0.998
mean-square error, mV	2.006	0.099

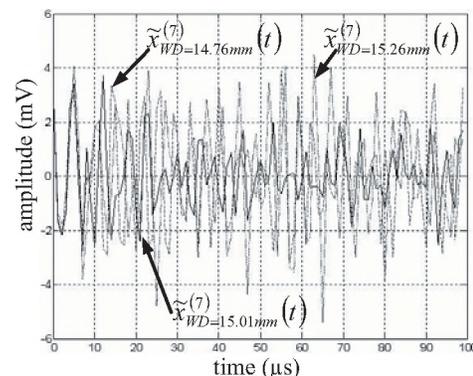


Figure 4. Comparison of the state vectors.

#### 5. SUMMARY AND CONCLUSIONS

This paper presents a study of applicability of the Kalman filtering methodology for on-line monitoring of the laser-material removal processes. Based on the experimental observations, a periodic signature with a period of applied laser pulses was selected for characterization of the LMRP's dynamic characteristic. An informational model of LMRP was proposed as a transfer function between applied laser pulse and measured AE signal. The system matrices of the state space representation of this model were identified that allowed calculation of the Kalman filters and evaluation of the measured, estimated, and predicted LMRP's dynamic characteristics obtained for working distances of {14.76, 15.01, 15.26} mm. The effect of the working distance on the state variables was studied. The following conclusions can be drawn from this study:

1. The proposed dynamic characteristic reliably characterizes LMRP.
2. The Kalman filtering methodology allows extraction of reliable information about the actual state of LMRP.
3. The state variables are significantly affected by the process parameters and therefore they can be used for on-line monitoring, diagnostic and control of LMRP.

#### REFERENCES

1. The Control Handbook (1996). Ed: Levine, W. S., CRC Press, p. 1548.
2. David Di Ruscio. (1997). A method for identification of combined deterministic stochastic systems. In Applications of Computer Aided Time Series Modeling. Lecture Notes in Statistics Series, Springer Verlag, pp. 181-235.

# COLLABORATIVE NOISE CONTROL – A CASE STUDY

Ramani Ramakrishnan<sup>1</sup> and Werner Richarz<sup>2</sup>

1 – Department of Architectural Science, Ryerson University, Toronto

2 – Aercoustics Engineering Ltd, Toronto

## 1.0 INTRODUCTION

Acoustic and noise control consulting covers a wide spectrum of potential projects. The noise control engineer is expected to provide the complete package of services in noise control from analysis, design and actual implementation even though some aspects may be outside his/hers area of specialty. One such instance will be explored and described through the following case study.

## 2.0 SITE DESCRIPTION

Dust collector fans are ubiquitous in any process industries that emit reasonably sized particulates into the atmosphere. These fans, usually centrifugal in designs, come in various sizes and specifications. A schematic detail of one such design in a cement plant in Northern USA is shown in Figure 1. The design is an ID fan, centrifugal, located downstream of the large dust-collection system. The fan's specifications are: 44,000 CFM, 1800 RPM, 250 HP, Pressure rise of 26" H<sub>2</sub>O. The outlet is a 38" diameter circular pipe and the exhaust stream temperature is 200° F. The fan produces a strong blade passing tone (around 296 Hz), which is further amplified by the constricted outlet design of the ID fan. The amplified tone of the fan is clearly audible at residences located, approximately, a kilometre away, depending on the wind directions. The residents lodged strong complaints and the local municipality issued orders to the industry to remove the offending noise.

The ID fan noise was measured at four locations, three (Locations A, B and C) in the plant and the fourth locations was in the yard of one of the residents. The plant measurement locations are identified in Figure 1. The results of the measurements are shown in Figures 2 and 3. The data shows that the noise is highly tonal and is purely from the outlet duct. Location C level in the 315 Hz band is 103 dB and is at least 20 dB more than the 315 Hz band levels for the other two locations. The narrowband spectrum of Figure 3 shows that the tone is clearly audible at the residential receptors.

The plant installed a short lined expansion as a possible passive silencer, also shown in Figure 1. The measure provided no noise attenuation. A simple solution would have been to install bullet silencer to provide a noise reduction of about 15-20 dB in the 315 Hz band. However, the fan did not have any headroom for the pressure drop of the silencer. Instead, the plant would have to redesign the fan without the noise concerns. However, the new fan design was not practical due to the large lead time, in excess of a year or so, required for the implementation of the new design.

Another possible solution was to install active noise control. However, one of the manufacturer of HVAC active noise control systems was unable to provide a suitable solution for the following reasons: the offending source produces a higher order duct mode (the cut-off frequency of the second mode inside the 38" diameter outlet duct at 200° F was less than 250 Hz); and the temperature of 200° is too high for the active noise control system components. Even though the first author has a strong theoretical understanding of active noise control, he has no actual design experience. Further, we were convinced an active noise control system can be suitably designed. Hence, we approached our colleagues in the consulting field with design expertise and requested their assistance. The resulting collaboration resulted in resolving the noise concerns of the ID fan. The details of the active noise control system implementation are described below.

## 3.0 ACTIVE NOISE CONTROL

The noise control objective was a substantial reduction of the fan blade passage tone. The nominal frequency of 300 Hz as well as the duct diameter put this at the upper limit of the capabilities of an active noise control system. In order to assure a high degree of stability a conventional dissipative silencer was interposed between the sensing microphone and the noise cancelling loudspeakers. In this manner the controller can be configured to operate in a 'feed-forward' mode wherein the signal to be controlled is measured by the sensing microphone and phase and amplitude adjustments are made based on feed-

back from the error microphone that is positioned well away from the duct discharge opening. The system details are shown in Figure 4.

The first configuration used a sensing microphone that was located close to the fan discharge. The sensing microphone was configured in a manner similar to that of a turbulence screen. Because of the hot environment a Knowles high temperature microphone was used. Unfortunately, the signal level blade passage frequency was not dominant, and the controller did not lock onto the reference signal. This interference problem was eliminated when the sensing microphone was moved further away from the fan. A sketch of the installed system is shown in Figure 5.

The performance of the active control is illustrated in Figure 6. The system reduces the blade passage tone by 12 dB. The system was first run during the evening hours. When the morning shift arrived for work, they thought that the plant was shut down, since the audibility of the familiar fan tone was greatly reduced. They were however disappointed when they found out that there was no provision for a parallel signal input so that the loudspeakers could play their favourite music.

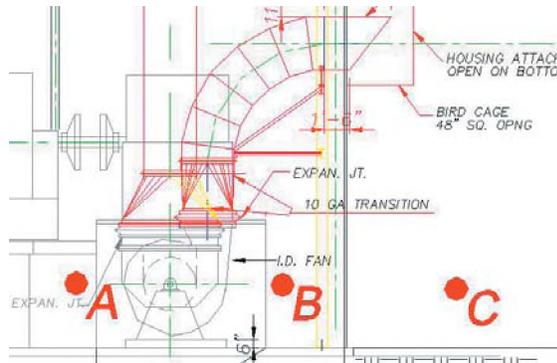


Figure 1. ID Fan Details

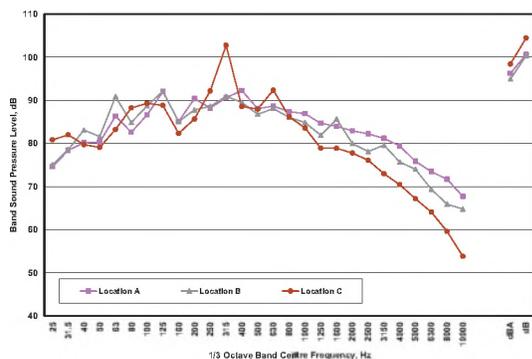


Figure 2. 1/3 Octave Band Spectra of ID Fan Noise.

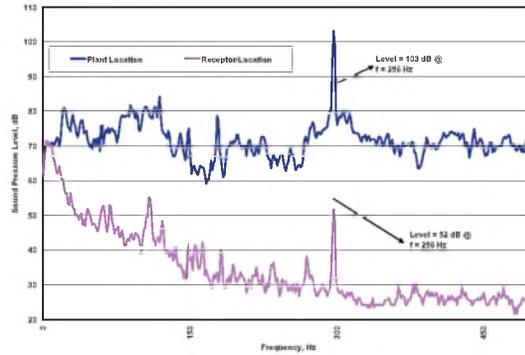


Figure 2. Narrow Band Spectra of ID Fan Noise.

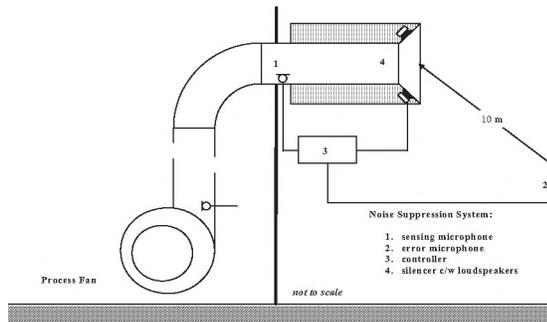


Figure 4. Active System Components



Figure 5. Active Noise Control Silencer

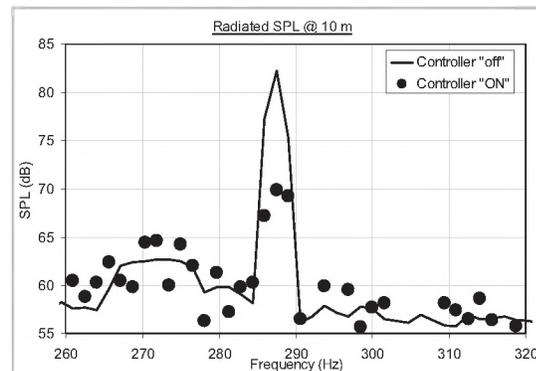


Figure 6. Performance of Active Noise Control System.

# USE OF PSYCHOACOUSTIC METRICS FOR THE ANALYSIS OF NEXT GENERATION COMPUTER COOLING FAN NOISE

Colin Novak<sup>1</sup>, Helen Ule<sup>1</sup>, Robert Gaspar<sup>1</sup>, and Robert Wiley<sup>2</sup>

<sup>1</sup>Dept. of Mechanical Automotive and Materials Engineering, University of Windsor  
401 Sunset Ave.; Windsor, Ontario; Canada; N9B 3P4

<sup>2</sup>ADDA Corporation – North American Thermal Division, 910 Campisi Way, Campbell, California; USA; 95008

## 1. INTRODUCTION

The competitive drive for performance improvements in computer processing abilities necessitate the inclusion of finned active cooling devices to dissipate the heat generated by the computer processing unit (CPU). In more recent years, the heat flux generated has increased with further performance improvements. The result being that much more complex forced air cooling solutions are required. One aspect that now challenges the industry is to provide this cooling while minimizing noise generation. A significant feature to this problem is missed if one only considers the problem to a one-dimensional sound level issue. From a consumer's perspective, the perceived quality of the noise emitted takes precedence over what traditional acoustical analysis techniques of this fan noise may imply. Here, sound quality metrics may be a more applicable analysis tool as it makes possible the quantification of these qualitative human impressions.

The present study investigates the validity of using several psychoacoustic metrics for acoustic analysis of two different cooling solutions. A discussion and comparison of measured results using traditional analysis techniques is also included. A discussion of the applicability of the various metrics along with justifications is presented.

## 2. MEASUREING COMPUTER FAN NOISE

The procedures, installation and operating conditions of acoustic testing for computer noise sources are regulated by several standards. These include ECMA-74, "Measurement of Airborne Noise Emitted by Information Technology and Telecommunication Equipment" and ISO 7779, "Measurement of Airborne Noise Emitted by Information Technology and Telecommunication Equipment". The method used to measure and calculate sound power level using free field sound pressure measurements is detailed in ISO 3745, "Determination of Sound Power Levels of Noise Sources".

Traditional analytical approaches serve well to quantify the amplitude of acoustic emissions, but they offer no suggestion as to the quality of the sound produced by the

cooling fans. Sound quality can significantly affect the acceptability of a product to the consumer. Given this, acoustic product evaluation of next generation computer cooling solutions must also include a sound quality or psychoacoustic analysis in order to truly determine the full acoustic impact that an active cooling solution will have on the end user.

The science of psychoacoustics involves the quantitative evaluation of these subjective sensations using sound quality metrics. Application of sound quality metrics allow for the visualization of the complicated relationship that exists between the physical and perceptual acoustic quantities. For this investigation, sound quality metrics including loudness, sharpness, roughness, prominent tone and articulation index were used to evaluate cooling fan noise of three different cooling fan-sink designs. The designs were chosen on the basis of the variety of acoustic characteristics they exhibited.

## 3. RESULTS

Figures 1 and 2 illustrate the results of the sound pressure level and sound power level versus fan RPM measurements respectively for each of the three fan designs. What is most relevant for this study is the realized amplitude range of acoustic emissions.

Inspection of both the sound pressure level and sound power level results illustrate a linear increase in acoustic emission levels. Observable from these figures is the fact that the smaller fan (design 1) spans the largest speed range and can run at the highest RPM when compared to a blower type fan (design 3) option; yet, they both generate approximately the same range of acoustic noise. The second design had the smallest range of both RPM and acoustic noise. A conclusion that one might draw from both Figures 1 and 2 is that fan RPM range has a greater impact on predicting noise emissions rather than the simple magnitude of the fan speed.

Figure 3 illustrates the loudness results versus fan speed for the three designs. The calculation of loudness involves an algorithm which is frequency dependant and includes other characteristics such as temporal masking effects. This

dimension of human perception adds to the meaningfulness of using this metric.

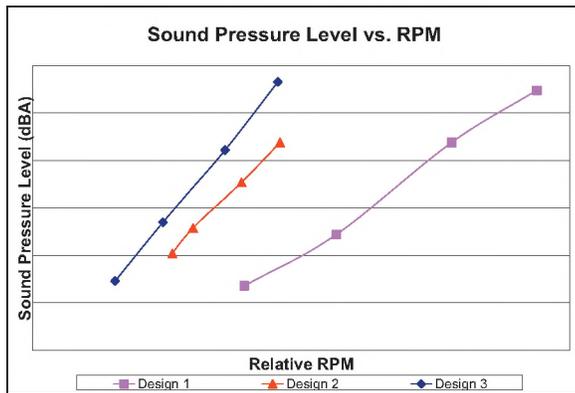


Figure 1: Sound Pressure Level vs. RPM for the Three Fan-Sink Design Options

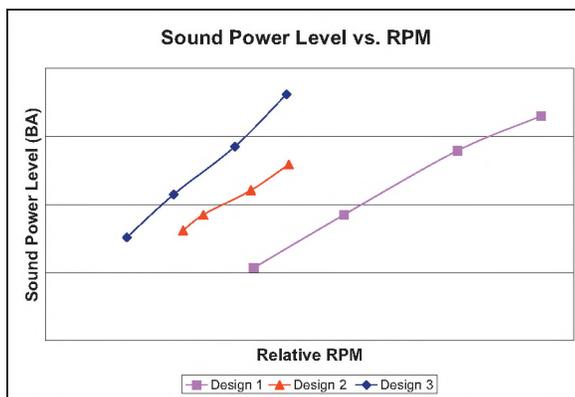


Figure 2: Sound Power Level vs. RPM for the Three Fan-Sink Design Options

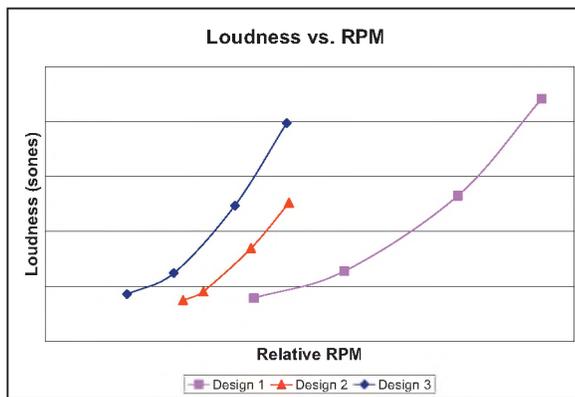


Figure 3: Loudness vs. RPM for the Three Fan-Sink Design Options

Table 1 illustrates the results of the remaining psychoacoustic metrics for each of the three fan designs at each of the operating speeds tested. Included are sharpness, roughness, prominent tone and articulation index.

Table 1: Psychoacoustic Results

<i>Design 1</i>				
RPM %	40 %	60 %	80 %	100 %
Sharpness	1.99	2.25	2.95	3.45
Roughness	3.70	3.85	4.11	4.02
Prominent Tone (dB)	6.9 @ 465 Hz	7.9 @ 727 Hz	6.9 @ 1288 Hz	10.3 @ 845 Hz
Articulation Index (%)	99.05	98.05	89.35	80.13
<i>Design 2</i>				
RPM %	44 %	62 %	84 %	100 %
Sharpness	1.56	1.86	2.69	3.72
Roughness	4.12	5.15	5.56	6.32
Prominent Tone (dB)	6.1 @ 656 Hz	7.3 @ 1845 Hz	4.5 @ 2088 Hz	5.1 @ 2250 Hz
Articulation Index (%)	99.45	99.45	99.4	95.32
<i>Design 3</i>				
RPM %	35 %	53%	80 %	100 %
Sharpness	1.1	1.2	1.61	1.75
Roughness	3.21	3.33	4.45	4.25
Prominent Tone (dB)	5.3 @ 669 Hz	10.8 @ 724 Hz	7.1 @ 419 Hz	8.7 @ 581 Hz
Articulation Index (%)	99.10	98.95	98.95	90.45

Inspection of the remaining psychoacoustic metrics provides valuable information. Inspection of sharpness, which is an indicator of high frequency annoyance, increases consistently with fan speed. This is not unexpected given that as the fan speed is increased, so is the amount of air noise and turbulence. A similar argument holds for the modulation metric of roughness perhaps indicating some imbalance. Such is not true for the prominent tone results for which high levels above 8 dB are more likely due to the fan blade passage frequency, which is a multiple of the fan speed, number of fan blades and any obstructions such as sink fins. For circumstances where a noise source is located in an area where the comprehension of speech is important, the measurement of articulation index is a very useful metric. In this case, all reported values are acceptable.

#### 4. CONCLUSIONS

The focus of this investigation was to investigate the validity of using several different psychoacoustic metrics for the analysis of fan cooling computer noise. Traditional noise analysis techniques were also conducted.

The results of most of the metrics were correlated with the operational fan speeds and all of the metrics demonstrated results useful for qualitative product analysis. It can further be concluded that while traditional noise analysis techniques provide useful information, they do not represent the entire impact that computer cooling fan-sink noise can have on the human ear.

# EFFECT OF NEXT GENERATION COMPUTER COOLING FAN SPEED ON ACOUSTIC NOISE EMISSIONS

**Helen Ule, Colin Novak and Robert Gaspar**

Dept. of Mechanical Automotive and Materials Engineering, University of Windsor  
401 Sunset Ave.; Windsor, Ontario; Canada; N9B 3P4

## 1. INTRODUCTION

Graphic Processor Units (GPUs) on the latest models of computer graphic cards generate significant amounts of heat. In fact, the required dissipation rate is so large that cooling fans mounted on heat-sinks must be used to maintain satisfactory temperatures. Where space limitations allow and heat transfer requirements dictate, blower type fans are implemented due to their ability to deliver higher flow rates compared to more traditional axial flow fans. The operation of these blower fans, particularly at high speeds, results in the generation of noise which is experienced by the user. Both computer manufacturers and consumers alike have deemed this noise to be excessive and annoying.

The purpose of this study was to investigate the realized acoustic performance of a blower style fan-sink mounted on an advanced graphics port (AGP) card. The goal of this investigation was to determine what thermal benefits of higher flow rate are realized by the blower fan at the expense of increased noise emissions.

## 2. METHOD

To determine the relationship between fan speed and acoustic performance, the RPM on the fan was adjusted to operate at five different speeds by varying the input voltage to the fan. All acoustic testing was performed with the fan-sink mounted on a video graphic card in a stand-alone arrangement where the video card was not placed inside a computer chassis.

The sound power levels of the fan-sinks were calculated from sound pressure level measurements collected in a hemi-anechoic room. The testing was carried out in compliance with the ECMA-74, "Measurement of Airborne Noise Emitted by Information Technology and Telecommunication Equipment". Sound power levels were calculated from the sound pressure measurements as detailed in ISO 3745, "Determination of Sound Power Levels of Noise Sources". This method involved the acquisition of ten sound pressure level measurements around the noise source in a hemi-spherical pattern specified in ISO 3745. Adjustments were also made to account for the

surface area of the test sphere as well as atmospheric conditions and ground reflections.

Sound pressure level results were also measured in the hemi-anechoic room using a binaural head manikin. The rationale for using a binaural head is to acquire data that best represents what would be perceived by an actual person for psychoacoustic analysis. The installation and operating conditions of the sound pressure level measurements were conducted in compliance with both ECMA-74 and ISO 7779, "Measurement of Airborne Noise Emitted by Information Technology and Telecommunication Equipment".

As part of this investigation, psychoacoustic metrics were used as part of the evaluation process. Psychoacoustics is a measure of the perception of how good or bad the character of a noise source is, as perceived by the human ear. For this study loudness and prominent tone (PR) were the evaluated psychoacoustic metrics. Loudness is a measure of the human perception of how loud a source is perceived to be as opposed to simply reporting a sound pressure level. Prominent tone provides an objective measure on the prominence of a tonal component in a sound. A tone is said to be prominent if its PR exceeds 7 dB and is usually reported with the frequency at which the prominent tone is located.

## 3. RESULTS

Figures 1 and 2 illustrate the results of both sound pressure level and sound power level vs. fan RPM measurements, respectively for the blower fan-sink. A linear increase in levels is seen for both with increases in fan speed. Figure 3 illustrates the loudness results versus fan speed for the blower heat-sink. Unlike the sound pressure and power level graphs, the loudness curve is not linear. Loudness is very frequency dependant. Its calculation is rather complicated and includes other characteristics such as temporal masking effects. As such, no simplistic relationship exists between the sound level metrics and loudness. Loudness, however is still a good indication of perceived noise emission and is therefore still expected to increase exponentially since it is not presented by

logarithmic values. This exponential trend is evident in Figure 3.

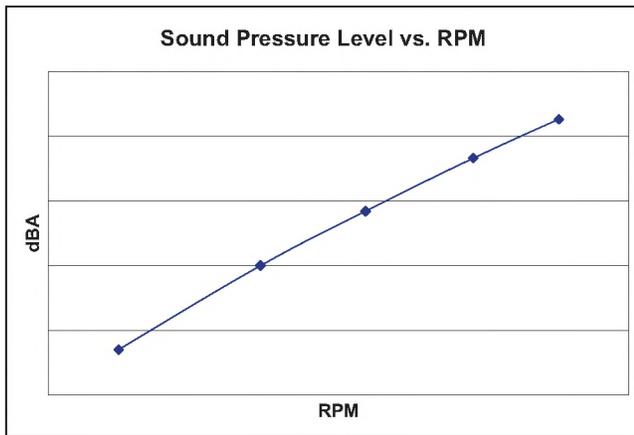


Figure 1: Sound Pressure Level vs. RPM for Blower Fan-Sink

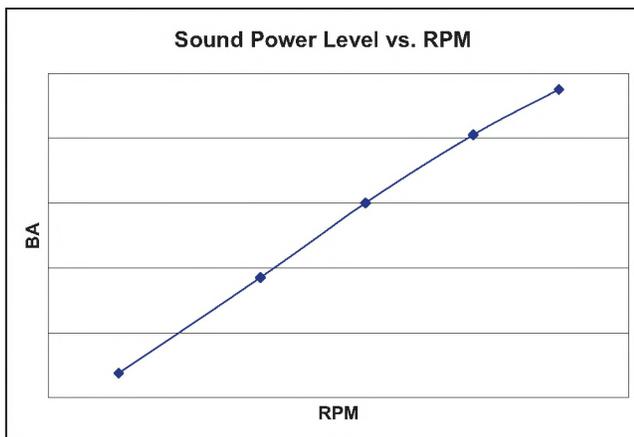


Figure 2: Sound Power Level vs. RPM for Blower Fan-Sink

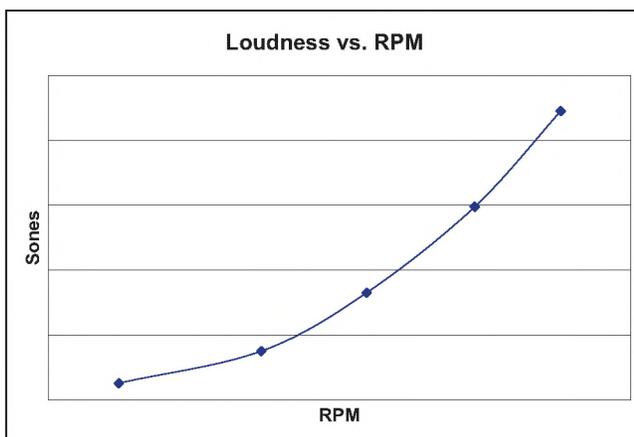


Figure 3: Loudness vs. RPM for Blower Fan-Sink

Table 1 illustrates the measured prominent tones and the respective frequencies at which they are found. Inspection of the results suggests that no direct relationship between

fan speed and prominent tone exists. Indirectly this is false. Aside from the case of a faulty operating fan, a prominent tone is often the result of spatially fixed irregularities that can produce a wake at either the inlet or outlet. These dynamic changes can cause prominent tones at the fan blade passage frequency, which is a multiple of the fan speed, number of fan blades and any obstructions such as sink fins. The amplitude of these resonances do not increase with fan RPM but instead increase and decrease like an incoming wave and will occur at frequencies that are multiples of each other.

Table 1: Prominent Tone Result for Blower Fan-Sink

Voltage	RPM %	Prominent Tone (dB)
6	48	8.4 @ 669 Hz
8	63	8.7 @ 594 Hz
10	77	9.6 @ 581 Hz
12	90	6.7 @ 581 Hz
14	100	4.8 @ 606 Hz

As part of a previous investigation, it was reported that the thermal performance of this fan-sink improved with an increase in fan RPM. However, this thermal performance improved at a lesser rate when the RPM was increased further. This, along with the acoustic results presented here, suggest that a diminished rate of thermal improvement with increased fan speeds are accompanied by negative increases in unwanted noise emissions. This was demonstrated using the metrics of sound pressure and power levels as well as loudness.

#### 4. SUMMARY AND CONCLUSIONS

Acoustic experiments were conducted on a next generation blower type fan-sinks to investigate the benefits of higher fan flow rate versus the effect on noise level. A previous study showed that thermal performance of the fan-sink improved with an increase in fan RPM, however, the performance improved at a lesser rate when the RPM was increased further. Noise on the other hand increased steadily. No direct correlation between an increase in fan speed and prominent tone was realized.

This study showed that while increasing fan speed did increase thermal performance, it also had a negative effect on noise emissions. Care must be taken in optimizing fan speed so as to gain the best thermal performance without producing detrimental noise levels.

# SUBBAND ADAPTIVE MODELING OF DIGITAL HEARING AIDS

Michael Wirtzfeld, Vijay Parsa

National Centre for Audiology, Faculties of Health Sciences & Engineering, Elborn College,  
University of Western Ontario, Ontario, Canada, N6G 1H1. Email: wirtzfeld@nca.uwo.ca, parsa@nca.uwo.ca

## 1. INTRODUCTION

The intelligibility and quality of processed speech are central concerns for designers and manufacturers of hearing aids, clinicians who prescribe and fit hearing aids, and importantly the end users. Conventional standardized and other widely-available test procedures [1] have been directed at manufacturing quality control, and provide very limited information on how hearing aid processing affects the intelligibility and quality of speech and other sounds of importance to the user. In particular, these procedures often neglect the fact that many modern hearing aids are adaptive, changing their electro acoustic characteristics in response to changes in the acoustic environment.

Researchers have recently begun to examine the potential for evaluating hearing aids using more complex stimuli that better approximate “real world” signals. For example, the ANSI standardized Speech Intelligibility Index (SII) [2] has been applied to predict the speech intelligibility improvement obtained through a hearing aid. Coherence measures, representing the linear relation between a broadband noise input and the resulting hearing aid output, have been studied as ways to predict the sound quality and speech intelligibility performance of hearing aids [3]. An alternative approach to quantifying the performance of digital hearing aids is to dynamically model their behavior using a system identification approach as this offers the flexibility of testing a hearing aid with speech and music stimuli [4, 5]. In this method, the hearing aid is modeled as a linear time-varying system and its response to speech and music stimuli is predicted using a linear adaptive filter. It is assumed that the model residual is mainly composed of distortion and noise components of the hearing aid under test [4]. The relative level of the distortion and noise can be quantified using a simple metric such as the Signal-to-Noise Ratio (SNR) or using a more sophisticated metric such as the Perceptual Evaluation of Speech Quality (PESQ) [6] which incorporates models of auditory perception.

Previous studies with analog hearing aids have shown that the speech quality metrics derived from using the system identification approach correlated well with perceptual judgments of speech quality, both by normal and hearing impaired listeners [4]. However, this method

has not been tested with modern digital hearing aids, majorities of which employ multi-channel compression among other advanced signal processing features. This necessitates the need for subband models to correctly model modern digital hearing aids [5].

## 2. SUBBAND ADAPTIVE MODEL

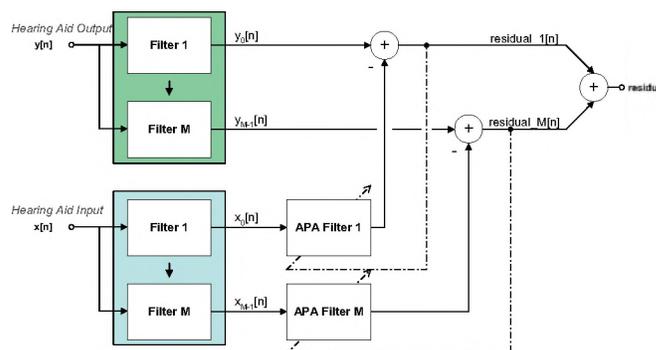


Figure 1. Block Diagram of the Subband Adaptive Model

Figure 1 illustrates the subband architecture considered in this paper. The recorded hearing aid output and input sequences,  $y[n]$  and  $x[n]$ , are filtered using uniform,  $M$ -band analysis filter banks. The resulting subband output sets,  $y_0[n], \dots, y_{M-1}[n]$  and  $x_0[n], \dots, x_{M-1}[n]$ , form the desired and reference sequences to the adaptive filter blocks (APA Filter 1, ... , APA Filter  $M$ ), respectively. Each of the constituent adaptive filter blocks is implemented as a finite impulse response (FIR) filter whose coefficients are updated using a complex affine projection algorithm based on recursive matrix updating.

## 3. METHOD

We have previously studied the performance of the subband adaptive model in characterizing the behavior of multi-channel compression hearing aids using computer simulations [5]. In this paper, we report initial results from subband modeling of two commercial hearing aids - Oticon Syncro and the Bernafon Symbio. The Syncro is a multi-channel compression hearing aid with eight “voice aligned” compression channels, while the Symbio is a “channel-free” digital hearing aid.

Each of these aids was first programmed to fit a steeply sloping, moderate to severe hearing loss profile. In order to focus on the compression characteristics of each hearing aid, all unrelated signal processing features of the device were deactivated. Each aid was placed in a Brüel & Kjaer anechoic test box, Type 4232, with accompanying microphones and preconditioning amplifiers. Ten Hearing In Noise Test (HINT – House Ear Institute of Los Angeles, CA, USA) speech sentences were concatenated and played back at 65 dB SPL. The hearing aid output was recorded through a 2 cc coupler, while a separate reference microphone was used to record the unprocessed speech. These two signals were then applied to the subband adaptive model where the number of analysis bands was altered over a range including the number of channels in the hearing aid being tested. For each analysis band, a Signal-to-Error Ratio (SER in dB) and the PESQ – Mean Opinion Score (MOS) were calculated.

#### 4. RESULTS & CONCLUSIONS

Figure 2 illustrates the results of subband modeling of Syncro and Symbio hearing aids for HINT 1-1 sentence. The FIR filter length in the adaptive filters was set to 256 taps and the projection order for the APA algorithm was 15. It can be seen that lower number of bands in the analysis filter bank results in lower SER values as the model cannot adequately characterize the complex, multi-channel compression of the hearing aid. An overall asymptotic trend can be observed for the SER for increasing analysis bands. It can also be noticed that an increase in the number of subbands results in much better performance for Syncro. Since Syncro has eight independent compression channels, at least 8 bands are required in adequately characterizing its dynamic behavior. On the other hand, Symbio is marketed as a channel-free compression hearing aid, and the results show that the performance improvement with an increasing number of subbands is not as significant.

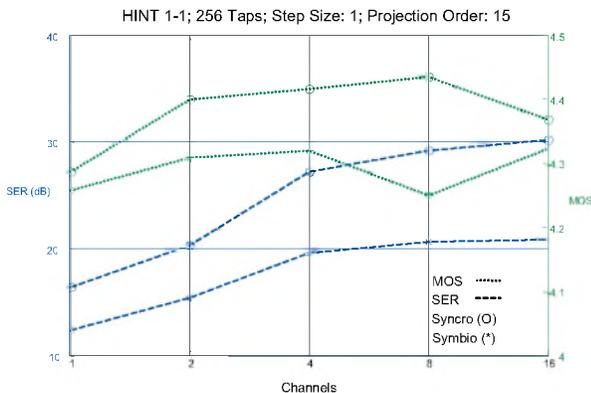


Figure 2. MOS and SER Values

Similar conclusions can be drawn from the PESQ MOS results. For the Syncro, the subband model does a

poor job of characterizing the compression behavior of the hearing aid when it has either too few or too many analysis bands. The largest MOS value occurs when the model has eight analysis bands. For the Symbio, the largest MOS value occurred with sixteen analysis bands.

In conclusion, to properly characterize the complex and dynamic behavior of multi-channel amplitude compression strategies used in current generation hearing aids, a subband adaptive model is necessary. Our previous simulation results [5] and the experimental results shown in this paper suggest that as the number of bands in the subband adaptive model increases to match or exceed the number of compression channels in the hearing aid being modeled, the effectual SER value improves in a positive, asymptotic manner. In addition, the PESQ MOS scores support the generalized behavior of the subband model. It also appears that the subband adaptive model is able to characterize devices which process speech in either in the temporal domain (Bernafon Symbio), or the frequency domain (Oticon Syncro).

#### 5. REFERENCES

- [1] American National Standards Institute, "Specification of Hearing Aid Characteristics" ANSI S3.22, New York, 1996.
- [2] American National Standards Institute, "Methods for the Calculation of the Speech Intelligibility Index", ANSI S3.5, New York, 1997.
- [3] J. Kates & K. Arehart, "Coherence and the Speech Intelligibility Index", *The Journal of the Acoustical Society of America*, Vol. 117, No. 4, pp. 2224–2237, April 2005.
- [4] V. Parsa & D.G. Jamieson, "Hearing Aid Distortion Measurement Using the Auditory Distance Parameter", Audio Engineering Society, Convention Paper 5434, September 2001.
- [5] M.R. Wirtzfeld & V. Parsa, "On Subband Adaptive Modeling of Compression Hearing Aids," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume 3, Page(s): 45-48, Philadelphia, PA, March 18-23, 2005.
- [6] International Telecommunications Union, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs", ITU-T Recommendation P.862, 2001.

#### 6. ACKNOWLEDGEMENTS

We gratefully acknowledge the financial support by the Oticon Foundation, Denmark, the Ontario Rehabilitation Technology Consortium, Canada, and the NSERC, Canada.

# AN EVALUATION OF THE IMPACT AND FEASIBILITY OF THE ANSI S3.22 2003 STANDARD FOR TESTING HEARING INSTRUMENTS

Andreas Seelisch<sup>1</sup> and Susan Scollie<sup>1</sup>

1) National Centre for Audiology, University of Western Ontario, London, ON, N6G 1H1

## 1. Introduction

The purpose of a standard is to maintain comparison and reproducibility of Hearing Instrument (HI) measures across different facilities (Staab, 2002.) This project evaluated the ANSI S3.22-2003 standard adopted this May 2005 to measure its feasibility and to document any electroacoustic changes it would produce.

The 1987 version of the ANSI s3.22 required that all tests on automatic gain control (AGC) hearing aids be done with the AGC set for maximum effect. In 1996, this was changed to allow the manufacturer to specify the settings for the tests. Manufacturer specified settings tend to vary widely. Consequently, the 2003 version reverted back to the 1987 wording for the two tests of AGC function. This is intended to indicate hearing aid function in compression settings. The new standard accomplishes this by requiring the HI be set for “minimal AGC effect” in the first 2 phases of the test, Full on Gain (FOG) and Reference Test Position (RTP), and in “maximal AGC effect” for the final phase (ANSI, 2003.)

## 2. Method

The goal of this project was to: (1) Test the feasibility of the new standard by: (A) Determining whether a naïve clinician can follow the new wording; (B) Determining the ease and time consumption involved in testing by the new standard; and (2) evaluating the general changes in electroacoustic performance attributable to the new standard. We evaluated fifteen different HI's from seven different manufacturers.

Each HI was set as required by section 5.2.5 of the ANSI s3.22-2003 standard. An AGC HIT test using the Audioscan Verifit© was then performed and adjustments were made where prompted. Without making changes, an input/output test was also performed. This procedure was then repeated for the 1996 standard.

## 3. Results

### 3.1 Electroacoustic Comparison

The Figures below display the raw measurements obtained using the ANSI 1996 and 2003 test sequences. Figure A shows the relationship between measured gain at

reference test position (RTP) and equivalent input noise (EIN) for each standard. Figure 2 shows the relationship between attack and release times measured with each standard. Repeated measures Analysis of Variance (ANOVA) was completed for the 12 measures obtained from each standard, to determine whether significant differences resulted from the changes between the 1996 and 2003 versions. Results indicated that some test sequence measures differed significantly across the two standards  $F(1.88,20.71)=4.59$ ,  $p=0.02$  (degrees of freedom adjusted using Greenhouse-Geisser epsilon for violation of sphericity). Post hoc pair-wise comparisons indicated significant differences for gain at RTP settings, and for attack and release times at certain frequencies (attack at 2000 Hz and 4000 Hz; release at 500 Hz and at 1000 Hz). In general, there was a trend for decreased gain at RTP with the 2003 standard, and for attack and release times to be longer. Some of these results can be explained by a few factors.

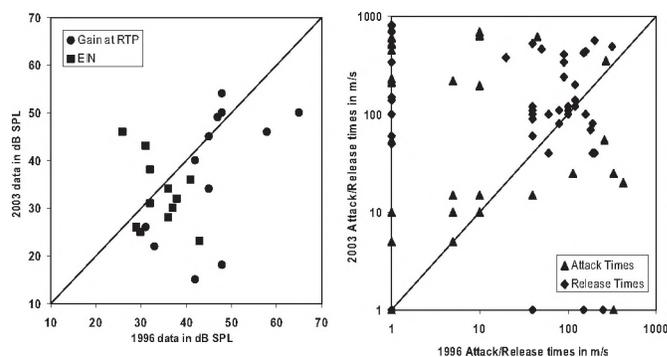


Figure A (Left): Equivalent Input Noise and gain at Reference Test Position with 1996 standard data being plotting against 2003 standard data. Figure B (right): Attack/release times for 1996 standard plotted against data from 2003 standard.

Firstly, many automated tests provided with the manufacturers software for the 1996 standard used a full-on gain setting as the reference test position. Essentially, these tests do not account for gain at RTP. This may account for the lower gain at RTP observed with the 2003 standard, which requires a non-full-on RTP. This is evident in cases that did provide true RTP gain: in these cases, the difference between the 1996 values and the 2003 values were less extreme. No real trend is visible with respect to EIN and no statistical significance was found.

Second, with respect to attack/release times, the old standard allowed the aids to be run at linear, in which case attack/release times were fast, as would be expected for a peak clipping device. The new standard requires use of compression, which, for many devices, resulted in appropriately longer attack and release times because the compression processing was active during ANSI testing.

### 3.2 Specification Sheets

The data sheets provided for any given hearing instrument are a necessity for verifying device performance. Data sheets provided by manufactures should provide complete documentation of HI test settings and expected test results; however, this is not always the case. Figure C below displays the frequency that standard test measures appear in specification data provided by manufacturers. While some measures like EIN appear frequently (14/15 cases) others such as attack/release times appear infrequently (8/15). This is problematic, as information is often required to replicate specification data. However, even when this is not the case (and a software mediated test mode is provided) data is important for quality assurance and device comparison.

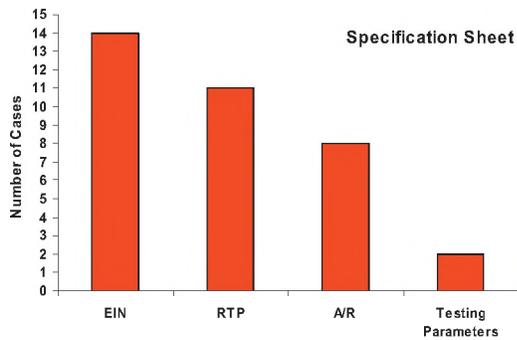


Figure C: Shows the frequency of various standard measures on manufacturer specification sheets

### 3.3 Software Mediated Test Modes

When properly implemented a test mode has the potential to eliminate error and increase reproducibility of results across different labs. Most manufactures (5 of 7 tested) use software mediated test modes. When not used, some (preferably all) test parameters need to be included. In practice, however, test modes are often not well labelled and lack critical stages in required by the standard (i.e. only 1 of the 5 manufactures using test modules included RTP setting.)

### 3.4 Confounds to Feasibility: Setting

Wording and issues of interpretation were a matter for concern. The 2003 standard asks for “minimal AGC effect”, various clinicians could interpret this differently (ANSI, 2003.) For example minimal AGC effect might be

construed to mean either linear or expansion. Another confound to feasibility was the association between HI setting controls. That is to say, a change in one control might inadvertently change another. This means that even when noticed, the clinician may have to choose which control is of greater importance (e.g. let max compression limit gain OR max gain limit compression.) Reproducibility may be compromised when a clinician must make a choice between two options.

Software flexibility proved another difficulty. The effects of having associated controls means one must learn the quirks of each setting program for each aid in order to “coax” the aid into maximal and minimal effects. The result is time consuming, frustrating and likely to be error prone. Some software also lacked control over critical features such as compression ratios, knee points and adaptive features that require adjustment according to the new standard.

### 3.5 Compliance

Compliance is a major issue for concern as it has been lacking in the past as was seen in the case of missing RTP settings, missing values on specification sheets and lack parameters being provided. For example, one manufacturer sent specification data on a new HI supposedly using the 2003 standard. However, it was run in a linear setting, which directly conflicts with the 5.2.5 wording of the test.

## 4. Conclusion

Some electroacoustic changes are present as a result of the change in wording, namely a decrease in gain at RTP and an increase in attack and release times at some frequencies. In addition, we found the feasibility of manually setting the 2003 standard to be low.

Therefore, it is our recommendation that: (A) Manufactures provide specification sheets including all fitting parameters and (B) Manufacturers include a test module that is clearly labelled and includes at least one stage for each of the three phases in the ANSI 2003 standard.

## 5. References

Staab, W. J. (2002) Characteristics and Use of Hearing Aids. In J. Katz (Ed.) *Handbook of Clinical Audiology*. New York: Lippincott Williams & Wilkins.

American National Standards Institute. (1996) Specification on Hearing aid characteristics (ANSI S3.22-1996). New York: Acoustical Society of America

American National Standards Institute. (2003) Specification on Hearing aid characteristics (ANSI S3.22-2003). New York: Acoustical Society of America .

# TUNED MASS DAMPERS FOR SOLDIER FIELD STADIUM GRANDSTAND VIBRATION

Mal P. Sacks<sup>1</sup>, Ramin Behboudi<sup>1</sup>, and John C. Swallow<sup>2</sup>

<sup>1</sup>Tacet Engineering Ltd., 45 Denver Cres., Toronto, Ontario, Canada, M2J 1G6, info@tacet.ca

<sup>2</sup>John Swallow Associates Ltd., 366 Revus Ave., Unit 23, Mississauga, Ontario, Canada, L5G 4S5, jswallow@jsal.ca

## 1. INTRODUCTION

This paper describes the design and testing of Tuned Mass Dampers (TMD's) for the west grandstand of the newly renovated Soldier Field stadium in Chicago as well as a Permanent Vibration Monitoring System. The stadium is to be used as a venue for both sports and entertainment events.

The state-of-the-art for the reduction of crowd-induced vibration of stadium structures has been extended to include very large systems. Twenty-one TMD's were designed and applied to the upper grandstand perimeter covering almost 1000 linear feet with a total sprung weight of 840,000 pounds. The TMD's are providing impressive reduction of vibration levels in the design frequency range of 1.5 Hz to 2.7 Hz and eliminate virtually all resonant peaks in the 2 Hz to 3 Hz frequency band. Vibration reduction factors are typically in the range of 3 to 10. The success of the project was based on a unique configuration of spring systems to achieve flexibility in the design and field tuning.

A purpose-built 64-channel Permanent Vibration monitoring System was also supplied for future TMD testing, on-going maintenance and to facilitate administrative control of vibration.

## 2. PROBLEM DEFINITION

The prior existing stadium was a designated historic building. This required that the existing façade be retained which, in turn, resulted in the top of the west grandstand being cantilevered out 40 feet beyond the column supports.

Finite element analysis of the upper west grandstand indicated vibration levels in excess of established limits due to crowd-induced dynamic loadings. Criteria were established for vibration level limits based on spectator comfort (i.e., structural integrity was not an issue).

Given the unique geometrical constraints of this grandstand, it was not possible to stiffen the structure sufficiently to achieve acceptable vibration levels. Consequently, the addition of a vibration damping system was required and TMD's were chosen. FE analysis indicated that TMD's would reduce vibration levels to below accepted values. Use of TMD's is typically cost effective because the same reduction in vibration amplitude through increased stiffness

would require very much larger structural framing which would not be practical and, in the case of this grandstand, not possible.

## 3. TMD DESIGN

Given the design constraints, the only way to achieve the required mass was in the form of a beam spanning the approximately 40 foot to 50 foot variable distance between rakers. A total of 21 TMD's are located at the top of the west grandstand spanning between the 22 raker girders. The TMD mass consists of a beam constructed as a steel box filled with concrete. The beam lengths vary from approximately 25 feet to 35 feet, are 17 inches deep and have a vertical height calculated to produce a total sprung weight of 40 kips; i.e. the height varies from approximately 4 feet to 6 feet. At each end of the mass beam is a yoke, i.e. a large horizontal U-shaped bracket, which mates with a cantilevered horizontal stub beam attached to the vertical upstand at the tip of the raker.

Support for the mass beam was provided by air springs and steel springs located between the stub beam and yoke. Energy dissipation is provided hydraulic cylinders supplied by Taylor Devices Inc., North Tonawanda, NY and are called Taylor dampers or dashpots. The dashpots are installed between the free end of the stub beam and upper arm of the yoke and can be adjusted to achieve desired TMD damping ratios. Two guide wheels are attached to the free ends of both the upper and lower arms of the yoke. These wheels run on vertical V-tracks attached to the raker upstand. These guide wheels provide stability for the air springs and constrain the mass beam vibration to vertical motion.

The spring system must provide the precise stiffness required and also support the dead load of the TMD while maintaining precise clearances required for TMD operation at tuning frequencies that may be changed in-situ at any time. These multiple requirements are resolved in a unique way by providing two spring systems. Industrial pneumatic springs, installed between the stub beam and upper yoke arm, support the dead load while providing a very low stiffness. The additional stiffness required to achieve the TMD tuning frequency is supplied by small non-load bearing helical steel springs installed between the stub beam and lower yoke arm, in parallel with the pneumatic springs. All TMD's were required to be tunable to all frequencies

between 1.5 Hz and 2.7 Hz in nominal increments of 0.1 Hz. Tests of a full scale prototype TMD demonstrated the flexibility and reliability of the tuning technique.

The 21 TMD's were tuned alternately to 1.5 Hz and 2.7 Hz with damping ratios from 0.06 to 0.13 of critical damping as determined by the FE analysis.

#### 4. TMD TESTING

##### 4.1 TMD Calibration Tests

A replica of TMD #12 was fabricated and assembled in a calibration test rig in order to be closely examined and tuned to various frequencies and damping ratios required in actual field operation. (Fig. 1)

A quick release mechanism was designed and used to displace the TMD and instantaneously release it in order to excite the system to vibrate at its natural frequency. Acceleration traces from accelerometers mounted on the TMD were digitized and processed in order to find system natural frequencies and damping ratios corresponding to various spring combinations and damper settings.

##### 4.2 Stadium Vibration Tests

Vibration tests were conducted on the west grandstand in June 2003 using a vibration shaker purpose built by Anco Engineers, Inc., Boulder, Colorado. The shaker is capable of producing a vertical force with a maximum amplitude of 2000 pounds at any frequency in the range of 1 Hz to 5 Hz. For comparison note that a 2000 pound dynamic force amplitude is equivalent to the force amplitude produced by about a dozen adults performing vigorous aerobic exercise, e.g. synchronized jumping jacks.

##### 4.3 Permanent Vibration Monitoring System

Stadium vibration measurements were obtained using the Permanent Vibration Monitoring System (PVMS) specifically designed for this project. This system consists of 64 accelerometers hard wired to a central data acquisition system. The data acquisition system consists of a 64 channel signal conditioner, 64 channel data acquisition board and a desktop computer that runs the purpose built software developed for this system to measure and display vibration magnitudes and phases for each accelerometer. The PVMS can be remotely operated through the internet.

One accelerometer is located at each end of each TMD (21 x 2) and one at each raker adjacent to the TMD's (22) for a total of 64. The accelerometers are aligned to measure vertical vibration which is consistent with the vertical operation of the TMD's.

The tests demonstrated that the TMD's are performing in accordance with their design criteria; i.e. the TMD's are providing significant reduction of vibration levels in the

design frequency range of 1.5 Hz to 2.7 Hz and eliminate virtually all resonant peaks in the 2 Hz to 3 Hz frequency band. Vibration reduction factors are typically in the range of 3 to 10. (Fig. 2)

#### REFERENCES

Sacks, M.P. and Swallow, J.C. (1993). Tuned Mass Dampers for Towers and Buildings. Proceedings of the ASCE Structures Congress '93, vol. 1, p. 640-645.

Burns, J.G. and McLean, D.P. (2004). Field Goals. Modern Steel Construction, July 2004.

Burns, J.G. and McLean, D.P. (2005). Adaptive Reuse of Soldier Field Chicago. Modern Steel Construction, May 2005.

Reynolds, P., Pavic, A. and Ibrahim, Z. (2003). A Remote Monitoring System for Stadia Dynamics. 21st International Modal Analysis Conference, Florida, USA, February 2003.



Fig. 1. Prototype TMD ready for calibration.

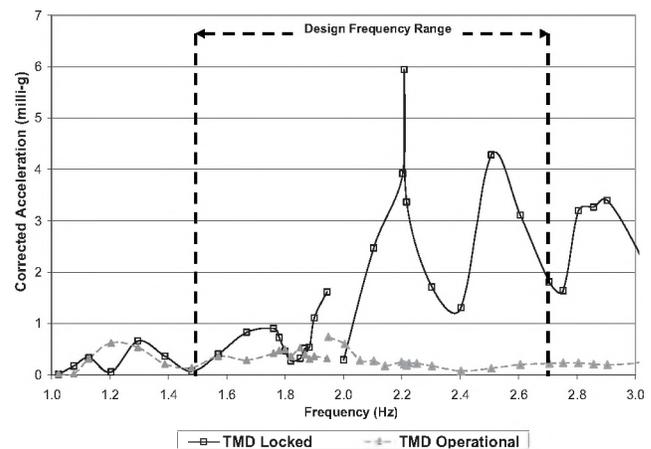


Fig. 2. Sample plot showing effectiveness of the installed TMD's.

# NORMATIVE THRESHOLD LEVELS FOR A CALIBRATED, COMPUTER-ASSISTED VERSION OF THE LING SIX-SOUND TEST

Julianne J. Tenhaaf<sup>1</sup> and Susan D. Scollie<sup>2</sup>

1.) Department of Applied Language Studies, Brock University, St. Catharines, ON. L2S 3A1.

2.) National Centre for Audiology, University of Western Ontario, London, ON. N6G 1H1.

## 1. Introduction

Traditionally, the Ling Six-Sound test has been used as an informal test of hearing ability (Ling, 1989). Consisting of the phonemes /m/, /i/, /a/, /u/, /sh/ and /s/, this test was designed to assess hearing across the speech frequencies (Ling, 1989). However, although convenient, the use of the Ling Six-Sound task in clinical practice is limited by its lack of a calibrated method of delivery, or normative data on detection thresholds.

When testing detection levels in sound field, wide band stimuli such as speech, warble tones, and random noise are commonly used (Arlinger & Jerlvall, 1987). Wide band signals are purported to be more reliable, as they generate a more uniform sound pressure level in the subject's region of sound field (Walker, Dillon & Byrne, 1984).

The present study combined the traditional Ling Six-Sound test with principles of sound field audiometry. The first aim of the study was to design and implement a calibrated, computer-assisted version of the Ling Six-Sound test in order to establish normative data on ten normal-hearing adults. In addition, test-retest reliability of phonemic thresholds was completed.

## 2. Method

The Ling Six phonemic stimuli were used to test auditory detection levels. The phonemes were recorded by a female speaker in a sound-treated booth, and were edited using Goldwave software to be of equal duration and peak level. Calibration on all stimuli was completed using concatenated sound files played at 70 dB HL through one speaker of a five-speaker array. Sound pressure at 0 degrees azimuth was analyzed using a B & K Signal Analyzer Unit type 2035. The resulting SPL values yielded the frequency distribution of each phoneme, as well as overall a-weighted levels in dB SPL.

### 2.1 Participants

Participants in this study were 10 normal hearing adults (3 male, 7 female) between the ages of 21 and 35. Auditory health was verified using tympanometry, otoacoustic emissions, case history questionnaires, and full inter-octave audiograms between 250 and 8000 Hz.

### 2.2 Procedure

Phonemic stimuli were inputted into a computer-assisted threshold bracketing procedure. Subjects were seated in sound field, and responded to sound presentations by selecting "heard it" or "didn't hear it" options on a computer screen. Based on these responses, attenuation levels were altered until threshold was bracketed. Subjects completed two full trials of this procedure for each phoneme. Thresholds were determined by subtracting attenuation levels at threshold from the a-weighted stimulus levels.

## 3. Results

Table 1 displays the mean, standard deviation, and 95% confidence intervals across phonemes. Values ranged from 1.9 to 5.5 dB SPL. Although there is wide variation in obtained values, it is important to note that threshold values of the stimuli /u/ and /sh/ were influenced by one subject who had threshold differences of over 10 dB between trials.

Phoneme	Mean Threshold (dB SPL)	SD	95% C.I. Upper Cutoff (dB SPL)
/m/	21.8	2.2	26.1
/u/	7.9	5.3	18.4
/i/	5.9	5.5	16.6
/a/	0.0	2.5	5.0
/sh/	-3.1	2.3	1.4
/s/	16.7	1.9	20.5

Table 1: Mean thresholds, standard deviation values, and upper cutoff threshold levels for the Ling Six phonemes.

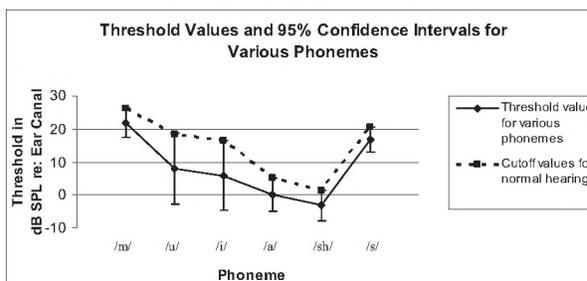


Figure 1: Normal-hearing threshold values and 95% confidence intervals for the Ling Six phonemes.

Using 95% confidence intervals, cutoff values for detection thresholds of the Ling Six phonemes were also established. These values are displayed in Figure 1. All cutoff thresholds were below 26 dB SPL, with some values as low as 1.4 dB SPL

The test-retest reliability of each phonemic stimulus was determined using Intraclass Correlation Coefficients (ICCs) and is shown in Table 2. Results from this analysis revealed a wide range of reliability coefficients with no effect of stimulus bandwidth.

Phoneme	ICC Values	Reliability*
/u/	0.92	Good
/m/	0.81	Good
/sh/	0.73	Moderate
/s/	0.70	Moderate
/i/	0.63	Moderate
/a/	0.47	Poor-Moderate
<b>GRAND AVG</b>	<b>0.71</b>	<b>Moderate-Good</b>

Table 2: Test-retest reliability ICC values for each phoneme.  
\* from Portney & Watkins (2000)

Finally, in order to assess the use of the Ling Six phonemes in hearing impaired subjects, an adult with a high-frequency sloping hearing loss was tested with the calibrated Ling Six-Sound test while aided. Figure 2 compares the aided hearing impaired thresholds with normal hearing averages. Results revealed thresholds above normal hearing limits on only the two highest-frequency phonemes, /sh/ and /s/.

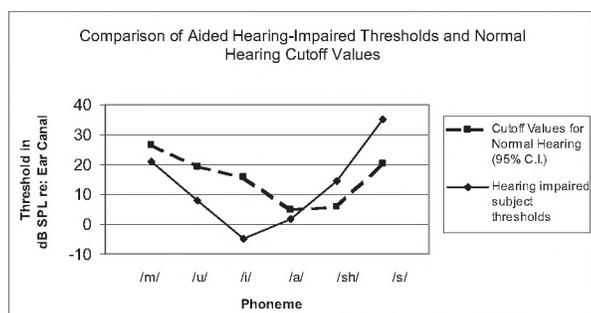


Figure 2: Comparison of aided hearing-impaired thresholds and normal-hearing cutoff values across the Ling Six phonemes.

#### 4. Discussion

While test-retest reliability analyses revealed variability in ICC values between stimuli, ratings of moderate to good for five of six stimuli suggest acceptable reliability for the Ling Six-Sound test. According to Portney and Watkins (2000), ICC values of 0.75 and over are considered to have good test-retest reliability. As the current ICC average value was 0.71, the Ling Six-Sound test appears to fall between moderate and good test-retest reliability.

Comparisons of 95% Confidence Intervals of the Ling Six phonemes and warble tones also indicated acceptable test-retest reliability. The current study found overall averaged 95% Confidence Interval for all phonemes to be +/- 7.1 dB, compared to +/- 15.1 dB for warble tones averaged across the frequency span (Hawkins, Montgomery, Prosek, & Walden, 1987). Thus, it appears that phonemes in sound field may be more reliable than traditional warble tone stimuli for detection tasks.

Furthermore, although the results of a single hearing-impaired subject cannot be generalized to a larger population, this preliminary assessment suggests that the Ling Six-Sound test may be sensitive to the effects of hearing loss, as thresholds were only outside normal limits for the high frequency phonemes in this subject with high frequency loss. Obviously, this requires further study.

Future development of the calibrated, computer-assisted Ling Six-Sound test may involve comparison of audiometric and Ling Six thresholds. Eventual clinical use of a calibrated Ling Six-Sound test may be warranted, as it is a fast and reliable method of testing phoneme detection in sound field.

#### 6. References

- Arlinger, S.D., & Jerlval, L.B. (1987). Reliability in warble-tone sound field audiometry. *Scandinavian Audiology*, 16, 21-27.
- Hawkins, D.B., Montgomery, A.A., Prosek, R.A., & Walden, B.E. (1987). Examination of two issues concerning functional gain measurements. *Journal of Speech and Hearing Disorders*, 52, 56-63.
- Ling, D. (1989). Foundations of spoken language for hearing-impaired children. Washington, D.C.: Alexander Graham Bell Association for the Deaf.
- Portney, L.G., & Watkins, M.P. (2000). Foundations of clinical research: Applications to Practice (2<sup>nd</sup> ed). Upper Saddle River, N.J: Prentice Hall Health.
- Walker, G., Dillon, H., & Byrne, D. (1984). Sound field audiometry: Recommended stimuli and procedures. *Ear and Hearing*, 5(1), 13-21.

# PERSISTENCE OF AUDITORY STREAMING PRESERVES RELEASE FROM MASKING

Kachina Allen<sup>1</sup>, Simon Carlile<sup>1</sup>, David Alais<sup>1</sup> and Karen Froud<sup>2</sup>

<sup>1</sup>Dept. of Physiology, University of Sydney, NSW. Australia 2106

<sup>2</sup>Dept. of Biobehavioural Sciences, Teacher's College, NY, USA, 10027

Correspondence: kachina\_allen@yahoo.com

## 1. INTRODUCTION

It is widely recognised that spatially separating target speech from masking sound results in a listening advantage characterised by improved speech reception thresholds (SRTs) (eg., Freyman, Balakrishnan et al. 2001).

Yost (1991) argued that binaural and spectral cues assisted in the creation of separate auditory objects, allowing the listener to attend to one object and filter out the others. While these cues may be important, other factors are also relevant. Freyman et al (1999), and Driver (1996) showed that unmasking can be elicited using the illusion of spatial separation. This suggests that processes in addition to the salient location cues, may underlie a proportion of improvements in SRT due to spatial unmasking.

Where there are a number of concurrent sound sources, the association of spectral components with individual external sources probably involves firstly a short term grouping process and secondly a streaming of the grouped elements over time (Bregman 1994). In this experiment we aimed to test the contributions that each of these separate processes make to spatial unmasking.

## 2. METHOD

Subjects were seated, facing forward, in a sound attenuated, semi-anechoic chamber (size = 3.5 x 4.6 x 2.4m). Three Tannoy active loudspeakers were placed 1.3m away on the subject's audiovisual horizon.

Subjects had normal hearing, spoke English as a main language and included 4 females and 1 male (mean age 32 yrs). All subjects carried out 100 unrecorded practice trials of the separated, co-located and start separated condition and 150 trials of each condition.

Stimuli were generated and presented using Matlab and a Hammerfall multiface sound card at a sampling rate of 44100 and a volume of 57dB. Stimulus sentences were taken from the Coordinate Response Measure (CRM) corpus (Bolia, Nelson et al. 2000) and consist of an identifier in the first half (the call sign) and two target words (a colour and a number) in the second half. The target talker was identified by the call sign Baron and the subject's task was to identify the two target words in the presence of two masker talkers with different call signs and target words. Subjects entered the target words on a laptop. The signal to noise ratio of the target in relation to the maskers was varied randomly for each trial.

There were 4 conditions.

\* **Co-located:** target and both maskers played from the central speaker.

\* **Separated:** target played from central speaker, 1 masker played from each of symmetrically spaced speakers 30° from the central speaker.

\* **Start Separated:** Target and masker start as in separated condition but collapsed to central speaker after 700 ms (just after the identifying call sign).

\* **Start Co-located:** Target and masker start as in co-located condition but move to locations as in separated condition after 700 ms.

## 3. RESULTS

Subject SRTs were calculated using maximum likelihood generation of cumulative Gaussians at the 50% intelligibility level. Release from masking (RFM) was calculated as

$$\text{RFM}(\text{condition}) = \text{SRT}(\text{condition}) - \text{SRT}(\text{co-located})$$

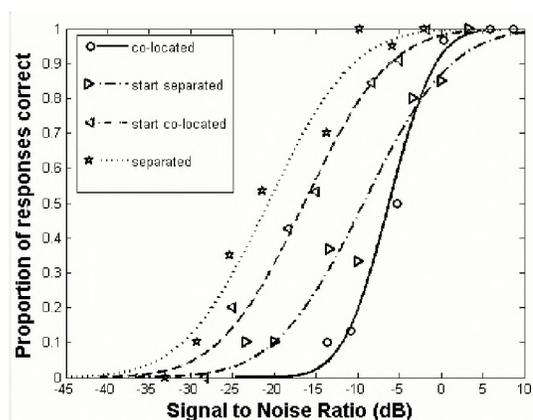


Figure 1 Representative psychometric functions (S4) for each condition. SRT calculated at 50% intelligibility

Bootstrapping was carried out with 500 repeats. Post-hoc t-tests ( $\alpha = 0.05$ ) with Bonferroni corrections were carried out on bootstrapped data.

All subjects showed significant release from masking for all conditions (Table 1). Release from masking in the start co-located condition was much less than in the separated condition.

Subject	Start Separated RFM (dB)	Separated RFM (dB)	Start Co-located RFM (dB)
S1	4.2 ± 1.5	13.6 ± 1.2	15.3 ± 1.2
S2	2.1 ± 0.5	7.0 ± 0.8	5.2 ± 0.8
S3	3.6 ± 1.2	13.9 ± 1.2	13.7 ± 1.3
S4	3.5 ± 0.9	14.5 ± 0.9	11.2 ± 0.8
S5	4.5 ± 0.7	12.0 ± 1.1	9.4 ± 0.7
Mean	3.6	12.2	10.1

Table 1 Release from masking for conditions. RFM given ± standard deviations calculated from bootstrapped SRTs.

In the start co-located condition, subject SRTs were similar to those obtained for the separated condition. While there was a slight trend to reduced SRTs in the start co-located condition, this difference was significant in only one subject. This suggests that subjects may have been employing the strategy of simply listening to the central speaker location after the talkers separated rather than identifying and following the target talker.

To test this, a condition, “start co-located-all move”, was added, in which subjects were forced to follow the target talker after separation. The target and masker co-located at the central speaker, then moved to locations where target and maskers were on different speakers after 700 ms, with the target randomly assigned to one of the three speakers.

Three subjects lost any release from masking in this condition. Where two subjects maintained significant release from masking, these subjects showed reduced performance (SRTs) on lateralised speakers indicating they were favouring the central speaker.

Subject	Start Co-located, all move RFM (dB)
S1	1.1 ± 1.3
S2	2.0 ± 0.8
S3	2.4 ± 1.2
S4	<b>2.3 ± 0.8</b>
S5	<b>5.9 ± 0.8</b>
Mean	2.1

Table 2 Release from masking for conditions. RFM given ± standard deviations calculated from bootstrapped SRTs. Those in bold show a significant ( $\alpha < 0.05$ ) release from masking.

#### 4. DISCUSSION

When target and masking voices are co-located, the listener relies entirely on cues such as voice characteristics (gender, tone, accent etc) to isolate the target talker. When separated, the listener can also use spatial information such as the binaural (inter-aural time and level differences ITDs and ILDs) and spectral cues to identify and maintain stream segregation of the talkers.

The spatial release from masking between the co-located and the 30° symmetrically separated condition (12.2 dB) is higher than that found in much of the literature (eg. 5dB, from Noble and Perrett 2002). The CRM corpus has the same carrier phrase for subject and maskers, with the

same onset time. This synchronisation maximises both the energetic and informational masking. Other studies have used dissimilar discourse for targets and maskers. Spatial release from masking is often more marked with high information masking (eg. Noble and Perrett 2002; Hawley, Litovsky et al. 2004) and this may explain the higher level of unmasking in this study compared with previous studies.

Previous studies have disagreed over the relative contributions of voice cues and location cues in grouping and streaming (Mondor, Zatorre et al. 1998; Darwin and Hukin 1999; Edmonds & Culling 2005). In the current study, all subjects showed a significant level of unmasking in the start-separated condition. This may be due to the initial separation allowing the listener to more easily create and identify the target streams and thus detect voice characteristics of the target talker. While spatial cues disappear as the streams become co-located, these identifying characteristics can be used to continue to attend to the target or to filter out masker talkers.

The differences in unmasking between the separated (12.2 dB) and start-separated (3.6 dB) conditions may indicate the proportion of spatial unmasking which relates to streaming due to location cues and that which relates to separation allowing identification of the target. Further research will be required to discover whether varying the reliability of other cues will affect the reliance on spatial information.

#### REFERENCES

- Bolia, R.S., Nelson, W.T. et al. (2000). A speech corpus for multitaler communications research. *JASA* 107(2): 1065-6.
- Bregman, A.S. (1994) *Auditory Scene Analysis: The Perceptual Organization of sound*, MIT Press, Cambridge.
- Darwin, C. J. & Hukin, R. W. (1999). Auditory objects of attention: the role of interaural time differences. *J Exp Psychol Hum Percept Perform.* 25(3): 617-29.
- Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature* 381:66-68.
- Edmonds, B. A. & Culling, J.F. (2005). The spatial unmasking of speech: evidence for within-channel processing of interaural time delay. *JASA* 117(5): 3069-3078.
- Freyman, R, Balakrishnan, U. et al. (2001). Spatial release from informational masking in speech recognition *JASA* 109(5):2112-22
- Freyman, R., Helfer, K. et al. (1999). The role of perceived spatial separation in the unmasking of speech. *JASA* 106(6): 3578 - 3588.
- Hawley, M., Litovsky, R. et al. (2004). The benefit of binaural hearing in a cocktail party: Effect of location & type of interferer. *JASA* 115(2): 833-843.
- Mondor, T., Zatorre, R. et al. (1998). Constraints on the Selection of Auditory Information. *J Exp Psychol Hum Percept Perform.* 24(1): 66-79.
- Noble, W. & Perrett, S. (2002). Hearing speech against spatially separate competing speech versus competing noise. *Percept Psychophys* 64(8): 1325-36.
- Yost, W. A. (1991). Auditory image perception & analysis: the basis for hearing. *Hear Res.* 56(1-2): 8-18.

# TOWARD BETTER AUTOMATIC SPEECH RECOGNITION

D. O'Shaughnessy, W. Wang, W. Zhu, V. Barreaud, T. Nagarajan, and R. Muralishankar

INRS-EMT, U. of Quebec, 800 de la Gauchetiere west, suite 6900, Montreal Quebec, Canada H5A 1K6 dougo@emt.inrs.ca

## 1. INTRODUCTION

Automatic speech recognition (ASR) performs best when there is a strong correspondence between system training and operating conditions, e.g., when one tests on speech data that is similar in style to that used for training. Mismatch (different environment speakers, or vocabulary) degrades performance. We have developed new model techniques able to adapt to various speech environments without modifying the basic ASR systems: an appropriate feature transformation scheme for the Mel-frequency cepstral coefficients (MFCC), a new speech-processing front-end feature that performs better than the existing MFCC, a log-energy dynamic range normalization technique for ASR in adverse conditions, and a continuous ASR method that exploits the advantages of syllable and phoneme-based sub-word unit models.

## 2. NEW ADAPTATION METHODS

Statistical Data Mapping (SDM) is a new approach for ASR model adaptation. SDM assumes that speech observations are generated by subsets of mutually related random sources. The relationship/model between random sources can be established by a maximum likelihood criterion. Thus speech observations can be mapped through the model to any desired environment without heavy loss of the original information. SDM is a non-linear approach and has the strength to handle non-time-invariant variations, e.g., bandwidth and speech context. SDM has a flexible framework, which can be varied in different applications. We evaluated our algorithms on the Visteon and Superman databases (Scansoft Inc.): a wide-band automobile embedded-speech database, recorded at automobile speeds of low, 50 km/h and 75 km/h, and a narrow-band network speech database. The TIMIT and NTIMIT databases were also used. Finally, we also used the Aurora 2 database, corresponding to TI-DIGITS training data down-sampled to 8 kHz and filtered with a G.712 characteristic. It includes clean and multi-condition training sets. The noisy utterances have SNR from -5 dB to 20 dB. ASR models adapted by our SDM technique have better performance than other models, e.g., non-adapted or adapted by MLLR. The SDM-adapted ASR models have improved performance in noisy environments.

We also dealt with the problem of speech enhancement when only a corrupted speech signal is available for processing. Kalman filtering is known as an effective speech enhancement technique, in which speech signal is usually modeled as autoregressive (AR) model and represented in the state-space domain. Various approaches based on the Kalman filter are presented in the literature. They usually

operate in two steps: first, additive noise and driving process variances and speech model parameters are estimated and second, the speech signal is estimated by using Kalman filtering. Sequential estimators are used for sub-optimal adaptive estimation of the unknown a priori driving process and additive noise statistics simultaneously with the system state. The estimation of time-varying AR signal model is based on weighted recursive least square algorithm with variable forgetting factor. The proposed algorithm provides improved state estimates at little computational expense.

In spectral subtraction for speech enhancement, we subtract the magnitude of the noise spectrum from that of the noisy speech, while keeping the noisy phase. Noise spectrum estimated during weak segments at the start of an utterance assumes that noise is stationary. Our approach is to have a sub-band-based speech detector to separate each signal frame into noise or speech. The estimated noise spectrum is updated at each frame with a forgetting factor, thus dealing with unstable noise. We utilize the noise suppression gain information in another way for noise reduction. Spectral valleys are usually more disturbed by noise than spectral peaks. So we emphasize spectral peaks (e.g., formants). We introduced a frequency masking filtering algorithm in the standard MFCC feature extraction algorithm.

## 3. FEATURE SPEAKER ADAPTATION WITH A SMALL FOOTPRINT

Most ASR systems are trained on a great variety of speakers and are thus called "Speaker Independent" (SI) systems. Yet training on data specific to the current user (Speaker dependent) gives better performance. We focused on the use of Speaker Normalization methods. They are useful for ASR that lacks memory and processing capacity such as embedded engines. Indeed, normalization amounts to a set of transformations of speech data (at different steps of the front-end). The light computation required by the normalization process can be done on-line. Moreover, the number of parameters describing this set of transformations is small and their set can be considered as a small user footprint. A speaker S uses a desktop dictation application (namely, Scansoft's Dragon Naturally Speaking, DNS). This application derives a speaker profile (the set of the normalization parameters), which would be used again to normalize S's speech when he uses DNS. One solution is to enroll portable speaker profiles on fully available dictation data with a desktop ASR and then use them for embedded recognition of command words. This architecture raises the problem of intra-speaker variability since the user-speaker used a speaking style that varies from enrollment to test.

We did a study of speaker profile portability. This work has been conducted on MREC, the DNS engine. The objective was to see if our profiles are efficient with small adaptation sets and if they can be used on command test data. Using profiles with MREC Engine, these results showed that speaker profiles could be ported from one task (sentence) to another (command). Their efficiency is reduced but the gain is still significant. This suggests that our framework will produce significant improvement in recognition.

#### 4. SEGMENTATION INTO SYLLABLES

Many public-domain speech corpora have only an orthographic transcription. In order to use these corpora to build syllable-level models and ASR systems, an efficient automatic speech segmentation algorithm is required. We made a new explicit segmentation algorithm that uses the orthographic transcription, which allows knowing the number of syllable segments in a speech signal a priori. Although the short-term energy (STE) function contains useful information about syllable boundaries, it cannot be directly used to do segmentation due to significant local energy fluctuations. We use an Auto-Regressive model-based algorithm to smooth the STE function using the knowledge of the number of syllable segments required. If the error is more than 40 ms, those segment points are considered as erroneous boundaries. Experiments on the TIMIT corpus show that the error in segmentation is at most 40 ms for 87.84% of the syllable segments.

We propose a new front-end feature, warped discrete cosine transform cepstrum (WDCTC), which achieves better vowel recognition and speaker-identification performances than the MFCCs. The WDCTC has a better performance than the MFCC in a 5-vowel recognition and speaker-identification task. Feature transformation aims to maximize the desired source of information for a speech signal in the front-end feature and to minimize undesired sources such as noise, speaker variability and other speech signals. Frequency domain interpretations of the feature transformation provide useful and interesting information by highlighting the importance of different frequency regions for a particular vowel. Tests are conducted with unknown vowel samples extracted from continuous speech in TIMIT. The average vowel-recognition performance with the feature transformation scheme was 71.2%.

#### 5. ENERGY RANGE NORMALIZATION

Cepstral mean normalization (CMN) and Cepstral Variance Normalization (CVN) are simple noise robust post-feature processing techniques. In CMN, the log-energy feature (or C0) is treated in the same way as other cepstral coefficients. Compared with cepstral coefficients, the log-energy feature has quite different characteristics. We try here to find a more effective way to remove the effects of additive noise for the log-energy feature. We propose a log-energy dynamic range normalization (ERN) method to

minimize mismatch between training and testing data. Comparing with that of clean speech, characteristics of the log-energy feature sequences of noisy speech are: elevated minimum value, and valleys that are buried by additive noise energy, while peaks are not affected as much. The larger difference for valleys leads to a mismatch between the clean and noisy speech. Obviously, mean normalization is not an optimized solution. We suggest an algorithm to scale the log-energy feature sequence of clean speech, in which we lift valleys while we keep peaks unchanged. The dynamic range of log-energy feature sequences of an utterance is normalized to a target dynamic range.

The proposed algorithm was evaluated on the Aurora 2.0 digit recognition task. The proposed log-energy dynamic range normalization algorithm had overall about a 31.56% relative performance improvement when systems were trained on a clean speech training set. This method does not require any prior knowledge of noise and level. It is effective to improve the performance of speech recognition for eight different noise conditions at various SNR levels. Histogram equalization and non-linear transformation techniques have been reported to achieve 51.81% or 49.71% performance gains in clean-condition training in the literature, but the proposed method can be easily combined with variance normalization to get a better result (54.21%). In addition, the proposed method does not need to estimate feature density functions or to direct transformation functions. It only needs a very small extra computation load.

#### 6. CONCLUSION

Feature transformation interpreted as a filter bank throws light on the relative significance of different frequency bands for a particular vowel and helps in understanding vowels from the frequency domain perspective. Inclusion of WDCTC as a front-end feature in the state-of-art ASR systems may improve the overall performance.

#### REFERENCES

- T. Nagarajan, D. O'Shaughnessy, Explicit segmentation of speech based on frequency-domain AR modeling, paper 1299, INTERSPEECH, Lisbon, Sept. 2005.
- V. Barreaud and D. O'Shaughnessy, Experiments on Speaker Profile Portability, # 1832, INTERSPEECH-05
- R. Muralishankar, A. Sangwan, D. O'Shaughnessy, Warped Discrete Cosine Transf Cepstrum: EUSIPCO, 2005.
- W. Zhu and D. O'Shaughnessy, Log-energy dynamic range normalization for robust speech recognition, ICASSP, Philadelphia, PA, SP-L.11.2, March 2005.
- W. Wang, D. O'Shaughnessy, Robust ASR Model Adaptation by Feature-Based Statistical Data Mapping, INTERSPEECH 2004, Oct. 2004, Jeju, Korea.

**ACKNOWLEDGMENTS :** We thank NSERC and PromptQuebec for their funding support.

# OUTPUT-BASED SOUND QUALITY EVALUATION USING STATISTICAL MODEL METHOD

Guo Chen, Vijay Parsa

National Centre for Audiology, Faculties of Health Sciences & Engineering, Elborn College,  
University of Western Ontario, Ontario, Canada, N6G 1H1. Email: sguo@nca.uwo.ca, parasa@nca.uwo.ca

## 1. INTRODUCTION

Good objective speech quality measurement is highly desirable and useful in the field of speech communication. The accuracy of the objective measurement is determined by correlating it with known subjective measurement of speech quality, such as the mean opinion score (MOS) defined in [1]. A majority of previously reported objective methods are based on input/output comparisons, which estimate the speech quality by measuring the "distortion" between the input and output signals, and mapping the distortion values to the predicted quality metric. But in some applications, a reference signal might not be available for an input/output comparison, e.g., evaluation of pathological voice. In such cases, an attractive alternative is to assess speech quality using only the output signal, i.e., output-based evaluation. Currently, a handful of output-based evaluation techniques have been reported in the literature. In [2], a vocal tract modeling based non-intrusive evaluation technique was presented to monitor telecommunication network distortion. In [3], an auditory non-intrusive quality estimation (ANIQUE) model, which is based on human auditory and articulation systems, was reported. In [4-6], the authors investigated output-based techniques by measuring perceptual spectral density distribution and by exploiting neuro-fuzzy techniques. In [7], a state-of-the-art non-intrusive evaluation standard was recommended by the ITU, which applies a complicated evaluation structure and myriad parameters. Clearly, it can be seen that all of the current output-based techniques reported in the literature are intended to provide an optimal continuous mapping from physical parameters to subjective speech quality scores (such as the MOS scores). But this is inconsistent with actual subjective listening-opinion tests [1]. As stated in [1], the subjective MOS listening-opinion test is a Category Judgment method (Page 3 in [1]), which can be thought of as a process of speech quality pattern classification. Therefore, it is intuitive to develop an evaluation method by employing statistical pattern classification to imitate subjects' behavior in a subjective listening-opinion test. In this paper, we propose a novel output-based evaluation technique using statistical modeling, which is in line with a subjective listening-opinion test. The proposed method was motivated by two facts: (i) the subjects in a listening test do in fact perform a quality pattern classification process in terms of their perception of the speech signal,

and (ii) in the process of a listening test, the subjects listen only to the output speech signal of the system under test, not hear the original input signal (i.e., the reference signal).

## 2. THE PROPOSED METHOD

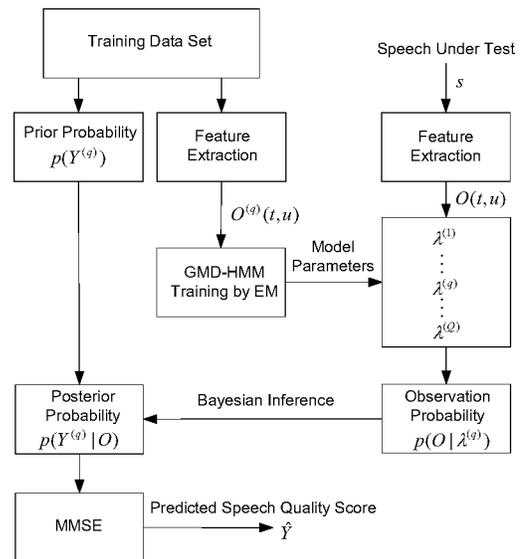


Fig.1. The Block Diagram of The Proposed Method

The block diagram of the proposed method is shown in Fig.1. We assume that subjective quality scores are classified into  $Q$  categories, denoted by  $Y^{(q)}$ ,  $q=1,2,\dots,Q$ . The observation features, representing the  $q$ -th quality category, denoted by  $O^{(q)}$ , are extracted by the pitch power density analysis. The observation features  $O^{(q)}$  are characterized by GMD-HMMs  $\lambda^{(q)}$ ,  $q=1,\dots,Q$ . The aim of the proposed method is to minimize the mean squared error between the predicted MOS values  $\hat{Y}$  and the true MOS values  $Y$  given observations, i.e., let  $E[(\hat{Y} - Y)^2 | O]$  approaches to minimum. The solution is a conditional expectation as  $\hat{Y} = E[Y | O] = \int Y p(Y | O) dY$ . Since there are  $Q$  quality categories, The conditional expectation becomes  $\hat{Y} = E[Y | O] = \sum_{q=1}^Q Y^{(q)} p(Y^{(q)} | O)$ . The conditional probability  $p(Y^{(q)} | O)$  can be calculated by

$$p(Y^{(q)} | O) = \sum_{i=1}^Q p(Y^{(q)} | \lambda^{(i)}) p(\lambda^{(i)} | O)$$

Using Bayesian inference, the posterior probability can be reformulated by the prior probability,  $p(\lambda^{(i)})$ , and the likelihood,  $p(O | \lambda^{(i)})$ , as below,

$$p(Y^{(q)} | O) = \sum_{i=1}^Q p(Y^{(q)} | \lambda^{(i)}) \frac{p(O | \lambda^{(i)}) p(\lambda^{(i)})}{\sum_{m=1}^Q p(O | \lambda^{(m)}) p(\lambda^{(m)})} \quad (1)$$

Since we assumed that the speech signal of the  $q$ -th quality category is generated by the corresponding statistical model  $i$ , we have  $p(Y^{(q)} | \lambda^{(i)}) = 1$  if  $i = q$  and  $p(Y^{(q)} | \lambda^{(i)}) = 0$  if  $i \neq q$ . Therefore, we may rewrite Eq.(1) as

$$p(Y^{(q)} | O) = \frac{p(O | \lambda^{(q)}) p(\lambda^{(q)})}{\sum_{m=1}^Q p(O | \lambda^{(m)}) p(\lambda^{(m)})} \quad (2)$$

In Eq.(2), the prior probability can be estimated from the training data, i.e., the probability of each existing quality category. The likelihood is estimated by a GMD-HMM.

In the proposed method, we employ the spread pitch power density values as the observation features for representing different speech quality categories. The feature extraction process consists of the following steps: (1) Voice Activity Detection; (2) Level normalization & IRS filtering; (3) Time-frequency mapping; (4) Outer and middle ear transfer function; (5) Transform to pitch (Bark) domain; (6) Adding internal noise; (7) Frequency domain spreading; Finally, the logarithmic values of the spread pitch power density are chosen as the observation features for each speech frame, denoted by  $O(t, u)$ .

The joint conditional observation densities  $p(O | \lambda^{(q)})$  are estimated by GMD-HMMs. There are  $Q$  such GMD-HMMs  $\lambda^{(q)}$ , where  $\lambda^{(q)} = \{\pi^{(q)}, A^{(q)}, B^{(q)}\}$  denotes the set parameters of the  $q$ -th  $N$ -state GMD-HMM used to characterize the  $q$ -th speech quality category, of which  $\pi^{(q)}$  represents the initial state distribution,  $A^{(q)}$  is the transition probability matrix, and  $B^{(q)}$  is the parameter vector composed of mixture parameters  $B_i^{(q)} = \{\omega_{ik}^{(q)}, \mu_{ik}^{(q)}, \sigma_{ik}^{(q)}\}$  for state  $i$ . The training of the parameters of the GMD-HMMs is performed with the *expectation-maximization* (EM) algorithm. A starting point is determined by clustering the training data with the  $K$ -means algorithm in our study.

### 3. EXPERIMENTAL RESULTS

The experimental data consist of seven subjective quality MOS databases obtained in two listening opinion tests as described in Experiment One and Three of the ITU-T P-Series Supplement 23[8]. We combined these

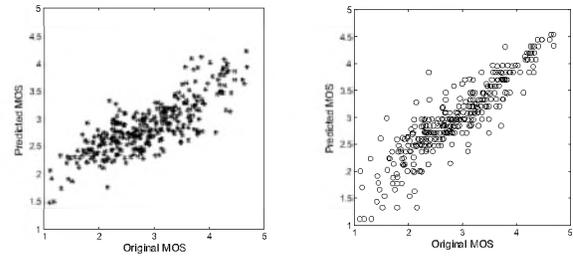


Fig.2. Predicted MOSs. \*: P.563; o: Proposed Method

seven databases into a global database, giving a total of 1328 MOS scores. Three-quarters of the global database was used for training, while the remaining was used as a validation data set. The correlation coefficient ( $\rho$ ) and standard error of estimate ( $\varepsilon$ ) were used to evaluate the performance. The results are shown in Fig.2. From the results, it can be observed that the correlation of the proposed method attained 0.9012 with a standard error of 0.3390 across the whole databases. This compares favorably with the P.563, which provides a correlation and standard error of 0.8422 and 0.4493, respectively. The results demonstrated the ability of the proposed method to predict speech quality scores without any reference signal.

### 4. REFERENCES

- [1] *Methods for subjective determination of transmission quality*, ITU-T Recommendation P.800, 1996.
- [2] P.Gray, M.P.Hollier, and R.E.Massara, "Non-intrusive speech quality assessment using vocal-tract models," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 147, no. 6, pp. 493–501, 2000.
- [3] D.-S. Kim, "Anique: An auditory model for single-ended speech quality estimation," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 4, 2005.
- [4] G. Chen and V. Parsa, "Output-based speech quality evaluation by measuring perceptual spectral density distribution," *IEE Electronics Letters*, vol. 40, no. 12, pp. 783–784, 2004.
- [5] G. Chen and V. Parsa, "Non-intrusive speech quality evaluation using an adaptive neuro-fuzzy inference system," *IEEE Signal Processing Letters*, vol. 12, no. 5, pp. 403–406, 2005.
- [6] G. Chen and V. Parsa, "Bayesian model based non-intrusive speech quality evaluation," in *Proc. IEEE of ICASSP 2005*, pp. 385–388, 2005.
- [7] *Single ended method for objective speech quality assessment in narrow-band telephony applicaitons*, ITU-T Recommendation P.563, 2004.
- [8] *ITU-T coded-speech database*, ITU-T P-series Supplement 23, 1998.

### 5. ACKNOWLEDGEMENTS

We gratefully acknowledge the financial support by the Oticon Foundation, Denmark, the Ontario Rehabilitation Technology Consortium, Canada, and the NSERC, Canada.

# AN EXPLORATORY STUDY ONTO THE ROLE OF PHONEMIC CATEGORIES IN SPEECH PERCEPTION

Yun Zheng<sup>1</sup>, Christian Giguère<sup>2</sup>

<sup>1</sup>Institute of Cognitive Science, Carleton University, 1125 Colonel By Drive, Ottawa, ON, K1S 5B6 zyun98@yahoo.com

<sup>2</sup>School of Rehabilitation Sciences, University of Ottawa, 451 Smyth Road, Ottawa, ON, K1H 8M5 cgiguere@uottawa.ca

## 1. INTRODUCTION

Stevens (1989; 2002) proposed that phonemic categories play a fundamental role in speech perception. According to this view, there is a process for specialized speech perception mechanisms to extract prototypical categories from variable acoustic inputs, and discard the variability to unmask the underlying phonetic or featural prototypes. One important support for this model is said to come from studies on categorical perception. For example, Cienfuegos et al. (1999) presented evidence that the deficit in categorizing acoustic properties of speech sounds may hinder the analysis of phonemes, and lead to the difficulty in language processing in schizophrenia.

Nonetheless, the assumed role of categorical perception in speech perception has been increasingly challenged in recent years. One major objection is related to the definition of categorical perception itself: true categorical perception is supposed to occur when discrimination is completely predicted by categorization, but this has never been demonstrated experimentally (McMurray et al., 2003). Schouten et al. (2003) pointed out that most discrimination tasks in categorical perception studies are subject to bias effects: the nature of the task usually compels subjects to use a labeling strategy. Subjects may prefer one response type to another, or they may refer the stimuli to an internal criterion of their own. When response bias is eliminated, categorical perception seems to fail. Furthermore, some studies have shown that listeners preserve sensitivity to fine-grained acoustic variation within phonetic categories in patterns of lexical activation (McMurry et al., 2002). Though small acoustic differences in Voice Onset Time (VOT) have minimal effects on phoneme identification, they have significant effects on lexical access.

Apart from the problems with categorical perception itself, the exact relationship between categorical perception and speech perception should be further studied. One intuitive inference from Stevens' model would be that speech perception is realized on the basis of the perception for individual phonemic categories. We will test such an inference in the present study.

## 2. METHOD

Three English words (IN, SIN, TIN) were read by a young native female English speaker, and digitally recorded using the VisiPitch IV system (Kay Elemetrics Corp.). The stimuli were further edited using the software Cool Edit Pro. Five categorically well-defined tokens (S = /S/ from SIN, T = /T/ from TIN, A1 = IN, A2 = /IN/ from SIN, and A3 = /IN/ from TIN) were extracted from the original stimuli. Six synthetic words were then constructed from these five tokens (A2, A3, SA1, SA3, TA1, TA2).

A 6x2 factorial design was adopted for the experiments. Factor one (Words) has six levels (the six synthetic words), and Factor two (Hearing Conditions) has two levels (normal-hearing and hearing-impaired). In Part one, the perception of the six synthetic words with the three natural words was compared under normal hearing conditions. An ABX task format was employed to reduce bias effects. The first two words (A, B) were a pair consisting of one natural word and one synthetic word, and X was always the natural word. The stimuli were delivered through high quality earphones. Subjects were prompted to report whether X was the same word as either A or B after each presentation. The serial position at which the natural word occurred within each task series was also controlled. Thus a total of 36 (6 x 3 x 2) task series were presented.

In Part two, the three natural words and the six synthetic words were processed through a hearing loss simulation system (University of Ottawa, 2004) to simulate the aging-induced hearing loss corresponding to a 65 year-old man (normalized data according to ISO 1999). The hearing thresholds were around 20dB for lower frequencies (125 Hz to 1 kHz), but significantly increased to about 80dB at higher frequencies (6 kHz to 8 kHz). The same ABX task format was employed in Part two, the 36 series of stimuli differed only from those in Part one due to the simulated hearing loss. As a result, some phonemes (e.g., /S/) of certain words were acoustically suppressed.

Preliminary data is available for 4 university students, all native English speakers with normal hearing. Presentation of the stimuli was controlled by a customized computer program (Visual C++). The order of presenting the 36 task series within each part was randomly determined by the

computer program. The order of completing Part one and two was also counterbalanced between subjects.

### 3. RESULTS

In most cases, native English speakers perceived the synthetic words to be distinguishable from the natural words, even when a synthetic word and a natural word consisted of tokens from the same categories (e.g., SIN and SA1). Specifically, they had little difficulty in perceiving the difference between IN and all synthetic words ( $F(5,84) = 1.260$ ,  $p = 0.289$ ), also the difference between SIN and all synthetic words ( $F(5,84) = 0.60$ ,  $p = 0.70$ ). But subjects seemed to be confused by TIN and some of the synthetic words ( $F(5,84) = 2.576$ ,  $p = 0.032$ ).

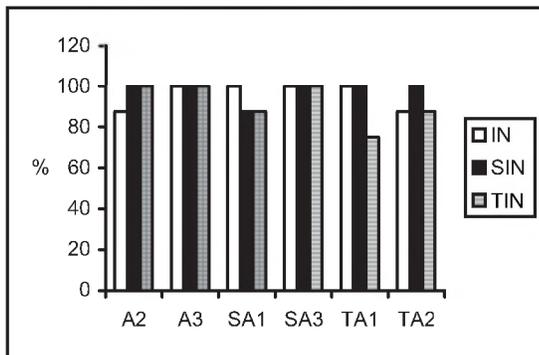


Fig. 1. The percentage of correctly identifying a natural word from a word-pair consisting of a natural word and a synthetic word under normal hearing conditions

When hearing loss was involved, the acoustic cue for the /S/ was largely suppressed. Yet this suppression had little effect on the performance of identifying the natural words from synthetic words, either for IN ( $F(1,84) = 0.70$ ,  $p = 0.405$ ), or SIN ( $F(1, 84) = 0.333$ ,  $p = 0.565$ ), or TIN ( $F(1,84) = 0$ ,  $p = 1.0$ ).

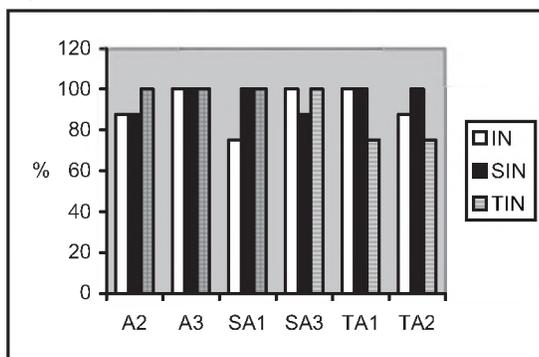


Fig. 2. The percentage of correctly identifying a natural word from a word-pair consisting of a natural word and a synthetic word when hearing loss was involved

No interactions between synthetic words and hearing conditions were significant. Interestingly, the suppression of /S/ by hearing loss made the acoustic waveforms for SA1 – IN, A2 – SIN hardly distinguishable. However, it resulted in

little perceptual disturbance for the subjects. A paired t test showed that the effect by this acoustic suppression was not significant,  $t(15) = 1.861$ , and  $p = 0.083$ .

### 4. DISCUSSION

Critics to Stevens' model had raised doubts about some problems associated with categorical perception. But the exact relationship between categorical perception and speech perception had barely been discussed. In the present study we tested whether speech perception is realized on the basis of the perception for individual phonemic categories in speech sounds.

In our experiments, all the six synthetic words consisted of well-defined phonemic categories (extracted from natural words). However, the synthetic words were perceived differently from the natural words, even when both consisted of tokens from the same categories (e.g., SIN and SA1). With the suppression of certain categories due to simulated hearing loss, a synthetic word and a natural word (though consisting of different categories) might also become acoustically similar (e.g., SIN and A2), yet subjects had little difficulty perceiving the difference. These results seemed to suggest that speech perception does not depend on the perception for single categories. This is consistent with findings from studies on co-articulation effects, whereby speech sounds are not produced as isolated gestures, but are superimposed in a context-dependent fashion (see Fowler, 2005). Thus, speech perception is more likely based on perception for co-articulated categories rather than single phonemic categories. For example, when information about one category was unavailable, listeners could rely on information from another category that co-articulated with it. This may have implications for hearing aid processing.

### REFERENCES

Cienfuegos, A., March, L., Shelley, A., and Javitt, D.C. (1999), Impaired categorical perception of synthetic speech sounds in schizophrenia, *Biol. Psychiatry*, 45, 82 – 88.  
 Fowler, C.A. (2005), Parsing coarticulated speech in perception: effects of coarticulation resistance, *Journal of Phonetics*, 33, 199–213.  
 McMurray, B., Tanenhaus, M. K., Aslin, R.N., Spivey, M.J. (2003), Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access, *Journal of Psycholinguistic Research*, 32(1), 77 - 97.  
 McMurray, B., Tanenhaus, M.K., Aslin, R.N. (2002), Gradient effects of within-category phonetic variation on lexical access, *Cognition*, 86, B33 – B42.  
 Schouten B., Gerrits, E., Hessen, van Hessen, A. (2003), The end of categorical perception as we know it, *Speech Communication*, 41, 71 – 80.  
 Stevens, K.N. (2002), Toward a model for lexical access based on acoustic landmarks and distinctive features, *Journal of the Acoustical Society of America*, 111(4), 1872 – 1891.  
 Stevens, K.N. (1989), On the Quantal Nature of Speech. *Journal of Phonetics*, 17(3), 3 – 45.

# MEASURING THE DYNAMIC PERFORMANCE OF VIDEO CONFERENCING SYSTEMS

David I. Havelock<sup>1</sup>, and David Green<sup>2</sup>

<sup>1</sup>Institute for Microstructural Science, NRC, 1200 Montreal Rd., Ottawa ON K1A 0R6, david.havelock@nrc.ca

<sup>2</sup>Institute for Information Technology NRC, 1200 Montreal Rd., Ottawa ON K1A 0R6, dave.green@nrc.ca

## 1. INTRODUCTION

The effectiveness of a video conferencing system is determined by diverse factors such as video and audio quality, interactive system control mechanisms, and the ability of the system to track conversation. System performance is often based on the subjective assessment of criteria that lack the rigor necessary to obtain definitive comparisons between different systems and strategies.

An automated video conferencing system emulating the skill of an expert videographer in capturing the visual and audio dynamics of a presentation is of great value for video conferencing [1], tele-presentations [2], meeting archiving [3], and possibly surveillance.

## 2. OBJECTIVE

We are concerned here with tracking talkers visually and acoustically during a conversation and obtaining performance metrics that can be conveniently measured, compared between systems, and related to subjective judgments of performance. As an initial step, a metric is defined for the ability of a video conferencing system to follow the talker transitions in conversation with multiple participants. The methodologies are applicable to a generic multimodal system [4] and are implemented in a system independent manner to determine the audio and video switching delays independently, and are demonstrated using a prototype system that combines independent audio and video talker localization [1].

## 3. METHOD AND APPARATUS

Pre-recorded ‘conversations’ are used with two talkers (or noise sources) speaking alternately. Audible and visible markers are inserted into the scene at a variable delay  $t_d$  from the onset of each transition (event), as in Fig. 1. If the steering mechanisms are able to switch fast enough to display the marker then the event is recorded as a ‘hit’, otherwise it is a ‘miss’. The probability of a hit as a function of marker delay gives the system latency probability distribution function (PDF).

The visible marker is an LED flash and the audible marker is a pair of 1 kHz tone bursts 15 ms long. The markers are synchronized as in Fig. 1 and are presented through speakers and an LED light system as shown in Fig. 2. Two sets of speakers and lights are positioned at the same distance (~90

cm) from the array and camera center, with 120° angular separation. The LED flash duration  $t_d$  is varied and if the system latency is less than  $t_d$  then the observer sees the flash (a ‘hit’). The first of a pair of tone bursts occurs at the event time, and the other occurs  $t_d$  later. The response of the microphone array is loudest when steered at the source so, if the system steers to the source fast enough, the second pulse is heard to be louder than the first and a ‘hit’ is declared for that event.

During a synthesized conversation scenario, the observer indicates when a hit or miss occur and the responses are collected on-line. About 30 runs, each with a range of 40 delays, were used to estimate each PDF. Multiple audio pulses and video flashes were used at each event to enhance the statistical significance of the observations. Both noise and recorded speech are used as audio sources. For speech, randomized phrases from a list of Harvard Sentences were used, with a male and a young female voice.

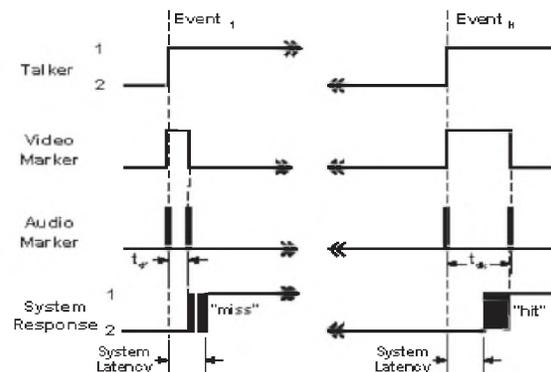


Figure 1 Timing Sequence. On the left (a ‘miss’), the video and audio markers are finished before the system (lower trace) steers to the active talker. On the right (a ‘hit’), the system responds fast enough to catch part of the video indicator and the second of the two audio markers.

## 4. OBSERVATIONS

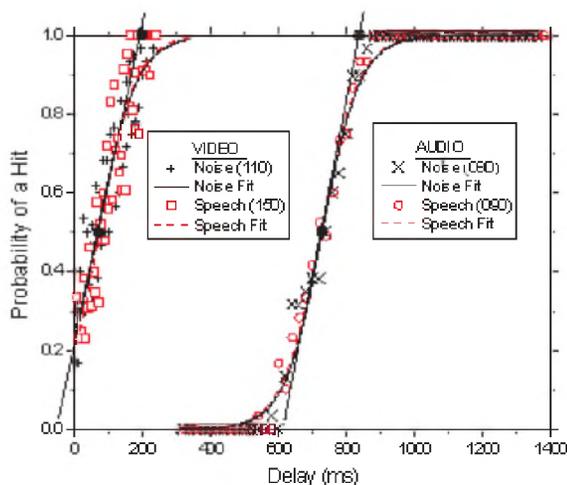
Figure 3 shows the video and audio latency PDF with a sigmoid function overlaid. The video and audio PDF have similar slope and spread but are offset by about 600 ms. The PDF data for noise and speech are similar in both cases. The video PDF curve has a positive zero-delay intercept because the system video and data buffers allow non-memoryless processing.

The mean system response delay is estimated by the midpoint

of the sigmoid and the maximum delay by projecting the tangent at the mid point to the top axis, as shown by the large solid dots in Fig. 3. The proposed metric comprises these four values (mean and maximum delay for audio and video system response).



**Figure 2 System setup.** The 16-microphone array, panoramic camera (center of array), speaker on tripod, and LED light system are shown. Another speaker and light system are located 120° about the array.



**Figure 3 System response probability distribution function (PDF) indicating the probability that the system will respond at least as fast as the ordinal delay. Estimated PDF values (points) from measurements are fit to a sigmoid function (lines). The metric for system response comprises the delays at the four solid dots. (Numbers in the legend are dataset codes.)**

## 5. CONCLUSIONS AND DISCUSSION

The metric comprising the mean and maximum systems delays has been measured in a system-independent manner and is suitable for intercomparisons. The relevance of the metric for subjective performance is not yet confirmed.

The sigmoid function is a reasonable data model when many factors contribute to the delay, but its asymptotic nature is

problematic for estimating the maximum delay accurately. The tangent method is a convenient approximation.

The shift of the audio PDF relative to the video is due mainly to the message packet queue used for steering the microphone array in the prototype Panocam system.

System processing is synchronous with the video frame rate (about 7 fps), yet the sloping PDF indicates variations in response time in excess of 250 ms. There are three identifiable contributors to this variation; the timing of the event relative to the start of the video frame, the asynchrony of the audio buffer and video frame, and jitter in the messaging system between the audio and video subsystems.

The video frame buffer memory allows a PDF with a non-zero value at zero delay. The video display, however, is delayed by the frame buffer interval and contributes to the system display latency.

The choice of sound source (noise, or male or female speech) has been observed to make minor differences in the PDF. Our scenario assumes a conversation is underway, with talkers identified, but the PDF may vary at the introduction to a conversation. A half-second of silence was inserted between events to model ‘polite’ conversation; other scenarios, such as interrupting or overlapping conversation, may give different results. The video and audio markers are thought to have little impact on the system operation.

## REFERENCES

- [1] M. Fiala, D. Green, and G. Roth, (2004) “A panoramic video and acoustic beamforming sensor for videoconferencing,” 3<sup>rd</sup> IEEE Int’l Workshop on Haptic, Audio, and Visual Environments and Their Applications, 2004:47-52.
- [2] Yong Rui, Anoop Gupta, and Jonathan Grudin (2003), “Videography for Telepresentations,” Proc. SIGCHI Conf. On Human Factors in Computing Systems 2003, Ft. Lauderdale, FL, 5-10 Apr. 2003, Vol. 5, Issue 1, pp. 457-464.
- [3] Eric Meurville and David Leroux (2004), “Collection and annotation of meeting room data,” m4- Multimodal Meeting Manager Report D1.2, 2 April 2004. [Website on 4 August 2005: <http://www.m4project.org/publicDelivs/D1-2.pdf>]
- [4] D. Lo, R.a. Goubran, R.M. Dansereau, “Multimodal talker localization in video conferencing environments,” 3<sup>rd</sup> IEEE Int’l Workshop on Haptic, Audio, and Visual Environments and Their Applications, 2004:195-200.

## ACKNOWLEDGEMENTS

We wish to thank Mark Fiala for help with the video subsystem.

# MISTAKES, ERRORS AND UNCERTAINTIES

Alberto Behar

IBBME, University of Toronto, Toronto, ON, [alberto.behar@utoronto.ca](mailto:alberto.behar@utoronto.ca)

## 1.0 INTRODUCTION

Most acoustical reports and even papers do not include margin of errors when reporting measurement results. Also, in many occasions assessments are made with only one measurement being taken. This paper examines the possible causes for the lack of statistical analysis in acoustics and makes some recommendations on how to provide more meaningful results.

## 2.0 SOME (BASIC) STATISTICS

Mistake is the common way of indicating that something that could have been prevented went wrong during a measurement. Examples could be: setting the SLM in dBA while performing frequency analysis, or measuring the noise from a source, while not paying attention to the background noise.

Error, on the other hand, is something inherent to the measurement. It cannot be prevented, but only minimized. Examples could be instrument error, errors due to reflecting surfaces in the vicinity, changing environmental conditions etc. Statistically, an error is the difference between the measured value and the “true” value that is not known.

Any book on statistics will make the difference between a systematic and a random error. The first shifts the mean value of the measurements in one or other direction. We can still have accuracy (good repetition of the results) but not precision (the mean value is shifted with respect to the “true” value). This is a typical error often caused by improper calibration of a measuring instrument. Random error, (probably caused by using instruments with poor accuracy) affects the variability of the measured values. We do get precision, but a large variability.

The most common way of calculating a “true” value is by repeating several times the measurement. Then, this “true” value is calculated as the mean value of those results. However, since the number of measurements is limited, there is a need to calculate the probability that this is really the “true” value. That is when the term of standard deviation comes into play. Its physical meaning is that there is 68% of probability that the “true” value is within the range of the mean value of the measurement results +/- one standard deviation. If two standard deviations are taken, then the probability increases to 92%. If the results of measurement is reported as 85 dB +/- 2.5 dB this will indicate that:

- a) We are dealing with a phenomenon where individual observations are normally distributed (as most physical phenomena are),

- b) The mean value of the measurements was 85 dBA, and
- c) There is a 68% probability that the “true” value will be any value between 82.5 and 87.5 dBA.

Does it mean then that this “true” value could be 82.5 or 87.5 dBA. The answer is definitely “YES”! Obviously, this opens the door to speculations, but this is the nature of the beast: we really cannot ascertain the mean value is the “true” one.

## 3.0 STATISTICS AND ACOUSTICS

Now, what is the situation in acoustics? Measurements are, or should be done following normalized procedures and instruments that fulfill requirements set in standards. In a typical sound level measurement, the accuracy of the calibrator is better than 0.1 dB, same as this of the SLM. However, the accuracy of a field sound level measurement is in the order of +/- 2 dBA. In the case of an audiometric test the situation is similar: the accuracies of the calibrator as well as of the audiometer are a fraction of a dB. However, the accuracy of the measurement itself is between 3 and 5 dB for the air conduction and between 4 and 8 for the bone conduction. The same can be applied to most field measurements, where the accuracy of the instrumentation greatly exceed this of the measurement itself. This is due mainly to the interface instrument – measured object (reflections, background noise, to name some of them).

Surprisingly, even instrument brochures and manuals seldom report the accuracy of the stated characteristics (sensitivity, frequency response, etc).

It should be mentioned, that not many standards include procedures for determining the accuracy of the measurements. In that respect, for instance, the CSA Z107.56 [1] states that:

- a) Measurements should be repeated until the differences fall within a certain range, and
- b) When measuring noise exposure of groups, a certain procedure should be followed to ensure a given degree of accuracy.

Some ISO standards (e.g., ISO 1996-1 [2], adopted by CSA) requires that the test report include “...uncertainty of the results and methods used to take them into account”. Interesting enough, the ISO 1996-2 [3] standard, also adopted by CSA, does not include any requirements regarding accuracy or uncertainty. However, the latest Draft ISO Standard (DIS) 1996-2 [4] does require uncertainties in prediction and in measurements to be reported.

Recently, the International Organization for Standardization (ISO) has issued a directive, that no standard should be produced unless it contains a section regarding the determination of the uncertainty of the measurement.

The issue of uncertainty has become so important that last June in Le Mans (France), the INCE/Europe and CIDB had organized the symposium MANAGING UNCERTAINTIES IN NOISE MEASUREMENT AND PREDICTION including over 125 papers.

From the above it follows that when the uncertainty is an issue, the measurement has to be repeated several times. This, of course, can be a problem in a field situation, where cost and time are an issue.

In summary, this author consider that uncertainty should be determined in the following situations:

- a) Laboratory measurements, and
- b) When the results are close to limits set by regulating authorities
- c) When compliance with a standard has to be demonstrated.

Also, when performing measurements, they should be repeated to confirm the results, especially when

- a) Results do not make sense
- b) Where personal perception is not backed-up by the measurement results (e.g., pure tones or low frequency rumbling)

Obviously, there is no need to underline the importance of the use of common sense applied to any particular situation.

#### 4.0 REFERENCES

1. Z107.56-94: Procedures for the Measurement of Occupational Noise Exposure. CSA 1994, R2004.
2. Subsection 8.2.1. in CAN/CSA-ISO 1996-1:05. Acoustics – Description and measurement of environmental noise – Part 1: Basic quantities and assessment procedures
3. CAN/CSA-ISO 1996-2:05. Acoustics – Description and measurement of environmental noise – Part 2: Acquisition of data pertinent to land use.
4. Section 12 in ISO/DIS 1996-2. Acoustics – Description, measurement and assessment of environmental noise – Part 2: Determination of environmental noise levels.

## Accuracy & Low Cost— Scantek Delivers Sound & Vibration Instruments

Scantek offers two integrating sound level meters and real-time octave-band analyzers from CESVA that make measurements quickly and conveniently. The easy to use SC-30 and SC-160 offer a single dynamic range of 100dB, eliminating any need for range adjustments. They simultaneously measure all the functions with frequency weightings A, C and Z. Other features include a large back-lit screen for graphical and numerical representation and a large internal memory.

The SC-30 is a Type 1 precision analyzer while the SC-160 Type 2 analyzer offers the added advantages of lower cost and NC analysis for real-time measurement of equipment and room noise. Prices starting under \$2,000, including software.

Scantek delivers more than just equipment. We provide solutions to today's complex noise and vibration problems with unlimited technical support by acoustical engineers that understand the complex measurement industry.

# Scantek

Sound and Vibration  
Instrumentation & Engineering

7060 Oakland Mills Road • Suite L  
Columbia, MD 21046  
800•224•3813  
www.scantekinc.com  
info@scantekinc.com

### SC-30 / SC-160 Applications

- Machinery Noise
- Community Noise
- HVAC Acoustics
- Room Acoustics & Reverb Time
- Noise Criteria (NC) (SC-160)

## CESVA

**We sell, rent, service, and calibrate sound and vibration instruments.**



# TIME-FREQUENCY SIGNAL DECOMPOSITIONS FOR AUDIO AND SPEECH PROCESSING

Karthikeyan Umamathy<sup>1</sup> and Sridhar Krishnan<sup>2</sup>

Dept. of Elec. and Comp. Engg., <sup>1</sup>The University of Western Ontario, London, ON N6A 5B8 kumapath@uwo.ca

<sup>2</sup>Ryerson University, Toronto, ON M5B 2K3 krishnan@ee.ryerson.ca

## 1. INTRODUCTION

Efficient analysis and processing of audio signals would lead to a better utilization of computer vision and machine learning technologies in automating audio related applications. Audio and speech are highly non-stationary signals with a time-varying spectrum. It is difficult to analyze them using simple signal processing tools. Most of the existing techniques segment the audio signals and assume the signal to be quasi stationary within the short periods and apply stationary signal processing tools. However these approaches suffer from fixed time-frequency resolution and cannot accurately model the time varying characteristics of the audio signals. An adaptive joint time-frequency (TF) approach would be the best way to analyze audio signals.

The two well-known time-frequency approaches are based on 1. Signal decomposition, and 2. Bilinear TF distributions (also known as Cohen's class) [1]. In order to perform an objective analysis and to extract useful parametric information, the TF decomposition based approach would be ideal. Hence, the proposed methodology uses an adaptive TF transform (ATFT) based on the matching pursuit (MP) algorithm with Gaussian TF functions [2].

Majority of the audio and speech applications perform some combination of the following operations: (i) Compression and (ii) Feature extraction (for pattern recognition), and (iii) Denoising. The output specification for each of the above operations is grossly different. This paper is an attempt to present the proposed adaptive TF technique as a unified methodology (block diagram shown in Fig. 1) in addressing all the above operations on audio and speech signals. The paper is organized as follows: Section 2 covers the methodology comprising the subsections of ATFT, audio compression and audio & speech classification. Section 3 covers the time-width versus frequency band mappings. Discussion and Conclusions are given in Section 4.

## 2. METHODOLOGY

### 2.1 Adaptive time-frequency transformation

The core of the proposed methodology lies in the adaptive TF transformation based on the MP algorithm. MP, when used with a dictionary of TF functions yields an

adaptive time-frequency transformation [2]. In MP any signal  $x(t)$  is decomposed into a linear combination of  $K$  TF functions selected from a redundant dictionary of TF functions  $g(t)$  as given by

$$x(t) = \sum_{n=0}^{K-1} \frac{a_n}{\sqrt{s_n}} g\left(\frac{t-p_n}{s_n}\right) \exp\{j(2\pi f_n t + \phi_n)\} \text{ where } a_n \text{ is}$$

the expansion coefficient, the scale factor  $s_n$  also called as octave or time-width parameter is used to control the width of the window function, and the parameter  $p_n$  controls the temporal placement. The parameters  $f_n$  and  $\phi_n$  are the frequency and phase of the exponential function respectively. The signal is projected over a redundant dictionary of TF functions with all possible combinations of scaling, translations and modulations. At each iteration, the best-correlated TF functions to the local signal structures are selected from the dictionary. The remaining signal called the residue is further decomposed in the same way subdividing them into TF functions.

### 2.2 Audio compression

In the audio compression application, we first modeled the audio signal (5s segments at 44.1k/s) with  $K$  number of TF functions that either captures 99.5 % of the signal energy or to a maximum of  $K=10,000$ . The TF decomposition parameters ( $a_n, s_n, p_n, f_n, \& \phi_n$ ) were analyzed and a novel TF psychoacoustics model was applied to discard the perceptually irrelevant TF functions. The perceptually filtered  $K'$  TF functions were then quantized using 54 bits/ TF function. A curve fitting technique was used on the energy  $a_n$  parameter to significantly further reduce the total number of bits. An audio database containing 8 stereo signals of 20s long were used for testing. Compression ratios as high as 40 were achieved with an average SDG (subjective difference grade) of -1.1. The proposed technique performed exceedingly well for classical type of music compared to the existing techniques.

### 2.3 Audio and Speech Classification

A database of 170 audio signals containing 6 groups (rock, classical, country, folk, jazz and pop) of music

---

Thanks to NSERC and MICRONET for funding this research work.

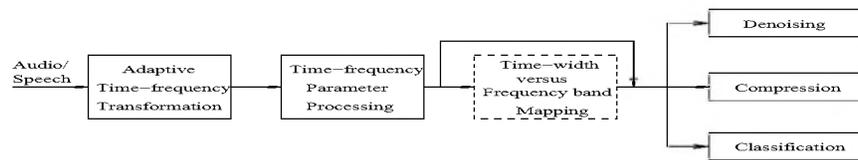


Fig. 1: Block diagram of the proposed methodology

signals were decomposed using the ATFT algorithm. The octave parameter  $s_n$  demonstrated high discrimination between the classes of signals. The octave distribution was computed over three frequency bands and used as features. A set of 42 features were extracted and used in classifying the six music groups. A classification accuracy of 97.6 % was achieved [3].

A database of 212 speech signals containing 51 normal and 161 pathological signals were used in the study. The distribution of octave parameter, the energy  $a_n$  capture rate and center frequency  $f_n$  of the TF functions demonstrated high discriminatory behavior between the normal and pathological signals. 3 features were derived and used for classifying the normal and pathological signals. A classification accuracy of 93.4% was achieved [4]. A sample energy  $a_n$  capture curve and octave  $s_n$  distribution are shown in Figs. 2(a) and 2(b). ATFT's inherent capability of denoising [2] helped in both the above discussed compression and classification applications to remove the insignificant signal components.

### 3. TIME-WIDTH VS FREQUENCY BAND MAPPING (TWFB)

Often when TF visualization of a signal is required, a TFD is constructed. Classical TFDs are non parametric and lacks the flexibility to relate visual patterns with a model or decomposition parameters. A more ideal choice of visualization would be that which preserves the parametric benefits of the decomposition. The idea is to generate a TF subspace mapping using the decomposition parameters that would serve as a (1) good parametric visualization tool, (2) an organized subspace mapping and (3) flexible TF subspace extractor that could address various denoising/ source separation problems. Of the five TF decomposition parameters discussed,  $(a_n, s_n \& f_n)$  would be more appropriate to generate a subspace mapping. We form a 3D visualization by accumulating the energy for every combination (tile) of  $(s_n \& f_n)$ . The step size of  $(s_n \& f_n)$  decides the resolution of the visualization. Figs. 2(c) to 2(f) show the spectrogram & TWFB mappings of a clean speech signal and a noisy speech signal (AWGN at 5dB). One can clearly see from Fig. 2(f), the tiles corresponding to the noise standing out separately. TWFB map is sensitive to signal structures; hence we could easily filter out or separate signal components that differ in structural content. This ability to segregate structural signal components can also be used to characterize different classes of (audio) signals.

### 4. DISCUSSION AND CONCLUSIONS

A unified adaptive TF decomposition based methodology for processing audio and speech signals was presented. The proposed non-stationary signal analysis tool performed well with diverse operations related to audio and speech applications. The proposed methodology is computationally expensive but considering the rapid hardware advancements this should not pose a problem in near future. A novel TWFB mapping was introduced which demonstrates high potential to form as a versatile parametric visualization/pattern recognition tool.

### REFERENCES

[1] Cohen, L (1989). Time-frequency distributions – a review. Proceedings of the IEEE., vol 77, no 7, pp 941-981.  
 [2] Mallat, S . G. and Zhang. Z (1993). Matching pursuits with time-frequency dictionaries. IEEE Transactions on signal processing, vol 41, no 12, pp 3397-3415.  
 [3] Umapathy, K., Krishnan, S. and Jimaa, S. (2005). Multigroup classification of audio signals using time-frequency parameters. IEEE Transactions on Multimedia, vol 7, no 2, pp 308-315.  
 [4] Umapathy, K., Krishnan, S., Parsa, V. and Jamieson, D. G. (2005) Discrimination of pathological voices using a time-frequency approach. IEEE Transactions on Biomedical Engineering, vol 52, no 3, pp 421-430

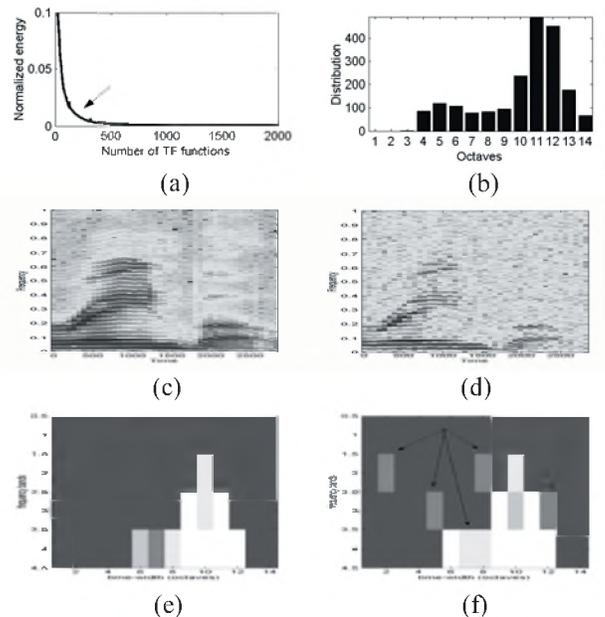


Fig. 2: (a) A sample energy capture curve, (b) A sample octave distribution, (c) Spectrogram of clean speech, (d) Spectrogram of a noisy speech (5dB), (e) TWFB map of a clean speech, (f) TWFB map of a noisy speech signal (5dB) showing distinctly the tiles corresponding to noise signal components.

# HEARING AID MODELING WITH ADAPTIVE NONLINEAR FILTERS

Ranga Rodrigo, Zhenhe Chen, Vijay Parsa and Jagath Samarabandu

Dept. of Electrical and Computer Engg., The University of Western Ontario, London, ON, N6A 5B9, Canada  
brodrigo@uwo.ca, zchen56@uwo.ca, jagath@uwo.ca, parsa@nca.uwo.ca

## 1. INTRODUCTION

In this paper, the performance of nonlinear adaptive filters in modeling devices with mild and moderately strong nonlinearities such as hearing aids is investigated. The Volterra series based nonlinear adaptive filter structure and bilinear adaptive filter structure are compared with their linear counterparts of equal number of tap weights. Experiments for two types of hearing aids using human voice sequences are carried out in a system identification fashion. The number of tap weights used range from 200 to 800 and the adaptation is done in normalized least mean squares (NLMS) form (Haykin 2002). Few graphical results of system identification with white noise input is shown for both Volterra and bilinear cases. The use of such filters for hearing aid modeling is then discussed. The results on the improvement attainable using nonlinear filters are presented. They verify the existence of nonlinearities in hearing aids and that nonlinear adaptive filters are better suited for hearing aid modeling compared to their linear counterparts with equal number of filter coefficients. The signal to noise ratio improvement in modeling, shows improvements in the range of 5% to 50%. The problem dependent but important aspect of the careful selection of the number of linear weights (coefficients of linear terms) and the number of nonlinear coefficients (coefficients of the nonlinear terms) is discussed.

## 2. NONLINEAR FILTERS

While linear filters are predominantly used in practice, there are many applications where nonlinear filters are required. Truncated Volterra series expansion based filters and bilinear filters relate the input signal of a (nonlinear) system to the output using a polynomial model of nonlinearity (Mathews 1991). Bilinear filters reduce the large number of coefficients required in Volterra form by using a recursive nonlinear difference equation. Bilinear filters have been successfully used to reduce the saturation effects of active noise cancellation systems (Kuo & We 2005). Hearing aids are known to be nonlinear devices. Therefore, the use of nonlinear adaptive filters for the modeling of hearing aids should produce better results than linear adaptive filters.

### 2.1 Volterra Series Expansion

Let  $x(n)$  and  $y(n)$  represent the input and output signals respectively. The Volterra series expansion for  $y(n)$  using

$x(n)$  is given by

$$y(n) = h_0 + \sum_{i_1=0}^{\infty} h_1(i_1)x(n-i_1) + \dots + \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \dots \sum_{i_p=0}^{\infty} (h_p(i_1, i_2, \dots, i_p) x(n-i_1)x(n-i_2)\dots x(n-i_p)) \quad (1)$$

where  $h_p(i_1, i_2, \dots, i_p)x(n-i_1)x(n-i_2)$  is called the  $p^{th}$  order Volterra kernel of the system.

### 2.2 Bilinear Filters

Volterra series representation requires a large number of coefficients to handle higher order nonlinear systems. An alternative representation is the recursive nonlinear difference equation including the simple model characterizing bilinear filters with input-output relationship,

$$y(n) = \sum_{i=1}^{N-1} c_i y(n-i) + \sum_{i=0}^{N-1} \sum_{j=1}^{N-1} b_{i,j} y(n-j)x(n-i) + \sum_{i=0}^{N-1} a_i x(n-i) \quad (2)$$

## 3. HEARING AID MODELING

The performance of hearing aids are usually modeled using human voice sequences rather than white noise. A reference signal (an audio recording) is fed to the hearing aid while the output of the hearing aid is recorded. Consequently, The reference signal and the hearing aid output should be aligned over time to compensate for the delays involved. This can be done by simply locating the maximum of the cross-correlation. For the purpose of comparison, the nonlinear adaptive filter and a linear filter of equal length should be compared. If the number of samples considered (and therefore the number of linear coefficients) is  $N$ , the total number of coefficients is  $N + N^2$ . For example if a Volterra filter is compared with a linear filter with 240 taps, the Volterra filter can comprise of 15 linear coefficients and  $15^2 = 225$  nonlinear coefficients. The small number of linear coefficients and the comparatively very large number of nonlinear coefficients are not desirable to model mild and slightly strong nonlinear filters. Therefore the coupling between the number of linear coefficients and the number of nonlinear coefficients is removed and filters with length  $M_{li} + M_{nl}$  are compared. Here  $M_{nl}$ , the number of nonlinear coefficients, is a square number and  $M_{nl} < M_{li}$ , which is the number of linear coefficients. In the implementation, the latest  $M = M_{li} + M_{nl}$  samples,  $\mathbf{u}_{li}$ ,  $(1 \times M)$  are selected

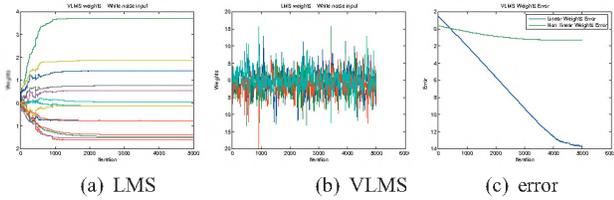


Fig. 1. LMS weights, VLMS weights, VLMS weights error.

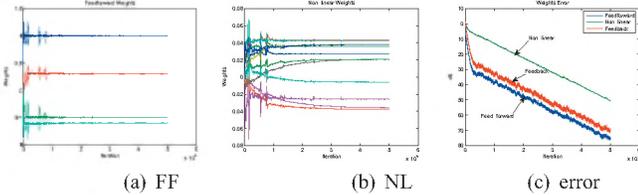


Fig. 2. Bilinear filter performance. FF: feedforward, NL: nonlinear weights

and used with the linear filter. The first  $M_{li}$  samples of  $\mathbf{u}_i$  and the mutual products of the first  $M_{nl}^{1/2}$  of  $\mathbf{u}_i$  are used to compose the input to the nonlinear filter  $\mathbf{u}_{nl}$ ,  $(1 \times M)$ .

#### 4. EXPERIMENTS AND RESULTS

##### 4.1 Volterra and Bilinear Filters

A second order Volterra filter was implemented for system identification of a nonlinear system with the following Volterra kernels. Resulting plots are shown in Figure 1.

$$h_1 = [-0.78 \quad -1.48 \quad -1.39 \quad 0.04]$$

$$h_2 = \begin{bmatrix} 0.54 & 3.72 & 1.86 & -0.76 & 3.72 & -1.62 & 0.76 & -0.12 \\ 1.86 & 0.76 & 1.41 & -1.52 & -0.76 & -0.12 & -1.52 & -0.13 \end{bmatrix}$$

A bilinear filter was tested for a system identification problem for a system with the parameters  $a$ , the feedforward weights,  $b$ , the nonlinear weights, and  $c$ , the feedback weights. Resulting plots are shown in Figure 2.

$$a = [1.0 \quad -0.5 \quad 0.3 \quad -0.6]^T$$

$$b = \begin{bmatrix} 0.027 & 0.036 & 0.043 & 0.042 & -0.036 & 0.038 \\ 0.021 & 0.038 & 0.020 & -0.038 & -0.006 & -0.026 \end{bmatrix}$$

$$c = [0.2 \quad -0.3 \quad 0.4]^T$$

Here,  $b$  is shown in a matrix form instead of the vector for convenience. The entries in  $b$  were deliberately selected to be an order of magnitude smaller to ensure convergence.

##### 4.2 Hearing Aid Modeling Results

The ability of linear and nonlinear filters of modeling hearing aids was compared for two different types of hearing aids: A with a mild nonlinearity and B with a moderately strong nonlinearity. The comparison measure used was the percentage SNR improvement gained by using the nonlinear filter as

$$\text{Improvement} = \frac{(S/N)_{nl} - (S/N)_{li}}{[(S/N)_{nl} + (S/N)_{li}]/2} \times 100\% \quad (3)$$

where  $(S/N)_{nl}$  is the SNR between the desired signal and the modeling error due to the use of the nonlinear NLMS

Table 1. Volterra based hearing aid modeling, SNR improvement for A

Length $M$	$M_{li}$	$M_{nl}$	$\mu$	Run	Improvement
249	249	49	0.8	50000	4.6324%
249	249	49	0.2	50000	9.0757%
500	500	100	0.2	50000	12.3211%

Table 2. Volterra based hearing aid modeling, SNR improvement for B

Length $M$	$M_{li}$	$M_{nl}$	$\mu$	Run	Improvement
249	249	49	0.8	50000	7.8183%
249	249	49	0.2	50000	6.1692%
500	500	100	0.2	50000	7.8698%
500	500	400	0.2	50000	55.7738%

filter and  $(S/N)_{li}$  is the same ratio for the linear NLMS filter.

Table 1 and Table 2 summarize the results for modeling hearing aid A and modeling hearing aid B, respectively, using Volterra series based non linear filters.  $M$  is the number of filter coefficients and  $M_{li}$  and  $M_{nl}$  are the linear and non-linear coefficients respectively. The convergence parameters  $\mu$  and the run length (iterations) are also listed.

Hearing aid modelling was done using bilinear filters and compared with the linear counterparts of equal length. Table 3 shows the results for modeling hearing aid B.  $M_{ff}$ ,  $M_{fb}$  and  $M_{nl}$  are the number of feed forward coefficients, number of feedback coefficients and the number of nonlinear coefficients respectively. The run length was selected to be 10000. Bilinear filters being not as exact as Volterra filters in approximating the nonlinearity could be a reason for weaker performance.

#### 5. DISCUSSION

Nonlinear adaptive filters perform significantly better than their linear counterparts, in approximating mild and moderately strong nonlinear hearing aids. As the nonlinearity of the device becomes stronger, a larger number of nonlinear coefficients should be used. The SNR improvement in modeling, showed improvements in the range of 5% to 50%. The match between the number of linear and the number of nonlinear coefficients should be selected to suit the system in question. The SNR seems to vary depending on the above coefficient numbers and the values of convergence parameters selected. These effects should further be investigated.

#### REFERENCES

- Haykin, S. (2002), *Adaptive Filter Theory*, 4 edn, Prentice Hall.
- Kuo, S. M. & We, H.-T. (2005), 'Nonlinear adaptive bilinear filters for active noise control systems', *IEEE Transactions on Circuits and Systems I* **52**(3), 617–624.
- Mathews, V. J. (1991), 'Adaptive polynomial filters', *IEEE Signal Processing Magazine* pp. 10–26.

Table 3. Bilinear hearing aid modeling, SNR improvement for A

$M$	$M_{ff}$	$M_{fb}$	$M_{nl}$	$\mu_a$	$\mu_b$	$\mu_c$	Impr.
205	103	102	0	0.2	0.005	0.01	0.88553%
205	100	99	6	0.2	0.005	0.01	
1379	690	689	0	0.2	0.005	0.01	2.4973%
1379	500	499	380	0.2	0.005	0.01	7.9348%
809	405	404	0	0.2	0.005	0.01	
809	300	299	210	0.2	0.005	0.01	

**NEW**  
4-channel system



# Pimento Rev. 5.

Simple, Powerful, Complete, Portable

## Simple

- 3-click operation - Acquire. Analyze. Report.
- Intuitive test procedure editor
- 101 pre-configured tests

## Powerful

- Supports industry standard tests: ISO, ANSI, EC, etc.
- Formula editor or customized analytics
- Expandable to 24 channels

## Complete

- LMS Software or acoustic and modal testing
- Exclusive LMS PolyMAX curve fitting utility
- Wide range of microphones, accelerometers, signal conditioners and accessories available

## Portable

- Compact, elegant and rugged design
- Battery or externally powered
- 58mm x 340mm x 295mm, 4.9kg

**Dalimar Instruments Inc.**

Tel. Montreal: 450-424-0033 / Tel. Toronto: 905-707-9000 / Tel. West: 403-288-4416  
E-mail : [info@dalimar.ca](mailto:info@dalimar.ca) / Web Site : [www.dalimar.ca](http://www.dalimar.ca)

pimento 

  
Larson Davis  
A PCB Group Co.

© 2004 Larson Davis, Inc. All trademarks are property of their respective owners.

# architecture | music | acoustics

International Cross-Disciplinary Conference Toronto, Canada June 8-10, 2006

In conjunction with the **soundaXis** festival of music, acoustics, and architecture, **Ryerson University** Department of Architectural Science is pleased to announce an international conference dedicated to a study of the same themes.

Among the critical fault-lines within architectural practice and discourse is that which privileges sight, conceiving of architecture as primarily a visual art form. Despite the multi-sensorial, embodied nature of our being in space, architectural discourse has been largely silent where senses other than the visual are impacted.

This conference proposes to move outside the visual paradigm to investigate the relationships between architecture, music, and acoustics. These intermeshed relationships have tended to one extreme or the other: either sound is understood by architects instrumentally, as an element to be controlled if not eliminated, or music is understood by architects as a metaphorical structure needing to be translated to visual terms before becoming available to architecture.

Framing the conference is the premise that Architecture, Music, and Acoustics have real relationships not requiring translation between visual and sonic terms, and not limited by the instrumental. The conference will involve both historical and critical studies of these relationships, focusing on the strategies and techniques used by architects in dealing with sound and with ideas borrowed from music. The conference will encourage an active mode of engagement with sound in architectural design.

**Session Themes:** Acoustic Ecology | Situated Sonic Practices | Spaces for Performance | Intersections of Music and Architecture | The Poetics of Closure | Sound in Architectural Education | The Architectural Representation of Sound .

**Call for Papers:** Abstracts due October 31, 2005. More information can be found on the conference website at <http://www.ryerson.ca/amaconf/> or by contacting the conference organizers at [amaconf@ryerson.ca](mailto:amaconf@ryerson.ca).

**soundaXis**  
Architecture • Music • Acoustics

RYERSON UNIVERSITY

## Why Purchase from a Single Manufacturer... ...When You Can Have the Best in the Industry From a Single Supplier?

Scantek is the company to call when you want the most comprehensive assortment of acoustical and vibration equipment. As a major distributor of the industry's finest instrumentation, we have the right equipment at the right price, saving you time and money. We are also your source for instrument rental, loaner equipment, product service, technical support, consulting, and precision calibration services.

Scantek delivers more than just equipment. Since 1985, we have been providing solutions to today's complex noise and vibration problems with unlimited technical support by acoustical engineers that understand the complex measurement industry.

### Suppliers of Instruments and Software:

- Norsonic
- RION
- CESVA
- DataKustik (Cadna & Bastian)
- KCF Technologies
- BSWA
- Castle Group
- Metra
- RTA Technologies
- G.R.A.S.

**Scantek**  
Sound and Vibration  
Instrumentation and Engineering

### Applications:

- Building Acoustics & Vibration
- Occupational Noise and Vibration
- Environmental and Community Noise Measurement
- Sound Power Testing
- Calibration
- Acoustical Laboratory Testing
- Loudspeaker Characterization
- Transportation Noise
- Mechanical Systems (HVAC) Acoustics

Scantek, Inc. • 7060 Oakland Mills Road • Suite L • Columbia, MD 21046 • 800•224•3813 • [www.scantekinc.com](http://www.scantekinc.com)

# ADVANCED SIGNAL PROCESSING TECHNOLOGIES FOR INTELLIGENT HEARING AIDS

Henry Luo<sup>1</sup> and Horst Arndt<sup>2</sup>

1 - Unitron Hearing Limited; 2 - Consultant

## Introduction

Digital signal processing technology has experienced a rapid evolution over the past three decades. DSP is central to a wide range of applications: speech coding and decoding for telecommunications, array signal processing for radar and sonar, spectral analysis, multimedia processing, as well as in devices used in daily life and home entertainment.

Approximately 10 years ago DSP began to be applied in digital hearing aids due to the development of practical low voltage, low power silicon technology [1][2][3][4].

## DSP Platform for DSP in Hearing Aids

DSP based amplifiers for hearing aids must be small, usually less than 20 mm<sup>2</sup> footprint and less than 3 mm thick. Operating voltage has to be between 1.0 and 1.3V. Current consumption has to be less than 1.0 mA, so that a standard battery can drive a hearing aid for 100 hours or more.

The first practical digital hearing aids were based on low power Application Specific Integrated Circuit (ASIC) DSP cores. Their designs were optimized for efficient execution of FFTs, iFFT, FIR filtering, and applications of gain. Long lead times are required to develop the ASIC hardware. It is difficult to alter its functionality because ASIC systems have a hard-wired, or fixed, functionality. But they do allow parametric changes via a programming interface to optimize gain, frequency response and compression characteristics of the device to best suit an individual's hearing loss.

Software controlled, or open, DSP systems with low voltage operation and low power consumption have also been developed. These can run various digital signal processing strategies, implemented in the form of algorithms. The time required for development and algorithm upgrades is significantly shorter than with ASIC hardware. [5].

The current DSPs in hearing aids integrate algorithm controlled cores with dedicated ASIC blocks. These blocks handle the repetitive computations of specific signal processing tasks

in the most efficient way thereby meeting the computational requirements within the power consumption constraints imposed by the hearing aid environment. The advantages of flexibility for design and algorithm upgrades, is maintained.

## DSP Algorithms

Adaptive filtering, noise reduction, beam-forming, signal enhancement and hearing loss compensation can be implemented in the time domain, the frequency domain or both. Therefore, time-to-frequency and frequency-to-time transforms such as fast Fourier Transforms or filter banks are usually required to convert the signal between the time-domain and the frequency-domain.

### A. Noise Reduction

The signal-to-noise ratio (SNR) is an important factor in hearing and speech perception. Whereas normal hearing persons have good sentence understanding with an SNR as low as -5 dB, hearing impaired persons may need +5 dB SNR to do as well.

Signal amplitude modulation has been widely used to differentiate between a desired or useful signal (speech or music) and noise in a hearing aid. Amplitude modulation of a speech or music signal in a frequency band is usually higher than amplitude modulation of stationary signals like pure tones or fan noise, office noise, traffic noise and multi-speaker babble.

Amplitude modulation does not on its own produce consistently reliable signal detection. Signal detection and noise reduction have been improved by using additional information such as temporal information and timing information about the signal and noise in combination with amplitude modulation [6]. Hearing aid NR algorithms now routinely reduce noise by up to 10 or 20 dB, depending on the type of noise. NR algorithms are known to improve signal quality and listener fatigue but efforts to improve speech understanding need to continue.

### B. Hearing Loss Compensation

The most important task of a hearing aid is to provide audibility, namely to amplify useful

input signals to a level that the user can hear. Multi-channel hearing instruments can produce gain requirements within 5 dB of target.

### C. Beam-forming

Monaural beam-forming and binaural beam-former processing are the major forms of applications which exploit spatial information about an acoustic signal and noise. Fixed delay-and-sum methods produce a fixed directional pattern. When the delay and sum parameters are varied an adaptive beam-former that allows the directional pattern to steer itself to the location of the noise source in order to maximize the attenuation of the noise is produced.

### D. Feedback Cancellation

Adaptive feedback detectors and active feedback cancellers are appearing in many new digital hearing aids. The terms “adaptive” and “active” usually imply that the feedback canceller responds to changes in the acoustic feedback path and may be on or off as required to ensure stable operation of the instrument. These systems adapt quickly to changes in the feedback path caused by distortion of the ear canal due to, for example, the motion of the jaw when eating or speaking. With active feedback cancellation, the maximum gain of the system will usually not be reduced as it is in a fixed or static feedback reduction system.

### Intelligent Hearing Aids

An intelligent hearing aid should include at least three fundamental capabilities: auditory scene adaptation, adaptive signal enhancement and adaptive hearing compensation.

Auditory scene adaptation gives the hearing aid the capability to detect, recognize and adapt to the acoustic environment [7]. Adaptive signal enhancement gives hearing aids the ability to optimize signal quality and to reduce or eliminate the interference of noise as it exists in the real world. It includes applying adaptive filtering, artificial intelligence, array signal processing, neural networks and various signal processing technologies to achieve signal enhancement, noise reduction or even separation of the signal from noise. Intelligent signal detection, adaptive noise reduction, beam-forming or adaptive beam-forming, feedback cancellation, and overlapping signal separation are key technologies used to achieve signal enhancement. Auditory scene adaptation and

adaptive signal enhancement can provide benefits for any listener. The same signal can be perceived differently depending on individual’s hearing loss. Adaptive hearing compensation can give more benefits for people with hearing loss. When adaptive signal compensation is integrated with auditory scene adaptation and adaptive scene enhancement hearing perception can be optimized. Reliable auditory scene adaptation and adaptive signal enhancement will allow hearing compensation to be adapted to different sound sources and will allow the enhanced signal to be optimized for the characteristics of individual hearing loss.

### Summary

More powerful and more efficient DSP platforms of small size and with low operating voltage and lower power consumption will be available in future for the development of intelligent digital hearing aids. Digital signal processing technologies and the associated algorithms will continue to improve in quality and performance. Intelligent digital hearing aids will make it possible to maximize hearing aid benefit by automatic adaptation to changes in the acoustic environment.

### References

1. Popelka GR: Computer and hearing aids: A Prediction of the future. *The Hearing Journal* 1998; 51(11);52, 57-62.
2. Fabry D: Do we really need digital hearing aids? *The Hearing Journal* 1998; 51(11); 30,32-33.
3. Murray DJ and Hanson JV: Application of Digital Signal Processing to Hearing Aids: A Critical Survey, *Journal of the American Academy of Audiology* 1992; 3(2)145-152.
4. Schum DJ: Artificial Intelligence: The New Advanced Technology in Hearing Aids, July, 2004, *Audiology Online* [http://www.audiologyonline.com/articles/arc\\_disp.asp?id=733](http://www.audiologyonline.com/articles/arc_disp.asp?id=733)).
5. Berg C, Canneyt M V: Hearing Instruments Become Core-centric, *EETIMES.com: The Technology Site for Engineers and Technical Management*, posted: 6/5/98.
6. Luo H, Arndt H: Apparatus and Method for Adaptive Signal Characterization and Noise Reduction in Hearing Aids and Other Audio Devices; US Patent Application, Publication No. 2002/0191804 A1, Dec. 19, 2002.
7. Bregman AS: Auditory Scene Analysis, *MIT Press* 1990.

# WAVELET PACKETS-BASED SPEECH ENHANCEMENT FOR HEARING AIDS APPLICATION

Jiming Yang and Sridhar Krishnan

Dept. of Electrical and Computer Engineering, Ryerson University, Toronto, ON, M5B 2K3  
[jyang, krishnan]@ee.ryerson.ca

## 1. INTRODUCTION

Hearing loss is the fastest growing chronic disability in Canada. Three million Canadian have certain degree of hearing impairment. Sensorineural hearing loss is the most common type of hearing impairments, which typically results in increased hearing thresholds and increased perceived loudness growth. As a result, impaired listeners frequently experience difficulty understanding speech, especially under noisy conditions [1].

Fourier domain methods such as spectral subtraction (S.S.) and Ephraim-Malah filter are effective reducing noise in speech processing. Wavelet-based methods are explored to reduce computational complexity and to achieve better noise reduction performance. Wavelet packet transform (WPT) has good time-frequency localization; it may not require overlapped signal windows thus process less data. Flexible multi-resolution analysis can be easily achieved to decompose signal according to critical bands. Some of the WPTs also have very reasonable computation complexity.

Loudness compression compensation uses the idea of hearing model. The purpose of a model based hearing aid is to prewarp the signal so that a hearing impaired listeners will hear it just as a normal hearing individual would hear it. It is accomplished by combining the normal auditory model with the inverse of the impaired model. The sound is processed with the model for a normal hearing subject followed by the inverse model for hearing impaired subjects. This enables close normal hearing perception for a impaired hearing.

In this paper, we propose a robust hearing aids system which will simultaneously conduct hearing compensation and noise removal by using critical band WPT. With a reasonable complexity, the new speech enhancement scheme can improve perception in noise for hearing-impaired listeners. The proposed speech enhancement scheme can be used to improve next generation hearing aids device performance

## 2. METHOD

Figure 1 shows the block diagram of the proposed speech enhancement scheme. First the noisy speech time series is decomposed according to psychoacoustic critical bands by using WPT. Wavelet coefficients (WCs) of noisy speech are perceptually weighted through a weighting function incorporating masking properties [2]. The denoised WCs are then compressed to compensate recruitment of

loudness problem of hearing impaired on the Compression stage. IWPT transform the processed signal back to time domain. To evaluate the processed speech on normal hearing subjects, Hearing loss simulation introduces effect of hearing loss to both original noisy signal and processed signal to simulate certain types of hearing impairment.

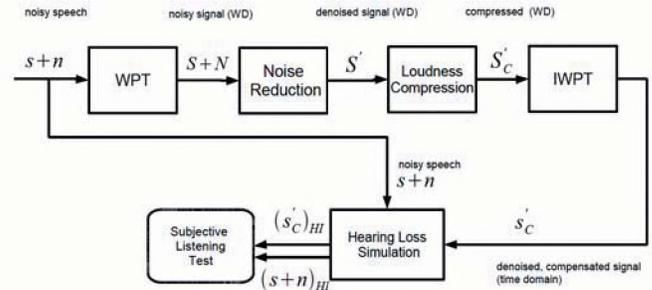


Figure 1. Block diagram of proposed method

### 2.1 Modified PTFS noise reduction

In wavelet domain, noisy input-to-noise ratio(INR) is estimated in each critical band. Based on the ratio, a weighting function  $H(m,n)=f(INR)$  is developed. The function incorporates properties of masking phenomenon. In a noise dominant band, minimum weight is given to eliminate noise. In a signal dominant band, WCs are not modified under the assumption that noise is masked by signal. In a signal noise co-dominant band, WCs are weighted by the multiplying the weighting function  $H(m,n)$ .

### 2.2 Loudness compression

Compression projects the WCs of input signal in normal listener's dynamic range to the impaired listener's. So that the loudness perceived by the impaired listener is equal to that of normal listener.  $C^* = T_{im} + (C - T_{nor})\Delta^*/\Delta$  ( $C$  and  $C^*$  are WCs before and after compression,  $T_{im}$  and  $T_{nor}$  are hearing thresholds,  $\Delta$  and  $\Delta^*$  are dynamic ranges of normal and hearing impaired) [3]. Wavelet-based compression can be easily modified to the subject specific model of loudness recruitment.

### 2.3 Hearing loss simulation

The objective of this stage is to emulate hearing loss effect on normal hearing subjects. Only normal hearing

subjects are needed to evaluate the enhancing results. Spectral smearing [4] can be used to simulate loss of frequency selectivity. Adding shaped noise raises the hearing threshold of certain bands. In this experiment, shaped noise is added to simulate the effect of high frequency hearing loss.

### 3. RESULTS

In the experiment, the noisy speech is obtained by corrupting the clean speech with white Gaussian noise of different SNR setting. Average segment SNR gain and subjective listening test are used to evaluate the processing results.

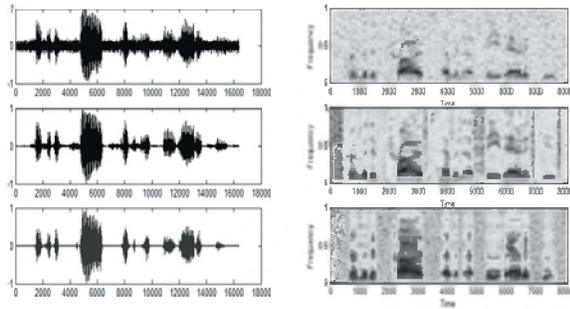


Figure 2. Waveform and spectrogram of noisy (5dB), clean and denoised signals (top to bottom)

Table 1. SegSNR improvement for enhanced speech by modified PTFS, conventional wavelet thresholding (soft), and spectral subtraction(SS)

Input SNR(dB)	SegSNR Gain(dB)		
	Modified PTFS	DWT	S.S.
-5	7.3	6.4	5.5
0	6.4	3.8	4.2
5	4.6	2.4	2.6

The waveforms (left) and spectrogram (right) of the enhanced speech are shown in Figure 2. It shows the algorithm efficiently removes the background noise. But the technique seems to attenuate the signal more in low SNR region.

The proposed noise reduction algorithm is compared with wavelet soft thresholding and spectral subtraction methods at various SNR level. Average Segment SNR gain in Table 1 demonstrates the modified PTFS method outperforms other two methods in all different SNR level. DWT and SS performances are relatively close.

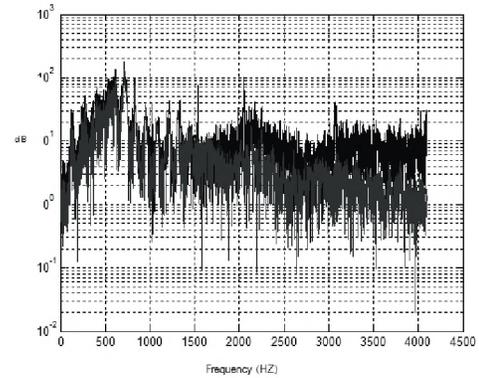


Figure 3 Spectral plot of signal before (light) and after (dark) loudness compression

Informal subjective listening test results with 11 subjects show the denoised speech from modified PTFS sounds natural, when removing the noise dramatically. S.S. is able to remove noise more than DWT but tends to introduce significant amount of musical noise.

Figure 3 illustrates that the compression stage introduces gain in high band thus compensating high frequency loss while keeps lower band signal relatively unchanged.

Listening test shows enhanced signal after hearing loss simulation produce much better speech quality than the unprocessed one. However, at low SNR level (-5dB), improvement is less significant.

### 4. CONCLUSION

Our proposed method removes background noise without introducing noticeable perceptual distortion. It also compensates frequency dependent loudness recruitment. This work advances WPT as an adequate and flexible choice for time-frequency speech processing. The complete system can serve as an efficient and effective way to develop and evaluate algorithm for hearing aids application.

### REFERENCES

- [1] Levitt, H(2001). Noise reduction in hearing aids: a review. Journal of rehabilitation research and development. Vol.38, No.1
- [2] Li, M, MacAllister, H.G., Black, N.D, De Perez T.A.(2001) Perceptual time frequency subtraction algorithm for noise reduction in hearing aids. IEEE transactions on biomedical engineering, Vol.48, No.9
- [3] Drake, L.A., Rutledge J.C. (1993). Wavelet analysis in recruitment of loudness compensation. IEEE Transactions on Signal Processing, Vol.41, No.12
- [4] Baer, T, and Moore, B (1993). Effects of spectral smearing on the intelligibility of sentences in noise. J. Acoust. Soc. Am, 94, 1229:1241

# VERTICAL FLANKING SOUND TRANSMISSION VIA THE WALL-FLOOR JUNCTION IN WOOD FRAMED CONSTRUCTION

**J. David Quirt, Trevor R.T. Nightingale, Robin E. Halliwell**  
Institute Research in Construction, National Research Council, Ottawa, K1A 0R6, Canada

## 1. INTRODUCTION

This paper presents factors controlling transmission of impact (footstep) sound from the floor of one room, to the room below, as indicated in Figure 1, and outlines how transmission data were used to develop a simplified design guide<sup>1</sup> that provides estimates of “**Apparent** Normalized Impact Sound Pressure Level” which is the combined transmission due to all direct and flanking paths.

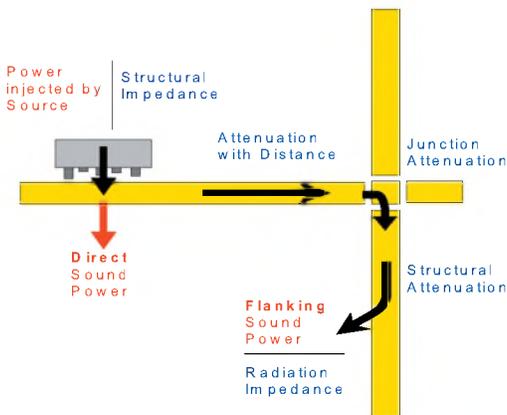


Figure 1: Transmission paths for impact sound.

The procedures to determine the “**Direct** Normalized Impact Sound Pressure Level” (due to transmission through just the floor-ceiling assembly) and the “**Flanking** Normalized Impact Sound Pressure Level” (due to flanking transmission via each wall below) are given elsewhere<sup>2</sup>.

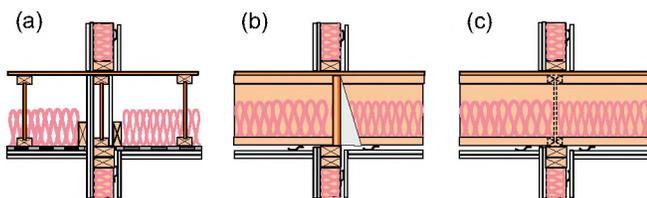


Figure 2: Construction details of the 3 wall/floor systems. Joists were oriented (a) parallel to the wall, (b) perpendicular to the wall, and (c) with joists continuous across the wall, perpendicular to it.

The basic procedure is to determine the path transmissions and then obtain estimates of the apparent sound insulation by summing the energy transmitted via the direct path through the floor-ceiling assembly and all the flanking paths involving the four wall-floor junctions. Results in this paper apply to wood-framed constructions, typical of those shown in Figure 2. Construction details are given elsewhere<sup>2</sup>.

## 2. RESULTS & DISCUSSION

Impact sound transmission was measured in accordance with ASTM E1007, except all source positions were 2.2 m from the wall. Additional tests were made to characterize structure borne propagation across the floor. Figure 3 shows that the Flanking NISPL depends on the joist orientation relative to the wall. However, the difference between single- and double-stud wall framing was small, so an average over the wall-framing cases can be used to create an estimate of flanking for the design guide.

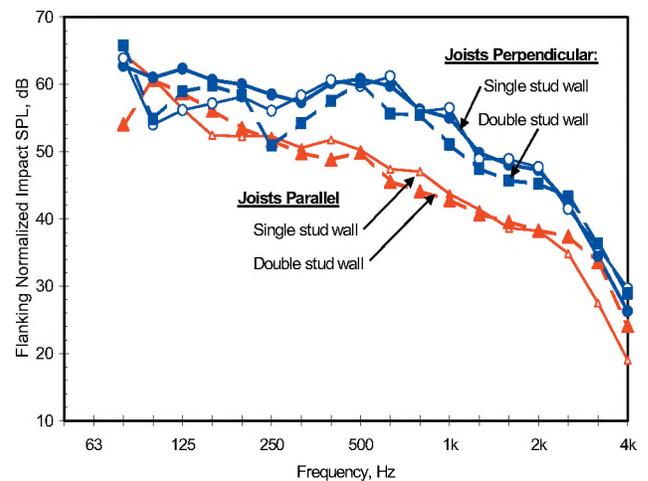


Figure 3: Flanking-NISPL for a single flanking wall with 2 layers of gypsum board, impact source 2.2 m from the wall.

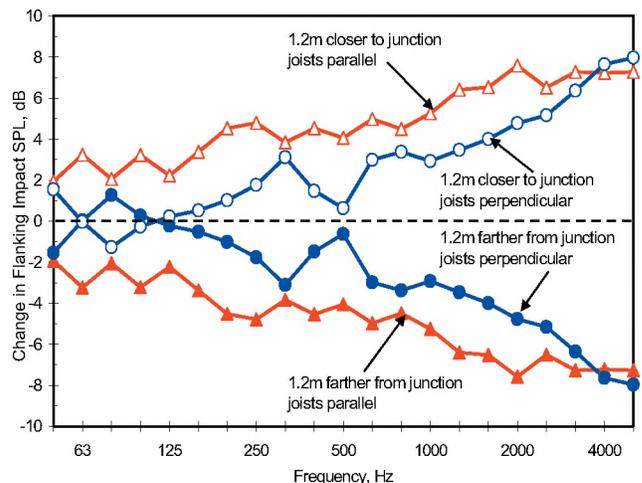


Figure 4: Change in Flanking-NISPL for one wall due to moving an impact source on the bare OSB floor surface

Flanking-NISPL for the path shown in Figure 1 increases as the source moves towards the floor-wall junction, as shown in Figure 4. The change is greater when the floor joists are parallel to the junction because attenuation across the floor is stronger in that direction.

A design estimate is obtained by considering the effect of all paths for the four walls in the room below. By evaluating a number of scenarios, it was found that different positions of the impact source on the floor above had little effect on total sound power transmitted by the combination of the four flanking paths. Figure 5 shows estimates of the Flanking-NISPL due to all four flanking walls, and compares these with direct transmission through the floor.

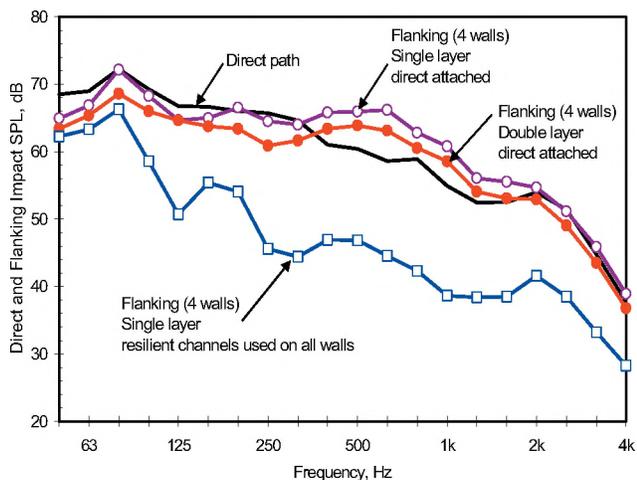


Figure 5: NISPL due to the direct path, or combined flanking paths for all four walls in the room below

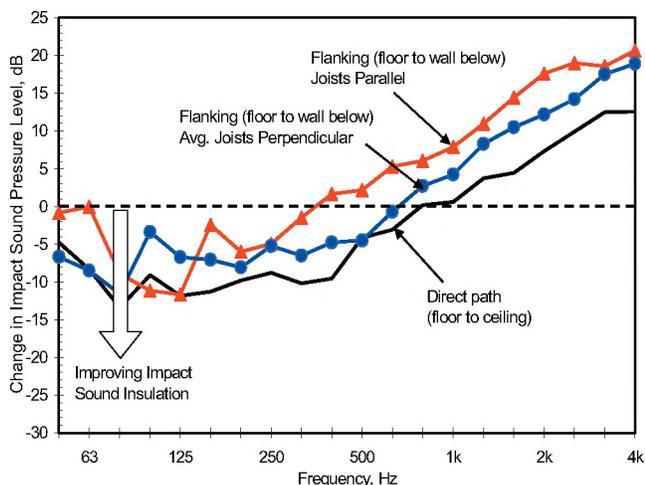


Figure 6: Effect of a 25 mm thick gypsum concrete topping bonded to the OSB subfloor

Changing the floor surface can alter impact sound transmission in two ways – changing the power injected into the floor by the impact, and changing transmission across the floor surface to the floor-wall junction. Figure 6 shows that adding a 25 mm layer of gypsum concrete reduces the

ISPL at the lower frequencies. The ISPL is significantly increased at higher frequencies because gypsum concrete is harder than OSB and more power is transferred from the tapping machine. Adding this topping also significantly changes vibration propagation across the floor surface (so in Figure 6, the change in direct transmission differs significantly from that for flanking and the flanking change is different for the two joist orientations).

In practice the gypsum concrete topping will have a floor covering that will add a compliant layer at the point of impact and there will be a significant reduction in the ISPL as shown in Figure 7. The basic vinyl flooring indicates the minimum improvement likely – carpet, or a resilient interlayer under the concrete topping could increase the attenuation for high frequencies.

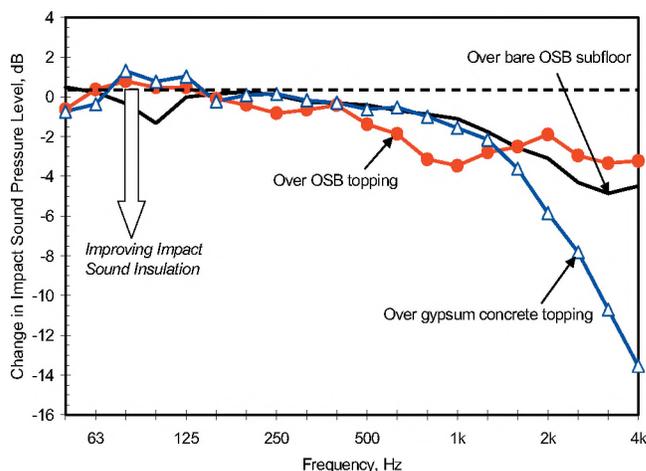


Figure 7: Change (for direct or flanking paths) due to adding vinyl floor covering over typical floor assemblies.

### 3. SUMMARY AND REFERENCES

This paper provides a very terse overview of how experimental characterization of the direct and flanking sound transmission paths in wood-framed construction can lead to a manageable set of path attenuation terms to represent the effect of specific design tradeoffs. By combining the energy transmitted via all paths allowing for the various attenuation mechanisms it is possible to arrive at estimates of the Apparent-NISPL for a range of constructions. Note that the reports are available on the IRC website at <http://irc.nrc-cnrc.gc.ca/ircpubs/>.

We wish to acknowledge the contributions of Ms. Frances King in performing many of the experimental measurements and the support of the industry partners<sup>1,2</sup>.

- 1 Quirt, J.D. Nightingale, T.R.T. Halliwell, R.E. Guide for Sound Insulation in Wood Frame Construction: Part 1, **RR193**, IRC, NRC Canada, March 2005.
- 2 Nightingale, T.R.T. Halliwell, R.E. Quirt, J.D., F. King, Flanking Transmission at the Wall/Floor Junction in Multifamily Dwellings, **RR168**, NRC Canada, March 2005.

# PREDICTING TRANSMITTED LEVELS AND AUDIBLE EFFECTS FOR MEETING ROOM SPEECH SECURITY

J.S. Bradley, M. Apfel and B.N. Gover

*Institute for Research in Construction, National Research Council, Montreal Rd., Ottawa, K1A 0R6*

## Introduction

The term ‘speech security’ is used to describe very high levels of speech privacy sometimes required for closed meeting rooms. A room is said to be completely speech secure when persons outside the room are not able to understand conversations from within the meeting room, or in more extreme cases are not able to detect any speech sounds from the meeting room. The audibility and intelligibility of transmitted speech sounds are related to the level of speech sounds relative to existing ambient noise levels (i.e. signal-to-noise ratios) at locations outside the meeting room. The levels of speech sounds from adjacent meeting rooms will depend on the loudness of the speech in the meeting room and the sound transmission characteristics of the room boundaries.

The conventional approach to measuring sound isolation between rooms is from the differences in space-average sound levels in each room. This is not appropriate for many situations where speech security must be verified. A new method is proposed in which attenuations between room-average sound levels in meeting rooms to spot-receiver positions near the outside of these rooms are used to evaluate the speech security of the room. The new method avoids problems with ill-defined receiving spaces, and spaces that do not have even approximately diffuse sound fields. It also provides assessments of speech security for more sensitive locations where an eavesdropper is more likely to be found.

## The new approach

The sound transmission characteristics of room boundaries are usually measured using standard tests in terms of room-average sound levels in the source and receiving room (e.g. ASTM E336 standard). (See also Fig. 1) According to such procedures, one can predict expected levels as a function of the  $1/3$  octave band frequency,  $f$ , in an adjacent space as follows,

$$L_R(f) = L_S(f) - TL(f) + 10 \log\{S/A(f)\}, \text{ dB} \quad (1)$$

$L_S(f)$  is the average source room sound level,  
 $L_R(f)$  is the average receiving room sound level,  
 $TL(f)$  is the sound transmission loss of the wall,  
 $S$  is the common wall area of the two rooms,  $\text{m}^2$ ,  
 $A(f)$  is the sound absorption in the receiving space,  $\text{m}^2$ .

This equation is derived assuming that the sound fields in both rooms are ideally diffuse. Although this may be approximated in the meeting room, the adjacent space

could be anything from a broom closet to an atrium or an open-plan office area. It is often difficult to apply equation (1) because the spaces are not diffuse and/or because it is not possible to define the dimensions of the receiving space. It is also more likely that an eavesdropper would be located close to the room boundary rather than in the middle of the receiving spaces. Therefore, the new method predicts transmitted sound levels 0.25 m from the outside of the meeting room,

$$L_{0.25}(f) = L_S(f) - TL(f) + k, \text{ dB} \quad (2)$$

If the receiving space is a free field,  $k \approx -3$  dB [1]. For conditions typical of meeting rooms, values of  $k$  were determined empirically.

## Experimental results

Even 0.25 m from the test wall, reverberant sound in the receiving space has a small effect on the measured  $L_{0.25}(f)$  values. Test walls were constructed between a pair of reverberation chambers and ASTM E90 sound transmission loss tests were first performed to obtain  $TL(f)$  values. The walls included wood and steel stud constructions and had STC values of 46, 53, and 56. Next, values of  $L_{0.25}(f)$  and  $L_S(f)$  in equation (2) were determined for varied amounts of sound absorption in the receiving space. A total of 3 different walls have been tested with 4 different amounts of absorption to give 12 estimates of  $k$ . These are plotted versus the reverberant sound level,  $L_{RV}$ , in Fig. 2.

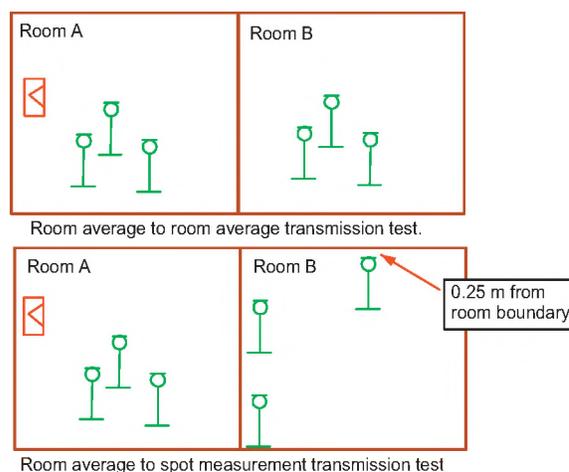


Fig. 1 Conventional room-average to room-average measurement approach (upper) and the new room-average to spot-receiver position approach (lower).

Using equation (2) and an estimate of  $k$  from Fig 2, one can predict received sound levels,  $L_{0.25}(f)$ . Fig. 3 compares measured and predicted  $L_{0.25}(f)$  for the three test walls. For the speech frequency range from 160 to 5,000 Hz, the average differences between measured and predicted  $L_{0.25}(f)$  were 0.24, 0.17 and -0.19 dB.

### Predicting the degree speech security

From a design meeting room speech level, we can predict transmitted speech levels at points 0.25 m from the outside of the meeting room using equation (2). Combining these speech levels at points 0.25 m from the outside of the meeting room with the ambient noise levels at those positions, we can estimate the degree of speech security in terms of previously determined signal-to-noise ratio speech security measures [2]. These indicate the fraction of listeners that would: understand at least one word, hear the cadence (or rhythm) of the speech, or hear any speech sounds at all. When mean subjective ratings are plotted versus the uniform weighted signal-to-noise ratio,  $SNR_{UNI32}$ , as illustrated in Fig. 4, 50% of listeners can understand at least one word for an  $SNR_{UNI32}$  of -16 dB, but for 50% of the listeners some speech sound is audible at or above  $SNR_{UNI32} = -22$  dB.

### Conclusions

The new procedure makes it possible to reliably rate the speech security of meeting rooms, even when adjacent spaces do not have diffuse sound fields and are difficult to define. The degree of speech security can be given in terms of signal-to-noise ratio measures that have been calibrated against subjective ratings in extensive previous work.

### Acknowledgements

This work was jointly funded by Public Works and Government Services Canada, the Royal Canadian Mounted Police and the National Research Council.

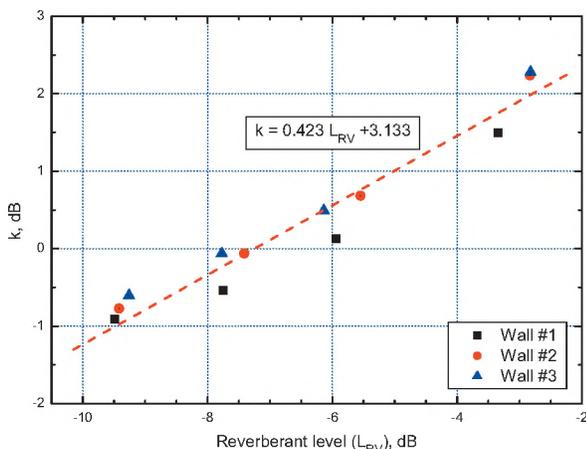


Fig. 2 Value of  $k$  versus reverberant level in the receiving space.

### References

1. Pierce, A.D., "Acoustics", McGraw-Hill (1981).
2. Gover, B.N. and Bradley, J.S., "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms", J. Acoust. Soc. Am. vol. 116, no. 6, pp. 3480-3490 (2004).

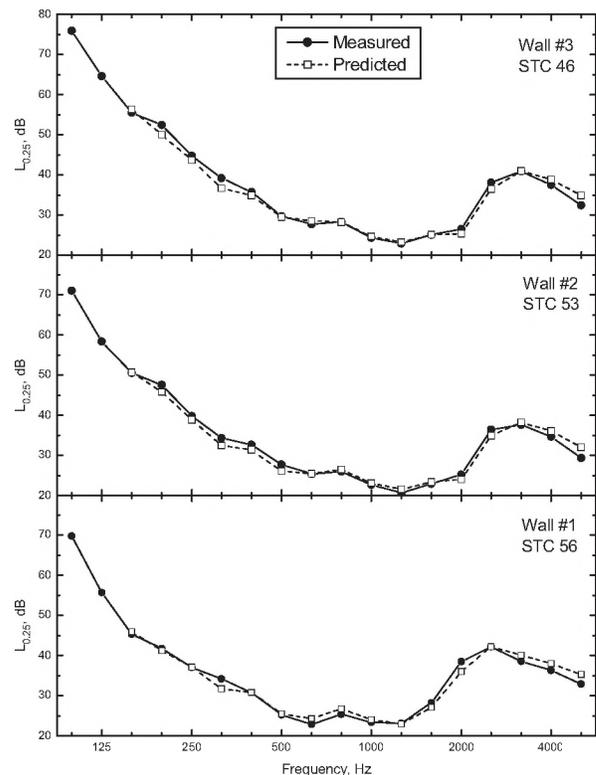


Fig. 3 Comparison of measured and predicted  $L_{0.25}$

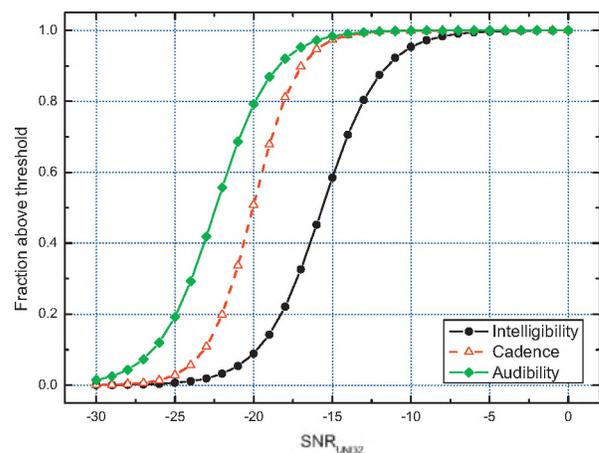


Fig. 4 Mean trends for the fraction of subjects above the thresholds: of the audibility of speech, of the audibility of the cadence of speech, and of the intelligibility of speech [2].

# INDIVIDUAL AND ROOST IDENTIFICATION USING THE ECHOLOCATION CALLS OF WILD BIG BROWN BATS, *EPTESICUS FUSCUS*

Elisabeth van Stam

Dept. of Biology, University of Western Ontario, 1151 Richmond St., London, Ontario, Canada, N6A 5B7

## 1. INTRODUCTION

The majority of bats use echoes of sound pulses to acquire three dimensional acoustic images of their environment, which allows them to navigate in darkness during flight (Griffin; 1958). This process is termed echolocation (Griffin; 1958) and can aid bats in avoiding objects, gaining access to roost sites and, locating, identifying and capturing prey.

Previous studies on microchiropteran bats have indicated that their echolocation calls contain variation and information unique to the individual and/or roost location that makes them identifiable to conspecifics (Burnett et al. 2001; Fenton et al. 2004; Kazial and Masters 2004; Masters et al. 1995 and Obrist et al. 1995). Studies on adult Big brown bats (*Eptesicus fuscus*), provided evidence of the vocal distinctiveness of individuals, sex, families, and age classes (Kazial et al., 2001; Kazial and Masters 2004; and Masters et al. 1995).

Despite the abundance of evidence and conjecture for distinctive individual call signals most supportive data has been collected in laboratory settings. However bats flying in natural conditions show more call variation as call features are altered depending on external factors such as proximity to obstacles or the ground and whether other bats are present (Burnett et al. 2001; and Surlykke, 2000). My research provides a unique opportunity to study the communication function of echolocation calls of wild Big brown bats under natural conditions

## 2. METHOD

The echolocation calls from emerging Big brown bats were recorded at 4 different maternity colonies outfitted with passive integrated transponder (PIT) tag readers throughout Fort Collins, Colorado. As a tagged bat enters or exits an equipped roost site a unique time stamp is created, recording the individual's identity, and the time and date that they passed through. The recordings were done with multi-array condenser microphones (UltraSoundgate 416) and Avisoft Recorder USG. This system allows the recording of a wide range of acoustic frequencies and can provide images of inaudible sounds in waveform, spectrogram and energy displays instantaneously. Using the

resulting visual displays the following 7 relevant call features were quantified: duration, frequency (kHz) with most energy (FME), lowest frequency (kHz) at both 5 and 10 dB below the FME, highest frequency (kHz) at both 5 and 10 dB below the FME, and Interpulse/Interval. Using only statistically significant call characteristics I attempted to identify the presence of individual voice signatures and assess call variation by running a Discriminant Function Analysis (DFA) using; individuals by roost location, sequence by roost location, and roost location as my grouping variables.

## 3. RESULTS

One hundred and twenty-eight useable call sequences from all roost locations, representing 99 individuals (29 repeated individuals) were identified. The number of individuals and sequences varied for each site.

A multivariate analysis of variance (MANOVA) determined that all seven call characteristics were statistically significant ( $P < 0.01$ ). Low levels of classification accuracy (both original and cross-validated) were determined using the DFA for both the grouping variables of sequence and individual. However, a DFA classifying roost location, using the same echolocation, calls yielded a reasonably high level of overall accuracy for both original (61.9%) and cross-validated (60.7%) classifications.

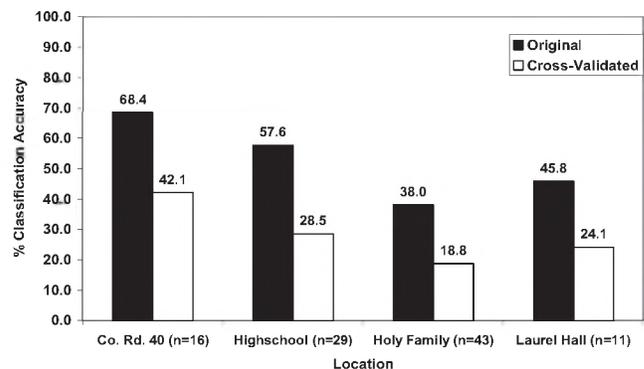


Fig. 1 Classification accuracy results using individual as grouping variable, varied between roost locations.

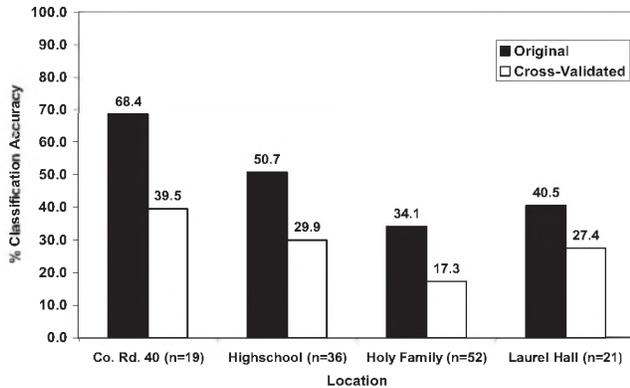


Fig. 2 Classification accuracy results using sequence as grouping variable, varied between roost locations.

In general, it can be seen that the smaller the roost sizes the higher the classification accuracy when running the DFA.

#### 4. DISCUSSION

There is currently no given threshold to denote an overall acceptable correct classification level in identifying echolocation calls using DFA. Previous literature has confidently stated classification accuracy with percentages ranging from 64-96% (Biscardi et al., 2004). My results although close fall below this range (Figures 1 and 2), indicating that it is likely that more than the echolocation calls of wild Big brown bats is needed to identify individuals during emergence. These classification results are likely result of a large the sample size, when using DFA to classify echolocation calls the correct classification accuracy decreases as the sample size increases (Biscardi et al, 2004). This tendency can clearly be seen in both Figures 1 and 2 when comparing the DFA for Co. Rd. 40 (the smallest colony) and Holy Family (the largest colony). In addition, identification of specific individuals is made even more difficult, as it appears there is higher classification accuracy when using sequence as the grouping variable in a DFA (Fig. 2). This data indicates that there is call variation of echolocation calls within the individual level and that a bat can actually sound slightly different each time it echolocates. Geographic, environmental, behavioural and body size variation have been listed as likely explanations for this (Barclay, 1999; and Obrist, 1995).

Bat calls are relatively simple high frequency sweeps that some believe are best suited for target detection and identification alone and that the physics of sound places constraints that would not allow the consistent and accurate identification between bat species (Barclay, 1999). Despite evidence opposing this claim (Kazial and Masters, 2004; Masters et al. 1995; Obrist, 1995), it may be reasonable to assume that there is not enough call variation within an

individual bat's echolocation calls to identify every bat in colonies that potentially contain hundreds of individuals. And it is the difference in sample size that explains the differences in conclusion. However, this does not preclude the idea that echolocation calls could contain enough variation for an individual to identify a fellow roost mate (Boughman, 1997; Fenton et al., 2004; Kazial et al., 2001; and Masters et al., 1995).

Big brown maternity colonies, such as the ones where I recorded, can be quite large containing hundreds of bats in which high degree of relatedness has been suggested. The ability to recognize ones roost mate or family members while outside the roost in flight situations where visual, olfactory, and even low frequency audible calls may be ineffective, may mean that useful information can be transferred that would be beneficial to both individuals. The capacity to recognize roost mates could play a significant role during foraging and hibernation; additionally it could help Big brown bats maintain group or pair cohesion as they alternate between roosts.

#### REFERENCES

- Barclay, R.M.R. (1999) Bats are not birds- A cautionary note on using echolocation calls to identify bats: a comment. *J. Mamm.*, 80(1), 290-296.
- Biscardi, S., Orprecio, J., Fenton, M.B., Tsoar, A., and Ratcliffe, J. (2004). Data, sample sizes and statistics affect recognition of species of bats by their echolocation calls. *Acta Chiropterologica*, 6(2), 347-363.
- Boughman, J.W. (1997). Greater spear-nosed-bats give group distinctive calls. *Behav. Ecol. Sociobiol.*, 40, 61-70.
- Burnett, S.C., Kazial, K. and Masters, M. (2001). Discriminating individual Big brown bat (*Eptesicus fuscus*) sonar vocalizations in different recording situations. *Bioacoustics*, 11, 189-210.
- Fenton, M.B., Jacobs, D.S., Richardson, E.J., Taylor, P.J. and White, W. (2004). Individual signatures in the frequency-modulated sweep calls of African large-eared, free-tailed bats *Otomops martiensseni* (Chiroptera: Molossidae). *J. Zool., Lond.*, 262, 11-19.
- Griffin, D.R. (1958). The acoustic orientation of bats. *In: Listening in the Dark*. (Yale University Press, USA). 57-80.
- Kazial, K., Burnett, S.C., and Masters, W.M. (2001). Individual and group variation in echolocation calls of Big brown bats, *Eptesicus fuscus* (Chiroptera: Vespertilionidae). *J. Mammal.*, 82(2), 339-351.
- Kazial, K., and Masters, M. (2004). Female big brown bats, *Eptesicus fuscus*, recognize sex from caller's echolocation signals. *Anim. Behav.*, 67, 855-863.
- Masters, W.M., Raver, K.A.S., and Kazial, K.A. (1995). Sonar signals of Big brown bats, *Eptesicus fuscus*, contain information about individual identity, age and family affiliation. *Anim. Behav.* 50, 1243-1260.
- Obrist, M.K. (1995). Flexible bat echolocation: the influence of individual, habitat and conspecifics on sonar signal design. *Behav. Ecol. Sociobiol.*, 36, 207-219.
- Surlykke, A. (2000) Echolocation behaviour of Big brown bats, *Eptesicus fuscus*, in the field and the laboratory. *J. Acoust. Soc. Am.*, 108(5), 2419-2429.

# MEASUREMENT AND MODELLING OF THE RESPONSE OF THE CAT EARDRUM TO LARGE STATIC PRESSURES

Hanif M. Ladak<sup>1,2</sup>, W. Robert J. Funnell<sup>3</sup>, Willem F. Decraemer<sup>4</sup> and Joris J. J. Dirckx<sup>4</sup>

<sup>1</sup>Depts. of Medical Biophysics and Electrical & Computer Engineering, The University of Western Ontario, London, Ontario, Canada, N6A 5C1 [hladak@uwo.ca](mailto:hladak@uwo.ca)

<sup>2</sup>Imaging Research Labs, Robarts Research Institute, London, Ontario, Canada, N6A 5K8

<sup>3</sup>Depts. of BioMedical Engineering and Otolaryngology, McGill University, Montréal, Québec, Canada, H3A 2B4

<sup>4</sup>Laboratory of BioMedical Physics, University of Antwerp, Antwerp, Belgium B-2020

## 1. INTRODUCTION

In tympanometry, a clinically important test of middle-ear function, static pressures are used which are so high that the responses become nonlinear. In current finite-element (FE) models of the middle ear, however, the relationship between applied pressure and the resulting displacements is assumed to be linear. Measurement and modelling of the response to high pressures will lead to a better understanding of the mechanics of tympanometry.

We have measured the shape and displacement patterns of the cat eardrum using phase-shift shadow-moiré topography, a non-contacting optical technique (Ladak *et al.*, 2004). The experiments involved cyclically pressurizing the middle ear up to  $\pm 2.2$  kPa after immobilizing the malleus. The immobile-malleus condition allows investigation of eardrum response in isolation from any possible nonlinearities of the cochlea and middle ear. The data indicate that the eardrum response is nonlinear at these pressures and that eardrum displacements are larger than its thickness. This suggests that, as a minimum, geometric nonlinearity must be included in FE models of the eardrum. The role of material nonlinearity is not clear from the data.

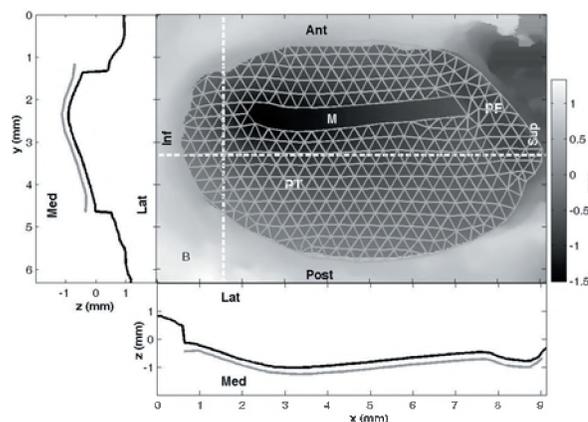
The objective of this work is to determine whether geometric nonlinearity alone can account for the observations, by incorporating it into subject-specific FE eardrum models and comparing simulation results to measurements.

## 2. FINITE-ELEMENT MODELS

### 2.1. Mesh geometry

Individualized FE meshes were constructed from unpressurized shape data using the procedure described by Funnell and Decraemer (1996). Here we present results for one cat. A sample image of the shape of the cat eardrum is shown in Figure 1. The grey levels in the image vary from black (points farthest from reader) to white (points closest to reader). A vertical profile through the pars tensa, just inferior to the umbo, is shown in black to the left of the image, and a horizontal profile through the posterior pars

tensa is shown below the image. The profiles are taken through the locations indicated by the dashed lines. Also shown is a lateral view of an FE mesh constructed from the data along with profiles (grey) through the mesh; the mesh profiles are offset so as not to overlap with the image profiles. Each triangle in the mesh represents an S3R shell element of the ABAQUS FE software package (Hibbit, Karlsson & Sorensen Inc., Pawtucket, RI, U.S.A.). No elements are generated for the manubrium as it is assumed to be completely immobile.



**Figure 1** Shape image of a cat eardrum with FE mesh superimposed. Anterior (Ant), posterior (Post), superior (Sup), inferior (Inf), medial (Med) and lateral (Lat) directions are approximate.

### 2.2. Mechanical properties

The eardrum is modelled as being isotropic, homogeneous throughout its thickness, and uniform across its surface. It is also assumed to be linearly elastic. The pars tensa has a Young's modulus of 20 MPa, a thickness of 40  $\mu\text{m}$  and a Poisson's ratio of 0.3; the pars flaccida has a Young's modulus of 1 MPa, a thickness of 80  $\mu\text{m}$  and a Poisson's ratio of 0.3 (Funnell and Decraemer, 1996). The periphery of the eardrum is assumed to be fully clamped.

### 2.3. Solution procedure

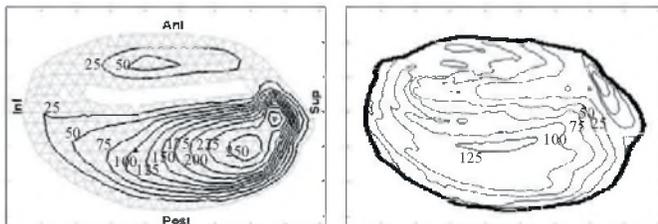
Displacements were computed for uniform static middle-ear pressures from 0 to +2.2 kPa in steps of +0.1 kPa

and from 0 to  $-2.2$  kPa in steps of  $-0.1$  kPa. A combined incremental/iterative solution procedure was used to compute the response at each step.

### 3. RESULTS

#### 3.1. Displacement patterns

Figure 2 shows simulated and measured iso-amplitude displacement patterns for a pressure of  $+2.2$  kPa. The simulated patterns exhibit a local maximum of  $255 \mu\text{m}$  in the supero-posterior pars tensa. The measured maximum, however, is smaller ( $140 \mu\text{m}$ ) and occurs more inferiorly than it does in the simulations.

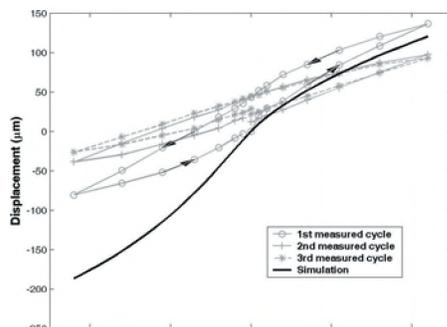


**Figure 2** (a) Simulated and (b) measured iso-amplitude displacement contours for  $+2.2$  kPa. Displacements are in  $\mu\text{m}$ .

For small negative pressures (e.g.,  $-0.1$  and  $-0.2$  kPa), simulated and measured displacements patterns are similar to those for positive pressures. For larger negative pressures, the location of the maximum is shifted inferiorly compared to that for positive pressures, for both simulated and measured displacements.

#### 3.2. Pressure-displacement curves

Figure 3 shows pressure-displacement curves for the point indicated by the triangle in Figure 2(a). Three cycles of measured data are shown in grey, and the simulated curve is shown in black. The measured curves vary from one cycle of pressurization to the next and exhibit hysteresis, with displacements being larger during unloading than during loading. There is only one simulated curve because the material is assumed to be elastic, with no hysteresis. Both simulated and measured curves exhibit nonlinearity, with displacements growing less than in proportion to the applied pressure.

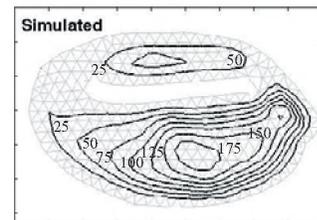


**Figure 3** Simulated and measured pressure-displacement curves.

For positive pressures, the simulated pressure-displacement curve agrees reasonably well with the measured data. For negative pressures, the calculated displacements are larger in magnitude than the measured ones.

#### 3.3. Effect of non-uniformity

The location of the maximum pars-tensa displacement in the simulated patterns can be shifted inferiorly by including non-uniformities in either the Young's modulus or the thickness. For example, Figure 4 shows the effect on the displacement patterns computed for  $+2.2$  kPa after increasing the thickness of the superior third of the posterior pars tensa from  $40 \mu\text{m}$  to  $80 \mu\text{m}$ .



**Figure 4** Simulated iso-amplitude displacement contours after including non-uniformity. Displacements are in  $\mu\text{m}$ .

### 4. DISCUSSION

The inclusion of geometric nonlinearity in an FE model allows simulation of the stiffening behaviour of the eardrum with increasing pressure.

The pars-tensa maximum in the simulated displacement patterns for positive pressures is located superior to the measured location. Increasing the stiffness or the thickness of the superior third of the posterior pars tensa in the model has the effect of shifting the maximum inferiorly and of decreasing its magnitude, improving the match with the experimental data. Such a thickness increase is consistent with recent experimental thickness measurements (Kuypers *et al.*, 2005).

### REFERENCES

- Funnell, W.R.J., and Decraemer, W.F. (1996). "On the incorporation of moiré shape measurements in finite-element models of the cat eardrum," *J. Acoust. Soc. Am.* **100**, 925–932.
- Kuypers LC, Decraemer WF, Dirckx JJJ & Timmermans J-P (2005). "Thickness distribution of fresh eardrums of cat obtained with confocal microscopy," *JARO* **6**, in press.
- Ladak, H.M., Decraemer, W.F., Dirckx, J.J.J., and Funnell, W.R.J. (2004). "Response of the cat eardrum to static pressures: Mobile versus immobile malleus," *J. Acoust. Soc. Am.* **116**, 3008–3021.

### ACKNOWLEDGEMENTS

Funding for this work was provided by the Canadian Institutes of Health Research (WRJF), the Natural Sciences and Engineering Research Council of Canada (HML) and the Québec-Belgium Exchange Programme (HML).

# EVALUATION OF PAVEMENT NOISE USING STATISTICAL TECHNIQUES

Yuen-Ting Fiona Leung<sup>1</sup>, Susan Tighe<sup>2</sup>, Scott Penton<sup>3</sup>, and Peter VanDelden<sup>4</sup>

<sup>1</sup>Civil Engineering Department, University of Waterloo, 200 University Ave. West, Waterloo, ON., Canada, N2L 3G1

<sup>2</sup>Civil Engineering Department, University of Waterloo, 200 University Ave. West, Waterloo, ON., Canada, N2L 3G1

<sup>3,4</sup>RWDI AIR Inc., 650 Woodlawn Road West, Guelph, ON., Canada, N1K 1B8

## 1. INTRODUCTION

Traffic noise is a growing concern as traffic demand increases with urban sprawl. The most common traffic noise mitigation methods include the construction of noise barriers and earth berms. Based on studies performed by different agencies around the world, the Regional Municipality of Waterloo (RMOW), Ontario, Canada, in partnership with the University of Waterloo Centre for Pavement and Transportation Technology (CPATT), have undertaken research into asphalt pavement, rubberized open friction course (rOFC) and rubberized open graded course (rOGC), to determine their noise reducing capability in the southern Ontario environment as compared to a typical pavement with a Hot-Laid 3 surface. This paper will also investigate the noise reducing capability of stone mastic asphalt mix, which is mainly used for heavy traffic conditions. Two types of sound level measurements were used in this study: Close-Proximity Method and Controlled Pass-By Method. A statistical analysis was performed utilizing the noise measurement results to determine if there is significant difference between mixes or within a mix.

## 2. DESCRIPTION OF THE STUDIED PAVEMENTS

In this study, the mix designs and gradations for the rOFC and rOGC are similar. The difference between these two mixes is related to the quality of the aggregates. The aggregates used in rOFC are premium grade and must meet much higher standards for the other aggregate tests. SMA also used a premium aggregate. The ones used in rOGC and HL-3 are local aggregates and do not meet all the requirements to qualify as a premium mix.

## 3. SITE DESCRIPTIONS

Four types of asphalt pavement surface courses were placed in a rural area surrounded by farmlands in Waterloo area. Pavement surface courses were placed in an order of rOFC, rOGC, SMA, and HL-3 from east to west. The length of each type of pavement section is approximately 600 m.

## 4. NOISE MEASUREMENT METHODS

Traffic noise measurements were taken a month after pavement placement. Thirteen testing vehicles were used for the noise testing: 5 light, 5 medium, and 3 heavy vehicles. Each noise measurement consisted of a single test vehicle passing through the test site. The driver of the testing vehicle drove through the centerline of the test road at

constant speeds of 60 km/h, 70 km/h, 80 km/h, and 90 km/h from east to west and then made a return trip. Two sound level measurement techniques were utilized in this analysis: the Close-Proximity Method (CPX) and the controlled Pass-By Method (PBM). During the noise testing, the entire area was closed and, since the study site is located in a rural, the ambient noise should be minimized and constant.

### 4.1 Close-Proximity Method

In this project, a microphone was mounted on the test vehicle and was located approximately 50 cm away from the centre of the front or rear wheel. The CPX measurement is designed to measure the direct noise generated from the interaction between vehicle tire and pavement and to avoid measuring the engine noise generated from the testing vehicle.

### 4.2 Controlled Pass-By Method

Pass-by Method (PBM) measured the sound as vehicles travel passed a stationary microphone. Four pass-by monitoring stations were set-up at the midway point of each asphalt pavement section. Each monitoring station was located 15 m away from the centreline of the road, 1.5 m above pavement, and was monitored by a technician. The maximum sound level ( $L_{max}$ ) was measured by the PBM.

## 5. NOISE MEASUREMENT RESULTS

All four pavements show that when the vehicle speed or size increases, the sound level increases in both measuring methods. The noise measurement range magnitudes, in terms of vehicle speed, are about 8 dBA (at 60 km/h) and 11 dBA (at 90 km/h) for CPX and PBM, respectively. Table 1 shows the amount of noise reduction as compared to HL-3 in terms of all vehicles, different vehicle speeds, and different vehicle sizes in a particular speed.

The noise reduction in this initial study for SMA has the worse performance of noise reduction among all pavement types. SMA did not reduce noise level in various categories. Both the rOFC and rOGC provide a significant amount of noise reduction as compared with HL-3 in all situations. The highest noise reductions for rOFC are 3.3 dBA and 2.5 dBA in CPX and PBM results, respectively. The highest reductions for OGC are 3.3 dBA and 2.8 dBA in CPX and PBM, respectively.

**Table 1: CPX – Average Sound Level Reduction**

Categories	Noise Reduction as compared to HL-3, dBA					
	CPX (Leq)			PBM		
	rOFC	rOGC	SMA	rOFC	rOGC	SMA
All Vehicle	-1.8	-2.0	0.0	-1.0	-1.7	+0.2
60 km/h	-1.4	-1.7	+0.5	-0.6	-1.2	+0.7
60 – L	-0.4	-0.6	+2.3	-0.1	-0.5	+2.1
60 – M	-1.8	-2.4	-0.7	-0.3	-1.9	-0.6
60 – H	-2.5	-2.5	-0.7	-1.9	-1.6	+0.1
70 km/h	-1.7	-1.8	+0.3	-0.8	-1.3	+0.4
70 – L	-0.7	-0.9	+1.8	-0.5	-0.5	+1.7
70 – M	-2.9	-2.8	-1.0	-1.2	-2.4	-1.2
70 – H	-1.8	-2.0	-0.3	-0.8	-1.7	-0.3
80 km/h	-2.2	-2.3	-0.2	-1.4	-2.2	0.0
80 – L	-1.0	-1.2	+1.6	-0.6	-1.6	+1.4
80 – M	-3.3	-3.3	-1.6	-1.8	-2.8	-1.6
80 – H	-2.3	-2.4	-0.7	-2.5	-2.4	+0.5
90 km/h	-1.9	-2.3	-0.5	-1.5	-2.3	-0.1
90 – L	-1.2	-1.8	+1.0	-1.2	-2.3	+1.2
90 – M	-2.7	-3.1	-1.6	-2.4	-2.8	-1.4
90 – H	-1.8	-1.9	-1.1	-1.2	-1.8	+0.7

L: Light Vehicle; M: Medium Vehicle; H: Heavy Vehicle

**6. STATISTICAL ANALYSIS**

A statistical analysis was carried out to examine if a significant difference existed between the asphalt mixes and within a particular asphalt mix in terms of vehicle speeds or sizes. The paired comparison method was used to compare the similarities between pavements by performing a t-distribution hypothesis test on the difference between two pavement mixes. A typical t-distribution hypothesis test was utilized to compare the similarity within pavement by comparing two sets of data in terms of vehicle sizes or speeds in the same mix. Four asphalt pavement mixes were studied in this project: rOFC, rOGC, SMA, and HL-3, therefore six comparisons in between mixes in terms of vehicle speed or size would be analyzed. Comparison within mix would also be analyzed in terms of vehicle speed or size.

The statistical analysis shows that there is no significant difference between rOFC and rOGC in terms of traffic speeds of 70 km/h and 80 km/h, and heavy vehicles for both sound measuring methods. No significant difference was observed between the SMA and HL-3 mixes for all vehicle speeds in both measuring methods, except for 60 km/h in PBM. In terms of vehicle sizes, significant difference was observed between the SMA and HL-3 mixes for all vehicle sizes in both measuring methods, except for heavy sized vehicle in PBM. The statistical analysis of both measuring methods also revealed that most of the other paired pavements, rOFC/SMA, rOFC/HL-3, rOGC/SMA, and rOGC/HL-3, were statistically significant different in terms of vehicle sizes or speeds.

Four types of pavements have the same statistical results in terms of the comparison between vehicle speeds within its

pavement type. In CPX statistical analysis results, all the pavements have no significant difference in the comparison between 70/80 km/h and 80/90 km/h. PBM statistical result shows that all the pavements have no significant difference in the comparison between 60/70 km/h, 70/80 km/h, and 80/90 km/h. Also, PBM statistical analysis has also shown that there are no significant differences in the comparison of 60/80 km/h in rOFC and rOGC, 70/90 km/h in rOGC and SMA.

**7. CONCLUSIONS**

It was found that vehicle noise increases when the vehicle speed or size increases for both test methods. Also, the SMA did not provide any noise reductions in most of the vehicle speeds and/or sizes in both test methods. rOFC and rOGC provided the highest amount of noise reduction in both testing methods. The statistical analysis for both measurement methods shows that there was no significant difference between rOFC and rOGC in most of the vehicle sizes or speeds comparison. It also had the same results between SMA and HL-3; there was no significant difference in some comparisons between vehicle sizes or speeds. In the statistical analysis of within mix, all pavements had no significant difference when comparing the vehicle speed of 80 km/h to 70 km/h and 90 km/h in the both measuring methods. In addition, all pavements also had no significant difference in PBM results when comparing the vehicle speeds of 60 km/h to 70 km/h; rOFC and rOGC had no significant difference when comparing the vehicle speeds of 60 km/h to 80 km/h; and rOGC and HL-3 had no significant difference when comparing the vehicle speeds of 70 km/h to 90 km/h.

**8. RECOMMENDATIONS**

It is recommended that additional sound level measurements be conducted in the future to monitor the pavement acoustical performance. Also, a comparison of the noise measurement results obtained from the two test methods should be performed and measurements of pavement acoustical absorption should also be conducted.

A life cycle cost for each pavement should also be performed for noise reducing pavement selection. Also, a further analysis of the correlation between mix design and sound level measurement should be performed which may be used to act as a noise reduction prediction model.

**REFERENCES**

Sanberg, U., Ejsmont, (2002) J.A., Tire/Road Noise Reference Book, First Edition, Poland

Duever, T., (2005) ChE 622 Statistics In Engineering Course Notes, Waterloo, ON, Canada

**ACKNOWLEDGEMENTS**

The authors gratefully acknowledge the support of the Regional Municipality of Waterloo and the Centre for Pavement and Transportation Technology.

# CROSS-LINGUISTIC INFLUENCES ON INFANT BABBLING

Karen Mattock, Susan Rvachew, And Linda Polka

School of Communication Sciences & Disorders, McGill University, 1266 Pine Ave West, Montreal, Qc, Canada, H3G 1A8  
karen.mattock@mail.mcgill.ca

## 1. INTRODUCTION

Access to phonetic structure through language input is critical for what it provides to language development. It is well known that infants' speech perception is progressively tuned to native language phonemes with increasing linguistic experience. The impact of linguistic experience is observed at a particularly early age for the perception of native language vowels (Kuhl et al., 1992). In contrast, little is known about the impact of linguistic experience on early vowel production. Most studies of vowel development have been concerned with English-learning infants: Kent and Murray (1982) observed a developmental increase in the range of first formant (F1) frequencies in vowels produced during the first year of life; Rvachew, Slawinski, Williams, & Green (1996) reported a small decrease in mean second formant (F2) frequencies and a steady increase in the standard deviation of F2 during the second year of life. Although some developmental changes in vowel production are undoubtedly due to anatomical and physiological changes in the vocal tract during this period, linguistic experience may also play a role. In a study of 10-month-old infants from English, French, Cantonese, and Algerian speaking families, de Boysson-Bardies et al., (1989) reported cross-linguistic effects on F1 and F2 frequencies. However, this study described only a single age group and a replication has never been published.

### 1.1. Purpose

The purpose of this study was to systematically examine (i) developmental changes in F1 and F2 frequencies in vowels produced by infants 10 through 18 months of age, and (ii) cross-linguistic differences in F1 and F2 in vowels produced by English- and French-learning infants during this period.

### 1.2. Hypotheses

It was hypothesized that there would be an initial overlap in the vowels produced by English and French infants, but that their respective vowel spaces would diverge linearly with age. Specifically, it was predicted that (i) mean F1 and standard deviation of F1 would remain stable with age for both groups; (ii) mean F2 would increase with age for English infants and decrease with age for French infants, and (iii) standard deviation of F2 would increase with age for both language groups.

## 2. METHOD

### 2.1. Participants and Design

Forty-three infants from Montreal, Canada, ranging in age from 10 to 18 months participated in this cross-sectional study - 23 from monolingual French- and 20 from monolingual English speaking families. All infants were born at full term, were reported to be healthy at the time of testing, and had no known hearing difficulties.

### 2.2. Babble Sample Recordings

The babble samples were recorded during a play session between mother and infant, which either took place in a sound proof booth in the laboratory, or in their homes. Mothers were instructed to interact with their child in the usual manner. The samples were obtained using a Sony portable DAT recorder and a Sennheiser microphone affixed to the infant's shirt at the shoulder. Each recording session continued until the infant produced 60 utterances, or until 30 minutes had lapsed, whichever came first.

### 2.3. Acoustic Analyses

Speech utterances were digitized at 22050Hz using Time Frequency Response software (AVAAZ) and PC hardware. Isolated vowels and vowels contained within canonical syllables (Oller, 1986) were selected if vowel or syllable duration was less than 500 ms and phonation and resonance were normal. Formant analysis on 1430 vowels was performed by an individual trained in speech acoustics, blind to the age and language background of the infant. F1 and F2 were measured at the middle of the steady state portion by LPC autocorrelation analysis with a window size of 256 points, 50% overlap, 98% preemphasis, Hanning window and model order of 12. Model order was increased or decreased for some vowels to obtain reliable measurements. Formant locations were confirmed with narrowband short-time FFT spectrograms (512 points).

### 2.4. Reliability Coding

Intraclass correlations between independently identified formant frequencies for 248 vowels were  $r = .918$  and  $r = .928$  ( $p = .01$  level) for F1 and F2 respectively.

## 3. RESULTS

### 3.1. Measures

The mean (M) and standard deviation (SD) of the formant values were calculated for each infant's vowel space, yielding four measures per sample:  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ ,  $SD_{F2}$ .

### 3.2. Regression Analyses

Linear regression analyses were used to examine the main effect of age, the main effect of language group, and the interaction of age and language group for each of the four measures. Mean F1 values and regression lines for English and French infants are plotted in Fig. 1. For  $M_{F1}$ , there was a trend towards an age by language group interaction, indicating a decline with age for French [ $F(1,21) = 7.62, p = .01$ ] but not English infants [ $F(1,18) = 2.22, p = .15$ ]. An interaction was also observed for  $SD_{F1}$  indicating an age-related decrease in dispersion of first formant frequencies for French [ $F(1,21) = 4.97, p = .04$ ] but not English [ $F(1,18) = 2.35, p = .14$ ] infants.

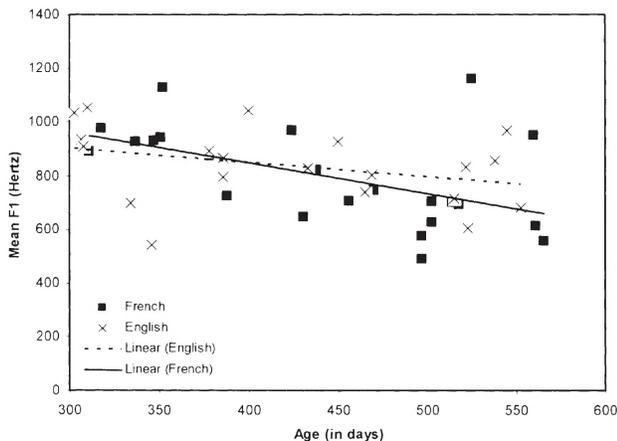


Fig. 1.  $M_{F1}$  by age and language group.

Mean F2 values and regression lines for English and French infants are plotted in Fig. 2.

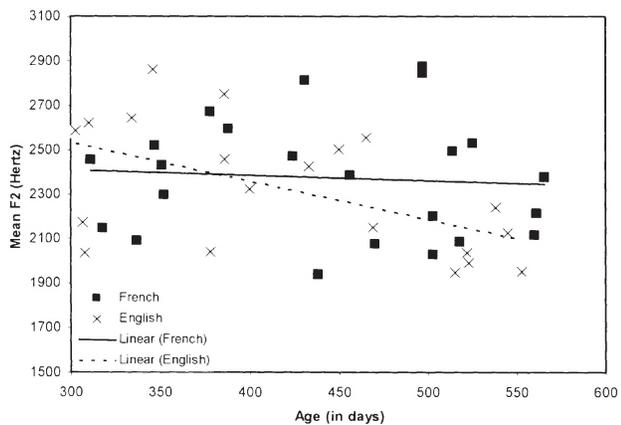


Fig. 2.  $M_{F2}$  by age and language group.

A significant interaction between age and language group was found and confirmed by the decline over age for English infants'  $M_{F2}$  [ $F(1,18) = 7.36, p = .01$ ], whereas French infants'

$M_{F2}$  remained stable across age [ $F(1,21) = .11, p = .74$ ]. An interaction was also observed for  $SD_{F2}$ , with increasing dispersion of F2 across age for English infants [ $F(1,18) = 5.57, p = .03$ ], but not for the French infants [ $F(1,21) = .17, p = .89$ ].

## 4. DISCUSSION

There is clear evidence for age by language group interactions. Mean F1 and standard deviation of F1 decreased with age for the French but not English infants. Mean F2 decreased and standard deviation of F2 increased with age for English-learning infants, but there was no change in either mean F2 or standard deviation of F2 for French-learning infants' vowels. These developmental changes are undoubtedly explained by a combination of anatomical, physiological, input, and intake factors. However, it is assumed that the anatomical and physiological constraints do not differ between language groups, and therefore these cross-linguistic differences must be explained by differences in the input provided to and received by the infants. Currently we are investigating differences in the input provided by French- and English-Canadian mothers to their infants, as well as their infant's ability to attend to and process cues for French and English vowels in their own speech.

## REFERENCES

- de Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.
- Kent, R. & Murray, S. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, 72, 353-365.
- Oller, D. K. (1986). Metaphonology and infant vocalizations. In B. Lindblom and R. Zetterstrom (Eds.), *Precursors of Early Speech* (pp. 21-36). New York: Stockton.
- Rvachew, S., Slawinski, E. B., Williams, M., and Green, C. L. (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics*, 24, 19-28.

## ACKNOWLEDGEMENTS

This research was supported by research grants from the Canadian Language and Literacy Research Network and the Natural Sciences and Engineering Research Council of Canada to the second author and a post-doctoral fellowship to the first author from the Centre for Research in Language, Mind and Brain. We are grateful to the infants and their families for their participation. We thank Sara Turner, Marie Desmarteau, Jade Heilmann, Pi-Yu Chiang, and Heather McKinnon for their assistance with data collection and analysis.

# RELATIONSHIP BETWEEN MEASURES OF INTELLIGIBILITY AND PHONETIC ACCURACY IN CHILDREN WITH AND WITHOUT CLEFT PALATE

Carrie L. Gotzke and Megan M. Hodge

Dept. of Speech Pathology and Audiology, University of Alberta, Edmonton, AB [megan.hodge@ualberta.ca](mailto:megan.hodge@ualberta.ca)

## 1. INTRODUCTION

*The Speech Intelligibility Probe for Children with Cleft Palate (SIP-CCLP) Ver. 3*<sup>1</sup> is a computer-administered, 124-word imitation measure of a speaker's ability to make phonetic contrasts of English understandable to unfamiliar listeners. The contrasts sample the expected speech error patterns for children with cleft palate<sup>1</sup>. These include manner and place preference, sibilant, voicing and glottal errors. For each error type, minimal word pairs that varied in their constituent consonants were selected as *SIP-CCLP* stimulus words. A child's recorded *SIP-CCLP* word productions can be judged using software-administered open- and closed-set tasks. In the open-set task, listeners identify the child's words, without knowledge of the target utterance, to yield an overall intelligibility score (percent words identified correctly) that reflects severity of the speech disorder. In the closed-set task, listeners first identify what sound is heard in a target position for each item and then rate the sound as "clear" or "distorted". The identification component captures errors of perceived sound substitutions or omissions, while the rating component captures perceived subphonemic differences that affect sound clarity. In the closed-set task, both identification judgments and distortion ratings contribute to a phonetic accuracy (PA) score. To guide treatment planning, error patterns that contribute to the child's intelligibility deficit can also be identified from an item analysis of listeners' responses on the closed-set task.

The purpose of this study was to determine how well *SIP-CCLP* PA scores predicted *SIP-CCLP* intelligibility scores and intelligibility scores obtained from a spontaneous speech sample for children with and without cleft palate. If *SIP-CCLP* and spontaneous sample intelligibility scores can be predicted with a high degree of confidence from *SIP-CCLP* PA scores, then it would more efficient clinically to have listeners perform only the *SIP-CCLP* closed-set task.

## 2. METHOD

### 2.1 Child Participants

Children age 3 years to 6 years, 11 months participated. The mean age of the 12 children with cleft palate was 51.2 mos (SD=11.7 mos) and of the 12 children without cleft palate was 50.1 mos (SD = 11.1 mos). Children with cleft palate had hearing within normal limits, no concomitant physical or cognitive impairments, and receptive and expressive language abilities within normal limits on a

standardized measure. Children had a variety of palatal impairments ranging from submucous cleft to bilateral cleft lip and palate. The 12 children without cleft palate all passed a hearing<sup>2</sup> and oral mechanism screening<sup>3</sup>, and scored at or above the 16<sup>th</sup> percentile on a standardized measure of receptive and expressive language and articulation.

Audio recordings were made in a quiet environment directly to a personal computer as digital audio files (SR 48 KHz; QS 16 bits). Recordings were made using a Shure WH20 unidirectional dynamic headset microphone, Audio Buddy Dual Mic Preamplifier and custom software<sup>1</sup>. The *SIP-CCLP* software created a unique order of the stimulus items for each administration. The child was instructed to repeat the name of the picture displayed on the computer monitor after the examiner modeled the word. The software recorded the child's production. A 15-minute spontaneous speech sample was also collected from each child using play scenarios, following Shriberg's<sup>4</sup> procedures. Each utterance in the child's sample was transcribed orthographically. Utterance boundaries were determined using Shriberg's conventions. A section with multiple child utterances longer than two words and few examiner turns was located in the transcript. Within this section, a subsection containing 100 consecutive words<sup>5</sup> and few utterance boundaries was selected for the listening task. During playback, each utterance was presented as an audio-file, in the order of occurrence in the transcript subsection. The transcript served as the key for scoring listeners' responses.

### 2.2 Listener Judges and Tasks

Seventy-two university students served as listener judges. All listeners had Canadian English as their first language and passed a hearing screening<sup>2</sup>. Three students were randomly assigned to judge each child's *SIP-CCLP* recordings. Three different listeners were randomly assigned to judge each child's spontaneous speech sample. All listening sessions took place in a Madsen OB822 sound booth. Stimuli were presented through a Technics Stereo Integrated Amplifier (model SU-V460) connected to ElectroVoice S-40 compact monitor speakers located in the booth. Listener judges were asked about the comfort level of the playback volume after the practice items for all three tasks (*SIP-CCLP* open- and closed-set, spontaneous sample) and adjustments made as necessary within a range of 55 – 65 dBA.

For the *SIP-CCLP* open-set identification task each listener was instructed to type the word or words heard, using the computer keyboard. If the words were not clear, the listener was instructed to guess. For the *SIP-CCLP* closed-set task listeners were told that they would see two words with a sound or sounds underlined, a “blank” button and a “can’t identify” button. They were instructed to focus on what was heard in the underlined position as they heard the child’s production of the item. If what was heard corresponded to one of the underlined choices, the judge was instructed to select it. If a different sound(s) was heard in the underlined position, judges were instructed to select the “blank” button and type in what was heard. If the sound(s) heard in the underlined position could not be identified, judges were instructed to select the “can’t identify” button. Judges were also instructed to rate what was heard in the underlined position as “clear” or “distorted” if one of the first three buttons (minimal contrast pair or the blank) was selected. The *SIP-CCLP* software randomly generated the order of item presentation for both open- and closed set judging tasks. For the spontaneous speech sample judging task, listeners were told that they would hear a series of utterances varying in word length. They were instructed to write down each word that they heard. Listeners heard each utterance twice if desired.

For the *SIP-CCLP* closed-set task, the responses for each listener judge were examined to determine the number of correct/clear (2 points), correct/distorted (1 point), and incorrect (0 points) items and then the assigned scores for all items were tallied. This sum was converted to a percentage of the total possible score (194 items x 2 = 388). The mean of the three listeners’ scores served as the child’s PA score. For the *SIP-CCLP* and spontaneous sample intelligibility scores, the number of words identified correctly by each judge was converted to a percentage. The mean of the three judges’ scores served as the child’s intelligibility score.

### 3. RESULTS

PA scores for the children with cleft palate ( $M = 73.4\%$ ,  $SD = 11.6$ ) were significantly lower ( $t = -4.41$ ,  $p < .000$ ) than those for the children without cleft palate ( $M = 89.4\%$ ,  $SD = 5.0$ ). *SIP-CCLP* open-set intelligibility scores for the children with cleft palate ( $M = 53.3\%$ ,  $SD = 15.9$ ) were significantly lower ( $t = -4.31$ ,  $p < .000$ ) than those for the children without cleft palate ( $M = 76.7\%$ ,  $SD = 10.1$ ). Spontaneous speech open-set intelligibility scores were significantly lower ( $t = -3.03$ ,  $p = 0.003$ ) for the children with cleft palate ( $M = 67.5\%$ ,  $SD = 13.3$ ) than for the children without cleft palate ( $M = 83.6\%$ ,  $SD = 12.7$ ). As shown in Table 1, Pearson correlation coefficients were high between *SIP-CCLP* PA scores and *SIP-CCLP* intelligibility scores. For the children with cleft palate, correlation coefficients were moderate and similar between *SIP-CCLP* PA and spontaneous speech intelligibility scores (.63) and between *SIP-CCLP* and spontaneous speech intelligibility scores (.62).

Interlistener reliability was examined for the three listeners for each child, for each score type, using the intraclass correlation coefficient (ICC). ICC values for *SIP-CCLP* PA scores and *SIP-CCLP* and spontaneous sample intelligibility scores for the children with cleft palate were .92, .95 and .87 respectively and, for the children without cleft palate, were .91, .96 and .93.

Table 1. Correlation matrix for *SIP-CCLP* phonetic accuracy (PA) and *SIP-CCLP* and spontaneous speech intelligibility scores.

	<i>SIP-CCLP</i> Intelligibility Scores		Spontaneous Speech Intelligibility Scores	
	With CP*	Without CP	With CP	Without CP
<i>SIP-CCLP</i> PA Scores	.93 ( $p < .000$ )	.87 ( $p < .000$ )	.63 ( $p = .03$ )	.48 ( $p = .113$ )
<i>SIP-CCLP</i> Intelligibility Scores			.62 ( $p = .016$ )	.67 ( $p = .008$ )

\*CP = Cleft Palate

### 4. DISCUSSION

ICC values of greater than .9 indicated that listeners achieved acceptable reliability in generating *SIP-CCLP* PA and *SIP-CCLP* intelligibility scores. PA scores over estimated spontaneous speech intelligibility scores on average by 5.9% and 5.8%, and also over-estimated *SIP-CCLP* intelligibility scores on average by 20.1% and 12.7%, respectively, for the children with and without cleft palate.  $R^2$  values (86% and 76%) were high between PA and *SIP-CCLP* intelligibility scores for children with and without cleft palate. These results suggest that listeners’ responses on the *SIP-CCLP* closed-set judging task can be used to generate PA scores that show a highly predictable relationship to *SIP-CCLP* intelligibility scores. In clinical settings, where time for listeners to perform speech identification tasks is limited, use of the same listener task to estimate an overall intelligibility score, as well as provide error specific information for treatment planning, is an advantage.

### REFERENCES

- Gotzke, C.L. (2005). *Speech intelligibility probe for children with cleft palate version 3: Assessment of reliability and validity*. Unpublished Master’s Thesis, University of Alberta, Edmonton, AB.
- American Speech and Hearing Association (ASHA). (1985). Guidelines for identification audiometry. *American Speech and Hearing Association Journal*, 27, 409-453.
- Dworkin, J. & Culatta, R. (1996). *Dworkin-Culatta Oral Mechanism Exam - Treatment*. Nicholasville, KY: Edgewood Press.
- Shriberg, L. (1986). *Programs to Examine Phonetic and Phonological Evaluation Records (P.E.P.P.E.R.) Version 4.0*. Madison, WI: University of Wisconsin-Madison.
- Weiss, C. (1980). *Weiss Comprehensive Test of Articulation*. Hingham, Massachusetts: Teaching Resources Corporation.

### ACKNOWLEDGEMENTS

Supported by a grant from the Canadian Language and Literacy Research Network and a graduate scholarship awarded to the first author from the National Science and Engineering Research Council of Canada.

# ACOUSTIC AND ARTICULATORY SPACE BEFORE AND AFTER LATERAL TONGUE RESECTIONS IN ORAL CANCER PATIENTS

Orchid Rastadmehr, Tim Bressmann, Chiang-Le Heng, Jonathan C. Irish

Dept. of Speech-Language Pathology, University of Toronto, 500 University Avenue, Ontario, Canada, M5G1V7  
tim.bressmann@utoronto.ca

## 1. INTRODUCTION

Tongue cancer is relatively rare but has particularly pernicious functional consequences for affected individuals. The tongue is the main organ of speech and swallowing. Problems regarding appearance, speech ability and nutrition often decrease the quality of life of glossectomy patients. The standard treatment of tongue cancer involves resection and reconstruction, often accompanied by radiotherapy. The surgical resection and reconstruction can cause speech deficits in the glossectomy patients. It is usually assumed that the loss of tissue will be somewhat proportional to a loss of function and movement range (Rentschler & Mann, 1980). For example, Davis et. al. (1987) argue that articulation errors result from a restricted range of tongue motion. Resections of the more actively moving anterior parts of the tongue dorsum and the tongue blade result in a greater loss in function than more posterior resections (Logemann et al., 1993). Obviously, the less the tongue body is modified during the surgery the less the speech production is altered. However, Savariaux et al. (2001) demonstrated that even patients with extended resections retain enough lingual motility in order to produce a differentiated range of vowels.

Despite the significant impact of tongue resections on speech and tongue motility, there have been very few studies on the relationship between the motility of the glossectomee tongue and the acoustic output. The quality of speech after a tongue resection and reconstruction is an important outcome measure for surgical success. The present research looks for an accurate description on the functional consequences of a partial glossectomy. In this paper, we investigated the impact of lateral tongue resections on midsagittal tongue movement and vowel production.

## 2. METHODS

Nine patients, male and female from 30 to 55 years of age underwent partial lateral tongue resections with different reconstructions. The patients were recruited from the Wharton Head and Neck Centre, Princess Margaret Hospital in Toronto. All participants were seen a few days before their surgery and one month postoperatively. The participants were asked to read one third of the Grandfather

passage (13-14 sec.) and repeat the syllables /aka/, /iki/, /uku/, /ata/, /iti/, and /utu/ five times each with stress on the second syllable. The patients' midsagittal tongue movement during these speech tasks was visualized using a General Electric Logiq Alpha 100 MP ultrasound scanner and a model E72 6.5MHz transducer with a 114° microconvex array. The ultrasound machine delivers a video frame rate of 30 fps. Video output of the ultrasound device was captured to digital video, together with high-quality audio. We used our Ultra-CATS software (Ultrasonographic Contour Analyzer for Tongue Surfaces) to measure the ultrasound image sequences. In this program, the distance from the ultrasound transducer to the tongue surface is measured along radiating gridlines in 5° intervals (Figure 1).

We used these measurements to establish the typical ranges of motion for the available points on the tongue before and after the operation. We calculated the average velocity of tongue movement at all available points on the tongue surface in m/sec by dividing the total distance traveled at each point by the total time of the sequence. We also evaluated the participants' vowel space by measuring the first and second formant frequencies during the steady-state phases of the second vowel in each of the VCV sequences. The measurements were made with the Kay MultiSpeech 3700 software.

## 3. RESULTS

Due to the page limitations, we only describe averaged results for the speaker group. The measurement values for the motion ranges at different grid angles demonstrated that the distance of the midsagittal tongue surface increased postoperatively. The boxplots with the average values for all 9 participants can be found in Figure 2. Multiple t-tests with a Bonferroni-adjusted alpha indicated that these postoperative increases were significant ( $p < 0.005$  in all cases). The postoperative velocities for all points on the tongue increased. A line graph with the average velocities for different points along the tongue can be found in Figure 3. We calculated mean travel speeds for all nine patients and compared the pre-and postoperative velocities using a t-test. The result indicated that the patients moved their tongues significantly faster following the operation ( $p < 0.05$ ). Finally, the formant measurements showed a slight decrease in vowel space for all speakers. These changes were most

marked for the second formants of the vowels /i/ and /a/, which dropped slightly in all speakers. When we calculated the average area for the vowel triangle for all speakers and input the results into a t-test, we found that the decrease in the vowel areas was not statistically significant.

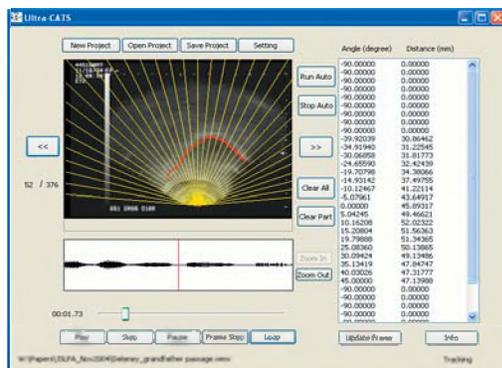


Fig. 1. Screenshot of the Ultra-CATS software with the measurement grid superimposed over a midsagittal ultrasound image with the tongue surface tracing. The anterior tongue is towards the right of the image.

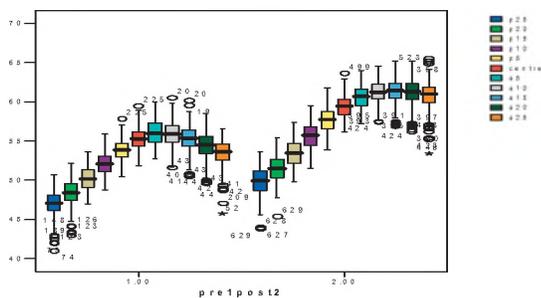


Fig. 2. Boxplots of the average movement ranges of different points on the tongue before and after the partial glossectomy surgery.

#### 4. DISCUSSION

Tongue movements of 9 patients with tongue cancer were analyzed using midsagittal ultrasound scans and parallel sound recordings. The range of tongue motion, the velocity of the tongue movement, and the vowel space were compared pre- and postoperatively. The results indicated larger ranges of motion and increases in the velocity of the tongue movement after resection. These larger and faster tongue movements are likely compensatory strategies that are employed by the patients to respond to the loss of lateral tongue tissue. As a result of this increased effort in tongue movement, almost all patients retained over 90% of their preoperative vowel space. These are interesting findings with regards to the ongoing debate among surgeons whether it is more desirable to replace as much lost tongue tissue as possible to preserve the overall volume of the tongue (Urken et al., 1994), or to opt for a reconstruction that maintains the motility of the residual tongue (Imai & Michi, 1992). Our future research about this topic will include more detailed articulatory profiles and speech intelligibility testing.

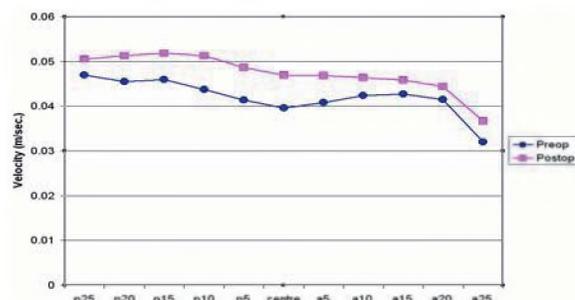


Fig. 3. Line graph of the average velocities of different points of the tongue before and after the partial glossectomy surgery.

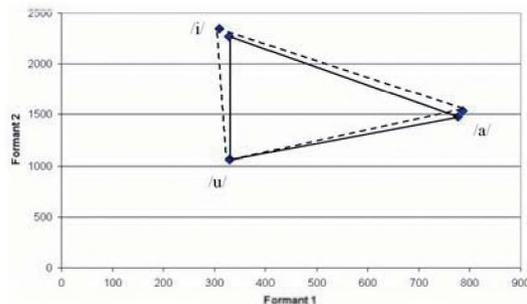


Fig. 4. Average vowel space for all patients (dashed line: preoperative vowel space; solid line: postoperative vowel space).

#### REFERENCES

Davis, J. W., Lazarus, C., Logemann, J.A., Hurst, P.S. (1987). Effects of a maxillary glossectomy prosthesis on articulation and swallowing. *J Prosthet Dent*, 57, 715-720.

Imai, S., Michi, K. (1992). Articulatory function after resection of the tongue and floor of the mouth: palatometric and perceptual evaluation. *Journal of Speech and Hearing Research*, 35, 68-78.

Logemann, J.A., Pauloski, B.R., Radermaker, A.W. (1993). Speech and swallowing function after tongue/base of tongue resection with primary closure. *Journal of Speech, Language and Hearing Research*, 36, 918-926.

Rentschler, G.J., Mann, M.B. (1980). The effects of glossectomy on intelligibility of speech and oral perceptual discrimination. *Journal of Oral Surgery*, 38, 348-354.

Savariaux, C., Perrier, P., Pape, D., Lebeau, J. (2001). Speech production after glossectomy and reconstruction lingual surgery. The 2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications. Italy.

Urken, M.L., Moscoso, J.F., Lawson, W., Biller, H.F. (1994). A systematic approach to functional reconstruction of the oral cavity following partial and total glossectomy. *Archives of Otolaryngology – Head and Neck Surgery*, 120, 589-601.

# FINITE-ELEMENT MODELLING OF ACOUSTIC WAVE SCATTERING FROM FLUID, RIGID AND ELASTIC OBJECTS

Omar Falou<sup>1</sup>, J. Carl Kumaradas<sup>1,2</sup>, and Michael C. Kolios<sup>1,2</sup>

<sup>1</sup>Dept. of Electrical and Computer Engineering, Ryerson University, 350 Victoria St., ON, Canada, M5B 2K3  
ofalou@ryerson.ca

<sup>2</sup>Dept. of Physics, Ryerson University, 350 Victoria St., ON, Canada, M5B 2K3

## 1. INTRODUCTION

Ultrasound has long been used as a diagnostic technique in which high frequency sound waves are directed into the body. Structures and tissues within the body reflect (scatter) the sound, and the resulting echoes are processed by a computer to produce an image. It has already been shown that high-frequency ultrasound (20MHz-60MHz) can be used to detect structural and physical changes in certain cell lines and cell ensembles during induced cell death, and in particular during apoptosis [1, 2].

In order to determine the proportion of responding cells in a tissue (to assess the effectiveness of the chemotherapy), a theoretical model of acoustic wave scattering by cell ensembles is required. The majority of ultrasound tissue models do not take into account the elastic nature of cells, which maybe significant at high frequencies. Therefore, a more complete model which incorporates the mechanical properties of cells is needed. The initial goal of this work is to develop a scattering model of a single cell.

Analytical solutions to the problem of wave scattering from three dimensional structures such as spheres have been studied extensively in the past. However, these formulations are not flexible enough to be extended to wave scattering from complex geometries such as those of cells. In this paper, we develop a finite element model of acoustic wave propagation through 3-D arbitrary-shaped structures to solve the problem of high-frequency acoustic scattering. The long term goal of this research is to understand the interaction of ultrasound with micrometer scale objects such as cells.

## 2. METHOD

The FEMLAB® software package (COMSOL AB, Stockholm) was used to develop and solve acoustic scattering models from fluid, rigid and elastic spheres immersed in a fluid medium, for which analytical solutions are available. The descriptions of these models are given in sections 2.1, 2.2 and 2.3 respectively.

### 2.1. Scattering from a Fluid Sphere

The entire computation domain is modelled as a sphere filled with fluid domains. A radiation (absorbing) boundary condition is applied on the domain boundary (surface of sphere). This boundary has a dual role: a) it approximates an infinite space b) it acts a wave source for an incident plane wave. The scatterer is modelled as a sphere centered in the middle of the domain. At the scatterer surface the particle normal acceleration is the same, thus a continuity boundary condition is required.

Inside and outside the spherical scatterer, wave propagation is governed by the Helmholtz equation of the acoustics application mode supplied by FEMLAB.

### 2.2. Scattering from a Fluid Immovable Sphere

The setup of this model is similar to that of the scattering from a fluid sphere described 2.1, except the fluid sphere is replaced by a rigid immovable one. In FEMLAB, this scenario is realized by disabling the acoustic wave equation (Helmholtz equation) inside the scatterer domain and by setting the scatter-surrounding medium interface boundary condition to “sound hard” condition which sets the acceleration at the surface to zero. This model used ~95,000 mesh elements and ~85,000 degrees of freedom.

### 2.3. Scattering from an Elastic Sphere

The two previous models assume only compressional waves. The Helmholtz equation is sufficient to describe this propagation. When a scatterer is elastic, in addition to compressional waves, shear and surface waves, may also exist inside and at the surface of the scatterer, respectively. In order to model all these phenomena, the FEMLAB 3D solid, stress-strain application mode is used inside the scatterer. The scatterer is surrounded by a fluid governed by the Helmholtz equation. Continuity of the normal component of acceleration is imposed as a boundary condition at the scatterer surface. ~93,000 mesh elements and ~44,000 degrees of freedom were used in this model.

### 3. RESULTS

In order to validate the three FEMLAB models, three incident wave frequencies (or  $ka$ 's) were used for each model.

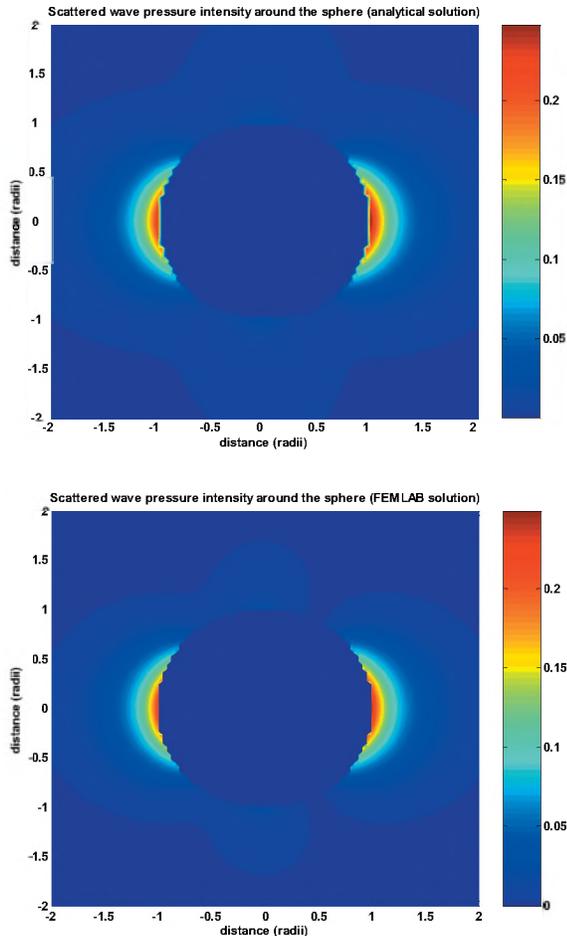


Fig. 1. Analytical (top) and FEMLAB (bottom) solution for the elastic scattering problem. The plots show the near-field pressure intensity for a slice parallel to the direction of wave propagation (left to right) through the origin. The solution inside the sphere was not calculated.

The results of these simulations were compared against analytical solutions derived by Anderson [3] (fluid scatterer) and Faran [4] (rigid immovable and elastic scatterer).

The scattered wave pressure intensity at 180 locations, situated at a distance at one and a half radii from the scatterer center was used to calculate the average error. Table 1 shows the calculated average error of the FEMLAB solution with respect to the analytical solution for each of the models.

Figure 1 compares the analytical and the FEMLAB solutions for an elastic sphere (polystyrene) immersed in water for  $ka = 1$ . It shows a slice parallel to the direction of

propagation of the incident plane wave, which travels from left to right. Due to the cylindrical symmetric nature of the model, the plot is independent of its angle of rotation around the axis of propagation.

Table 1. Average error for each model

Scatterer Type	$ka$	% Error
Low Contrast Fluid (density: 1060 kg/m <sup>3</sup> and sound speed: 1530 m/s)	0.8378	1.43
	1.0	1.47
	3.0	12.08
Rigid Immovable	0.8378	7.30
	1.0	5.77
	3.0	7.99
Elastic (Polystyrene)	0.8378	2.05
	1.0	4.24
	3.0	6.42

### 4. DISCUSSION AND CONCLUSION

The average error of the FEMLAB solution with respect to the analytical solution generally increases as  $ka$  increases. This could be explained by the variation of the mesh density with respect to the wavelength of the incident wave. A higher value of  $ka$ , implies a smaller wavelength, which leads to a lower mesh density per wavelength. A lower mesh density causes a decrease in the accuracy of the solution.

The finite-element models presented above provide accurate predictions of wave scattering by three types of scatterers, fluid, rigid immovable and elastic. These models can be extended to solve acoustic scattering problems from arbitrarily shaped and complex structures such as cells.

### REFERENCES

- [1] G. J. Czarnota, M. C. Kolios, H. Vaziri, S. Benchimol, F. P. Ottensmeyer, M. D. Sherar, and J. W. Hunt, "Ultrasonic biomicroscopy of viable, dead and apoptotic cells," *Ultrasound in Medicine and Biology*, vol. 23, pp. 961-965, 1997.
- [2] M. C. Kolios, G. J. Czarnota, A. Worthington, A. Giles, A. S. Tunis, and M. D. Sherar, "Towards Understanding the Nature of High Frequency Backscatter from Cells and Tissues: An Investigation of Backscatter Power Spectra from Different Concentrations of Cells of Different Sizes," presented at 2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control Joint 50th Anniversary Conference, 2004.
- [3] V. C. Anderson, "Sound Scattering from a Fluid Sphere," *Journal of the Acoustical Society of America*, vol. 22, pp. 426-431, 1950.
- [4] J. J. Faran, "Sound Scattering by Solid Cylinders and Spheres," *Journal of the Acoustical Society of America*, vol. 23, pp. 405-418, 1951.

# FIRST-ORDER SPECKLE STATISTICS OF ULTRASOUND BREAST IMAGES SYNTHESIZED FROM A COMPUTATIONAL ANATOMY MODEL

Yi-Ting Shen<sup>1,3</sup> and James C. Lacefield<sup>1,2,3</sup>

Depts. of <sup>1</sup>Electrical & Computer Engineering and <sup>2</sup>Medical Biophysics, University of Western Ontario,  
279 Thompson Engineering Building, London, Ontario, Canada N6A 5B9

<sup>3</sup>Imaging Research Laboratories, Robarts Research Institute, 100 Perth Drive, London, Ontario, Canada N6A 5K8  
jlacefield@eng.uwo.ca

## 1. INTRODUCTION

Focus aberration produced by heterogeneous tissues can significantly reduce the spatial and contrast resolution of ultrasound imaging systems. Aberration can be particularly severe in breast imaging, which has limited the role of ultrasound in breast cancer diagnosis. Techniques based on adaptive inverse filtering<sup>1,2</sup> are now the state of the art in ultrasound aberration correction. Compared to thin phase screen methods, inverse filtering techniques are less dependent on assumptions about the spatial organization of aberrating structures in tissue, but realistic aberration models are still needed to evaluate new focusing methods and guide improvement of their performance.

This paper introduces a three-dimensional breast anatomy model and image synthesis method designed for future use in numerical studies of aberration correction and lesion detection in ultrasound imaging. First-order speckle statistics of synthetic B-mode ultrasound images are analyzed as an initial step toward validation of the model. The objective of this analysis is to demonstrate that the synthetic images possess the non-Rayleigh speckle statistics that are characteristic of clinical ultrasound breast images.<sup>3</sup>

## 2. METHODS

### 2.1 Breast anatomy model

A three-dimensional breast anatomy model was implemented using polyhedrons, spline curves, and fractal structures to represent skin, fat lobules, connective tissue, and lactiferous ducts in a manner similar to recently developed models for x-ray mammography simulation.<sup>4,5</sup> An unlimited number of realizations of breast anatomy are possible via random variation of user-specified structural parameters.

### 2.2 Ultrasound image synthesis

A two-dimensional cross-section with 0.015 mm sample spacing was extracted from the anatomy model and input to a  $k$ -space (spatial frequency domain) ultrasound propagation simulator.<sup>6</sup> Density, compressibility, and

frequency-dependent attenuation parameters were assigned to each tissue type based on literature values, and compressibility variations with root-mean-square magnitude equal to 1% of the nominal values were added to produce randomly distributed scattering. B-mode imaging with a 192-element linear array was simulated using a synthetic aperture method. The simulated array had a 5 MHz centre frequency, a 3 MHz  $-6$  dB bandwidth, and a one-wavelength element pitch. An image of a region without glandular tissue was synthesized with a  $3.56 \times 3.25$  cm<sup>2</sup> field of view, six transmit focal zones, and dynamic receive focusing. A control image of a statistically homogeneous fat-mimicking phantom was also synthesized using the same procedure.

### 2.3 Image analysis

Each image was partitioned in sixteen  $2.0 \times 4.5$  mm<sup>2</sup> regions of interest (ROI). The point signal-to-noise ratio (SNR), which is equal to the mean divided by the standard deviation of the echo signal magnitudes, was computed for each ROI. Wilcoxon signed-rank tests were performed to test null hypotheses that the point SNR of the fat-mimicking phantom and breast model images were equal to the theoretical value of 1.91 for Rayleigh-distributed speckle.  $P$ -values less than 0.05 were considered significant.

Empirical distribution function (EDF) tests using the Anderson-Darling  $A^2$  statistic<sup>7</sup> were also performed to evaluate the goodness of fit of the Rayleigh probability density function (PDF) to histograms of echo magnitudes in the synthesized images.  $A^2$  statistics for each ROI were compared to 1.321, which is the critical value at significance level  $\alpha = 0.05$  for an EDF test of the Weibull distribution<sup>7</sup> (the Rayleigh PDF is a limiting case of the Weibull PDF). An  $A^2$  statistic less than the critical value is considered evidence that the data were sampled from the hypothesized PDF.

## 3. RESULTS

The cross section through the breast anatomy model and the resultant synthetic B-mode image are shown in Fig. 1. Corresponding features are labelled in both panels.

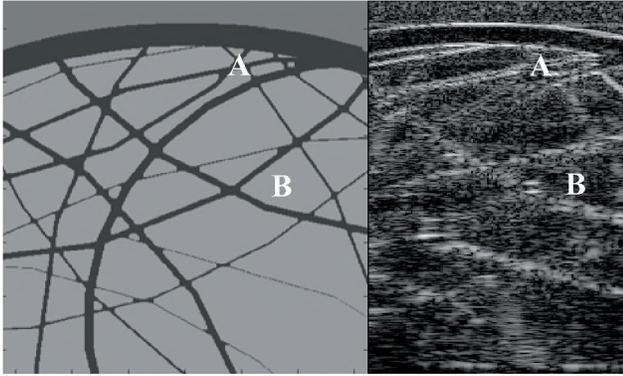


Fig. 1. Cross section from the breast anatomy model (left panel) and synthesized B-mode image (right panel) displayed using a 40-dB logarithmic grey scale. Dark pixels in the model represent skin and connective tissue, and lighter pixels in the interior of the model represent fat.

The point SNR of the breast model image was  $1.50 \pm 0.30$  (mean  $\pm$  standard deviation over 16 ROI) and was significantly less than the theoretical value of 1.91 for Rayleigh-distributed speckle ( $p < 0.001$ ). The point SNR in the control image of the fat-mimicking phantom was  $1.89 \pm 0.06$ , which was not significantly different from 1.91 ( $p = 0.20$ ).

Figure 2 summarizes the  $A^2$  statistics obtained for each ROI in the two images. In the breast model image,  $A^2$  was less than the  $\alpha = 0.05$  critical value of 1.321 in only two out of 16 ROI. For comparison,  $A^2$  was less than the critical value in eight out of 16 ROI in the control image.

#### 4. DISCUSSION

Fully developed Rayleigh speckle arises when a medium contains a large number of statistically homogeneous scatterers (*i.e.*, at least ten scatterers per resolution cell) that are randomly and uniformly positioned throughout the field of view, whereas a medium that violates one or more of these conditions produces non-Rayleigh speckle. The only scattering sites in the fat-mimicking phantom are random compressibility variations on the scale of the spatial sampling. The breast model also produces stronger scattering from the interfaces between the fat and the connective tissue septa, which should alter the speckle statistics of the resulting image.

The expected products of the simulations were a Rayleigh-distributed control image and a non-Rayleigh distributed image from the breast anatomy model, so the point SNR data are encouraging. The point SNR significantly less than 1.91 obtained in the breast model image is a characteristic of the speckle distributions observed in clinical ultrasound breast images.

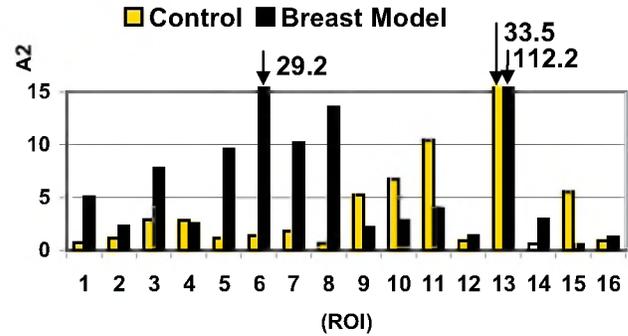


Fig. 2.  $A^2$  goodness-of-fit statistics comparing the echo amplitude distribution in each ROI of the breast model and control images to the Rayleigh PDF. Speckle in an ROI is considered Rayleigh distributed at significance level 0.05 if  $A^2$  is less than the critical value of 1.321.

The EDF test results provide further evidence that the breast model produced non-Rayleigh speckle. However, the observation that speckle in the control image was not consistently Rayleigh distributed indicates a need for further characterization of the image synthesis algorithm. The results may be improved by inclusion of additional realizations of both types of image. In addition, implementation of a goodness of fit test for the  $K$  distribution that describes clinical breast images<sup>3</sup> is desirable.

#### REFERENCES

- <sup>1</sup>M. Tanter, J.-F. Aubry, J. Gerber, J.-L. Thomas, and M. Fink, *J. Acoust. Soc. Am.* **110**, 37-47 (2001).
- <sup>2</sup>F. Lin and R. C. Waag, *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **49**, 739-755 (2002).
- <sup>3</sup>R. C. Molthen, P. M. Shankar, J. M. Reid, F. Forsberg, E. J. Halpern, C. W. Piccoli, and B. B. Goldberg, *Ultrasound Med. Biol.* **24**, 93-100 (1998).
- <sup>4</sup>P. R. Bakic, M. Albert, D. Brzakovic, and A. D. Maidment, *Med. Phys.* **29**, 2131-2139 (2002).
- <sup>5</sup>K. Bliznakova, Z. Bliznakov, V. Bravou, Z. Kolitsi, and N. Pallikarakis, *Phys. Med. Biol.* **48**, 3699-3719 (2003).
- <sup>6</sup>M. Tabei, T. D. Mast, and R. C. Waag, *J. Acoust. Soc. Am.* **111**, 53-63 (2002).
- <sup>7</sup>R. B. D'Agostino and M. A. Stephens, *Goodness-of-Fit Techniques* (Marcel Dekker, Inc., New York, 1986), Chap. 4.

#### ACKNOWLEDGEMENTS

The authors thank Dr. Makoto Tabei, Dr. T. Douglas Mast, and Prof. Robert C. Waag for sharing the source code for the  $k$ -space propagation simulator. This research was supported by NSERC Discovery Grant 261323-03.

# DEVELOPMENT OF PROTOTYPE PHASED ARRAY ULTRASOUND SYSTEMS FOR HYPERTHERMIA AND TARGETED DRUG DELIVERY

Don Chorman and Robert J. McGough

Dept. of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, USA, 48824

## 1. INTRODUCTION

Thermal therapy utilizes heat to kill malignant cancer cells. Thermal therapies generally fit into one of two general categories, namely 1) moderate temperature hyperthermia or 2) high temperature thermal ablation [1]. Moderate temperature hyperthermia is characterized by the use of elevated temperatures in the range of 42-45 degrees Celsius, with treatment times of 15-60 minutes [1]. Moderate temperature hyperthermia produces cytotoxic effects in malignant tissue when used in multiple treatments. High temperature thermal ablation is characterized by the use of temperatures greater than 60 degrees Celsius, with treatment times of 4-6 minutes [1].

Two modalities that can be utilized to noninvasively generate moderate temperature hyperthermia in malignant tissue are ultrasound (US) and radiofrequency (RF) electromagnetic radiation. High power phased array ultrasound systems can be used to generate and direct both moderate temperature hyperthermia and high temperature thermal ablation.

In addition to creating hyperthermic effects, high intensity US can also generate cavitation induced cellular lysis that is useful for destroying malignant tissue [2, 3]. The effects of hyperthermia and cavitation can also be utilized as activation mechanisms in targeted drug delivery [4, 5]. Targeted drug delivery is a noninvasive form of treatment, which results in drug activation that is limited to a prescribed region of the body for a controlled amount of time.

The phased array ultrasound systems described in this research usually consist of five components: (1) an array of ultrasound transducers, (2) electrical interconnects (3) resonant filters, (4) driving electronics, and (5) computer control software. The ultrasound prototype applicators described in this paper are 1D arrays constructed from Lead Zirconate Titanate (PZT) elements. The electrical interconnect serves as the electrical connection for the filtered output of the power electronics. The resonant filters pass only the fundamental frequency component in the square-wave output of the power electronics. The resonant filters also provide a voltage gain 3-10 times the input voltage. The driving electronics consist of a digital signal generator and Class-D power amplifiers. The digital signal

generator boards utilize field programmable gate arrays to generate low-level (0-3.3V) frequency and phasing signals. Class-D power amplifiers amplify these low-level signals. A personal computer is interfaced with the power electronics to control the frequency, phase, and mode of operation. Through software control, either continuous-wave (CW) or pulsed output signals are produced by the driving electronics.

Commercially available, custom built high power ultrasound arrays and the associated power electronics for use with hyperthermia are expensive and inflexible when considering operational characteristics such as the number of transducers, number of electrical channels, frequency of operation, modes of operation, and phasing. These issues motivate this research, which generates robust and operationally flexible phased array ultrasound systems that are cost-effective for thermal therapy applications.

## 2. METHODS

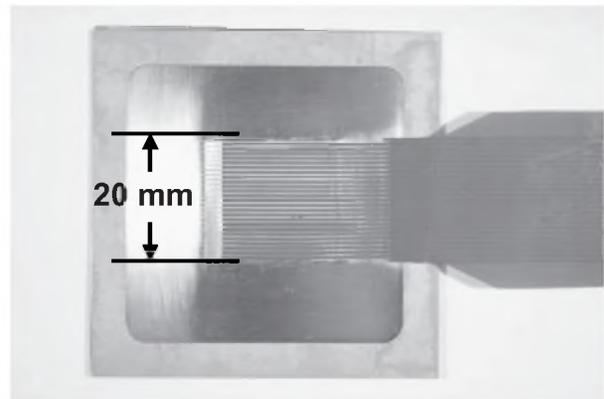


Fig. 1: 1D ultrasound array with a flex circuit interconnect.

### 2.1 Ultrasound Array and Electrical Interconnect

The one-dimensional prototype ultrasound arrays were constructed from PZT (Navy Type III) and double-sided copper-clad Pyralux baffles. For a single array, a PZT plate was bonded to a baffle and a flexible interconnect circuit was bonded to the top of the PZT plate. A completed array is shown below in Figure 1.

The adhesive used for bonding was Cho-Bond 584 silver electrically conductive epoxy. After curing the epoxy

bonds, a wafer-dicing saw was used to cut the PZT plate into individual elements, forming a 1D ultrasound phased array with an attached electrical interconnect. The flex circuit interconnect was made of single-sided copper-clad Pyralux. The interconnect circuit was designed using Orcad Capture and Layout software.

## 2.2 Resonant Filters

The resonant filters were comprised of capacitors and hand-wound toroidal inductors. These form a low-pass filter that has the same electrical resonant frequency as the elements in the array.

## 2.3 Driving Electronics

The digital signal generator boards use XILINX Spartan field programmable gate arrays to generate low-level (0-3.3V) frequency and phasing signals. These boards were programmed using the XILINX ISE Webpack software. The digital signal generator operates in continuous-wave or pulsed mode with a frequency Range of 50kHz-20MHz. The phase resolution at 1MHz is 0.6 degrees. This design can be expanded to synthesize 1024 channel signals. The Class-D power amplifiers are from a previous research project. The amplifiers are capable of driving 20 watts per channel, with a frequency range of 500-1000 kHz.

## 2.4 Computer Control

A personal computer controls the digital signal generator through a Matlab interface. A simple script adjusts the frequency, phase, and mode (CW or pulsed) of operation.

## 3. RESULTS

The system was tested after the ultrasound phased array was mounted in a water tank. The computer-controlled power electronics steer a focus at specified locations within the tank. The resulting focal patterns were mapped using a computer controlled positioning system and a hydrophone. The hydrophone was traced through a rectilinear grid, measuring the pressure generated by the ultrasound. Figures 2 and 3 show two experimentally measured pressure fields.

## 4. DISCUSSION

The experimentally measured pressure fields demonstrate that the 1D ultrasound phased array is capable of generating and steering a focus within a water tank. The shape of the measured field in the direction of the array sampling is similar to the results obtained from computer simulations; however, measurements and simulations are dissimilar along the long axis of the array elements. This suggests that the particle velocity is not uniform across each active element, so more detailed computer models are required. This is confirmed in comparisons of simulations and measurements of single element pressure fields.

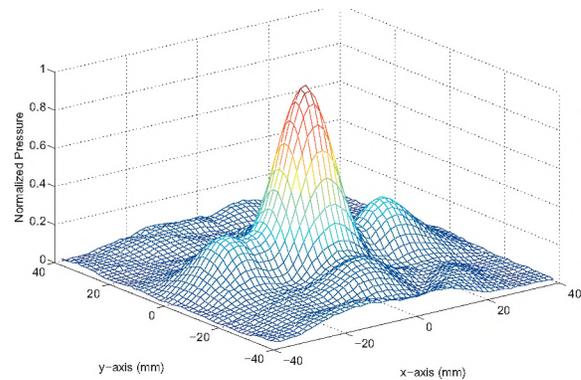


Fig. 2: Normalized pressure field generated by the 22-element 1D array shown in Figure 1. This array, which operates at 971.5kHz, is focused on-axis at  $[0, 0, 304.8]$  (mm).

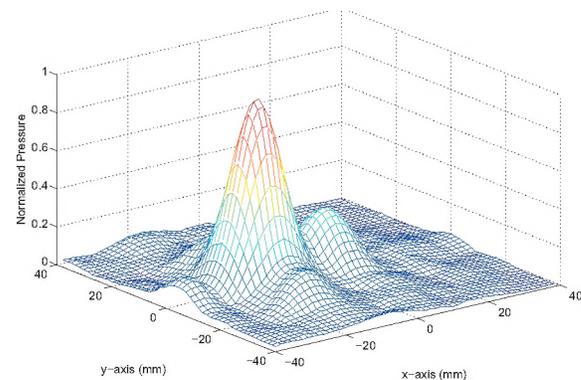


Fig. 3: Normalized pressure fields generated by the 22-element 1D array shown in Figure 1. This array, which operates at 971.5kHz, is steered off-axis with a focus located at  $[-15, 0, 304.8]$  (mm).

## 5. REFERENCES

- 1) P.R. Stauffer, S.N. Goldberg. Introduction: Thermal Ablation Therapy, *International Journal of Hyperthermia*, 2004; 20(7): 671-677
- 2) B.C. Tran, Seo Jongbum, T.L. Hall, J.B. Fowlkes, and C.A. Cain. Microbubble enhanced cavitation for noninvasive ultrasound surgery. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, October 2003; 50(10): 1296-1304
- 3) F.J. Fry, N.T. Sanghvi, R.S. Foster, R. Bihrl, and C. Hennige. Ultrasound and microbubbles – Their generation, detection, and potential utilization. *Ultrasound in Medicine and Biology*, 1995; 21(5): 1227-1237
- 4) Neru Munshia, Natalya Rapoport and William G. Pitt. Ultrasonic activated drug delivery from Pluronic P-105 micelles. *Cancer Letters*, September 1997; 118(1):13-19
- 5) Jared L. Nelson, Beverly L. Roeder, John C. Carmen, Friederike Roloff and William G. Pitt, Ultrasonically Activated Chemotherapeutic Drug Delivery in a Rat Model. *Cancer Research*, December 2002; 62: 7280-7283

## ACKNOWLEDGEMENTS

The authors would like to thank F. Foust from the Department of Electrical and Computer Engineering at Michigan State University. This work was supported in part by NIH R01CA093669.

# MODE SCANNING FOR ULTRASOUND PHASED ARRAYS WITH SIX PLANES OF SYMMETRY

Duo Chen, and Robert J. McGough

Dept. of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, 48824

## 1. ABSTRACT

Modes are symmetric focal patterns that exploit the symmetry of ultrasound phased arrays and cancel the complex pressure field in multiple planes that contain the central axis. The mode scanning approach, which was originally derived for symmetric planar arrays populated with rectangular elements, is extended to include triangular and hexagonal elements. This focusing method is very useful for generating heating patterns in hyperthermia. An example of mode scanning with 6-fold symmetry shown for a flat, two-dimensional ultrasound phased array populated with equilateral triangles. The pressure field of a single triangular element is computed with the fast nearfield method, which generates pressure fields with relatively small errors. The derivation of the element driving signals that produce different model focal patterns for any mode combines the characteristic of geometric symmetry with the method of gain maximization.

## 2. INTRODUCTION

Various scanning techniques generate useful heating patterns for hyperthermia with ultrasound phased array applicators. Spot scanning, which is effective for small and superficial tumors, is a focusing approach that is often employed, but for large and deep tumors this method generally produces axial hot spots [1]. Another technique is mode scanning [2], which has been demonstrated using a square array that produces a four-focus mode. Mode scanning cancels axial pressure fields and therefore effectively heats larger tumors. This paper demonstrates the results of mode scanning with 6-fold symmetry shown for a flat, two-dimensional ultrasound phased array populated with equilateral triangles.

## 3. METHODS

### 3.1 Mode definition

Modes are symmetric focal patterns that exploit the symmetry of ultrasound phased arrays and cancel the complex pressure field in multiple planes that contain the central axis.

### 3.2 Excitation scheme

The gain maximization method [3] generates useful driving signals for multiple focus patterns. The complex focal point pressures,  $\mathbf{p}$ , relate to the complex

column vector of array element surface particle velocities,  $\mathbf{u}$ , through the matrix relation

$$\mathbf{p} = \mathbf{H}\mathbf{u}, \quad (1)$$

where  $\mathbf{H}$  is a complex forward propagation matrix that describes the transfer relationship between all array elements and all distinct focal points. With the gain maximization method, the normalized intensity gain,  $\mathbf{G} = \langle \mathbf{p}, \mathbf{p} \rangle / \langle \mathbf{u}, \mathbf{u} \rangle$ , is maximized in order to minimize the total energy needed by the phased array to achieve given focal intensity values. For a single focus, a phase-only excitation scheme adopted from the gain maximization method computes the phase vector through the formula

$$\theta = \arg(H^*) + \arg(p), \quad (2)$$

and the amplitudes of the array elements are determined through the formula

$$A = \frac{|p|}{\langle h, \exp(j\theta) \rangle}. \quad (3)$$

### 3.3 Mode derivation

The derivation of the element driving signals that produce different focal patterns for any mode combines the characteristic of geometric symmetry with the method of gain maximization. The symmetric element arrangement of the array corresponding to six foci is shown in Fig. 1. Using this ordering scheme, the partitioned complex excitation vector can then be written as

$$\begin{aligned} \mathbf{u} &= [\tilde{u}^1 \tilde{u}^2 \tilde{u}^3 \tilde{u}^4 \tilde{u}^5 \tilde{u}^6]^T \\ &= [\tilde{u}^1 \tilde{u}^1 e^{j\pi} \tilde{u}^1 \tilde{u}^1 e^{j\pi} \tilde{u}^1 \tilde{u}^1 e^{j\pi}]^T \end{aligned} \quad (4)$$

In order to match the partitioning of the input excitation  $\mathbf{u}$ , the forward propagation vector can be separated as

$$\mathbf{h} = [\tilde{h}^1 \tilde{h}^2 \tilde{h}^3 \tilde{h}^4 \tilde{h}^5 \tilde{h}^6], \quad (5)$$

so equation (1) can be rewritten as

$$p = \tilde{h}^1 \tilde{u}^1 + \tilde{h}^2 \tilde{u}^2 + \tilde{h}^3 \tilde{u}^3 + \tilde{h}^4 \tilde{u}^4 + \tilde{h}^5 \tilde{u}^5 + \tilde{h}^6 \tilde{u}^6. \quad (6)$$

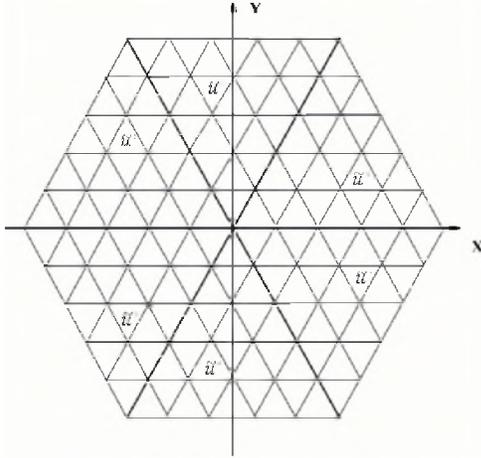


Fig. 1: Flat planar 2D ultrasound phased array comprised of triangular pistons.

For six plane symmetric foci, the positions of the six focal points mimic the symmetry of the element ordering scheme. The six complex pressures are expressed as  $p_1, p_2, p_3, p_4, p_5,$  and  $p_6$ . All of these relationships are combined into a single matrix expression according to

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix} = \begin{bmatrix} \tilde{h}^1 & \tilde{h}^2 & \tilde{h}^3 & \tilde{h}^4 & \tilde{h}^5 & \tilde{h}^6 \\ \tilde{h}^2 & \tilde{h}^1 & \tilde{h}^6 & \tilde{h}^3 & \tilde{h}^4 & \tilde{h}^3 \\ \tilde{h}^3 & \tilde{h}^6 & \tilde{h}^1 & \tilde{h}^2 & \tilde{h}^3 & \tilde{h}^4 \\ \tilde{h}^4 & \tilde{h}^3 & \tilde{h}^2 & \tilde{h}^1 & \tilde{h}^6 & \tilde{h}^5 \\ \tilde{h}^5 & \tilde{h}^4 & \tilde{h}^5 & \tilde{h}^6 & \tilde{h}^1 & \tilde{h}^2 \\ \tilde{h}^6 & \tilde{h}^5 & \tilde{h}^4 & \tilde{h}^3 & \tilde{h}^2 & \tilde{h}^1 \end{bmatrix} \begin{bmatrix} \tilde{u}^1 \\ \tilde{u}^2 \\ \tilde{u}^3 \\ \tilde{u}^4 \\ \tilde{u}^5 \\ \tilde{u}^6 \end{bmatrix} \quad (7)$$

When the amplitude at each focus is equal, that is,  $|p_1| = |p_2| = |p_3| = |p_4| = |p_5| = |p_6|$ , then

$$p_1 = p_2 e^{j\pi} = p_3 = p_4 e^{j\pi} = p_5 = p_6 e^{j\pi} \\ = [\tilde{h}^1 + e^{j\pi} \tilde{h}^2 + \tilde{h}^3 + e^{j\pi} \tilde{h}^4 + \tilde{h}^5 + e^{j\pi} \tilde{h}^6] \tilde{u}^1. \quad (8)$$

Thus, the forward propagation matrix H of a six focus mode can be reduced to a single vector,

$$\tilde{h} = [\tilde{h}^1 - \tilde{h}^2 + \tilde{h}^3 - \tilde{h}^4 + \tilde{h}^5 - \tilde{h}^6] \quad (9)$$

In order to obtain the driving signals for the entire array, the  $\tilde{h}$  vector is computed first using equation (9), and then equation (2) computes the excitation  $\tilde{u}^1$ .

## 4. RESULTS

Mode scanning simulations are evaluated for a flat, two-dimensional ultrasound phased array populated with 150 equilateral triangles. The length of each element is 1 wavelength. The pressure field of a single triangular element is computed with the fast nearfield method [4][5], which generates pressure fields with relatively small errors.

Figure 2 shows the pressure field generated by this array of triangular sources in the focal plane. The focal plane is 80 wavelengths from the surface of the phased array. Fig. 2 contains six symmetric foci and three planes of symmetry in which the pressure field cancels perfectly.

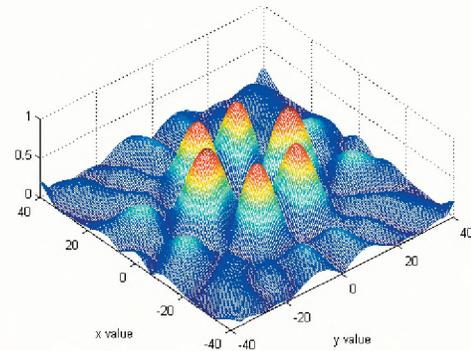


Fig. 2: Pressure field computed in the focal plane with mode scanning.

## 5. CONCLUSION

These results show that, with the combination of geometric symmetry and the method of gain maximization, element driving signals can be selected to generate symmetric patterns that cancel the pressure field in symmetric planes that intersect along the array normal and therefore eliminate axial hot spots.

## REFERENCES

- [1] Moros E., Roemer R. and Hynynen, K., 1990, Pre-focal plane high-temperature regions induced by scanning focused ultrasound beams. *International Journal of Hyperthermia*, 6, 351-366.
- [2] McGough, R. J., Wang, H., Ebbini, E. S. and Cain, C. A. 1994. Mode scanning: heating pattern synthesis with ultrasound phased arrays. *International Journal of Hyperthermia*, 3, 433-442.
- [3] Ebbini, E. S. and Cain, C. A., 1991, Optimization of the intensity gain of multiple-focus phased array heating patterns. *International Journal of Hyperthermia*, 7, 953-973.
- [4] McGough, R. J., Samulski, T. V., and Kelly, J. F. 2004. An efficient grid sectoring method for calculations of the nearfield pressure generated by a circular piston. *J. Acoust. Soc. Am.* 115 (5), 1942-1954.
- [5] McGough, R. J. 2004, "Rapid calculations of time-harmonic nearfield pressures produced by rectangular pistons," *J. Acoust. Soc. Am.*, 115(5).

# COMPUTER MODELING OF A HYBRID RF/US PHASED ARRAY SYSTEM FOR HYPERTHERMIA CANCER TREATMENTS IN THE INTACT BREAST

Liyong Wu, Xiaozheng Zeng, and Robert J. McGough

Dept. of Electrical Computer Engineering, Michigan State University, East Lansing, MI, USA, 48823  
wuliyong@msu.edu

## 1. INTRODUCTION

Among women living in the United States, breast cancer is the most common malignancy diagnosed and the second leading cause of death from cancer [1]. Better therapies are clearly needed for breast cancer, which motivates the development of new adjuvant hyperthermia treatments. External heating devices appropriate for deep hyperthermia in the intact breast include ultrasound phased arrays [2] and radio-frequency (RF) electromagnetic phased arrays [3]. Ultrasound (US) is an appropriate local modality for heating small targets in the breast (up to about 2cm diameter [4]), whereas heat generated by RF electromagnetic devices is delivered regionally across a much larger area. Ultrasound phased arrays have been evaluated for ablation of fibroadenomas in the breast [2], and RF phased arrays have been developed previously for deep hyperthermia in the pelvis [3] and in the extremities [5], respectively.

When these two modalities are combined, significantly improved temperature distributions are observed in simulated hyperthermia treatments of locally advanced breast cancer (LABC). Power distributions are simulated for a hybrid applicator consisting of a planar square US phased array and an RF phased array, and temperature responses are evaluated for hyperthermia treatments in the intact breast.

## 2. METHODS

### 2.1 Applicator and patient model

A 16.2cm wide and 16.2cm high 2D 1MHz US phased array is placed on one side of a water tank that defines the patient interface. This US phased array consists of 5329 circular sources, where each source has a diameter of one wavelength with  $\lambda/2$  spacing between adjacent elements. The RF phased array, which operates at 140 MHz, consists of four end-loaded dipole antennas mounted on a Lexan water tank. The thickness of the lexan tank is about 0.5 cm. To facilitate coupling of US and RF into the patient model, the tank is filled with deionized water.

The 3D patient model, including breast, tumor, chest wall, heart, and lung structures, are extracted from MR or CT images. The contours that define these structures are manually extracted from patient images with a Matlab based program and reconstructed in the finite element modeling software package HFSS (Ansoft Corp. Pittsburgh, PA).

### 2.2 Fast Near Field method US pressure computations

The pressure field generated by a single source is calculated with the Fast Near field Method (FNM) [4]. The mode scanning technique [6] generates 2 or 4 focal spots while canceling the fields in some planes to reduce unwanted heat accumulated between the array and the tumor region. In order to reduce intervening tissue heating, focal spots are placed at the distal half of the tumor. The power deposition from each focal pattern is computed and multiplied by a weighting factor.

### 2.3 E field simulation and optimization

An edge based 3D finite element (FE) model calculates the electric (E) field distribution inside the breast and tumor region. The permittivity and conductivity values for human tissue are from Joines [6]. The FE mesh is truncated with a radiation boundary condition. The mesh node spacing is adapted such that the electric field is densely sampled near material interfaces. The E-field, based on Maxwell's equations, is computed for each antenna with the finite element (FE) method using the commercial software HFSS.

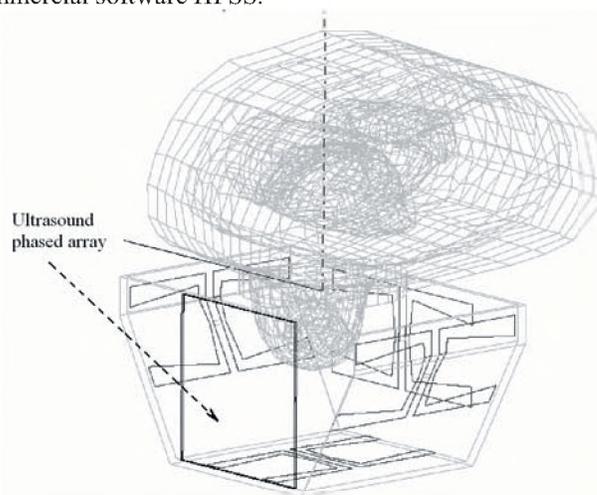


Fig 1: Hybrid applicator model and patient model.

### 2.4 Temperature calculation and optimization

The total specific absorption rate (SAR) is the sum of the SAR from the E-field and the SAR produced by the pressure field. The temperature increase is calculated from the SAR

using a finite difference formulation of the linear bio-heat transfer equation (BHTE). To obtain a more uniform temperature distribution inside the tumor, a Powell search algorithm determines the optimal relative weighting of the RF- and US-induced powers.

### 3. RESULTS

Based on the computed E-field and pressure fields, the SAR and temperature values are evaluated in a  $16\text{cm} \times 16\text{cm} \times 10\text{cm}$  volume that contains the breast and surrounding water. The fields produced by the RF and US arrays are calculated separately and then the SAR values are combined for temperature calculations. All fields are computed in 3D. Temperatures are demonstrated for individual and combined modalities in the central portion of the tumor, namely the  $xz$  plane with  $y = 5\text{mm}$ . Fig. 2 shows the temperature distribution produced by the RF phased array in this sample plane. The input power for each antenna is the same, and the phases of the four RF antennas channels are  $70^\circ$ ,  $-63^\circ$ ,  $113^\circ$ ,  $0^\circ$ . The US array likewise targets a spherical region located at  $(-2.5, 1.7, -3)(\text{cm})$  with 15 focal spots distributed in the back portion of the tumor. The temperature produced by the US array in the same plane as Fig. 2 is shown in Fig. 3.

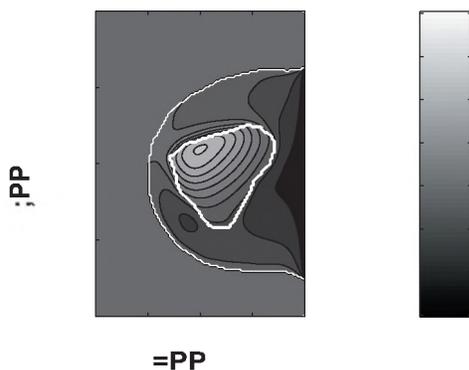


Fig 2: The temperature distribution obtained with RF alone.

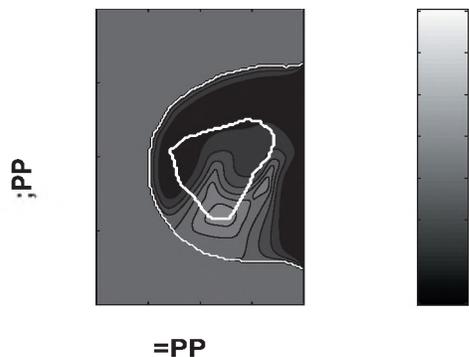


Fig 3: The temperature distribution obtained with US alone.

The total SAR of the hybrid RF/US phased array is the sum of the SAR computed from E field and the pressure field. The temperature calculated from the total hybrid SAR is shown in Fig. 4.

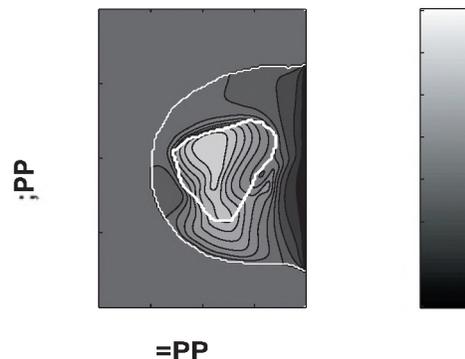


Fig 4: Temperature distribution obtained with the RF/US hybrid approach.

### 4. DISCUSSION

In the Fig. 4, the RF and US components heat separate portions of the tumor. As shown in Figs. 2-4, the hybrid method achieves a better temperature distribution inside the tumor than either modality applied alone. The hybrid method achieves this result because of the overlapping power contribution of the RF and US in the tumor, and problems with hot spots are therefore reduced.

### REFERENCES

- [1] R. T. Greenlee, T. Murray, S. Bolden, and P. A. Wingo, (2000) "Cancer statistics, 2000," *CA-Cancer J. Clin.*, vol. 50, no. 1, pp. 7-33.
- [2] K. Hynynen, O. Pomeroy, D. N. Smith, P. E. Huber, N. J. McDannold, J. Kettenbach, J. Baum, S. Singer, and F. A. Jolesz, (2001) "Mr imaging-guided focused ultrasound surgery of fibroadenomas in the breast: A feasibility study," *Radiology*, vol. 219, no. 1, pp. 176-185.
- [3] P. Wust, J. Nadobny, R. Felix, P. Deuflhard, A. Louis, and W. John, (1991) "Strategies for optimized application of annular-phased-array systems in clinical hyperthermia," *Int. J. Hypertherm.*, vol. 7, no. 1, pp. 157-173.
- [4] M. Malinen, T. Huttunen, K. Hynynen, and J. P. Kaipio, (2004) "Simulation study for thermal dose optimization in ultrasound surgery of the breast," *Med. Phys.*, vol. 31, no. 5, pp. 1296-1307.
- [5] Y. Zhang, W. T. Joines, R. L. Jirtle, and T. V. Samulski, (1993) "Theoretical and measured electric field distributions within an annular phased array: Consideration of source antennas," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 8, pp. 780-787.
- [6] W. T. Joines, Y. Zhang, C. X. Li, et al. (1994) "The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz" *Medical. Phys.* 21 (4).
- [7] R. J. McGough *et al.*, (2004) An efficient grid sectoring method for calculations of the nearfield pressure generated by a circular piston. *Journal of Acoustic Society of America*, 115(5).
- [8] R. J. McGough, H. Wang, E. S. Ebbini and C. A. Cain, (1994) Mode scanning: heating pattern synthesis with ultrasound phased arrays, *International Journal of Hyperthermia*, 10(3).

### ACKNOWLEDGEMENTS

This work was funded in part by NIH grant 1R01CA093669.

# EFFECTS OF MULTI-TALKER BACKGROUND NOISE ON THE INTENSITY OF SPOKEN SENTENCES IN PARKINSON'S DISEASE

Scott Adams, Ph.D.<sup>1</sup>, Olga Haralabous, M.Sc.<sup>1</sup>, Allyson Dykstra, M.Sc.<sup>2</sup>, Kayla Abrams, B.Sc.<sup>1</sup>, Mandar Jog, M.D.<sup>3</sup>

<sup>1</sup>School of Communication Sciences and Disorders, <sup>2</sup>Doctoral Program in Rehabilitation Sciences,

<sup>3</sup>Department of Clinical Neurological Sciences

University of Western Ontario, London, Ontario, Canada, N6G 1H1

## 1. INTRODUCTION

Hypophonia or low speech intensity is one of the most frequent and commonly treated speech symptoms in Parkinson's disease (PD) (Adams, 1997). The moment-to-moment impact of hypophonia on communication is often observed to be dramatically influenced by the intensity of the surrounding background noise. In general, the louder the background noise the more difficult it is for the person with hypophonia to communicate. Unfortunately, the relationship between speech intensity and background noise level has rarely been systematically examined in previous studies of hypophonia. A preliminary report, by Adams and Lang (1992), found that 90 dB SPL of white noise produced a marked increase in speech intensity in 10 PD subjects. In contrast, Ho, Bradshaw, Iansek and Alfredson (1999), found that pink noise, presented at 10-30 dB above threshold, produced minimal or no increase in speech intensity in a group of 12 PDs. These inconsistencies may be related to a number of factors such as, the severity of hypophonia, the type background noise, the noise levels, the speech tasks, the intensity measures, and the methods of stimulus presentation. Defining the relationship between speech intensity and background noise has important implications for the understanding, assessment and treatment of hypophonia in PD. The purpose of the present study was to attempt to define the relationship between speech intensity and background noise in individuals with hypophonia and Parkinson's disease.

## 2. METHOD

This study included 10 idiopathic PD subjects (8 male, 2 female), (64-78 years, M=69, SD=5.3) with hypophonia and 10 age-equivalent controls (9 male, 1 female), (55-80 years, M=73, SD=4.4). All subjects with PD were reported by a Neurologist (MJ) to demonstrate reduced speech intensity or hypophonia. All subjects with PD were stabilized on their anti-parkinsonian medication and were tested at approximately 1 hour after taking their regularly scheduled anti-parkinsonian medication. Normal and PD subjects passed a 40 dB hearing screening.

All subjects were tested in an audiometric booth. During all conditions, subjects sat in a chair facing a wall of the audiometric booth. A loudspeaker was placed 72 inches in

front of the subjects. Subjects wore a headset microphone (AKG-C420) positioned a constant 6 cm distance from the mouth. The experimenter presented a standard tape-recording of multi-talker noise (Audiotech – 4 talker noise) through the loudspeaker, adjusting the dB level of the noise via a diagnostic audiometer (GSI 61). The speech of each subject was recorded using a digital audio tape recorder (Tascam DA-01). Subjects were required to repeat the sentences "I owe you a yo-yo. I owe you a yo-yo" (Goldinger, Pisoni, & Luce, 1996) in each noise condition. The sentence "I owe you a yo-yo" was selected because all segments within this sentence are voiced and therefore the sentence is fairly easy to segment and it has a fairly constant intensity contour.

The multi-talker noise was presented randomly to each subject in five dB increments ranging from 50-70 dB. Each dB level was presented twice in the following order: 50, 65, 60, 70, 55 dB then 55, 70, 60, 65, 50 dB. The subjects' recorded test sentences were digitized using Kay Elemetrics' Visipitch program. The average intensity (dB) for each test sentence was determined using the Visipitch intensity analysis routine. A two-factor repeated measures ANOVA was used in the statistical analysis.

## 3. RESULTS

The results for the PD and control groups are shown in Figure 1. Both PDs and controls showed a significant increase in speech intensity across increases in the level of background noise ( $p=.0001$ ). PD subjects were approximately 2-3 dB lower than controls across all noise levels. This PD versus control group difference approached significance ( $p=.065$ ). Hypophonic PD subjects appeared to show a speech intensity versus background noise relationship that was parallel to the controls but at a consistently lower speech intensity (see regression lines in Figure 1). The speech intensity versus background noise regression lines for each of the PD subjects are shown in Figure 2. These regression lines all show a positive slope. Interestingly, the most severe hypophonic PD can be seen (Figure 2) to have a fairly strong positive slope. Figure 2 suggests that there is no relationship between the severity of the subject's hypophonia and the slope of the regression line.

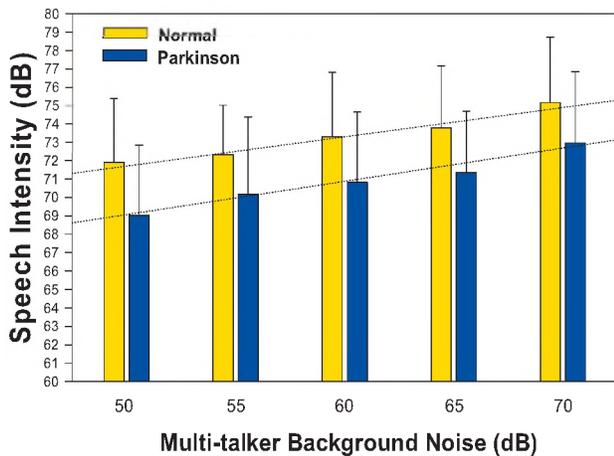


Fig. 1. Average speech intensity obtained from Parkinson disease and control subjects during 5 levels of multi-talker noise (50, 55, 60, 65, 70 dB). Corresponding regression lines are shown for each group.

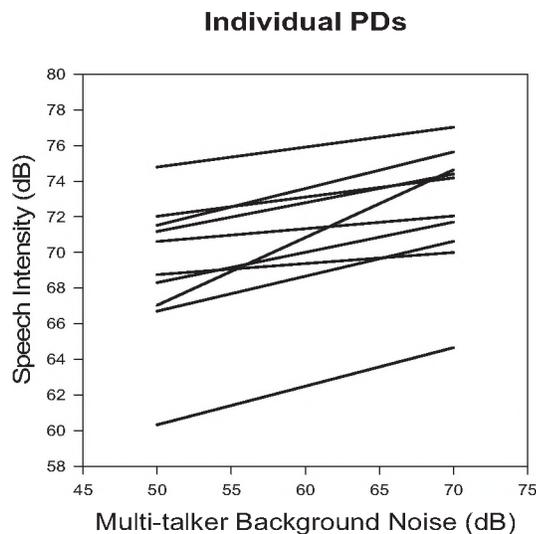


Fig. 2. Individual speech intensity versus background noise regression lines for the subjects with Parkinson's disease (PD).

#### 4. DISCUSSION

The results of this study suggest that subjects with PD show a systematic increase in speech intensity across increases in background noise. These findings agree with the first published demonstration of the "Lombard effect" in PD (Adams & Lang, 1992). The present study also showed that the positive relationship between speech intensity and background noise is approximately parallel to that of controls. Interestingly, despite this positive Lombard relationship, the PD subjects' speech intensity was consistently below that of the controls for each of the noise levels examined. Thus, relative to controls, the PD subjects

showed a parallel but reduced speech versus noise intensity relationship.

The results of the present study are in contrast to one previous study by Ho et al. (1999) that failed to observe a strong positive Lombard relationship in their PD subjects. When comparing the methods of the Ho et al. (1999) study to those of Adams and Lang (1992) and to the present study, the main inconsistency appears to be related to the level of the background noise that was used. The Ho et al. (1999) study appears to have used noise levels that were below 50 dB while the present study and the Adams & Lang (1992) study used noise levels that were above 50 dB SPL. In an earlier, study of the Lombard effect in normals, Lane and Travel (1971) warned that it may be difficult to demonstrate a strong and consistent Lombard effect at very low levels of background noise. It is suggested that this difficulty may be even more pronounced in hypophonic subjects. Future studies involving a wider range of background levels may be required to resolve this issue.

The present study highlights the potentially powerful and positive effects of background noise on speech intensity production in hypophonia. Future studies are required to determine if the Lombard effect can be incorporated into novel methods of treatment for individuals with PD and hypophonia.

#### REFERENCES

- Adams, S.G. (1997). Hypokinetic dysarthria in Parkinson's disease. In M.R. McNeil (Ed). *Clinical Management of Sensorimotor Speech Disorders* (p. 261-285). NY: Thieme Pub.
- Adams, S.G. & Lang, A.E. (1992). Can the Lombard effect be used to improve low voice intensity in Parkinson's disease? *European Journal of Disorders of Communication*, 27, 121-127.
- Goldinger, S.D., Pisoni, D.B., & Luce, P.A. (1996). Speech perception and word recognition: research and theory. In Lass, N.J. (Ed.), *Principles of Experimental Phonetics* (p. 277-327). St. Louis: Mosby.
- Ho, A., Bradshaw, J., Ianssek, R., & Alfredson, R. (1999). Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*, 37, 1453-1460.
- Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 14, 677-709.

#### ACKNOWLEDGEMENTS

This research was funded by a grant from the University of Western Ontario's Academic Development Fund that was awarded to the first author.

#### AUTHOR NOTES

This work was conducted while Olga Haralabous was completing her M.Sc. Degree in Communication Sciences and Disorders (Speech-Language Pathology) at the University of Western Ontario.

# THE IMPACT OF EMOTIONAL PROSODY ON NASALANCE SCORES

Tim Bressmann and Jayanthi Sasisekaran

Dept. of Speech-Language Pathology, University of Toronto, 500 University Avenue, Ontario, Canada, M5G1V7  
tim.bressmann@utoronto.ca

## 1. INTRODUCTION

Nasalance is a commonly used acoustic measure that allows a speech-language pathologist to validate and quantify the perceptual assessment of a speaker's nasal resonance. It is particularly useful for the assessment of the hyper- or hyponasal resonance disorders that are often associated with conditions such as cleft palate or neurogenic dysarthria. The nasalance score is calculated as a ratio of the nasal sound pressure level to the combined nasal and oral sound pressure level (Fletcher, 1978). There are currently three commercially available instruments that measure nasalance: the Kay Nasometer 6200/ 6300 (Kay Elemetrics, Lincoln Park NJ), the NasalView (Tiger Electronics, Seattle WA), and the OroNasal System (Glottal Enterprises Inc., Syracuse NY).

Researchers and clinicians have noted that there is some variability between repeated nasalance measurements in the same patients (Bressmann, 2005). Watterson et al. (1994) studied the influence of speaker loudness on nasalance values. While the authors did not find significant differences in nasalance magnitudes for different loudness conditions they noted that nasalance scores tended to be lower when the speakers used a louder voice. Lewis et al. (2000) demonstrated that the vowel content in test sentences can significantly impact the nasalance magnitudes.

A factor that has not been investigated but that may contribute to the variability of nasalance scores is the emotional prosody of the speaker during an utterance. The purpose of the present study was to investigate the impact of different types of emotional prosody on nasalance scores. It was our hypothesis that different emotions would be associated with consistently different nasalance scores. As a co-variable, we investigated the impact of emotional prosody on different phonetic materials. Nasalance testing for hypernasality is usually carried out with non-nasal test sentences because these materials have a higher diagnostic value (Dalston & Seaver, 1992). We hypothesized that the impact of emotional prosody would be more pronounced for phonetically balanced materials because they allow a higher level of vocal plasticity than non-nasal stimuli.

Previous studies on emotional prosody have employed professional or amateur actors (Banse & Scherer, 1996). The advantage of using actors is that these individuals are trained to portray different emotional states on command.

We therefore decided to conduct our experiment with a group of speakers who had acting experience.

## 2. METHODS

The participants were 6 females and 4 males with a median age of 29 years (SD 7.37 years, range 18-42 years), who were recruited from a university drama club. All recordings were made with the NasalView System in the same quiet office environment. The test items were recorded directly to hard disk at a sampling rate of 44.1 kHz stereo (16 bit). The noise filter threshold was set to 18dB to reduce the effect of ambient noise on the nasalance recordings. The mean nasalance values and the corresponding standard deviations were recorded from the computer screen. Two sentences were used as stimuli for this study. The sentence 'Hat sundig pron you venzy' was taken from Banse & Scherer (1996). This sentence contains syllable structures that may be found in several indogermanic languages. It is a nonsense sentence, which has been found to be advantageous for the portrayal of emotions. The second sentence that we used was the sentence 'He had two rock lizards'. This sentence is balanced for vowel content, contains no nasal sounds and has been used in previous research (Lewis et al., 2000). The emotional states that were portrayed by the speakers were hot anger, sadness, happiness, interest, boredom, and contempt, as well as neutral intonation. These particular emotions were chosen because Banse & Scherer (1996) demonstrated that they had the most salient perceptual cues for listeners and could be portrayed most convincingly by the actors. The order of presentation was randomized.

## 3. RESULTS

The participants read two repetitions of the two sentences in six portrayed emotions and in a neutral intonation, resulting in a total of 28 recordings for each participant. We calculated mean values for the two repetitions of each token. The nasalance scores for all sentences in the different emotional states can be found in Table 1 and in Figures 1 and 2. For the non-nasal sentence 'He had two rock lizards', there was not much variability between the different portrayed emotions. The nonsense sentence 'Hat sundig pron you venzy' showed more variability. The highest nasalance values were observed for the 'neutral' and the 'sad' conditions, the lowest values were found for 'anger' and 'happiness'. In order to determine whether any of the

observed changes in nasalance scores were significant, we calculated two series of t-tests between the neutral condition for each sentence and the corresponding six portrayed emotions. The alpha was Bonferroni-adjusted accordingly to  $p < 0.008$ . We did not find any significant differences between the neutral and the emotional conditions for the sentence 'He had two rock lizards'. For the sentence 'Hat sundig pron you venzy', we found significantly lower nasalance values for the emotions hot anger ( $p < 0.002$ ), happiness ( $p < 0.005$ ) and interest ( $p < 0.002$ ).

Tab. 1. Mean nasalance scores and standard deviations for the sentences 'He had two rock lizards' and 'Hat sundig pron you venzy', spoken with different portrayed emotions.

	'He had two rock lizards'	'Hat sundig pron you venzy'
Neutral	Mean 22.0 SD 3.11	Mean 46.34 SD 2.76
Hot anger	Mean 23.68 SD 2.24	Mean 37.79 SD 3.09
Sadness	Mean 24.25 SD 5.72	Mean 45.86 SD 8.57
Happiness	Mean 22.95 SD 2.24	Mean 38.79 SD 4.96
Interest	Mean 22.32 SD 3.42	Mean 39.78 SD 5.07
Boredom	Mean 22.85 SD 5.11	Mean 44.98 SD 6.23
Contempt	Mean 24.70 SD 4.21	Mean 43.21 SD 6.40

#### 4. DISCUSSION

The amateur actors who participated in this study were able to portray the different emotions with relatively little variability in nasalance values across repeated measurements. The nasalance magnitudes for the non-nasal sentence 'He had two rock lizards' demonstrated that different portrayed emotions did not affect the nasalance values. The level of variability was higher for the nonsense sentence 'Hat sundig pron you venzy'. The paired t-tests demonstrated that the portrayed emotions 'anger', 'happiness' and 'interest' had significantly lower nasalance values. A number of subjects shared the feedback that they found it considerably easier to portray the different emotions when they were reading the nonsense sentence 'Hat sundig pron you venzy'. It is possible that the vocal and articulatory adjustments that are made in the vocal tract to convey an emotional state also affect the oral-nasal balance of the speaker. The conclusion drawn from the present study is that some types of affective prosody can impact on nasalance magnitudes. However, this impact is more pronounced for phonetically balanced materials than it is for non-nasal stimuli. This is good news for clinicians because diagnostic materials usually consist of non-nasal phonetic materials (Dalston & Seaver, 1992). When recording phonetically balanced materials, it is important to remember that affective prosody can confound the nasalance measurements. In order to obtain reliable nasalance

recordings, subjects should be instructed to read or repeat passages in an intonation that is as emotionally neutral as possible.

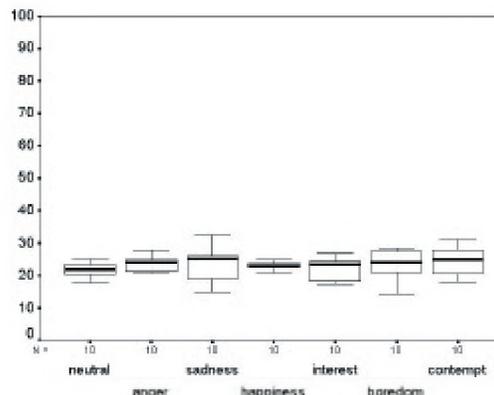


Fig. 1. Boxplots of the nasalance scores for the sentence 'He had two rock lizards', spoken with different portrayed emotions.

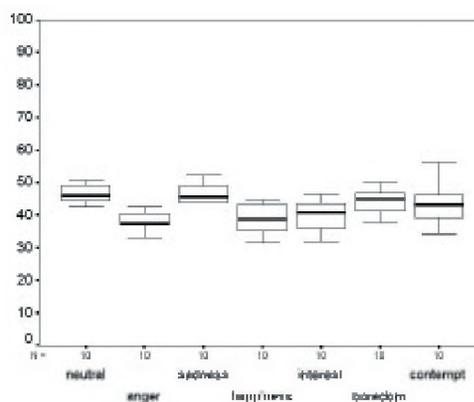


Fig. 2. Boxplots of the nasalance scores for the sentence 'He had two rock lizards', spoken with different portrayed emotions.

#### REFERENCES

- Banse, R., Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*; 70: 614-636.
- Bressmann, T. (2005). Comparison of nasalance scores obtained with the Nasometer, the NasalView, and the OroNasal System. *Cleft Palate Craniofacial Journal*; 42: 423-433
- Dalston, R.M., Seaver, E.J. (1992). Relative values of various standardized passages in the nasometric assessment of patients with velopharyngeal impairment. *Cleft Palate Craniofacial Journal*; 29: 17-21.
- Fletcher, S. (1978). *Diagnosing speech disorders from cleft palate*. New York, Grune & Stratton.
- Lewis, K.E., Watterson, T., Quint, T. (2000). The effect of vowels on nasalance scores. *Cleft Palate Craniofacial Journal*; 37: 584-589.
- Watterson, T., York, S.L., McFarlane, S.C. (1994). Effects of vocal loudness on nasalance measures. *Journal of Communication Disorders*; 27: 257-262.

# LINGUAL PROTRUSION AND ELEVATION IN LINGUAL DYSTONIA: A HYPOTHESIS

Allyson D. Dykstra, M.Sc.<sup>1</sup>, Scott G. Adams, Ph.D.<sup>2</sup>, and Mandar Jog, M.D.<sup>3</sup>

<sup>1</sup>Doctoral Program in Rehabilitation Sciences, <sup>2</sup>School of Communication Sciences and Disorders,

<sup>3</sup>Department of Clinical Neurosciences

University of Western Ontario, London, Ontario Canada

## INTRODUCTION

Dystonia is a neurological movement disorder characterized by sustained or tonic muscle contractions that can result in abnormal posturing and positioning of specific body regions. These tonic muscle contractions frequently are associated with abnormal and sometimes painful posturing and positioning, twisting and repetitive movements (Duffy, 1995; Fahn, Marsden & Calne, 1987). When dystonia affects the tongue it is known as lingual dystonia. Lingual dystonia, a hyperkinetic dysarthria, is characterized by abnormal contractions of the intrinsic or extrinsic muscles of the tongue that result in sustained tongue postures. In severe cases, lingual dystonia can have devastating effects on speech production and intelligibility.

There is no cure for lingual dystonia or any dystonia. The principle goals of therapy focus on reducing abnormal postures of the tongue, improving orofacial aesthetics, and ultimately restoring functional speech, masticatory and swallowing capabilities. One contemporary management technique that has produced generally favourable results is Botulinum toxin type A (BtA) injections to the intrinsic and/or extrinsic muscles of the tongue (Blitzer, Brin & Fahn, 1991; Charles, Davis, Shannon, Hook & Warner, 1997; Dykstra, Adams & Jog, 2004). Side effects of BtA are usually well tolerated but can include mild dysarthria, difficulty chewing, and mild dysphagia (Goldman & Comella, 2003; Munchau & Bhatia, 2000).

In 1969, Darley, Aronson and Brown studied 30 patients with hyperkinetic dysarthria associated with dystonia. These researchers established the most deviant speech dimensions of dystonia from most to least severe to be imprecise consonant articulation, vowel distortion, harsh voice, irregular articulatory breakdown, strained-strangled voice quality, monopitch and monoloudness. Speech rate in dystonia also was found to be generally slow, with abnormal direction and rhythm of movement. It should be noted, however, that Darley, Aronson and Brown's 30 subjects with dystonia included individuals with oromandibular dystonia (OMD) and individuals with spasmodic dysphonia.

In a study examining the speech characteristics of patients with Meige's syndrome (OMD and blepharospasm), Golper, Nutt, Rau and Coleman (1983) described one patient with severe lingual dystonia and a severe speech intelligibility deficit as having difficulty with forward tongue postures. As an example, the vowel /a/ was fronted, raised and diphthongized to become /aI/. A similar

fronting of lingual fricatives was noted in Dykstra, Adams and Jog's (2004) acoustic and perceptual study of lingual dystonia.

Intelligibility deficits were associated with palatal fricatives being perceived as more fronted alveolar fricatives (i.e. /ʃ →s/. These phonetic errors were clearly associated with a corresponding abnormal increase in the frequency of the fricative spectra. An informal review of published audio and videotaped samples of OMD (Darley, Aronson & Brown, 1975; Klawans, Goetz & Tanner, 1988) suggests that abnormal lingual protrusion or fronting and lingual elevation or raising are fairly common sources of imprecise consonants, vowel distortions and intelligibility deficits in subjects with OMD. For example, two OMD patients in the Darley, Aronson and Brown audiotapes (#10 and #11) show very clear fronting and raising during connected speech and also during prolonged vowels (i.e. shift from /a → i/ or /a → æ/) and syllable repetitions (i.e. tΛ tΛ → tI tI).

The nature of dystonic lingual contractions is poorly understood and described in the literature. The purpose of this paper is to examine and describe the speech deficits associated with lingual dystonia using acoustic analyses. It will be proposed that the most common features of lingual dystonic contractions involve lingual protrusion and lingual elevation during speech production. Evidence from a case study of lingual dystonia is presented to support this hypothesis.

## 2. METHOD

Retrospective acoustic analyses of a previously published case study of lingual dystonia (see Dykstra et al., 2004) was undertaken to determine and describe the nature of lingual dystonic contractions. This case-study involved an examination of the effect of BtA injections to the intrinsic muscles of the tongue on speech intelligibility. In the present investigation, acoustic analyses focusing on lingual-fricative spectra was conducted using spectrographic displays produced by the Time Frequency Response (TFR) software program (Avaaz, 1999) to investigate the hypothesis of abnormal lingual fronting and elevation.

## 3. RESULTS

Acoustic and phonetic analyses of this case suggest that lingual fronting and elevation are predominate features during lingual dystonic contractions. A prominent phonetic

error was the misperception of the palatal fricative /ʃ/ for the alveolar fricative /s/. Upon examination of these fricatives acoustically, the frequency of /ʃ/ was elevated (6639 Hertz (Hz)) relative to a normal range of above 3000 Hz (Kent, 1997). The increased frequency of /ʃ/ likely caused an incorrect perception of /ʃ/ as /s/ in the pre-treatment condition. Following treatment, the frequency of /ʃ/ approached a normal range and decreased to approximately 4300 Hz. The decrease in the frequency of /ʃ/ was reflected in the correct perception of the /s—ʃ/ contrast post-BtA treatment (Figure 1).

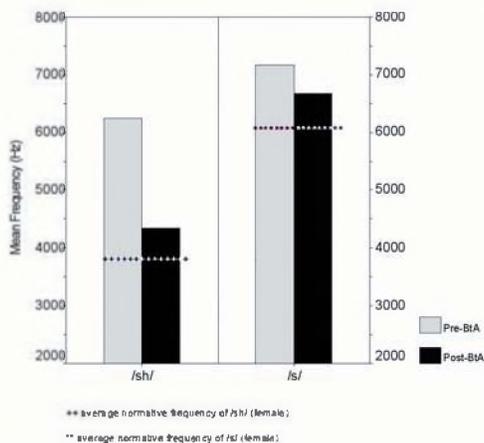


Fig. 1. Mean Frequency of fricatives /s/ and /ʃ/ - Pre and Post BtA Treatment.

Another prominent phonetic error was the misperception of /θ/ for /ʃ/ prior to treatment with BtA. Acoustically, the frequency of /ʃ/ was in the 7000-9000 Hz range pre-BtA which is much higher than the normal range of above 3000 Hz. This elevated frequency likely contributed to the misperception of /θ/ for /ʃ/. After treatment with BtA, the frequency of /ʃ/ approached a normal range of 4300 Hz, reflecting the correct perception of the /θ—ʃ/ contrast.

#### 4. DISCUSSION

Results of the above acoustic analysis, in addition to an informal perceptual evaluation of published audio and videotapes (Darley et al., 1975; Golper et al., 1983; Klawans et al., 1988) suggest that a common feature of lingual dystonic contractions involve lingual fronting and lingual elevation during speech production. A possible explanation for this phenomenon may be biomechanical in nature. The tongue is a muscular hydrostat in that it lacks a rigid skeletal framework (Kent, 1997). Muscular hydrostats are unique because any decrease in one dimension results in a corresponding increase or change in another dimension. Since the tongue is fixed posteriorly, there may be less resistance to movement in more anterior dimensions, than in posterior dimensions, perhaps accounting for the observed tendency for lingual protrusion. Similarly, the tongue is attached to the floor of the mouth, and lingual elevation may

predominate because of less resistance to upward (superior) versus downward (inferior) movement.

This examination lends preliminary support to the hypothesis of lingual fronting and elevation in lingual dystonia. Future directions include larger scale studies that systematically evaluate the exact mechanisms contributing to the speech deficit associated with lingual dystonia. Additional scientific investigation is merited to help evaluate this hypothesis.

#### REFERENCES

- Blitzer, A., Brin, M.F. & Fahn, S. (1991). Botulinum Toxin Injections for Lingual Dystonia. *Laryngoscope*, 101, 799.
- Charles, P., Davis, T., Shannon, K., Hook, M. & Warner, J. (1997). Tongue Protrusion Dystonia: Treatment with Botulinum Toxin. *Southern Medical Journal*, 90, 522-525.
- Darley, F.L., Aronson, A.E., & Brown, J.R. (1969b). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech, Hearing and Research*, 12, 462-496.
- Darley, F.L., Aronson, A.E., & Brown, J.R. (1975). *Audio Seminar in Motor Speech Disorders*. Philadelphia: Saunders
- Duffy, J. (1995). *Motor speech disorders*. Toronto: Mosby.
- Dykstra, A.D., Adams, S., & Jog, M. (2004, November). *The Effect of Botulinum toxin type A on Speech Intelligibility in Lingual Dystonia*. Poster session presented at the ASHA Annual Convention, Philadelphia, PA.
- Fahn, S., Marsden, C., & Calne, D. (1987). Classification and investigation of dystonia. In C. Marsden & S. Fahn (Eds.) *Movement Disorders 2*. London: Butterworths.
- Goldman, J.G. & Comella, C.L. (2003). Treatment of Dystonia. *Clinical Neuropharmacology*, 26(2), 102-108.
- Golper, L., Nutt, J., Rau, M., & Coleman, R. (1983). Focal Cranial Dystonia. *Journal of Speech and Hearing Disorders*, 48, 128-134.
- Kent, R. (1997). *The Speech Sciences* (pp. 350-360). San Diego, CA: Singular.
- Kent, R., Weismer, G., Kent, J., & Rosenbek, J. (1989). Toward explanatory intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, 54, 482-489.
- Klawans, H.L., Goetz, C.G. & Tanner, C.M. (1988). *Common Movement Disorders - A video presentation* [Videotape]. Lippincott Williams & Wilkins.
- Munchau, A. & Bhatia, K.P. (2000). Uses of botulinum toxin injection in medicine today. *British Medical Journal*, 320(7228), 161-165.
- TFR: Version 2.1, (1999) [Computer software]. London, ON: Avaaz Innovations.

# THE EFFECT OF PACKING ORDER ON ULTRASOUND BACKSCATTER FROM CELLS AT DIFFERENT VOLUME FRACTIONS

Ralph E. Baddour<sup>1</sup> and Michael C. Kolios<sup>2</sup>

<sup>1</sup>Dept. of Medical Biophysics, University of Toronto, 610 University Ave., Toronto, ON, Canada, M5G2M9; E-mail: rbaddour@uhnres.utoronto.ca <sup>2</sup>Dept. of Physics, Ryerson University, 350 Victoria St., Toronto, ON, Canada, M5B2K3

## 1. INTRODUCTION

Recently, new high frequency ultrasound devices have emerged with better system signal-to-noise ratio characteristics, which make it possible to measure ultrasound scattering in tissues so that small variations in scatterer volume fraction, which can result from changes in tissue microstructure due to cancer therapies, may have a significant impact on the backscattered signal. Previous studies have shown that the backscattered ultrasound power from suspensions of scatterers is related to the fraction of the total volume occupied by the scatterers. Although this volume fraction effect has been well studied for certain biological scatterers, in particular red blood cells (Shung *et al.*, 1984; Mo *et al.*, 1994), the sensitivity of these responses to packing order, assuming no aggregation, have not been investigated. In this study, a three-dimensional computer model was used to study the effect of position randomness on the ultrasound backscatter from cell suspensions at several volume fractions.

## 2. COMPUTER MODEL

Although producing cell suspensions *in vitro* with any desired volume fraction is a feasible undertaking, it is impossible to create suspensions in which cells have a precise degree of freedom of position. Therefore, a computer model was developed to study the effect of these two factors on ultrasound backscatter. To calculate the backscatter frequency response from a simulated suspension of cells, the total backscattered pulse must be computed. This is calculated by summing the individual backscattered pressure pulses generated from each cell (all cells are assumed to be identical), accordingly time-delayed by depth (frequency-dependent attenuation in the medium is assumed to be negligible) and weighted by a transducer aperture

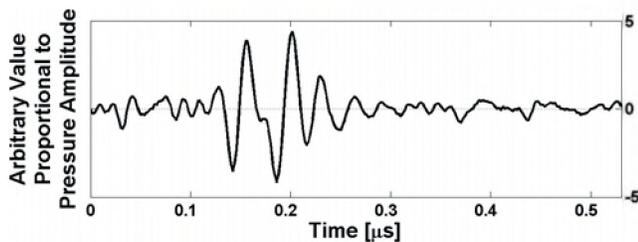


Fig. 1. Backscattered pulse from a single AML cell interrogated with a broadband 20 MHz transducer (Baddour *et al.*, 2005).

function. As there is currently no satisfactory scattering model for single, nucleated cells, a representative backscattered pulse measured from a single acute myeloid leukemia (AML) cell (13.4  $\mu\text{m}$  mean diameter) with a broadband 20 MHz focussed transducer (Baddour *et al.*, 2005) was used as the elemental scattered pulse (see Fig. 1).

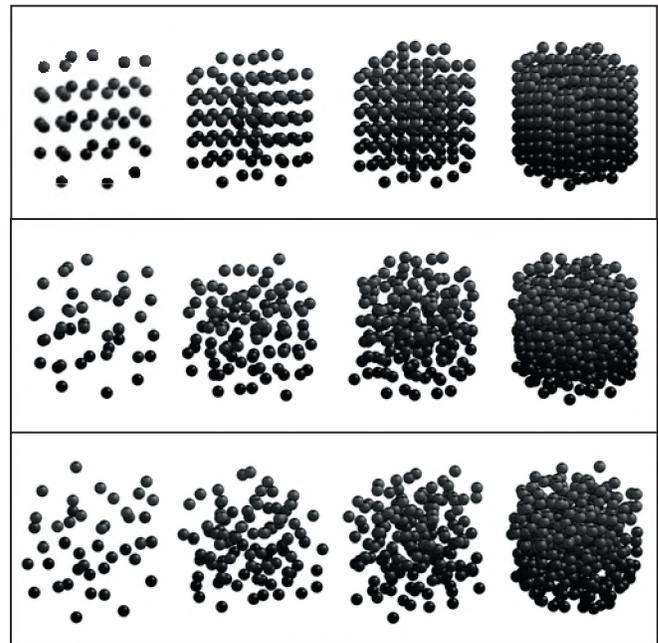


Fig. 2. Orthographic projections of realisations of some simulated cell suspensions, varying the volume fraction from left to right (0.02, 0.06, 0.10, 0.30) and varying the value of *cellLocRange* from top to bottom (0.00, 0.25, 0.50). Note that these are simply meant to be illustrative; the pack dimensions do not reflect the cylindrical region employed to obtain the simulation results.

The spatial arrangement of cells in the simulated suspensions can be varied, for example to simulate perfect, crystal-like sets or more random arrangements (some examples are presented in Fig. 2). The model extends a previously proposed two-dimensional model (Baddour *et al.*, 2002) using the parameter *cellLocRange* to control the randomness of the cell positions in the suspension. Here, *cellLocRange* is defined as a fraction of the inter-cell spacing, for a given volume fraction, assuming an ordered lattice packing. When *cellLocRange* is zero, the cells are packed in a perfectly staggered lattice. As *cellLocRange*

increases, cells are allowed to uniformly randomly take positions farther away (up to *cellLocRange* times the lattice-based inter-cell spacing) from their assigned lattice position.

### 3. RESULTS AND DISCUSSION

Cell suspensions were simulated for several values of volume fraction (up to 0.74, the limit for face-centered cubic packing of spheres) and *cellLocRange*. In each case, the total backscattered pulse returned by the model was used to calculate the integrated backscatter (IB) (Raju and Srinivasan, 2001). The IB was calculated for the range of 10 - 30 MHz, corresponding to the 6 dB bandwidth of the transducer employed to acquire the backscattered pulse from the single AML cell. A reflected pulse, using the same acquisition setup as the AML cell measurement (Baddour *et al.*, 2005), from a flat, polished SiO<sub>2</sub> crystal, placed perpendicular to the beam at the transducer's focus, was used as the reference in the IB calculation.

To enhance the confidence of the results, the IB was calculated for 10 independent suspension realisations at each parameter condition. The mean IBs are presented in Fig. 3. Although the mean IB was also calculated for the *cellLocRange* = 0 case, it is not presented in the figure as it fluctuates greatly with volume fraction, with no apparent trend. This behaviour is due to the perfect lattice nature of these packs; as the volume fraction is varied, the uniform spacing of the cells will inevitably correspond to an ideally constructive or destructive interference condition for the pulse, shown in Fig. 1, used as the elemental scattered pulse from individual cells in the model.

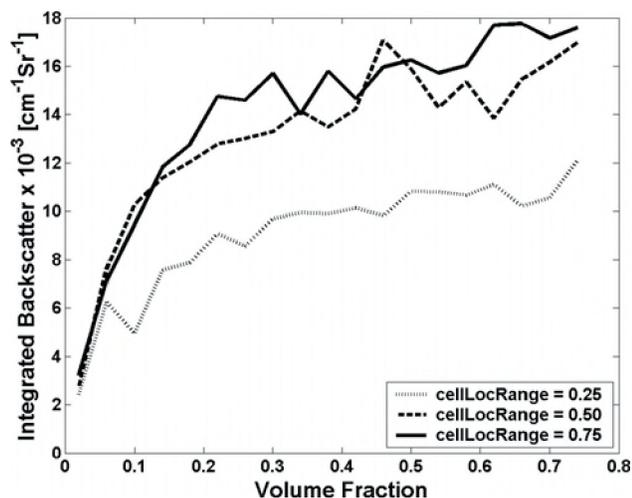


Fig. 3. The IB from simulated suspensions of cells with different degrees of position randomness (lower *cellLocRange* corresponds to more ordered, less random packing) for a wide range of volume fractions. Each IB curve represents the mean of 10 suspension realisations.

As expected, and previously documented (Hunt *et al.*, 1995), increasing the randomness of the scattering source positions increases the backscatter. In this model, because

the randomisation parameter *cellLocRange* is relative to the volume fraction, one might expect the gap in mean IB between the *cellLocRange* = 0.25 and *cellLocRange* = 0.50 conditions to narrow for higher volume fractions, as the effective freedom of cell positions is reduced. However, this trend does not appear to exist. Nevertheless, for very low volume fractions, below 0.06, it is interesting to note that the degree of packing disorder does not appear to have a significant impact on backscatter.

The relation between IB and the fraction of the total observation volume occupied by cells, although appearing to plateau for high volume fractions, is generally positive. This is not what was seen with red blood cells (Shung *et al.*, 1984; Mo *et al.*, 1994), where both the measured and theoretical backscatter coefficients decreased for volume fractions (hematocrit) above 0.2. The authors of these studies concluded that this effect was due to increasing correlation between cell locations as the volume fraction increases; with more ordered packing, destructive interference between individual scattered pulses becomes more likely. Although the model presented here does not prevent cell overlap, which becomes possible with higher *cellLocRange* values, it is not likely the only reason that a decrease in backscatter is not observed. The red blood cell models (e.g. hard sphere, continuum) assume a simple scattering function, whereas it is clear from the single cell pulse in Fig. 1 that the scattering from a nucleated cell is not trivial, with possibly some resonant component.

It is anticipated that this model will become a useful tool when attempting to interpret experimentally measured backscatter from suspensions of different cell types.

### REFERENCES

- Baddour, R. E., Sherar, M. D., Czarnota, G. J., Hunt, J. W., Taggart, L., Giles, A., Farnoud, N. R. and Kolios, M. C. (2002). "High frequency ultrasound imaging of changes in cell structure including apoptosis," Proc. IEEE Ultrason. Symp. 1598-1603.
- Baddour, R. E., Sherar, M. D., Hunt, J. W., Czarnota, G. J. and Kolios, M. C. (2005). "High-frequency ultrasound scattering from microspheres and single cells," J. Acoust. Soc. Am. **117**, 934-943.
- Hunt, J. W., Worthington, A. E. and Kerr, A. T. (1995). "The subtleties of ultrasound images of an ensemble of cells - simulation from regular and more random distributions of scatterers," Ultrasound Med. Biol. **21**, 329-341.
- Mo, L. Y. L., Kuo, I. Y., Shung, K. K., Ceresne, L. and Cobbold, R. S. C. (1994). "Ultrasonic scattering from blood with hematocrits up to one hundred percent," IEEE Trans. Biomed. Eng. **41**, 91-95.
- Raju, B. I. and Srinivasan, M. A. (2001). "High-frequency ultrasonic attenuation and backscatter coefficients of in vivo normal human dermis and subcutaneous fat," Ultrasound Med. Biol. **27**, 1543-1556.
- Shung, K. K., Yuan, Y. W., Fei, D. Y. and Tarbell, J. M. (1984). "Effect of flow disturbance on ultrasonic backscatter from blood," J. Acoust. Soc. Am. **75**, 1265-1272.

# THE FAST NEARFIELD METHOD APPLIED TO AXISYMMETRIC RADIATORS

James F. Kelly, Xiaozheng Zeng, and Robert J. McGough<sup>1</sup>

<sup>1</sup>Dept. of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, 48824, USA  
[kellyja8@msu.edu](mailto:kellyja8@msu.edu)

## 1. INTRODUCTION

The pressure fields radiated by baffled circular pistons have been studied extensively by the acoustics community [1]. Since many realistic transducers have a velocity profile that varies in the radial direction, the field generated by axisymmetric radiators is an important application [2,3]. In an earlier work, the authors derived a fast nearfield method (FNM) for pistons that have a spatially uniform velocity [4]; and in this paper, the FNM is extended to radiators with a radially varying normal velocity. This fast nearfield method is implemented and compared with the standard Rayleigh integral approach.

## 2. THEORY

Consider a circular piston of radius  $a$  lying in the  $x$ - $y$  plane surround by an infinite, rigid baffle. Assume a lossless, homogeneous medium with density  $\rho$  and sound speed  $c$ .

The single-frequency pressure  $P_a(r, z; \omega)$  generated by this piston with a spatially uniform velocity of unity is given by the fast nearfield method (FNM)

$$P_a(r, z; \omega) = \frac{\rho c a}{\pi} \int_0^\pi \frac{r \cos \psi - a}{r^2 + a^2 - 2ar \cos \psi} \times \left( e^{-jk\sqrt{r^2 + a^2 + z^2 - 2ar \cos \psi}} - e^{-jkz} \right) d\psi \quad (1)$$

where  $r$  and  $z$  are the radial and axial observation coordinates, respectively. Eq. (1) is derived and analyzed in terms of error and speed in [4]. To extend Eq. (1) to the more general apodized piston, define an apodization function  $q(s)$  the gives the normal velocity of the piston as function of radius  $s$ . As noted by [5], the pressure field generated by this apodized piston is synthesized by decomposing the circle of radius  $a$  into evenly spaced concentric annuli. Fig 1 illustrates this concept. By allowing the number of annuli to approach infinity, the following expression is derived:

$$P(r, z; \omega) = \int_0^a \frac{\partial P_s(r, z; \omega)}{\partial s} q(s) ds \quad (2)$$

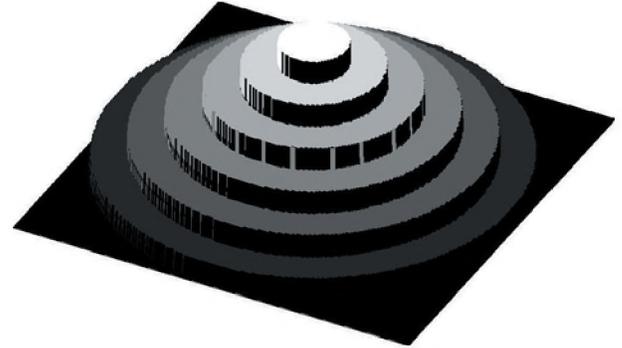


Fig. 1: Schematic illustrating how an apodized piston is decomposed into contributions from concentric annuli.

By assuming that the normal velocity vanishes on the boundary of the piston ( $q(a) = 0$ ), Eq. (2) is integrated by parts, yielding

$$P(r, z; \omega) = - \int_0^a \frac{\partial q}{\partial s} P_s(r, z; \omega) ds \quad (3)$$

Substituting Eq. (1) into Eq. (3) yields the final result

$$P_a(r, z; \omega) = - \frac{\rho c}{\pi} \int_0^a q'(s) s \int_0^\pi \frac{r \cos \psi - a}{r^2 + a^2 - 2ar \cos \psi} \times \left( e^{-jk\sqrt{r^2 + a^2 + z^2 - 2ar \cos \psi}} - e^{-jkz} \right) d\psi ds \quad (4)$$

## 3. METHODS

Eq. (4) is valid for any radial apodization. One choice for the aperture function is

$$q(s) = 1 - (s/a)^n \quad (5)$$

where  $n$  is a fixed integer. A rigid piston corresponds to  $n = 0$ , whereas a simply supported piston is modeled by  $n = 2$  [3]. Although Eq. (4) is difficult to evaluate analytically, the double integral can be evaluated numerically via Gauss quadrature or other standard quadrature rules.

## 4. RESULTS

A reference pressure field generated by a circular radiator of radius 1.5 mm operating at a frequency of 2.5 MHz is evaluated for a parabolic apodization ( $n=2$ ). In order to compare Eq. (4) with the Rayleigh integral, a reference field is evaluated with 160,000 point Gauss quadrature rule on an axially offset 21 by 75 point grid with quarter-wavelength spatial sampling. Figure 2 shows the normalized pressure amplitude. The apodized beam pattern is more spatially bandlimited than a rigid piston beam pattern; in addition, the beam pattern in Fig. 2 lacks the on-axis nulls that characterize the unapodized piston's pressure field.

The Rayleigh-Sommerfeld integral representing the apodized pressure field was also evaluated using the point source method [6]. Both the FNM and point source approaches were implemented in the C programming language and executed on a 3.0 GHz Pentium IV processor running Red Hat Linux. The FNM and point source method are evaluated with varying numbers of quadrature points. The resulting peak normalized errors and computation times are summarized in Tables 1 and 2.

## 5. DISCUSSION

Unlike the Rayleigh-Sommerfeld integral, the FNM embodied in Eq. (4) does not experience any numerical difficulty near the piston surface. This smooth behavior leads to a more rapid convergence with respect to number of quadrature points. As evinced by Tables 1 and 2, the FNM requires 0.0064 seconds to achieve 10% peak error, whereas the point source approach requires 0.0258 seconds; thus the FNM achieves a speedup by a factor of 4 at 10 % error level. At 1 % error, FNM achieves a speedup factor is about 4.9 relative to the point source approach. Reduced computation times may be possible by implementing the grid-sectoring method described in [4].

A time-domain analog of Eq. (4) can also be obtained for transient excitations. This time-domain expression may prove useful in evaluating scattered and pulse-echo fields generated by imaging transducers.

Table 1. Quadrature points need for specified peak error.

	10 % Error	1% Error
FNM	40	105
Point Source	140	442

Table 2. Computation times need for specified peak error.

	10 % Error	1% Error
FNM	0.0064 s	0.0165 s
Point Source	0.0258 s	0.0809 s

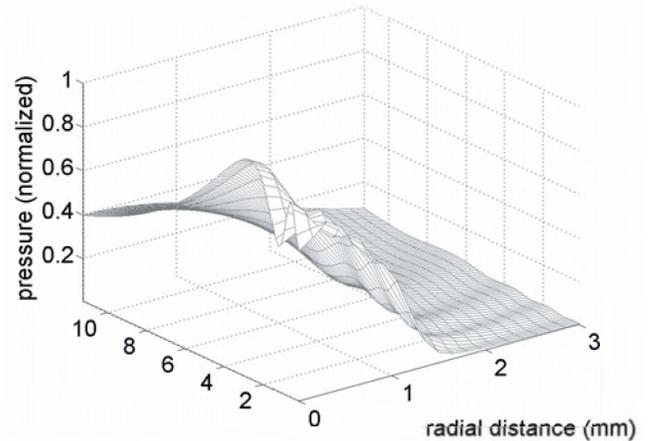


Fig. 2: Normalized pressure amplitude for a parabolic radiator ( $n=2$ ) with radius  $a=2.5$  mm, or 2.5 acoustic wavelengths. Eq. (3) was evaluated with a 1000 by 1000 Gauss quadrature on an axially offset 21 by 75 point grid with quarter-wavelength spatial sampling.

## 6. CONCLUSION

A fast nearfield method has been derived for radially apodized circular pistons. Since this method is numerically well-behaved at all observation points, convergence is accelerated as compared to the point-source approach. A speedup on the order of 4 is achieved at 10 and 1 % peak errors.

## REFERENCES

- Harris, G. R. (1981). Review of transient field theory for a baffled planar piston. *J. Acoust. Soc. Am.*, 70(1), 10-20.
- Greenspan, M. (1979) Piston radiator: Some extensions of the theory. *J. Acoust. Soc. Am.* (65), 608-62.
- Christopher, P. T. and K. J. Parker (1991). New approaches to the linear propagation of acoustic fields. *J. Acoust. Soc. Am.*, 90(1), 507-521.
- McGough, R. J., T. V. Samulski, and J. F. Kelly. An efficient grid sectoring method for calculations of the near-field pressure generated by a circular piston. *J. Acoust. Soc. Am.*, 115(5), 1942-1954.
- Hutchins, D.A., H. D. Mair, P. A. Puhach, and A. J. Osei (1986). Continuous -wave pressure fields of ultrasonic transducers. *J. Acoust. Soc. Am.* 80(1), 1-12.
- J. Zemanek (1971). Beam behavior within the nearfield of a vibrating piston. *J. Acoust. Soc. Am.*, 49(1), 181-191.

## ACKNOWLEDGEMENTS

The authors would like to thank Donald Chorman for helpful discussions. This research was funded in part by NIH grant 5R01CA093669.

# MULTI-PLANAR ANGULAR SPECTRUM APPROACH APPLIED TO PRESSURE FIELD CALCULATIONS OF SPHERICALLY FOCUSED PISTONS

Xiaozheng Zeng, James F. Kelly, and Robert J. McGough

Dept. of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, zengxiao@msu.edu

## 1. INTRODUCTION

The angular spectrum approach (ASA) is essentially a Green's function method that computes the field by multiplying the source and the Green's function, or propagation operator, in the spectral domain. Some ASA applications use the analytical Fourier transform of Green's function as the propagator [1]-[3]; however, this method suffers from wrap-around error and aliasing. In the far field, large errors are produced because of the truncation of the spatial Green's function. This paper presents a multi-planar angular spectrum approach based on the analysis of the errors. The additional source planes at the path of propagation reformulate the spatial propagator and compensate for the truncation errors.

Most investigations apply the ASA to planar sources, including single transducers and arrays. The ASA for cylindrically curved radiators has also been studied by Wu [4]. This paper will focus on the application of ASA on spherically focused circular pistons.

## 2. THEORY

### 1.1 Angular spectrum approach

The convolution relationship for the waves diffracted from finite apertures is

$$p(x, y, z) = p_0(x, y, z_0) \otimes h(x, y, z), \quad 1$$

where  $p_0(x, y, z_0)$  is the source pressure in plane  $z_0$  and  $p(x, y, z)$  is the pressure in other planes. For time harmonic wave excitation, the spatial propagator  $h(x, y, z)$  can be represented by [5]

$$h(x, y, z) = \frac{\Delta z}{2\pi d^3} (1 + jkd) e^{-jkd}, \quad 2$$

where  $k=2\pi/\lambda$  is the spatial wavenumber,  $\Delta z = z - z_0$  and

$d = \sqrt{x^2 + y^2 + \Delta z^2}$ . If  $p(x, y, z)$ ,  $p_0(x, y, z)$ , and  $h(x, y, z)$  are transformed into spectral domain by 2D FFT, and Eq. 1 becomes

$$P(k_x, k_y, z) = P_0(k_x, k_y, z) H(k_x, k_y, z), \quad 3$$

where  $k_x$ ,  $k_y$ , and  $k_z$  are angular frequencies of the decomposed plane waves. The wave number

$k = \sqrt{k_x^2 + k_y^2 + k_z^2}$  is constructed from these angular frequency values.

Numerical implementations of these expressions sample the spatial fields at an interval of  $\delta$  in both  $x$  and  $y$  dimensions. Each pressure field input is confined to an  $L \times L$  square area and sampled by an  $M \times M$  grid.  $M$  is preferably an odd integer so that the source is symmetric with respect to the origin. To avoid problems with circular convolution, the source plane is zero padded and enlarged to an  $N \times N$  matrix. In general, the accuracy of angular spectrum method is influenced by the spatial sampling rate, the size of the computational grids, and the angular resolution of spectra.

### 1.2. Spatial propagator and propagation distance

For a uniform sampling rate and field size, the errors produced by the ASA vary as a function of depth. The errors generated by the spatial propagator are encountered in two distinct regions. First, as  $\Delta z \rightarrow 0$  a singularity appears at the origin of  $h(x, y, z)$ , the steep slope of the signal makes adequate sampling impossible. Thus, the computation of the ASA in the region  $\Delta z < 1\lambda - 2\lambda$  consistently produces a relatively large error. Second, as  $\Delta z$  increases, the wavefront spreads over larger region. A plane that captures most of the energy in the near field often includes much less energy in the far field. If a significant portion of the power is lost due to truncation, the resulting spectrum is distorted.

## 3. METHODS

### 2.1 Spherically focused piston

Accurate fields of spherically focused pistons can be computed as reference by impulse response algorithm [6] with a high sampling rate. The excitation frequency is 1MHz and acoustic velocity is 1500m/s for the simulation. Lossless media is assumed in these simulations.

### 2.2 Error evaluation

The normalized peak error is defined as the amplitude of the maximum error normalized by the peak pressure in one plane according to the expression

$$\eta_{\max} = \frac{\max_{x,y} |p(x, y, z) - p_{\text{ref}}(x, y, z)|}{\max_{x,y} |p_{\text{ref}}(x, y, z)|}, \quad 4$$

where  $p_{\text{ref}}(x, y, z)$  is the reference pressure, and  $p(x, y, z)$  is the pressure computed by the ASA. The

notation  $\max_{x,y}$  indicates that the error is evaluated in a transverse plane.

### 2.3 Multi-planar scheme

An adaptive computing algorithm propagates the fields along  $z$  direction and places multiple source planes according to a threshold. First, the errors are evaluated as a function of depth and the position where the error first exceeds the threshold is identified. This determines the location of the next source plane, and the fields in subsequent planes are computed from this source plane until the error threshold is again exceeded. The same process is repeated until all of the source plane locations are determined. The error within a short distance of the first source plane may exceed the threshold, but if the error drops below the threshold after one or two wavelengths, the points adjacent to the initial plane are discarded.

## 4. RESULTS

The multi-planar ASA is demonstrated for a spherically focused piston with diameter  $2a=16\text{cm}$  and radius of curvature  $R=16\text{cm}$ . The normal evaluated at the center of the spherical shell, which is coincident with the  $z$ -axis, defines the origin of the Cartesian coordinate system. The computational volume is  $16.2\text{cm} \times 16.2\text{cm} \times 32.75\text{cm}$  sampled at an interval of  $\delta=\lambda/2=0.075\text{cm}$ . Each plane is zero-padded to  $512 \times 512$  points, and then a 2D FFT is evaluated. The adaptive multi-planar algorithm allocates five source planes, and this restricts the peak error to values below 5%. Figure 1 is an illustration of the pressure field in the  $y=0$  plane. This result is obtained from source planes positioned at  $z=2.25\text{cm}$ ,  $4.5\text{cm}$ ,  $11.1\text{cm}$ ,  $21.45\text{cm}$  and  $29.4\text{cm}$ .

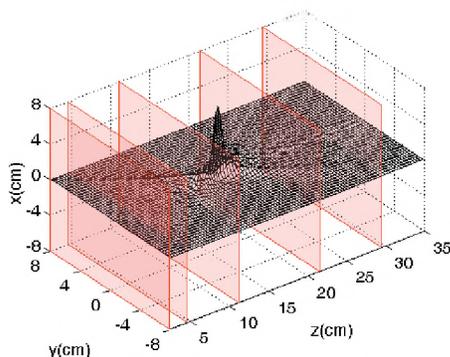


Figure 1: Illustration of the pressure generated by a spherically focused piston computed with five source planes.

Figure 2 shows the peak errors as a function of depth. The solid line is computed by the initial source plane at  $z=2.25\text{cm}$ . With a single source plane (solid line), the error grows rapidly beyond the focus. Shortly thereafter, the error reaches 100%. The dashed line is the multi-planar result computed from five source planes. This result shows that the multi-planar approach maintains an error of 5% or less throughout the computational grid.

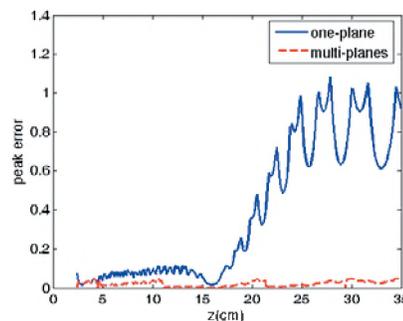


Figure 2: Peak errors as a function of depth. The solid line represents the error produced by one source plane at  $z=2.25\text{cm}$ , and the dashed line describes the error obtained with five source planes.

## 5. DISCUSSION

Since the inadequate truncation of the spatial propagator is one source of error, increasing the size of the sampled grid is expected to reduce the peak error. However, in order to include 90% of the energy in  $h(x,y,z)$  at depths of  $z=10\text{cm}$  and  $30\text{cm}$ , the extent of the sampled grid should be  $27.2\text{cm}$  and  $81.6\text{cm}$ , respectively. This calculation would require an excessive amount of computer memory. The multi-planar scheme reduces numerical errors by efficiently utilizing memory resources without adding a significant computational cost.

## 6. CONCLUSION

The calculations of pressure fields from spherically focused piston using ASA often suffer from truncation errors in the far field. The multi-planar ASA compensates for the error by adding additional source planes and achieves much greater accuracy.

## REFERENCES

1. Stepanishen, P. R. and Benjamin, K. C.(1982), Forward and backward projection of acoustic fields using FFT methods, *J. Acoust. Soc. Am.*, vol. 71, no. 4, pp. 803-812.
2. Clement, G. T. and Hynynen, K.( 2000), Field characterization of therapeutic ultrasound phased arrays through forward and backward planar projection, *J. Acoust. Soc. Am.*, vol. 108, no. 1, pp. 441-446.
3. Wu, P., Kazys, R. and Stepinski, T.(1996), Analysis of the numerically implemented angular spectrum approach based on the evaluation of two-dimensional acoustic fields I & II, *J. Acoust. Soc. Am.*, vol. 99, no. 3, pp. 1339-1359.
4. Wu, P., Kazys, R. and Stepinski, T. (1999), Extension of the angular spectrum approach to curved radiators, *J. Acoust. Soc. Am.*, vol. 105, no. 5, pp. 2618-2627.
5. Liu, D. and Wagg, R. C.( 1997), Propagation and backpropagation for ultrasonic wavefront design, *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 44, no. 1, pp. 1-13.
6. Arditi, M., Foster, F. S. and Hunt, J. W. (1981), Transient fields of concave annular arrays, *Ultrason. Imaging*, vol. 3, no. 1, pp. 37-61.

[This work was supported in part by NIH R01CA093669.]

# A TIME-SPACE DECOMPOSITION METHOD FOR FAST CALCULATIONS OF TRANSIENT PRESSURE FIELDS GENERATED BY ULTRASOUND PHASED ARRAYS

James F. Kelly and Robert J. McGough<sup>1</sup>

<sup>1</sup>Dept. of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, 48824, USA  
[kellyja8@msu.edu](mailto:kellyja8@msu.edu)

## 1. INTRODUCTION

Computing the transient pressure field generated by large phased arrays is helpful in designing ultrasound imaging systems. Methods for computing the transient near-field pressure of a planar aperture include the point-source method [1] and the spatial impulse response (SIR) method [2,3,4]. Recently, a rapid single integral approach has been developed for computing the time domain pressure generated by a baffled circular piston [5]. This solution is applied to a 129 element focused phased array and compared to a similar computation made using Field II [6].

## 2. THEORY

A time domain solution to the lossless wave equation can be derived subject to an input pulse  $v(t)$ , which models the uniform normal velocity of the piston. Consider a baffled rigid piston with radius  $a$  radiating into a homogeneous fluid with density  $\rho_0$  and sound speed  $c$ . Solving the wave equation in cylindrical coordinates  $(r,z)$  yields the single-integral solution:

$$p(r, z, t) = \frac{\rho_0 a c}{\pi} \int_0^\pi \frac{r \cos \psi - a}{r^2 + a^2 - 2ar \cos \psi} \times \left[ v \left( t - \sqrt{r^2 + z^2 + a^2 - 2ar \cos \psi} / c \right) - v(t - z/c) \right] d\psi \quad (1)$$

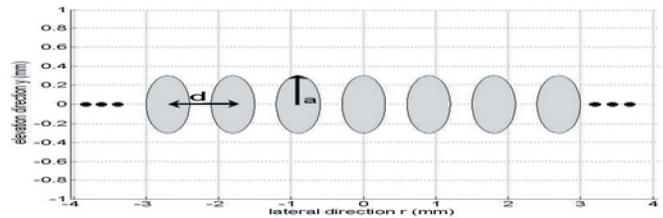
Although Eq. (1) can be directly evaluated in the time-domain using Gauss quadrature, the computational complexity is significantly reduced by decoupling the temporal and spatial dependence in the integrand of Eq. (1). Consider a Hanning-weighted pulse  $v(t)$  with duration  $W$ , which is decomposed via

$$v(t - \tau) = \text{rect} \left( \frac{t - \tau}{W} \right) \sum_{n=1}^6 f_n(\tau) g_n(t) \quad (2)$$

where the delay  $\tau$  depends on the spatial coordinates  $(r,z)$  and the variable of integration  $\psi$ . Inserting Eq. (2) into Eq. (1) allows the temporal and spatial dependence to be separated. The expansion functions in Eq. (2) are computed via trigonometric expansions. This decomposition reduces the number of integrations per observation point from the number of time samples to 6 without introducing any additional error.

## 3. METHODS

A linear array of 129 circular elements is simulated with pistons of radius  $a = 0.30$  mm (half-wavelength) and inter-element spacing  $d = 0.90$  mm. Fig. 1 shows the array geometry. A computational grid using quarter-wavelength (0.15 mm) spatial sampling in both the lateral and axial dimensions was employed. The Hanning excitation pulse has a central frequency of 2.5 MHz and duration of 1.2 microseconds. Beam steering and focusing is achieved by application of temporal delays to each element [7]. The total pressure is then synthesized via superposition. Beam steering and focusing are achieved by applying temporal delays. To focus on axis at distance  $F = 50$  mm, quadratic delays are employed.



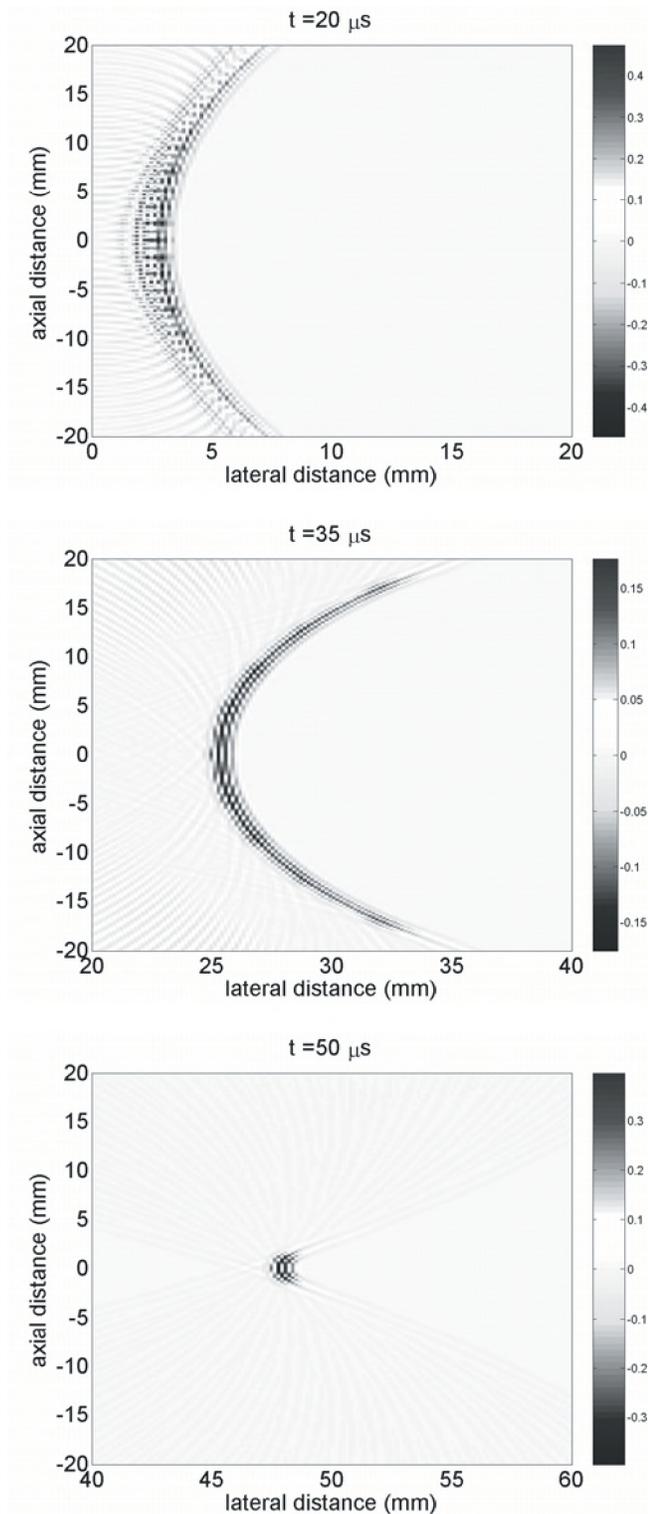
**Fig 1:** Densely sampled linear array. The array used in these computations has 129 circular pistons with radius  $a = 0.30$  mm (half-wavelength) and inter-element spacing  $d = 0.90$  mm on a computational grid extending 288 wavelengths in lateral direction by 167 wavelengths in axial direction.

## 4. RESULTS

To determine the correct number of quadrature points to apply to the decomposition technique given by Eq. (1), an error analysis is shown in Table 1. As Table 1 shows, Eq. (1) requires 4 Gauss abscissas to achieve a 1 % peak error. Field II, which subdivides the aperture into rectangular sub-elements, requires 484 sub-elements to achieve this error level.

Simulation pressure fields are shown in Fig. 2 at three successive times, where the pressure has been normalized with respect to peak pressure. The total computation time for this array system was 11 minutes. In comparison, similar computations using Field II software<sup>a</sup> [6] took approximately 8 hours to achieve commensurate accuracy. The peak field error is computed relative to a 1000 point

<sup>a</sup> Field II version 2.86 for MATLAB, [www.es.oersted.dtu.dk/staff/jaj/field/](http://www.es.oersted.dtu.dk/staff/jaj/field/)



**Fig 6:** Normalized pressure field at three successive times for the phased array defined in Fig. 1. Pressure is normalized with respect to the peak value. On-axis focusing at 50 mm is employed via quadratic time delays.

Table 1. Single-element error analysis.

	10 % Error	1 % Error
Decomposition	3 abscissas	4 abscissas
Field II	12 sub-elements	484 sub-elements

sampling frequency of 32 Hz (compared to Field's 100 reference field; a 4-point quadrature yields a peak error below 1% at all points in the computational grid. Since the present decomposition technique utilizes a temporal MHz sampling), less memory is used.

## 5. DISCUSSION

Fast and accurate incident pressure field computations are necessary in several applications. Of particular importance is the iterative design of imaging arrays (both 1 and 2D). Array geometry and parameters can be optimized by computing transmitted and pulse-echo pressure fields. Large-scale modeling of wave propagation and scattering can also benefit from the fast method presented. Time-domain scattering methods, such as generalizations of the fast multipole method (FMM) [8] require incident field data on large, unstructured grids as an input.

## 6. CONCLUSION

A simulation scheme for pulsed computations with linear phased arrays has been proposed. Unlike previous methods [6], far field and aperture approximations are not used; instead, an exact time-domain solution forms the basis for array simulations, which is accelerated by decomposing the spatial and temporal dependence of the integrand.

## REFERENCES

1. Zemanek, J. Beam behavior within nearfield of a vibrating piston. *J. Acoust. Soc. Am.* 49(1), 181-191.
2. Lockwood, J. C. and J. G. Willette (1973). High-speed method for computing exact solution for pressure variations in nearfield of a baffled piston. *J. Acoust. Soc. Am.* 53 (3), 735-741.
3. Oberhettinger, F. (1961). On transient solutions of the 'baffled piston' problem. *Journal of Research of the National Bureau of Standards, Section B*, 65B, 1-6.
4. Stepanishen, P. R. (1970). Transient radiation from pistons in an infinite planar baffle. *J. Acoust. Soc. Am.* 49 (5), 1629-1638.
5. Kelly, J. F. and R. J. McGough (2005, submitted). A time-space decomposition method for calculating the near field pressure generated by a pulsed circular piston. *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*
6. Jensen, J. A. and N. B. Svendsen (1992). Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, 39 (2), 262-267.
7. Von Ramm, O. T. and S. W. Smith (1983). Beam steering with linear arrays. *IEEE Trans. Biomed. Eng.* 30 (8), 438-52.
8. Ergin, A. A., B. Shanker, and E. Michielssen (1999). The plane-wave time-domain algorithm for the fast analysis of transient wave phenomena. *IEEE Antennas and Propagation Magazine*, 41(4), 39-51.

## ACKNOWLEDGEMENTS

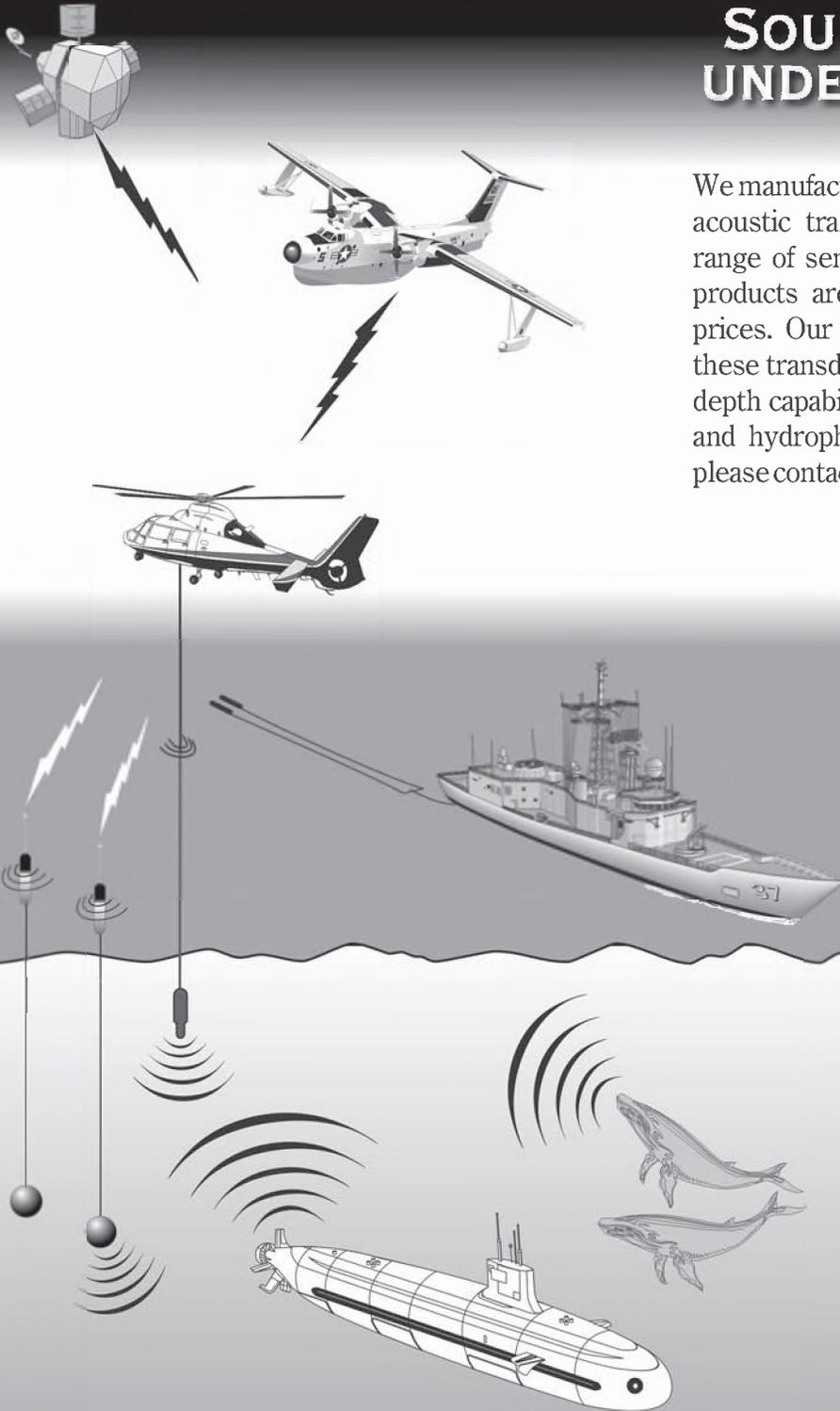
This research was funded in part by NIH grant 5R01CA093669. The authors thank Xiaozheng Zeng, Shanker Balasubramaniam, and Liyong Wu for helpful discussions and comments.



# ACOUSTIC TRANSDUCERS HYDROPHONES

## SOUND SOLUTIONS IN UNDERWATER SENSORS

We manufacture commercial and military underwater acoustic transducers and hydrophones in a broad range of sensitivities and sizes. These high-quality products are delivered in volumes at competitive prices. Our engineers can also customize many of these transducers to meet specific requirements and depth capability. To discuss your acoustic transducer and hydrophone needs, and prices and availability please contact us.



**SENSOR TECHNOLOGY LIMITED**

PO Box 97, Collingwood, Ontario, Canada L9Y 3Z4  
www.sensortech.ca email: techsupport@sensortech.ca

*PLACE*

*ECKEL*

*AD HERE*

## ABSTRACTS FOR PRESENTATIONS WITHOUT SUMMARY PAPERS

### UNDERWATER ACOUSTICS

#### **A Smart Hydrophone For Underwater Acoustics**

**N. Somayajula, B. Yan, S.E. Prasad, and R. Blacow**

Sensor Technology Limited, Collingwood, Ontario, Canada

The paper will describe a smart, networkable hydrophone for underwater acoustics. The sensor module also has other integrated sensors that provide information on operational parameters such as temperature and pressure. The hydrophone can use the additional sensory information to self-compensate for variations in operating conditions, thereby providing more stable operational data. Such sensors will be able to identify themselves on a network and will include multi-channel sensing capabilities, a built-in amplifier and possible data multiplexing. The hydrophone will be able to store critical information such as sensitivity, capacitance etc. The additional sensor information can be multiplexed into a single output data stream. This paper describes the construction, architecture and applications for smart hydrophones in underwater acoustics.

### PHYSICAL ACOUSTICS/ULTRASOUND

#### **Mri Measurements Of Acoustically Cavitated Fluid In A Standing Wave**

**Igor Mastikhin\*, Benedict Newling**

MRI Centre, Department of Physics, P.O. Box 4400, University of New Brunswick, Fredericton, New Brunswick, Canada E3B 5A3

The prevalent methods used in studies of cavitation are optical and acoustical. Given their nature, the techniques are sensitive to changes in optical or acoustical transparency respectively. The bubbles in a cavitating fluid reflect light and absorb and re-radiate acoustic signals. The shielding effect of surrounding bubbles makes it difficult to obtain information about the central zones of a cavitation cloud. In this work, Magnetic Resonance Imaging was applied, for the first time, to studies of cavitating fluid in a standing acoustic wave at 31kHz sound frequency. A spin echo Pulsed Field Gradient sequence was employed to sensitize the measurement to motion. Velocity spectra, kinetic energy maps and maps of dispersion coefficient were obtained for air-saturated water, water with surfactant (sodium dodecyl sulphate, 1mM) and water with sodium dodecyl sulphate / NaCl ([NaCl]=0.1M). Cavitation bubbles cause an increase in dispersion coefficient and acoustic streaming. These effects are not observed in degassed sample. Streaming was most developed in samples with surfactants which also demonstrate a pronounced anisotropy of the dispersion coefficient parallel and normal to the direction of the sound wave propagation. A stabilization of bubble surface and concomitant reduction of bubble coalescence by the surfactant can explain the observed differences.

### ENGINEERING ACOUSTICS & NOISE CONTROL - I

#### **A Tribute to Herbert S. Ribner**

**Dr. Werner Richarz, P.Eng., FASA**

Aeroustics Engineering Ltd., 50 Ronson Dr. Suite 165, Toronto ON, M9W 1B3

Herbert Spencer Ribner, Professor Emeritus, University of Toronto Institute for Aerospace Studies passed away on May 8, 2005. Over a period spanning seven decades Herbert S. Ribner made notable contributions in the field of aerodynamics and acoustics. His interest in acoustics was sparked by the noise from the test-sections of supersonic windtunnels while he was a section head at NASA Lewis (Cleveland). After moving to the then University of Toronto Institute for Aerophysics, he became one of the leading researchers in the field of aerodynamic noise generation and propagation. Under his guidance the 'Aeroacoustics Group' made many contributions towards a better understanding of boundary layer noise, jet noise and sonic booms. The presentation will give an overview of Ribner's approach of combining theory and experiment, and provide some insight into the generation from sound and other sources.

### HEARING AIDS

#### **Using Active Control to Reduce the Occlusion Effect in Hearing Aids**

**Jim Ryan, Steve Armstrong and Betty Rule**

Gennum Corporation

The hearing-aid occlusion effect occurs when the ear canal is occluded by a snug-fitting ear mold. Patients experiencing the occlusion effect often complain about the sound of their own voice when wearing their hearing aids. The traditional solution to this problem is to introduce a vent or opening through the hearing aid from the canal end to the faceplate. Unfortunately, significant amounts of venting can result in feedback problems. This paper describes an alternate solution for the occlusion problem that is based on active noise control techniques. The advantage of this approach is that occlusion can be reduced without the feedback problems caused by large vents.

#### **Combined Impact of a Rapid Feedback Manager and Telephone Use on Acoustic Insertion Gain**

**Donald Hayes, David Eddins, and Carrie Secor**

Unitron Hearing

Abstract: In this study a group of 18 adults were each fitted with 3 hearing aids. The shell styles were equally distributed between Behind-the-Ear, Full Shell, In-the-Canal and Completely-in-Canal. The participants placed a telephone

handset over the aided ear while a series of insertion gain measurements were made of a signal generated through the handset. The stimuli consisted of several recordings of two talker babble filtered into 1000 Hz bands matching the bandpass frequencies of the aids under test. The results show that activating the feedback manager causes frequency specific (1000 Hz increments) increases in available aided insertion gain before the onset of acoustic feedback. The impact on available gain of different microphones, omnidirectional vs directional, and changes to venting, opened vs. closed will also be presented.

## VIBRATION CONTROL

### **Tranverse Vibrations of Centrifugally Stiffened Tapered Beams using Mathematica.**

**Dominic Jackson and S. Olutunde Oyadiji**

University of Manchester, UK

In this paper, Mathematica is used symbolically to derive the governing differential equation of motion of rotating tapered Euler-Bernoulli beams in free vibration by Hamilton's principle and solved by the Frobenius power series method respectively. Two cases of the beam geometry are considered each with a root offset from its axis of rotation. A symmetric cross sectional area is assumed in both cases but one beam tapers linearly along the span in one plane whereas the other tapers linearly in two planes. The non dimensional natural frequencies of vibration are evaluated in the exact sense for a few boundary conditions using a simple Mathematica code. The effect on the stiffness and natural frequencies due to variation of the root offset parameter and taper ratio is also investigated. Due to the very few publications available on the subject, a finite element routine is also used to validate the results obtained and the agreement is excellent. The results demonstrate the reliability and the ease with which symbolic and numerical operations are carried out using Mathematica. The paper serves as a building block for the use of symbolic computation in the vibration analysis of rotating beams, plates and composite structures.

## HEARING SCIENCE & HEARING CONSERVATION

### **Noise exposures caused by respirators in neonatal intensive care**

**A. Mayrand, S. D. Scollie, H. Roukema, and M. Cheesman**

National Centre for Audiology, University of Western Ontario, London, Canada N6G 1H1.

The focus of this study was noise exposure in the neonatal intensive care unit at St. Joseph's hospital in London, Ontario. Noise levels produced by five different modes of ventilation were measured in order to determine if any may be a contributing factor in permanently shifting hearing thresholds in neonates. Noise levels produced across the frequency

spectrum were measured with a probe tube microphone in the ear canal and in the nasopharynx of neonates. Dosimetric measurements and transfer functions were also determined and used to estimate noise exposure from ventilated related sources and resonant properties of low birth weight infants' ear canals.

### **Sound Sense: Saving Our Students' Hearing**

**Gael Hannan**

The Hearing Foundation of Canada, Toronto, Canada.

Research shows that children are experiencing Noise-Induced Hearing Loss which can profoundly impact their academic, social and career success. *Sound Sense: Save Your Hearing for the Music! / Oui à l'Ouïe: Ménagez Vos Oreilles Pour La Musique!* is a hearing conservation program developed by The Hearing Foundation of Canada that will be introduced to 3,000 Ontario elementary schools over the next two school years. The music-themed program complements the core health curriculums and teaches students the 'why', 'when' and 'how' of protecting their hearing. The presentation provides an overview of this entertaining and innovative program that we hope to make available to every middle school student in Canada.

### **Physiologic indicators of auditory dysfunction in children with suspected auditory processing disorders**

**P. Allen, C. Allan, and N. Chelehmalzadeh**

National Centre for Audiology, University of Western Ontario, London, Canada.

The integrity of the auditory system in cases of auditory processing disorders (APD) is often the subject of controversy and uncertainty. From an ongoing study of children with suspected APD, indicators of physiologic dysfunction in the auditory nerve and/or brainstem will be reported. While hearing threshold levels are well within normal limits, a large number of the children participating present with abnormalities in acoustic reflex thresholds (measured bilaterally both ipsi- and contra-laterally to stimulation) and/or abnormal Auditory Brainstem Responses. Findings from over 50 clinical cases will be presented.

## SPEECH SCIENCES I

### **Auditory Perceptual Evaluation Voice Acceptability and Listener Comfort in Non-Normal Vocal Signals**

**Philip C. Doyle<sup>1</sup>, Michelle Durdle<sup>2</sup>, Paul G. Beaudin<sup>1</sup>, and Tanya L. Eadie<sup>3</sup>**

<sup>1</sup>Voice Production and Perception Laboratory, Doctoral Program in Rehabilitation Sciences; <sup>2</sup>Department of Psychology, The University of Western Ontario, London, ON, N6G 1H1; <sup>3</sup>Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, 98195

The use of voice acceptability and listener comfort as auditory-perceptual dimensions of voice and speech signals may serve

as important measures for both laryngeal and non-laryngeal speech communication. This study sought to determine the psychophysical character of listener ratings of acceptability and listener comfort for non-normal speaker samples using direct magnitude estimation (DME) and equal appearing interval (EAI) scaling procedures. The potential relationship across the auditory-perceptual ratings of voice acceptability and listener comfort was also assessed. Experimental voice samples were generated by individuals who had undergone surgical removal of the larynx and who used pulmonary driven “esophageal” vibration as an alternative voicing source. This method of communication is termed tracheoesophageal (TE) speech. Thirty-two normal-hearing naïve listeners assessed 22 adult male TE speech samples for perceived acceptability and listener comfort. Results yielded a significant relationship between auditory-perceptions of both voice attributes. It was also determined that listeners judge acceptability on a metathetic continuum and rate listener comfort on a prothetic continuum. Therefore, auditory-perceptual evaluation of TE voice and speech must be made using DME scale for listener comfort, however, acceptability can be validly scaled using either EAI or DME methods.

## **SIGNAL PROCESSING & NUMERICAL METHODS I**

### **Acoustic Ray Tracing For 3D Environment Simulation.**

**Lu Meng and David Gerhard**

Department of Computer Science, University of Regina,  
Regina, SK S4S 0A2

Inspired by ray tracing algorithms used for 3D graphic rendering, this paper describes progress toward the implementation of a 3D sound environment simulation using acoustic ray tracing. Current 3D sound simulations focus on 3D panning, and on sound effects such as parameter-based reverberation, which often has little or no direct relationship to the simulated space but is chosen from a set of standardized templates. Our approach seeks to simulate a 3D sound environment using only the geometric layout of a virtual room, the location of sound source in that room and the location and orientation of the listener. Ray tracing assumes that light travels in a straight line until reflected or refracted by objects or materials. Acoustic ray tracing makes use of the same principal, taking into account that the frequency of sound is much lower than that of light, and that low frequency sounds propagate differently than middle to high frequency. [work supported by Natural Sciences and Engineering Research Council of Canada]

## **SIGNAL PROCESSING & NUMERICAL METHODS II**

### **A Modal analysis of an Indian Gong**

**Paresh G. Shrivage, S. Parmeswaran, and K.V.deSa**

Electro-Acoustics Research Lab, Dept. of Physics, N.Wadia College, Pune, Maharashtra, India – 411001

Indian gong is a Percussion instrument. It is used in many

*Canadian Acoustics / Acoustique canadienne*

concerts and religious processions for its musical tones. It is made from brass metal with some iron content to get hard metallic tones. It is hammered at centre to get hard tone after repeated intervals. It is approximately 30 cm in diameter and 5-10 mm in thickness. The study has been done for modal analysis of the gong to check modal shapes and tonal qualities. The paper relates to modal analysis of the gong by accelerometer and holographic analysis of the gong for validation of the modal shapes. The methodology of the work is, a circular grid of circles is plotted on the gong and an accelerometer is mounted on every point to check the response of the surface by spectrum analyzer. The Gong is excited by an impact hammer at particular frequency. This procedure is repeated for other points for different frequencies and response is checked for those frequencies. A Time-average holographic study is also done to see the modes of vibration. The modal shapes are found by post processing of the recorded data using software. In conclusion the paper will cover modal analysis of the gong, holographic aspect as a non-destructive tool for vibration analysis and use of whole study for designing musical gong for different materials.

## **PHYSIOLOGICAL AND BIOACOUSTICS**

### **Characteristics of chimney swift in-flight vocalizations.**

**Jenn Bouchard**

Department of Biology, Biological and Geological Sciences Building, University of Western Ontario, London, Ontario N6A 5B7

Among the most aerial of land birds, the highly gregarious chimney swift (*Chaetura pelagica*), is also well known for its unmistakable vocalizations. Despite being a defining feature of the species, the context in which these acoustic signals are given as well as their characteristics remains poorly understood. From June until the end of September 2005, the images and sounds of chimney swifts in flight were simultaneously recorded at various locations in London Ontario. Call characteristics were analyzed using Avisoft-SASLab Pro (Version 4.2). The use of two digital cameras enabled the re-creation of flight paths allowing changes in call characteristics to be observed along with changes in the position of group members relative to each other. In addition, a catalogue of in-flight behaviour and associated vocalizations was developed. Finally, chimney swift call characteristics were compared to those of other bird and bat species.

### **Distress calls in Neotropical bats**

**Juliet J. Nagel**

University of Western Ontario, Department of Biology, Biological & Geological Sciences Building, London, Ontario, N6A 5B7

“Distress calls” are typically viewed as sounds produced by an animal in situations of extreme stress, such as being captured by a predator or held by a researcher. With their notable acoustical abilities, bats provide a good opportunity

to study the form and function of distress calls. By recording the sound produced and examining the situation the bat was in, the aim of this project was to define what a distress call is. Research was conducted at Lamanai Outpost Lodge in Belize, Central America, during May to August, 2005. Distress calls were recorded from bats tangled in mist nets, placed in a wire cage, and held by a researcher. While few bats called during the first two situations, on average almost half of all bats called when held by a researcher, though this varied greatly with species. Calls were then analyzed with BatSound Pro (Pettersson Elektronik AB), and distress calls from the different species were compared. While calls were usually broadband and often showed similar structure, each species had a unique call. Finally, using an ultrasonic speaker, distress calls were played at mist nets. Recordings of echolocation and static were played as controls. Capture rates and species caught were then evaluated to determine if distress calls served to attract or repel other bats from the area.

## ENGINEERING ACOUSTICS II

### Evaluating the Accuracy of Determining Hourly Road Traffic Noise Levels Using Typical Versus Actual Road Traffic Distributions

**Peter VanDelden, Jakub Wrobel, Marcus Li, Chris Palis, and Scott Penton**

RWDI Air Inc., 650 Woodlawn Road West, Guelph, Ont., N1K 1B8

The determination of background sound level in Ontario is frequently determined from road traffic modelling. Limits are based on the lowest hourly value, and therefore road traffic noise modelling requires hourly traffic counts for the entire day. Many jurisdictions do not have complete hourly traffic counts available; instead, only 8-hour counts, peak AM/PM hour counts, or estimated AADT values may be obtained. The Institute for Transportation Engineering (ITE) has published a typical hourly traffic distribution for arterial roadway networks, which may be used as a proxy for the hourly distribution of traffic. The validity of using the ITE distribution for determination of ambient noise levels from rural, secondary, and major arterial roads is assessed. Hourly noise levels predicted using the ITE distribution versus actual traffic distributions from a number of different regions and communities are compared. Availability of appropriate traffic data in Ontario is also surveyed.

### Validation of the STEAM Rail Traffic Noise Prediction Model – Initial Results

**Peter VanDelden, Chris Palis, Darron Chin-Quee, and Scott Penton**

RWDI Air Inc., 650 Woodlawn Road West, Guelph, Ont., N1K 1B8

The “Sound from Trains Environmental Analysis Method (STEAM) algorithms, published by the MOE, is the standard

method for rail traffic noise prediction within Ontario, and is used extensively in other jurisdictions. The algorithms are based on older federal models also still in use (CHMC). A recent project conducted by RWDI allowed for the direct comparison of measured pass-by and overall long-term  $L_{eq}$  levels at various setback distances for a well travelled section of railway. The results showed that the STEAM prediction method consistently over-predicts sound exposures by as much as 7 dB. This paper presents the initial results of the RWDI’s analysis.

## Environmental Considerations for Noise Barriers

**Jason Tsang<sup>1</sup>, Craig Vatcher<sup>2</sup>, Scott Penton<sup>1</sup>**

<sup>1</sup> Rowan Williams Davies & Irwin Inc., 650 Woodlawn Road West, Guelph, Ontario, N1K 1B8

<sup>2</sup> RWDI Air Inc., Suite 1000, 736 - 8th Avenue SW, Calgary, Alberta, T2P 1H4

Noise barrier walls are often used as a low-cost, low-tech solution for industrial noise control. However, wind loading on the barrier and snow accumulation due to the barrier can affect the required structural design of roof systems and the operation of mechanical equipment. This paper illustrates the wind and snow loading issues for noise barriers, and the acoustical effects of mitigation measures used to reduce snow loading effects.

## BIOMEDICAL ULTRASOUND I

### Exploring the mathematical relationship between acoustic and visual speech for facial animation.

**Matt Craig, Pascal van Lieshout, and Willy Wong**

Oral Dynamics Lab, University of Toronto, 160-500 University Avenue, Toronto, M5G 1V7

Speech perception involves two human senses, hearing and seeing. The synchrony between acoustic and visual speech is a strong driving force for the development of facial animators, which attempt to predict facial motion given an acoustic speech input. A number of different models have been proposed for such “speech-driven facial animation”, which vary both in methodological approach and computational expense. This presentation will focus on a computationally and conceptually simple model of audio-visual speech production, one that relates acoustic speech and visual speech through a linear transformation. The model is tested on an unprecedented large number of subjects (N=16), with a large set of generic sentences, as well as sentences composed of highly controlled contexts to elicit specific types of articulator motion. On average, the transformation can predict 70% of facial motion for any input acoustic sentence, but this result is influenced by the context of the sentence. Preliminary evaluations of predicted facial animations show them to be quite realistic for most sentences, suggesting that the simple approach taken here may provide a time-efficient alternative for facial animation as opposed to more computational methods. This research is relevant to animating talking faces for multimedia

or telecommunications, and further may be useful for the rehabilitation of patients who have suffered forms of facial paralysis, such as those suffering from Parkinson's disease.

### **Perceptual correlates of compensation and adaptation to biteblock perturbation in people who stutter**

**Aravind K. Namasivayam and Pascal H. H. M. van Lieshout**

Oral Dynamics Laboratory, University of Toronto, 160-500 University Avenue, Toronto, ON M5G 1V7

The purpose of the study was to test the hypothesis that persons who stutter (PWS) would differ from control speakers in their ability to compensate and adapt to the presence of a bite-block perturbation. To test this hypothesis, speech samples from 5 PWS and 5 fluent speakers spoken at 2 rates of speech (normal and fast) under 3 conditions viz., with and without the bite-block and after 10 minutes of speaking practice with the bite-block in place, were paired up in various combinations to form the test stimuli. These pairs of stimuli were then presented to 17 listeners using an AX paradigm. The listeners rated the speech quality of the test stimulus (X) on how closely it resembles the exemplar (A) on a computerized 10-cm visual analog scale. The results indicated that compensation to the bite-blocks were partial, immediate (no difference between practice and no practice conditions), and similar across groups at normal rates. However, at fast rates of speech, PWS were not able to compensate to the same extent as fluent speakers. These results will be discussed in terms of potential group differences in the use of specific speech motor control strategies.

## **SPEECH SCIENCES III**

### **Reconstructing 3D tongue movement trajectories from multiplanar paced ultrasound scans**

**Heather Flowers, Tim Bressmann, Brent Carmichael, Chiang-Le Heng**

Department of Speech-Language Pathology, University of Toronto, 500 University Avenue, Toronto ON M5G 1V7

Investigations of tongue movement in speech are usually limited to the midsagittal plane (electromagnetic articulography, videofluoroscopy, B-mode ultrasound). However, tongue movement in speech is a three-dimensional process, and the position of the lateral free margins of the tongue cannot be automatically inferred from the midsagittal plane. We present a new method of reconstructing 3D tongue movement of running speech from multiple 2D ultrasound scans. Seven normal participants read a poem to a metronome beat. The tongue surface for each speaker was reconstructed from ultrasound scans in 4 coronal planes. The goal of our research was to identify global indicators of tongue movement that differentiate between individuals. Our preliminary results suggest that normal speakers can be differentiated mainly on the basis of their midsagittal tongue grooving. Some speakers maintain midsagittal grooving (contraction of the GG) throughout a lengthy speech stream

while others adopt a flatter tongue position with a possible center of control (stability) determined by concavity at the tongue blade. Principal component analyses revealed fluctuating boundaries between the functional segments of the tongue in different speakers.

### **Acoustic and psychophysical relationships and the classification of dysphonic voice**

**Philip C. Doyle<sup>1</sup> and Tanya L. Eadie<sup>2</sup>**

<sup>1</sup>Voice Production and Perception Laboratory, Doctoral Program in Rehabilitation Sciences, The University of Western Ontario, London, ON, N6G 1H1

<sup>2</sup>Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, 98195

Efforts seeking to identify relationships between acoustic measures of voice and auditory-perceptual assessments of voice quality have been longstanding in the literature. This study assessed this relationship through perceptual judgments of 30 adults with dysphonic voices and six normal participants by 12 naïve adult listeners. Perceptual features of voice severity and pleasantness were evaluated using direct magnitude estimation (DME) scaling tasks. Additionally, multiple acoustic measures of the vocal signal were obtained for long-term average spectra, glottal noise, as well as measures based on linear prediction. Classification capacity of the two measures was assessed using logistic regression analysis. The data obtained suggest that high levels of classification accuracy may be achieved when both measures are considered. Differential profiles related to classification will be discussed relative to this analysis and the application of such measures to improved classification of normal and dysphonic voice. [Work supported through CIHR].

# Award-winning *Innovation*

## Created for You

With over 60 years as pioneers within the world of sound and vibration, Brüel & Kjær presents its innovative, award-winning, 4th generation of hand-held instruments for sound and vibration measurement. Development of this latest generation – Type 2250 – was instigated and inspired entirely by the requirements of users participating in in-depth workshops around the world. The hardware has been designed to meet the specific ergonomic requirements of users, and the application software covers everything from environmental noise, troubleshooting, and occupational health, to quality control. The software packages can be licensed separately, so you can get what you need when you need it and won't get left behind if your requirements change. This way, the platform ensures the safety of your investment now and in the future. Created, built and made for you personally, you'll find it will make a difference to your work and all your measurement tasks.

## New Sound Recording Option

The Sound Recording Option works with all other software modules, and lets you record measurement signals in order to identify and document sound sources. Recordings are automatically attached to the measurement and kept with it, even after transfer of the data to a PC.

If you need more information please go to [www.type2250.com](http://www.type2250.com)

HEADQUARTERS: DK-2850 Nærum · Denmark · Telephone: +45 45 80 05 00  
Fax: +45 45 80 14 05 · [www.bksv.com](http://www.bksv.com) · [info@bksv.com](mailto:info@bksv.com)

USA: 2815 Colonnades Court · Norcross, GA 30071  
Toll free (800) 332-2040 · [www.BKHome.com](http://www.BKHome.com) · [bkinfo@bksv.com](mailto:bkinfo@bksv.com)

Australia (+61) 2 9889-8888 · Austria (+43) 1 865 74 00 · Brazil (+55) 11 5188-8166  
Canada (+1) 514 695-8225 · China (+86) 10 680 29906 · Czech Republic (+420) 2 6702 1100  
Finland (+358) 9 521300 · France (+33) 1 69 90 71 00 · Germany (+49) 421 17 87 0  
Hong Kong (+852) 2548 7486 · Hungary (+36) 1 215 83 05 · Ireland (+353) 1 807 4083  
Italy (+39) 0257 68061 · Japan (+81) 3 5715 1612 · Korea (+82) 2 3473 0605  
Netherlands (+31) 318 55 9290 · Norway (+47) 66 771 155 · Poland (+48) 22 816 75 56  
Portugal (+351) 21 4711 4 53 · Korea (+82) 2 3473 0605 · Singapore (+65) 377 4512  
Slovak Republic (+421) 25 443 0701 · Spain (+34) 91 659 0820 · Sweden (+46) 8 449 8600  
Switzerland (+41) 44 880 7035 · Taiwan (+886) 2 2502 7255 · United Kingdom (+44) 14 38 739 000  
USA (+1) 800 332 2040 · Local representatives and service organisations worldwide

Type 2250 – *Inspired by Users*

Brüel & Kjær 

## NEWS / INFORMATIONS

### CONFERENCES

*If you have any news to share with us, send them by mail or fax to the News Editor (see address on the inside cover), or via electronic mail to [stevenb@aciacoustical.com](mailto:stevenb@aciacoustical.com)*

#### 2005

12-14 October: Acoustics Week in Canada. London, Ontario, Canada. Web: <http://caa-aca.ca>

17-18 October: Wind Turbine Noise: Perspectives for Control. Berlin, Germany. Web: [www.windturbinenoise2005.org](http://www.windturbinenoise2005.org)

17-21 October: 150<sup>th</sup> Meeting of the Acoustical Society of America JOINT with NOISE-CON 2005, Minneapolis, Minnesota. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

19-21 October: 36th Spanish Congress on Acoustics and 2005 Iberian Meeting on Acoustics. Terrassa-Barcelona, Spain. Web: [www.ia.csic.es/sea/index.html](http://www.ia.csic.es/sea/index.html)

25-26 October: UK Institute of Acoustics Autumn Conference 2005. Oxford, UK. Web: [www.ioa.org.uk](http://www.ioa.org.uk)

27-28 October: Autumn Meeting of the Swiss Acoustical Society. Aarau, Switzerland. Web: [www.sga-ssa.ch](http://www.sga-ssa.ch)

27-29 October: Meeting of the International Society on Therapeutic Ultrasound. Boston, MA, USA. Web: [www.istu2005.org](http://www.istu2005.org)

04-05 November: Reproduced Sound 21. Oxford, UK. Web: [www.ioa.org.uk](http://www.ioa.org.uk)

09-11 November: Australian Acoustical Society Meeting "Acoustics in a Changing Environment". Busselton, WA, Australia. Web: [www.acoustics.asn.au/divisions/2005-conference.shtml](http://www.acoustics.asn.au/divisions/2005-conference.shtml)

14-18 November: XVI Session of the Russian Acoustical Society. Moscow, Russia. Web: [www.akin.ru](http://www.akin.ru)

07-09 December: Symposium on the Acoustics of Pro-Elastic Materials. Lyon, France. Web: <http://v0.intelligence.eu.com/sapem2005>

#### 2006

15-19 May: IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2006). Toulouse, France. Web: <http://icassp2006.org>

5-7 June: 6th European Conference on Noise Control (Euronoise2006). Web: [www.acoustics.hut.fi/asf](http://www.acoustics.hut.fi/asf)

5-9 June: 151<sup>st</sup> Meeting of the Acoustical Society of America, Providence, Rhode Island. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

26-28 June: 9<sup>th</sup> Western Pacific Acoustics Conference. Seoul, Korea. Web: [www.wespac8.com/WespacIX.html](http://www.wespac8.com/WespacIX.html)

### CONFÉRENCES

*Si vous avez des nouvelles à nous communiquer, envoyez-les par courrier ou fax (coordonnées incluses à l'envers de la page couverture), ou par courriel à [stevenb@aciacoustical.com](mailto:stevenb@aciacoustical.com)*

#### 2005

12-14 octobre: Acoustics Week in Canada. London, Ontario, Canada. Web: <http://caa-aca.ca>

17-18 octobre: Wind Turbine Noise: Perspectives for Control. Berlin, Germany. Web: [www.windturbinenoise2005.org](http://www.windturbinenoise2005.org)

17-21 octobre: 150<sup>e</sup> rencontre de l'Acoustical Society of America AVEC NOISE-CON 2005, Minneapolis, Minnesota. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

19-21 octobre: 36th Spanish Congress dur Acoustics et 2005 Iberian Meeting sur Acoustics. Terrassa-Barcelona, Spain. Web: [www.ia.csic.es/sea/index.html](http://www.ia.csic.es/sea/index.html)

25-26 octobre: UK Institute d'Acoustics Autumn Conference 2005. Oxford, UK. Web: [www.ioa.org.uk](http://www.ioa.org.uk)

27-28 octobre: Autumn Meeting de l'Acoustical Society Swiss. Aarau, Switzerland. Web: [www.sga-ssa.ch](http://www.sga-ssa.ch)

27-29 octobre: Meeting de l'International Society sur Therapeutic Ultrasound. Boston, MA, USA. Web: [www.istu2005.org](http://www.istu2005.org)

04-05 novembre: Reproduced Sound 21. Oxford, UK. Web: [www.ioa.org.uk](http://www.ioa.org.uk)

09-11 novembre: Australian Acoustical Society Meeting "Acoustics in a Changing Environment". Busselton, WA, Australia. Web: [www.acoustics.asn.au/divisions/2005-conference.shtml](http://www.acoustics.asn.au/divisions/2005-conference.shtml)

14-18 novembre: XVI Session de l'Acoustical Society de Russian. Moscow, Russia. Web: [www.akin.ru](http://www.akin.ru)

07-09 decembre: Symposium sur l'Acoustics de Pro-Elastic Materials. Lyon, France. Web: <http://v0.intelligence.eu.com/sapem2005>

#### 2006

15-19 mai: IEEE Conference Internationale sur Acoustics, Speech, et Signal Processing (IEEE ICASSP 2006). Toulouse, France. Web: <http://icassp2006.org>

5-7 juin: 6th European Conference on Noise Control (Euronoise2006). Web: [www.acoustics.hut.fi/asf](http://www.acoustics.hut.fi/asf)

5-9 juin: 151<sup>e</sup> rencontre de l'Acoustical Society of America, Providence, Rhode Island. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

26-28 juin: 9<sup>e</sup> Conférence Western Pacific Acoustics. Seoul, Korea. Web: [www.wespac8.com/WespacIX.html](http://www.wespac8.com/WespacIX.html)

3-7 July: 13th International Congress on Sound and Vibration (ICSV13). Vienna, Austria. [Http://info.tuwien.ac.at/icsv13](http://info.tuwien.ac.at/icsv13)

17-19 July: 9th International Conference on Recent Advances in Structural Dynamics. Southampton, UK. Web: [www.isvr.soton.ac.uk/sd2006/index.htm](http://www.isvr.soton.ac.uk/sd2006/index.htm)

13-15 September: Autumn Meeting of the Acoustical Society of Japan. Web: [www.asj.gr.jp/index-en.html](http://www.asj.gr.jp/index-en.html)

17-21 September: Interspeech 2006 - ICSLP. Web: [www.interspeech2006.org](http://www.interspeech2006.org)

18-20 September: International Conference on Noise and Vibration Engineering (ISMA2006). Leuven, Belgium. Web: [www.isma-isaac.be](http://www.isma-isaac.be)

18-21 September: INTERSPEECH 2006 - ICSLP. Pittsburgh, PA, USA. Web: [www.interspeech2006.org](http://www.interspeech2006.org)

28 November – 2 December: 152<sup>nd</sup> meeting, 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii. Contact: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel: 516-576-2360; Fax: 516-576-2377; E-mail: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

3 - 6 December: INTER-NOISE 2006, Honolulu HA, USA (Same Hotel at ASA meeting the week preceeding)

## 2007

17-20 April. IEEE International Congress on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2007). Honolulu, HI, USA

9-12 July: 14th International Congress on Sound and Vibration (ICSV14). Cairns, Australia. Email: [n.kessissoglou@unsw.edu.au](mailto:n.kessissoglou@unsw.edu.au)

27-31 August: Interspeech 2007. E-mail: [conf@isca-speech.org](mailto:conf@isca-speech.org)

2-7 September 19<sup>th</sup> International Congress on Acoustics (ICA2007), Madrid Spain. (SEA, Serrano 144, 28006 Madrid, Spain; Web: [www.ia.csic/sea/index.html](http://www.ia.csic/sea/index.html))

9-12 September: ICA2007 Satellite Symposium on Musical Acoustics (ISMA2007). Barcelona, Spain. Web: [www.ica2007madrid.org](http://www.ica2007madrid.org)

## 2008

23-27 June: Joint Meeting of European Acoustical Association, Acoustical Society of America, and Acoustical Society of France. Paris, France E-mail: [phillipe.blanc-benon@ec-lyon.fr](mailto:phillipe.blanc-benon@ec-lyon.fr)

28 July - 1 August: 9th International Congress on Noise as a Public Health Problem. Mashantucket, Pequot Tribal Nation, (CT, USA). Web: [www.icben.org](http://www.icben.org)

3-7 juillet: 13th Congress Internationale sur Sound et Vibration (ICSV13). Vienna, Austria. [Http://info.tuwien.ac.at/icsv13](http://info.tuwien.ac.at/icsv13)

17-19 juillet: 9th International Conference sur Recent Advances in Structural Dynamics. Southampton, UK. Web: [www.isvr.soton.ac.uk/sd2006/index.htm](http://www.isvr.soton.ac.uk/sd2006/index.htm)

13-15 septembre: Autumn Meeting de l'Acoustical Society du Japan. Web: [www.asj.gr.jp/index-en.html](http://www.asj.gr.jp/index-en.html)

17-21 septembre: Interspeech 2006 - ICSLP. Web: [www.interspeech2006.org](http://www.interspeech2006.org)

18-20 septembre: International Conference sur Noise et Vibration Engineering (ISMA2006). Leuven, Belgium. Web: [www.isma-isaac.be](http://www.isma-isaac.be)

18-21 septembre: INTERSPEECH 2006 - ICSLP. Pittsburgh, PA, USA. Web: [www.interspeech2006.org](http://www.interspeech2006.org)

28 novembre – 2 décembre: 152<sup>e</sup> rencontre, 4<sup>e</sup> Rencontre acoustique jointe de l'Acoustical Society of America, et l'Acoustical Society of Japan, Honolulu, Hawaii. Info: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tél.: 516-576-2360; Fax: 516-576-2377; Courriel: [asa@aip.org](mailto:asa@aip.org); Web: [asa.aip.org](http://asa.aip.org)

3 - 6 December: INTER-NOISE 2006, Honolulu HA, USA (Same Hotel at ASA meeting the week preceeding)

## 2007

17-20 avril. IEEE Congress Internationale sur Acoustics, Speech, et Signal Processing (IEEE ICASSP 2007). Honolulu, HI, USA

9-12 juillet: 14th Congress Internationale sur Sound et Vibration (ICSV14). Cairns, Australia. Email: [n.kessissoglou@unsw.edu.au](mailto:n.kessissoglou@unsw.edu.au)

27-31 août: Interspeech 2007. E-mail: [conf@isca-speech.org](mailto:conf@isca-speech.org)

2-7 septembre 19<sup>e</sup> Congrès international sur l'acoustique (ICA2007), Madrid Spain. (SEA, Serrano 144, 28006 Madrid, Spain; Web: [www.ia.csic/sea/index.html](http://www.ia.csic/sea/index.html))

9-12 septembre: ICA2007 Satellite Symposium sur Musical Acoustics (ISMA2007). Barcelona, Spain. Web: [www.ica2007madrid.org](http://www.ica2007madrid.org)

## 2008

23-27 juin: Rencontre jointe de l'European Acoustical Association, l'Acoustical Society of America, et l'Acoustical Society of France. Paris, France E-mail: [phillipe.blanc-benon@ec-lyon.fr](mailto:phillipe.blanc-benon@ec-lyon.fr)

28 juillet - 1 août: 9th International Congress sur Noise as a Public Health Problem. Mashantucket, Pequot Tribal Nation, (CT, USA). Web: [www.icben.org](http://www.icben.org)

## NEWS

We want to hear from you! If you have any news items related to the Canadian Acoustical Association, please send them. Job promotions, recognition of service, interesting projects, recent research, etc. are what make this section interesting.

## AMERICAN NATIONAL STANDARD ON CLASSROOM ACOUSTICS NOW AVAILABLE AT NO COST

Thanks to a unique partnership of industry supporters: [Armstrong Ceiling Systems](#), [illbruck Architectural Products](#), and [Trane](#), the Acoustical Society of America (ASA) announced today that its popular classroom acoustics standard is now available online at no cost to the user. Officially known as **ANSI S12.60-2002 American National Standard Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools**, this standard provides helpful guidance to design professionals, educational facilities planners, and the general public, as well. This Standard was developed by Accredited Standards Committee S12, Noise, under American National Standards Institute (ANSI) approved operating procedures. The Standard, which brings the US into agreement with the requirements of the World Health Organization as well as many countries around the world, will assist school districts, architects, and building planners in designing classrooms to optimize the ability of children to learn and prosper. The Standard is available through the ASA Standards Store at <http://asastore.aip.org/>. Users will be asked to establish a username and password and accept the End User License before downloading the standard.

### EXCERPTS FROM “WE HEAR THAT”, IN ECHOS, ASA

ASA Fellow **Jan Achenbach**, Walter P. Murphy Professor and Distinguished McCormick School Professor of the Departments of Mechanical Engineering and Civil and Environmental Engineering at Northwestern University was awarded the National Medal of Technology. He was presented this award, the nation’s highest honor for technological innovation, by President Bush in Washington D.C. at the White House on March 14, 2005.

ASA Fellow **Clive Dym**, Professor of Engineering at Harvey Mudd College, has been awarded the Ruth and Joel Spira Outstanding Design Educator Award for “exceptional contributions to design education through widely-cited authorship on engineering design, through sponsorship of the ASA workshops and conference panels, and through enthusiastic mentoring of engineering students in the art and science of design.”

ASA Fellow **Ira Hirsh** has received a Life Achievement Award from the American Auditory Society. The announcement was made at the Society’s 2005 Scientific and Technology Meeting in Scottsdale, Arizona. ASA Fellow **H. Vincent Poor**, George Van Ness Lothrop Professor in Engineering at Princeton University, has been named the 2005 winner of the Distinguished Alumnus Award by the Tau Beta Pi engineering honor society at Auburn University. A \$2000 scholarship will be given in Dr. Poor’s name to a deserving student member of Tau Beta Pi.

The American Institute of Physics **State Department Science Fellowship** represents an opportunity for scientists to make a contribution to U.S. foreign policy. At least one Fellow annually will be chosen to spend a year working in a bureau of the State Department, providing scientific and technical expertise to the Department while becoming directly involved in the foreign policy process. Fellows are required to be U.S. citizens and members of at least one of the 10 AIP Member Societies at the time of application. Please visit <http://www.aip.org/gov/sdf.html> for details. All application materials must be postmarked by November 1, 2005.

### EXCERPTS FROM “SCANNING THE JOURNALS”, IN ECHOS, ASA

Theory predicts abundant production of **acoustic waves in subsurface layers of the Sun**, and such waves are believed by many to constitute the dominant heating mechanism of the chromosphere, according to a paper in the 16 June issue of *Nature*. Such waves are difficult to detect because of disturbances in the Earth’s atmosphere. This paper reports the detection of such waves and numerical simulations to show that the acoustic energy flux of these waves is too low, by a factor of at least ten, to balance the radiative losses in the solar chromosphere. Acoustic waves therefore cannot constitute the dominant heating mechanisms of the solar chromosphere.

There is widespread belief among players and listeners alike that **violins improve with age** and/or playing. Although mechanical measurements show noticeable differences between two violins built from the same wood samples, rankings of the instruments by experienced playing and listening panels showed no statistical differences in the finished instruments, according to a paper in the April issue of *Acoustics Australia*. One instrument had been played regularly and the other had been kept in museum condition.

Tree frog embryos have a remarkable ability to **sense and interpret vibrations**, according to a paper in the July issue of *Animal Behaviour*. Eggs of the red-eyed tree frog usually hatch after seven days, but the embryos can emerge up to 30% earlier to escape a predator’s attack. Upon hatching they drop into the water and, as tadpoles, swim away to safety. They are more likely to hatch when exposed to vibrations recorded from a snake attack than when exposed to recordings of heavy rain. The embryos must therefore be able to distinguish between these different kinds of motion.

The **Australian didgeridoo** is a simple musical instrument that is capable of a spectacular variety of timbres, according to a brief paper in the 7 July issue of *Nature*. Simultaneous measurement of the didgeridoo sound and the acoustic impedance of the player’s vocal tract just inside the lips indicated that the maxima in the envelope of the sound spectrum are associated with minima in the impedance of the vocal tract. This acoustical effect is similar to the production of vowel sounds made during human speech or singing, although the mechanism is different, and leads to the conclusion that experienced players are subconsciously using their glottis to accentuate the instrument’s tonal variation.

New evidence for **bubble fusion** or sonofusion is reported in the May issue of *Nuclear Engineering and Design*. Engineers at Purdue University used the same test chamber filled with deuterated acetone as in previous experiments (see Summer 2004 issue of *ECHOES*) but with californium-252 as a continuous source of neutrons instead of the pulsed source previously used. The acetone was exposed to the neutron source and then bombarded with ultrasound to produce tiny bubbles that expand before imploding.

The March issue of *Acoustical Science and Technology* is a special issue on **Room Acoustics** in honor of RADS 2004 (International Symposium on Room Acoustics: Design and Science 2004) held on the Island of Awaji in April 2004. The symposium, which was a satellite symposium of ICA 2004, included 20 invited oral presentations and 63 poster presentations, of which twelve papers, one technical report, and twelve acoustical letters appear in this special issue.

A novel method of controlling **reflections in a listening room**, using flat panel loudspeakers, is described in a paper in the May issue of the *Journal of the Audio Engineering Society*. Models and implementations are presented for single-channel, two-channel, and five-channel arrangements. The results of a pilot listening test showed that differences in reflection patterns were readily detected by a panel of experienced listeners.

*EURASIP Journal on Applied Signal Processing*, Volume 2005, Issue 9 is a special issue on **Anthropomorphic Processing of Audio and Speech**. Some papers are paid for by the authors and can be downloaded free of charge at <http://www.hindawi.com/journals/asp/volume-2005/issue-9.html>.

### EDITORIAL BOARD / COMITÉ EDITORIAL

ARCHITECTURAL ACOUSTICS: ACOUSTIQUE ARCHITECTURALE:	<b>Vacant</b>		
ENGINEERING ACOUSTICS / NOISE CONTROL: GÉNIE ACOUSTIQUE / CONTROLE DU BRUIT:	<b>Vacant</b>		
PHYSICAL ACOUSTICS / ULTRASOUND: ACOUSTIQUE PHYSIQUE / ULTRASONS:	<b>Werner Richarz</b>	Aeroustics Engineering Inc.	(416) 249-3361
MUSICAL ACOUSTICS / ELECTROACOUSTICS: ACOUSTIQUE MUSICALE / ELECTROACOUSTIQUE:	<b>Annabel Cohen</b>	University of P. E. I.	(902) 628-4331
PSYCHOLOGICAL ACOUSTICS: PSYCHO-ACOUSTIQUE:	<b>Annabel Cohen</b>	University of P. E. I.	(902) 628-4331
PHYSIOLOGICAL ACOUSTICS: PHYSIO-ACOUSTIQUE:	<b>Robert Harrison</b>	Hospital for Sick Children	(416) 813-6535
SHOCK / VIBRATION: CHOCS / VIBRATIONS:	<b>Li Cheng</b>	Université de Laval	(418) 656-7920
HEARING SCIENCES: AUDITION:	<b>Kathy Pichora-Fuller</b>	University of Toronto	
HEARING CONSERVATION: Préservation de L'Ouïe:	<b>Alberto Behar</b>	A. Behar Noise Control	(416) 265-1816
SPEECH SCIENCES: PAROLE:	<b>Linda Polka</b>	McGill University	(514) 398-4137
UNDERWATER ACOUSTICS: ACOUSTIQUE SOUS-MARINE:	<b>Garry Heard</b>	DRDC Atlantic	(902) 426-3100
SIGNAL PROCESSING / NUMERICAL METHODS: TRAITMENT DES SIGNAUX / METHODES NUMERIQUES:	<b>David I. Havelock</b>	N. R. C.	(613) 993-7661
CONSULTING: CONSULTATION:	<b>Corjan Buma</b>	ACI Acoustical Consultants Inc.	(780) 435-9172
ADVISOR: MEMBER CONSEILLER:	<b>Sid-Ali Meslioui</b>	Pratt & Whitney Canada	(450) 647-7339

**Marine Acoustics: Direct and Inverse Problems**  
**James L. Buchanan, Robert P. Gilbert, Armand Wirgin**  
**and Yongzhi S. Xu**  
**SIAM Publishing, 2004**  
**ISBN: 0-89871-547-4, US\$80 (paperback)**

I should state at the outset that as a geophysicist working in ocean acoustics research and a former defence scientist, I may not be the ideal reviewer for a text written by eminent mathematicians and published by SIAM (Society of Industrial and Applied Mathematics). However, my attention was captured by claims on the book's back cover that the text would be of "keen interest" to "engineers and scientists working in ocean acoustics, military scientists interested in submarine detection and long-range communications, and geophysicists involved in locating underwater oil fields." The book further claimed to "address many areas of practical interest related to underwater acoustical imaging, including ecological survey and cleanup, protection of open water harbours, maintenance of offshore petroleum and gas enterprises, and other areas of environmental and military concern." The Preface even includes a paragraph on seabed gas hydrate deposits and their possible impact on global climate. Overall, the impression was that of a broad and accessible text on ocean acoustics. In fact, I found the book to be an esoteric and highly mathematical treatise focused on acoustic scattering theory with almost no mention of practical applications after the Introduction. While this text portrays the state of the art for mathematical specialists in underwater acoustic scattering, in reality I expect it will be of limited interest to most others.

Chapter 1, entitled *The Mechanics of Continua*, is designed to provide all the physics necessary for the remainder of the text. The Introduction and Survey of Previous Work sections at the start of this chapter set the rather heavy-going tone of this book, with literally hundreds of references and a profusion of parenthetical comments. The remainder of the chapter covers continuum mechanics in a succinct but thorough manner. However, despite the fact that this section is billed as an "excellent introduction for the reader unacquainted with theoretical acoustics," those without a strong background in this material will likely find it largely impenetrable as presented, stressing formal mathematical development rather than physical understanding (some schematic diagrams would be welcome here).

Chapters 2 and 3 cover direct and inverse scattering theory in ocean environments with a perfectly-reflecting seabed, a somewhat artificial construct, but one amenable to rigorous mathematical treatment. The goals here are to establish existence/uniqueness of the acoustic forward scattering problem for oceans containing an embedded (idealized) inclusion, and to examine the problem of discerning properties of such an inclusion from the scattered fields. The style of these two

chapters differs notably from that of the first chapter, and they appear to be written explicitly for mathematicians in a formal lemma/theorem-proof format and with a level of abstraction that will not appeal to non-mathematicians. However, these chapters do provide a thorough and rigorous treatment of the subject which is otherwise difficult to glean from the (somewhat dispersed) research literature, and hence should provide a valuable resource to specialists in the field.

Chapter 4 considers a more realistic treatment of the seabed as a penetrable elastic medium, and considers the problem of acoustic-field modelling and of determining the elastic properties of the seabed from scattered fields. Notable here is the lack of mention of the large and diverse literature on geoacoustic inversion methods and experiments, a closely related subject.

Chapter 5 considers a more complex representation of the seabed as a poroelastic medium, and concentrates on the forward scattering problem. This chapter contains some practical and insightful examination and discussion of the increasingly-popular Biot sediment model and variations thereon. Finally, Chapter 6 presents a general (and complicated) mathematical treatment of poroelastic media that includes Biot-like materials as a special case.

Overall, my main concern with this book is that it is a highly specialized mathematical treatment portrayed as a practical generalist text. The book would have been better titled *Marine Acoustics: Direct and Inverse Scattering Problems*, since scattering is only one aspect of the larger fields of direct and inverse problems in ocean acoustics. Finally, while the book claims to be "the first to investigate inverse problems in an ocean environment" it paradoxically includes within its reference list a number of previous texts on precisely that subject (Tolstoy: *Matched-field Processing for Underwater Acoustics*, World-Scientific, 1993; Taroudakis and Makrakis: *Inverse Problems in Underwater Acoustics*, Springer-Verlag, 2001, among others).

**Stan Dosso, Professor**  
**School of Earth and Ocean Sciences**  
**University of Victoria**  
**Victoria BC Canada**

**Managing Noise and Vibration at Work**  
**By Tim South**  
**Elsevier Butterworth-Heinemann, 2004**  
**IX + 268 pp. Price: US\$ 33.73 (paperback)**  
**ISBN: 0-7506-6342-1**

The title of the book has a following sub-title: "A practical guide to assessment, measurement and control." This reviewer will agree that this is a book on management of noise, but certainly not a practical guide to assessment, measurement and control. If it would offer a guide on all those topics, then it will have to be much, much longer and the extent of the content will have to be much wider.

The author has been teaching courses for the Institute of Acoustics' Certificate of Competence in Workplace Noise Assessment and also in Management of Occupational Exposure to Hand-arm Vibration. In a way, this book can be seen as a text-book for students in those courses. As such, it requires some previous knowledge of the basics of noise and vibrations, since many topics are dealt with in a superficial manner. It is definitely good for someone who just needs to know what to do, without digging too much into the "why."

This book will attract, mainly, health and safety professionals and it will be especially useful to Industrial Hygienists. It is easy on the math, something that makes it particularly attractive for someone interested in practical advice. A drawback is that the main focus of the book is on public from the UK – as a matter of fact the entire legislation in the book is from this country. However there is enough material to make it useful even on this side of the ocean.

The book is well presented, with clean and clear graphs and photographs and includes many practical examples that helps understand the different calculations (fortunately not too many, nor too complicated).

There are four parts to the book. The first one, Noise, takes almost a third of the book. It starts with the physics of noise, and continues through human response to noise, noise measurements and noise exposure and exposure assessment. An entire chapter, dedicated to the development of the noise legislation in the UK, will be of little use for our professionals. This reviewer was pleased to see that the exposure is expressed in terms of LEX and Leq, something missing in many books on the same subject. It even includes the calculations for the long term, weekly average exposure. The dose concept is also included and several examples show how to convert dose into Leq and vice versa. Many practical recommendations regarding how to perform a noise exposure assessment, the precautions to be taken and how to manage the resulting data are also included.

Unfortunately, there is no recommendation about how to deal with large group of workers: how many to be tested, for how long and how to treat the measurement results. This is a

rather unfortunate omission (probably there is nothing in the UK legislation or standards on this issue), since in most situations there will be several workers exposed to the same noise environment and testing each and every one of them can be an expensive and time-consuming endeavor. In that respect, we in Canada are lucky to have the CSA Standard Z107.56-94 Procedures for the Measurement of Occupational Noise Exposure that provide guidance on how to proceed in those cases.

The second part of the book deals with Hand-arm vibrations. This is an issue that has been largely ignored or overlooked in many textbooks. Here it is examined in detail and hence, constitutes a useful reference. After examining the vibration phenomenon, the author deals with the hand-arm vibration syndrome. Then, he gets into the subject of the measurement of the vibration. Instrumentation, their use and precautions during the measurements are included in detail. Finally the use of the results for the assessment of the risk and time limits is included. Some examples help understand the process better.

The third part is Whole Body Vibration (WBV). This part is somewhat shorter, due to the fact that there is not much knowledge in that subject. Also, WBV assessment is quite complex since there are two estimates (A and VDV) and no uniform answer of which one to use.

The final part of the book is entitled Reducing noise and vibration risks. It deals with management of noise and vibrations, including some basic noise control. That is where the subject of hearing protection can be found. Being such a popular way of controlling noise exposure, this subject is dealt with quite extensively.

Four appendices and a list of bibliographic references are also included.

In summary, this is a low cost, practical book that will be quite useful to the practicing occupational health and safety professional in his everyday practice.

**Alberto Behar**  
**Adjunct Professor**  
**University of Toronto**  
**Toronto, ON, Canada**

## INSTRUCTIONS TO AUTHORS FOR THE PREPARATION OF MANUSCRIPTS

**Submissions:** The original manuscript and two copies should be sent to the Editor-in-Chief.

**General Presentation:** Papers should be submitted in camera-ready format. Paper size 8.5" x 11". If you have access to a word processor, copy as closely as possible the format of the articles in *Canadian Acoustics* 18(4) 1990. All text in Times-Roman 10 pt font, with single (12 pt) spacing. Main body of text in two columns separated by 0.25". One line space between paragraphs.

**Margins:** Top - title page: 1.25"; other pages, 0.75"; bottom, 1" minimum; sides, 0.75".

**Title:** Bold, 14 pt with 14 pt spacing, upper case, centered.

**Authors/addresses:** Names and full mailing addresses, 10 pt with single (12 pt) spacing, upper and lower case, centered. Names in bold text.

**Abstracts:** English and French versions. Headings, 12 pt bold, upper case, centered. Indent text 0.5" on both sides.

**Headings:** Headings to be in 12 pt bold, Times-Roman font. Number at the left margin and indent text 0.5". Main headings, numbered as 1, 2, 3, ... to be in upper case. Sub-headings numbered as 1.1, 1.2, 1.3, ... in upper and lower case. Sub-sub-headings not numbered, in upper and lower case, underlined.

**Equations:** Minimize. Place in text if short. Numbered.

**Figures/Tables:** Keep small. Insert in text at top or bottom of page. Name as "Figure 1, 2, ..." Caption in 9 pt with single (12 pt) spacing. Leave 0.5" between text.

**Line Widths:** Line widths in technical drawings, figures and tables should be a minimum of 0.5 pt.

**Photographs:** Submit original glossy, black and white photograph.

**Scans:** Should be between 225 dpi and 300 dpi. Scan: Line art as bitmap tiffs; Black and white as grayscale tiffs and colour as CMYK tiffs;

**References:** Cite in text and list at end in any consistent format, 9 pt with single (12 pt) spacing.

**Page numbers:** In light pencil at the bottom of each page. Reprints: Can be ordered at time of acceptance of paper.

## DIRECTIVES A L'INTENTION DES AUTEURS PREPARATION DES MANUSCRITS

**Soumissions:** Le manuscrit original ainsi que deux copies doivent être soumis au rédacteur-en-chef.

**Présentation générale:** Le manuscrit doit comprendre le collage. Dimensions des pages, 8.5" x 11". Si vous avez accès à un système de traitement de texte, dans la mesure du possible, suivre le format des articles dans *l'Acoustique Canadienne* 18(4) 1990. Tout le texte doit être en caractères Times-Roman, 10 pt et à simple (12 pt) interligne. Le texte principal doit être en deux colonnes séparées d'un espace de 0.25". Les paragraphes sont séparés d'un espace d'une ligne.

**Marges:** Dans le haut - page titre, 1.25"; autres pages, 0.75"; dans le bas, 1" minimum; latérales, 0.75".

**Titre du manuscrit:** 14 pt à 14 pt interligne, lettres majuscules, caractères gras. Centré.

**Auteurs/adresses:** Noms et adresses postales. Lettres majuscules et minuscules, 10 pt à simple (12 pt) interligne. Centré. Les noms doivent être en caractères gras.

**Sommaire:** En versions anglaise et française. Titre en 12 pt, lettres majuscules, caractères gras, centré. Paragraphe 0.5" en alinéa de la marge, des 2 cotés.

**Titres des sections:** Tous en caractères gras, 12 pt, Times-Roman. Premiers titres: numéroter 1, 2, 3, ..., en lettres majuscules; sous-titres: numéroter 1.1, 1.2, 1.3, ..., en lettres majuscules et minuscules; sous-sous-titres: ne pas numéroter, en lettres majuscules et minuscules et soulignés.

**Equations:** Les minimiser. Les insérer dans le texte si elles sont courtes. Les numéroter.

**Figures/Tableaux:** De petites tailles. Les insérer dans le texte dans le haut ou dans le bas de la page. Les nommer "Figure 1, 2, 3,..." Légende en 9 pt à simple (12 pt) interligne. Laisser un espace de 0.5" entre le texte.

**Largeur Des Traits:** La largeur des traits sur les schémas technique doivent être au minimum de 0.5 pt pour permettre une bonne reproduction.

**Photographies:** Soumettre la photographie originale sur papier glacé, noir et blanc.

**Figures Scanées:** Doivent être au minimum de 225 dpi et au maximum de 300 dpi. Les schémas doivent être scannés en bitmaps tif format. Les photos noir et blanc doivent être scannées en échelle de gris tifs et toutes les photos couleurs doivent être scannées en CMYK tifs.

**Références:** Les citer dans le texte et en faire la liste à la fin du document, en format uniforme, 9 pt à simple (12 pt) interligne.

**Pagination:** Au crayon pâle, au bas de chaque page. Tirés-à-part: Ils peuvent être commandés au moment de l'acceptation du manuscrit.



**Application for Membership**

CAA membership is open to all individuals who have an interest in acoustics. Annual dues total \$60.00 for individual members and \$20.00 for Student members. This includes a subscription to *Canadian Acoustics*, the Association's journal, which is published 4 times/year. New membership applications received before August 31 will be applied to the current year and include that year's back issues of *Canadian Acoustics*, if available. New membership applications received after August 31 will be applied to the next year.

**Subscriptions to *Canadian Acoustics* or Sustaining Subscriptions**

Subscriptions to *Canadian Acoustics* are available to companies and institutions at the institutional subscription price of \$60.00. Many companies and institutions prefer to be a Sustaining Subscriber, paying \$250.00 per year, in order to assist CAA financially. A list of Sustaining Subscribers is published in each issue of *Canadian Acoustics*. Subscriptions for the current calendar year are due by January 31. New subscriptions received before August 31 will be applied to the current year and include that year's back issues of *Canadian Acoustics*, if available.

Please note that electronic forms can be downloaded from the CAA Website at [caa-aca.ca](http://caa-aca.ca)

**Address for subscription / membership correspondence:**

Name / Organization \_\_\_\_\_  
 Address \_\_\_\_\_  
 City/Province \_\_\_\_\_ Postal Code \_\_\_\_\_ Country \_\_\_\_\_  
 Phone \_\_\_\_\_ Fax \_\_\_\_\_ E-mail \_\_\_\_\_

**Address for mailing *Canadian Acoustics*, if different from above:**

Name / Organization \_\_\_\_\_  
 Address \_\_\_\_\_  
 City/Province \_\_\_\_\_ Postal Code \_\_\_\_\_ Country \_\_\_\_\_

**Areas of Interest:** (Please mark 3 maximum)

- |  |   |   |
|--|---|---|
| 1. Architectural Acoustics               | 5. Psychological / Physiological Acoustic | 9. Underwater Acoustics                   |
| 2. Engineering Acoustics / Noise Control | 6. Shock and Vibration                    | 10. Signal Processing / Numerical Methods |
| 3. Physical Acoustics / Ultrasound       | 7. Hearing Sciences                       | 11. Other                                 |
| 4. Musical Acoustics / Electro-acoustics | 8. Speech Sciences                        |   |

For student membership, please also provide:

\_\_\_\_\_  
 (University) (Faculty Member) (Signature of Faculty Member) (Date)

I have enclosed the indicated payment for:  
 CAA Membership \$ 60.00  
 CAA Student Membership \$ 20.00  
 Institutional Subscription \$ 60.00  
 Sustaining Subscriber \$ 250.00  
 includes subscription (4 issues /year)  
 to *Canadian Acoustics*.

Payment by:  Cheque  
 Money Order  
 VISA credit card (Only VISA accepted)

For payment by VISA credit card:

Card number \_\_\_\_\_  
 Name of cardholder \_\_\_\_\_  
 Expiry date \_\_\_\_\_

**Mail application and attached payment to:**

\_\_\_\_\_  
 (Signature) (Date)

**D. Quirt, Secretary, Canadian Acoustical Association, PO Box 74068, Ottawa, Ontario, K1M 2H9, Canada**



**Formulaire d'adhésion**

L'adhésion à l'ACA est ouverte à tous ceux qui s'intéressent à l'acoustique. La cotisation annuelle est de 60.00\$ pour les membres individuels, et de 20.00\$ pour les étudiants. Tous les membres reçoivent *l'Acoustique Canadienne*, la revue de l'association. Les nouveaux abonnements reçus avant le 31 août s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de *l'Acoustique Canadienne* de cette année. Les nouveaux abonnements reçus après le 31 août s'appliquent à l'année suivante.

**Abonnement pour la revue *Acoustique Canadienne* et abonnement de soutien**

Les abonnements pour la revue *Acoustique Canadienne* sont disponibles pour les compagnies et autres établissements au coût annuel de 60.00\$. Des compagnies et établissements préfèrent souvent la cotisation de membre bienfaiteur, de 250.00\$ par année, pour assister financièrement l'ACA. La liste des membres bienfaiteurs est publiée dans chaque issue de la revue *Acoustique Canadienne*. Les nouveaux abonnements reçus avant le 31 août s'appliquent à l'année courante et incluent les anciens numéros (non-épuisés) de *l'Acoustique Canadienne* de cette année. Les nouveaux abonnements reçus après le 31 août s'appliquent à l'année suivante.

Pour obtenir des formulaires électroniques, visitez le site Web: [caa-aca.ca](http://caa-aca.ca)

Pour correspondance administrative et financière:

Nom / Organisation \_\_\_\_\_  
 Adresse \_\_\_\_\_  
 Ville/Province \_\_\_\_\_ Code postal \_\_\_\_\_ Pays \_\_\_\_\_  
 Téléphone \_\_\_\_\_ Téléc. \_\_\_\_\_ Courriel \_\_\_\_\_

Adresse postale pour la revue *Acoustique Canadienne*

Nom / Organisation \_\_\_\_\_  
 Adresse \_\_\_\_\_  
 Ville/Province \_\_\_\_\_ Code postal \_\_\_\_\_ Pays \_\_\_\_\_

**Cocher vos champs d'intérêt:** (maximum 3)

- |   |                               |  |
|---|-------------------------------|--|
| 1. Acoustique architecturale                | 5. Physio / Psycho-acoustique | 9. Acoustique sous-marine                        |
| 2. Génie acoustique / Contrôle du bruit     | 6. Chocs et vibrations        | 10. Traitement des signaux / Méthodes numériques |
| 3. Acoustique physique / Ultrasons          | 7. Audition                   |  |
| 4. Acoustique musicale / Electro-acoustique | 8. Parole                     | 11. Autre  |

Prière de remplir pour les étudiants et étudiantes:

\_\_\_\_\_ (Université) (Nom d'un membre du corps professoral) \_\_\_\_\_ (Signature du membre du corps professoral) \_\_\_\_\_ (Date)

Cocher la case appropriée:

- Membre individuel \$ 60.00  
 Membre étudiant(e) \$ 20.00  
 Abonnement institutionnel \$ 60.00  
 Abonnement de soutien \$ 250.00  
 (comprend l'abonnement à *L'acoustique Canadienne*)

Méthode de paiement:

- Chèque au nom de l'Association Canadienne d'Acoustique  
 Mandat postal  
 VISA (Seulement VISA)

Pour carte VISA: Carte n° \_\_\_\_\_

Nom \_\_\_\_\_

Date d'expiration \_\_\_\_\_

**Prière d'attacher votre paiement au formulaire d'adhésion. Envoyer à :**

\_\_\_\_\_  
 (Signature)

\_\_\_\_\_  
 (Date)

**Secrétaire exécutif, Association Canadienne d'Acoustique, Casier Postal 74068, Ottawa, K1M 2H9, Canada**

# The Canadian Acoustical Association l'Association Canadienne d'Acoustique



**PRESIDENT  
PRÉSIDENT**

**Stan Dosso**  
University of Victoria  
Victoria, British Columbia  
V8W 3P6

(250) 472-4341  
sdosso@uvic.ca

**PAST PRESIDENT  
PRÉSIDENT SORTANT**

**John Bradley**  
IRC, NRCC  
Ottawa, Ontario  
K1A 0R6

(613) 993-9747  
john.bradley@nrc-cnrc.gc.ca

**SECRETARY  
SECRÉTAIRE**

**David Quirt**  
P. O. Box 74068  
Ottawa, Ontario  
K1M 2H9

(613) 993-9746  
dave.quirt@nrc-cnrc.gc.ca

**TREASURER  
TRÉSORIER**

**Dalila Giusti**  
Jade Acoustics  
545 North Rivermede Road, Suite 203  
Concord, Ontario  
L4K 4H1

(905) 660-2444  
dalila@jadeacoustics.com

**ADVERTISING  
PUBLICITÉ**

**VACANT**

**EDITOR-IN-CHIEF  
RÉDACTEUR EN CHEF**

**Ramani Ramakrishnan**  
Dept. of Architectural Science  
Ryerson University  
350 Victoria Street  
Toronto, Ontario  
M5B 2K3

(416) 979-5000 #6508  
rramakri@ryerson.ca  
ramani@aiolos.com

**DIRECTORS  
DIRECTEURS**

**Alberto Behar**  
**Corjan Buma**  
**Mark Cheng**  
**Nicole Collison**  
**Christian Giguère**  
**Megan Hodge**  
**Raymond Panneton**  
**Vijay Parsa**

(416) 265-1816  
(780) 435-9172  
(604) 276-6366  
(902) 426-3100  
613-562-5800 Ext 3071  
(780) 492-5898  
(819) 821-8000  
(519) 661-2111 Ex. 88947

**WORLD WIDE WEB HOME  
PAGE:**

**<http://www.caa-aca.ca>  
Dave Stredulinsky**

(902) 426-3100

# SUSTAINING SUBSCRIBERS / ABONNES DE SOUTIEN

The Canadian Acoustical Association gratefully acknowledges the financial assistance of the Sustaining Subscribers listed below. Annual donations (of \$250.00 or more) enable the journal to be distributed to all at a reasonable cost. Sustaining Subscribers receive the journal free of charge. Please address donation (made payable to the Canadian Acoustical Association) to the Secretary of the Association.

L'Association Canadienne d'Acoustique tient à témoigner sa reconnaissance à l'égard de ses Abonnés de Soutien en publiant ci-dessous leur nom et leur adresse. En amortissant les coûts de publication et de distribution, les dons annuels (de \$250.00 et plus) rendent le journal accessible à tous nos membres. Les Abonnés de Soutien reçoivent le journal gratuitement. Pour devenir un Abonné de Soutien, faites parvenir vos dons (chèque ou mandat-poste fait au nom de l'Association Canadienne d'Acoustique) au secrétaire de l'Association.

## **ACI Acoustical Consultants Inc.**

Mr. Steven Bilawchuk - (780) 414-6373  
stevenb@aciacoustical.com - Edmonton, AB

## **Acoustik GE Inc.**

M. Gilles Elhadad - (514) 487 7159  
ge@acoustikge.com - Cote St Luc, QC

## **Dalimar Instruments Inc.**

Mr. Daniel Larose - (514) 424-0033  
daniel@dalimar.ca - Vaudreuil-Dorion, QC

## **Eckel Industries of Canada Ltd.**

Mr. Blake Noon - (613) 543-2967  
eckel@eckel.ca - Morrisburg, ON

## **HGC Engineering Ltd.**

Mr. Bill Gastmeier - (905) 826-4044  
info@hgcengineering.com - Mississauga, ON

## **J.E. Coulter Associates Ltd.**

Mr. John Coulter - (416) 502-8598  
jcoulter@on.aibn.com - Toronto, ON

## **JASCO Research Ltd.**

Mr. Scott Carr - (902) 491-4489  
scott@jasco.com - Halifax, NS

## **MJM Conseillers en Acoustique Inc.**

MJM Acoustical Consultants Inc.  
M. Michel Morin - (514) 737-9811  
mmorin@mjm.qc.ca - Montréal, QC

## **OZA Inspections Ltd.**

Mr. Frank Westaway - (800) 667-8263 ex28  
oza@ozagroup.com - Grimsby, ON

## **Pinchin Environmental**

Mr. Weidong Li  
(905) 507-4855  
wli@pinchin.com - Mississauga, ON

## **SILEX Innovations Inc.**

Mr. Mehmood Ahmed - (905) 612-4000  
mehmooda@silex.com - Mississauga, ON

## **Solutions Acoustiques Inc.**

Sylvain Larivière - (514) 793-3767  
sylvainlariviere@solutionsacoustiques.com -  
Laval, Québec

## **Tacet Engineering Ltd.**

Dr. M.P. Sacks - (416) 782-0298  
mal.sacks@tacet.ca - Toronto, ON

## **Wakefield Acoustics Ltd.**

Mr. Clair Wakefield - (250) 370-9302  
nonoise@shaw.ca - Victoria, BC

## **ACO Pacific Inc.**

Mr. Noland Lewis - (650) 595-8588  
acopac@acopacific.com - Belmont, CA

## **Aercoustics Engineering Ltd**

Mr. John O'Keefe - (416) 249-3361  
aercoustics@aercoustics.com - Rexdale, ON

## **Dodge-Regupol**

Mr. Paul Downey - (416) 440-1094  
pcd@regupol.com - Toronto, ON

## **H.L. Blachford Ltd.**

Mr. Dalton Prince - (905) 823-3200  
amsales@blachford.ca - Mississauga, ON

## **Hydro-Quebec**

M. Blaise Gosselin - (514) 840-3000 ex5134  
gosselin.blaise@hydro.qc.ca - Montréal, QC

## **J.L.Richards & Assoc. Ltd.**

Mr. Fernando Ribas - (613) 728-3571  
mail@jlrichards.ca - Ottawa, ON

## **John Swallow Associates Ltd.**

Mr. John Swallow - (905) 271-7888  
jswallow@jsal.ca - Mississauga, ON

## **Novel Dynamics Test Inc.**

Mr. Andy Metelka - (519) 853-4495  
metelka@aztec-net.com - Acton, ON

## **Peutz & Associés**

M. Marc Asselineau - +33 1 45230500  
marc.asselineau@club-internet.fr  
Paris, FRANCE

## **Pyrok Inc.**

Mr. Howard Podolsky - (914) 777-7770  
info@pyrokinc.com - Mamaroneck, NY

## **SNC/Lavalin Environment Inc.**

M. Jean-Luc Allard - (514) 651-6710  
jeanluc.allard@snclavalin.com - Longueuil, QC

## **Spaarg Engineering Ltd.**

Dr. Robert Gaspar - (519) 972-0677  
gasparr@kelcom.igs.net - Windsor, ON

## **Valcoustics Canada Ltd.**

Dr. Al Lightstone - (905) 764-5223  
solutions@valcoustics.com  
Richmond Hill, ON

## **West Caldwell Calibration Labs**

Mr. Stanley Christopher - (905) 624-3919  
info@wccal.com - Mississauga, ON

## **Acoustec Inc.**

Dr. J.G. Migneron - (418) 682-2331  
courrier@acoustec.qc.ca - Québec, QC

## **Bruel & Kjaer North America Inc.**

Mr. Andrew Khoury - (514) 695-8225  
- Pointe-Claire, QC

## **Earth Tech Canada Inc.**

Ms. Deborah Olsen - (905) 886-7022 ex2209  
noisevibration@earthtech.ca - Markham, ON

## **Hatch Associates Ltd.**

Mr. Tim Kelsall - (905) 403-3932  
tkelsall@hatch.ca - Mississauga, ON

## **Integral DX Engineering Ltd.**

Mr. Greg Clunis - (613) 761-1565  
au741@ncf.ca - Ottawa, ON

## **Jade Acoustics Inc.**

Ms. Dalila Giusti - (905) 660-2444  
dalila@jadeacoustics.com - Concord, ON

## **Mc SQUARED System Design Group**

Mr. Wade McGregor - (604) 986-8181  
info@mcsquared.com - North Vancouver, BC

## **Owens-Corning Canada Inc.**

Mr. Keith Wilson - (800) 988-5269  
keith.wilson@owenscorning.com - Milton, NS

## **Michel Picard**

(514) 343-7617; FAX: (514) 343-2115  
michel.picard@umontreal.ca - Brossard, QC

## **Scantek Inc.**

Mr. Richard J. Peppin - (410)-290-7726  
peppinr@scantekinc.com - Columbia, MD

## **Soft dB Inc.**

M. André L'Espérance - (418) 686-0993  
contact@softdb.com - Sillery, Québec

## **State of the Art Acoustik Inc.**

Dr. C. Fortier - (613) 745-2003  
sota@sota.ca - Ottawa, ON

## **Vibro-Acoustics**

Mr. Tim Charlton - (800) 565-8401  
tcharlton@vibro-acoustics.com  
Scarborough, ON

## **Wilrep Ltd.**

Mr. Don Wilkinson - (905) 625-8944  
www.wilrep.com - Mississauga, ON