AES 38th INTERNATIONAL CONFERENCE

Sound Quality Evaluation

Luleå University of Technology

Piteå, Sweden

June 13–15, 2010

If you care about sound quality and are involved in its evaluation, then you cannot afford to miss the AES 38th International Conference, *Sound Quality Evaluation*, to be held June 13–15 in Piteå in northern Sweden near to the Arctic Circle.

This conference, being chaired by Jan Berg, will be entirely dedicated to the perceptual evaluation of sound quality, a complex field that involves a number of diverse disciplines including engineering, psychology, statistics, aesthetics, and language. Classical psychoacoustics has tended to concentrate on studying relatively simple stimuli and their auditory response, whereas sound quality evaluation requires methods that connect more directly with everyday listening experience, perhaps using "real world" program material and contexts. Is listening in a car essentially different from listening in the kitchen or on an iPod? How, for example, can the scientist account for the contextual effects in listening? Is it possible to represent an "average" listener, or should we try to find out what different groups perceive? Should we care what people like to hear, or are we only concerned with what is "correct" sound reproduction?

Increasingly, attempts are being made to emulate the human perception of sound quality using computational models. The challenges of these are considerable, and it is necessary to determine the strengths and weaknesses of these so-called "artificial listeners." How, for example, can we build in the necessary cognitive and behavioral traits of human listeners? What sort of trade-offs can be made between speed and accuracy for different applications, and how can automatic quality evaluation be built into product development and testing processes? There is considerable debate about the value of such models, and data is needed that enables the scientist to understand the limits and benefits of the technology. We can ask whether methods used widely in other areas of sensory evaluation, such as food and beverage, are applicable to audio engineering, or whether we can learn from the work going on in picture quality analysis. Finally there is the question of the contribution of audio to multimodal quality perception and the so-called "Quality of Experience" that arises from the entire user experience of a product or service.

THE TECHNICAL PROGRAM

Francis Rumsey and Søren Bech, papers cochairs, have planned a program of seven sessions: Speech Quality Assessment and Prediction, Quality of Experience, Hearing Aid Sound Quality, Spatial Quality Evaluation, Advances in Test Methodology and Analysis, Pedagogy and Sound Design, and Predictive Models.

Workshops, tutorials, demos, and social events are also being planned. The complete program with abstracts follows on pages 317–323. For online registration go to www.aes.org/events/38/registration.

THE VENUE, AND THE RIVIERA OF THE NORTH

The Department of Music and Media at Luleå University of Technology (see photos) provides one of Sweden's most popular education programs for audio engineers. Its mix of artists, engineers, producers, and scientists produces a creative learning environment, the perfect locale for a technically challenging, enlightening, and highly enjoyable conference. In addition to challenging technical programs, another enormously valuable component of AES conferences is the networking and social interaction that takes place during coffee and lunch breaks, where discussions begun during the technical sessions are continued and expanded upon.

If you have time before or after the conference you can enjoy a visit to the nearby beaches along the coast or sail to the islands in the Piteå Archipelago. During the summer months this area of Sweden gets close to 24 hours of daylight, with temperatures ranging from 20–25°C (68–77°F), so don't forget your sunglasses and sunscreen. In transit to the conference you should visit Stockholm, a spectacular city of 14 islands connected by 57 bridges. The *Love Stockholm 2010* festival of music, food, and a wide range of cultural activities begins on June 6 with National Day celebrations and continues through June 17 when the city will celebrate the wedding of Crown Princess Victoria. Of course, as long as you are in Scandinavia, you may also want to attend the AES 39th Conference, *Audio Forensics: Practices and Challenges*, June 17–19 in nearby Denmark (see conference preview starting on page 324).

Be sure to register early to guarantee your participation at the AES 38th International Conference, June 13–15. Check for conference updates at www.aes.org/events/38.

Photographs courtesy of Piteå Turistbyrå / Piteå Tourist Information

AES 38th INTERNATIONAL CONFERENCE

Sound Quality Evaluation

Piteå, Sweden

2010 June 13–15

This preliminary program is accurate as of press time. See updates at www.aes.org/events/38

Sunday, June 13 **09:00**

OPENING REMARKS

Opening Remarks

Sunday, June 13 **09:15**

PAPER SESSION: SPEECH QUALITY ASSESSMENT AND PREDICTION

- **Speech Quality Assessment for Listening-Room Compensation**—*Stefan Goetze,*[1] *Eugen Albertin,*[1] *Jan Rennies,*[1] *Emanuël A. P. Habets,*[2] *Karl-Dirk Kammeyer*[3]
[1]Fraunhofer Institute for Digital Media Technology (IDMT), Oldenburg, Germany
[2]Imperial College of London, London, UK
[3]University of Bremen, Bremen, Germany

  In this paper objective measures for quality assessment of speech signals are evaluated for listening-room compensation algorithms. Dereverberation of speech signals by means of equalization of the room impulse response and reverberation suppression has been an active research topic within the last several years. However, no commonly accepted objective quality measures exist for assessment of the enhancement achieved by those algorithms. This paper discusses several objective quality measures and their applicability for dereverberation of speech signals focusing on algorithms for listening-room compensation.

- **Sound Quality Evaluation of a Speech Enhancement Algorithm for Hearing Impaired Listeners**—*Young Woo Lee,*[1] *In Yong Kim*[2]
[1]Samsung Electronics Co., Ltd., Suwon, Korea
[2]Hanyang University, Seoul, Korea

  Difficulty in understanding speech in noisy environments has been one of the most common complaints of hearing aid users. Various speech enhancement algorithms have been applied to digital hearing aids to improve speech perception in noisy environments. In this paper a speech enhancement algorithm using both spectral subtraction and companding is proposed for digital hearing aids. Objective evaluation, spectrogram comparison, and subjective evaluation under various environmental conditions were performed. We tested Segmental SNR and LLR (Log Likelihood Ratio) for objective evaluation. For subjective evaluation, mean opinion score (MOS) that represented the global perception score was performed to 10 normal listeners and 9 hearing aid users. Also, we measured the speech reception threshold (SRT) of hearing aid users based on the 50-percent correct point in terms of SNR. Statistical comparisons of all methods were performed using the Wilcoxon signed rank test with 0.05 significant levels. All of the results using objective and subjective evaluation show that the proposed speech enhancement algorithm is beneficial for hearing aid users in diverse noisy environments.

- **Hybrid Model for Non-Intrusive Speech Quality Evaluation in Telephony Applications**—*Adrien Leman,*[1,2] *Julien Faure,*[1] *Etienne Parizet*[2]
[1]Orange Labs / France Telecom R&D, Lannion France
[2]Laboratoire Vibrations Acoustique, INSA de Lyon, France

  The evaluation of the speech quality is based on a multidimensional approach that consists in selecting some perceptive attributes according to degradations caused by the type of telecommunication. The paper presents a model assessing speech quality as a combination of three perceptive features: noisiness, coloration, and continuity. Each one of these dimensions is estimated without reference by hybrid indicators that are a combination between parametric (codec, network statistics) and signal-based indicators, depending on the type of information available at the measuring point.

- **A Framework for Predicting Speech Quality Using Detectability of Multiple Distortions**—*D. Sen, Wenliang Lu*, University of New South Wales, Sydney, NSW, Australia

  This paper proposes a framework for predicting overall speech quality using a multidimensional analysis of individual distortions. The algorithm makes use of a physiologically motivated hydro-mechanical Cochlear Model to convert the speech signal into ➥

a domain that is more representative of what is perceived. Salient features are extracted and compared between the original and degraded representations to analyze the detectability of individual distortions. These are subsequently combined to predict overall quality.

- **On the Need for a Theory for Voice Quality—***Nikolai Kouznetsov, Leigh Thorpe*, Research In Motion, Waterloo, Ontario, Canada

  Starting with Lord Rayleigh's and George Campbell's experiments on phoneme identification 100 years ago, Telecommunications Voice Quality as a discipline has developed some approaches and procedures of evaluating the perceived quality of the sound. However, it is still in its infancy. One of the reasons for this thought to exist is that there is no theory that could explain phenomena in voice quality research. In this paper an attempt is made to delineate briefly an example of such a theory and show how it could help voice quality researchers to think and speak about voice quality-related phenomena.

- **Stationarity Assumption and Frame Segmentation in Objective Quality Evaluation Systems: A Language Dependency—***Sonia Djaziri Larbi, Faten Ben Ali, Meriem Jaïdane*, Ecole Nationale d'Ingénieurs de Tunis, Tunis, Tunisia

  In this paper we use a time-frequency based measure, the stationarity index, to characterize the non stationary behavior of languages. We show that different languages have different non stationary behavior. This is done in order to explain why speech processing systems relying on frame-by-frame analysis in the frequency domain, have a language dependent behavior. This dependency is demonstrated using the case of PESQ, the Perceptual Evaluation of Speech Quality model, which enables objective quality assessment of speech quality.

- **Speech Intelligibility in Binaural Reproductions Compared to Real Life Listening—***Gediminas Vasiliauskas*,[1] *Arne Nykänen*,[2] *Johan Odelius*,[2] *Roger Johnsson*[2]
  [1]Lithuanian University of Agriculture, Noreikišk, Lithuania
  [2]Luleå University of Technology, Luleå, Sweden

  A good reproduction of speech should allow the listener to suppress noise and reverberation as if the sounds were heard in real life. An experiment was designed where room properties and reproduction techniques were varied in a way that allowed evaluation of noise and reverberation suppression based on speech intelligibility measurements. Speech intelligibility was considerably better in real life compared to artificial head recordings presented through headphones. The headphone reproductions did not provide enough information to allow the listener to suppress noise and reverberation as well as in real life. It was found that subjects were very consistent. This makes the method precise and it should be useable for making comparisons of different reproduction techniques.

Sunday, June 13 **14:15**

WORKSHOP 1

Sunday, June 13 **16:15**

PAPER SESSION: QUALITY OF EXPERIENCE

- **Computational Quality Model for IP-Based Audio —***Eugene S. Myakotnykh, U. Peter Svensson*, Norwegian University of Science and Technology, Trondheim, Norway

  This paper proposes a model for the computational assessment of IP-based audio quality in the absence of a reference signal. It is similar in type to the E-model for speech signals and includes two types of impairments causing quality degradation: audio stream encoding and network packet loss. Effects of the impairments were parameterized using a subjective experiment. Because subjective tests for impairment quantification are complex and time consuming, the paper also proposes a methodology for using the objective full-reference PEAQ algorithm to quantify parameters for the developed unreferenced model.

- **QoE Assessment for Broadcast Audio Contribution over IP (ACIP)—***Maxim Graubner, Parag S. Mogre*,[1] *Thorsten Lorenzen*[2]
  [1]Technische Universitaet Darmstadt, Darmstadt, Germany
  [2]IT Systemservice–Radio, Hessischer Rundfunk, Frankfurt, Germany

  Broadcasters will increasingly use audio contribution over Internet protocol (ACIP). So far, no dedicated Quality of Experience (QoE) prediction framework exists. In this paper we present a novel non-intrusive parametric QoE rating framework for such a professional broadband audio communication using Voice over IP (VoIP) technology, based on the extended E-model for telephone networks (ITU-T Rec. G.107). For this, we propose an R-factor scale extension to a maximum value of 157, instead of 129 as is used for wideband. Finally, our QoE rating model provides separated impairment factors for the delay, loss, coding, and bandwidth impairment. The model was developed based on our proposal for instrumental evaluation as well as subjective experience by experts.

Monday, June 14 **09:00**

PAPER SESSION: HEARING AID SOUND QUALITY

- **Validation of Objective Sound Quality Models for Hearing Aids—***Lars Bramsløw, Marcus Holmberg*, Oticon A/S, Smørum, Denmark

  This study investigated three objective sound quality models developed for hearing aid tests with hearing-impaired listeners. The models had been developed and trained before this validation study. Here we present validation results on a range of new signals and new types of signal processing using both normal-hearing and hearing-impaired listeners. The results of the validation were mixed, with some predictions highly correlated with subjective ratings and others poorly correlated. On average, the models seemed to perform quite equally, although specific models performed better on specific data sets.

- **Objective Measures to Quantify the Perceptual Effects of Noise Reduction in Hearing Aids—***Karolina Smeds*,[1] *Florian Wolters*,[1,2] *Anders Nilsson*,[1,3] *Sara Båsjö*,[1] *Sofia Hertzman*,[1]

Arne Leijon[3]
[1]Widex A/S ORCA Europe, Stockholm, Sweden
[2]University of Applied Sciences, Oldenburg, Germany;
[3]KTH, Stockholm, Sweden

Twenty listeners with hearing impairments evaluated three noise-reduction algorithms using paired comparisons of speech clarity, noise loudness, and preference. The subjective test produced results in terms of physical signal-to-noise ratios that corresponded to equal subjective performance with and without the noise-reduction algorithms. This facilitated a direct test of how well a number of objective performance measures corresponded with the subjective test results.

Monday, June 14 **10:30**

PAPER SESSION: SPATIAL QUALITY EVALUATION—PART 1

- **Intuitive Hand Gestures in Measurement of the Perceived Size of an Auditory Image of a Symphony Orchestra—***Tapio Lokki, Heikki Vertanen, Samuel Siltanen*, Aalto University School of Science and Technology, Aalto, Finland

  The perceived size of a symphony orchestra was studied with a novel method by tracking the hands of an assessor with an optical tracking system. The task of the assessor was to draw in the air the apparent auditory source. An experiment with three musical pieces, three concert halls, and three listening positions, was completed by 12 assessors. The results suggest that perception of the size of an orchestra is highly individual. In addition, the proposed nonverbal elicitation technique revealed that a concert hall has a significant effect on the perceived size of the auditory source.

- **Measuring Perceived Distance of Violins—A Direct Scaling Test—***Robert Mores*, University of Applied Sciences Hamburg (HAW), Hamburg, Germany

  Specific violins are attributed to be acoustically intimate. A blind test is designed to answer the question, whether such intimacy can be measured in terms of perceived distance. The perceived distance is measured on some 24 subjects in a blind listening test while two violins are played that have already revealed some unspecific differences in terms of acoustical intimacy. A professional musician plays the violins on discrete positions of a physical scale, while subjects guess the sound origin in a blind test. To explore test design options and violins a few parameters are randomized such as the physical room, the loudness, and the duration of samples. Additionally, intermediate voice references and continuous pink noise are investigated on whether these would possibly boost perceptual differences between violins. Subjects are screened and selected by quality measures for unreliability, discrimination, and disagreement. The test delivers general results for human listening, as well as results for the usability of the test design. In terms of the investigated violins, there is little evidence to support the presumed differences. In conclusion, the perceived physical distance is not a prominent component of the acoustical intimacy of a violin.

- **Gaze as a Measure of Sound Source Localization—***Robert Schleicher, Sascha Spors, Dirk Jahn, Robert Walter*, Deutsche Telekom Laboratories, TU Berlin, Berlin, Germany

  We present a study on utilizing eye movements for acoustic source localization tests. Test subjects had to indicate the presumed location of a hidden sound source with their head unconstrained by either fixating or additionally pointing with a laser pointer. Stimuli varied only in the horizontal plane from +45° (left) to –45° (right). Fixation error was always smaller than error in pointing and remained constant for all source positions, whereas pointing error showed a clear relation to source position with more eccentric positions leading to a higher error. Based on these results we conclude that gaze constitutes a useful measure for sound localization tests.

Monday, June 14 **13:00**

WORKSHOP 2

Monday, June 14 **15:00**

PAPER SESSION: SPATIAL QUALITY EVALUATION—PART 2

- **Improved Prediction of Multichannel Audio Quality-Based on Envelope ITD of High Frequency Sounds—***Jeong-Hun Seo, Inyong Choi, Sang Bae Chon, Koeng-Mo Sung*, Seoul National University, Seoul, Korea

  In the assessment of the multichannel audio coding systems, spatial factors are important as well as timbral factors. A prediction model by Choi that extended the ITU-R Rec. BS.1387-1 to the multichannel audio coding systems with the use of spatial features, ITDDist (Interaural Time Difference Distortion), ILDDist (Interaural Level Difference Distortion), and IACCDist (InterAural Cross Correlation Distortion), is an example. In that implementation, the ITDDists were computed only for the low frequency (below 1500 Hz) sounds and ILD distortions were computed only for the high frequency components. That implementation is reasonable under classical duplex theory. However, in high frequency range, the interaural difference in temporal envelopes is also important in spatial perception, especially in sound localization. In order to investigate the role of such ITD on prediction of perceived spatial quality in a quantitative way, a new model to compute the ITD distortions of temporal envelopes in high frequency components is introduced in this paper. The computed ITD distortions of temporal envelopes in high frequency components were highly correlated with perceived sound quality. Moreover, when the proposed envelope ITD distortion was included in the prediction model as one of the multiple features to predict overall sound quality, it enhanced the overall performance of sound quality prediction compared with the model in [Choi, et al., J. Audio Eng. Soc., vol. 56, pp. 3-17, 2008].

- **Flexible and Intuitive Pointing Method for 3-D Auditory Localization Experiments—***Matthias Frank, Ludwig Mohr, Alois Sontacchi, Franz Zotter*, University of Music and Performing Arts Graz, Austria

  This paper presents an intuitive pointing method for measuring the perceived direction in 3-D localiza- ➥

tion experiments. The method uses a motion tracked toy-gun as a pointing device and can be used from any position in any nearly convex surrounding hull or loudspeaker setup, as the pointed direction is computed from the piercing point of the gun's direction and the surrounding surface. The reference point for the pointed direction can be chosen freely. The computation is implemented in real time open-source software and works with any 6 degrees of freedom tracking system. In this paper the accuracy of the method is measured for the static and hand-held case using 2 tracking systems in 2 rooms/surrounding hulls.

- **State of the Art on the Subjective Assessment of the Spatial Sound Quality—***Sarah Le Bagousse,*[1] *Catherine Colomes*,[1] *Mathieu Paquier*[2]
[1]Orange Labs - France Télécom R&D, Cesson Sévigné, France
[2]Laboratoire d'Informatique des Systèmes Complexes (LISyC), Brest, France

  Sound technologies aim at spatial reproduction, raising new questions about their quality assessment. This paper is a state of the art about the spatial audio quality evaluation. For the audio coding assessment, two subjective ITU-R test methodologies are identified. However they are limited at the overall quality. Several studies have highlighted some features specific to surround sound that could be included in a new quality assessment methodology.

Monday, June 14 **16:30**

PAPER SESSION: PEDAGOGY AND SOUND DESIGN

- **Sound Quality Education: Pedagogical Issues, Concepts, and Practices—***Christopher Reba*, University of New Haven, New Haven, CT, USA

  The high degree of variability in terms of the many aspects of sound quality evaluation can make formation of effective pedagogical practices for the audio educator seem like an onerous task. This paper attempts to identify certain practices and concepts that may aid in more effective teaching methods for audio educators.

- **PSST! Product Sound Sketching Tool—***Reinier J. Jansen, Elif Özcan, René van Egmond*, Delft University of Technology, Delft, The Netherlands

  Design and evaluation of product sounds during the conceptual design phase may lead to a more efficient and effective design process. Sketching is a characteristic activity for this phase, but existing sound design tools are arguably not supporting this activity for product sounds. A tool has been developed that allows inexperienced product sound designers to sketch product sounds during conceptualization.

Tuesday, June 15 **09:00**

PAPER SESSION: ADVANCES IN TEST METHODOLOGY AND ANALYSIS

- **Assessor Selection and Behavior in Individual Vocabulary Profiling of Concert Hall Acoustics—***Antti Kuusinen, Heikki Vertanen, Tapio Lokki*, Aalto University School of Science and Technology, Aalto, Finland

  The selection process of assessors for individual vocabulary profiling is presented. The assessors were selected with a four-phase screening procedure consisting of an online questionnaire, a pure tone audiometry, a test for vocabulary skills, and a triangle test for the discriminative skills of audio stimuli. In addition, the implementation of individual vocabulary profiling method for assessing concert hall acoustics is presented. The multidimensionality of the elicited attributes and the behavior of the assessors is analyzed with principal component analysis. The feedback after the test from assessors is also introduced. Finally, the suitability of direct elicitation of attributes for assessing concert hall acoustics is discussed.

- **eGauge—A Measure of Assessor Expertise in Audio Quality Evaluations—***Gaëtan Lorho*,[1] *Guillaume Le Ray*,[2] *Nick Zacharov*[2]
[1]Nokia Corporation, Helsinki, Finland
[2]DELTA SenseLab, Hørsholm, Denmark

  The eGauge measure for assessor expertise in listening tests is presented. This ANOVA-based approach aims to objectively qualify the goodness of assessors within an experiment in terms of their discrimination and reliability skill and their overall agreement with the panel. A number of datasets resulting from experiments using different standardized mean opinion score (MOS) methodologies are then tested to illustrate this approach.

- **Analysis of Subjective Data from The MPEG Unified Speech and Audio Coding Call for Proposals—***Schuyler Quackenbush*,[1] *Alan Gross*[2]
[1]Audio Research Labs, Scotch Plains, NJ, USA
[2]Alan M. Gross Consulting, Somerset, NJ, USA

  In October 2007 ISO/IEC MPEG issued a Call for Unified Speech and Audio Coding. Eleven technologies were proposed and each was evaluated at nine operating points (i.e., bit rate for either mono or stereo signals) using the MUSHRA test methodology. The test was conducted at seven sites comprising a total of 64,500 individual subjective scores. The paper contrasts a simple 95 percent confidence interval on the mean scores and an analysis of variance (ANOVA) model for the set of data. A deeper analysis validated the assumption of Gaussian distribution for the model errors and the ANOVA methodology gave 30 percent greater power than the simple grand mean analysis. In other words, the ANOVA reduced the 95 percent confidence interval on the grand mean by 30 percent or, conversely, would have permitted fewer listeners in the testing effort in order to get the same results as the simple grand mean analysis.

- **Sound Quality Evaluation of Car Interior Noise Using Brain Magnetic Field—***Shunsuke Ishimitsu,*[1] *Hiromi Nishikawa*,[1] *Kenji Takami*,[2] *Seiji Nakagawa*[3]
[1]Hiroshima City University, Hiroshima, Japan
[2]University of Hyogo, Hyogo, Japan
[3]Advanced Industrial Science and Technology, Osaka Japan

  The production concept of car engine sound has been changing from finding a solution to noise to designing sound. Although many studies have been conducted on creating comfortable car-engine sound, the psy-

choacoustic effects of time-varying rate for accelerating-engine sounds are still unclear. Thus, we investigated the effects of increasing the frequency rate of car interior noise on auditory impressions using psychological and neurophysiological methods. Harmonic complex tones that simulate acceleration noise were used as stimuli. First, relationship between "sporty" feeling from dynamic characteristics from the engine sound and brain magnetic fields was investigated. In this investigation, subjective evaluations were examined using the semantic differential (SD) method. And neuronal activities of the auditory cortex evoked by these stimuli were measured by magnetoencephalography (MEG). The results indicated that that has a significant effect on subjective impressions and on neuronal activities of the auditory cortex. Second, we investigated the relationships between subjective preference and brain magnetic fields for accelerating car interior noise. Subjective evaluations were examined using the paired-comparison method. At the same time, the MEG alpha-waves range (8 to 13 Hz) measurements were made and analyzed using the autocorrelation function (ACF). The results indicate that the effective duration of the ACF of the MEGs of between 8 and 13 Hz lengthens after the presentation of a preferred sound.

- **Perceptual Validation of Binaural Recordings for Mobile Multimedia Loudspeaker Evaluations**—*Gaëtan Lorho*,[1] *Søren Vase Legarth*,[2] *Nick Zacharov*[2]
[1]Nokia Corporation, Helsinki, Finland
[2]DELTA SenseLab, Hørsholm, Denmark

  To study the perceptual validity of binaural recordings from stereo mobile device loudspeakers, a twofold experiment was performed using a descriptive analysis approach. Sixteen listeners developed their own set of attributes to describe nine real devices and the associated binaural recordings replayed over headphones for three musical items. The results show that the perceptual characteristics of the systems tested in the two scenarios are comparable.

Tuesday, June 15 **13:00**

PAPER SESSION: PREDICTIVE MODELS

- **Estimates of Perceived Spatial Quality across the Listening Area**—*Philip Jackson*,[1] *Martin Dewhirst*,[1] *Rob Conetta*,[2] *Slawomir Zielinski*[1]
[1]University of Surrey, Guildford, UK
[2]London South Bank University, London, UK

  This paper describes a computational model for the prediction of perceived spatial quality for reproduced sound at arbitrary locations in the listening area. The model is specifically designed to evaluate distortions in the spatial domain such as changes in source location, width, and envelopment. Maps of perceived spatial quality across the listening area are presented from our initial results.

- **Developing Content-Specific Parameters of Room Acoustical Quality Using a Binaural, Nonlinear Auditory Model—***Jasper van Dorp Schuitman, Diemer de Vries*, Delft University of Technology, Delft, The Netherlands

  Objective acoustical parameters were developed using the outputs of a binaural, nonlinear auditory model. Results from various listening tests show that these parameters can have a high correlation with the human perception of various aspects of room acoustics. Furthermore, using the proposed framework the acoustics of a room can be assessed by evaluating arbitrary binaural recordings instead of measured impulse responses. This way, the temporal and spectral properties of the signal as well as the absolute sound pressure level are automatically taken into account in the assessment.

- **Multi-Criteria Subjective and Objective Evaluation of Audio Source Separation—***Valentin Emiya,*[1] *Emmanuel Vincent*,[1] *Niklas Harlander,*[2] *Volker Hohmann*[2]
  [1]INRIA-IRISA, Rennes Cedex, France
  [2]Carl von Ossietzky-Universität Oldenburg
  Oldenburg, Germany

  In this paper we address the problem of assessing the perceived quality of estimated source signals in the context of audio source separation. These signals involve different kinds of distortions depending on the considered separation algorithm, including distortion of the target source, interference from other sources or musical noise artifacts. A new MUSHRA-based subjective test protocol is proposed to assess the perceived quality with respect to each kind of distortion and collect the scores of 20 subjects over 80 sounds. Subsequently, the contribution of each type of distortion to the overall quality is analyzed. We propose a family of objective measures aiming to predict the subjective scores based on a decomposition of the estimation error into several distortion components. We conclude by discussing possible implications of this work in the field of 3-D audio quality assessment.

- **Sound Quality Assessment of Earphones: A Subjective Assessment Procedure and an Objective Prediction Model—***Sang Bae Chon, Inyong Choi, Koeng-Mo Sung*, Seoul National University, Seoul, Korea

  This paper describes a subjective assessment procedure and an objective prediction model for sound quality assessment of earphones. The proposed subjective assessment procedure is designed to provide an instantaneous switching from one stimulus to another or reference stimulus. Based on the result from subjective assessment, an objective prediction model was designed using Average Distorted Block (ADBB). Eight earphones were assessed by the proposed subjective assessment procedure and the perceived Mean Opinion Score (MOSs) were used in the regression process in ADBB. The result shows that the proposed objective prediction model achieves an outstanding performance with a cross correlation of 0.9 and root mean square error of 7.57 in 100 point scale.

Tuesday, June 15 **15:00**

CLOSING REMARKS