

AES 38th International Conference

13–15 June 2010
Piteå, Sweden



Not far from the Arctic Circle, in Northern Sweden, lies Piteå – a small but thriving town that is home to the prominent School of Music and Media. As part of Luleå University of Technology, the school runs courses in topics such as sound engineering, radio journalism, and “experience production” alongside more traditional conservatory options and jazz musician training. With an active AES student section and dedicated sound engineering staff, this was an ideal venue for the AES 38th International Conference, *Sound Quality Evaluation*, June 13–15. Being close to the summer solstice, the sky stayed light the entire night at this latitude, which added to the novelty of the experience for those present.

Chaired by Jan Berg, the conference hosted delegates from countries as far afield as Korea, the USA, Japan, Tunisia, and European nations (see seekers of improved sound quality in background photo). Søren Bech and Francis Rumsey, papers cochairs, had selected 28 papers from a large number of submissions. The papers covered topics such as speech quality, spatial quality, quality of experience, hearing aid quality, and predictive models.

Lars Hallberg led a dedicated facilities crew consisting of students from the School who worked long hours to ensure the smooth running of the entire event. Of particular note was the sterling work of Lania Sitepu, who took responsibility for the onsite coordination of the conference.

Two workshops included in the program had been the responsibility of Nick Zacharov. These workshops enabled delegates to witness some table-top demonstrations and to discuss future challenges in quality evaluation. Special thanks go to Delta SenseLab of Denmark for its sponsorship of the conference.

Key themes emerging from the conference concerned the reliability of so-called objective models for quality prediction, the contextual validity of listening tests, and the relevance of standard scales, as well as the importance of choosing relevant quality targets or references when conducting experiments. For those who spent three or four days in this intimate setting there was ample opportunity to make new friends and business contacts, and to learn from each other’s widely varying expertise.



PAPERS

Speech quality

Speech communication technology is gradually becoming more sophisticated, and there is a degree of convergence between this and high-quality audio applications these days. Seven papers on the first day tackled topics on speech quality evaluation. The first was given by Stefan Goetze, who considered how to evaluate the quality of listening room compensation (LRC) or dereverberation algorithms. Goetze concluded that channel-based measures of the dereverberation filters employed resulted in higher correlations between objective and subjective metrics. However, if such measures could not be employed, then those incorporating sophisticated auditory models should be used, such as the Perceptual Similarity Measure (PSM). Further speakers in the field of speech quality looked at nonintrusive evaluation models, language dependency of quality prediction, multidimensional prediction, and the effects of binaural reproduction on speech intelligibility. The last of these concluded that binaural reproduction can impair some of the important spatial cues needed for reverberation suppression and source segregation, thereby making intelligibility lower than in real-life situations. Nikolai Kouznetsov's presentation on a theory of voice quality contrasted the properties of the signal



Jan Berg
conference chair



Francis Rumsey (left), papers cochair, and Diemer de Vries, AES president

with the capabilities and preferences of the listener, suggesting that in some cases voice quality seems to be judged independently of content intelligibility, whereas in others there seems to be a strong relationship between the two. This seems to depend on the type of person doing the evaluation, making it difficult to devise universally objective models.

Quality of experience

It would appear that quality of experience (QoE) may or may not be different from other concepts of audio quality—a topic that also received some attention in the second conference workshop. In network audio systems QoE seems to be used to evaluate the overall quality effects of signal transmission and reception. In the two papers presented on this topic, the authors attempted to evaluate the effects of packet loss when audio is transferred over IP networks, in addition to the effects of low bit-rate encoding. The so-called E model is a computational quality model that has been applied in these contexts, which assumes that psychological factors affecting quality perception are additive. The E model combines the effects of various network and terminal parameters, as well as audio encoding effects. In another paper Eugene Myakotnykh described a reference-free approach capable of evaluating audio stream encoding and network packet loss, which appeared to result in good mapping between objective and subjective ratings on a MUSHRA scale. However, he suggested that the traditional MUSHRA low anchor (a stimulus low-pass filtered to 3.5 kHz) might not be the most relevant in these sorts of tests. (Others also questioned the relevancy of the standard anchor stimuli in tests of different kinds, because the choice of anchors is so influential on the results of experiments and because they can have characteristics very different from the types of impairment presented.) Maxim Graubner also discussed a nonintrusive parametric predictor for audio contribution over IP, suggesting that better models should include additional factors such as loss burstiness, packet size, and coding bitrate. He also proposed that in certain extreme use situations the E-model advantage factor A could be included, to take account of the fact that audio quality can take a back seat when someone in an emergency situation wants to contribute an audio commentary over an IP network, say using a mobile phone. Under such circumstances the con-

text suggests the need for a different emphasis in the metrics that contribute to QoE, prioritizing immediacy and actuality over absolute audio quality. This pointed up another important outcome of the conference—the increasing need for contextual awareness in quality prediction.

Hearing aid quality

There was considerable interest in the quality of sound for hearing impaired listeners at the 38th Conference, with at least three presenters tackling some direct aspects of the problem, both in the speech quality session and the hearing aid quality session. Until recently

the research in this area has been more concerned with improving intelligibility than with quality per se, but now that devices are becoming more advanced there is an increasing interest in optimizing sound quality from a broader point of view. Two representatives from Scandinavian hearing aid research centers were present and able to offer their considerable experience of the subject. Yong Woo Lee from Samsung described a speech-enhancement algorithm based on modified spectral subtraction and companding, which was designed to reduce background noise in hearing aids without introducing undesirable distortion. He and his colleagues had attempted to perform both objective and subjective measurements of quality in different types of noise background such as white noise, speech babble, and car interior noise—all of which have differing challenges. Both MOS (mean opinion score) and SRT (speech recognition threshold) tests were conducted with hearing aid users and the new method showed good noise-suppression results, with low “musical” noise artifacts and minimal distortion.



Karolina Smeds poses a question to a presenter.

Lars Bramslow of Oticon evaluated a number of different quality-prediction models for hearing aids, attempting to validate their performance on new types of hearing aid processing not used in the original model calibration phase. Such processing included methods for frequency shifting parts of the spectrum to make them more accessible to the hearing impaired. A predictive quality model known as PHAQM, which is based on the Perceptual Audio Quality Measure (PAQM) by Beerends and Steemerink, was found to perform much better than two other models, PEMO-HA and MCHI-R. However, Bramslow concluded that none of the models tested was yet sufficiently mature for hearing aid development purposes. Also the training stimuli for the validated models had not employed the sort of frequency processing used in the current tests, making this a very tough test.

Following in a similar vein, Karolina Smets, director of research for Widex A/S ORCA Europe, looked into objective measures that can characterize perceptual effects in hearing aids. Twenty listeners with hearing impairments rated noise loudness, speech clarity, and preference, and these results were compared to a number of physical measures. The standard speech quality evaluation model, PESQ, was tried, as well as a composite measure designed to evaluate overall quality, signal distortion, and background noise distortion. A third measure of signal-to-distortion ratio (SDR) was also compared. Although used outside their intended scope, the first two proved good at ranking three hearing aid noise-reduction algorithms, when compared with subjective ratings, although the absolute ratings were hard to interpret.

Spatial quality

Six papers on spatial quality evaluation dominated the second day of the conference, with the first group of authors looking into novel ways of measuring spatial scene components. Tappio Lokki and his colleagues from Finland evaluated the perceived size of a symphony orchestra based on intuitive hand gestures of assessors, while Robert Schleicher presented a work on utilizing eye movements as a measure of sound source localization. Robert Mores found that listeners in an experiment designed to measure the perceived distance of violins exhibited quite good abilities at judging absolute distance when using the direct scaling method he employed, somewhat contradicting received wisdom. However, perceived distance seemed not to be closely related to the perception of an acoustically intimate sound.

Pedagogy and sound design

Christopher Reba from the University of New Haven, talked about his experiences of trying to teach sound record-

ing students about sound quality, having found that spectrum visualization tools were of considerable help in this respect. Product sound design was another aspect of the quality debate covered by Reiner Jansen from Delft University, who discussed a product sound sketching tool (PSST) intended to enable inexperienced sound designers to sketch product sounds during conceptualization.

Advances in test methodology

New approaches to selecting quality assessors can help experimenters choose subjects for their panels, and Antti Kuusinen showed how a screening procedure can be used in experiments based on individual vocabulary profiling. Subjects develop their own attribute lists for describing stimuli (in this case concert hall acoustics) and can be screened for their descriptive and discriminative skills. Multidimensional analysis can help to show the underlying perceptual structure of their responses, and a limited number of attribute groups appear to emerge. In order to quantify the reliability of listening panelists, a tool known as eGauge can be used, according to Gaëtan Lorho of Nokia. By employing an underlying ANOVA model on listener MOS ratings, relatively independent metrics for reliability, discrimination, and agreement can be generated, but poor discrimination ability cannot be easily distinguished from a lack of differences between the stimuli in question. ANOVA models also came in for detailed scrutiny by Schuyler Quackenbush of Audio Research Labs, who showed ways in which such models can improve the ability to

1nV Low Noise Dual JFET

Low Noise < 1nV
 Monolithic Dual
 Tighter Idss Matching
 Narrow Idss Grades
 Low Capacitance: 20pf
 Functional Replacement for 2SK389

www.linearsystems.com 800-359-4023

discriminate between codecs in MPEG subjective tests. Apparently the MPEG community has relied heavily on “grand mean” comparisons in the past, which are somewhat crude from a statistical point of view. ANOVA reveals interesting interactions in the data set and can lead to more powerful discriminations between stimuli, or the need for fewer observations to achieve a certain confidence interval. Multiple comparison tests can be controversial in such a hotly contested arena, owing to the number of possible corrections that can be applied in such tests and the difficulty of choosing the right one for the job.

If you don't trust the listener to give you a truthful verbal answer to a question about preference, or regard conventional approaches to quality evaluation as unreliable, then perhaps the answer is brain magnetic field measurements. Shunsuku Ishimitsu of Hiroshima City University attempted to evaluate the effect of different vehicle acceleration sounds on brain alpha wave generation. Quite high correlations were found between various alpha wave features and subject preference, suggesting that such measures can be a useful additional tool in the evaluation of industrial processes. Alpha waves are normally said to correspond to a state of mental well-being or relaxation and have traditionally been associated with the occipital lobe and visual perception. In these experiments he found a notable correlation between certain alpha wave features and the “sporty” feeling generated by car acceleration sounds, which was a curious outcome in the view of some. Further tests are planned, looking at other types of brain magnetic field response.

When evaluating binaural reproduction as a means of presenting stimuli in an evaluation of mobile multimedia loudspeakers, Gaëtan Lorho found that while attributes relating to timbre, discrimination, and artifacts were well represented, spatial features such as width were somewhat compromised in binaural reproduction without head tracking. Because listeners used individual vocabularies to describe perceived attributes of audio quality, it was necessary to use a multivariate data analysis method known as hierarchical multiple factor analysis (HMFA) to interpret and compare the results. This proved to be a powerful approach capable of generating a global sensory description of the stimuli under test.

Predictive models

Measuring the perceived spatial quality of a reproduced sound scene was the theme of Martin Dewhirst's paper on Tuesday, starting off a session on predictive models. Martin explained how test signals can be employed to probe the background and foreground elements of spatial distortion when comparing an impaired process with a reference process, leading to a prediction of perceived spatial quality that has good correlation with subjective ratings. Maps of spatial quality across the listening area can also be generated, showing how off-center listening positions affect it for different types of audio processing. Jasper van Dorp Schuitmann went on to show how he had been able to derive measures of room acoustical quality from ordinary musical signals using a binaural auditory model. This enabled acoustical quality of, say, a concert hall to be evaluated during a live performance, and he had found that relatively simple peak detection and measurement methods



Author Shunsuku Ishimitsu presents his paper on the effect of vehicle acceleration on brain magnetic fields.



Author Martin Dewhirst speaks on spatial quality prediction.

can be used to separate direct and reverberant parts of the soundfield to derive relevant metrics.

Source-separation algorithms often give rise to unpleasant results, including distortion of the target source, interference, and artifacts. In the work of Valentin Emiya, four new rating categories were developed for evaluating the quality of such algorithms, likely to be less ambiguous to naive subjects—namely global quality, preservation of target source, suppression of other sources, and absence of additional artificial noise. Objective measures of distortion, analyzed using the recently proposed PEMO-Q audio quality measure, were then extracted and a mapping function was trained by the sub-



Delegates explore the rapids at Storforsen (above) and experience the variable acoustics of the new concert hall recently built for the School of Music and Media in Piteå (right).



The student support crew at Piteå: from left, Linda Iro Näsström, Erik Sikström, Johannes Oscarsson, Erik Johansson, Björn Karlsson, and Lania Sitepu (conference coordinator).

AES 38th International Conference

jective data. Overall quality was well predicted, whereas the target distortion and additional noise criteria were harder to predict from the metrics.

Sang Bae Chon from the Seoul National University explained a way of assessing earphones by making binaural recordings of sound reproduced over them using a dummy head. Various corrections for the ear canal response and method of mounting were used and the result monitored using Sennheiser HD650 headphones, so that the listener could switch freely between the stimuli in a MUSHRO-style experiment. An objective quality-assessment model for these earphones was developed based on an auditory model and one of the model output variables, known as Average Distorted Block, used in the IUD-R 1387 (PEBA) quality model. This was used to

estimate the perceived differences between the test and reference stimuli, and the resulting correlation between actual and predicted scores was high (0.98) while the prediction error appeared low; however, only eight points were available for this correlation plot.

WORKSHOPS AND TOURS

A table-top demonstration on the first day enabled three conference participants to explain their systems to delegates, during which time other groups were offered a chance to visit the facilities at the conference venue. Stefan Goltz demonstrated his Speech Intelligibility Prediction (SIP) toolbox, designed to enable quick and easy prediction of speech quality in all acoustics conditions. Schuyler Quackenbush explained his STEP program for subjective quality testing, while Torben Pedersen showed Delta's SenseLabOnline system, which is a rapid solution for web-based listening tests.

The studio facilities of the School of Music and Media were appreciated by everyone as examples of high-quality resources for students in Piteå, and delegates were stunned by the magnificent new concert hall built in the Studio Acusticum a few years ago. The hall is in



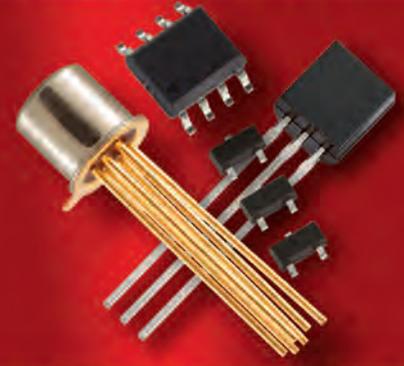
Torben Pedersen (right) of Delta SenseLab (conference sponsor) demonstrates his company's online subjective testing tools.



Francis Rumsey takes a bow at the end of his organ recital.

JFETS

Second Source Replacement for Siliconix



Monolithic Duals-JFETS

- SST/U401-406 Series
- SST/U441 Series
- U421-423 Series
- U430-431 Series

Single-JFETS

- SST/2N4391-4393 Series
- SST/J174-177 Series
- SST/2N4117A-4119A Series
- SST/2N4416

Current Regulating Diodes

- J500-J511 Series
- SST502-SST511 Series

Low Leakage Diodes

- DPAD1-50 Series
- JPAD/PAD1-50 Series
- SSTPAD5-100 Series
- SSTDPAD5-100 Series

Features

Surface Mount & Thru Hole
Special Selections Available
Lead-Free/Rohs Compliant

1-800-359-4023

LINEAR SYSTEMS

www.linearsystems.com



Delegates enjoy an excellent dinner hosted by a representative of the town of Piteå at the Bryggargartan Restaurant.

the process of having a large concert organ installed (see <http://acousticmorgan.com>), built by the German company Woehl, which will fill a large hole at the back of the stage. Moreover, the hall has variable acoustics using a movable ceiling and absorbent panels, which conference delegates were able to experience. The volume of the hall can be altered by almost one third and the reverberation time changed from some 2.7 seconds down to something in the region of one second, for multipurpose event management.

Francis Rumsey chaired a workshop on the second day, during which panellists Nikolai Kouznetsov, Gaetan Lohro, Tapio Lokki, Karolina Smeds, and Schuyler Quackenbush debated questions about challenges in sound quality evaluation. They discussed the use of laboratory conditions versus real-world, field, or behavioral observations when trying to find out people's genuine responses to sound quality; also whether methods adopted by the food and beverage industries are appropriate for audio and what we might understand by the term quality of experience. The future of predictive models was also briefly examined, considering how to make them more context- and task-aware. Interestingly, the hearing aid industry already uses some field trial methods in which they are able to monitor usage of the hearing aid, which settings are used for how long, and so forth, although there are ethical questions surrounding issues such as recording environment sounds and geographical mapping of users.

SPECIAL EVENTS

A trip to the grand rapids, known as Storforsen in Swedish, provided an excellent escape from indoor environments on the evening of the first day. This roaring monster carries millions

of liters of water per hour down from the mountains on the Piteå river, which was used for transporting logs. The dinner at a nearby hotel was enjoyed by all, and the student team wrote an amusing "audio geek" quiz as after-dinner entertainment. Diemer de Vries, AES president, offered a vote of thanks to all who had made the conference such a great success.

Taking advantage of the conference's presence in such a well equipped music conservatory, Francis Rumsey's organ recital on Monday evening provided delegates with a chance to listen to some live music for half an hour. Demonstrating the wide range of sounds of the fine neoclassical instrument installed in the Orgelsal, the recital included Bach's "Toccatà, Adagio and Fugue in C", and Lionel Rogg's "Partita on Nun Freut Euch," after which all repaired to a fine local restaurant (Restaurant Bryggargatan, Brewer Street Restaurant) for a tasty buffet dinner provided by a local, prize-winning Icelandic chef. Hosted by a representative of the town of Piteå, this dinner provided a chance for friendships to be cemented and good conversation to be exchanged.

Closing the conference, Jan Berg thanked all who had contributed to making it such a success, noting both the breadth and depth of the discussions that had taken place over the past few days. He expressed his hope that the AES will once again visit Piteå in a few years to address the evolution of this important topic.

Editor's note: The conference papers are available for purchase as a book or as a downloadable PDF at www.aes.org/publications/conferences. Individual conference papers can also be obtained from the AES Electronic Library at www.aes.org/e-lib.