



John Oh
conference chair



Bob Moses
AES president



Byungki Oh
keynote speaker



AES 34th INTERNATIONAL CONFERENCE

Jeju Island, Korea

August 28–30, 2008

New Trends in Audio for Mobile and Handheld Devices

The AES 34th International Conference, *New Trends for Mobile and Handheld Devices*, was held at Hana Hotel on Jeju Island, Korea, August 28-30, 2008. This conference was a follow up to the successful 29th Conference, *Audio for Mobile and Handheld Devices*, held in Seoul, Korea in 2006. 67 participants from 14 countries listened to 17 paper presentations and 3 invited lectures on topics ranging from advanced speech coding technologies to new user interface paradigms. Many of the hottest consumer electronics products on the

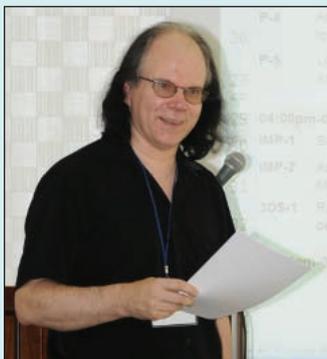
market today are mobile devices such as portable music players, multifunction mobile telephones, and game players. As these devices shrink in size, the computing power necessary to support their expanding multimedia features and expectations for higher performance are increasing exponentially. Audio quality, in particular, is a paramount concern as music delivery continues to be the “killer app.” Thus, mobile devices represent one of the greatest challenges in the engineering profession. Consumers demand new products that are faster, better, cheaper, smaller, and that have longer battery life. This conference brought together many leaders in this field to share the latest advances in mobile technologies and business models.

Known as the “Island of the Gods,” Jeju is the largest of Korea’s islands and is a popular destination for vacationers throughout Asia. The tropical paradise offers visitors a staggering number of outdoor activities and culinary delights. Relatively isolated from the rest of the world, Jeju’s natural environment has been largely preserved since ancient times. Mt. Halla rises 1950 meters above sea level in the center of the island, and 368 secondary craters are scattered throughout its landscape. The island’s volcanic geology, frequent rains, and temperate climate, are very comparable to the Hawaiian Islands.





Authors



Among the authors presenting papers were, from left, top row, Kewei Yang, Myung-suk Song, Joachim Thiemann, and Yangwon Jung; middle row, Juha Backman, Mohammed Chalil, Choong Sang Cho, and Te-won Lee; bottom row, Nathan Bentall, Sang-keun Oh, and Sang Ha Park.

Besides tourism, Jeju's main industries are fishing and agriculture. Of particular note are the legendary Haenyo (diving women) who swim to incredible depths (up to 20 meters) in search of shellfish, abalone, seaweed, and other marine life with nothing more than a hoe, spear, and gourd. They often stay under water for three or four minutes without any breathing apparatus, appearing on the surface just long enough to take in a deep breath of air before diving underwater again. The Haenyo tradition has been passed down from generation to generation and is one of Jeju's most celebrated customs. Thus, the setting for this conference was an interesting juxtaposition of natural wonders, ancient culture, and modern electronics.

DAY 1

Conference chair John Oh kicked off the conference on Thursday by welcoming attendees and reassuring participants that they would be rewarded with an excellent program if they could avoid the allure of the nearby beaches. AES president Bob Moses followed by thanking the conference organizers

and presenters for their hard work and the sponsors and exhibitors for their financial support. Moses noted that this was only the second AES international conference ever held in Asia and pledged that the AES will strive to better serve its Asian membership in the future.

Byungki Oh of Korean mobile operator KTF kicked off the technical program with a keynote address exploring the company's suite of award-winning 3-G mobile services known as SHOW. Oh noted that Korea is a world telecommunications leader, having been ranked as the top country in the International Telecommunication Union (ITU) Digital Opportunity Index (DOI) for three consecutive years. As of April 2008, 91.3% of the Korean population was subscribed to mobile service. As one of the leading mobile service providers in Korea, KTF has a customer base of over 14.5 million customers. Its W-CDMA/HSDPA SHOW service aims to deliver high-speed multimedia, live television, a music download portal, video on demand, Internet, global roaming, banking, and other services. Oh's presentation inspired conference attendees with the magnificent promise of future mobile services.



Papers cochairs Eunmi Oh, left, and Young-Cheol Park

Following the keynote address, Kewei Yang of Analogix Semiconductor gave an invited lecture on the fledgling DisplayPort Digital Multimedia Display Interface standard. DisplayPort aims to replace existing interface standards in the consumer and personal computing marketplaces in “box-to-box” and internal (such as notebook PC LCD panel to motherboard) applications. Yang characterized the DisplayPort architecture as an extension of the PCI Express high-speed serial interface standard, with the addition of copy protection and digital audio based on the IEC60958 standard. Yang noted that most of the people and businesses behind DisplayPort are not audio experts and invited audio professionals to join the standardization effort to ensure that audio is handled properly.

The Signal Processing, Part 1 session began with the paper “On Evaluation of Blind Audio Source Separation” presented by Mingu Lee of Seoul National University. Blind audio source separation (BASS) is a process of “unmixing” a compound signal into its component signals. Many BASS algorithms have been published but methods to evaluate them are not consistent, making it difficult to compare them. Lee presented a unified performance measurement procedure for evaluating BASS algorithms based on estimated source decomposition and the MUSHRA test method.

The next paper, “Discrimination of Music Signals for Mobile Broadcasting Receivers,” was presented by Myung-suk Song of Yonsei University. Song proposed a Gaussian mixture model (GMM)-based music-discrimination system for mobile receivers that can automatically archive music from broadcasts that have interference from human voice, acoustic noise, commercial advertisements, and other sources. After describing the GMM method, Song showed experimental results on Korean and English pop music programs with approximately 80% success.

The final paper of this session was “Non-Linear Signal Processing for Low Frequency Enhancement,” presented by Jakob Birkedal Nielsen of AM3D A/S. Nielsen explained how the AM3D algorithm performs bass enhancement on small loudspeakers with limited low-frequency response by applying a novel dynamics-processing algorithm. This algorithm boosts low frequencies when the

original signal has ample headroom and generates harmonics that create the perception of a missing fundamental when available headroom is low.

Following Nielsen’s presentation, attendees enjoyed a coffee break in the exhibit area. The exhibitors were APT, Electronics and Telecommunication Institute (ETRI), LG Electronics, Oxford Digital, and Pulsus. APT provided information about its apt-x coding technologies that enable real time-delivery of professional quality audio over wireless and Bluetooth links. ETRI demonstrated its MUSIC2.0 interactive music service that allows users to control the mix of prerecorded music and to view lyrics, music, videos, slide shows, and other data. LG Electronics showed its ClearVoice technology, which enhances human voices over background sounds, and MP3Remix technology that allows multitrack MP3 files to be mixed by the user (the subject of a presentation later in the conference). Oxford Digital demonstrated its Tiny DSP core, recently integrated into a number of products: a new general-purpose DSP chip from Sony, schematic-based software development tools for this DSP, and Oxford Digital’s Easy-Tune GUI-based software tool for tuning sound quality on digital audio devices. Pulsus provided information about its family of digital-to-digital converter (DDC) ICs, which convert PCM audio sources into PWM signals that can be digitally amplified.

The Coding for Audio and Speech, Part 1 session followed the coffee break. The session began with a presentation by Eunmi Oh of the Samsung Advanced Institute of Technology on “Framework for Unified Speech and Audio Coding.” Oh explained that algorithms optimized for speech are usually poor at coding music, and vice versa. She advocated a new coding scheme known as frequency-varying modulated lapped transform (FV-MLT). It offers flexible time/frequency resolution, perfect reconstruction even when time/frequency resolution is adaptively changed, and no loss of frequency response compared to conventional methods. Oh presented experimental results that showed improved performance at low bit rates of 16 to 24 kbps versus other methods.

The next presentation, “Using Salient Envelope Features for Audio Coding,” was from Joachim Thiemann of McGill University. Thiemann described a work-in-progress in which



Authors who presented papers in the poster session had an opportunity to give in-depth answers to the questions posed by attendees.



During coffee breaks exhibitors APT, Electronics and Telecommunication Institute (ETRI), LG Electronics, Oxford Digital, and Pulsus offered information on their products and services.

he is investigating a new method for perceptual audio coding using perceptually-salient envelope features. These features are found by passing the audio through a series of gammatone filters and computing Hilbert envelopes. Relevant points on these envelopes are then transmitted to a decoder that reconstructs the original audio signal iteratively. Initial experiments show promise with moderate bitrate reduction on speech signals, but further work is required to improve audio quality with tonal signals.

The final presentation in this session was “Personalized Music Service Based on Parametric Object-Oriented Spatial Audio Coding” by Yangwon Jung of LG Electronics. Jung described a method for parametric music control (mix, EQ, etc.) called parametric object-oriented spatial audio coding (POC), based on the spatial audio object coding (SAOC) work by the ISO/MPEG standardization group. POC carries control data in side information with the MP3 data, allowing the user to remix MP3 audio objects. Jung showed a prototype player called MP3Remix that allows real-time mixing of music (level and pan), as well as preset mixes of varying styles (flat, karaoke, a cappella, etc.). Jung suggested that this service enhances the music experience by allowing people to personalize it to their tastes.

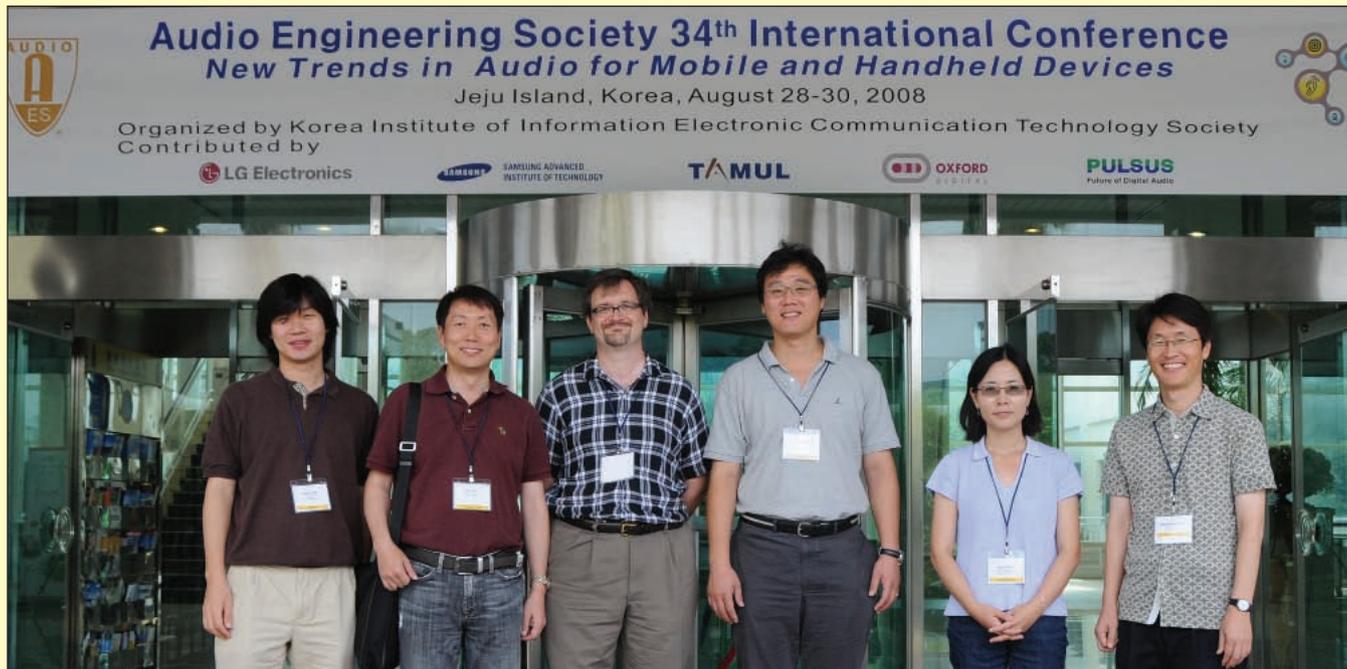
The Coding for Audio and Speech session was followed by a poster session with five papers. One paper, “A Bit Reduction Algorithm for Spectral Band Replication Using the Masking Effect” by Sang Bae Chon, et al., discussed a spectral band replication (SBR) method that reduces bit rate by modifying

envelope data such that the reduction can not be perceived subjectively. Experimental results show a 10 to 12% reduction in envelope data at a bit rate of 24 kbps with no reduction in audio quality. Another paper, “Bit-Rate Reduction Using Efficient Difference Coding of Sinusoid Amplitude” by Namsuk Lee, described a new method for sinusoid amplitude coding. Lee proposes a new difference method of sinusoid amplitude in birth that achieves nearly 16% reduction in bit rate in experimental results. Another poster, “Implementation of 3-D Sound Using Grouped HRTF” by Seo Bo-Kug, et al., provided a new method for head-related transfer functions (HRTF) that groups and averages HRTFs in order to improve sound localization and moving sound effects. The fourth paper, “An Improved Weighting Curve Based on Equal-Loudness Contour” by Inseok Heo and Koeng-Mo Sung described a new loudness weighting curve. The authors point out that some weighting curves are good at low levels and others are good at high levels. Their proposed curve claims to be useful at any signal level. The fifth poster presentation, “Low Carrier Frequency Noise-Shaper for Digital Amplifier” by Park Kyoungsoo and Mo Sung, offered a new noise-shaping method optimized for low oversampling ratios.

Next came the Implementations and 3-D Audio session. Mohammed Chalil of Analog Devices led off with a presentation on “Smooth PCM Clipping of Audio.” Chalil discussed the basic concepts of clipping in digital audio systems and how his team was able to minimize distortion in an AAC codec by looking ahead and detecting clipping and applying



The workshop *Audio in the IT Industry* featured the insights of, from left, moderator Yonhong Jung, John Oh, Juha Backman, Te-won Lee, and Eunmi Oh.



AES 34th Conference Committee: from left, Hyen-O Oh, treasurer; John Oh, conference chair; Bob Moses, AES president; Yonhong Jung, workshops chair; and Eunmi Oh and Young-Cheol Park, papers cochairs.

gain reduction. The Attack Release Sustain Decay characteristics of the original audio signal were preserved in this method, thereby maximizing sound quality.

Hyun Wook Kim of Samsung Electronics followed with a presentation titled “Advanced Terrestrial DMB System Structure for Multichannel Audio Services.” Kim described how two object-descriptor (OD) streams can be employed to carry multichannel side information with elementary streams in an MPEG stream. This enables multichannel digital multimedia broadcasts with backward compatibility with stereo systems.

Young-Cheol Park of Yonsei University closed the session with a presentation on “Robust Crosstalk Cancellation Based on Energy Density Control.” Park described how crosstalk cancellation filters can be used to widen the perceived stereo field and a method of using the energy density function in order to improve stability compared to conventional pressure-control methods

At the conclusion of the first day, attendees enjoyed an excellent Korean buffet followed by an enchanting evening in the nearby gardens and beaches.

DAY 2

Day 2 began with an invited lecture by Juha Backman of Nokia titled “Mobile Phone Audio: The Shape of Things to Come.” Backman predicted the key innovations in the mobile field will be increases in processing power, storage capacity, data transfer rates, the number of connected devices, advances in amplifier, signal processing, and transducer technologies, and acoustic modeling. Backman explained how mobile-device designers face challenges such as energy management, consumer expectations for better/faster/cheaper devices, and long development cycles in an impatient marketplace. He discussed the evolution from circuit switched (such as plain old telephone service, or POTS) to packet switched systems (such as Voice over IP,

or VoIP). He noted that services such as Skype have brought VoIP into the mainstream, offering a radical improvement in network-capacity utilization, better audio quality, no need for echo cancellation, and other advantages. Another interesting innovation is in the area of microelectromechanical systems (MEMS) microphones and multi-microphone array technologies that offer improved voice quality. Transducers and headphones are getting smaller and are more efficiently using available space. Improvements in magnetic materials and digital signal processing are improving fidelity. Backman ended his presentation by noting that 10% of the population is hearing impaired but less than 2% wear an assisted-hearing system. Backman warned that all of today’s mobile users will grow older and their hearing will naturally get worse. Thus, manufactures must address this growing problem in future devices so that customers can continue to enjoy them.

Next came the Evaluation and Testing Session. The first presentation, “A Study of Evaluating the Button Sounds using Wavelets,” was given by Shunsuke Ishimitsu of Hiroshima City University. Ishimitsu described how button sounds emitted by 11 commercial automotive audio units were analyzed both subjectively and by using wavelet transforms to determine which sounds were most pleasing to users. The study found that metallic sounds with high force were less pleasing than nonmetallic, low-force sounds. In addition, low-frequency sounds were considered more favorable than high-frequency sounds. The study also concluded that tactile feedback (touch impression) improved the user experience.

The next session was Coding for Audio and Speech, Part 2. Javier Tapia of PacketVideo Corporation began with a presentation on “Introduction to the OpenCore Audio Components Used in the Android Platform.” Android is a free and open platform for mobile devices. Originally created by Google, Android has evolved into a collaborative effort ➡



Attendees visited Yakchunsa Buddhist Temple and Jeju Folk Village Museum.

of 34 companies. The operating system, middleware, and sample applications will be available as open source. Tapia described the OpenCore framework within Android for combining independent media-processing components such as file formats, codecs, streaming protocols, and rendering. He also discussed implementation issues such as optimization, 3rd party add-ons, etc. Simulation results were given showing performance and CPU utilization for several codecs running on an ARM9E processor.

Choong Sang Cho of Korea Electronic Technology Institute (KETI) followed with the presentation “An Efficient Forward Prediction Order Selection Method for MPEG-4 Audio Lossless Coding.” Cho gave an overview of the MPEG-4 ALS audio lossless coding structure and discussed the relationship between complexity and compression ratio. He proposed a mean squared error (MSE)-based order-selection method to achieve superior compression ratios.

The final presentation in this session was “Segmented Dimensionality Reduction Coding on Frequency Domain Signal” by Minje Kim of ETRI. Kim proposed methods to compress frequency-domain signals using dimensionality-reduction methods.

Following the lunch break, attendees were treated to a tour of three sites on the island. The first excursion was to Jusangjeolli Cliff, one of Jeju Island’s natural wonders with 20-meter-high hexagon-shaped stone pillars formed when lava from Mt. Hallasan erupted into the Sea of Jungmun. Attendees spent a half hour exploring and photographing these pillars. As a bonus, a Haenyo (diving woman) was spotted near the cliffs selling freshly caught shellfish and sea vegetables “fresh from the bucket.”

The next stop on the tour was Yakchunsa Buddhist Temple. With a height of 30 meters, the 3-story Yakchunsa is the largest temple in Asia. Famous for its medicinal waters, attendees were invited to have a drink and make a wish. Sadly, the wishes of many attendees that they could remain in this tropical paradise forever did not come true.

The final stop of the tour was Jeju Folk Village Museum, a 40-acre site with over 100 restored buildings and 8,000 folk artifacts showcasing Jeju life in the 1890s. Among the attrac-

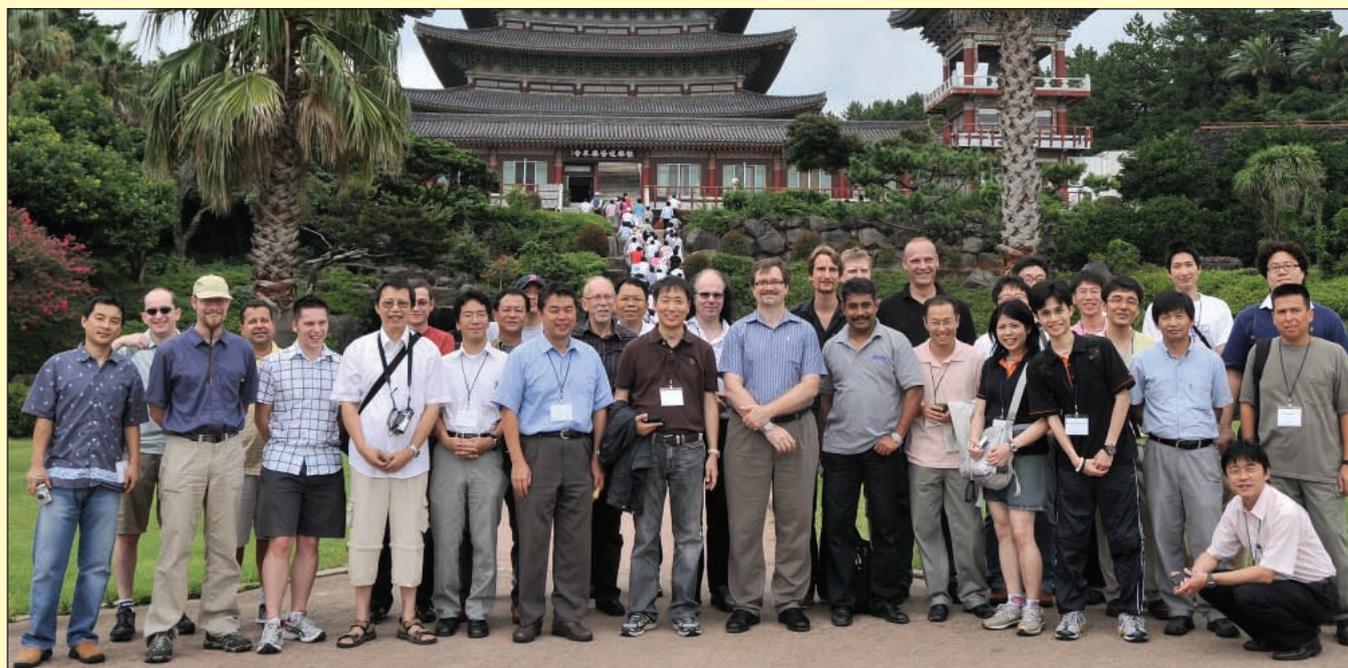
tions are a reconstructed mountain village with farms and a horse-driven mill, and a hill-country village with a wealthy family’s house, blacksmith shop, and other public buildings. A restored fishing village gave a glimpse of a Haenyo’s house and her diving equipment, and a shamanism village displayed a local shaman’s shrine, a fortune teller’s house, witch house, virgin shrine, and shrines for mountain and sea gods.

The conference banquet followed the afternoon field trips. Held at the Jeju Korea House, attendees were treated to a traditional Korean BBQ with grilled seafood, “black pork,” numerous side dishes, and capped off with a few rounds of Korean Shoju and toasts to the people who put the event together. An eclectic multimedia video/musical performance was provided by students from the Korean National University of Arts, led by Professor Jaeho Chang. A stunning sunset provided the backdrop for the outdoor stage, placing an exclamation mark on an exceptional day.

DAY 3

The conference technical program resumed on Day 3 with an invited lecture by Te-won Lee of Qualcomm on “Speech Enhancement for Mobile Applications.” Lee explained the importance of noise reduction in mobile systems and challenges to reducing noise, such as the loud ambient environments in which mobile devices are commonly used (cars, night clubs, etc.) and the relatively long distance between the mouth and microphones worn on the ear. He described single-channel and multichannel methods for speech enhancement such as spectral enhancement, beam forming with multiple microphones, computational auditory scene analysis, and blind-source separation (BSS). Lee explained how Qualcomm has implemented these techniques to provide noise reduction, echo cancellation, wind noise reduction, and speech enhancement. He played several audio demonstrations showing how a cacophony of unintelligible voices can be separated into separate channels, how voice drowned out by loud music can be isolated from the music, and how wind noise can be significantly reduced.

The Signal Processing, Part 2 session followed, beginning with a presentation by Nathan Bentall of Oxford Digital on



Attendees posed for a group photo at the entrance to the Jeju Folk Village Museum.

“Tiny DSP: DSP Core, Algorithm Development and Device Mastering.” Bentall noted that by their very nature mobile devices have very small loudspeakers with poor frequency response, posing difficult acoustic challenges. Despite amazing advances in mobile devices (processing power, battery life, video streaming, etc.), audio quality seems to be getting worse, not better. Thus, audio quality is an opportunity for product differentiation. Bentall explained his company’s concept of “device mastering” in which device fidelity can be improved at the final stage of development by a series of “tweaks,” analogous to mastering in the recording process. He showed several techniques that can be applied during device-mastering such as equalization and compression, and described a software tool called EasyTune that can be used in the process. Then Bentall introduced the features and architecture of a small synthesizable DSP core called TinyDSP. Requiring only 30,000 gates in a typical implementation, the TinyDSP core is being licensed to semiconductor companies and embedded into new mobile devices. A schematic-driven graphical user interface methodology is used to program the DSP.

Sang-keun Oh of LG Electronics wrapped up the session with the presentation “Simple High-Band Extension Method Using Wavelet for Mobile Device.” Oh explained that digital audio must be compressed in order to fit within the data bandwidth constraints imposed by mobile systems. This compression degrades performance in the higher-frequency band. Oh reviewed conventional methods such as spectral folding and spectral translation to copy low-frequency information into the high-frequency band. He then proposed a new method using the digital wavelet transform (DWT).

The next session was 3-D Audio and Synthetic Audio. Till Schafers of Deutsche Telekom Laboratories began with the paper “Designing Low-Dimensional Interaction for Mobile

Navigation in 3-D Audio Spaces.” Schafers presented several uses for mobile spatial audio: teleconferencing, gaming, notification systems, and navigation. He discussed gesture-based interaction, posing the intriguing question, “Why not apply spatial audio techniques to mobile devices and interact in a virtual space?” Schafers summarized five interaction schemes: head tracking using a gyroscope attached to headphones, tracking device, orientation using a compass or gyroscope, using a camera to track device movement, sensing tilt using accelerometers, and a keypad. He also discussed challenges such as training users, making gestures while walking, and sensor fidelity.

The second presentation in this session was given by Sang Ha Park of Seoul National University on “A Consonance-Maximization Tuning Algorithm in Equal-Tempered Synthesized Tones.” Park began with an overview of musical tuning theories and some of the problems with each of them (for example, dissonance). She also gave a background on consonance theory and proposed a tuning system that maximizes consonance. Park proposed this new tuning for commercial applications such as ringtone synthesis.

The final event of the conference was the roundtable workshop “Audio in the IT Industry,” moderated by Yonhong Jung of Tamul Multimedia. The four panelists were John Oh, Juhah Backman, Eumni Oh, and Te-won Lee. Backman began the discussion by pointing out that engineers may find the technology behind mobile devices fascinating but most consumers do not care about technology as much as they care about compelling services such as music delivery. Backman noted that many other offerings, no matter how technologically superior they were to their predecessors, have failed due to lack of music content. Thus, he suggested, winning mobile products and services must be associated with an excellent music-distribution channel.



Eunmi Oh followed with a discussion about trends in speech and audio coding technologies and a review of standards from MPEG, ITU-T, and 3GPP.

Te-won Lee made a pitch for new modalities and sensors in addition to microphones to enhance the user experience.

John Oh gave a short presentation on the customer's view of handheld audio. He noted that many interesting new technologies were presented in the conference, but questioned whether they will actually be used. Does the general public know that excellent mobile audio exists today? Do they care? And, can they figure out how to use it? Oh noted that he personally has a video phone but has never used the video features. We are still in the early market for mobile multimedia and have a long way to go before it is mature. He posed the questions: What makes good audio relevant? and What will be the killer app? He also proposed a test for determining the value of any new technology: would it be a good holiday gift for your grandmother?

Backman followed with an overview of digital rights management (DRM) technologies and noted how they are often cumbersome and counter-productive. Bob Moses responded with a brief explanation of a DRM technology that has been implemented for online music sharing called Weedshare, which operates in a manner similar to a pyramid scheme in which music enthusiasts who distribute music files are rewarded with a small portion of the sale price when the music they share is purchased by others. If someone pirates a song, they circumvent the pyramid and forfeit the opportunity to profit from further sales. Thus, there is a strong incentive to pay for the music as an investment toward higher gains. Moses proposed that as mobile services promote new applications allowing users to share photos, videos, ring tones, and music, Weedshare might provide a compelling method for compensating rights holders.

Te-won Lee talked about the importance of timing the introduction of new technologies to when people are ready to purchase and use them. John Oh said we need to create products that people need and want, not just what engineers enjoy inventing.

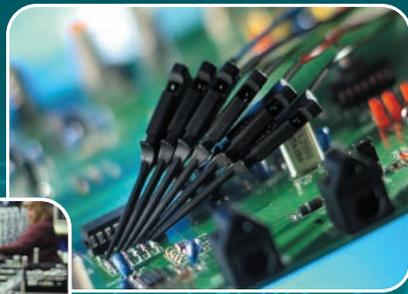
The group ended the conference by posing some thought-provoking questions: With mobile technologies rapidly evolving, which of these technologies offer consumers value that they will be willing to pay for,

and which of them are functions dreamed up by engineers that no one else will understand or want? What is the difference between a mobile device TOY and a mobile device TOOL? Is music delivery the killer app? How can rights be managed in the mobile environment without the rampant piracy experienced online? Given the fierce competition and innovation in the mobile industry, it will be fascinating to watch where this compelling technology takes us next. We can hope that it will bring us back to beautiful Jeju Island in a couple years for another compelling AES conference.

Editor's note: The CD-ROM of conference papers can be purchased at <www.aes.org/publications/conf.cfm>.

dScope Series III

audio analyzer



Product design



to production line

The dScope Series III audio analyzer includes built-in automation tools for production-line testing.

- VBScript IDE for rapid test development
- ActiveX control for 3rd party automation
- Multi-tone tools for increased throughput
- Switching options for multi-channel testing
- Comprehensive results for rapid diagnostics




Contact us now to arrange your demo

Email: sales@prismsound.com

+1-973-983-9577
 +44 (0)1223 424988

www.prismsound.com

