

AES 54th International Conference **Audio Forensics**

**12–14 June 2014
London, UK**

CONFERENCE REPORT



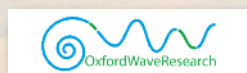
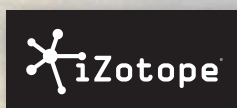
In June 2014 the city of London welcomed the audio community for the latest international audio forensics conference. Participants from all over the world gathered to share information on research and practice in forensic science. The 54th AES International Conference, *Audio Forensics—Techniques, Technologies, and Practice*, was the most recent AES event focusing on the field of audio forensic analysis and interpretation. The sequence of AES audio forensic conferences began in 2005 with the 26th AES Conference held in Denver. The 33rd Conference returned to Denver in 2008, followed in 2010 by the 39th Conference in Hillerød, Denmark, and back to Denver for the 46th Conference in 2012.

As was true for the prior AES audio forensics meetings, the 54th Conference brought together an outstanding combination of practitioners, researchers, law enforcement professionals, attorneys, and many other experts and students all sharing an interest in the latest developments and contributions to audio forensic science made by AES members. The 62 registered participants included representatives from 15 countries.

Platinum Sponsor



Sponsors



The two years of planning prior to the 54th Conference involved members of the AES Technical Committee on Audio Forensics and the AES headquarters staff. Mark Huckvale and Jeff M. Smith, were conference cochair, while Mike Brookes and Durand R. Begault served as papers cochair. Workshop sessions were organized by Catalin Grigoras, and Patrick Naylor organized the exhibition of audio forensics products. The committee's work was supplemented by student volunteers from University College London who assisted with registration and presentation logistics.

The conference was hosted at the Holiday Inn – Bloomsbury, located in central London near Russell Square. The Bloomsbury district of London is noted for its connections to education, the arts, literature, and medicine. Former residents of the area include Charles Dickens, T.S. Eliot, Virginia Woolf, John Maynard Keynes, and Dorothy Sayers, among many others. This vibrant part of the city is also home to University College London, the British Museum, the British Library, and the National Hospital for Neurology and Neurosurgery. The hotel meeting area and accommodations were ideally suited to the conference, and enabled the delegates to enjoy one-on-one interaction and small-group discussions that are among the key features of all AES international conferences. The meeting schedule included group lunches, coffee breaks, and other unstructured social events that were filled with lively conversations.

CONFERENCE OPENING



Mark Huckvale welcomes everyone to the event.



Jeff Smith thanks the conference sponsors.

The conference opened on Thursday morning, 12 June, with clear skies and very comfortable temperatures in London. Conference cochair Mark Huckvale of University College London welcomed the delegates and provided an overview of the conference. Huckvale also shared some historical insights about the Bloomsbury district of London, and invited the delegates to enjoy the sights, sounds, and flavors of the community. Cochair Jeff Smith, associate director of the National Center for Media Forensics of the University of Colorado Denver, added his words of welcome to the participants, and thanked the sponsors and exhibitors, led by platinum sponsor CEDAR Audio Ltd. Other exhibitors and sponsors included AGNITiO, iZotope, Oxford Wave Research, Salient Sciences, and Voxalys. The exhibitors were present throughout the conference for demonstrations and discussion. The cochair also thanked the local organizing committee and the volunteers for arranging the conference venue and special events.

KEYNOTE LECTURE

Itiel Dror, a noted cognitive neuroscience researcher from University College London, presented an engaging, provocative, and entertaining keynote address for the 54th Conference entitled "Cognitive Bias in the Interpretation of Forensic Evidence." Dror performs research in the performance of experts in real interpretive situations, and his work reinforces the fact that even experts can show unintended bias when presented with evidence from an

investigation. While much of Dror's work has involved tactical decision-making in real-time by doctors, soldiers, and others who must make split-second decisions, he pointed out that forensic science is generally carried out without immediate time pressure, so bias in forensic investigations can be addressed in a more comprehensive and systematic manner.

Dror explained that the human brain is built to interpret a highly filtered and inherently biased version of the world around us, and that "expertise" can actually represent a very specialized type of bias: the ability to focus on specific details and meanings while ignoring many other aspects of a particular circumstance or scene.

In forensic analysis, one of the ways in which cognitive bias often occurs is by a detective or another investigator sharing side-information about the case that is inevitably prejudicial to the expert's interpretation. Dror emphasized that audio forensic experts should never see themselves as "part of the team," trying to "make the case" for one side or the other. His advice is for experts to acknowledge that cognitive bias is always present, and rather than treating this as a fault or weakness, consider ways to help the court or jury understand the potential uncertainty that is present even in an expert's testimony.

EXHIBITOR INTRODUCTIONS

Among the features of many AES conferences are exhibits and hands-on demonstrations provided by companies and product developers. The 54th Conference was no exception, and the end of the Thursday morning technical session included some time for the exhibitors to introduce themselves and the products being shown at the event.

CEDAR Audio Ltd. of Cambridge, UK, platinum sponsor of the conference, was represented by Gordon Reid, who explained the



Itiel Dror engages delegates to the conference during his keynote.



Jeff Smith, left, thanks Gordon Reid of CEDAR, platinum sponsor.



CONFERENCE EXHIBITION IN FULL SWING

history of CEDAR and its noise-reduction and quality-enhancement products, and described a wide variety of applications in the entertainment and audio forensics fields.

Oxford Wave Research, of Oxford, U.K., was introduced by Anil Alexander. Alexander explained his company's work in automatic speaker diarization (processing a recorded dialog into separate segments containing the utterances of each participant in the recorded conversation), time alignment and removal of known audio, and automatic talker recognition.

Jonas Lindh introduced Voxalys AB, of Göteborg, Sweden. Voxalys is a forensic speech and acoustic lab that offers consultation and services regarding forensic phonetics, audio, training, and software.

Salient Sciences, Durham, North Carolina, USA, was represented by Don Tunstall. Salient Sciences was formed recently by a merger of Digital Audio Corporation and Salient Stills. The company provides audio/video analysis and enhancement services for law enforcement and forensic analysis, including hardware, VST plugins, and other software components. The company also provides training, as well as custom engineering and system development.

Rounding out the outstanding group of exhibitors was Antonio Moreno, technical sales director of AGNITiO, a company

from Madrid, Spain, specializing in voice-identification products and technology.

The exhibition area was located in the lobby immediately adjacent to the main meeting room for the conference, providing numerous opportunities for the delegates to examine the products and services offered by the sponsoring companies.

TECHNICAL PROGRAM—DAY 1

Following a break for lunch, the first afternoon of the conference included a tutorial session and presentation of the first group of technical papers.

TUTORIAL 1: SPEAKER COMPARISON

Gordon Reid of CEDAR served as the chair for a tutorial session covering current aspects of forensic automatic speaker recognition from degraded speech material, and also from recorded speech that had been passed through a stage of noise reduction or quality enhancement processing.

Reid introduced Hermann Künzle, University of Marburg, Germany, who since 2009 has studied the question of whether or not audio enhancement processing can be simultaneously suitable for audition by human listeners and for automatic speaker recognition by software. Künzle also introduced Paul Alexander of CEDAR, who is involved with testing and evaluating forensic audio algorithms.

Antonio Moreno of AGNITiO described the basic principles of automatic speaker recognition using the classical method of phonetic analysis. The approach used in AGNITiO's Batvox software is to use a closed comparison of a particular example speech recording to a set of potential matches. The analysis uses mel-frequency cepstral coefficients (MFCCs) and a Gaussian Mixture Model (GMM) as the matching method.

One important issue is that a given speech recording will include not only the spectro-phonetic characteristics of the talker, but also be convolved acoustically with the characteristics of the recording system and communications channel, plus additive noise and distortion. This means that it is necessary to attempt to separate the speaker variability from the channel variability so that the comparison is of the talker's characteristics and not the channels' attributes.

Künzle then described an experiment conducted to evaluate the speaker recognition performance for undisturbed (clean) speech and speech degraded with various types of noise and reverberation added. The experiment further involved three different types of signal enhancement for the degraded speech samples so that recognition performance could be compared following the enhancement processes. The results showed that the enhancement improved the recognition performance with some types of noise, but also caused worse performance with



Antonio Moreno explains automatic speaker recognition.



Hermann Künzle looks into signal enhancement for improving speech recognition.



From left, Doug Lacey, Jeff Smith, and Rob Maher

other types of noise and enhancement settings. A paper describing the experiment was published recently in the *AES Journal* (Künzel, Hermann, and Alexander, Paul, "Forensic Automatic Speaker Recognition with Degraded and Enhanced Speech," vol. 62, no. 4, April 2014, pp. 244-253).

PAPER SESSION 1: FORENSIC SPEAKER COMPARISON

Following the afternoon break for tea, coffee, discussion, and conversation in the exhibition area, the Thursday afternoon technical paper session was introduced by Durand Begault, papers cochair.

The first paper, "A Study of F_0 as a Function of Vocal Effort," by Eddy B. Brixen of EBB-consult, Denmark, described his investigation of the fundamental frequency (F_0) of speech under variations in vocal effort such as casual conversation, loud talking, and yelling. The empirical work showed a general trend toward increasing F_0 with increasing vocal effort, which may lead to future investigation of how an examiner might compensate for such changes when comparing utterances recorded under circumstances of differing vocal exertion.

Michael Jessen of the Bundeskriminalamt (Federal Criminal Police Office), Germany, presented a paper entitled "Forensic Voice Comparisons in German with Phonetic and Automatic Features Using VOCALISE Software." Jessen described his work in collaboration with Anil Alexander and Oscar Forth of Oxford Wave Research Ltd. to use the commercial forensic acoustical analysis package

VOCALISE (Voice Comparison and Analysis of the Likelihood of Speech Evidence) for their experiments. Jessen described how the German Federal Police use the software for forensic voice comparisons. The software is found to be particularly convenient for extracting spectral and phonetic parameters from speech examples, although the issues of how to determine the various analysis settings and interpret the results in the presence of channel effects remain as challenges for the forensic examiner.

The concluding paper of the first session was "Evaluation Results of Speaker Verification for VoIP Transmission with Packet Loss," presented by Jörg Bitzer of Fraunhofer IDMT. Bitzer examined the degree to which speaker verification could be affected by the common packet loss and error concealment strategies employed in unreliable communications systems such as Voice over Internet Protocol (VoIP). Under the conditions used in the experiment, Bitzer found that the influence of packet loss on automatic speaker verification can be neglected. Challenges still remain, however, for other types of degradation, such as background noise, echoes, and reverberation.

Upon the conclusion of the successful first day of the conference the attendees were invited to reconvene for an informal reception at the Grant Museum of Zoology and Comparative Anatomy, located on the campus of University College London, just a short walk from the conference site. The Grant Museum dates back to its founding in 1827 by Robert Edmond Grant (1793–1874), who set about creating a vast collection of animal specimens from around the world for use in his university courses. While the modern fields of zoology and biology increasingly rely upon DNA and other molecular analysis techniques for categorizing species, the Grant zoological collection provides a fascinating look at micro and macro fauna gathered over nearly 200 years. The 54th Conference delegates enjoyed champagne, hors d'oeuvres, and conversation while strolling among exhibits including the bones of the extinct Dodo and Quagga, thousands of microscopic specimen slides, and the entire intact skeleton of a five-meter-long Anaconda snake. The participants relished the opportunity to relax and unwind in the Museum's unique surroundings while discussing the day's topics and presentations.

TECHNICAL PROGRAM—DAY 2

The second day of the conference dawned in London with comfortable temperatures and clear skies. Delegates from Brazil were happy about the news overnight of their team winning the opening game of the World Cup in Sao Paulo, while those from Mexico, Spain, and the Netherlands were anticipating word of their own team's results during the Friday competitions.

TUTORIAL 2: DEREVERBERATION

Catalin Grigoras introduced the opening event of the day, a tutorial session on experimental methods for dereverberation. Patrick Naylor of Imperial College London provided a clear and complete introduction to the topic, including the acoustical origins of reverberation and



A reception at the Grant Museum enables the assembled company to discuss audio forensics in the company of stuffed animals.



Patrick Naylor gives his tutorial on the various methods for removing reverberation. Patrick also organized the exhibition.

the various approaches that have been developed over the years to try to “de-convolve” the original sound from a reverberant recording. Naylor and his students, Alastair Moore and Christine Evers, described several experimental techniques for enhancement of reverberant speech and for blind system identification. Dereverberation is considered an unsolved problem, and many of the difficult theoretical and practical considerations remain interesting topics for ongoing exploration.



Christine Evers describes techniques for the enhancement of reverberant speech and blind system identification.

PAPER SESSION 2: ACOUSTICS

The chair of the morning papers session, Mike Brookes, introduced Ephraim Gower of the Botswana International University of Science and Technology, Palapye, Botswana. Gower spoke about “Exploiting Sparsity for Source Separation Using the Sliding Ratio Signal Algorithm,” a time-domain approach to solving convolutional mixing. The technique approaches the signal separation problem by exploiting the low likelihood that the energy from multiple concurrent signals is present at the same time and same frequency. To the extent that the signals are disjoint in this fashion, the concept is to create an approximation to the separated signals in each time segment of a sliding discrete Fourier transform (SDFT).

The next paper, presented by Hamza Javed of Imperial College London, was entitled “Evaluation of an Extended Reverberation

Decay Tail Metric as a Measure of Perceived Reverberation.” Javed explained the need for a quantitative way to estimate the reverberation level in recorded speech and the potential usefulness of the reverb “decay tail” in comparing the improvement produced by dereverberation algorithms.

The morning session concluded with a paper by Alastair Moore, also of Imperial College London, entitled “Room Identification Using Roomprints.” The goal of the work was to have a way to identify characteristics of the room in which a recording was



Alastair Moore speaks on “roomprints.”

made based on the recording itself. Moore’s experiment involved 22 room impulse responses that were filtered with five different filter-bank resolution settings ($\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, and $\frac{1}{12}$ octave) to determine if they could be discriminated reliably. The results indicated that $\frac{1}{4}$ -octave resolution with lowest frequency between 100 Hz and 300 Hz is sufficient to achieve at least 95.7% identification accuracy on the test database.

TUTORIAL 3: FORENSIC AUDIO AUTHENTICATION WORKSHOP

Eddy Brixen introduced Catalin Grigoras, 54th Conference workshops chair and director of the National Center for Media Forensics at the University of Colorado Denver. Grigoras described many key principles for handling digital evidence, especially in the context of authentication. The audio forensic examiner must take great care to avoid contamination of the digital material through careless or inadvertent handling. Grigoras recommended that the examiner carefully document the chain of custody and the processing steps involved so that there would be no question later about the integrity of the audio material or the conclusions drawn from it.



Catalin Grigoras explains the principles of handling digital evidence.

PAPER SESSION 3: AUDIO AUTHENTICATION

Following a break for tea and conversation, the afternoon technical program resumed with a paper session chaired by Daniel Rappaport of CACI Digital Forensics Lab, Alexandria, Virginia, USA. Rappaport introduced Bruce Koenig of BEK TEK LLC, Clifton, Virginia, who presented the paper “Forensic Authenticity Analyses of the Metadata in Re-Encoded WAV files.” Koenig explained the basic format of a Resource Interchange File Format (RIFF) file containing audio material, typically referred to as a WAV file. The WAV file uses embedded tags and logical information containers, commonly called “chunks,” to hold not only the audio data itself, but also descriptive information (metadata) about the data format, audio parameters, and possibly proprietary manufacturer-specific



Bruce Koenig deals with metadata in RIFF files.



A crowded conference hall with delegates listening attentively during one of the many fascinating papers sessions.



Delegates gather in front of a statue of Gandhi during a walking tour of Bloomsbury.

information inserted by the particular recording device or software package used to create or edit the WAV file. Koenig explained that when an original WAV file is opened by a file editing program and simply saved again without deliberate alteration, it is still possible that the editing program may insert new or altered metadata into the newly saved WAV file. Based on an experiment with nine different audio recording devices, Koenig suggested that audio forensic examiners become familiar with the particular WAV file chunks produced by various devices and file editors as one way to identify possible tampering or copying, since simply opening a file and saving it unaltered can result in changes to the chunk contents in some cases.

Next, Luca Cuccovillo of Fraunhofer IDMT, Germany, presented his group's work entitled "A Multi-Codec Audio Dataset for Codec Analysis and Tampering Detection." The work has involved creating a dataset of "tampered" and untampered

audio excerpts, where the tampering consists of a change in the encoding parameters, such as encoder type, encoding bit rate, and the framing alignment of the encoder. The alterations are made in such a manner that the changes should not be audible, but yet may still be detected by identifying underlying changes in the encoding that result in various encoding-dependent artifacts. The alterations of interest are situations in which an original encoded stream is decoded, edited, and then re-



Luca Cuccovillo speaks on codec analysis.

encoded by the same or a different codec. Full annotation is provided with the dataset so that the type and extent of the tampering is known. The dataset is available for use by anyone under the Creative Commons license, and the developers are hopeful that it will be a useful resource for developing tampering-detection strategies and for testing software analysis systems.

Anibal Ferreira of the University of Porto, Portugal, spoke about "Real-Time Monitoring of ENF and THD Quality Parameters of the Electrical Grid in Portugal." The experiment

was to measure the electrical grid power quality at several separate locations in Portugal's power grid. As expected, the Electrical Network Frequency (ENF) was consistent across the entire grid, showing drift and macro changes due to load and generation management. The quality of the sinusoidal electrical power waveform varied from place to place on the grid, as measured by Total Harmonic Distortion (THD). While the ENF is governed by exactly balancing the electrical generation capacity with the instantaneous electrical load on the grid, the THD depends upon distortion introduced by nonlinearities in the power system, such as solid-state electrical switching power supplies, triac power controls, and industrial switching devices.

Concluding Friday's technical papers was "Quantization Level Analysis for Forensic Media Authentication," presented by Catalin Grigoras. Grigoras proposed that one way to detect possible audio file modification in the case of an original low-resolution file (e.g., 8-bit quantization) that has been converted for editing at a higher resolution (e.g., 16-bit), is to look for tell-tale indications that there are discrete quantization levels remaining in the data stream. Similarly, PCM files that have had a change in quantization level followed by amplitude fading or interpolation may show tell-tale linear interpolation effects between the original discrete levels. The quantization level analysis is appropriate when the file is normalized when converted to higher resolution, meaning that the least-significant bits are zero. If the file is converted while maintaining absolute value so that the most-significant bits are filled with the sign bits, the detection strategy will not be effective.

FRIDAY SOCIAL EVENT: BLOOMSBURY WALKING TOUR AND WINE BAR VISIT

With the delegates ready to unwind after two fine days of audio forensic content, the conference organizing committee thoughtfully provided the opportunity for an informal social outing on Friday evening. The social time began with two expert guides leading groups on a fascinating walking tour of the Bloomsbury area, including Russell, Tavistock, Red Lion, Bloomsbury, and Garden Squares, plus Coram's Fields and the British Museum. The weather cooperated beautifully, and the tour guides' insights and observations made London's 18th-century streetscapes and modern plazas meld into a charming display of what makes London a world-class city. All too soon the tour came to an end, but, thankfully, the culminating destination was Truckles Wine Bar, located just down the block from the British Museum. The conference delegates were treated to a fine selec-



Anibal Ferreira looks into the electrical network frequency.

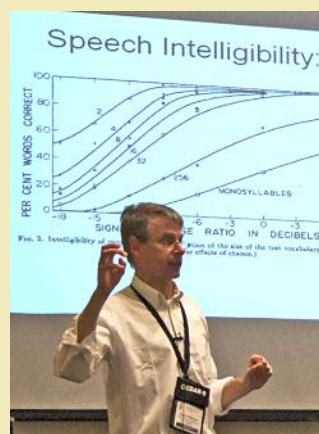
tion of hors d'oeuvres, ales, wines, and other beverages derived from fruits and grains, while mingling in the spacious courtyard with Londoners preparing for the World Cup matches to be televised later that evening.

TECHNICAL PROGRAM—DAY 3

The final day of the 54th Conference featured a tutorial and three technical paper sessions. Prior to the session, the hallway discussions among some of the delegates included the Spain vs. Netherlands World Cup match, in which the Dutch prevailed 5–1.

TUTORIAL 4: SPEECH INTELLIGIBILITY

Speech intelligibility is an important aspect of audio forensics. Many audio forensic investigations involve understanding and



Gaston Hilkuysen tackles the topic of speech intelligibility.

transcribing speech utterances, and understanding the ways in which audio processing algorithms can positively or negatively affect intelligibility is a key area of interest. The tutorial, presented by Gaston Hilkuysen of the University of Marseille, France, covered the experimental results of human testing with speech presented as isolated words and as meaningful sentences. Speech-enhancement algorithms tend to work better at high signal-to-noise ratios, but with high SNR the intelligibility is often very good to begin with. In some studies it appears that

simple spectral shaping, such as light high-pass filtering, can give as much of an intelligibility improvement as more complicated methods. Hilkuysen explained that this may be due to the high-pass filtering simply limiting the upward spread of masking, since from a psychoacoustical standpoint, noise and interfering sounds tend to mask high-frequency content to a greater degree than low-frequency content.

Hilkuysen emphasized that most speech intelligibility and speech quality tests show that performance is better with steady, stationary noise like the background hum within a moving automobile or train, and performance is worse for fluctuating or time-varying noise such as speech babble or interfering music. Considerable effort continues in the intelligibility and enhancement fields, and both theoretical and practical breakthroughs are needed.

PAPER SESSION 4: SPEECH ENHANCEMENT

Richard Stanton of Imperial College, London, presented an interesting paper entitled “A Differentiable Approximation to Speech Intelligibility Index with Applications to Listening Enhancement.” The Speech Intelligibility Index (SII) is a standardized objective measurement that is intended to be a reliable estimate of intelligibility using a weighted-sum of signal-to-noise calculations in a set of subbands. The SII requires separate access to both the “clean” speech and to the “noise,” and then produces an estimate of the expected intelligibility for the

speech and noise mixture. Stanton explained that the SII standard results in a discontinuous function across frequency, which is therefore not mathematically differentiable. Determining a differentiable version of the SII would provide a way to do mathematical minimization and optimization, so his work has been to develop an approximation to the SII that is smooth and continuous and analytically differentiable. The results have been very good, and have been shown to allow an optimized filter that improved the SII of noisy speech segments while simultaneously constraining signal power.

PAPER SESSION 5: FORENSIC MUSICOLOGY

Forensic musicology deals with disputes about similarities in musical works that could be considered copyright infringement, authentication of musical recordings, determining the origin of melodic material, and similar questions that can benefit from systematic audio forensic investigation. Durand Begault of the Audio Forensic Center of Charles M. Salter Associates, San Francisco, provided a description of recent developments in the field in his paper “Forensic Musicology – An Overview.” Begault explained that the field of forensic musicology is traditionally highly subjective, and forensic experts in the field are, more often than not, asked to provide subjective analyses based primarily on their own impressions or “golden ear” assessments rather than a clear and objective scientific approach. Begault warned against a reliance on subjective impressions, since human observers are innately able to locate patterns and similarities even in unstructured and random shapes such as clouds and ink blots. He also recommended that forensic musicological examiners eschew reliance on elaborate sequences of sophisticated steps and transformations that tend to exaggerate the potential similarity of musical material. In conclusion, Begault recommended that forensic examiners avoid the highly subjective and pseudo-scientific approaches, and continue to develop and publish new techniques in forensic musicology that will benefit the forensic community and, ultimately, the legal system.

PAPER SESSION 6: GUNSHOT ACOUSTICS

Following the final lunch break and active discussions in the exhibition area, Papers cochair Durand Begault rounded up the delegates to return to the main conference room for the last technical paper session of the conference, joking that it was “time to conclude the meeting with a bang: two papers on forensic gunshot acoustics.”



From left, Mark Huckvale, Jeff Smith, Mike Brookes, Durand Begault, and Catalin Grigoras of the conference committee



Jos Vermeulen, left, and Antonio Moreno get down to brass tacks over a thorny point in audio forensics.

The first paper was “Gunshot Recordings from Digital Voice Recorders,” presented by Rob Maher of Montana State University, Bozeman, Montana. Maher explained that the audio gunshot recordings commonly presented for forensic examination are



Rob Maher explains some issues about gunshot recordings.

increasingly made using digital voice recorders whose microphones, amplifiers, and speech coding algorithms are not designed to capture the high intensity and brief impulses of gunfire. Maher described the characteristics of typical muzzle blast sounds from firearms, emphasizing that the duration of the muzzle blast itself is only a few milliseconds, while most forensic gunshot recordings include significant amounts of reflected and reverberant sonic energy that lingers for hundreds

or thousands of milliseconds depending upon the acoustical surroundings. Amplitude clipping and temporal distortion from the coding algorithm and the recorder’s automatic gain control are often present in gunshot recordings, making it difficult to be confident about waveform integrity. He presented several experimental results obtained under controlled conditions, and also several examples of gunshot audio from actual forensic investigations. Maher recommended that audio forensic examiners become aware of the strengths and weaknesses of gunshot recordings obtained from digital voice recorders, and utilize the evidence while keeping in mind the inherent limitations of the recording devices.

Douglas Lacey of BEK TEK LLC presented the final technical paper of the conference, entitled “The Effect of Sample Length on Cross-Correlation Comparisons of Recorded Gunshot Sounds.” Lacey reported on his team’s investigation of techniques to determine a quantitative measure of similarity between two recorded gunshots. A crosscorrelation technique has previously been explored for this purpose, and Lacey compared correlations calculated with time segment lengths varying between 2 and 50 milliseconds to an expert’s subjective

impression of similarity. The results showed that the objective crosscorrelation results were in general agreement with the subjective comparison, and that observing the crosscorrelation for a range of segment lengths gave useful insights into the details and acoustical characteristics of the gunshot samples.

CONFERENCE BEST PAPER AWARD

The organizing committee arranged a Best Paper Award competition with a cash prize made possible by conference Platinum Sponsor, CEDAR Audio Ltd., and selected by CEDAR’s Engineering Director Christopher Hicks. The paper selected for the award was “A Differentiable Approximation to Speech Intelligibility Index with Applications to Listening Enhancement,” authored by Richard Stanton, Nikolay Gaubitch, Patrick Naylor, and Mike Brookes. Richard Stanton received a hearty ovation and congratulations as he accepted the award on behalf of the authors.

AES AUDIO FORENSICS: CHARTING THE FUTURE

The AES 54th Conference continued the tradition established at the four prior AES forensics conferences by treating the participants to a superior slate of audio forensics papers and workshops. AES has proven its place as the leading professional group in the field of forensic audio analysis and interpretation. Conference cochairs Jeff Smith and Mark Huckvale, together with AES deputy director Roger Furness, concluded the conference by thanking the volunteer conference committee and commending all of the conference delegates for their effort in presenting their work for the AES audience.

It continues to be an outstanding time to be involved in the growing field of audio forensics. The delegates at the 54th Conference left the wonderful city of London with many new and exciting concepts, and all look forward to the next opportunity to attend a future AES conference on Forensic Audio.



The Best Paper Award is presented to Richard Stanton, right, by Chris Hicks

Editor’s note: a USB drive or downloadable PDF of the conference papers can be purchased online at www.aes.org/publications/conf.cfm