AES 33rd INTERNATIONAL CONFERENCE

Audio Forensics—Theory and Practice

Denver, Colorado • June 5–7, 2008

The AES 33rd International Conference, Audio Forensics—Theory and Practice, will be held June 5–7 in Denver, Colorado, at the Sheraton Denver Hotel (formerly the Adam's Mark). This is a follow-up conference to the very successful AES 26th Conference on forensics which was also held in Denver in 2005. The 33rd will explore the history, hardware, and techniques of forensic investigation of audio materials. The field has gone through significant advances with the advent of digital audio recording, signal processing, and computer-assisted evaluation. The appropriate use of analog and digital processes provides the contemporary audio engineer with powerful tools for quality audio investigation in support of the law enforcement, legal, archival, and restoration communities.

TUTORIAL DAY

The conference will get underway on Thursday afternoon with a number of tutorials by forensic audio experts. This will introduce the basic concepts to those attendees who are new to the field and prime them for the papers sessions on the next two days. Tom Owen will lead off with a presentation on the history of audio forensics. Gordon Reid will cover the basics of audio enhancement, explaining the need for noise reduction to increase speech intelligibility for transcripts and critical listening. Next the trio of Fausto Poza, Eddy Brixen, and Durand Begault will tackle the subject of voice identification, which is used to determine scientifically whether a voice present in an audio recording matches the voice of a suspected individual. Finally, Greg Stutchman will give an overview of photogrammetry, which extends forensic investigations beyond audio into optics. There will be a meet-and-greet reception following the tutorials.

TECHNICAL SESSIONS

Durand Begault and Eddy Brixen, papers cochairs, have organized the papers into four sessions on Friday and Saturday. Session one will have three papers—by Brixen, Cooper, and Sanders—on the use of ENF (Electric Network Frequency) to facilitate forensic investigations. Because of the unique patterns inherent in power-supply frequency over time, it is possible to use these patterns to determine, for example, whether recordings have been edited, mixed, or whether a recoding was made at the time claimed. Sometimes skilled examiners can pinpoint the exact time of a recording.

The other sessions will cover such topics as audio recovery and enhancement, authenticating analog and digital media, and voice recognition and identification. Charles Van Winkle will present the paper "Audio Analysis and Spectral Restoration Workflows Using Adobe Audition." Rob Maher will present research on deciphering gunshot recordings, work that might help crime-scene investigators to nail down a shooter's location. Begault will present work on the analysis of female screams. Pappas will present the paper "Magneto-Resistive Field Mapping of Analog Audio Tapes for Forensics Imaging," a technique that could allow examiners to link a certain recording device to a given recording. Kent Gibson will present a paper that looks at covert recording techniques that can be used to gather incriminating evidence when suspects are placed in a room together after arrest. They usually can't resist trying to "get their stories straight" before the police question them. In some states, such as California, they have no legal expectation of privacy, so a recording of their conversation could be used as evidence against them. The abstracts of all the papers are listed in the preliminary program on the following pages.

FORENSIC ENHANCEMENT CONTEST

In the months leading up to the conference, forensic specialists from around the world requested a CD with five specially prepared samples of typical noise problems (such as telephone hum, near/far party problems, waterfall noises, restaurant recordings, moving multiple tones, etc.). They processed the samples and returned their enhanced recording, which will be judged by a committee of experts. The winners will receive $500 prizes at the conference, and they will explain the techniques they used.

SOCIAL EVENTS, ROCKY MOUNTAIN TOUR

Social events are still being planned; they will include a banquet on Friday night. There will also be an optional Rocky Mountain tour on Sunday. See you in Denver June 5–7. Go to <www.aes.org/events/33> to register online.

Photo of Denver airport courtesy Denver Metro Convention & Visitors Bureau

Technical Sessions*

*This preliminary program is accurate as of press time.

Thursday, June 5

TUTORIAL DAY

1:00 pm

History of Audio Enhancement and Authenticity: From Edison to the Present—*Tom Owen*

2:30 pm

Basics of Audio Enhancement—*Gordon Reid*

4:00 pm

Voice ID—*Fausto Poza, Eddy Brixen, Durand Begault*

Eddy Brixen and Durand Begault will conduct an informative interview with Mr. Poza on the important technical and historical issues of voice identification, followed by a question-answer period with the audience.

Fausto "Tito" Poza (Poza Consulting Systems) has worked as an expert in the field of forensic audio for over 30 years. He is a nationally-recognized expert in the area of voice identification and elimination. He served as the technical consultant to the National Academy of Sciences committee that reported on voice ID in 1979, which included such luminaries as Richard Bolt, David Green, Franklin Cooper, and Jay McKnight . He has also co-authored two encyclopedic articles that provide attorneys with valuable information and insights in the areas of Voice Identification and Tape Authentication in particular and more generally in the broader areas of Forensic Communications and Forensic Acoustics. Tito received his BS and MA from MIT in 1959, and has been a forensic consultant since 1971; he was formerly a Research Engineer at Hughes Aircraft and a Senior Research Engineer in the Speech Research Group at Stanford Research Institute.

5:30 pm

Photogrammetry—*Greg Stutchman*

Friday, June 6 **9:30 am**

OPENING REMARKS BY RICHARD SANDERS

Friday, June 6

PAPER SESSION 1

1-1 **ENF: Quantification of the Magnetic Field**— *Eddy B. Brixen, EBB-consult,* Smorum, Denmark

Authentication of digital recordings by tracing the electric network frequency (the ENF criterion) has shown promising results. This methodology is possible due to the presence of electromagnetic fields radiated from all kinds of electric equipment and connected power lines. In the ENF range—around the 50/60 Hz—the magnetic component is dominating. In this paper the results of magnetic field strength measurements carried out in different relevant environments are presented. Also presented are the results of an experiment performed in order to quantify the thresholds by which the magnetic fields leave traceable ENF in various battery powered digital audio/ video recording equipment. In this experiment no battery powered digital recording device equipped with electret microphones were susceptible to the magnetic field.

1-2 **The Electric Network Frequency (ENF) as an Aid to Authenticating Forensic Digital Audio Recordings—An Automated Approach**—*Alan J. Cooper*, Metropolitan Police Service, London, UK

A recent forensic technique developed to establish the authenticity of recorded digital audio evidence is the Electric Network Frequency (ENF) Criterion. This paper confirms the applicability of the ENF criterion for use in mainland UK and introduces an automated approach to matching ENF estimates taken from a questioned recording to a database of ENF values. The signal processing procedures described have been used successfully by the Metropolitan Police Forensic Audio Laboratory in London to extract and match ENF data from evidential recordings.

1-3 **Digital Audio Authenticity Using the Electric Network Frequency**—*Richard W. Sanders*, National Center for Audio/Video Forensics, College of Arts & Media, University of Colorado Denver, Denver CO, USA

With the advent of personal computers, hand-held electronic devices, and the huge transition to all things digital Digital Audio Authenticity (DAA) has become an increasingly important issue. With digital audio it is easy for anyone with a computer to copy, alter, mix together or otherwise change the recording, and it is usually difficult to detect such changes. There is often no correct indication of when the recording was made, altered or recorded over. This paper is largely a test report on the Electric Network Frequency (ENF) variations in the three United States power grids. It will also show how using the ENF in digital audio can authenticate a recording, determine if it is a copy, time stamp the original and a copy, and possibly determine stops and starts in a recording. This research is partially funded by the Department of Justice.

Friday, June 6 **11:00 am**

EXHIBITOR INTRODUCTIONS

The following exhibitors will hold demonstrations throughout the Conference: Speech Technology Center, Russia; Cedar, Great Britain; Avalon, Boulder; IAS Forensics, Jonathan Broyles, Parker Colorado; University of Colorado Denver, Masters Degree in Audio/Video Forensics; various demonstrations and exhibitions.

Friday, June 6 **1:30 pm**

PAPER SESSION 2

2-1 **Quantifying the Speaking Voice: Further Investigation into Speaker Identification by a Simple Code-Matching Technique**—*Jeff M. Smith, Richard W. Sanders*, National Center for Audio/Video Forensics, College of Arts & Media, University of Colorado Denver, Denver CO, USA

This paper reports on the techniques refined for a method of speaker identification through the automated comparison of spectral, timbral, and temporal features unique to an individual's speech production. This method was first described in Convention Paper 7274 presented by the co-author of this paper, Richard Sanders, at the 123rd Convention of the Audio Engineering Society (New York). Since its first publication, the system (now referred to as SIDNI or Speaker Identification by Numerical Imprint) has improved from 79 percent correct identifications in 78 comparisons from the speech of 26 males to 100 percent correct identifications in 150 comparisons from the speech of 50 males. This paper will provide more information on these results and the results of several other tests while also elaborating on the specific speech characteristics exploited by the system and their potential for identification. Some characteristics include: average fundamental speaking frequency, ratio of spectral densities below 1 kHz to those above 1 kHz (Alpha ratio), average rate of vowels, jitter, and shimmer.

2-2 **The Role of Transcriptions in the Courtroom: A Scientific Evaluation**—*Fausto Poza,*[1] *Durand R. Begault*[2]
[1]Poza Consulting Services, Menlo Park, CA, USA
[2]Audio Forensic Center, Charles M. Salter Associates, San Francisco, CA, USA

The past thirty-five years have seen a heated debate in both scientific and legal venues as to the proven accuracy of

Forensic Voice Identification, which is now admissible in some jurisdictions but not in others. It is notable that there has been little, if any, scientific reporting on the common practice of allowing the use of transcriptions of difficult to understand recordings as aids to the trier of fact in court proceedings. The use of such transcripts, whether produced by "experts" or laypersons is discussed in the context of their potential for anonymously biasing the trier of fact. Signal Detection Theory shows that when subjective judgments are made in the presence of uncertainty, as is the case when trying to understand marginally intelligible recordings, the criterion of the decision maker can be significantly affected by external factors. When ruling on whether to allow transcripts of marginally intelligible recordings to be used as "aids" to the trier of fact, the Court should consider whether such "aids," rather that the recordings themselves will effectively become the evidence.

Saturday, June 7 **9:30 am**

PAPER SESSION 3

3-1 Audio Analysis and Spectral Restoration Workflow Using Adobe Audition—*Charles van Winkle*, Adobe Systems Incorporated, Seattle, WA, USA

To the audio professional, many advantages are given when working in the frequency-domain audio spectrum. The workflow is much like working in an image-editing environment; however many of the tools available to the audio professional are rudimentary compared to true image editing interfaces. This paper gives an overview of the new spectral audio restoration technologies included in Adobe Audition 3. Technologies to be covered include audio spectrogram conversion to and from bitmaps, image editing-like *Spot Healing-* & *Effects Paint-* brushes, audio editing in the stereo-panoramic domain, and audio editing in the phase domain. Best practice recommendations are included whenever possible.

3-2 Deciphering Gunshot Recordings—*Robert C. Maher, Steven R. Shaw*, Digital audio Signal Processing Laboratory, Department of Electrical and Computer Engineering, Montana State University, Bozeman, MT, USA

Audio gunshot recordings can be helpful for crime scene reconstruction, estimation of the shooter's location and orientation, and verification of eyewitness accounts. The audio evidence can include the muzzle blast, the shock wave signature if the projectile is traveling at supersonic speed, and possibly even the characteristic sound of the firearm's mechanical action if the recording is obtained close to the shooting position. To investigate the acoustical phenomena associated with gunshot evidence, a systematic set of rifle shots were made from distances ranging from 10 meters to nearly 800 meters away from the recording microphone. This paper summarizes the primary acoustical evidence derived from these recorded gunshots and suggests several strengths and weaknesses of gunshot analysis for forensic purposes.

3-3 Forensic Analysis of the Audibility of Female Screams—*Durand R. Begault*, Audio Forensic Center, Charles M. Salter Associates, San Francisco, CA, USA

Acoustical engineers and forensic acoustical experts are sometimes called upon to render opinions on the audibility of specific sounds at a given distance. Such sounds include speech, gunfire, warning signals such as fire alarms or locomotive horns, and in certain cases, human screaming. The audibility of female screaming has been questioned in several cases, where the expert can use both analytical and demonstrative techniques in order to form an opinion. The determination of audibility may be refined in terms of detection, discrimination and identification. This paper addresses measurement and typical levels of female screams, and reports on two different audibility analyses.

3-4 Magneto-Resistive Field Mapping of Analog Audio Tapes for Forensics Imaging—*David P. Pappas,*[1] *Ken Marr*[2]
[1]National Institute of Standards and Technology, Boulder, CO, USA
[2]FBI Engineering Research Facility, Audio Laboratory, Quantico, VA, USA

Magneto-resistive imaging of cassette tapes for forensic analysis applications is presented. Sample tapes were recorded on various tape recorders with events typically encountered in investigations. A comparison of the images with those obtained using ferrofluid is presented, along with comparisons of multiple instances of the same event with the various recorders

3-5 "Caught Red-Handed"—Forensic Audio Workflow: Covert Recording of Co-Conspirators —*Kent Gibson*, Forensicaudio.org, Los Angeles, CA, USA

Covert recording of newly arrested suspects yields fertile results. Suspects cannot resist "getting their stories straight." Advanced recording techniques and equipment make the acquisition easier and more decipherable. Enhancement using new forensic audio hardware and software vastly improves intelligibility. Diagrams of typical workflows are presented: (1) Planning; (2–4) Recording and Legalities; (5) Chain of Custody; (6, 8) Assessment and Plan; (7, 9–12) Enhancement; (13) Delivery and Documentation; (13,14) Court Testimony.

Saturday, June 7 **1:30 pm**

PAPER SESSION 4

4-1 Methods of Digital Media Authentication—*John Dinos; Randy Haines*

In the field of audio forensics, Digital Media Authenticity (DMA) continues to gain importance. Current methods of authentication use the Electric Network Frequency (ENF). The ENF signal, present with varying fluctuations in all three U.S. power grids, is embedded in just about any recording made. For authentication purposes, it is necessary to extract it from the recording in question and compare it to the raw ENF signal recorded directly from the power stations. Such a comparison can reveal a number of characteristics about the recording. One can determine, for example, whether the recording is an original or a copy, if it has been edited or recorded over, or if it contains starts or stops. What is the most efficient means of extracting the ENF? How does the extracted signal compare to the raw ENF? What problems are encountered when conducting such research?

4-2 SWGDE Best Practices for Forensic Audio—*Michael Piper,*[1,3] *David Hallimore*[2,3]
[1]U.S. Secret Service, Washington, D.C. USA
[2]Houston Police Department, Houston,TX, USA
[3]Scientific Working Group on Digital Evidence, USA

The Audio Committee of the Scientific Working Group on Digital Evidence (SWGDE) recently published the document "SWGDE Best Practices for Forensic Audio." This document provides forensic audio practitioners recommendations for the handling and examination of forensic audio evidence in a manner suitable for introduction into a court of law. An overview of this document, the SWGDE organization, and its Audio Committee will be presented. The document is available online at www.swgde.org/documents.

4-3 Speaker ID on Language Unknown to an Expert—*Polina Zubova*, Speech Technology Center, St. Petersburg, Russia

Speaker identification is one of the most important tasks of forensic audio investigations. It becomes more difficult if a person (suspect) speaks a language or dialect unknown to an expert. In our opinion the only solution to this problem is to give an expert the identification methodology that contains number of methods and tools of investigation, sufficient for decision on identity/difference of speakers and not dependent on speakers' language or nationality. Speech Technology Center, Ltd. devoted several years to work out this methodology, to test it in several real cases, and to train the experts to use it in forensic investigations.

AES 33rd INTERNATIONAL CONFERENCE

Audio Forensics—Theory and Practice

2008 June 5–7 • Denver, Colorado

THURSDAY, JUNE 5
TUTORIAL DAY

Opening/Welcome
1:00 pm

History of Audio Enhancement and Authentication: From Edison to the Present
Tom Owen
1:00 pm

Basics of Audio Enhancement
Gordon Reid
2:30 pm

Voice Identification
Fausto Poza, Eddy Brixen, Durand Begault
4:00 pm

Photogrammetry
Greg Stutchman
5:30 pm

Meet-and-Greet Party (refreshments)
6:30 pm

FRIDAY, JUNE 6
CONFERENCE DAY 1

Registration
8:30 am–9:30 am

Welcome / Paper Session 1
9:30 am

Exhibitor Introductions
Brief explanations about demos that will be given throughout the conference.
11:00 am

Lunch
12:00 noon

Exhibits Open
1:00 pm

Paper Session 2
1:30 pm

Exhibits Open
3:30 pm

Paper Session 2, continued
4:00 pm

Western BBQ
7:00 pm

SATURDAY, JUNE 7
CONFERENCE DAY 2

Paper Session 3
9:30 am

Lunch
12:00 noon–1:30 pm

Exhibits Open
1:00 pm

Paper Session 4
1:30 pm

Forensic Contest Awards
Winners will explain their methods.
3:00 pm

Closing Remarks
4:00 pm

Exhibits Open
4:15 pm

Sunday, June 8—Optional Rocky Mountain Tour
A bus tour of the Rocky Mountain area nearby Denver is planned for attendees and guests. It will include visits to a gold-mining town, internationally known skiing areas, the scenic region, and the Continental Divide. A separate fee will be established to cover bus transportation and lunch. Admission fees to attractions such as gondola rides, gold mine admission, etc., are not covered.

Schedule is subject to change. Check www.aes.org/events/33/ for updates.