AES 26TH INTERNATIONAL CONFERENCE
Audio Forensics in the Digital Age
Denver, Colorado, USA • July 7–9, 2005

The AES 26th International Conference, *Audio Forensics in the Digital Age,* will explore this fascinating topic that is so important today with many of our artistic, cultural, and political recordings being archived as digital files. Chair Roy Pritts and the other members of the conference committee have planned a unique event in mile-high Denver, which is just a short drive from the spectacular Colorado Rockies.

Denver's downtown, I. M. Pei-designed Adam's Mark Hotel will serve as the conference center. It has such amenities as a heated outdoor pool, fitness center, sauna, and high-speed Internet access. The hotel is one block away from the free bus service running all day on the 16th Street Mall, which connects with Denver's two light rail lines, museums, shopping, and the restaurants and night spots in historically preserved Larimer Square. Thursday's tutorial sessions will be held at the University of Colorado at Denver, just a short trip from the hotel.

TUTORIAL DAY
Thursday, July 7, will be devoted to four tutorials that will provide background information and a historical perspective to technical papers and workshops to follow on Friday and Saturday. After an introductory tutorial, Tom Owen will discuss authentication. After lunch Gordon Reid will cover enhancement in the third tutorial. The fourth and final tutorial, to be given by Rich Sanders, will look at voice identification. The tutorials may be attended as a single-day option or as part of the full conference program. See the registration form on the last page for more information about prices.

PAPERS AND WORKSHOPS
Papers Cochairs Richard Sanders and Tom Owen have planned a program of six sessions over two days. The first paper on Friday morning will be "Voice Identification and Elimination Using Aural-Spectrographic Protocols," presented by Durand Begault. This will be followed by Andrew Harper presenting "A Comparison of Digital and Analog Voice Spectrographs." Following a coffee break the morning's second session will begin with Bradd Fanberg presenting "Spectrographic Analysis of Vocal Alterations." Next up will be Sanja Grubesa presenting "Speaker Recognition Method Combining FFT, Wavelet Functions, and Neural Networks."

Following lunch there will be a panel discussion and workshop devoted to voice identification. Later in the the afternoon the third papers session will begin with "Validity of Bit-Compressed Digital Voice Recordings for Spectrographic Analyses: Input for a Database; a Preliminary Test," presented by Eddy Brixen. The final paper on Friday will be "Methods for Detection and Removal of Parasitic Frequency Modulation in Audio Recordings," presented by Bozena Kostek.

Andrzej Czyzewski will lead off Saturday morning's session when he presents "Multitask Noisy Speech Enhancement System." He will then be followed by Rob Maher presenting "Audio Enhancement Using Nonlinear Time-Frequency Filtering." After the coffee break there will be a panel discussion and workshop on audio enhancement.

Brian Brustad will lead off the papers sessions on Saturday afternoon with his presentation of "Tape Analysis and Authentication Using Multitrack Recorders." The next presentation, "Fiber Optical Accelerometer for Structure-Borne Audio," will be presented by Yuvi Kahana. The concluding papers session of the conference will begin with "A Survey of Audio Forensic Gunshot Investigations," presented by Brian Brustad. The final papers will be "Using Current Digital Watermarking Techniques to Authenticate Audio Recordings," presented by Pat Vaisvil, and "Proof for Origin and Ownership by Watermarking," presented by Daniel Stocker. See the next page for the complete technical program.

SOCIAL EVENTS
The social interaction and networking among attendees and presenters is always an added attraction in addition to the technical content of AES conferences. This personal contact will be enhanced by a cocktail party Thursday evening and an outdoor barbeque Friday night at a nearby dude ranch for an authentic taste of the food, hospitality, and music of the American West. On Sunday there will be an optional bus tour into the nearby Rocky Mountains to see the Continental Divide, a gold-mining town, and some of Colorado's world-famous ski resorts. Baseball fans may be interested to know that the Colorado Rockies are playing at home at downtown Coors Field this weekend against San Diego.

Plan to attend this exciting event in one of America's most dynamic and fast-paced cities with a back-drop of the majestic beauty of the Rocky Mountains. For more details and online registration see http://www.aes.org/events/26/.

AES 26th INTERNATIONAL CONFERENCE PROGRAM

Audio Forensics in the Digital Age
2005 July 7–9 • Denver, Colorado

Technical Sessions*
This preliminary program is accurate as of press time.

Thursday, July 7　　　　　　　　**9:00 am–10:15 am**

TUTORIAL 1:　INTRODUCTION AND HISTORY

Abstract Not Available at Press Time

　　　　　　　　　　　　　10:45 am–12:00 noon

TUTORIAL 2:　AUTHENTICATION

How to Conduct an Audio Authenticity Examination

Presenter:　　　**Tom Owen**, Owl Investigations, Inc.,
　　　　　　　Colonia, NJ, USA

Utilizing the 12-Step Authenticity Methodology, this author will illustrate in a step-by-step manner, the process to determine the authenticity of a tape. An authenticity examination includes determining whether or not the tape is an original; whether it has been edited in any way; whether any material has been omitted, deleted, inserted, or altered in any way, so as to change what actually occurred at the time of the original recording.

　　　　　　　　　　　　　1:30 pm–3:00 pm

TUTORIAL 3:　ENHANCEMENT

**The Selection and Use of Adaptive Filters
and Associated Signal Processing**

Presenter:　　　**Gordon Reid**, CEDAR Audio Ltd.,
　　　　　　　Cambridge, UK

Speech enhancement has come a long way in the digital era, but it is still far from the "magic wand" depicted in Hollywood movies. Adaptive filters have traditionally been the basis of much forensic audio work, but it is not always obvious which is the correct choice of filter. Furthermore, we have found that a combination of techniques—including broadband noise reduction, buzz removal, equalization, and background noise suppression—can provide superior results when compared with any single approach. This tutorial, illustrated using real examples processed on a CEDAR Cambridge Forensic system, aims to shed light on these choices, demonstrating how signal processing can aid investigators in areas including criminal investigation, terrorism countermeasures, and air accident investigation.

　　　　　　　　　　　　　3:30 pm–5:00 pm

TUTORIAL 4:　VOICE ID

Voice Identification

Presenter:　　　**Rich Sanders**

Abstract Not Available at Press Time

Friday, July 8　　　　　　　　**9:00 am–10:15 am**

PAPER SESSION 1

1-1　**Voice Identification and Elimination Using Aural-Spectrographic Protocols—**_Fausto Poza,_[1] _Durand R. Begault_[2]
[1]Poza Consulting Services;
[2]Charles M. Salter Associates, San Francisco, CA, USA

The use of spectrographic pattern matching and aural comparison in forensic voice identification requires careful control of examiner bias and an awareness of the principals of signal detection theory. This paper briefly reviews the history of the aural-spectrographic method and the experimental results of Oscar Tosi, and then summarizes the voice elimination and identification protocols addressed by Gruber and Poza (American Jurisprudence 54, 1995) and other authors. Given suitable voice exemplars, and prudent examination protocols, voice examiners can provide useful probative findings in the forensic setting. Opinions expressed in reports and in courtroom testimony should set forth the limitations of the technique so that the trier of fact can properly evaluate the examiner's findings.

1-2　**A Comparison of Digital and Analog Voice Spectrographs—**_Andrew Harper, Esther Spence, Kevin Garland_, University of Colorado at Denver, Denver, CO, USA

Analog spectrographic representations of voice recordings have had a fluctuant role in audio forensics. Although recorded evidence has been admitted in court since 1959, forensics scientists still struggle to establish analog spectrographic evidence used for voice identification as a solid, reliable scientific method. Media across the board continue to converge on the digital format, and the spectrograph is no exception. Various computer programs already exist to create spectrographs digitally. Their accuracy, usefulness, and viability as evidence in courtrooms remain in question, however. This paper is an investigation into these issues—both the digital spectrograph as it relates to its analog counterpart and as it relates to other

digital representations.

Friday, July 8   **10:45 am–12:00 noon**

PAPER SESSION 2

2-1   Spectrographic Analysis of Vocal Alterations—
Bradd Fanberg, Jeff Smith, Rebecca Wright,
University of Colorado Denver, Denver, CO, USA

With advances in digital spectrographic analysis, the possibility of real-time voice identification for security purposes can be realistically considered. In developing technologies of this sort, however, falsification of identity has to be considered as well. If the spectrographic signature of a person's voice is the result of unique nuances in their physical build, can this vocal "fingerprint" be altered by manipulating one's voice? To what degree can spectrographic data change when a subject's vocalizations are altered?

To answer these questions, this paper presents comparisons in spectrographic data gathered from various individuals. Samples from each individual include: (1) spectrographic analysis of individual's natural speech; (2) spectrographic analysis of individuals consciously raising the pitch of their voice; (3) spectrographic analysis of individuals with intentionally restricted airflow; and (4) spectrographic analysis of individuals mimicking one of two control samples: male or female.

2-2   Speaker Recognition Method Combining FFT, Wavelet Functions, and Neural Networks—*Sanja Grubesa, Tomislav Grubesa, Hrvoje Domitrovic*,
University of Zagreb, Zagreb, Croatia

The method of speaker recognition based on wavelet functions and neural networks is presented in this paper. The wavelet functions are used to obtain the approximation function and the details of the speaker's averaged spectrum in order to extract the speaker's voice characteristics from the frequency spectrum. The approximation function and the details are then used as input data for decision-making neural networks. In this recognition process, not only the decision on the speaker's identity is made, but also the probability that the decision is correct can be provided.

Friday, July 8   1:30 pm–3:00 pm

WORKSHOP A

Voice Identification: A Panel Discussion

Abstract Not Available at Press Time

Friday, July 8   3:30 pm–5:00 pm

PAPER SESSION 3

3-1   Validity of Bit Compressed Digital Voice Recordings for Spectrographic Analyses: Input for a Database, a Preliminary Test—*Eddy Brixen*,[1]
Durand R. Begault[2]
[1]EBB Consult., Smorum, Denmark
[2]Charles M. Salter Associates, San Francisco, CA, USA

In the spectrographic analysis of voice exemplars, there is a question of the impact of commercially available compression schemes on the visual representation of the voice. Specifically, there is a question whether or not the codec reconstruction of the time-domain waveform might alter features of the

wide-band spectrogram, including formant shape and position. This paper contributes a collection of comparative analyses of Danish and English speech exemplars that are recorded simultaneously via both linear and compression-based recording methods. G.723.1-encoded recordings (ACELP-based algorithm, max 6.3 kbps) used in IC-recorders and surveillance recording equipment and the Sony Memory stick format using MSV LPEC-SP compression are evaluated.

3-2   Frequency-Domain Processors for Efficient Removal of Noise and Unwanted Audio Events—
Christoph M. Musialik, Ulrich Hatje, Algorithmix GmbH, Waldshut-Tiengen, Germany

Filtering heavily polluted audio signals with low-cut, high-cut or notch filters helps only if the spectrum of disturbances does not overlap the desired signal. Unfortunately, this is rarely the case in real situations. Therefore the only way to achieve dramatic improvements when cleaning up noisy signals is by using frequency-domain methods. Spectral denoising works well in the case of broadband, smoothly changing disturbances. However, the situation becomes hopeless when the noise character changes abruptly and the S/N ratio is around zero decibels. We present a unique tool, reNOVAtor, working in 3-D domain (time, frequency, amplitude) allowing efficient removal of tenacious disturbances. An intuitive man-machine interface allows localization, identification, and very precise removal of unwanted audio events. Special functions like automatic detection of clicks, tones, and harmonics significantly accelerate the workflow. A special difference function allows monitoring of the portion being removed.

3-3   Methods for Detection and Removal of Parasitic Frequency Modulation in Audio Recordings—
Andrzej Czyzewski, Bozena Kostek, Przemyslaw Maziewski, Marek Dziubinski, Andrzej Ciarkowski, Maciej Kulesza, Jozef Kotus, Gdansk University of Technology, Gdansk, Poland

The paper presents several methods devoted to wow defect estimation and compensation in audio recordings. The presented algorithms utilize time- and spectral-domain audio processing routines. The time-domain algorithm employs autocorrelation analysis of short-term pitch variations. The spectral methods include power-line hum and bias frequency tracking as well as sinusoidal components fluctuation analysis. The proposed algorithms were tested based on some archival audio recording samples allowing individual procedures effectiveness evaluation in practice.

Saturday, July 9   9:00 am–10:15 am

PAPER SESSION 4

4-1   Multitask Noisy Speech Enhancement System—
Andrzej Czyzewski,[1] *Jozef Kotus,*[1] *Grzegorz Szwoch,*[1] *Andrzej Rypulak,*[2] *Arkadiusz Pawlik*[2]
[1]Gdansk University of Technology, Gdansk, Poland
[2]Air Force Academy, Deblin, Poland

This paper includes a general characteristic of the designed and implemented Multitask Noisy Speech Enhancement System providing a specialized software suite designed for recording speech and for improving quality and intelligibility of a recorded speech signal based on some innovative digital signal processing algorithm applications. The system con-➡

sists of the following applications: restorer, recorder, and browser. The software may be used in all cases when speech intelligibility is important, but it is not possible to obtain high quality speech recordings.

4-2   Audio Enhancement Using Nonlinear Time-Frequency Filtering—_Robert Maher_, Montana State University, Bozeman, MT, USA

Forensic audio recordings may contain undesired noise that can impair source identification, speech recognition, and other audio processing requirements. In this paper several custom analysis/synthesis algorithms are presented based on a time-varying spectral representation of the noisy signal. The enhancement process adapts to the instantaneous signal behavior and alters the noisy signal so that the enhanced output signal is higher in quality than the unprocessed noisy input signal. Nonlinear and time-varying filters operate on the spectral representation in order to retain features that are attributable to the desired signal, such as human speech, while removing the features that are more likely to be due to the noise contamination. Audio examples are presented.

Saturday, July 9   10:45 am–12:00 noon

WORKSHOP B

Audio Enhancement: A Panel Discussion

Abstract Not Available at Press Time

Saturday, July 9   1:30 pm–3:00 pm

PAPER SESSION 5

5-1   Tape Analysis and Authentication Using Multitrack Recorders—_Brian Brustad, Durand R. Begault_, Charles M. Salter Associates, San Francisco, CA, USA

The technique of offset multitrack (> quarter track) playback heads for forensic analysis of standard monophonic and stereophonic cassette and microcassette tape recordings is discussed. The time-domain waveform for recorded signatures can be analyzed in terms of relative timing offset for determining azimuth of the record head used in making a specimen tape. Additionally, excursion of the erase signature into the guard band region is a reliable indicator of an original versus a copied erase signature. Relocation of the playback head allows capturing "record on" signatures that are not normally captured. Comparative advantages over magnetic development techniques are discussed.

5-2   Fiber Optical Accelerometer for Structure-Borne Audio—_Yuvi Kahana, Alexander Kots_, Phone-Or Ltd., Or-Yehuda, Israel

The fiber optical accelerometer (FOA) is a MEMS (Micro Electro-Mechanical Systems) transducer whose transduction principle is based on the intensity modulation of light. The sensor includes three main parts: the optical accelerometer head and its membrane; the optical fibers; and at a remote location, the electro-optic unit which includes the source of light (LED), the photodetector, and their associated electronics.

    In this paper, we first present the physical mechanisms, sensors' construction, and performance. The

acoustical characteristics of the FOA are reviewed and studied. Wave files of audio will be presented for various and difficult scenarios such as listening to speech through thick concrete walls, frozen conditions, walls masked by stationary noise, and others.

5-3   Law and the Expert Witness—The Admissibility of Recorded Evidence—_Tom Owen_,[1] _Michael McDermott_,[2] _Jennifer Owen_[1]
[1]Owl Investigations, Inc., Colonia, NJ, USA
[2]Frank McDermott Ltd., Great Falls, VA, USA

Audio tapes, video tapes, voice Identification, and other matters related to the admissibility of recorded media must pass muster with the courts before it is accepted as evidence. This presentation will be a step-by-step paper on the applicable law, the correct manner of presentation, and the rules that need to be followed when writing reports for trial when dealing with recorded media.

Saturday, July 9   3:30 pm–5:00 pm

PAPER SESSION 6

6-1   A Survey of Audio Forensic Gunshot Investigations—_Jack Freytag_, Charles M. Salter Associates, San Francisco, CA, USA

Audio recordings of gunshots may be used with other forensic evidence to identify (or eliminate) gunshot source locations, to establish certain attributes of the weapon fired, and/or elucidate or eliminate certain physical parameters of the gunshot surroundings. Under very controlled circumstances audio recordings have been able to generally identify or eliminate the type of weapon fired. Audio analyses of gunshots cannot, however, generally identify the specific weapon fired nor specific ballistic parameters due to the gain control electronics invariably built into audio recorders typically in use around crime scenes. A survey of firearm acoustics is given, along with several case studies.

6-2   Using Current Digital Watermarking Techniques to Authenticate Audio Recordings—_Pat Vaisvil, Adam Olson, Jose Quinones_, University of Colorado at Denver, Denver, CO, USA

Techniques to verify the authenticity of an analog tape have been in existence for quite some time, for example taking advantage of anomalies with respect to the starting and stopping of recorders through magnetic development. With the advent of digital technology, many of these techniques are not applicable. In this paper we will examine several techniques used in other facets of the industry (primarily in terms of copyright protection) and examine their potential usefulness in validating whether or not a digital recording is authentic.

6-3   Proof of Origin and Ownership by Watermarking—_Daniel Stocker, Jean-Christophe Kummer_, NOA Audio Solutions, Vienna, Austria

The presented paper examines the general concept behind watermarking set against the requirements of audio forensics and concludes that watermarking—being a convenient way to mark and track recordings—can be turned into an effective tool in the hands of investigation only when the detection channel is secured. It is not enough to have an embedding and detection algorithm on their own: the input and output processes must be backed up by an individual identifier allowing or prohibiting access to the data hosted within the audio.

AES 26th INTERNATIONAL CONFERENCE
Audio Forensics in the Digital Age
2005 July 7–9 • Denver, Colorado

THURSDAY, JULY 7
TUTORIAL DAY

- **Tutorial 1**
  Introduction and History
  9:00 am–10:15 am

- **Tutorial 2**
  Authentication
  10:45 am–12:00 noon

- **Lunch**
  12:00 noon–1:30 pm

- **Tutorial 3**
  Enhancement
  1:30 pm–3:00 pm

- **Tutorial 4**
  Voice Identification
  3:30 pm–5:00 pm

- **Conference Registration**
  6:00 pm–8:00 pm

- **Cocktail Party**
  6:00 pm–8:00 pm

FRIDAY, JULY 8
CONFERENCE DAY 1

- **Conference Registration**
  8:00 am–9:00 am

- **Welcome / Paper Session 1**
  9:00 am–10:15 am

- **Paper Session 2**
  10:45 am–12:00 noon

- **Lunch**
  12:00 noon–1:30 pm

- **Workshop A**
  Voice Identification
  1:30 pm–3:00 pm

- **Paper Session 3**
  3:30 pm–5:00 pm

- **BBQ Western Evening**
  5:30 pm–10:00 pm

SATURDAY, JULY 9
CONFERENCE DAY 2

- **Paper Session 4**
  9:00 am–10:15 am

- **Workshop B**
  Audio Enhancement
  10:45 am–12:00 noon

- **Lunch**
  12:00 noon–1:30 pm

- **Paper Session 5**
  1:30 pm–3:00 pm

- **Paper Session 6 / Closing Remarks**
  3:30 pm–5:00 pm

Sunday, July 10—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 http://www.aes.org/events/26/ for updates.