AFEKA TEL- AVIV ACADEMIC
COLLEGE OF ENGINEERING

APPLIED VOICE INPUT/OUTPUT SOCIETY

2011 Speech Processing Conference
June 21-22

www.speechprocessingday2011.com

ATHENA
Security Implementations Ltd.

AudioCodes

AUDIOZOOM
pure sound technology

BINATIX
Machine Intelligence

FST21

GM

NICE · Intent. Insight. Impact.™

NUANCE

OPTIMUMGROUP

VERINT
POWERING ACTIONABLE INTELLIGENCE®

2011 Speech Processing Conference

Conference Chairman

Prof. Ami Moyal, ACLP, Afeka
SPD-Chairman@afeka.ac.il

Organizing Committee

Dr. Vered Aharonson, ACLP, Afeka
vered@afeka.ac.il

Ms. Adina Caspi, VP Public Affairs, Afeka
SPD-Sponsorship@afeka.ac.il

Prof. Ami Moyal, ACLP, Afeka
SPD-Chairman@afeka.ac.il

Dr. Nava Shaked, Chairman, AVIOS Israel
nava@business-tech.co.il

Technical Program Committee

Dr. Vered Aharonson, ACLP, Afeka, Israel
Prof. Itamar Arel, University of Tennessee, USA
Dr. Sara Basson, IBM, India
Dr. Khalid Choukri, ELRA, France
Prof. Israel Cohen, Technion, Israel
Prof. Sharon Gannot, Bar Ilan University, Israel
Dr. Dina Goren-Bar, Afeka and Optimum Group, Israel
Prof. Yossi Keshet, Toyota Technological Institute at Chicago, USA
Dr. Itshak Lapidot, Sami Shamoon College, Israel
Prof. David Malah, Technion, Israel
Dr. Joseph Mariani, LIMSI-CNRS and IMMI, France
Dr. William Meisel, TMA Associates, USA
Prof. Paul Mermelstein, University of Miami, USA
Prof. Asunción Moreno, Universitat Politecnica De Catalunya, Spain
Prof. Ami Moyal, ACLP, Afeka, Israel
Dr. David Nahamoo, IBM, USA
Prof. Lawrence Rabiner, Rutgers University, USA
Dr. Nava Shaked, BBT Ltd., Israel
Prof. Ilan Shallom, AudioCodes, Israel
Dr. Lorin Wilde, MIT and Wilder Communications, USA
Dr. Matthew Yuschik, MultiModal Interfaces, LLC, USA

Technical Committee Secretariat

Ms. Michal Gishri, ACLP, Afeka
speechprocessingday2011@afeka.ac.il

Registration Fees

Tuesday, June 21: Pre-Conference Seminars

Four pre-conference seminars will be presented by:

- Prof. Alan Oppenheim, MIT
- Dr. Sara Basson, IBM and Prof. Ami Moyal, Afeka
- Prof. Lawrence Rabiner, Rutgers University
- Dr. Matthew Yuschik, Multimodal Interfaces

Seminar Fee: 1000 NIS ($270) per seminar

Those registering for two seminars will be admitted free of charge to the main conference on June 22.

Wednesday, June 22: Speech Processing Conference

Keynote Speaker: **Dr. Michael Cohen, Google**

"Scaling Up Speech Technology"

Sessions:
- Human-Machine Interaction
- Speech Recognition
- Speaker Diarization and Verification
- Applications
- Speech Processing in the Security Market
- Speaker Traits
- Feature Extraction

Main Conference Fee: 250 NIS ($70)

Participants may register for either day or both.
50% discount on a limited number of seats for students.

See conference website for additional details on registration procedures
www.speechprocessingday2011.com

Contact Information:

Registration:
SPD-Registration@afeka.ac.il or +972-03-768-8663
Paper Submission:
speechprocessingday2011@afeka.ac.il
Sponsorship:
SPD-Sponsorship@afeka.ac.il

FREE one-year AVIOS membership for all participants!

Conference Overview

The Afeka Center for Language Processing (ACLP) and AVIOS Israel have joined together to create the first annual "Speech Processing Conference" intended to bring together the local Israeli speech community from both academia and the industry. Researchers and developers from any Speech Processing field are invited to attend. On June 21, four pre-conference seminars will be presented by **Prof. Alan Oppenheim**, **Dr. Sara Basson**, **Prof. Lawrence Rabiner** and **Dr. Matthew Yuschik**. The main conference on June 22 will feature **Dr. Michael Cohen of Google** as the keynote speaker and a range of sessions on various speech processing topics.

Words from Prof. Ami Moyal, Chairman, 2011 Speech Processing Conference

Welcome to the 2011 Speech Processing Conference at Afeka!

Following the success of last year's local conference, we have decided to establish a yearly conference at Afeka on Speech Processing. This year's conference will provide another excellent opportunity for the industry and academia in Israel to meet, network, exchange ideas and catch up on recent research and technology trends. I am confident the conference will contribute to the establishment of a connected and productive speech processing community in Israel.

I would like to take this opportunity and give thanks to our international guests for their willingness to join us and share their viewpoints and their wealth of experience, to our industry sponsors for their support, to the AVIOS organization for their collaboration, and particularly to AVIOS Israel chairman, Dr. Nava Shaked, for her contribution to the success of this event.

Words from Dr. K.W. Scholz, President, AVIOS

It has been long recognized that a most effective vehicle for facilitating communication among professionals is a national or international conference where participants meet face-to-face to share their latest technical achievements from the podium or the conference room. As a long-time believer in the importance of such communication, AVIOS has sponsored regional, national, and international conferences for over a quarter of a century. We have conducted these throughout the US as well as Europe and Israel, and as plans unroll for our 2011 conference in Israel, we have chosen to combine Afeka College resources with our own to create an even more significant offering. The AVIOS-Afeka collaboration establishes credentials which will not only attract key presenters and participants, but will also appeal to significant corporate sponsors, ensuring the event's financial solvency. Hence, after 29 years of conference sponsorship including 6 in Israel, we are confident that the 2011 AVIOS-Afeka conference will be an event worthy of participation by leading speech technology professionals from around the world.

ACLP - Afeka Center for Language Processing

The Afeka Center for Language processing (ACLP) is part of the Afeka Academic College of Engineering and is a unique research, development and instruction laboratory located in Israel and specifically dedicated to the field of language processing. The center's activities include: a research and instruction laboratory for speech and text processing; joint R&D projects with the industry; grant funded research projects; consulting services to corporate, government and consumer industries; specialized courses for language processing professionals; and project opportunities for students in their final year of studies. The center is lead by Prof. Ami Moyal, previously CEO of NSC, who has over 20 years experience in the field of Speech Recognition. The research team includes Dr. Vered Aharonson, a specialist in text processing and emotion detection, linguists specializing in phonetics and discourse analysis, signal processing and speech recognition engineers and research associates.

AVIOS - Applied Voice Input Output Society

AVIOS was launched as a non-profit organization for speech technology professionals nearly 25 years ago. AVIOS sees great value in bringing together all of these application sectors, customers and professional interests, in a single speech forum, for continued opportunities to share best practices, to identify synergies, and to leverage advances in one application area across a range of other areas. The Israeli local chapter of AVIOS was officially announced in 2005. Dr. Nava Shaked is the chairperson of the local chapter and is also an active member of the AVIOS board. Recent years have witnessed the growth of the speech community in Israel. Israeli companies and startups have raised the interest of the international community with their ingenuity and creativity. AVIOS Israel leverages the best practices, and shares them across our speech community. For the past 5 years the annual AVIOS Israel conference has attracted hundreds of participants - both clients and suppliers.

TUESDAY, JUNE 21

PRE-CONFERENCE SEMINARS

10:00-13:00 Morning Session: Parallel Seminars

Fundamentals of Phase, Group Delay and Cepstral Analysis
Prof. Alan Oppenheim, MIT

In this seminar the interpretation of spectral phase through the concepts of phase unwrapping and group delay will be discussed along with their relationship to minimum-phase and minimum-energy-delay signals and systems. This interpretation also provides the basis for defining the cepstrum and understanding its properties. The use of the cepstrum for speech analysis and synthesis will also be explored.

Ficus Building, Room 203

Speech Transcription Overview - Markets, Applications & Technologies
Dr. Sara Basson, IBM and Prof. Ami Moyal, Afeka

Speech Transcription is core to ASR capabilities and thus to the potential markets targeted by Speech Recognition technologies. To this effect, it is important to gain an in-depth understanding of the overall importance of developing technologies in the field, as well as, the challenges involved in Speech Transcription and possible methods for overcoming these challenges. This Seminar will begin with Prof. Ami Moyal presenting a technical overview of Speech Transcription and will continue with Dr. Sara Basson presenting target applications and markets.

Ficus Building, Room 204

13:00-15:00 Lunch Break - *Ficus Building, Upper Floor Lobby*
Morning and afternoon participants are invited to attend

15:00-18:00 Afternoon Session: Parallel Seminars

Fundamentals of Hidden Markov Models and their Application to Speech Recognition
Prof. Lawrence Rabiner, Rutgers University and University of California, Santa Barbara

The Hidden Markov Model has become the most dominant technology used in speech recognition today. This lecture will begin with a review of the theory of Markov Models (both discrete Markov processes and hidden Markov processes), and continue with a discussion on solutions to the three basic problems when using HMMs, namely computation of observation probabilities, determination of "optimal state sequences", and optimal training of model parameters. We will then discuss variations of elements of the HMM such as different model types, and the use of continuous Gaussian mixture densities for modeling the probability density functions of speech parameters. We will spend some time on the key implementation issues, namely scaling in the forward-backward algorithm, the use of multiple observation sequences for non-ergodic HMMs, the impact of different methods of making initial parameter estimates, and the effect of having insufficient training data for the chosen size of the HMM.

Ficus Building, Room 203

Multimodal Human-Machine Interaction
Dr. Matthew Yuschik, Multimodal Interfaces, LLC

Human-Computer behavior has evolved so that end-users expect computer devices to support various input and output modes. The flexibility that multimodal interfaces provide on products and services has already begun to show the value it has added to the interface. Consequently, it is important to understand the concepts of multimodality to see how it can be best utilized in products and services. This seminar begins with a definition of multimodality, and shows examples of modality usage for mobile devices, games and call-center workstations. The pragmatics of business case components are then addressed through a discussion of end-user needs assessment and identification of opportunities in key market segments.

Ficus Building, Room 204

SPEECH PROCESSING CONFERENCE

8:30-9:00 **Networking and Registration**

9:00-9:30 **Welcoming Remarks** - *Hakirya Building, Main Auditorium*
Prof. Ami Moyal - Conference Chairman
Dr. Nava Shaked - AVIOS Israel Chairman
Mr. Eden Bar Tal - General Manager, Ministry of Communication

9:30-10:30 **Keynote Speaker**
Scaling Up Speech Technology
Dr. Michael Cohen, Google, USA

10:30-11:00 **Coffee Break** - *Hakirya Building, Room G-5*

11:00-12:00 **Parallel Sessions**

• **Speech Recognition** - *Hakirya Building, Main Auditorium*
Session Chair: Prof. Lawrence Rabiner

11:00-11:15 **The Power of Fusing Speech Recognition & Voice Biometrics Technologies**
Almog Aley-Raz - Nuance Communications

11:15-11:30 **Acoustic Spatiotemporal Modeling using Deep Machine Learning for Robust Phoneme Recognition**
Itamar Arel - The University of Tennessee, USA
Shay Berant and Tsvi Slonim - Binatix, USA
Ami Moyal - Afeka Academic College of Engineering

11:30-11:45 **Robust Speech Recognition using an Auxiliary Laser-Diode**
Yekutiel Avargel, Tal Bakish, Gabi Horovitz, Yechiel Kurtz - Audiozoom

11:45-12:00 **Natural Speech Interface in Domain-Specific Dialogue**
Dikla Shechnik, Noam Lotner, Michal Gishri, Ami Moyal, Vered Aharonson -
ACLP, Afeka Academic College of Engineering

• **Human-Machine Interaction** - *Hakirya Building, Hall D*
Session Chair: Dr. Nava Shaked

11:00-11:15 **Hands Free Technical Manual**
Dina Goren-Bar - Afeka Academic College of Engineering and Optimum Group
Tal Sasson - Afeka Academic College of Engineering

11:15-11:30 **Multimodal Interfaces for Micro User Actions**
Eran Aharonson - Intuitive User Interfaces
Vered Aharonson - Afeka Academic College of Engineering

11:30-11:45 **Horror as Never Before! The First Interactive Cine Film**
Detlev Artelt - Aixvox, Germany
Nava Shaked - Brit Business Technologies

11:45-12:00 **Identifying Cultural Aspects in Use of In-Vehicle Speech Applications**
Ute Winter, Omer Tsimhoni - General Motors, Advanced Technical Center, Israel

12:00-13:00   **Parallel Sessions**

• **Speaker Diarization and Verification** - *Hakirya Building, Main Auditorium*
Session Chair: Dr. Matthew Yuschik

12:00-12:15   **Speaker Diarization using Unsupervised Compensation of Within-Speaker Variability**
Hagai Aronowitz - IBM, Haifa Research Lab

12:15-12:30   **Unsupervised Speaker Separation for Contact Center Interactions**
Tzach Ashkenazi, Shimrit Artzi, Ronen Lapardon, Merav Ben-Asher, Gil Dobry,
Moshe Wasserblat, Oren Pereg - NICE

12:30-12:45   **Detecting Goats in Speaker Verification Systems**
Orith Toledo-Ronen - IBM, Haifa Research Lab

12:45-13:00   **Future Challenges in Speaker Diarization**
Itshak Lapidot - Sami Shamoon College
Hugo Guterman - Ben Gurion University

• **Applications** - *Hakirya Building, Hall D*
Session Chair: Dr. Sara Basson

12:00-12:15   **Real-Time Speech Analytics**
Moshe Wasserblat, Oren Pereg, Tzach Ashkenazi, Shimrit Artzi, Ronen Lapardon
Merav Ben-Asher - NICE

12:15-12:30   **TTS Application: Earth Surface as an Info-Sphere for Traveler**
Alexei Cherenkov - Toozla, Russia

12:30-12:45   **Implementation of Speech Recognition Technology in Smart-Home Applications**
Zohar Dvir - Zohar Dvir Yazamut

12:45-13:00   **Building a Dynamic Corpus for Hebrew: Observations, Challenges & Solutions**
Ron Hasson - Aharon Speech Technologies

13:00-14:30   **Lunch Break** - *Ficus Building, Upper Floor Lobby*

14:30-15:20   **Joint Session**

• **Speech Processing in the Security Market** - *Hakirya Building, Main Auditorium*
Session Chair: Prof. Ami Moyal

14:30-14:55   **Safe Rise – Challenging the Ability to Fusion**
Shahar Belkin - FST21

14:55-15:20   **The Requirements of an Intelligence Organization from Speech Processing Technologies**
Shabtai Shavit - Athena
Former Director General of the Israeli Mossad

15:20-15:40   **Coffee Break** - *Hakirya Building, Room G-5*

15:40-16:40   **Parallel Sessions**

- **Speaker Traits** - *Hakirya Building, Main Auditorium*
  Session Chair: Dr. Vered Aharonson

15:40-15:55   **Selecting Training Data for Cross-Corpus Speech Emotion Recognition: Prototypicality vs. Generalization**
Björn Schuller, Zixing Zhang, and Gerhard Rigoll - Institute for Human-Machine Communication, Germany

15:55-16:10   **Speaker Normalization in Detecting Angry Speech**
Noam Amir - Department of Communication Disorders, Sackler Faculty of Medicine, Tel-Aviv University

16:10-16:25   **Age Group Classification Based on Introduction Sentences**
Ron M. Hecht, Aharon Bar Hillel, Dan Levi, Ute Winter, Ran Y. Gazit, Omer Tsimhoni - General Motors, Advanced Technical Center, Israel

16:25-16:40   **Cepstral and Wavelets Methods for Pitch Tracking in Speech Processing of Tonal Languages**
Victor Katsnelson, Eduard Polityko - Speech Modules

- **Feature Extraction** - *Hakirya Building, Hall D*
  Session Chair: Prof. Alan Oppenheim

15:40-15:55   **Non–Stationary Analysis of Speech Applied to Perceptual Linear Prediction**
Oron Gamliel - Department of Electrical and Computer Engineering, Ben Gurion University
Ilan D. Shallom - NSC Natural Speech Communication

15:55-16:10   **Glottal Inverse Filtering: a New Road-Map and First Results**
Sandra Dias, Ricardo Sousa, Aníbal Ferreira - University of Porto, School of Engineering, Portugal

16:10-16:25   **Advances in Spectral Parameterization for Statistical (HMM-Based) TTS**
Slava Shechtman - IBM, Haifa Research Lab

16:25-16:40   **On Feature Extraction for Voice Pathology Detection from Speech Signals**
Aharon Satt, Zvi Kons, Ron Hoory - IBM, Haifa Research Lab

16:40-17:00   **Closing Remarks** - *Hakirya Building, Main Auditorium*
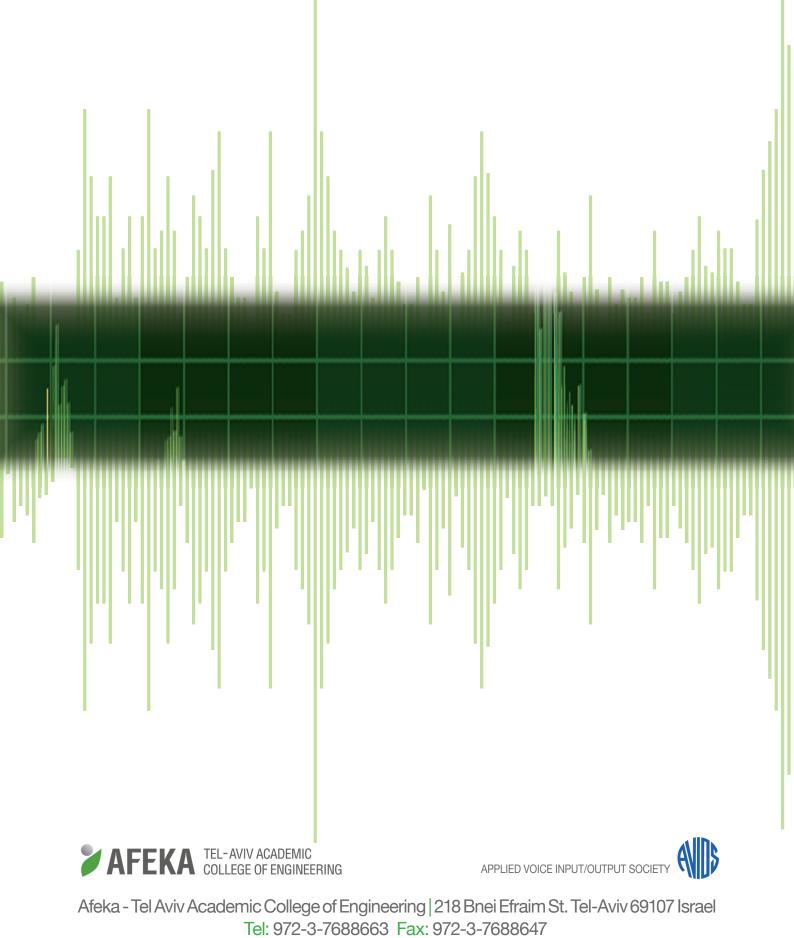

CONFERENCE VENUE

The 2011 Speech Processing Conference will be held at the:

Afeka - Tel-Aviv Academic College of Engineering

- Hakirya building, 218 Benei Efraim St.
- Ficus building, 38 Mivtsa Kadesh St.

Free Parking: 218 Benei Efraim St.

AFEKA TEL-AVIV ACADEMIC
COLLEGE OF ENGINEERING

APPLIED VOICE INPUT/OUTPUT SOCIETY

Afeka - Tel Aviv Academic College of Engineering | 218 Bnei Efraim St. Tel-Aviv 69107 Israel
Tel: 972-3-7688663 Fax: 972-3-7688647
www.speechprocessingday2011.com | speechprocessingday2011@afeka.ac.il

ATHENA
Security Implementations Ltd.

AudioCodes

AUDIOZOOM
pure sound technology

BINATIX
Machine Intelligence

FST21

GM

NICE Intent. Insight. Impact.

NUANCE

OPTIMUMGROUP

VERINT
POWERING ACTIONABLE INTELLIGENCE