The Speaker and Language Recognition Workshop

Odyssey 2012
Singapore

http://www.odyssey2012.org

25-28 June, 2012
Matrix Building, 30 Biopolis Street, Singapore 138671
Organizing Committee

Haizhou Li, Institute for Infocomm Research, Singapore, Chair
Kay Berkling, Inline Internet Online Dienste GmbH, Germany
Jean-François Bonastre, University of Avignon, France
Niko Brümmer, Agnitio, South Africa
Lukáš Burget, SRI, USA
Joseph Campbell, MIT Lincoln Lab, USA
Jan “Honzá” Černocký, Brno University of Technology, Czech Republic
Alvin Martin, NIST, USA
Douglas Reynolds, MIT Lincoln Lab, USA

Scientific Committee

Eliathamby Ambikairajah, U. of New South Wales, Australia
Tim Anderson, AFRL, USA
Walt Andrews, BBN, USA
Aladdin Ariyaeeinia, U. of Hertfordshire, UK
Hagai Aronowitz, IBM Haifa Research Lab, Israel
Roland Auckenthaler, NMS Comm., USA
Claude Barras, LIMSI, France
Laurent Besacier, CLIPS-IMAG, France
Bill Campbell, MIT Lincoln Lab, USA
Mauro Falcone, FUB, Italy
Kevin Farrell, Nuance, USA
Sadaoki Furui, Tokyo Tech, Japan
Ran Gazit, General Motors, Israel
Joaquin Gonzalez-Rodriguez, U.A. de Madrid, Spain
Craig Greenberg, NIST, USA
Hynek Hermansky, JHU, USA and BUT, Czech Republic
Patrick Kenny, CRIM, Canada
Tomi Kinnunen, U. of East Finland, Finland
Tina Kohler, DoD, USA
Pietro Laface, Politecnico di Torino, Italy
Tomoko Matsui, Institute of Statistical Mathematics, Japan
Jiri Navratil, IBM, USA
Jason Pelecanos, IBM, USA
Daniel Ramos, U. A. de Madrid, Spain
Nicolas Scheffer, SRI, USA
Andreas Stolcke, Microsoft Speech Labs, USA
Doug Sturim, MIT Lincoln Lab, USA
Pedro Torres-Carrasquillo, MIT Lincoln Lab, USA
David van Leeuwen, Radboud University Nijmegen, The Netherlands
Robbie Vogt, Queensland U.of Technology, Australia
Michael Wagner, U. of Canberra, Australia
Local Organizing Committee

Haizhou Li, Institute for Infocomm Research, Singapore
Nancy Chen, Institute for Infocomm Research, Singapore
Anthony Larcher, Institute for Infocomm Research, Singapore
Kong Aik Lee, Institute for Infocomm Research, Singapore
Cheung-Chi Leung, Institute for Infocomm Research, Singapore
Bin Ma, Institute for Infocomm Research, Singapore
Swee Lan See, Institute for Infocomm Research, Singapore
Hanwu Sun, Institute for Infocomm Research, Singapore
Rong Tong, Institute for Infocomm Research, Singapore
Xiong Xiao, Temasek Laboratories, Nanyang Technological University, Singapore
Changhuai You, Institute for Infocomm Research, Singapore

Reviewing Committee

Eliathamby Ambikairajah, Tim Anderson, Walt Andrews, Aladdin Ariyaeinia, Hagai Aronowitz, Roland Auckenthaler, Claude Barras, Kay Berkling, Jean-François Bonastre, Niko Brümmer, Lukáš Burget, Bill Campbell, Joseph Campbell, Jan “Honza” Černocký, Nancy Chen, Nicholas Evans, Mauro Falcone, Kevin Farrell, Sadaoki Furui, Ran Gazit, Joaquin Gonzalez-Rodriguez, Craig Greenberg, Ville Hautamaki, Hynek Hermansky, Chien-Lin Huang, Patrick Kenny, Tomi Kinnunen, Tina Kohler, Pietro Laface, Anthony Larcher, Kong Aik Lee, Cheung-Chi Leung, Haizhou Li, Bin Ma, Alvin Martin, John Mason, Driss Matrouf, Tomoko Matsu, Jiri Navratil, Raymond W. M. Ng, Jason Pelecanos, Daniel Ramos, Douglas Reynolds, Nicolas Scheffer, Andreas Stolcke, Doug Sturim, Hanwu Sun, Rong Tong, Pedro Torres-Carrasquillo, David van Leeuwen, Michael Wagner, Changhuai You
Welcome Message

Welcome to Odyssey 2012: The Speaker and Language Recognition Workshop, organized by the Chinese and Oriental Languages Information Processing Society (COLIPS), the ISCA Speaker and Language Characterization Special Interest Group (SpLC), and the Institute for Infocomm Research (I²R) in Singapore, on 25-28 June 2012

Odyssey 2012 received overwhelming response from the speaker and language recognition community. We accepted 51 papers out of 65 submissions, which we organized into a 4-day technical program consisting of 11 sessions. Researchers will present their latest endeavours and insights from multiple aspects, covering speaker and language characterization, modelling, evaluation, and applications. In addition, Odyssey 2012 also features 3 invited speakers: Dr. Niko Brümmer (Agnitio Corporation) will discuss how proper scoring rules adopted from weather prediction can be applied to pattern recognition to make probabilistic decisions; Dr. Li Deng (Microsoft Research) will share with us how new generation models such as deep belief networks and dynamic Bayesian networks can revamp the traditional framework of Gaussian mixture model and hidden Markov model in speech technology; and Dr. Alvin Martin (National Institute of Standards and Technology) will talk about the past, present and future of the NIST Speaker Recognition Evaluation Series.

With a unique history and geographical location, Singapore is one of the few places on earth to experience the extreme blending of diverse languages. In the streets and neighbourhoods of Singapore, you can easily hear people talking in a rich mix of languages ranging from Chinese varieties (e.g., Mandarin, Hokkien, Cantonese) to other Asian languages (e.g., Tamil, Malay), and to English (e.g., British and Singaporean dialects). It is our hope to celebrate the diversity of languages by hosting the Odyssey 2012 Workshop in Singapore. I therefore wish you an enjoyable stay with a multi-cultural experience!

Many volunteers have contributed significantly to organizing this workshop. I would like to take this opportunity to show my gratitude to the organizing committee for valuable advice, and the scientific committee for their thorough paper reviews. Great thanks also go to my colleagues at Institute for Infocomm Research, especially to: Bin Ma and Kong Aik Lee (technical program and website), Swee Lan See (finance), Nancy Chen and Minghui Dong (publicity and sponsorship), Anthony Larcher and Hanwu Sun (social activities), Cheung-Chi Leung (venue), Rong Tong (registration), Changhuai You (publication). I also appreciate the financial support from the Lee Foundation, Temasek Laboratories at Nanyang Technological University, and International Speech Communication Association (ISCA).

Last but not least, I sincerely thank the reviewers, authors, invited speakers, presenters, exhibitors, local supporting staff, and all the workshop attendees for your support and enthusiasm. It is you that makes this conference possible. Thank you!

Haizhou Li
Odyssey 2012 Chair
# Program Overview

**Monday 25 June 2012**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 - 9:30</td>
<td>Registration</td>
</tr>
<tr>
<td>9:30 - 10:00</td>
<td>Opening</td>
</tr>
<tr>
<td>10:00 - 11:00</td>
<td><strong>Plenary Session 1</strong>: <a href="#">The Role of Proper Scoring Rules in Training and Evaluating Probabilistic Speaker and Language Recognizers</a></td>
</tr>
<tr>
<td>11:00 - 11:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>11:30 - 13:10</td>
<td><strong>Session 1</strong>: <a href="#">Speaker Recognition – Compact Representation</a></td>
</tr>
<tr>
<td>13:10 - 14:00</td>
<td>Lunch</td>
</tr>
<tr>
<td>14:00 - 16:05</td>
<td><strong>Session 2</strong>: <a href="#">Speaker Recognition – Generative modeling</a></td>
</tr>
<tr>
<td>16:05 - 16:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>16:30 - 17:45</td>
<td><strong>Session 3</strong>: <a href="#">Forensic Speaker Recognition</a></td>
</tr>
<tr>
<td>17:45 - 22:00</td>
<td>Welcome Reception @ Tiger Brewery</td>
</tr>
</tbody>
</table>

**Tuesday 26 June 2012**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 - 10:00</td>
<td><strong>Plenary Session 2</strong>: <a href="#">Being Deep and Being Dynamic – New-Generation Models and Methodology for Advancing Speech Technology</a></td>
</tr>
<tr>
<td>10:00 - 10:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>10:30 - 12:35</td>
<td><strong>Session 4</strong>: <a href="#">Neural Network for Speaker Recognition</a></td>
</tr>
<tr>
<td>12:35 - 13:30</td>
<td>Lunch</td>
</tr>
<tr>
<td>13:30 - 15:35</td>
<td><strong>Session 5</strong>: <a href="#">Speaker Diarization</a></td>
</tr>
<tr>
<td>15:35 - 16:00</td>
<td>Coffee break</td>
</tr>
<tr>
<td>16:00 - 18:05</td>
<td><strong>Session 6</strong>: <a href="#">Speaker Recognition – Channel Robustness</a></td>
</tr>
</tbody>
</table>
**Wednesday 27 June 2012**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session/Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 - 10:00</td>
<td><strong>Plenary Session 3:</strong> The NIST Speaker Recognition Evaluations</td>
</tr>
<tr>
<td>10:00 - 10:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>10:30 - 12:35</td>
<td><strong>Session 7:</strong> Language Recognition Evaluation</td>
</tr>
<tr>
<td>12:35 - 13:30</td>
<td>Lunch</td>
</tr>
<tr>
<td>13:30 - 15:35</td>
<td><strong>Session 8:</strong> Features for Speaker Recognition</td>
</tr>
<tr>
<td>15:35 - 22:00</td>
<td>Social Event + Banquet @ Megu</td>
</tr>
</tbody>
</table>

**Thursday 28 June 2012**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session/Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:15 - 10:00</td>
<td>Coffee</td>
</tr>
<tr>
<td>10:00 - 12:05</td>
<td><strong>Session 9:</strong> Speaker Recognition Evaluation</td>
</tr>
<tr>
<td>12:05 - 13:00</td>
<td>Lunch</td>
</tr>
<tr>
<td>13:00 - 15:05</td>
<td><strong>Session 10:</strong> Speaker Recognition – Application</td>
</tr>
<tr>
<td>15:05 - 15:30</td>
<td>Coffee break</td>
</tr>
<tr>
<td>15:30 - 17:35</td>
<td><strong>Session 11:</strong> Language Recognition – Feature, Classifier and Fusion</td>
</tr>
</tbody>
</table>
Plenary Session 1: The Role of Proper Scoring Rules in Training and Evaluating Probabilistic Speaker and Language Recognizers

Speaker: Dr Niko Brümmer, Chief Scientist at AGNITIO Corp., South Africa

Session Chair: Douglas Reynolds, MIT Lincoln Laboratory, USA

It is obvious how to evaluate the goodness of a pattern classifier that outputs hard classification decisions --- you count the errors. But hard classification decisions are implicitly dependent on fixed priors and costs, so that they are applicable only in a narrow range of applications. A classifier can widen its range of applicability by outputting instead soft decisions, in the form of class probabilities or likelihoods. However, it is much less obvious how to evaluate the goodness of such probabilistic outputs. To evaluate the goodness of recognized classes, they can simply be compared to the true class labels in a supervised evaluation database. But we simply don't have a similar truth reference for probabilistic outputs.

A solution to this problem, originally from weather prediction, called "proper scoring rules", has been known for several decades, but has enjoyed only limited attention in pattern recognition and machine learning. This talk will explain how they work, how they generalize error-rate, how they measure information and how to use them for both training and evaluation of probabilistic pattern recognizers.

Short Biography

Niko Brummer received B.Eng (1986), M.Eng (1988) and Ph.D. (2010) degrees, all in electronic engineering, from Stellenbosch University. He worked as researcher at DataFusion (later called Spescom DataVoice) and is currently chief scientist at AGNITIO. Most of his research for the last two decades has been applied to automatic speaker and language recognition and he has been participating in most of the NIST SRE and LRE evaluations in these technologies, from the year 2000 to the present. He has been contributing to the Odyssey Workshop series since 2001 and was organizer of Odyssey 2008 in Stellenbosch. His FoCal Toolkit is widely used for fusion and calibration in speaker and language recognition research.

His research interests include development of new algorithms for speaker and language recognition, as well as evaluation methodologies for these technologies. In both cases, his emphasis is on probabilistic modeling. He has worked with both generative (eigenchannel, JFA, i-vector PLDA) and discriminative (system fusion, discriminative JFA and PLDA) recognizers. In evaluation, his focus is on judging the goodness of classifiers that produce probabilistic outputs in the form of well calibrated class likelihoods.
Semantic information embedded in the speech signal --- not only the phonetic/linguistic content but also a full range of paralinguistic information including speaker characteristics --- manifests itself in a dynamic process rooted in the deep linguistic hierarchy as an intrinsic part of the human cognitive system. Modeling both the dynamic process and the deep structure for advancing speech technology has been an active pursuit for over more than 20 years, but it is not until recently (since only a few years ago) that noticeable breakthrough has been achieved by the new methodology commonly referred to as "deep learning". Deep Belief Net (DBN) is recently being used to replace the Gaussian Mixture Model (GMM) component in HMM-based speech recognition, and has produced dramatic error rate reduction in both phone recognition and large vocabulary speech recognition while keeping the HMM component intact. On the other hand, the (constrained) Dynamic Bayesian Net (referred to as DBN* here) has been developed for many years to improve the dynamic models of speech while overcoming the IID assumption as a key weakness of the HMM, with a set of techniques and representations commonly known as hidden dynamic/trajectory models or articulatory-like models. A history of these two largely separate lines of "DBN/DBN*" research will be critically reviewed and analyzed in the context of modeling deep and dynamic linguistic hierarchy for advancing speech (as well as speaker) recognition technology. Future directions will be discussed for this exciting area of research that holds promise to build a foundation for the next-generation speech technology with human-like cognitive ability.

Short Biography: Li Deng received the Ph.D. from Univ. Wisconsin-Madison. He was an Assistant (1989-1992), Associate (1992-1996), and Full Professor (1996-1999) at the University of Waterloo, Ontario, Canada. He then joined Microsoft Research, Redmond, where he is currently a Principal Researcher and where he received Microsoft Research Technology Transfer, Goldstar, and Achievement Awards. Prior to MSR, he also worked or taught at Massachusetts Institute of Technology, ATR Interpreting Telecom. Research Lab. (Kyoto, Japan), and HKUST. He has published over 300 refereed papers in leading journals/conferences and 3 books covering broad areas of human language technology, machine learning, and audio, speech, and signal processing. He is a Fellow of the Acoustical Society of America, a Fellow of the IEEE, and a Fellow of the International Speech Communication Association. He is an inventor or co-inventor of over 50 granted US, Japanese, or international patents. He served on the Board of Governors of the IEEE Sig. Proc. Soc. (2008-2010). More recently, he served as Editor-in-Chief for IEEE Signal Processing Magazine (2009-2011), which, according to the Thompson Reuters Journal Citation Report released 2010 and 2011, ranked first in both years among all 127 IEEE publications and all 247 publications within the Electrical and Electronics Engineering Category worldwide in terms of its impact factor, and for which he received the 2011 IEEE SPS Meritorious Service Award. He currently serves as Editor-in-Chief for IEEE Transactions on Audio, Speech and Language Processing. His recent tutorials on deep learning at APSIPA (Oct 2011) and at ICASSP (March 2012) received the highest attendance rate at both conferences.
Plenary Session 3: The NIST Speaker Recognition Evaluations

Speaker: Dr Alvin Martin, Mathematician at National Institute of Standards and Technology, USA

Session Chair: Eliathamby Ambikairajah, University of New South Whales, Australia

Since 1996 the National Institute of Standards and Technologies has coordinated a series of annual or bi-annual open evaluations of automatic speaker recognition technology. These have concentrated on the task of single speaker detection in the context of spontaneous speech of a conversational telephone or one-on-one interview situation, recorded over ordinary telephone channels or room microphones. System performance has been assessed in relation to a variety of factors, including notably the quantity of training and test speech supplied, the speech styles being used, and the types and variability of the recording channels. While English has been the primary language employed, several of the evaluations have included substantial quantities of speech by multi-lingual speakers to allow examination of language and cross-language effects. More recently, initial efforts have been made to consider the effects of voice aging and varying vocal effort on performance. We discuss the considerations that have gone into planning and organizing these and a few related evaluations, the performance metrics that have been employed, the considerable progress observed over time, and the ongoing plans for further evaluation in 2012 and beyond.

Short Biography

Alvin Martin served as a mathematician in the Multimodal Information Group at the National Institute of Standards and Technology from 1991 through 2011. He has coordinated NIST’s series of evaluations since 1996 in the areas of speaker recognition and of language and dialect recognition, and has contributed to its evaluations of large vocabulary continuous speech recognition. This work has involved the collection, selection, and pre-processing of appropriate speech data, the writing of evaluation plans, the specification of metrics and charts for the scoring, presentation, and analysis of results, the implementation of statistical tests for determining the significance of performance differences, and the organization of workshops to review evaluation results.

He received a Ph.D. in mathematics from Yale University (1977), has taught mathematics and computer science at the college level, and worked on the development of automatic speech recognition and speech processing systems before coming to NIST.
Technical Program

Session 1: Speaker Recognition – Compact Representation
Monday 25 June 2012
Chair: Lukáš Burget, SRI, USA

11:30 - 11:55  A Small Footprint i-Vector Extractor  
Patrick Kenny

11:55 - 12:20  Memory and Computation Effective Approaches for i–Vector Extraction  
Sandro Cumani, Pietro Laface and Vasileios Vasilakakis

12:20 - 12:45  A Hybrid Factor Analysis and Probabilistic PCA-based system for Dictionary Learning and Encoding for Robust Speaker Recognition  
Srikanth Madikeri

12:45 - 13:10  On Exploring the Similarity and Fusion of i-Vector and Sparse Representation based Speaker Verification Systems  
Haris B C and Rohit Sinha

Session 2: Speaker Recognition – Generative modeling
Monday 25 June 2012
Chair: Patrick Kenny, CRIM, Canada

14:00 - 14:25  PLDA based Speaker Recognition on Short Utterances  
Ahilan Kanagasundaram, Robbie Vogt, David Dean and Sridha Sridharan

14:25 - 14:50  PLDA based Speaker Verification with Weighted LDA Techniques  
Ahilan Kanagasundaram, David Dean, Sridha Sridharan and Robbie Vogt

14:50 - 15:15  Dataset Shift in PLDA based Speaker Verification  
Carlos Vaquero

15:15 - 15:40  Bayesian Adaptation of PLDA Based Speaker Recognition to Domains with Scarce Development Data  
Jesus Villalba and Eduardo Lleida

15:40 - 16:05  Source Normalization for Language-Independent Speaker Recognition using i-Vectors  
Mitchell McLaren, Miranti Indar Mandasari and David A. van Leeuwen
Session 3: Forensic Speaker Recognition

Monday 25 June 2012

Chair: David van Leeuwen, Radboud University Nijmegen, the Netherlands

16:30 - 16:55 Database Selection for Forensic Voice Comparison
Geoffrey Stewart Morrison, Felipe Ochoa, Tharmarajah Thiruvaran

16:55 - 17:20 Voice Source Features for Forensic Voice Comparison - an Evaluation of the GLOTTEX Software Package
Ewald Enzinger, Cuiling Zhang and Geoffrey Stewart Morrison

17:20 - 17:45 Comparison of Speaker Recognition Systems on a Real Forensic Benchmark
Yosef Solewicz, Timo Becker, Jardine Gaelle and Stefan Gfroerer

Session 4: Neural Network for Speaker Recognition

Tuesday 26 June 2012

Chair: Hynek Hermansky, JHU, USA and BUT, Czech Republic

10:30 - 10:55 Factor Analysis of Mixture of Auto-Associative Neural Networks for Speaker Verification
Sivaram Garimella and Hynek Hermansky

10:55 - 11:20 Adaptation Transforms of Auto-Associative Neural Networks as Features for Speaker Verification
Samuel Thomas, Sri Harish Mallidi, Sriram Ganapathy and Hynek Hermansky

11:20 - 11:45 Bottleneck Features for Speaker Recognition
Sibel Yaman, Jason Pelecanos and Ruhi Sarikaya

11:45 - 12:10 Preliminary Investigation of Boltzmann Machine Classifiers for Speaker Recognition
Themos Stafylakis, Patrick Kenny, Mohammed Senoussaoui and Pierre Dumouchel

12:10 - 12:35 First attempt of Boltzmann Machines for Speaker Verification
Mohammed Senoussaoui, Najim Dehak, Patrick Kenny, Rêda Dehak and Pierre Dumouchel
Session 5: Speaker Diarization
Tuesday 26 June 2012
Chair: Andreas Stolcke, Microsoft Speech Labs, USA
13:30 - 13:55   Online Two Speaker Diarization
               Hagai Aronowitz, Yosef Solewicz and Orith Toledo-Ronen
13:55 - 14:20   On the use of Agglomerative and Spectral Clustering in Speaker Diarization of Meetings
               Jordi Luque and Javier Hernando
14:20 - 14:45   Generalized Viterbi-based Models for Time-Series Segmentation Applied to Speaker Diarization
               Itshak Lapidot and Jean-Francois Bonastre
14:45 - 15:10   A Global Optimization Framework For Speaker Diarization
               Mickael Rouvier and Sylvain Meignier
15:10 - 15:35   Cisco’s Speaker Segmentation and Recognition System
               Sachin Kajarekar, Aparna Khare, Matthias Paulik, Neha Agrawal, Panchi Panchapagesan, Ananth Sankar and Satish Gannu

Session 6: Speaker Recognition – Channel Robustness
Tuesday 26 June 2012
Chair: Honza Cernocky, BUT, Czech Republic
16:00 - 16:25   Variance-Spectra based Normalization for I-vector Standard and Probabilistic Linear Discriminant Analysis
               Pierre-Michel Bousquet, Anthony Larcher, Driss Matrouf, Jean-Francois Bonastre and Oldrich Plchot
16:25 - 16:50   Utterance Partitioning with Acoustic Vector Resampling for I-Vector based Speaker Verification
               Wei Rao and Man-Wai Mak
16:50 - 17:15   Study on the Effects of Intrinsic Variation using i-Vectors in Text-Independent Speaker Verification
               Sheng Chen, Mingxing Xu and Emlyn Pratt
17:15 - 17:40 Exploring the Impact of Advanced Front-End Processing on NIST Speaker Recognition Microphone Tasks
William Campbell, Doug Sturim, Jonas Borgstrom, Robert Dunn, Alan McCree, Tom Quatieri and Doug Reynolds

17:40 - 18:05 Linear Prediction Modulation Filtering for Speaker Recognition of Reverberant Speech
Bengt Borgstrom and Alan McCree

Session 7: Language Recognition Evaluation
Wednesday 27 June 2012
Chair: Tomi Kinnunen, U. of East Finland, Finland

10:30 - 10:55 Evaluation of Spoken Language Recognition Technology Using Broadcast Speech: Performance and Challenges
Luis J. Rodriguez-Fuentes, Amparo Varona, Mireia Diez, Mikel Penagarikano and German Bordel

Stephanie Strassel, Kevin Walker, Karen Jones, Dave Graff and Christopher Cieri

11:20 - 11:45 The MITLL NIST LRE 2011 Language Recognition System
Elliot Singer, Pedro Torres-Carrasquillo, Douglas Reynolds, Alan McCree, Fred Richardson, Najim Dehak and Doug Sturim

11:45 - 12:10 Description and analysis of the Brno276 system for LRE2011
Niko Brummer, Sandro Cumani, Ondrej Glembek, Martin Karafiat, Pavel Matejka, Jan Pesan, Oldrich Plchot, Mehdi Soufifar, Edward de Villiers and Jan Cernocky

12:10 - 12:35 A Linguistic Data Acquisition Front-End for Language Recognition Evaluation
Gang Liu, Chi Zhang and John Hansen

Session 8: Features for Speaker Recognition
Wednesday 27 June 2012
Chair: John Hansen, University of Texas at Dallas, USA

13:30 - 13:55 Feature Extraction Using 2-D Autoregressive Models For Speaker Recognition
Sriram Ganapathy, Samuel Thomas and Hynek Hermansky
13:55 - 14:20 Regularization of All-Pole Models for Speaker Verification Under Additive Noise
Cemal Hanilci, Tomi Kinnunen, Rahim Saeidi, Jouni Pohjalainen, Paavo Alku and Figen Ertas

14:20 - 14:45 Factor Analysis of Acoustic Features using a Mixture of Probabilistic Principal Component Analyzers for robust Speaker Verification
Taufiq Hasan and John Hansen

14:45 - 15:10 Exemplar-based Sparse Representation and Sparse Discrimination for Noise Robust Speaker Identification
Rahim Saeidi, Antti Hurmalainen, Tuomas Virtanen and David A. van Leeuwen

15:10 - 15:35 On the use of Asymmetric-shaped Tapers for Speaker Verification using l-vectors
Md Jahangir Alam, Patrick Kenny and Douglas O'Shaughnessy

Session 9: Speaker Recognition Evaluation
Thursday 28 June 2012
Chair: William Campbell, MIT Lincoln Lab, USA

10:00 - 10:25 The Effect of Target/Non-Target Age Difference on Speaker Recognition Performance
George Doddington

10:25 - 10:50 Variational Bayes Logistic Regression as Regularized Fusion for NIST SRE 2010
Ville Hautamäki, Kong Aik Lee, Anthony Larcher, Tomi Kinnunen, Bin Ma and Haizhou Li

10:50 - 11:15 The 2011 BEST Speaker Recognition Interim Assessment
Craig Greenberg, Alvin Martin and Mark Przybocki

11:15 - 11:40 The REPERE Challenge: finding people in a multimodal context
Juliette Kahn, Olivier Galibert, Matthieu Carré, Aude Giraudel, Philippe Joly and Ludovic Quintard

11:40 - 12:05 The RATS Radio Traffic Collection System
Kevin Walker and Stephanie Strassel
Session 10: Speaker Recognition – Application  
Thursday 28 June 2012  
Chair: Man-Wai Mak, Hong Kong Polytechnic University, Hong Kong  

13:00 - 13:25  
Effects of Audio and ASR Quality on Cepstral and High-level Speaker Verification Systems  
Andreas Stolcke, Martin Graciarena and Luciana Ferrer

13:25 - 13:50  
Audio Context Recognition in Variable Mobile Environments from Short Segments using Speaker and Language Recognizers  
Tomi Kinnunen, Rahim Saeidi, Jussi Leppaanen and Jukka P. Saarinen

13:50 - 14:15  
Text Dependent Speaker Verification Using a Small Development Set  
Hagai Aronowitz

14:15 - 14:40  
A Unified Approach for Audio Characterization and its Application to Speaker Recognition  
Luciana Ferrer, Lukas Burget, Oldrich Plchot and Nicolas Scheffer

14:40 - 15:05  
Mean Shift Algorithm for Exponential Families with Applications to Speaker Clustering  
Themos Stafylakis, Vassilis Katsouros, Patrick Kenny and Pierre Dumouchel

Session 11: Language Recognition – Feature, Classifier and Fusion  
Thursday 28 June 2012  
Chair: Michael Wagner, University of Canberra, Australia  

15:30 - 15:55  
Speaker Vectors from Subspace Gaussian Mixture Model as Complementary Features for Language Identification  
Oldrich Plchot, Martin Karafiat, Niko Brummer, Ondrej Glembek, Pavel Matejka, Edward de Villiers and Jan Cernocky

15:55 - 16:20  
Complementary Combination in i-Vector Level for Language Recognition  
Zhi-Yi Li, Wei-Qiang Zhang, Liang He and Jia Liu

16:20 - 16:45  
Bhattacharyya-based GMM-SVM System with Adaptive Relevance Factor for Pair Language Recognition  
Changhuai You, Haizhou Li, Eliathamby Ambikairajah, Kong Aik Lee and Bin Ma

16:45 - 17:10  
Fusing Language Information from Diverse Data Sources for Phonotactic Language Recognition  
Mohamed Fauzi BenZeghiba, Jean-Luc Gauvain and Lori Lamel
**Social Program**

**Welcome Reception**  
**Monday 25 June 2012, 17:45 – 22:00**

The welcome reception will take place at TAVERN @ THE TIGER BREWERY, immediately after the last technical session on Monday. After a short tour of the Tiger Brewery, there will be a buffet dinner with a tasting session of 7 types of freshly-brewed beer. Karaoke, pool table and dart board will also be available.

The Welcome Reception is included as part of the Odyssey 2012 registration. On the way back, buses are chartered to reach four MRT stations (Buona Vista, Outram Park, City Hall and Bugis). There are two departures from the brewery, the first at 21:00 and the second at 22:00.

**Odyssey 2012 Social Event – Singapore Marina Sightseeing & Banquet**  
**Wednesday 27 June 2012, 15:35 - 22:00**

The social event will take place at Singapore central area. The event will begin right after the last technical session on Wednesday. Buses will take you to the Singapore River side, from where you will board the boats to cruise along the Singapore River all the way to the Singapore Flyer – the world largest observation wheel. After enjoying the panoramic view over the Marina Bay, dinner will be served at the Megu Restaurant located next to the Flyer. There will also be live Jazz music played by Fuchsia. The tentative schedule is as follows:

15:35   End of technical session
16:00   Depart from the Matrix building at Biopolis
16:30   Arrive at the Merlion Esplanade
17:00   Board Singapore River cruise boats
17:45   Arrival at Singapore Flyer, ride the giant observation wheel
18:30   Welcome drink at Megu restaurant (within the Flyer's building), enjoy the view at the Marina Bay
19:00   Banquet at the Megu Restaurant
22:00   Depart from the Megu Restaurant, buses are chartered to reach four MRT stations (Bugis, City Hall, Outram Park and Buona Vista)

The social event is included as part of the Odyssey 2012 registration. Accompanying persons are welcomed to attend the banquet at SGD 150.00 per person.
Dining Places near Biopolis

**Biopolis Food Court**
1st level, Matrix Building  
*Besides Singaporean food, there is also Chinese, Japanese, Indonesian, and Western food options.*

**Khansama Tandoori Castle**
http://www.khansama.net/default.aspx  
8A Biomedical Grove, #01-20, 21, 22 Immunos building, Biopolis Way  
(+65) 6299 0300; (+65) 6774 0300  
*Indian Buffet*

**The Lawn Grill and Salad Cafe**
http://www.thelawn.com.sg/  
#01-07 Nanos Building, 31 Biopolis Way  
(+65) 6478 9739  
*Serves western style breakfast. For lunch/dinner: salad with sizzling hot grilled meat/seafood/mushrooms.*

**Raj**
http://www.rajrestaurant.com.sg/  
#01-03 Centros Block, 20 Biopolis Way  
(+65) 6478-9495  
*Vegetarian Indian Restaurant*
About Singapore

Singapore is a Southeast Asian city-state island off the southern tip of the Malay Peninsula. Located at the crossroads of the East and West, Singapore is a cultural and economic hub: a harmonious society created by people of diverse races, religions, and languages. Singapore is a tourist hotspot, offering a blend of Malay, Chinese and Indian delicacies, futuristic architecture, sunny get-away resorts, vibrant shopping malls, and dazzling nightlife.

Below we list some practical information and travel recommendations (nature, culture, and food) of how you can discover the city.

**Emergency Phone Numbers**

Singapore Country Code: 65

Police: 999 (toll-free)

Emergencies/Ambulance/Fire: 995

Non-emergency Ambulance: 1777

**Transportation**

People here refer to the subway as MRT (Mass Rapid Transit), which is the major publication transportation system. You can also call for taxis:

Comfort and City Cab: +65-6552-1111
Comfort and City Premier Cabs: +65-6552-2828
Prime Taxi: +65-6778-0808
Smart Cab: +65-6485-7777
Yellow-Top Taxi: +65-6293-5545
**NATURE**

**Singapore Zoo**
http://www.zoo.com.sg
Address: 80 Mandai Lake Rd
Highlights: Elephant show, Night Safari, white tiger, gazillion different types of monkeys!

**Singapore Botanical Gardens**
http://www.sbg.org.sg
Address: 1 Cluny Rd, Singapore Botanic Gardens
Getting there: Botanic Garden MRT Station (Circle Line)
Highlight: National Orchid Garden

**MacRitchie Reservoir Park**
Getting there: Marymount MRT (Circle Line) then buses 855,52,167 (2 stops)
Highlight: multiple hiking trails; monkeys are easily seen around the area (get your cameras ready!)

**Sentosa Island Resort**
http://www.sentosa.com.sg
Address: 40 Imbiah Rd
Highlights: sunny beaches, adventure rides, British colonial military remnants

**CULTURE**

**Asian Civilisations Museum**
http://www.acm.org.sg/home/home.asp
Address: 1 Empress Pl Asian Civilisations Museum
Getting there: Raffles Place MRT Station (North South Line/East West Line)
*Discover how Southeast Asian culture is influenced by East Asia (e.g., China), Southern Asia (e.g., India) and the Middle East.*

**Peranakan Museum**
http://www.peranakanmuseum.sg
Address: 39 Armenian St
Getting there: City Hall MRT Station (North South Line/East West Line); Bras Basah MRT Station (Circle Line)
*Peranakan Chinese or Baba-Nyonya are the descendants of the 16th-century Chinese immigrants to the Indonesian archipelago during the Colonial era.*
**FOOD & DRINKS**

Due to Singapore’s geographic advantage and historical heritage, the Lion City has become a paradise for foodies. Throughout the centuries, Singapore cuisine has integrated recipes from many diverse cultures including Chinese, Indian, Indonesian, and Malaysian, creating its own uniquely Singapore flavor. In addition, Singapore also offers authentic ethnic food from Asia (e.g., Japan) and Europe (e.g., French).

In Singapore, food is everywhere! Wherever you turn, there is a place to eat. To prevent you from getting lost amongst the countless options, as local committee members, we wrote this section to help you find your way.

Although Geylang is Singapore's red light district, it is also well known for its plentiful gourmet food options. Along Geylang Road, there are numerous east Asian restaurants. Danny Lee of Sin Huant Eating House is known as the “food Nazi of Geylang”. Despite his snobbish demeanor and quick temper, his famous *crab bee hoon* is to die for, according to Anthony Bourdain’s list of the 13 places to eat before you die. Yong He Eating House is known for its Taiwanese snacks such as soya bean milk, you tiao (deep fried dough with fritters), tau huay (sweet soya bean curd), and pork-floss filled sticky rice balls. Like many other shops around the area, it’s open 24 hours, making it great for after-party comfort food.  
*Getting there: Aljunied MRT (East West Line)*

**Little India** is the best place to experience Singapore’s Tamil heritage. The crowded streets are filled with lively music and food stalls that open till late night. You can find Indian restaurants and canteens along the entire way from Little India MRT to Farrer Park MRT (North East Line). Choosing what to eat among the rich variety of North and South Indian cuisine is not easy, but you will never be disappointed.  
*Getting there: Little India / Farrer Park MRT (North East Line) [Around C8 on Map]*

For shopaholics, **Orchard Road** is the place to be! If strolling down Orchard Road under the blazing sun is too hot for you, enjoy the air-conditioned extensive network of underpasses connecting the shopping malls. Each shopping mall also homes a number of good restaurants and food courts.  
*Getting there: Orchard/Somerset MRT (North South Line) [I11 on Map]*

**Boat Quay** offers a variety of restaurants mixing Chinese, Western, Indian or Thai cuisine, where the hip Singaporeans like to hang out.  
*Getting there: Raffles City MRT (North South Line) [N6 on Map]*

Want to party the night away? Head to **Clarke Quay**, the centre of Singapore’s nightlife! Clarke Quay offers a wide range of bars, restaurants, pubs, cafes and dance clubs. Have a drink while enjoying the colorfully renovated shop-houses along the Singapore River. Looking for something funkier? How about the Clinic Bar? Tasty cocktails are served in test tubes and IV bags, while you tour around the labyrinthine of pill-shaped rooms in your wheel chair!  
*Getting there: Clarke Quay MRT (North East Line) [M8 on Map]*

**Chijmes** is a Catholic Covent converted into restaurants and bars, a good place to relax in a cozy atmosphere. (Address: 30 Victoria Street).  
*Getting there: City Hall MRT (North South Line/ East West Line) [I6 on Map]*
Kopitiam (Coffee Shop)

*kopi* is Malay for *coffee* (borrowed from Portuguese) and *tiam* is the Hokkien word for *shop*. Also referred to as *hawker centers* and *food courts*.

**Popular Dishes** (~ $3-6/dish)

- **Hainanese Chicken Rice**: National dish of Singapore
- **Bak Kut Teh** (*meat bone tea* in Hokkien): meaty pork ribs simmered in a broth of herbs and spices
- **(Roti) Prata**: Northern Indian pancake served with vegetable or meat based curry.
- **Kaya toast + soft-boiled eggs + tea/coffee**: The classic Singaporean breakfast. *Kaya* is coconut egg jam flavored by pandan leaf.
- **Curry puff**: deep-fried pastry filled with curry, chicken, and potatoes. Found in food stalls such as Old Chang Kee (老曾記). A great snack when you are on the go.

**Newton Food Center**
Must try: BBQ Seafood, Fried Carrot Cake, Oyster Omelette
24 hours outside venue very lively
*Getting there: Newton MRT (North South Line) [E17 on Map]*

**East Coast Lagoon Food Village**
Located on the South East part of Singapore by the sea
must try: Oyster Omelet, Beef Noodles, Zi Char
*Getting there: taxi*

**Golden Mile Complex**
*Address: 505, Beach Road*  
*Golden Mile Complex (don't confuse with the Golden Mile Food Center) offers a simulated trip to Thailand at a low price.*
*Must try: green mango salad, all Thai dishes*
*Getting there: Lavender MRT (East West Line) then bus 100 towards Ghim Moh Ter (2 stops) [C1 on Map]*

**Lao Pa Sat** (literally means *old market* in Hokkien)
Must try: Satay (marinated skewered grilled meat of chicken, lamb, beef or prawns)
*Getting there: City Hall MRT (North South Line/ East West Line) exit i [Q4 on Map]*

**People’s Park Food Center**
Located right in the heart of the bustling Chinatown at New Market Road Blk 32
Must try: Frog Clay Pot
*Getting there: Chinatown MRT (North East Line) [Q9 on Map]*

**Singapore Flyer Food Center**
*Must try: Oyster Omelet, Char Kway Teoh, Carrot Cake, Fried Noodles*
Getting there: Promenade MRT (Circle Line); the food center is located straight below the Flyer [J1 on Map]

**Hong Lim Food Center**
*Address: 531A Upper Cross Street*
*Getting there: Chinatown MRT (North East Line) [P7 on Map]*
Zi Char Restaurants

Zi Char 煮炒 (literally means cook-fry in Hokkien). Unlike the hawker centers where a lot of the dishes are pre-prepared, zi char style cooks every dish fresh on the spot. Typically each dish costs around $8-15 SGD.

**Popular Dishes:**
- Chili Crab
- Kang Kong Blachan (fried water spinach with spicy dried shrimp paste)

**Sin Huat Eating House (新發海鮮館)**
Must try: crab bee hoon (spicy crab rice noodles) is 1 of the 13 places to eat before you die (Note: it is pricy: ~$60 SGD)
([http://www.menshealth.com/nutrition/must-visit-restaurants](http://www.menshealth.com/nutrition/must-visit-restaurants))
Address: 659 Geylang Road
Tel: (+65) 6748 9495
*Getting there:* Aljunied MRT Station (*East West Line*)

**Fatty Wengs**
([http://food.ingsing.com/business/fatty-weng-restaurant/geylang-aljunied/id-065a0200](http://food.ingsing.com/business/fatty-weng-restaurant/geylang-aljunied/id-065a0200))
Address: 94 Guillemard Road
Tel: (+65) 6345-2702
*Getting there:* Aljunied MRT Station (*East West Line*)
Chinese Food

**Putien**
Address: 127 Kitchener Road
Tel: (+65) 6295 6358
*Putien traces its roots back to Putian, a beautiful coastal town located in Fujian, China. (Many Chinese immigrants (including those of Singapore) were originally from Fujian.) Putien presents delightful surprises by combining traditional Chinese Putian dishes with a touch of Southeast Asian flavors and spices.*
*NOT to be confused with “Putian Seafood”*
*Getting There: Farrer Park MRT (North East Line)*

**Paradise Dynasty**
Address: 2 Orchard Turn, #04-12A ION Orchard
Tel: (+65) 6509-9118
*Signature Dynasty Xiao Long Bao features 8 types of Shanghainese soup dumpling, ranging from herbal ginseng, to garlic, to black truffles, to spicy Szechuan. They also serve silky handmade Japanese noodles (La Mian).*
*Getting there: Orchard MRT (North South Line) [J18 on Map]*

**Prima Tower Revolving Restaurant**
Address: 201 Keppel Road
*From level 9 of this flour mill you got an amazing view over Singapore’s harbor must try: Peking duck*
*Getting there: Harbour Front MRT (Circle Line/ North East Line) [V11 on Map]*

**Dim Sum** (from breakfast to lunch time)

**Red Star**
Address: 54 Chin Swee Road
*Take the lift in this parking bock until you reach the 7th floor
*Authentic Hong Kong style Dim Sum with carts loaded with Cantonese dishes.*
*Getting there: Chinatown MRT (North East Line) [O11 on Map]*
**Cathay Restaurant**
Address: 2 Handy road
Hong Kong style Dim Sum
Must try: wasabi fish and nai huang bao (奶黄包), steamed buns filled with creamy custard made of milk and egg yolk.
*Getting there: Dhoby Ghaut MRT (North East Line/North South Line/Circle Line)*

**Peranakan Food**
*True Blue Cuisine*
http://www.truebluecuisine.com
Address: 49 Armenian St
*Getting there: City Hall MRT (North East Line/North South Line) [K8 on Map]*

**Indian Food**
*Sakunthala’s*
Address: 151 Dunlop Street
A very popular place where you can try various types of chapati, prata and other Indian dishes
*Getting there: Little India MRT (North East Line) [D9 on Map]*

*Sankranti*
Address: 100 Syed Alwi Rd (in Front of the Mustafa Center)
must try: bindi masala, butter chicken, kebab
*Getting there: Farrer Park MRT (North East Line) [B8 on Map]*

*Lagnaa Barefoot*
http://www.lagnaacom
Address: 6 Upper Dickson rd
Comfortable place with great choice of Indian dishes
Must try: all but ask for the spicy level that suits you!
*Getting there: Little India MRT (North East Line) [D9 on Map]*
**Western food**

**Russian:** Buyan

[http://buyan.sg](http://buyan.sg)

Address: 9/10 Duxton Hill

*Getting there: Outram Park MRT* *(East West Line/North East Line)* [S8 on Map]

**French – L’angelus**

Address: 85 Club Street

*Getting there: Chinatown MRT* *(North East Line)* [Q7 on Map]

**Italian – Da Paolo**

Address: 80 club Street

*Getting there: Chinatown MRT* *(North East Line)* [Q7 on Map]
Where to have a drink

Having a Tiger beer or a coconut milkshake in a Coffee Shop is a popular yet inexpensive pastime. (Bars/pubs are relatively costly in Singapore.)

Screening Room
Address: 12 Ann Siang Road
*Rooftop bar from which you can appreciate a view of the central business district and the Buddha Tooth Relic Temple.*
*Getting there: Chinatown MRT* (*North East Line*) [R7 on Map]

Zsofi
Address: 68 Dunlop Street
*Rooftop bar. One free tapas for every drink, large variety of beers (including Belgium ones).*
*Getting there: Little India MRT* (*North East Line*) [E8 on Map]

New Asia
Raffles City, access through the Swiss Hotel
*Located at the 71th floor of the Stamford Tower, the bar offers an amazing view over the island and the Marina Bay.*
*Getting there: City Hall MRT* (*North South Line/East West Line*) [J6 on Map]

One Altitude
Address: 1 Raffles Place
*One of the world's highest roof top bars located in the middle of Singapore's central business district.*
*Getting there: Raffles Place MRT* (*North South Line/East West Line*) [O4 on Map]

Kudeta
Address: 1 Bayfront Avenue, Marina Bay Sands Sky Garden
*Located at the top of the emblematic Marina Bay Sands building.*
*Getting there: Bayfront MRT* (*Circle Line*) [N1 on Map]

Bali Lane
*This small street near the Kampung Glam homes cozy shisha bars and live music.*
*Getting there: Bugis MRT* (*East West Line*) [F4 on Map]

Circular Road
*Just behind Boat Quay; this street lights up when offices close and goes to sleep at dawn.*
*Getting there: Raffles Place MRT* (*North South Line/East West Line*) [O6 on Map]

Emerald Hill Road
*The glittering nightlife of Orchard Road*
*Getting there: Somerset MRT* (*North South Line*) [I15 on Map]
Getting to Biopolis, Matrix Building

By MRT (metro/subway)

Alight at Buona Vista MRT station (EW 21/CC 22) and use Exit B or Exit D (B1 CCL Concourse) at the MRT station heading towards Ministry of Education/Biopolis, walk towards North Buona Vista Road. It is approximately a 10 min walk to Matrix building from the exit at the Ministry of Education.

By Bus

- Bus Services along North Buona Vista Road (Bus Stops 1, 2 and 3): 74, 91, 92, 95, 191, 196, 198, and 200.
- Bus Services along Commonwealth Ave West heading towards Dover (Bus Stop 4): 74, 91, 95, 100, 105, 111, 147, 191, 196, 198 and 200.
- Bus Services along Commonwealth Ave West heading towards Commonwealth (Bus Stop 5): 32, 74, 91, 95, 100, 105, 111, 145, 147, 191, 196, 198 and 200.

By Taxi

Taxis can be hailed at various taxi-stands, along major roads as well as most hotels’ main entrances. All fares are charged according to the taxi meters and are based on a flag down rate and the distance travelled. Additional surcharges will apply during the morning and evening peak hour times.

- Midnight surcharge at 50% of final metered fare from 12.00am to 6.00am (daily)
- Peak hour surcharge at 25% of final metered fare:
  - Monday – Friday: 6.00am to 9.30am
  - Monday – Sunday & Public Holidays: 6.00pm to 11.59pm

One specific pointer that you might give to the taxi driver: "Biopolis is right behind the Ministry of Education (MOE), along the South Buona Vista Road."
Singapore Mass Rapid Transport (MRT) Map