Odyssey 2010: The Speaker and Language Recognition Workshop

Brno, Czech Republic
28 June – 1 July 2010

# TABLE OF CONTENTS

## SESSION 1: SPEAKER RECOGNITION – LVCSR AND HIGH LEVEL FEATURES

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained Subword Units for Speaker Recognition</td>
<td>1</td>
</tr>
<tr>
<td>Computationally Efficient Speaker Identification for Large Population Tasks using MLLR and Sufficient Statistics</td>
<td>7</td>
</tr>
<tr>
<td>Modeling Prosody for Speaker Recognition: Why Estimating Pitch May Be a Red Herring</td>
<td>12</td>
</tr>
</tbody>
</table>

Doris Baum, Daniel Schneider, Timo Mertens, Joachim Köhler

Achintya Kumar Sarkar, S. Umesh, Shakti Prasad Rath

## SESSION 2: FEATURES FOR SPEAKER RECOGNITION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectionist Transformation Network Features for Speaker Recognition</td>
<td>20</td>
</tr>
<tr>
<td>An i-vector Extractor Suitable for Speaker Recognition with both Microphone and Telephone Speech</td>
<td>28</td>
</tr>
<tr>
<td>Investigation of Spectral Centroid Magnitude and Frequency for Speaker Recognition</td>
<td>34</td>
</tr>
<tr>
<td>Temporally Weighted Linear Prediction Features for Speaker Verification in Additive Noise</td>
<td>40</td>
</tr>
</tbody>
</table>

Alberto Abad, Jordi Luque

Mohammed Senoussaoui, Patrick Kenny, Najim Dehak, Pierre Dumouchel

Jia Min Karen Kua, Tharmarajah Thiruvanan, Mohaddeseh Nosrati, Eliathamby Ambikairajah, Julien Epps

## SESSION 3: BACKGROUND MODELING IN SPEAKER RECOGNITION, FORENSICS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Background Models for Speaker Verification</td>
<td>47</td>
</tr>
<tr>
<td>Training Universal Background Models for Speaker Recognition</td>
<td>52</td>
</tr>
<tr>
<td>SPES: The BKA Forensic Automatic Voice Comparison System</td>
<td>58</td>
</tr>
<tr>
<td>Estimating the Precision of the Likelihood-Ratio Output of a Forensic-Voice-Comparison System</td>
<td>63</td>
</tr>
<tr>
<td>Investigation of Speaker-Clustered UBMs based on Vocal Tract Lengths and MLLR Matrices for Speaker Verification</td>
<td>71</td>
</tr>
</tbody>
</table>

Wei-Qiang Zhang, Yuxiang Shan, Jia Liu

Mohamed Omar, Jason Pelecanos

Timo Becker, Michael Jessen, Sebastian Alsbach, Franz Broß, Torsten Meier

Geoffrey Stewart Morrison, Tharmarajah Thiruvanan, Julien Epps

Achintya Kumar Sarkar, S. Umesh

## SESSION 4: SPEAKER AND LANGUAGE RECOGNITION – SCORING, CONFIDENCES AND CALIBRATION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosine Similarity Scoring without Score Normalization Techniques</td>
<td>79</td>
</tr>
<tr>
<td>Unsupervised Speaker Adaptation based on the Cosine Similarity for Text-Independent Speaker Verification</td>
<td>84</td>
</tr>
<tr>
<td>Experiments in SVM-Based Speaker Verification Using Short Utterances</td>
<td>91</td>
</tr>
<tr>
<td>Detection Target Dependent Score Calibration for Language Recognition</td>
<td>99</td>
</tr>
</tbody>
</table>

Najim Dehak, Reda Dehak, James Glass, Douglas Reynolds, Patrick Kenny

Stephen Sham, Najim Dehak, Reda Dehak, James Glass

Mitchell McLaren, Robbie Vogt, Brendan Baker, Sridha Sridharan

Raymond W. M. Ng, Cheung Chi Leung, Tan Lee, Bin Ma, Haizhou Li

## SESSION 5: SPEAKER RECOGNITION – INTER-SESSION VARIABILITY

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted Nuisance Attribute Projection</td>
<td>105</td>
</tr>
<tr>
<td>On the Use of Factor Analysis with Restricted Target Data in Speaker Verification</td>
<td>111</td>
</tr>
</tbody>
</table>

William Campbell

Javier Gonzalez-Dominguez, Brendan Baker, Robbie Vogt, Joaquin Gonzalez-Rodriguez, Sridha Sridharan
Intra-speaker Variability Effects on Speaker Verification Performance .................................................. 117
Juliette Kahn, Nicolas Audibert, Solange Rossato, Jean-François Bonastre

Joint Factor Analysis for Speaker Recognition Reinterpreted as Signal Coding Using Overcomplete
Dictionaries ............................................................................................................................................... 125
Daniel García-Romero, Carol Y. Espy-Wilson

SESSION 6: DIARIZATION

Online Diarization of Telephone Conversations .................................................................................. 133
Oshry Ben-Harush, Itshak Lapidot, Hugo Guterman

Factor Analysis-based Approaches Applied to the Speaker Diarization Task of Meetings: A
Preliminary Study .................................................................................................................................. 139
Pavel Tomasek, Corinne Fredouille, Driss Matrouf

Unsupervised Compensation of Intra-Session Intra-Speaker Variability for Speaker Diarization .......... 146
Hagai Aronowitz

On the Use of GSV-SVM for Speaker Diarization and Tracking ......................................................... 154
Viet Bac Le, Claude Barras, Marc Ferras

SESSION 7: SPEAKER AND LANGUAGE RECOGNITION - EVALUATIONS AND
PERFORMANCE TESTING

Evaluation of the Vulnerability of Speaker Verification to Synthetic Speech ........................................ 159
Phillip DeLeon, Michael Pucher, Junichi Yamagishi

Hunting for Wolves in Speaker Recognition ....................................................................................... 167
Lara Stoll, George Doddington

The 2009 NIST Language Recognition Evaluation .............................................................................. 173
Alvin Martin, Craig Greenberg

The Albayzin 2008 Language Recognition Evaluation ........................................................................ 180
Luis Javier Rodríguez-Fuentes, Mikel Penagarikano, German Bordel, Amparo Varona

SESSION 8: HUMAN PERFORMANCES IN SPEAKER RECOGNITION, SPEAKER
CLUSTERING AND PARTITIONING

Human Assisted Speaker Recognition In NIST SRE10 ......................................................................... 188
Craig Greenberg, Alvin Martin, Linda Brandschlain, Joseph Campbell, Christopher Cieri, George Doddington,
John Godfrey

Speaker Clustering Via the Mean Shift Algorithm .............................................................................. 194
Themos Stafylakis, Vassilis Katsouras, George Carayannis

The Speaker Partitioning Problem ...................................................................................................... 202
Niko Brümmer, Edward de Villiers

Speaker Linking in Large Data Sets ..................................................................................................... 210
David van Leeuwen

SESSION 9: LANGUAGE RECOGNITION – GENERAL AND DATA

Estimating and Exploiting Language Distributions of Unlabeled Data ................................................ 217
Alan McCree

Data Selection and Calibration Issues in Automatic Language Recognition - Investigation with BUT-
AGNITIO NIST LRE 2009 System ........................................................................................................ 223
Zdenek Jancik, Oldrich Pichot, Niko Brümmer, Lukas Burget, Ondrej Glembek, Valiantsina Hubeika, Martin
Karafiat, Pavel Matejka, Tomas Mikolov, Albert Strasheim, Jan “Honzá” Cernocký

Comparison of Large-Scale SVM Training Algorithms for Language Recognition .......................... 230
Sandro Cumani, Fabio Castaldo, Pietro Laface, Daniele Colibro, Claudio Vair

Coping with Two Different Transmission Channels in Language Recognition .................................... 238
Florian Verdet, Driss Matrouf, Jean-François Bonastre, Jean Hennebert
## SESSION 10: LANGUAGE RECOGNITION – PHONOTACTICS

**Improved Modeling of Cross-Decoder Phone Co-occurrences in SVM-Based Phonotactic Language Recognition**

Mikel Penagarikano, Amparo Varona, Luis Javier Rodríguez-Fuentes, German Bordel

Page: 246

**Parallel Acoustic Model Adaptation for Improving Phonotactic Language Recognition**

Cheung Chi Leung, Bin Ma, Haizhou Li

Page: 254

**PCA-based Feature Extraction for Phonotactic Language Recognition**

Tomáš Mikolov, Oldrich Plichot, Ondrej Glembek, Lukáš Burget, Jan “Honza” Cernocký

Page: 259

**Improving Language Recognition with Multilingual Phone Recognition and Speaker Adaptation Transforms**

Andreas Stolcke, Murat Akbacak, Luciana Ferrer, Sachin Kajarekar, Colleen Richey, Nicolas Scheffer, Elizabeth Shriberg

Page: 264

## SESSION 11: DIALECT RECOGNITION

**Discriminative Phonotactics for Dialect Recognition Using Context-Dependent Phone Classifiers**

Fadi Biadsy, Hagen Soltau, Lidia Mangu, Jiri Navratil, Julia Hirschberg

Page: 271

**Suprasegmental Acoustic Cues of Foreignness in Czech English**

Jan Volin, Radek Skarnitzl

Page: 279

**Exploiting Variety-Dependent Phones in Portuguese Variety Identification**

Oscar Koller, Alberto Abad, Isabel Trancoso

Page: 287

**Author Index**