VOLUME 3

Post-Aspiration in Standard Italian: Some First Cross-Regional Acoustic Evidence
Mary Stevens, John Hodge

Articulatory Grounding of Southern Salentino Harmony Processes
Mirko Grimaldi, Andrea Calabrese, Francesco Sgona, Luigina Garrapa, Bianca Sisinni

Effects of Accent Typicality and Phonotactic Frequency on Nonword Immediate Serial Recall Performance in Japanese
Yuki Tanida, Taiji Ueno, Satoshi Saito, Matthew A. Lambon Ralph

How Abstract is Phonetics?
Osamu Fujimura

TUE-SES3-P4: SPEECH PRODUCTION II: VOCAL TRACT MODELING AND IMAGING

Data-Driven Analysis of Realtime Vocal Tract MRI Using Correlated Image Regions
Adam C. Lammert, Michael I. Proctor, Shrikanth S. Narayanan

Rapid Semi-Automatic Segmentation of Real-Time Magnetic Resonance Images for Parametric Vocal Tract Analysis
Michael I. Proctor, Daniel Bone, Athanasias Katsamanis, Shrikanth S. Narayanan

Improved Real-Time MRI of Oral-Velar Coordination Using a Golden-Ratio Spiral View Order
Taon-Chiel Kim, Shrikanth S. Narayanan, Krishna S. Nayak

Statistical Multi-Stream Modeling of Real-Time MRI Articulatory Speech Data
Erik Bresch, Athanasias Katsamanis, Louis Goldstein, Shrikanth S. Narayanan

Predicting Unseen Articulations from Multi-Speaker Articulatory Models
G. Ananthakrishnan, Pierre Bouin, Julian Andres Faisal Vergas, Glen Engelwalt

Estimating Missing Data Sequences in X-Ray Microbeam Recordings
Chao Qin, Miguel A. Correia-Perpinan

Adaptation of a Tongue Shape Model by Local Feature Transformations
Chao Qin, Miguel A. Correia-Perpinan, Mohsen Farhadloo

Vocal Tract Contour Analysis of Emotional Speech by the Functional Data Curve Representation

Locally-Weighted Regression for Estimating the Forward Kinematics of a Geometric Vocal Tract Model
Adam C. Lammert, Louis Goldstein, Khalil Iskarous

Identifying Articulatory Goals from Kinematic Data Using Principal Differential Analysis
Michael Reimer, Frank Rudzicz

Estimation of Speech Lip Features from Discrete Cosinus Transform
Zuheng Ming, Denis Broustemp, Gang Feng, Sebastien Schmider

Autoregressive Modelling for Linear Prediction of Ultrasonic Speech
Furamaxe Ahmed, Ian V. McLoughlin, Hamid R. Shafizadeh

WED-SES1-S1: SPECIAL SESSION: SPEECH INTELLIGIBILITY ENHANCEMENT FOR ALL AGES, HEALTH CONDITIONS AND ENVIRONMENTS

Enhanced Speech Yielding Higher Intelligibility for All Listeners and Environments
Takayuki Araki, Nao Hodoshima
Quality Conversion of Non-Acoustic Signals for Facilitating Human-to-Human Speech Communication Under Harsh Acoustic Conditions

Seyed Omid Sadjadi, Sanjay A. Palil, John H. L. Hansen

The Use of Air-Pressure Sensor in Electrolaryngeal Speech Enhancement Based on Statistical Voice Conversion

Keigo Nakanuma, Tomoki Toda, Hidetsugu Sato, Koichiro Shikano

A New Binary Mask Based on Noise Constraints for Improved Speech Intelligibility

Gihak Kim, Philippas C. Loizou

Energy Reallocation Strategies for Speech Enhancement in Known Noise Conditions

Yen Tang, Martin Cooke

Effects of Enhancement of Spectral Changes on Speech Quality and Subjective Speech Intelligibility

Jing Chen, Thomas Baer, Brian C. J. Moore

**WED-SES1-O1: ASR: ACOUSTIC MODEL ADAPTATION**

Prior Information for Rapid Speaker Adaptation

C. Broth, D. Gilches, M. J. F. Gates, K. B. Knoll, K. Xu

Discriminative Adaptation for Log-Linear Acoustic Models

Jian Lu, Ralf Schlueter, Hernan Ney

Automatic Speech Recognition of Multiple Accented English Data

 Dimitrou Vargyi, Liet Lenu, Jean-Luc Gauvain

Shrinkage Model Adaptation in Automatic Speech Recognition

Jinyu Li, Yu Tan, Chong-Hui Lee

Unsupervised Transform with Online Distortion Estimation for HMM Adaptation

Jinyu Li, Dongyu Yi, Yifen Gong, L. Deng

HMM Adaptation Using Linear Spline Interpolation with Integrated Spline Parameter Training for Robust Speech Recognition

Michael L. Seltzer, Alex Acero

**WED-SES1-O2: SLP SYSTEMS FOR INFORMATION EXTRACTION/RETRIEVAL**

CRF-Based Stochastic Pronunciation Modeling for Out-of-Vocabulary Spoken Term Detection

Chen-Ping Chen, Heng-Yi Lee, Ching-Feng Yeh, Lin-Shao Lee

Towards Spoken Term Discovery at Scale with Zero Resources

Aven Jansen, Kenneth Church, Hynek Hermansky

Vocabulary Independent Spoken Query: A Case for Subword Units

Evandro Gouvea, Tony Ezaz

Extractive Speech Summarization --- From the View of Decision Theory

Shih-Hsiang Lin, Yau-Ming Yeh, Berlin Chen

The Impact of ASR on Abstractive vs. Extractive Meeting Summaries

Gabriel Murray, Giuseppe Caruolo, Raymond Ng

**WED-SES1-O3: SPEECH REPRESENTATION**

Binary Coding of Speech Spectrograms Using a Deep Auto-Encoder

Li Deng, Michael L. Seltzer, Dong Yu, Alex Acero, Abdal-Rahman Mohamed, G. Hinton

A Super-Resolution Spectrogram Using Coupled PLCA

Julien Nam, Guanghua J. M. Moore, Joaochin Cuninman, Kyung Lee, Jonathan S. Abel

Fast Least-Squares Solution for Sinusoidal, Harmonic and Quasi-Harmonic Models

Georgios Tzedakis, Younis Fantazi, Ollefar Rossoc, Younis Syliou

Sparse Component Analysis for Speech Recognition in Multi-Speaker Environment

Afiyaw Aslul, Harve Bourlard, Thomas N. Gerner

Intra-Frame Variability as a Predictor of Frame Classifiability

Trond Engstad, Torbjorn Svendsen

Autocorrelation and Double Autocorrelation Based Spectral Representations for a Noisy Word Recognition System

Tetsuya Shimamura, Nguyen Dinh Nguyen

**WED-SES1-O4: VOICE CONVERSION**

Maximum a posteriori Voice Conversion Using Sequential Monte Carlo Methods

Elina Helander, Hanna Sillan, Joaochin Cuninman, Moncef Gabbouj

Dynamic Model Selection for Spectral Voice Conversion

Pierre Lanchantin, Xavier Ridel

Speaker-Independent HMM-Based Voice Conversion Using Quantized Fundamental Frequency

Takahiro Nao, Takanobu Kobayashi
Probabilistic Integration of Joint Density Model and Speaker Model for Voice Conversion .................................................. 1728
Daisuke Saito, Shinji Watanabe, Atsushi Nakamura, Nobuaki Minematsu
Text-Independent F0 Transformation with Non-Parallel Data for Voice Conversion .................................................. 1732
Zhi-Zheng Wu, Tonti Kinnunen, Eng Siang Chong, Haizhou Li
A Minimum Converted Trajectory Error (MCTE) Approach to High Quality Speech-to-Lips Conversion .......................... 1736
Xiaodan Zhang, Lijian Wang, Frank K. Soong, Mark Hasegawa-Johnson

WED-SES1-P1: PROSODY: LANGUAGE-SPECIFIC MODELS

Influence of Lexical Tones on Intonation in Kammu ............................................ 1740
Anastasia Karkson, David House, Jan-Ohf Svaneveen, Damrong Tayapin
Phonetic Realization of Second Occurrence Focus in Japanese ............................................ 1744
Sasuki Nambu, Yong-Chool Lee
Prosodic Grouping and Relative Clause Disambiguation in Mandarin ............................................ 1748
Jianyi Kuang
Text-Based Unstressed Syllable Prediction in Mandarin ............................................ 1752
Yu Li, Jiuhua Tao, Meng Zhang, Shijing Pan, Xueying Xu
"Flat Pitch Accents" in Czech ............................................ 1756
Tomas Dubeda

WED-SES1-P2: ASR: LANGUAGE MODELING AND SPEECH UNDERSTANDING I

Within and Across Sentence Boundary Language Model ............................................ 1800
Saeedeh Momtazi, Friedrich Pasdel, Diereich Klukow
Impact of Word Classing on Shrinkage-Based Language Models ............................................ 1804
Rudi Sarkarya, Stanley F. Chen, Athinao Sety, Bhuvana Ramabhadran
Combination of Probabilistic and Possibilistic Language Models ............................................ 1808
Stamtsis Oger, Vladimir Ponomr, Georges Linares
On-Demand Language Model Interpolation for Mobile Speech Input ............................................ 1812
Tim Schlippe, Chenfei Zhu, Jan Gebhardt, Tanja Schultz
Efficient Estimation of Maximum Entropy Language Models with N-Gram Features: An SRILM Extension ............................................ 1820
Christian Gillot, Christophe Cerisara, David LanglLOT, Jean-Paul Haon

WED-SES1-P3: FIRST AND SECOND LANGUAGE ACQUISITION

Cantonese Tone Word Learning by Tone and Non-Tone Language Speakers ............................................ 1840
Angela Cooper, Yue Wang
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation of a Training Method for L2 Continuous-Speech Segmentation</td>
<td></td>
</tr>
<tr>
<td>Anne Cutler, Jane Shalley</td>
<td>1844</td>
</tr>
<tr>
<td>Linguistic Rhythm in Foreign Accent</td>
<td></td>
</tr>
<tr>
<td>Juijiao Yuan</td>
<td>1848</td>
</tr>
<tr>
<td>The Effect of a Word Embedded in a Sentence and Speaking Rate Variation on the Perceptual Training of Geminate and Singleton Consonant Distinction</td>
<td></td>
</tr>
<tr>
<td>Mee Sun, Keichi Tajima, Hiroaki Kato, Yoshinori Sagisaka</td>
<td>1850</td>
</tr>
<tr>
<td>Foreign Accent Matters Most When Timing is Wrong</td>
<td></td>
</tr>
<tr>
<td>Chi harsh Tsuratani</td>
<td>1854</td>
</tr>
<tr>
<td>Effects of Korean Learners' Consonant Cluster Reduction Strategies on English Speech Recognition Performance</td>
<td></td>
</tr>
<tr>
<td>Hyejin Hong, Jinsu Kim, Minhyo Chung</td>
<td>1858</td>
</tr>
<tr>
<td>The Effects of EMA-Based Augmented Visual Feedback on the English Speakers' Acquisition of the Japanese Flap: A Perceptual Study</td>
<td></td>
</tr>
<tr>
<td>June S. Levitt, William F. Katz</td>
<td>1862</td>
</tr>
<tr>
<td>Perception of Voiceless Fricatives by Japanese Listeners of Advanced and Intermediate Level English Proficiency</td>
<td></td>
</tr>
<tr>
<td>Hinako Masuda, Tokayuki Ari</td>
<td>1866</td>
</tr>
<tr>
<td>Perception of Estonian Vowel Categories by Native and Non-Native Speakers</td>
<td></td>
</tr>
<tr>
<td>LySa Metzler, Elmar Metzler</td>
<td>1870</td>
</tr>
<tr>
<td>Spoken English Assessment System for Non-Native Speakers Using Acoustic and Prosodic Features</td>
<td></td>
</tr>
<tr>
<td>Qiao Shi, Kan Li, Shihui Zhang, Stephen M. Chu, Xiu Xia, Zhihan Ou</td>
<td>1874</td>
</tr>
<tr>
<td>Russian Infants and Children's Sounds and Speech Corporuses for Language Acquisition Studies</td>
<td></td>
</tr>
<tr>
<td>Elena S. Lyakso, Olga V. Frolova, Anna V. Karachova, Julia S. Gaikova</td>
<td>1878</td>
</tr>
<tr>
<td>Language-Specific Influence on Phoneme Development: French and Drehu Data</td>
<td></td>
</tr>
<tr>
<td>Julia Monnin, Helen Leenhanbuck</td>
<td>1882</td>
</tr>
<tr>
<td>Did you Say Susi or Shushi? Measuring the Emergence of Robust Fricative Contrasts in English- and Japanese-Acquiring Children</td>
<td></td>
</tr>
<tr>
<td>Jeffrey J. Holliday, Mary E. Beckman, Chandlee Myia</td>
<td>1886</td>
</tr>
</tbody>
</table>

**WED-SES1-P4: SPOKEN LANGUAGE RESOURCES, SYSTEMS AND EVALUATION I**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Empirical Comparison of the T3, Juicer, HDecode and Sphinx Decoders</td>
<td></td>
</tr>
<tr>
<td>Josef R. Novak, Paul R. Dixon, Sadaoki Furai</td>
<td>1890</td>
</tr>
<tr>
<td>Tracter: A Lightweight Dataflow Framework</td>
<td></td>
</tr>
<tr>
<td>Philip H. Garner, John Dienes</td>
<td>1894</td>
</tr>
<tr>
<td>Verifying Pronunciation Dictionaries Using Conflict Analysis</td>
<td></td>
</tr>
<tr>
<td>Morelia H. Duval, Febe Du Wet</td>
<td>1898</td>
</tr>
<tr>
<td>Automatic Estimation of Transcription Accuracy and Difficulty</td>
<td></td>
</tr>
<tr>
<td>Braham K. Ray, Sarvath Vazhaghah, Deb Roy</td>
<td>1902</td>
</tr>
<tr>
<td>Creating a Linguistic Plausibility Dataset with Non-Expert Annotators</td>
<td></td>
</tr>
<tr>
<td>Benjamin Lambers, Rita Singh, Bhikshu Raj</td>
<td>1906</td>
</tr>
<tr>
<td>Construction and Evaluations of an Annotated Chinese Conversational Corpus in Travel Domain for the Language Model of Speech Recognition</td>
<td></td>
</tr>
<tr>
<td>Xiaohui Hu, Ryucoo Iizumi, Hitachi Kanai, Satoshi Nohama</td>
<td>1910</td>
</tr>
<tr>
<td>Building Transcribed Speech Corpora Quickly and Cheaply for Many Languages</td>
<td></td>
</tr>
<tr>
<td>Thad Hughes, Koichi Nishijima, Hiromu Aki, Pedro J. Moreno, Mike Leibens</td>
<td>1914</td>
</tr>
<tr>
<td>The CHIME Corpus: A Resource and a Challenge for Computational Hearing in Multisource Environments</td>
<td></td>
</tr>
<tr>
<td>Heidi Christiansen, Jon Barker, Wang Ma, Phil D. Green</td>
<td>1918</td>
</tr>
<tr>
<td>Developing a Chinese L2 Speech Database of Japanese Learners with Narrow-Phonetic Labels for Computer Astistic Pronunciation Training</td>
<td></td>
</tr>
<tr>
<td>Wen Cao, Dongqing Wang, Jinsong Zhang, Ziyi Xiong</td>
<td>1922</td>
</tr>
<tr>
<td>How Children Acquire Situation Understanding Skills?: A Developmental Analysis Utilizing Multimodal Speech Behavior Corpus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1926</td>
</tr>
<tr>
<td>The Influence of Expertise and Efficiency on Modality Selection Strategies and Perceived Mental Effort</td>
<td></td>
</tr>
<tr>
<td>Ina Wechsung, Stefan Schaffert, Robert Schielicher, Anna Hansmann, Sebastian Moller</td>
<td>1930</td>
</tr>
<tr>
<td>Parameters Describing Multimodal Interaction — Definitions and Three Usage Scenarios</td>
<td></td>
</tr>
<tr>
<td>Christine Kubler, Benjamin Weiss, Sebastian Moller</td>
<td>1934</td>
</tr>
<tr>
<td>Repair Strategies on Trial: Which Error Recovery Do Users Like Best?</td>
<td></td>
</tr>
<tr>
<td>Alexander Zgorzelski, Alexander Schmidt, Tobias Heintroth, Wolfgang Minkler</td>
<td>1938</td>
</tr>
</tbody>
</table>

**WED-SES2-01: ASR: SEARCH, DECODING AND CONFIDENCE MEASURES II**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRF-Based Combination of Contextual Features to Improve a posteriori Word-Level Confidence Measures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1942</td>
</tr>
<tr>
<td>Recognition of Spontaneous Conversational Speech Using Long Short-Term Memory Phoneme Predictions</td>
<td></td>
</tr>
<tr>
<td>Marin Wellner, Florian Eyben, Bjorn Schuller, Gerhard Rigoll</td>
<td>1946</td>
</tr>
<tr>
<td>Improving ASR Error Detection with Non-Decoder Based Features</td>
<td></td>
</tr>
<tr>
<td>Thomas Pellegnini, Isabel Trancano</td>
<td>1950</td>
</tr>
</tbody>
</table>
Mask Estimation in Non-Stationary Noise Environments for Missing Feature Based Robust Speech Recognition
Shirin Badiezadegan, Richard C. Rose

Robust Automatic Speech Recognition with Decoder Oriented Ideal Binary Mask Estimation
Lee-Hoon Kim, Kyoung-Tae Kim, Mark Haugens-Johnston

A Robust Speech Recognition System Against the Ego Noise of a Robot
Gohkan Ince, Kasahiro Nozakidai, Tobias Redemann, Hiroshi Tagawa, Jun-Ichi Inoue

Empirical Mode Decomposition for Noise-Robust Automatic Speech Recognition
Kuo-Hao Wu, Chia-Ping Chen

An Effective Feature Compensation Scheme Tightly Matched with Speech Recognizer Employing SVM-Based GMM Generation
Woosil Kim, Jun-Won Soh, John H. L. Hansen

Artificial and Online Acquired Noise Dictionaries for Noise Robust ASR
Jori F. Gemmeke, Thomas Virtanen

Voice Activity Detection Based on Conditional Random Fields Using Multiple Features
Akira Soito, Yoshitake Nankake, Akihiko Lee, Ketsch Tokuda

A Comparative Study of Noise Estimation Algorithms for VTS-Based Robust Speech Recognition
Yong Zhao, Bing-Hwao Jiang

On Using Missing-Feature Theory with Cepstral Features — Approximations to the Multivariate Integral
Frank Stolte, Pat Zhao

Using a DBN to Integrate Sparse Classification and GMM-Based ASR
Yang Sun, Jori F. Gemmeke, Bert Cranen, L. Tim Bosch, Lou Bouwes

WED-SES2-P2: SPEAKER CHARACTERIZATION AND RECOGNITION IV

Transcript-Dependent Speaker Recognition Using Mixer 1 and 2
Fred S. Richardson, Joseph P. Campbell

On the Potential of Glottal Signatures for Speaker Recognition
Thomas Drugman, Thierry Dutot

Acoustic Feature Diversity and Speaker Verification
R. Padmanabhan, Hema A. Murthy

A Discriminative Performance Metric for GMM-UBM Speaker Identification
Omid Dehzangi, Bin Ma, Eng Siong Cheng, Haizhou Li

A Novel Speaker Binary Key Derived from Anchor Models
Xavier Anguera, Jean-Francois Bonastre

Variant Time-Frequency Cepstral Features for Speaker Recognition
Wei-Qiang Zhang, Yan Deng, Liang He, Jia Liu

Exploitation of Phase Information for Speaker Recognition
Rong Wang, P. C. Cheng, Yan Lee

Effects of the Phonological Relevance in Speaker Verification
Yashua Long, Lirong Dai, Bin Ma, Wu Guo

Topological Representation of Speech for Speaker Recognition
Gabriel H. Sierra, Jean-Francois Bonastre, Dries Matroos, Jose R. Calvo

Assessment of Single-Channel Speech Enhancement Techniques for Speaker Identification Under Mismatched Conditions
Sayed Omid Dadjordi, John H. L. Hansen

Speaker Recognition Using the Resynthesized Speech via Spectrum Modeling
Xiang Zhang, Chuan Cen, Lin Yang, Huihui Sun, Jianping Zhang, Yonghong Yan

WED-SES2-P3: VOICE CONVERSION AND SPEECH SYNTHESIS

Shape-Invariant Speech Transformation with the Phase Vocoder
Axel Robel

A Phonetic Alternative to Cross-Language Voice Conversion in a Text-Dependent Context: Evaluation of Speaker Identity
Kyoake Tonomori, Mark Hecke

Evaluation of Speaker Mimic Technology for Personalizing SGD Voices
Esther Klaiber, Alexander Kain, Jan P. H. Van Santen

Adaptive Voice-Quality Control Based on One-to-Many Eigenvoice Conversion
Kum Ohia, Tomoki Toeda, Yasuo Okuani, Hiroshi Saruwatari, Kiyohiro Shikano

Applying Voice Conversion to Concatenative Singing-Voice Synthesis
Fernando Villavicencio, Jordi Bonada

Improved Generation of Fundamental Frequency in HMM-Based Speech Synthesis Using Generation Process Model
Ming Lei, Yiyan Wu, Frank K. Soong, Zhen-Hua Ling, Lirong Dai

A Hierarchical F0 Modeling Method for HMM-Based Speech Synthesis
Ming Lei, Yiyan Wu, Frank K. Soong, Zhen-Hua Ling, Lirong Dai

Training a Parametric-Based Log-F0 Model with the Minimum Generation Error Criterion
Xavier Anguera, M. J. F. Gales, Haizhou Li
Improving Mandarin Segmental Duration Prediction with Automatically Extracted Syntax Features
Miaomiao Wen, Miaomiao Wang, Kelsichi Hirose, Nobuaki Minematsu

An Intonation Model for TTS in Sepedi
Daniel R. Van Niekerk, Elienne Barnard

Synthesis of Fast Speech with Interpolation of Adapted HSMNs and Its Evaluation by Blind and Sighted Listeners
Michael Puehler, Dietmar Schabus, Junichi Yamagishi

A Comparison of Pronunciation Modeling Approaches for HMM-TTS
Gabriel Webster, Saso Krsuljovic, Kate Knill

HMM-Based Text-to-Articulatory-Movement Prediction and Analysis of Critical Articulators
Zhi-Wei Ling, Kurtis Richmond, Junichi Yamagishi

WED-SES2-P4: DETECTION, CLASSIFICATION, AND SEGMENTATION

Audio-Based Sports Highlight Detection by Fourier Local Auto-Correlations
Jiasing Ye, Takuori Kobayashi, Tetsuya Higuchi

Automatic Excitement-Level Detection for Sports Highlights Generation
Hynek Bartl, Abhijeet Sangwan, Toufik Hassan, John R. L. Hansen

Detecting Novel Objects in Acoustic Scenes Through Classifier Incongruence
Jorg-Hendrik Bach, Juris Anemuller

A Multidomain Approach for Automatic Home Environmental Sound Classification
Savero Mabomphira, Hynek Potamitis, Nikos Fakotakis

Content-Based Advertisement Detection
Patrick Cardinal, Vithwa Gupta, Giltas Baslamas

Identification of Abnormal Audio Events Based on Probabilistic Novelty Detection
Shreves Nalaprapit, Hynek Potamitis, Nikos Fakotakis

Lightly Supervised Recognition for Automatic Alignment of Large Coherent Speech Recordings
Norbert Brownsholtz, M. J. P. Gales, Sabine Buchholz

Incremental Diarization of Telephone Conversations
Osthy Ben-Hur, Tichaok Lopidet, Hugo generman

Audio Analytics by Template Modeling and 1-Pass DP Based Decoding
Srikanth Cherla, V. Ramamohanam

Perceptual Wavelet Decomposition for Speech Segmentation
Martha Ziolko, Jakab Galka, Bartosz Ziolko, Tomasz Drozdow

A Comparative Study of Constrained and Unconstrained Approaches for Segmentation of Speech Signal
Venkatish Kari, Kishore Prabhu

Automatic Discriminative Measurement of Voice Onset Time
Morgan Sanderegger, Joseph Keshet

Selective Gammatone Filterbank Feature for Robust Sound Event Recognition
Yi Ren Long, Hui Dat Tran, Norshida Kiasa, Hezhu Li

WED-SES3-S1: SPECIAL SESSION: COMPRESSIVE SENSING FOR SPEECH AND LANGUAGE PROCESSING

Towards a Robust Face Recognition System Using Compressive Sensing
Allen T. Yang, Zihan Zhou, Yi Mu, S. Shankar Sastry

Sparse Representation Features for Speech Recognition
Avinash Sathy, Sara S. Sainath, Bhuvana Ramabhadran, Dimitri Kaneval

Data Selection for Language Modeling Using Sparse Representations
Abhinav Sathy, Sara N. Sainath, Bhuvana Ramabhadran, Dimitri Kanevsky

Observation Uncertainty Measures for Sparse Imputation
Joon F. Gemmeke, Utpal Ramesh, Kalle J. Palomaki

Sparse Representations for Text Categorization
Tara N. Sainath, Sameer R. Musaik, Dimitri Kanevsky, Bhuvana Ramabhadran, David Nahamoo, Julia Hirschberg

Sparse Auto-Associative Neural Networks: Theory and Application to Speech Recognition
G. S. V. S. Sivaram, Sriram Ganapathy, Hynek Hermansky

WED-SES3-01: ASR: LEXICAL AND PRONUNCIATION MODELING

FSM-Based Pronunciation Modeling Using Articulatory Phonological Code
Chih-Hsuan Chiu, Xiongden Zhuang, Mark Hargreaves-Johnston

Detailed Pronunciation Variant Modeling for Speech Transcription
Denis Jousse, Dominique Fohr, Erina Blome

A Minimum Classification Error Approach to Pronunciation Variation Modeling of Non-Native Proper Names
Line Addt, Bert Revell, Jean-Pierre Martens, Torbjorn Svensen

Acoustics-Based Phonetic Transcription Method for Proper Nouns
Antoine Laurent, Sylvain Meignier, Tova Martin, Paul Deligne

Wiktionary as a Source for Automatic Pronunciation Extraction
Tim Schlimke, Sebastian Ochs, Tanja Schultz
Learning New Word Pronunciations from Spoken Examples
Ibrahim Balci, Ian McGregor, James Glass

WED-SES3-Q2: SPEAKER RECOGNITION AND DIARIZATION

Phonetic Subspace Mixture Model for Speaker Diarization
I-Fan Chen, Shih-Sian Cheng, Hsiao-Min Wang

Overlap Detection for Speaker Diarization by Fusing Spectral and Spatial Features
Martin Zelenak, Carlos Segura, Javier Hernando

Floor Holder Detection and End of Speaker Turn Prediction in Meetings
Alfred Diehmann, Giulio Garau, Herve Bearlard

Confidence Measures for Speaker Segmentation and Their Relation to Speaker Verification
Carlos Vaquero, Alfonso Ortega, Juma Villalba, Antonio Miguel, Eduardo Lleida

Decoupling Session Variability Modelling and Speaker Characterisation
Anthony Larcher, Christophe Ley, Driks Matrouf, Jean-Francois Bonezare

Incorporating MAP Estimation and Covariance Transform for SVM Based Speaker Recognition
Cheung-Chi Leung, Dongli Zhao, Kong Ai Lee, Bin Ma, Haizhou Li

WED-SES3-Q3: SPEECH AND AUDIO CLASSIFICATION

Single-Speaker/Multi-Speaker Co-Channel Speech Classification
Stephanie Rossignol, Olivier Pietquin

Discriminative Training for Hierarchical Clustering in Speaker Diarization
Oriol Vinyals, Gerald Friedland, Nelson Morgan

GMM-UBM Based Open-Set Online Speaker Diarization
Jurgen Geiger, Frank Wallhoff, Gerhard Rigoll

A Segment-Based Non-Parametric Approach for Monophone Recognition
Ladan Golipour, Dangson O'Shaughnessy

A Fast One-Pass-Training Feature Selection Technique for GMM-Based Acoustic Event Detection with Audio-Visual Data
Taras Butko, Climent Nadeu

Effects of Modelling Within- and Between-Frame Temporal Variations in Power Spectra on Non-Verbal Sound Recognition
Nobuhide Yamakawa, Tetsuro Kikuchi, Toru Takahashi, Kazuyoshi Koidan, Tetsuya Ogata, Hiroshi Okuno