11th Annual Conference of the International Speech Communication Association 2010

(INTERSPEECH 2010)

Makuhari, Chiba, Japan
26-30 September 2010

Volume 4 of 4

WED-SES3-04: EMOTION RECOGNITION

On the Importance of Glottal Flow Spectral Energy for the Recognition of Emotions in Speech
Ling He, Margaret Lech, Nicholas Allen

Real-Life Emotion-Related States Detection in Call Centers: A Cross-Corpora Study
Laurence Devillers, Christophe Fauve, Clement Chastagnol

Multi-Class and Hierarchical SVMs for Emotion Recognition
Ali Hatton, Robert I. Dumner

Determining Optimal Features for Emotion Recognition from Speech by Applying an Evolutionary Algorithm
David Hubner, Bogdan Flaszko, Tobias Grosser, Andreas Ristenpart

Context-Sensitive Multimodal Emotion Recognition from Speech and Facial Expression Using Bidirectional LSTM Modeling
Martin Wolfner, Angeliki Metallinou, Florian Eyben, Bjorn Schuller, Shrikanth S. Narayanan

Data-Dependent Evaluator Modeling and Its Application to Emotional Valence Classification from Speech
Karlik Audhkhasi, Shrikanth S. Narayanan

WED-SES3-P1: SPEECH CODING, MODELING, AND TRANSMISSION

Modelling Speech Line Spectral Frequencies with Dirichlet Mixture Models
Zhanyu Ma, Arne Leijon

PDF-Optimized LSF Vector Quantization Based on Beta Mixture Models
Zhanyu Ma, Arne Leijon

Non-Linear Predictive Vector Quantization of Feature Vectors for Distributed Speech Recognition
Jose Enrique Garcia, Alfonso Ortega, Antonio Miguel, Eduardo Latido

Superwideband Extension of G.718 and G.729.1 Speech Codecs
Lasse Laalaonen, Mikko Tammi, Vladimir Malenovsky, Tommy Vallanceart, Mi Suk Lee, Tamefumi Yamamata, Masahiro Cahlker, Claude Lanhbhin, Rahul Koesia, Lei Min, Deming Zhang, Jon Gibbs, Holly Francois

A Multi-pulse FEC Scheme Based on Amplitude Estimation for CELP Codecs Over Packet Networks
Jose L. Carmona, Angel M. Gomez, Antonio M. Pelmado, Jose L. Perez-Cordoba, Jose A. Gonzalez
Voice Quality Evaluation of Recent Open Source Codecs

Anssi Ramo, Henri Tootkaboni

Efficient HMM-Based Estimation of Missing Features, with Applications to Packet Loss Concealment

Boo J. Borgstrom, Per H. Borgstrom, Abhij Alvar

Speech Inventory Based Discriminatory Training for Joint Speech Enhancement and Low-Rate Speech Coding

Xiaoqiang Xiao, Robert M. Nickal

Quality-Based Playout Buffering with FEC for Conversational VoIP

Qipeng Gong, Peter Kobal

Sub-Band Basis Spectrum Model for Pitch-Synchronous Log-Spectrum and Phase Based on Approximation of Sparse Coding

Masasuke Tamura, Takahiro Kogaibhina, Masumi Akamine

A Multimodal Density Function Estimation Approach to Formant Tracking

Harshvardhan S., Chandra Sekhar Selvamani, Y. V. Srinivas

Estimation Studies of Vocal Tract Shape Trajectory Using a Variable Length and Lossy Kelly-Lochbaum Model

Helkki Rastio, Unko K. Lahti, Oikka Rananen

WED-SES3-P2: ASR: LANGUAGE MODELING AND SPEECH UNDERSTANDING II

Improving Back-Off Models with Bag of Words and Hollow-Grams

Benjamin Lecouteux, Raphael Rubino, Georges Linares

Study on Interaction Between Entropy Pruning and Kneser-Ney Smoothing

Ciprian Cheliba, Thorsten Branta, Will Hovais, Feng Xu

Dynamic Language Model Adaptation Using Keyword Category Classification

Akhil Hadkar, Douglas O'Shaugnnessy

Automatic Speech Recognition System Channel Modeling

Arun Deng, Ramoni De Meli

Novel Weighting Scheme for Unsupervised Language Model Adaptation Using Latent Dirichlet Allocation

Abdul Haslo, Douglas O'Shaugnnessy

Round-Robin Discrimination Model for Reranking ASR Hypotheses

Qun Feng Tan, Karthik Aitdhkhasi, Panayiotis G. Georgian, Emit Ettelaie, Shrikanth S. Narayannam

On-the-Fly Lattice Rescoring for Real-Time Automatic Speech Recognition

Hasan Seh, Murat Saraclar, Tung Gung

WED-SES3-P3: SPEECH PERCEPTION III: PROCESSING AND INTELLIGIBILITY

A Feature Extraction Method for Automatic Speech Recognition Based on the Cochlear Nucleus

Sreasal Huque, Roberta Tognari

A Phonoeme Recognition Framework Based on Auditory Spectro-Temporal Receptive Fields

Samuel Thomas, Kailash Patili, Srijan Goswadhi, Nina Murgaram, Jyoti Hermaikey

Perceptual Compensation for Effects of Reverberation in Speech Identification: A Computer Model Based on Auditory Effferent Processing

Amy Y. Bresnat, Guy J. Brown

Predicting Human Perception and ASR Classification of Word-Final [t] by Its Acoustic Sub-Segmental Properties

Barbara Schaeper, Milyan Ernestus, Tim Van Dommelen, Jacques Keremain

Multichannel Source Separation Based on Source Location Cue with Log-Spectral Shaping by Hidden Markov Source Model

Takahiko Nakatani, Shoko Araki, Takuya Yoshioka, Masakiyo Fujimoto

A Speech-in-Noise Test Based on Spoken Digits: Comparison of Normal and Impaired Listeners Using a Computer Model

Takahiro Nakatani, Shoko Araki, Takuya Yoshioka, Masakiyo Fujimoto

Evaluation of Bone-Conducted Ultrasonic Hearing-Aid Regarding Transmission of Paralinguistic Information: A Comparison with Cochlear Implant Simulator

Takahiro Nakatani, Shoko Araki, Takuya Yoshioka, Masakiyo Fujimoto

Challenging the Speech Intelligibility Index: Macroscopic vs. Microscopic Prediction of Sentence Recognition in Normal and Hearing-Impaired Listeners

Trampi N. Uslar, Thomas Brand, Marko Ho, Rebecca Carroll, Esther Buigendijk, Carmelia Hammar, Birger Kollmeier

Does Sentence Complexity Interfere with Intelligibility in Noise? Evaluation of the Oldenburg Linguistically and Audiologically Controlled Sentence Test (OLACS)

Juan-Pablo Ramirez, Homed Ketabdar, Alexander Raake

Intelligibility Predictions for Speech Against Fluctuating Masker

Juan-Pablo Ramirez, Homed Ketabdar, Alexander Raake

An Effect of Formant Amplitude in Vowel Perception

Masami Hio, Koji Obara, Akioh Ito, Masafumi Yano
WED-SES3-P4: SPOKEN LANGUAGE UNDERSTANDING AND SPOKEN LANGUAGE TRANSLATION I

Functional Imaging of Brain Regions Sensitive to Communication Sounds in Primates .................................................. 2498
Christopher J. Petkov, Benjamin Wilson

Strategies for Statistical Spoken Language Understanding with Small Amount of Data — An Empirical Study .......................... 2502
Ye-Yi Wang

Investigating Multiple Approaches for SLU Portability to a New Language ................................................................. 2506
Bassam Jabaian, Laurent Benacier, Fabrice Lefevre

Learning Naturally Spoken Commands for a Robot ................................................................. 2510
Asja Austermann, Solji Yamada, Kiyoto Fushikazu, Miki Nakano

A-Semi-Supervised Cluster-and-Label Approach for Utterance Classification .......................................................... 2514
Amgara Abidane, Aparna Suchindranath, David Austermann, Wolfgang Moker

Classifying Dialog Acts in Human-Human and Human-Machine Spoken Conversations ........................................ 2518
Silvia Querzini, Giuseppe Riccardi

Exploring Speaker Characteristics for Meeting Summarization .................................................................................. 2522
Pai Lin, Tony Liu

Semi-Supervised Extractive Speech Summarization via Co-Training Algorithm ...................................................... 2526
Shashu Xin, Hui Lin, Yung Liu

Extractive Summarization Using a Latent Variable Model ....................................................................................... 2530
Asli Celikyilmaz, Dilek Hakkani-Tür

Hierarchical Classification for Speech-to-Speech Translation ............................................................................. 2534
Emre Etel, Panayiotis G. Georgiou, Shrikant S. Narayanan

Rapid Development of Speech Translation Using Consecutive Interpretation .............................................................. 2538
Matthias Paukli, Alex Wabbel

THU-SES1-S1: SPECIAL SESSION: SOCIAL SIGNALS IN SPEECH

Combining Many Alignments for Speech to Speech Translation .................................................................................. 2542
Sameer R. Mokady, Steven J. Kermile, Bowen Zhou

Detecting Politeness and Efficiency in a Cooperative Social Interaction ........................................................................ 2546
Paul M. Brunet, Marcela Chatzifilian, Rodrick Cowie, Marc Schneider, Hastings Donman, Elin Douglas-Cowie

Comparing Measures of Synchrony and Alignment in Dialogue Speech Timing with Respect to Turn-Taking Activity ...................................................................................................................... 2550
Nick Campbell, Stefan Scherer

Resources for Turn Competition in Overlap in Multi-Party Conversations: Speech Rate, Pausing and Duration ............ 2554
Dina Khord, Guy J. Brown, Bill Wells

Disambiguating the Functions of Conversational Sounds with Prosody: The Case of 'Yeah' ........................................... 2558
Kha Ph. Truong, Dirk Hoores

Prosody and Voice Quality of Vocal Social Signals: The Case of Dominance in Scenario Meetings ............................ 2562
Marcela Chatzifilian, Marc Schneider, Imen Steiner

THU-SES1-O2: PHYSIOLOGY AND PATHOLOGY OF SPOKEN LANGUAGE

Online SLU Model Adaptation with a Partial Oracle ...................................................................................................... 2566
Pierre Gofeh, Geraline Dammari, Frederic Bechet, Lionel Delphina-Poulet

The Prosody of Swedish Conversational Grunts ....................................................................................................... 2570
D. Neber, J. Gustafson

Reliable Tracking Based on Speech Sample Salience of Vocal Cycle Length Perturbations .............................................. 2574
C. Martina, Francis Grenou, Line Crevier-Buchman, Jean Schoonhagen

Longitudinal Changes of Selected Voice Source Parameters .................................................................................... 2578
Iriki Kasain, Hajime Yushida, Satsuki Ishikari, Hiroshi Morit

Automatic Perceptual Categorization of Disordered Connected Speech ........................................................................... 2582
A. Ajar, Jean Schoonhagen, T. Maryn, Francis Grenou

Kinematic Analysis of Tongue Movement Control in Spastic Dysarthria ............................................................. 2586
Heeja Kim, Pampring Rong, Torrey M. Lawek, Mark Hasegawa-Johnson

THU-SES1-O3: PITCH AND GLOTTAL-WAVEFORM ESTIMATION AND MODELING II

Pre- and Short-Term Posttreatment Vocal Functioning in Patients with Advanced Head and Neck Cancer Treated with Concomitant Chemoradiotherapy................................................................................................. 2590
Irene Jacobs, Lisette Van Der Molen, Maya Van Rossum, Frans Higges

Acoustic Analysis of Infection in Parkinson's Disease ........................................................................................................ 2594
Joa K. T. Mo, Rutger Hoffmann

SAFE: A Statistical Algorithm for F0 Estimation for Both Clean and Noisy Speech ........................................... 2598
Wei Chi, Abier Alwah

Robust and Efficient Pitch Estimation Using an Iterative ARMA Technique .......................................................... 2602
Jung Ook Hong, Patrick J. Wolfe
Statistical Modeling of FO Dynamics in Singing Voices Based on Gaussian Processes with Multiple Oscillation Bases

Yasunori Ohishi, Hirokazu Kameoka, Daichi Mochihashi, Hidehisa Nagano, Kazuo Kashino

Applying Geometric Source Separation for Improved Pitch Extraction in Human-Robot Interaction

Martin Heckensta, Claudius Glaser, Frank Joublin, Kazuhiro Nakadai

THU-SES1-Q4: ASR: FEATURE EXTRACTION II

A Spectral I.F. Model Based Approach to Voice Source Parameterisation

John Kane, Mark Kane, Christian Gohb

Glotto-Based Analysis of the Lombard Effect

Thomas Drangnag, Thierry Dutot

Hidden Logistic Linear Regression for Support Vector Machine Based Phone Verification

Bo Li, Kae Chai Sim

Jointly Optimized Discriminative Features for Speech Recognition

Tim Pen, Bing Zhang, Long Nguyen

Invariant Integration Features Combined with Speaker-Adaptation Methods

Florian Muller, Alfred Merlins

Multi Resolution Discriminative Models for Subvocalic Speech Recognition

Mark Raujas, Vinak Kumar Rangarajan Sridhar, Rohit Prasad, Praveen Natarajan

THU-SES1-P1: SPEAKER DIARIZATION

A Comparative Large Scale Study of MLP Features for Mandarin ASR

Fabio Valente, Muthu Makhir Doss, Christian Mohls, Saman V. Kavari, Wen Wang

Recognizing Cochlear Implant-Like Spectrally Reduced Speech with HMM-Based ASR: Experiments with MFCCs and PLP Coefficients

Cong-Thanh Do, Denique Pasier, Gao Le Lan, Andre Goalic

An Integrated Top-Down/Bottom-Up Approach to Speaker Diarization

Simon Biaouene, Nicolas Evans, Christienn Freidell

A Hybrid Approach to Online Speaker Diarization

Carlo Tauscher, Oriol Vinyals, Gerald Friedland

Advances in Fast Multistream Diarization Based on the Information Bottleneck Framework

Gisela Garau, Alfred Dielmann, Herve Bourlard

An Improved Cluster Model Selection Method for Agglomerative Hierarchical Speaker Clustering Using Incremental Gaussian Mixture Models

Ky H. Huan, Skarthwaith N. Narayanam

Audio-Visual Synchronisation for Speaker Diarisation

Gisela Garau, Alfred Dielmann, Herve Bourlard

Audio-Visual Anticipatory Coarticulation Modeling by Human and Machine

Louis H. Terry, Karen Livstein, Janet B. Diermeckers, Aggelis E. Katagiannis

Impact of Lack of Acoustic Feedback in EMG-Based Silent Speech Recognition

Manthar Janke, Michael Wand, Tanja Schultz

Using Prosody to Improve Mandarin Automatic Speech Recognition

Chong-Jia Ni, Wenju Liu, Bo Xu

Speaker Tracking in an Unsupervised Speech Controlled System

Tobias Herbig, Prasa Guel, Wolfgang Minker

MultiBIC: An Improved Speaker Segmentation Technique for TV Shows

Paul Lopez-Otero, Laura Dario-Fernandez, Carmen Garcia-Maeso

Automatic Speech Recognition for Assistive Writing in Speech Supplemented Word Prediction

Nakay Karpov, Andrey Roshin, Konstantin Markov, Milos Zelazny

Wisme-Dependent Weight Optimization for CHMM-Based Audio-Visual Speech Recognition

Matthew Janke, Michael Wand, Tanja Schultz

Efficient Manycore CHMM Speech Recognition for Audiovisual and Multistream Data

Dorothea Kohla, Jike Chong, Steffen Zeiler, Kurt Keutzer
### THU-SES1-P3: SPEAKER AND LANGUAGE RECOGNITION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takami Yoshida, Kazuhiro Nakadai</td>
<td></td>
</tr>
<tr>
<td>Non-Audible Murmur Recognition Based on Fusion of Audio and Visual Streams</td>
<td>2714</td>
</tr>
<tr>
<td>Panumon Herangonua, Naruhito Hagaia</td>
<td></td>
</tr>
<tr>
<td>Improved N-Gram Phonotactic Models for Language Recognition</td>
<td>2718</td>
</tr>
<tr>
<td>Mohamed Fouzi Benromdhana, Jean-Luc Gauvain, Lori Lamel</td>
<td></td>
</tr>
<tr>
<td>A Study of Term Weighting in Phonotactic Approach to Spoken Language Recognition</td>
<td>2722</td>
</tr>
<tr>
<td>Sirinoot Roosna, Donglai Zhu, Bivu Ma, Atitwong Suchan, Phoonprasop Panyaburana, Nattanan Theptithakkhal, Chai Watwattanarat</td>
<td></td>
</tr>
<tr>
<td>Exploiting Context-Dependency and Acoustic Resolution of Universal Speech Attribute Models in Spoken Language Recognition</td>
<td>2726</td>
</tr>
<tr>
<td>Sabato Marco Siniscalchi, Jeremy Reed, Torbjorn Sundsen, Chito-Hui Lee</td>
<td></td>
</tr>
<tr>
<td>Hierarchical Multilayer Perceptron Based Language Identification</td>
<td>2730</td>
</tr>
<tr>
<td>David Intang, Matthew Mogenti Doss, Herve Boudard</td>
<td></td>
</tr>
<tr>
<td>The NIST 2010 Speaker Recognition Evaluation</td>
<td>2734</td>
</tr>
<tr>
<td>Avin F. Martin, Craig S. Greenberg</td>
<td></td>
</tr>
<tr>
<td>Bayesian Speaker Recognition Using Gaussian Mixture Model and Laplace Approximation</td>
<td>2738</td>
</tr>
<tr>
<td>Shih-Sian Cheng, J-Fan Chen, Hei-Min Wang</td>
<td></td>
</tr>
<tr>
<td>What Else is New Than the Hamming Window? Robust MFCCs for Speaker Recognition via Multitapering</td>
<td>2742</td>
</tr>
<tr>
<td>Torni Kinnunen, Rahim Saeidi, Johan Sandberg, Merla Hanssen-Sandsten</td>
<td></td>
</tr>
<tr>
<td>Fast Computation of Speaker Characterization Vector Using MLR and Sufficient Statistics in Anchor Model Framework</td>
<td>2746</td>
</tr>
<tr>
<td>A. K. Sarkar, S. Umesh</td>
<td></td>
</tr>
<tr>
<td>Graph-Embedding for Speaker Recognition</td>
<td>2750</td>
</tr>
<tr>
<td>Zahi N. Karam, W. M. Campbell</td>
<td></td>
</tr>
<tr>
<td>A Hybrid Modeling Strategy for GMM-SVM Speaker Recognition with Adaptive Relevance Factor</td>
<td>2754</td>
</tr>
<tr>
<td>Chang Hui You, Heichou Li, Song Anf Lee</td>
<td></td>
</tr>
</tbody>
</table>

### THU-SES1-P4: SOURCE LOCALIZATION AND SEPARATION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust Mixture Modeling Using T-Distribution: Application to Speaker ID</td>
<td>2758</td>
</tr>
<tr>
<td>Harshavardhan S., T. V. Sreenuvva</td>
<td></td>
</tr>
<tr>
<td>A Variable Frame Length and Rate Algorithm Based on the Spectral Kurtosis Measure for Speaker Verification</td>
<td>2762</td>
</tr>
<tr>
<td>Chi-Sang Song, Kyo-J. Han, Hyun-Soon Seo, Seok-Dum S. Narayanan, Hang-Goo Kang</td>
<td></td>
</tr>
<tr>
<td>Near Field Sound Source Localization Based on Cross-Power Spectrum Phase Analysis with Multiple Microphones</td>
<td>2766</td>
</tr>
<tr>
<td>Rohel Hayashi, Masanori Murase, Takamasa Nishimura</td>
<td></td>
</tr>
<tr>
<td>A Maximum a Posteriori Sound Source Localization in Reverberant and Noisy Conditions</td>
<td>2770</td>
</tr>
<tr>
<td>Jinho Choi, Chong D. Yoo</td>
<td></td>
</tr>
<tr>
<td>A DOA Estimation Algorithm Based on Equalization-Cancellation Theory</td>
<td>2774</td>
</tr>
<tr>
<td>Dae Toh Chai, Junfeng Li, Maasao Akiya</td>
<td></td>
</tr>
<tr>
<td>Concurrent Speaker Localization Using Multi-Band Position-Pitch (M-PoPi) Algorithm with Spectro-Temporal Pre-Processing</td>
<td>2778</td>
</tr>
<tr>
<td>Tania Hall, Harald Roosdhofer</td>
<td></td>
</tr>
<tr>
<td>On Using Gaussian Mixture Model for Double-Talk Detection in Acoustic Echo Suppression</td>
<td>2782</td>
</tr>
<tr>
<td>Ji-Hyun Song, Kyu-Ho Lee, Yun-Seok Park, Sang-Ick Kang, Jeon-Hyuk Chang</td>
<td></td>
</tr>
<tr>
<td>Catalog-Based Single-Channel Speech-Music Separation</td>
<td>2786</td>
</tr>
<tr>
<td>Cedal Domir, A. Toymen Cemgil, Murali Sarackar</td>
<td></td>
</tr>
<tr>
<td>Unvoiced Speech Segregation Based on CASA and Spectral Subtraction</td>
<td>2790</td>
</tr>
<tr>
<td>Ke Hu, Deliang Wang</td>
<td></td>
</tr>
</tbody>
</table>

### THU-SES2-S1: SPECIAL SESSION: INTERSPEECH 2010 PARALINGUISTIC CHALLENGE

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsupervised Sequential Organization for Cochannel Speech Separation</td>
<td>2794</td>
</tr>
<tr>
<td>Ke Hu, Deliang Wang</td>
<td></td>
</tr>
<tr>
<td>The INTERSPEECH 2010 Paralinguistic Challenge</td>
<td>2798</td>
</tr>
<tr>
<td>Bjorn Schiller, Stefan Steidl, Anton Badiner, Felix Burkhardt, Laurence Devillers, Christian Muller, Shrikanth S. Narayanan</td>
<td></td>
</tr>
<tr>
<td>Age and Gender Classification from Speech Using Decision Level Fusion and Ensemble Based Techniques</td>
<td>2802</td>
</tr>
<tr>
<td>Florian Lingenfahler, Johannes Wagner, Thorid Pfitz, Jonghwa Kim, Elisabeth Andre</td>
<td></td>
</tr>
<tr>
<td>Level of Interest Sensing in Spoken Dialog Using Multi-Level Fusion of Acoustic and Lexical Evidence</td>
<td>2806</td>
</tr>
<tr>
<td>Jie Han Jiao, Rui Xia, Yong Liu</td>
<td></td>
</tr>
<tr>
<td>Fuzzy Support Vector Machines for Age and Gender Classification</td>
<td>2810</td>
</tr>
<tr>
<td>Phuc Nguyen, Trung Le, Du Tran, Xu Huang, Dharmendra Sharma</td>
<td></td>
</tr>
<tr>
<td>Gender and Affect Recognition Based on GMM and GMM-UBM Modeling with Relevance MAP Estimation</td>
<td>2814</td>
</tr>
<tr>
<td>Rick Gujik, James Zibert, Tadej Justins, Viviene Struc, Basjan Vucicic, France Mihelic</td>
<td></td>
</tr>
<tr>
<td>Age Recognition Based on Speech Signals Using Weights Supervector</td>
<td>2818</td>
</tr>
<tr>
<td>Royi Perot, Dan Lange, Yanli Zigel</td>
<td></td>
</tr>
</tbody>
</table>
Age and Gender Classification Using Fusion of Acoustic and Prosodic Features
Hugo Meinedo, Isabel Trancoso
Brno University of Technology System for Interspeech 2010 Paralinguistic Challenge

Combining Five Acoustic Level Modeling Methods for Automatic Speaker Age and Gender Recognition
Ming Li, Chi-Sang Jung, Kyu J. Han

Age and Gender Recognition Based on Multiple Systems — Early vs. Late Fusion
Tobias Boeckel, Georg Stemmer, Viktor Zetzsche, Elmar Noth

THU-SES2-O1: NEW PARADIGMS IN ASR II
Automated Speaker Age and Gender Recognition in the Car for Tailoring Dialog and Mobile Services
Michael Feld, Felix Burkhardt, Christian Muller

Improved Topic Classification and Keyword Discovery Using an HMM-Based Speech Recognizer Trained Without Supervision

An Analysis of Sparseness and Regularization in Exemplar-Based Methods for Speech Classification
Dimitri Kenerer, Tara N. Sainath, Bhavana Ramakrishnan, David Nahamoo

Investigation of Full-Sequence Training of Deep Belief Networks for Speech Recognition
Abdel-Rahman Mohamed, Dong Yu, L. Deng

Mandarin Tone Recognition Using Affine-Invariant Prosodic Features and Tone Posteriorgram
Yow-Bang Wang, Lin-Shan Lee

Continuous Speech Recognition with a TF-IDF Acoustic Model
Geoffrey Zweig, Patrick Nguyen, Jasha Droppo, Alex Acero

THU-SES2-O2: SPOKEN LANGUAGE UNDERSTANDING AND SPOKEN LANGUAGE TRANSLATION II
SCARF: A Segmental Conditional Random Field Toolkit for Speech Recognition
Geoffrey Zweig, Patrick Nguyen

Role of Language Models in Spoken Fluency Evaluation

Social Role Discovery from Spoken Language Using Dynamic Bayesian Networks
Sibel Yaman, Dilrak Hakani-Tur, Gokhan Tur

Domain Adaptation and Compensation for Emotion Detection
Michelle Hewitt Sanchez, Gokhan Tur, Luciana Ferrer, Dilrak Hakani-Tur

Phrase Alignment Confidence for Statistical Machine Translation
Sankaranarayanan Ananthakrishnan, Rohit Prasad, Prem Narayanan

Named-Entity Projection and Data-Driven Morphological Decomposition for Field Maintainable Speech-to-Speech Translation Systems
Ian R. Leme, Alex Waibel

THU-SES2-O3: SIGNAL PROCESSING FOR MUSIC AND SONG
Acoustic Correlates of Voice Quality Improvement by Voice Training
Kyoaki Aikawa, Junko Uemama, Tomoko Akita

Phonetic Segmentation of Singing Voice Using MIDI and Parallel Speech
Meigui Dong, Paul Chan, Jing Cen, Hutchinson Li, Jason Tao, Ping Jin Xa

A Singing Style Modeling System for Singing Voice Synthesizers
Kazuko Sato, Masato Tachibana, Hideki Kawasaki

A Fast Query by Humming System Based on Notes
Jing X. Yang, Jie Liu, Wei-Qiang Zhang

Melody Pitch Estimation Based on Range Estimation and Candidate Extraction Using Harmonic Structure Model
Seokhwan Jo, Shyin Ju, Chang D. Yoo

Modified Spatial Audio Object Coding Scheme with Harmonic Extraction and Elimination Structure for Interactive Audio Service
Jhoon Park, Kwangil Kim, Jeongil Seo, Minsoo Hahn

THU-SES2-O4: MODELING FIRST LANGUAGE ACQUISITION
Modelling the Effect of Speaker Familiarity and Noise on Infant Word Recognition
Christina Bergmann, Michele Gablon, Lou Boves

Unsupervised Learning of Vowels from Continuous Speech Based on Self-Organized Phoneme Acquisition Model
Koaki Miyazawa, Hideo Coto, Reiko Matsuoka

Learning Speaker Normalization Using Semisupervised Manifold Alignment
Andrew R. Plummer, Mary E. Beckman, Mikhail Belkin, Eric Foster-Luisset, Benjamin Munson
Fully Unsupervised Word Learning from Continuous Speech Using Transitional Probabilities of Atomic Acoustic Events

Okko Johannes Rasanen

Language Acquisition and Cross-Modal Associations: Computational Simulation of the Result of Infant Studies

L. Ten Bosch, Lou Boves

Active Word Learning Under Uncertain Input Conditions

Maarten Versleegh, L. Ten Bosch, Lou Boves

THU-SES2-P1: ASR: ACOUSTIC MODELS III

Parallel Training of Neural Networks for Speech Recognition

Karel Patočky, Lukas Berger, František Grezl

The Use of Sense in Unsupervised Training of Acoustic Models for ASR Systems

Ruo Singh, Benjamin Lambert, Bhiksha Raj

Boosted Mixture Learning of Gaussian Mixture HMMs for Speech Recognition

Jan Du, Yu Hu, Hui Jiang

On the Exploitation of Hidden Markov Models and Linear Dynamic Models in a Hybrid Decoder Architecture for Continuous Speech Recognition

Yalter Levent, Reinhold Hirsch-Uslebach

Context Dependent Modelling Approaches for Hybrid Speech Recognizers

Alberto Amed, Thomas Pellegrini, Isabel Trancoso, Jose Neto

A Regularized Discriminative Training Method of Acoustic Models Derived by Minimum Relative Entropy Discrimination

Yutaro Koba, Shinji Watanabe, Atsushi Nakamura, Tetsuomi Kobayashi

Decision Tree State Clustering with Word and Syllable Features

Hank Liu, Chris Albright, Michael Zuechianti, Glorion Sihan

A Duration Modeling Technique with Incremental Speech Rate Normalization

Haruki Fujimori, Takashi Mano, Mito Oyamada, Tsuchida

Long Short-Term Memory Networks for Noise Robust Speech Recognition

Martin Wolfmann, Yong San, Florian Eyben, Bjorn Schuller

One-Model Speech Recognition and Synthesis Based on Articulatory Movement HMMs

Tran Anh, Takashi Onaka, Masahiko Kimura, Yurie Iriha, Koichi Kato

Acoustic Modeling with Bootstrap and Restructuring for Low-Resourced Languages

Xiaolong Cui, Jian Xue, Pierre L. Dogmin, Upendra V. Chaudhari, Bowen Zhou

Lecture Speech Recognition by Combining Word Graphs of Various Acoustic Models

Tetsu Konaka, Katsuki Goto, Takashi Ito, Masaharu Kato

Semi-Parametric Trajectory Modelling Using Temporally Varying Feature Mapping for Speech Recognition

Khe Choi Sun, Shihin Liu

Deep-Structured Hidden Conditional Random Fields for Phonetic Recognition

Deep Supervised Learning for Improved Expression of Uncertainty in Discriminative Classifiers

Jonathan Maldon, Jeff Ibarretxe

Modeling Posterior Probabilities Using the Linear Exponential Family

Peder Olsen, Valihova Gaoi, Charles Micchelli, John R. Hershey

THU-SES2-P2: SPOKEN DIALOGUE SYSTEMS II

New Technique to Enhance the Performance of Spoken Dialogue Systems Based on Dialogue States-Dependent Language Models and Grammatical Rules

Ramon Lopez-Cazorla, David Girod

A Stochastic Finite-State Transducer Approach to Spoken Dialog Management

Lino-F. Harizides, Joaquin Planelles, Encarna Sagrip, Emilio Santula, David Girod

Enhanced Monitoring Tools and Online Dialogue Optimisation Merged into a New Spoken Dialogue System Design Experience

Romain Laroch, Philippe Breier, Ghislain Parent

Optimising a Handcrafted Dialogue System Design

Romain Laroch, Ghislain Parent, Philippe Breier

Utterance Selection for Speech Acts in a Cognitive Tourguide Scenario

Felix Putze, Tanja Schiltz

Lexical Entrainment of Real Users in the Let's Go Spoken Dialog System

Gabriel Parent, Maxime Ekenel

Combining User Intention and Error Modeling for Statistical Dialog Simulators

Silvia Quaronee, Meritxell Gonzalez, Giuseppe Riccardi, Sebastian Vargas

Parallel Processing of Interruptions and Feedback in Companions Affective Dialogue System

Jindra Hablanov, Markku Tammari, Ruon Sama De La Cima, Nigel Creek

Dynamic Language Modeling Using Bayesian Networks for Spoken Dialog Systems

Antoine Roas, Neville Mehta, Dinesh Ramachandran, Rakesh Gupta

Automatic Detection of Task-Incomplete Dialog for Spoken Dialog System Based on Dialog Act N-Gram

Satos Hara, Norihide Kiioka, Kazuya Tomono
Wei-Bin Liang, Chung-Hsien Wu, Yu-Cheng Hsiao
Detection of Hot Spots in Poster Conversations Based on Reactive Tokens of Audience
Tatsuya Kanakura, Kosuke Sumi, Zhe-Qing Chang, Kanayo Takahashi
Psychological Evaluation of a Group Communication Activation Robot in a Party Game
Yoichi Matsumiya, Shinpei Fujie, Hikaru Taniyama, Tetsunori Kobayashi
Analyzing User Utterances in Barge-in-Able Spoken Dialogue System for Improving Identification Accuracy
Kyoko Matsumiya, Katsunori Komatani, Ryo Tokeda, Tetsu Takahashi, Tatsuya Opote, Hitoshi G. Okuno
Pitch Similarity in the Vicinity of Backchannels
Matthias Heidler, Jens Edlund, Julia Hirschberg
A Rule-Based Backchannel Prediction Model Using Pitch and Pause Information
Khai P. Truong, Ronald Poppe, Dirk Heylen

THU-SES2-P3: DISCOURSE AND DIALOGUE
Combining Text Categorization and Dialog Modeling for Speaker Role Identification on Call Center Conversations
Remi Lavalley, Chloe Clavel, Patrick Brilet, Marc El-Base
Topic-Dependent N-Gram Models Based on Optimization of Context Lengths in LDA
Akira Nakamura, Satoshi Hayami
Expectations for Discourse Genre Identification: A Prosodic Study
Nicolas Ohin, Volker Dalbo, Anne Lachere, Xavier Rodel
Dialogue Act Tagging and Segmentation with a Single Perceptron
Ramon Granell, Stephen Pulman, Carlos D. Marques-Huergas, Jose Miguel Benedi
Improving the Readability of Class Lecture ASR Results Using a Confusion Network
Yamhisa Fujii, Katsumaru Yamamoto, Setichi Nakagawa

THU-SES2-P4: VOICE ACTIVITY AND TURN DETECTION
Toward Detecting Voice Activity Employing Soft Decision in Second-Order Conditional MAP
Song-Kyun Kim, Jeong-Chun Cho, Song-Ick Kang, Ji-Hyun Song, Joon-Hyuk Chung
Voice Activity Detection in a Regularized Reproducing Kernel Hilbert Space
Xianggong Lu, Masahiro Usuki, Ryo Yoneke Isitami, Hitoshi Komori, Satoshi Nakamura
A New VAD Framework Using Statistical Model and Human Knowledge Based Empirical Rule
Ji Wu, Xue-Lai Zhang, Wei Li
Adaptive High Accuracy Approaches to Speech Activity Detection in Noisy and Hostile Audio Environments
Mark Huggins, Brett Smolenski, Aaron Lawson
Robust Voice Activity Detection in Stereo Recording with Crosstalk
Prasanta Kumar Ghosh, Andreas Tsiriruas, Panayiota G. Georgiou, Shrikanth S. Narayanan
Voice Activity Detection Using Frame-Wise Model Re-Estimation Method Based on Gaussian Pruning with Weight Normalization
Masakiyo Fujimoto, Shinya Watanabe, Tomohiro Nakatani
Spectral Entropy-Based Voice Activity Detector for Videoconferencing Systems
Kwon Lee, Debarkha Mukherjee
The QUT-NOISE-TIMIT Corpus for the Evaluation of Voice Activity Detection Algorithms
David Dean, Sridha Sridharan, Robert Vogt, Michael Masin
A Bayesian Approach to Voice Activity Detection Using Multiple Statistical Models and Discriminative Training
Tao Yu, John H. L. Hansen
Noise Robust Voice Activity Detection Using Features Extracted from the Time-Domain Autocorrelation Function
Houman Osniangham, Brendan Baker, Robert Vogt, Sridha Sridharan
VAD-Measure-Embedded Decoder with Online Model Adaptation
Tanaka Otsuki, Koji Iwano, Sadashki Murai
Robust Statistical Voice Activity Detection Using a Likelihood Ratio Sign Test
Shiwen Deng, Jingqin Han
Automatic Turn Segmentation in Spoken Conversations
Alex V. Ivanov, Giuseppe Riccardi
Turn Taking-Based Conversation Detection by Using DOA Estimation
Yoko Kenmochi, Masahito Yagoami, Yasunari Obuchi

Author Index