2012 8th International Symposium on Chinese Spoken Language Processing

(ISCSLP 2012)

Kowloon Tong, China
5 – 8 December 2012

IEEE

IEEE Catalog Number: CFP12583-PRT
ISBN: 978-1-4673-2506-6

O1 : Automatic Speech Recognition

1 **Alternative Hypothesis Generation Using a Weighted Kernel Feature Matrix for ASR Substitution Error Correction**
(Chao-Hong Liu, Chung-Hsien Wu, David Sarwono)

6 **Speaker-Ensemble Hidden Markov Modeling for Automatic Speech Recognition**
(Guoli Ye, Brian Mak)

11 **A Synchronized Pruning Composition Algorithm of Weighted Finite State Transducers for Large Vocabulary Speech Recognition**
(Zhiyang He, Ping Lv, Wei Li, Ji Wu)

16 **Context Dependant Phone Mapping for Cross-Lingual Acoustic Modeling**
(Van Hai Do, Xiong Xiao, Eng Siong Chng, Haizhou Li)

21 **A Comparative Study of fMPE and RDLT Approaches to LVCSR**
(Jian Xu, Zhi-Jie Yan, Qiang Huo)

O2 : Phonetics, Phonology, Linguistics, and Speech Analysis

25 **A Cross-Dialect Comparison of Vowel Dispersion and Vowel Variability**
(Wai-Sum Lee)

30 **Analyzing Semantic Orientation of Terms Using Affinity Propagation**
(Yan Li, Si Li, Weiran Xu, Jun Guo)

35 **Effects of Excitation Spread on the Intelligibility of Mandarin Speech in Cochlear Implant Simulations**
(Fei Chen, Tian Guan, Lena L.N. Wong)

40 **Acoustic and Articulatory Analysis on Japanese Vowels in Emotional Speech**
(Mengxue Cao, Aijun Li, Qiang Fang, Jianguo Wei, Chan Song, Jianwu Dang)

45 **Articulatory and Spectral Characteristics of Cantonese Vowels**
(Wai-Sum Lee)

O3 : Speech Synthesis and Voice Conversion

50 **Exploring Mutual Information for GMM-Based Spectral Conversion**
(Hsin-Te Hwang, Yu Tsao, Hsin-Min Wang, Yih-Ru Wang, Sin-Horng Chen)

55 **Incorporating Dynamic Features into Minimum Generation Error Training for HMM-Based Speech Synthesis**
(Duy Khanh Ninh, Masanori Morise, Yoichi Yamashita)

60 **Cross Validation and Minimum Generation Error for Improved Model Clustering in HMM-Based TTS**
(Feng-Long Xie, Yi-Jian Wu, Frank K. Soong)

64 **Perceptual Clustering Based Unit Selection Optimization for Concatenative Text-to-Speech Synthesis**
(Tao Jiang, Zhiyong Wu, Jia Jia, Lianhong Cai)

69 **Voice Conversion Using Bayesian Mixture of Probabilistic Linear Regressions and Dynamic Kernel Features**
(Na Li, Yu Qiao)

O4 : Speech Synthesis

74 **Effective Sentence Selection Based on Phone/Model Coverage Maximization for Speaker Adaptation in HMM-Based Speech Synthesis**
(Cheng Hsien Lin, Po Kai Huang, Cheng-Yuan Lin, Chih-Chung Kuo)

79 **Hierarchical Prosodic Pattern Selection Based on Fujisaki Model for Natural Mandarin Speech Synthesis**
(Yi-Chin Huang, Chung-Hsien Wu, Sz-Ting Weng)

84 **Cross-Stream Dependency Modeling Using Continuous F0 Model for HMM-Based Speech Synthesis**
(Xin Wang, Zhen-Hua Ling, LiRong Dai)

88 **Resonance-Based Spectral Deformation in HMM-Based Speech Synthesis**
(Jinfu Ni, Yoshinori Shiga, Hisashi Kawai, Hideki Kashioka)

93 **Detection and Emphatic Realization of Contrastive Word Pairs for Expressive Text-to-Speech Synthesis**
(Chunrong Li, Zhiyong Wu, Fanbo Meng, Helen M. Meng, Lianhong Cai)

ISCSLP 2012 Table of Contents

O5 : General Topics on ASR

98 **Spoken Term Detection for OOV Terms Based on Triphone Confusion Matrix**
(Yong Xu, Wu Guo, Shan Su, LiRong Dai)

103 **Hierarchical Clustering and Robust Identification for Block-Based Autoregressive Speech Parameter Estimation**
(Ruofei Chen, Cheung-Fat Chan)

108 **Phonotactic Spoken Language Recognition: Using Diversely Adapted Acoustic Models in Parallel Phone Recognizers**
(Cheung-Chi Leung, Bin Ma, Haizhou Li)

112 **A New Confidence Measure Combining Hidden Markov Models and Artificial Neural Networks of Phonemes for Effective Keyword Spotting**
(Su Jun Leow, Tze Siong Lau, Alvina Goh, Han Meng Peh, Teck Khim Ng, Sabato Marco Siniscalchi, Chin-Hui Lee)

117 **Two Objective Measures for Speech Distortion and Noise Reduction Evaluation of Enhanced Speech Signals**
(Huijun Ding, Tan Lee, Ing Yann Soon)

O6: Robust Speech Recognition

122 **Synthesized Stereo-Based Stochastic Mapping with Data Selection for Robust Speech Recognition**
(Jun Du, Qiang Huo)

126 **TDOA Information Based VAD for Robust Speech Recognition in Directional and Diffuse Noise Field**
(Kuan-Lang Huang, Tai-Shih Chi)

131 **An Analysis of Vector Taylor Series Model Compensation for Non-Stationary Noise in Speech Recognition**
(Duc Hoang Ha Nguyen, Xiong Xiao, Eng Siong Chng, Haizhou Li)

136 **Structured Modeling Based on Generalized Variable Parameter HMMs and Speaker Adaptation**
(Yang Li, Xunying Liu, Lan Wang)

141 **A Study on Cepstral Sub-Band Normalization for Robust ASR**
(Syu-Siang Wang, Jeih-weih Hung, Yu Tsao)

P1: Speech Synthesis, Pronunciation Learning, and Language Modeling

146 **Statistical Modification Based Post-Filtering Technique for HMM-Based Speech Synthesis**
(Zhengqi Wen, Jianhua Tao, Hao Che)

150 **A Study of F0 Modelling and Generation with Lyrics and Shape Characterization for Singing Voice Synthesis**
(S.W. Lee, Minghui Dong, Haizhou Li)

155 **Experiments on Unsupervised Statistical Parametric Speech Synthesis**
(Jinfu Ni, Yoshinori Shiga, Hisashi Kawai, Hideki Kashioka)

160 **Improved Unit Selection Speech Synthesis Method Utilizing Subjective Evaluation Results on Synthetic Speech**
(Xian-Jun Xia, Zhen-Hua Ling, Chen-Yu Yang, LiRong Dai)

165 **A Unified Trajectory Tiling Approach to High Quality TTS and Crosslingual Voice Transformation**
(Yao Qian, Frank K. Soong)

170 **mENUNCIATE: Development of a Computer-Aided Pronunciation Training System on a Cross-Platform Framework for Mobile, Speech-Enabled Application Development**
(Pengfei Liu, Ka-Wa Yuen, Wai-Kim Leung, Helen M. Meng)

174 **Analysis on Mispronunciations in CAPT Based on Computational Speech Perception**
(Jia Jia, Wai-Kim Leung, Ye Tian, Lianhong Cai, Helen M. Meng)

179 **Perceptually-Motivated Assessment of Automatically Detected Lexical Stress in L2 Learners' Speech**
(Kun Li, Helen M. Meng)

184 **Improve Mispronunciation Detection with Tandem Feature**
(Hua Yuan, Junhong Zhao, Jia Liu)

P1 continued…

188 **Bayesian Nonparametric Language Models**
(Ying-Lan Chang, Jen-Tzung Chien)

193 **Phrase-Based Data Selection for Language Model Adaptation in Spoken Language Translation**
(Shixiang Lu, Wei Wei, Xiaoyin Fu, Lichun Fan, Bo Xu)

197 **Collecting Sentences from Web Resources for Constructing Spontaneous Chinese Language Model**
(Xinhui Hu, Youzheng Wu, Shigeki Matsuda, Chiori Hori, Hideki Kashioka)

ISCSLP 2012 Table of Contents

P2 : Speech Enhancement and Robust ASR

201 **Controlling the Tradeoff Property in a Regularization Framework for Noise Reduction**
(Xugang Lu, Masashi Unoki, Shigeki Matsuda, Chiori Hori, Hideki Kashioka)

206 **A Fast Two-Microphone Noise Reduction Algorithm Based on Power Level Ratio for Mobile Phone**
(Jian Zhang, Risheng Xia, Zhonghua Fu, Junfeng Li, Yonghong Yan)

210 **The Lossless Adaptive Arithmetic Coding Based on Context for ITU-T G.719 at Variable Rate**
(Xuan Ji, Jing Wang, Hailong He, Jingming Kuang)

215 **Unified Denoising and Dereverberation Method Used in Restoration of MTF-Based Power Envelope**
(Masashi Unoki, Xugang Lu)

220 **Noise-Robust Whispered Speech Recognition Using a Non-Audible-Murmur Microphone with VTS Compensation**
(Chen-Yu Yang, Georgina Brown, Liang Lu, Junichi Yamagishi, Simon King)

224 **Power-Normalized PLP (PNPLP) Feature for Robust Speech Recognition**
(Lichun Fan, Dengfeng Ke, Xiaoyin Fu, Shixiang Lu, Bo Xu)

229 **A Feature-Transform Based Approach to Unsupervised Task Adaptation and Personalization**
(Jian Xu, Zhi-Jie Yan, Qiang Huo)

233 **Keyword-Specific Normalization Based Keyword Spotting for Spontaneous Speech**
(Weifeng Li, Qingmin Liao)

238 **Enhanced Lengthening Cancellation Using Bidirectional Pitch Similarity Alignment for Spontaneous Speech**
(Po-Yi Shih, Bo-Wei Chen, Jhing-Fa Wang, Jhing-Wei Wu)

P3 : Speech Prosody

243 **Information Allocation and Prosodic Expressiveness in Continuous Speech: A Mandarin Cross-Genre Analysis**
(Chiu-yu Tseng, Chao-yu Su)

247 **Automatic Pitch Accent Detection Using Auto-Context with Acoustic Features**
(Junhong Zhao, Wei-Qiang Zhang, Hua Yuan, Jia Liu, Shanhong Xia)

252 **An Improved Tone Labeling and Prediction Method with Non-Uniform Segmentation of F0 Contour**
(Xingyu Na, Xiang Xie, Jingming Kuang, Yaling He)

256 **Break Index Labeling of Mandarin Text via Syntactic-to-Prosodic Tree Mapping**
(Xiaotian Zhang, Yao Qian, Hai Zhao, Frank K. Soong)

261 **Prosody-Based Sentence Boundary Detection in Chinese Broadcast News**
(Lei Xie, Chenglin Xu, Xiaoxuan Wang)

266 **Pitch Accent Detection and Prediction with DCT Features and CRF Model**
(Wenping Hu, Yao Qian, Frank K. Soong)

271 **More Targets? Simulating Emotional Intonation of Mandarin with Penta**
(Aijun Li, Qiang Fang, Yuan Jia, Jianwu Dang)

N/A **Research on the Pitch Contour and Prosodic Phrase in Mongolian**
(Aomin, Ziyu Xiong, Bayarmend Borjigin)

281 **A Syllable-Based Prosody Modeling for L1 and L2 English Speeches**
(Wei-Fan Chen, Chin-Kuan Kuo, Yih-Ru Wang, Sin-Horng Chen)

286 **A Simple and Effective Pitch Re-Estimation Method for Rich Prosody and Speaking Styles in HMM-Based Speech Synthesis**
(Cheng-Yuan Lin, Chien-Hung Huang, Chih-Chung Kuo)

ISCSLP 2012 Table of Contents

P3 continued…

291 **Diachronic Contrastive Analysis on Read Speech in Broadcast News: Evidence from Pitch and Duration**
(Yu Zou, Yan Wang, Wei He)

296 **Phonetic Realization of Accent from Chinese English Learners in Various Dialectal Regions**
(Yuan Jia, Aijun Li)

P4 : Speech and Speaker Recognition

301 **Investigation of Deep Neural Networks (DNN) for Large Vocabulary Continuous Speech Recognition: Why DNN Surpasses GMMs in Acoustic Modeling**
(Jia Pan, Cong Liu, Zhiguo Wang, Yu Hu, Hui Jiang)

306 **An Improved Steady Segment Based Decoding Algorithm by Using Response Probability for LVCSR**
(Zhanlei Yang, Wenju Liu, Hao Chao)

311 **Acoustic Space Partition Based on Broad Phonetic Class for Ensemble Acoustic Modeling**
(Xugang Lu, Yu Tsao, Shigeki Matsuda, Chiori Hori, Hideki Kashioka)

315 **A Study on Cross-Language Knowledge Integration in Mandarin LVCSR**
(Chen-Yu Chiang, Sabato Marco Siniscalchi, Yih-Ru Wang, Sin-Horng Chen, Chin-Hui Lee)

320 **Minimum Phone Error Model Training on Merged Acoustic Units for Transcribing Bilingual Code-Switched Speech**
(Ching-Feng Yeh, Yiu-Chang Lin, Lin-Shan Lee)

325 **Acoustic Modeling for Native and Non-Native Mandarin Speech Recognition**
(Xin Chen, Jian Cheng)

330 **Intra-Conversation Intra-Speaker Variability Compensation for Speaker Clustering**
(Kui Wu, Yan Song, Wu Guo, LiRong Dai)

335 **Alleviating the Small Sample-Size Problem in I-Vector Based Speaker Verification**
(Wei Rao, Man-Wai Mak)

340 **Text-Dependent Speaker Recognition with Long-Term Features Based on Functional Data Analysis**
(Chenhao Zhang, Thomas Fang Zheng, Ruxin Chen)

ISCSLP 2012 Table of Contents

P4 continued…

345 **Efficient Feature Extraction of Speaker Identification Using Phoneme Mean F-Ratio for Chinese**
(Chen Zhao, Hongcui Wang, Songgun Hyon, Jianguo Wei, Jianwu Dang)

349 **Discriminant Local Information Distance Preserving Projection for Text-Independent Speaker Recognition**
(Liang He, Jia Liu)

353 **Acoustic Analysis of Disguised Voices with Raised and Lowered Pitch**
(Cuiling Zhang)

P5 : General Topics in Spoken Language Processing

358 **Boundary-Expanding Locality Sensitive Hashing**
(Qiang Wang, Zhiyuan Guo, Gang Liu, Jun Guo)

363 **Adaptive Named Entity Recognition Based on Conditional Random Fields with Automatic Updated Dynamic Gazetteers**
(Xixin Wu, Zhiyong Wu, Jia Jia, Lianhong Cai)

368 **Nesting Hierarchical Phrase-Based Model for Speech-to-Speech Translation**
(Xiaoyin Fu, Wei Wei, Lichun Fan, Shixiang Lu, Bo Xu)

373 **A Phone Segmentation Method and Its Evaluation on Mandarin Speech Corpus**
(Dac-Thang Hoang, Hsiao-Chuan Wang)

378 **A Hybrid Fragment/Syllable-Based System for Improved OOV Term Detection**
(Yong Xu, Wu Guo, LiRong Dai)

383 **Tongue Shape Synthesis Based on Active Shape Model**
(Chan Song, Jianguo Wei, Qiang Fang, Shen Liu, Yuguang Wang, Jianwu Dang)

387 **Perceptual Similarity Between Audio Clips and Feature Selection for Its Measurement**
(Qinghua Wu, Xiaolei Zhang, Ping Lv, Ji Wu)

392 **Self Documentation of Endangered Languages**
(Sagun Dhakhwa, Jens Allwood)

396 **Reconstruction of Vocal Tract Based on Multi-Source Image Information**
(Song Wang, Shen Liu, Jianguo Wei, Qiang Fang, Jianwu Dang)

400 **Robust Voice Activity Detection Using Empirical Mode Decomposition and Modulation Spectrum Analysis**
(Yasuaki Kanai, Masashi Unoki)

405 **A Real-Time Tone Enhancement Method for Continuous Mandarin Speeches**
(Ye Tian, Jia Jia, Yongxin Wang, Lianhong Cai)

ISCSLP 2012 Table of Contents

P6 : Phonetics and Phonology; Speech Production and Perception

409 **A Preliminary Study on the Interlanguage Speech Intelligibility Benefit for English-Mandarin Bilingual L2 Learners**
(Guo Li, Peggy Mok)

413 **Detailed Morphological Analysis of Mandarin Sustained Steady Vowels**
(Yuguang Wang, Hongcui Wang, Jiaqi Gao, Jianguo Wei, Jianwu Dang)

417 **Effects of Carriers on Mandarin Tone Categorical Perception**
(Dazuo Wang, Xiuxiu Wang, Gang Peng)

422 **Tones in Whispered Mandarin**
(Bin Li, Rong Rong)

426 **A Study on the Coarticulation of Bi-Syllabic Words in Chinese**
(Maolin Wang, Shengnan Xiong, Jiayun Li, Ziyu Xiong)

431 **A Comparative Study of Perception of Tone 2 and Tone 3 in Mandarin by Native Speakers and Japanese Learners**
(Ting Zou, Jinsong Zhang, Wen Cao)

436 **A Preliminary Investigation of the Third Tone Sandhi in Standard Chinese with a Prosodic Corpus**
(Hongwei Ding, Daniel Hirst)

440 **Locus of Orthographic Facilitation Effect in Spoken Word Production: Evidence from Cantonese Chinese**
(I-Fan Su, Sin-Ting Yeung, Brendan S. Weekes, Sam-Po Law)

445 **The Temporal Effect of Speaking Rate, Focus and Prosody in Chinese**
(Maolin Wang, Wei Shi, Ruixian Huang, Ziyu Xiong)

P6 continued…

450 **How to Describe Speech Emotion More Completely — An Investigation on Chinese Broadcast News Speech**
(Yingying Gao, Weibin Zhu)

454 **The Coarticulation Resistance of Consonants in Standard Chinese — An Electropalatographic and Acoustic Study**
(Yinghao Li, Jinghua Zhang, Jiangping Kong)