Table of Contents

Simultaneous Acoustic, Prosodic, and Phrasing Model Training for TTs Conversion Systems ......................... 1  
Keiichiro Oura, Yoshihiko Nankaku, Tomoki Toda, Keiichi Tokuda, Rannierry Maia, Shinsuke Sakai and S. Nakamura

Cross-Stream Dependency Modeling for HMM-Based Speech Synthesis .............................................................. 5  
Zhen-Hua Ling, Wei Zhang and Ren-Hua Wang

Cross-Lingual Speaker Adaptation for HMM-Based Speech Synthesis ................................................................. 9  
Yi-Jian Wu, Simon King and Keiichi Tokuda

HMM-Based Mixed-Language (Mandarin-English) Speech Synthesis .............................................................. 13  
Yao Qian, Houwei Cao and Frank K. Soong

Improving HMM Based Speech Synthesis by Reducing Over-Smoothing Problems ............................................ 17  
Meng Zhang, Jianhua Tao, Huibin Jia and Xia Wang

Pronunciation Space Models for Pronunciation Evaluation .................................................................................... 21  
Si Wei, Yi-Qian Pan, Guo-Ping Hu, Yu Hu and Ren-Hua Wang

Decision Fusion for Improving Mispronunciation Detection Using Language Transfer Knowledge and Phoneme-Dependent Pronunciation Scoring .............................................................. 25  
W. K. Lo, Alissa M. Harrison, Helen Meng and Lan Wang

Mandarin Learning Using Speech and Language Technologies: A Translation Game in the Travel Domain ............. 29  
Yushi Xu and Stephanie Seneff

Word Order Correction for Language Transfer Using Relative Position Language Modeling .............................. 33  
Chao-Hong Liu, Chung-Hsien Wu and Matthew Harris

Improving Automatic Evaluation of Mandarin Pronunciation with Speaker Adaptive Training (SAT) and MLLR Speaker Adaption .............................................................. 37  
Chao Huang, Feng Zhang and Frank K. Soong

Automatic Assessment of Language Proficiency through Shadowing .............................................................. 41  
Dean Luo, Nobuaki Minematsu, Yutaka Yamauchi and Keikichi Hirose

What’s in the F0 of Mandarin Speech: Tones, Intonation and Beyond .............................................................. 45  
Chiu-Yu Tseng and Zhao-Yu Su

A Perceptual Study of Approximated Cantonese Tone Contours ........................................................................ 49  
Yujia Li and Tan Lee

A New Prosodic Strength Calculation Method for Prosody Reduction Modeling .................................................. 53  
Honglei Cong, Zhiyong Wu, Lianhong Cai and Helen M. Meng

Prosody Study with Context-Dependent Acoustic Models .................................................................................... 57  
Yue-Ning Hu and Min Chu

Intonational Prominence of "SHI...(DE)" Construction in Standard Chinese .......................................................... 61  
Yuan Jia, Aijun Li and Ziyu Xiong

Entropy-Based Analysis of the Prosodic Features of Chinese Dialects .............................................................. 65  
Raymond W. M. Ng and Tan Lee
Improvements on Mel-Frequency Cepstrum Minimum-Mean-Square-Error Noise Suppressor for Robust Speech Recognition .................................................................69  
Dong Yu, Li Deng, Jian Wu, Yifan Gong and Alex Acero

Effect of Feature Smoothing for Robust Speech Recognition ......................................................73  
Xiong Xiao, Eng Siong Chng and Haizhou Li

Reference Eigen-Environment and Speaker Weighting for Robust Speech Recognition ..................77  
Yuan-Fu Liao, Hung-Hsiang Fang and Chih-Min Yang

Evaluation of a Feature Compensation Approach Using High-Order Vector Taylor Series Approximation of an Explicit Distortion Model for Aurora2, Aurora3, and Aurora4 Tasks ....................81  
Jun Du, Qiang Huo and Yu Hu

Deriving MFCC Parameters from the Dynamic Spectrum for Robust Speech Recognition ...........85  
Nengheng Zheng, Xia Li, Houwei Cao, Tan Lee and P. C. Ching

Discriminative Output Coding Features for Speech Recognition ..............................................89  
Omid Dehzangi, Bin Ma, Eng Siong Chng and Haizhou Li

Investigation on Adaptation Using Different Discriminative Training Criteria Based Linear Regression and Map .................................................................93  
Bo Zhu, Zhi-Jie Yan, Yu Hu, Zhi-Guo Wang, Li-Rong Dai and Ren-Hua Wang

Utilization of Huge Written Text Corpora for Conversational Speech Recognition ..................97  
Xinhui Hu, Hirofumi Yamamoto, Jinsong Zhang, Keiji Yasuda, Youzheng Wu and Hideki Kashioka

Position Information for Language Modeling in Speech Recognition .......................................101  
Hsuan-Sheng Chiu, Guan-Yu Chen, Chun-Jen Lee and Berlin Chen

An Investigation of Phonological Feature Systems Used in Detection-Based ASR ......................105  
I-Fan Chen and Hsin-Min Wang

An HMM Compensation Approach for Dynamic Features Using Unscented Transformation and its Application to Noisy Speech Recognition .........................109  
Yu Hu and Qiang Huo

Mandarin Language Understanding in Dialogue Context ..........................................................113  
Yushi Xu, Jingjing Liu and Stephanie Seneff

Frequency Modulation Technique for Prosodic Modification ....................................................117  
Jinfu Ni, Shinsuke Sakai, T. Shimizu and S. Nakamura

Modeling and Generating Tone Contour with Phrase Intonation for Mandarin Chinese Speech .......121  
Zhizheng Wu, Yao Qian, Frank K. Soong and Bo Zhang

A Three-Stage Text Normalization Strategy for Mandarin Text-to-Speech Systems .................125  
Tao Zhou, Yuan Dong, Dezhi Huang, Wu Liu and Haila Wang

Multi-Layer F0 Modeling for HMM-Based Speech Synthesis ..................................................129  
Cheng-Cheng Wang, Zhen-Hua Ling, Bu-Fan Zhang and Li-Rong Dai

Predicting Spectral and Prosodic Parameters for Unit Selection in Speech Synthesis ...............133  
Minghui Dong and Haizhou Li
Heteronym Verification for Mandarin Speech Synthesis ................................................................. 137
Heng Lu, Zhenhua Ling, Si Wei, Yu Hu, Lirong Dai and Renhua Wang

The Improved TS-BASE Approaches with Interference Compensation and Their Evaluations for
Speech Enhancement ................................................................................................................. 141
Junfeng Li, Shuichi Sakamoto, Satoshi Hongo, Masato Akagi and Yoiti Suzuki

Pitch Tracking for Model-Based Speech Separation ................................................................. 145
S. W. Lee, Frank K. Soong, P. C. Ching and Tan Lee

Improved Linear Discriminant Analysis Considering Empirical Pairwise Classification Error Rates .... 149
Hung-Shin Lee and Berlin Chen

CityBrowser II: A Multimodal Restaurant Guide in Mandarin .................................................. 153
Jingjing Liu, Yushi Xu, Stephanie Seneff and Victor Zue

Evaluation and Analysis of Minimum Phone Error Training and its Modified Versions for Large
Vocabulary Mandarin Speech Recognition .................................................................................. 157
Yung-Jen Cheng, Che-Kuang Lin and Lin-Shan Lee

A Two-Stage Algorithm for Multi-Speaker Identification System ........................................... 161
Yong Guan and Wen-Ju Liu

Double Gauss Based Unsupervised Score Normalization in Speaker Verification ......................... 165
Wu Guo, Li-Rong Dai and Ren-Hua Wang

Discriminative Feedback Adaptation for GMM-UBM Speaker Verification ................................. 169
Yi-Hsiang Chao, Wei-Ho Tsai and Hsin-Min Wang

Using Pseudo-Key for Language Recognition System Design .................................................. 173
Hanwu Sun, Bin Ma and Haizhou Li

Self-Organized Clustering for Feature Mapping in Language Recognition ................................. 177
Chang Huai You, Kong Aik Lee, Bin Ma and Haizhou Li

An Efficient Feature Selection Method for Speaker Recognition ............................................... 181
Hanwu Sun, Bin Ma and Haizhou Li

PLSA Based Topic Mixture Language Modeling Approach ....................................................... 185
Shuanhu Bai and Haizhou Li

A Trellis Based Fast Lattice Generating Algorithm ..................................................................... 189
Wei Li, Ji Wu and Zhiguo Wang

Order Adaptation of the Fractional Fourier Transform Using the Intraframe Pitch Change Rate for
Speech Recognition .................................................................................................................... 193
Hui Yin, Climent Nadeu, Volker Hohmann, Xiang Xie and Jingming Kuang

Large Vocabulary Continuous Speech Recognition in Uyghur: Data Preparation and
Experimental Results .................................................................................................................. 197
Nasirjan Tursun and Wushour Silamu

An Improvement for Training Efficiency of Semi-Tied Covariance ........................................... 201
Si-Bao Chen, Yu Hu, Bin Luo and Ren-Hua Wang
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved Semi-Parametric Mean Trajectory Model Using Discriminatively Trained Centroids</td>
<td>205</td>
</tr>
<tr>
<td>Ran Xu, Jielin Pan and Yonghong Yan</td>
<td></td>
</tr>
<tr>
<td>Local Mismatch Phone for Confidence Measure in Standard and Accented Chinese Speech Recognition</td>
<td>209</td>
</tr>
<tr>
<td>Wenxiao Cao, Yi Liu and Thomas Fang Zheng</td>
<td></td>
</tr>
<tr>
<td>A Combined Task Analysis Method for Data Selection in Mandarin Isolated Word Recognition System</td>
<td>213</td>
</tr>
<tr>
<td>Z. Y. He, Z. G. Wang, W. Li and J. Wu</td>
<td></td>
</tr>
<tr>
<td>Mandarin Speech Recognition for Nonnative Speakers Based on Pronunciation Dictionary Adaptation</td>
<td>217</td>
</tr>
<tr>
<td>Jian Yang, Peishan Wu and Dan Xu</td>
<td></td>
</tr>
<tr>
<td>A New Similarity Measure Between HMMS</td>
<td>221</td>
</tr>
<tr>
<td>Yih-Ru Wang</td>
<td></td>
</tr>
<tr>
<td>Recognition of Syllable-Contracted Words in Spontaneous Speech Using Word Expansion and Duration Information</td>
<td>225</td>
</tr>
<tr>
<td>Wei-Bin Liang, Chung-Hsien Wu and Yu-Kai Kang</td>
<td></td>
</tr>
<tr>
<td>Exploiting Non-Target Region Information for Confidence Measure Based on Bayesian Information Criterion</td>
<td>229</td>
</tr>
<tr>
<td>Cong Liu, Yu Hu, Xiong-Guo Lei, Zhi-Guo Wang, Li-Rong Dai and Ren-Hua Wang</td>
<td></td>
</tr>
<tr>
<td>Prosodic Modeling for Isolated Mandarin Words and its Application</td>
<td>233</td>
</tr>
<tr>
<td>Hung-Kuang Shih, Chen-Yu Chiang, Yih-Ru Wang and Sin-Horng Chen</td>
<td></td>
</tr>
<tr>
<td>A CSI and Rate-Distortion Based Packet Loss Recovery Algorithm for VoIP</td>
<td>237</td>
</tr>
<tr>
<td>Zhongbo Li, Shenghui Zhao, Jing Wang and Jingming Kuang</td>
<td></td>
</tr>
<tr>
<td>Mandarin Stops Classification Based on Random Forest Approach</td>
<td>241</td>
</tr>
<tr>
<td>Chi-Yueh Lin and Hsiao-Chuan Wang</td>
<td></td>
</tr>
<tr>
<td>A Pitch Synchronous Method for Speech Modification</td>
<td>245</td>
</tr>
<tr>
<td>Chih-Ting Kuo and Hsiao-Chuan Wang</td>
<td></td>
</tr>
<tr>
<td>Speech Database Compacted for an Embedded Mandarin TTS System</td>
<td>249</td>
</tr>
<tr>
<td>Qing Guo, Bin Wang and Nobuyuki Katae</td>
<td></td>
</tr>
<tr>
<td>Prosody Modification on Mixed-Language Speech Synthesis</td>
<td>253</td>
</tr>
<tr>
<td>Yi Zhang and Jianhua Tao</td>
<td></td>
</tr>
<tr>
<td>A Maximum Entropy Based Hierarchical Model for Automatic Prosodic Boundary Labeling in Mandarin</td>
<td>257</td>
</tr>
<tr>
<td>Fangzhou Liu, Huibin Jia and Jianhua Tao</td>
<td></td>
</tr>
<tr>
<td>Tone Evaluation of Chinese Continuous Speech Based on Prosodic Words</td>
<td>261</td>
</tr>
<tr>
<td>Yi-Qian Pan, Si Wei and Ren-Hua Wang</td>
<td></td>
</tr>
<tr>
<td>The Pitch Analysis of Imperative Sentences in Standard Chinese</td>
<td>265</td>
</tr>
<tr>
<td>Jia Sun, Jilun Lu, Aijun Li and Yuan Jia</td>
<td></td>
</tr>
<tr>
<td>Word Alignment Based on Multi-Grain Model</td>
<td>269</td>
</tr>
<tr>
<td>Yanqing He, Yu Zhou and Chengqing Zong</td>
<td></td>
</tr>
</tbody>
</table>
Word Reordering Alignment for Combination of Statistical Machine Translation Systems ........................................... 273
Maoxi Li and Chengqing Zong

An EMD Based Approach to Transliteration Unit Alignment between English and Chinese ........................................... 277
Muyun Yang, Shujie Liu, Lei Wang, Sheng Li, Jufeng Li, Teijun Zhao and Haoliang Qi

Analysis and Modeling of Affective Audio Visual Speech Based on PAD Emotion Space ........................................... 281
Shen Zhang, Yingjin Xu, Jia Jia and Lianhong Cai

Noise Reduction Based Random Matrix Theory ........................................................................................................... 285
X. Lu, S. Matsuda, T. Shimizu and S. Nakamura

Language Model Adaptation for Relevance Feedback in Information Retrieval .......................................................... 289
Ying-Lang Chang and Jen-Tzung Chien

Predicting and Tagging Dialog-Act Using MDP and SVM .......................................................................................... 293
Keyan Zhou, Chengqing Zong, Hua Wu and Haifeng Wang

A Synchronous Method for Automatic Scoring of Language Learning ........................................................................ 297
Bin Dong and Yonghong Yan

Using Reference to Tune Language Model for Detection of Reading Miscues ............................................................ 302
Changliang Liu, Fuping Pan, Fengpei Ge, Bin Dong and Yonghong Yan

How Syllables Group in Chinese .................................................................................................................................. 306
Maolin Wang and Yi Xu

Simplified Deformation Compensation for Emotional Speaker Recognition ................................................................. 310
Yingchun Yang, Tian Wu and Hongbing Lv

Interfusing the Confused Region Score of Speaker Verification Systems ................................................................. 314
Yanhua Long, Wu Guo and Lirong Dai

Parallel Phone Recognizer based MLLR Speaker Recognition ....................................................................................... 318
Eryu Wang, Wu Guo and Lirong Dai

Eigenchannel Compensation and Symmetric Score for Robust Text-Independent Speaker Verification .......................... 322
Yuan Dong, Jian Zhao, Liang Lu, Jiqing Lui, Xianyu Zhao and Haila Wang

A Sample and Feature Selection Scheme for GMM-SVM Based Language Recognition ............................................ 326
Yan Song and Lirong Dai

Speaker Recognition using a Kind of Novel Phonotactic Information ........................................................................ 330
Xiang Zhang, Xiang Xiao, Haipeng Wang, Hongbin Suo, Qingwei Zhao and Yonghong Yan

The Adaptation Schemes In PR-SVM Based Language Recognition ........................................................................... 334
Bing Xu, Yan Song and Lirong Dai

Mandarin Tone Perception with Temporal Envelope and Periodicity Cues from Different Frequency Regions .......................... 338
Meng Yuan, Tan Lee and Sigfrid D. Soli

Prosodic Variation in Cantonese-English Code-Mixed Speech ...................................................................................... 342
Wentao Gu, Tan Lee and P. C. Ching
Pronunciation Error Detection for Computer Assisted Pronunciation Teaching in Mandarin ........................................... 346
Min-Siong Liang, Jian-Yung Hung, Ren-Yuan Lyu and Yuang-Chin Chiang

A Two-Stage Multi-Feature Integration Approach to Unsupervised Speaker Change Detection in Real-Time News Broadcasting .................................................................................................. 350
Lei Xie and Guangsen Wang

Automatic Prosody Boundary Labeling of Mandarin Using Both Text and Acoustic Information ............................................ 354
Chongjia Ni, Wenju Liu and Bo Xu

Subword Latent Semantic Analysis for Texttiling-Based Automatic Story Segmentation of Chinese Broadcast News ......................................................................................................................................... 358
Yulian Yang and Lei Xie

Multipitch Detection Based on Weighted Summary Correlogram ........................................................................................ 362
Xueliang Zhang, Wenju Liu, Peng Li and Bo Xu

Efficient System Combination for Syllable-Confusion-Network-Based Chinese Spoken Term Detection ...... 366
Jie Gao, Qingwei Zhao, Yonghong Yan and Jian Shao

The Use of Dynamic Deformable Templates for Lip Tracking in an Audio-Visual Corpus with Large Variations in Head Pose, Face Illumination and Lip Shapes ................................................................................................................................. 370
Zhiyong Wu, Jiying Wu and Helen M. Meng

Microphone Array Post-Filter Based on Auditory Filtering ................................................................................................. 374
Peng Li, Fengchai Liao, Ning Cheng, Bo Xu and Wenju Liu

Exploring Tone Variations in Chinese Dialects Using Context Dependent Tone Models .................................................. 378
Wei Guo and Min Chu

Author Index .......................................................................................................................................................................................... follows page 381