# Table of Contents

## Plenary

Interactive Computer Aids for Acquiring Proficiency in Mandarin  
*Stephanie Seneff*  
1

The Affective and Pragmatic Coding of Prosody  
*Klaus R. Scherer*  
13

Challenges in Machine Translation  
*Franz Josef Och*  
15

Automatic Indexing and Retrieval of Large Broadcast News Video Collections - The TRECVID Experience  
*Tat-Seng Chua*  
16

## Tutorial

An HMM-Based Approach to Flexible Speech Synthesis  
*Keiichi Tokuda*  
17

Text Information Extraction and Retrieval  
*Hang Li*  
18

## Topics in Speech Science

Mechanisms of Question Intonation in Mandarin  
*Jiahong Yuan*  
19

Comparison of Perceived Prosodic Boundaries and Global Characteristics of Voice Fundamental Frequency Contours in Mandarin Speech  
*Wentao Gu, Keikichi Hirose, Hiroya Fujisaki*  
31

Linguistic Markings of Units in Spontaneous Mandarin  
*Shu-Chuan Tseng*  
43

Phonetic and Phonological Analysis of Focal Accents of Disyllabic Words in Standard Chinese  
*Yuan Jia, Ziyu Xiong, Aijun Li*  
55
# Table of Contents

Focus, Lexical Stress and Boundary Tone: Interaction of Three Prosodic Features ........................................... 67  
Lu Zhang, Yi-Qing Zu, Run-Qiang Yan

## Speech Analysis

A Robust Voice Activity Detection Based on Noise Eigenspace Projection ......................................................... 76  
Dongwen Ying, Yu Shi, Frank Soong, Jianwu Dang, Xugang Lu

Pitch Mean Based Frequency Warping ............................................. 87  
Jian Liu, Thomas Fang Zheng, Wenhu Wu

A Study of Knowledge-Based Features for Obstruent Detection and Classification in Continuous Mandarin Speech ........................................... 95  
Kuang-Ting Sung, Hsiao-Chuan Wang

Speaker-and-Environment Change Detection in Broadcast News Using Maximum Divergence Common Component GMM ........................................... 106  
Yih-Ru Wang

UBM Based Speaker Segmentation and Clustering for 2-Speaker Detection ......................................................... 116  
Jing Deng, Thomas Fang Zheng, Wenhu Wu

Design of Cubic Spline Wavelet for Open Set Speaker Classification in Marathi ............................................. 126  
Hemant A. Patil, T.K. Basu

## Speech Synthesis and Generation

Rhythmic Organization of Mandarin Utterances — A Two-Stage Process ......................................................... 138  
Min Chu, Yunjia Wang

Prosodic Boundary Prediction Based on Maximum Entropy Model with Error-Driven Modification ........................................... 149  
Xiaonan Zhang, Jun Xu, Lianhong Cai

Prosodic Words Prediction from Lexicon Words with CRF and TBL Joint Method ............................................. 161  
Heng Kang, Wenju Liu
Prosodic Word Prediction Using a Maximum Entropy Approach .......... 169
  Honghui Dong, Jianhua Tao, Bo Xu

Predicting Prosody from Text ............................................ 179
  Keh-Jiann Chen, Chiu-yu Tseng, Chia-hung Tai

Nonlinear Emotional Prosody Generation and Annotation ................. 189
  Jianhua Tao, Jian Yu, Yongguo Kang

A Unified Framework for Text Analysis in Chinese TTS .................. 200
  Guohong Fu, Min Zhang, GuoDong Zhou, Kang-Kuong Luke

Speech Synthesis Based on a Physiological Articulatory Model .......... 211
  Qiang Fang, Jianwu Dang

An HMM-Based Mandarin Chinese Text-To-Speech System .................. 223
  Yao Qian, Frank Soong, Yining Chen, Min Chu

HMM-Based Emotional Speech Synthesis Using Average Emotion Model ......................................................... 233
  Long Qin, Zhen-Hua Ling, Yi-Jian Wu, Bu-Fan Zhang,
  Ren-Hua Wang

A Hakka Text-To-Speech System ........................................... 241
  Hsiu-Min Yu, Hsin-Te Hwang, Dong-Yi Lin, Sin-Horng Chen

Speech Enhancement

Adaptive Null-Forming Algorithm with Auditory Sub-bands .............. 248
  Heng Zhang, Qiang Fu, Yonghong Yan

Multi-channel Noise Reduction in Noisy Environments .................. 258
  Junfeng Li, Masato Akagi, Yōiti Suzuki

Acoustic Modeling for Automatic Speech Recognition

Minimum Phone Error (MPE) Model and Feature Training
  on Mandarin Broadcast News Task ........................................ 270
  Jia-Yu Chen, Chia-Yu Wan, Yi Chen, Berlin Chen, Lin-shan Lee

State-Dependent Phoneme-Based Model Merging for Dialectal Chinese
Speech Recognition ......................................................... 282
  Linquan Liu, Thomas Fang Zheng, Wenhu Wu
Non-uniform Kernel Allocation Based Parsimonious HMM ........................ 294
   Peng Liu, Jian-Lai Zhou, Frank Soong

Consistent Modeling of the Static and Time-Derivative Cepstrums
for Speech Recognition Using HSPTM ........................................... 303
   Yiu-Pong Lai, Man-Hung Siu

Robust Speech Recognition

Vector Autoregressive Model for Missing Feature Reconstruction ......... 315
   Xiong Xiao, Haizhou Li, Eng-Siong Chng

Auditory Contrast Spectrum for Robust Speech Recognition ................. 325
   Xugang Lu, Jianwu Dang

Signal Trajectory Based Noise Compensation for Robust Speech
Recognition ................................................................................. 335
   Zhi-Jie Yan, Jian-Lai Zhou, Frank Soong, Ren-Hua Wang

An HMM Compensation Approach Using Unscented Transformation
for Noisy Speech Recognition ...................................................... 346
   Yu Hu, Qiang Huo

Noisy Speech Recognition Performance of Discriminative HMMs .......... 358
   Jun Du, Peng Liu, Frank Soong, Jian-Lai Zhou, Ren-Hua Wang

Distributed Speech Recognition of Mandarin Digits String .................. 370
   Yih-Ru Wang, Bo-Xuan Lu, Yuan-Fu Liao, Sin-Horng Chen

Speech Adaptation/Normalization

Unsupervised Speaker Adaptation Using Reference Speaker
Weighting ...................................................................................... 380
   Tsz-Chung Lai, Brian Mak

Automatic Construction of Regression Class Tree for MLLR Via
Model-Based Hierarchical Clustering ............................................. 390
   Shih-Sian Cheng, Yeong-Yuh Xu, Hsin-Min Wang, Hsin-Chia Fu

General Topics in Speech Recognition

A Minimum Boundary Error Framework for Automatic Phonetic
Segmentation .................................................................................. 399
   Jen-Wei Kuo, Hsin-Min Wang
Large Vocabulary Continuous Speech Recognition

Advances in Mandarin Broadcast Speech Transcription at IBM Under the DARPA GALE Program ....................................................... 410
  Yong Qin, Qin Shi, Yi Y. Liu, Hagai Aronowitz, Stephen M. Chu,
  Hong-Kwang Kuo, Geoffrey Zweig

Improved Large Vocabulary Continuous Chinese Speech Recognition by Character-Based Consensus Networks .......................... 422
  Yi-Sheng Fu, Yi-Cheng Pan, Lin-shan Lee

All-Path Decoding Algorithm for Segmental Based Speech Recognition ................................................................. 435
  Yun Tang, Wenju Liu, Bo Xu

Improved Mandarin Speech Recognition by Lattice Rescoring with Enhanced Tone Models ......................................................... 445
  Huanliang Wang, Yao Qian, Frank Soong, Jian-Lai Zhou, Jiqing Han

On Using Entropy Information to Improve Posterior Probability-Based Confidence Measures ......................................................... 454
  Tzan-Hwei Chen, Berlin Chen, Hsin-Min Wang

Vietnamese Automatic Speech Recognition: The FLaVoR Approach ........... 464
  Quan Vu, Kris Demuynck, Dirk Van Compernolle

Multilingual Recognition and Identification

Language Identification by Using Syllable-Based Duration Classification on Code-Switching Speech ............................................ 475
  Dau-cheng Lyu, Ren-yuan Lyu, Yuang-chin Chiang, Chun-nan Hsu

Speaker Recognition and Characterization

CCC Speaker Recognition Evaluation 2006: Overview, Methods, Data, Results and Perspective ......................................................... 485
  Thomas Fang Zheng, Zhanjiang Song, Lihong Zhang,
  Michael Brasser, Wei Wu, Jing Deng

The IIR Submission to CSLP 2006 Speaker Recognition Evaluation ........ 494
  Kong-Aik Lee, Hanwu Sun, Rong Tong, Bin Ma, Minghui Dong,
  Changhuai You, Donglai Zhu, Chin-Wei Eugene Koh, Lei Wang,
  Tomi Kinnunen, Eng-Siong Chng, Haizhou Li
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Novel Alternative Hypothesis Characterization Using Kernel Classifiers for LLR-Based Speaker Verification</td>
<td>506</td>
</tr>
<tr>
<td>Yi-Hsiang Chao, Hsin-Min Wang, Ruei-Chuan Chang</td>
<td></td>
</tr>
<tr>
<td>Speaker Verification Using Complementary Information from Vocal Source and Vocal Tract</td>
<td>518</td>
</tr>
<tr>
<td>Nengheng Zheng, Ning Wang, Tan Lee, P.C. Ching</td>
<td></td>
</tr>
<tr>
<td>ISCSLP SR Evaluation, UVA-CS.cs.es System Description. A System Based on ANNs</td>
<td>529</td>
</tr>
<tr>
<td>Carlos E. Vivaracho</td>
<td></td>
</tr>
<tr>
<td>Evaluation of EMD-Based Speaker Recognition Using ISCSLP2006 Chinese Speaker Recognition Evaluation Corpus</td>
<td>539</td>
</tr>
<tr>
<td>Shingo Kuroiwa, Satoru Tsuge, Masahiko Kita, Fuji Ren</td>
<td></td>
</tr>
<tr>
<td>Integrating Complementary Features with a Confidence Measure for Speaker Identification</td>
<td>549</td>
</tr>
<tr>
<td>Nengheng Zheng, P.C. Ching, Ning Wang, Tan Lee</td>
<td></td>
</tr>
<tr>
<td>Discriminative Transformation for Sufficient Adaptation in Text-Independent Speaker Verification</td>
<td>558</td>
</tr>
<tr>
<td>Hao Yang, Yuan Dong, Xianyu Zhao, Jian Zhao, Haila Wang</td>
<td></td>
</tr>
<tr>
<td>Fusion of Acoustic and Tokenization Features for Speaker Recognition</td>
<td>566</td>
</tr>
<tr>
<td>Rong Tong, Bin Ma, Kong-Aik Lee, Changhuai You, Donglai Zhu, Tomi Kinnunen, Hanwu Sun, Minghui Dong, Eng-Siong Chng, Haizhou Li</td>
<td></td>
</tr>
<tr>
<td>Spoken Language Understanding</td>
<td></td>
</tr>
<tr>
<td>Contextual Maximum Entropy Model for Edit Disfluency Detection of Spontaneous Speech</td>
<td>578</td>
</tr>
<tr>
<td>Jui-Feng Yeh, Chung-Hsien Wu, Wei-Yen Wu</td>
<td></td>
</tr>
<tr>
<td>Human Language Acquisition, Development and Learning</td>
<td></td>
</tr>
<tr>
<td>Automatic Detection of Tone Mispronunciation in Mandarin</td>
<td>590</td>
</tr>
<tr>
<td>Li Zhang, Chao Huang, Min Chu, Frank Soong, Xianda Zhang, Yudong Chen</td>
<td></td>
</tr>
<tr>
<td>Towards Automatic Tone Correction in Non-native Mandarin</td>
<td>602</td>
</tr>
<tr>
<td>Mitchell Peabody, Stephanie Seneff</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Spoken and Multimodal Dialog Systems</td>
<td></td>
</tr>
<tr>
<td>A Corpus-Based Approach for Cooperative Response Generation in a</td>
<td>614</td>
</tr>
<tr>
<td>Dialog System</td>
<td></td>
</tr>
<tr>
<td>Zhiyong Wu, Helen Meng, Hui Ning, Sam C. Tse</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>A Cantonese Speech-Driven Talking Face Using Translingual Audio-to-</td>
<td>627</td>
</tr>
<tr>
<td>Visual Conversion</td>
<td></td>
</tr>
<tr>
<td>Lei Xie, Helen Meng, Zhi-Qiang Liu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>The Implementation of Service Enabling with Spoken Language of a</td>
<td>640</td>
</tr>
<tr>
<td>Multi-modal System Ozone</td>
<td></td>
</tr>
<tr>
<td>Sen Zhang, Yves Laprie</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoken Correction for Chinese Text Entry</td>
<td>648</td>
</tr>
<tr>
<td>Bo-June Paul Hsu, James Glass</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech Data Mining and Document Retrieval</td>
<td></td>
</tr>
<tr>
<td>Extractive Chinese Spoken Document Summarization Using Probaibotic</td>
<td>660</td>
</tr>
<tr>
<td>Ranking Models</td>
<td></td>
</tr>
<tr>
<td>Yi-Ting Chen, Suhan Yu, Hsin-Min Wang, Berlin Chen</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Meeting Segmentation Using Two-Layer Cascaded Subband Filters</td>
<td>672</td>
</tr>
<tr>
<td>Manuel Giuliani, Tin Lay Nwe, Haizhou Li</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>A Multi-layered Summarization System for Multi-media Archives by</td>
<td>683</td>
</tr>
<tr>
<td>Understanding and Structuring of Chinese Spoken Documents</td>
<td></td>
</tr>
<tr>
<td>Lin-shan Lee, Sheng-yi Kong, Yi-cheng Pan, Yi-sheng Fu, Yu-tsun</td>
<td></td>
</tr>
<tr>
<td>Huang, Chien-chih Wang</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Experiments on Automatic Story Segmentation in Chinese</td>
<td>693</td>
</tr>
<tr>
<td>Spoken Documents Using Lexical Cohesion of Extracted Named Entities</td>
<td></td>
</tr>
<tr>
<td>Devon Li, Wai-Kit Lo, Helen Meng</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Machine Translation of Speech</td>
<td></td>
</tr>
<tr>
<td>Some Improvements in Phrase-Based Statistical Machine Translation</td>
<td>704</td>
</tr>
<tr>
<td>Zhendong Yang, Wei Pang, Jinhua Du, Wei Wei, Bo Xu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic Spoken Language Translation Template Acquisition Based on</td>
<td>712</td>
</tr>
<tr>
<td>Boosting Structure Extraction and Alignment</td>
<td></td>
</tr>
<tr>
<td>Rile Hu, Xia Wang</td>
<td></td>
</tr>
</tbody>
</table>
Spoken Language Resources and Annotation

HKUST/MTS: A Very Large Scale Mandarin Telephone Speech Corpus ........................................... 724
   Yi Liu, Pascale Fung, Yongsheng Yang, Christopher Cieri, Shudong Huang, David Graff

The Paradigm for Creating Multi-lingual Text-To-Speech Voice Databases ........................................ 736
   Min Chu, Yong Zhao, Yining Chen, Lijuan Wang, Frank Soong

Multilingual Speech Corpora for TTS System Development ..................................................... 748
   Hsi-Chun Hsiao, Hsiu-Min Yu, Yih-Ru Wang, Sin-Horng Chen

Construct Trilingual Parallel Corpus on Demand ........................................................................ 760
   Muyun Yang, Hongfei Jiang, Tiejyun Zhao, Sheng Li

The Contribution of Lexical Resources to Natural Language Processing of CJK Languages ................... 768
   Jack Halpern

Multilingual Spoken Language Corpus Development for Communication Research ....................... 781
   Toshiyuki Takezawa

Development of Multi-lingual Spoken Corpora of Indian Languages ........................................... 792
   K. Samudravijaya

Author Index ............................................................................................................................. 803