

Communications in Computer and Information Science

807

Commenced Publication in 2007

Founding and Former Series Editors:

Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu, Dominik Ślęzak,
and Xiaokang Yang

Editorial Board

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Phoebe Chen

La Trobe University, Melbourne, Australia

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation of the Russian
Academy of Sciences, St. Petersburg, Russia*

Krishna M. Sivalingam

Indian Institute of Technology Madras, Chennai, India

Takashi Washio

Osaka University, Osaka, Japan

Junsong Yuan

Nanyang Technological University, Singapore, Singapore

Lizhu Zhou

Tsinghua University, Beijing, China

More information about this series at <http://www.springer.com/series/7899>

Jianhua Tao · Thomas Fang Zheng
Changchun Bao · Dong Wang
Ya Li (Eds.)

Man-Machine Speech Communication

14th National Conference, NCMMSC 2017
Lianyungang, China, October 11–13, 2017
Revised Selected Papers

Editors

Jianhua Tao
Institute of Automation
Chinese Academy of Sciences
Beijing
China

Thomas Fang Zheng
Computer Science and Technology
Tsinghua University
Beijing
China

Changchun Bao
Beijing University of Technology
Beijing
China

Dong Wang
Tsinghua University
Beijing
China

Ya Li
Institute of Automation
Chinese Academy of Sciences
Beijing
China

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-981-10-8110-1 ISBN 978-981-10-8111-8 (eBook)
<https://doi.org/10.1007/978-981-10-8111-8>

Library of Congress Control Number: 2018930522

© Springer Nature Singapore Pte Ltd. 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

NCMMSC (the National Conference on Man–Machine Speech Communication) is organized by the Speech Information Committee of the Chinese Information Processing Society and co-organized by the Chinese Phonetics Association under CLP (Chinese Language Society) as well as the Language, Audio, and Music Acoustic Committee under ASC (Acoustical Society of China). Representing an important stage for experts, researchers, and practitioners to share their ideas, NCMMSC significantly promotes research and technical innovations in the corresponding fields. The papers in these proceedings mainly address challenging issues in speech recognition and enhancement, speaker and language recognition, speech synthesis, corpus and phonetic in speech technology, speech generation, speech analyzing and modelling, speech processing of ethnic minorities, speech emotion recognition, and audio signal processing among others.

This year, NCMMSC received 133 submissions, including Chinese and English submissions. After a thorough reviewing process, 98 papers were selected for presentation as regular papers. The overall acceptance rate of all submissions reached 73.7%. For English papers, after two rounds of reviewing, we selected 13 papers from 39 submissions for presentation. The acceptance rate is 33%. Furthermore, two papers were selected for the Best Student Paper Award.

Many thanks to the authors for choosing NCMMSC to present their work and for their contribution to the high level of the conference. We also express our heartfelt thanks to the Program Committee members and Organizing Committee members, who put a tremendous amount of effort into soliciting and selecting qualified research papers with a good balance between quality and creativity.

We hope you enjoy reading and benefit from the proceedings of NCSMMSC 2017.

October 2017

Jianhua Tao
Thomas Fang Zheng

Organization

NCMMSC 2017 was organized by the Speech Information Committee of the Chinese Information Processing Society and co-organized by the Chinese Phonetics Association of CLP (Chinese Language Society) as well as the Language, Audio, and Music Acoustic Committee of the ASC (Acoustical Society of China), and was hosted by Huaihai Institute of Technology.

Organizing Committee

General Chairs

Jianhua Tao	Institute of Automation, Chinese Academy of Sciences, China
Thomas Fang Zheng	Tsinghua University, China
Xiaoming Ning	Huaihai Institute of Technology, China

Program Chairs

Cunhua Li	Huaihai Institute of Technology, China
Changchun Bao	Beijing University of Technology, China
Dong Wang	Tsinghua University, China

Organization Chairs

Hongwei Dai	Huaihai Institute of Technology, China
Jianguo Wei	Tianjin University, China

Publication Chairs

Jiqing Han	Harbin Institute of Technology, China
Mingxing Xu	Tsinghua University, China

Special Session Chairs

Jinsong Zhang	Beijing Language and Culture University, China
Jia Jia	Tsinghua University, China
Lei Xie	Northwestern Polytechnical University, China
Weibin Zhu	Beijing Jiaotong University, China

Program Committee

Huaiqiao Bao	National Institute of Chinese Academy of Social Sciences, Beijing
Changchun Bao	Beijing University of Technology, Beijing
Lianhong Cai	Tsinghua University, Beijing
Jianfen Cao	Institute of Linguistics, CASS, Beijing
Min Chu	Alibaba (China) Technology Co., Ltd., Hangzhou
Liyong Dai	University of Science and Technology of China, Hefei
Jianwu Dang	Tianjin University, Hefei
Minghui Dong	Institute for Infocomm Research, Singapore
Richard LiMin Du	Beijing Voxeasy Technology Co., Ltd., Beijing
Ditang Fang	Tsinghua University, Beijing
Zhonghua Fu	Northwestern Polytechnical University, Xi'an
Wentao Gu	Nanjing Normal University, Nanjing
Wu Guo	University of Science and Technology of China, Hefei
Jiqing Han	Harbin Institute of Technology, Harbin
Yufeng Hao	Beijing Speechocean Phonetic Technology Co. Ltd., Beijing
Lei He	Toshiba Medical System, Beijing
Wei He	Communication University of China, Beijing
Qingyang Hong	Xiamen University, Xiamen
Yu Hu	Iflytek Co., Ltd., Hefei
Jia Jia	Tsinghua University, Beijing
Lei Jia	Baidu Online Network Technology (Beijing) Co., Ltd., Beijing
Dongmei Jiang	Northwestern Polytechnical University, Xi'an
Qin Jin	Renmin University of China, Beijing
Jiangping Kong	Center for Chinese Linguistics PKU, Beijing
Jiasong Kong	Tsinghua University, Beijing
Tan Lee	The Chinese University of Hong Kong, Hong Kong
Haizhou Li	Institute for Infocomm Research, Singapore
Aijun Li	Institute of Linguistics, CASS, Beijing
Honglian Li	Beijing Information Science and Technology University, Beijing
Ya Li	Institute of Automation, Chinese Academy of Sciences, Beijing
Wei Li	Fudan University, Shanghai
Weiqian Liang	Tsinghua University, Beijing
Maocan Lin	Institute of Linguistics, CASS, Beijing
Zhenhua Ling	University of Science and Technology of China, Hefei
Jia Liu	Tsinghua University, Beijing
Shinan Lu	Beijing Sinovoice Phonetic Technology Co. Ltd., Beijing
Kim Teng Lua	National University of Singapore, Singapore
Helen Meng	The Chinese University of Hong Kong, Hong Kong
Jielin Pan	Institute of Acoustics of the Chinese Academy of Sciences, Beijing

Yong Qin	IBM Research, Beijing
Gan Woon Seng	Nanyang Technological University, Singapore
Yuanyuan Shi	Samsung Electronics, Beijing
Geping Song	Microsoft Research Asia, Beijing
Jianhua Tao	Institute of Automation, Chinese Academy of Sciences, Beijing
Feng Tong	Xiamen University, Xiamen
Chiu-yu Tseng	Academia Sinica, Taipei
Aerdukelimu Tuerhongjiang	Xinjiang Normal University, Urumqi
Dong Wang	Tsinghua University, Beijing
Kunlun Wang	Hefei University, Hefei
Renhua Wang	University of Science and Technology of China, Hefei
Zuoying Wang	Tsinghua University, Beijing
Xia Wang	Nokia (China) Investment Co. Ltd., Beijing
Hsin-Min Wang	Academia Sinica, Taipei
Jianguo Wei	Tianjin University, Tianjin
Zhengqi Wen	Institute of Automation, Chinese Academy of Sciences, Beijing
Zhiyong Wu	Tsinghua University, Beijing
Ji Wu	Tsinghua University, Beijing
Xinhong Wu	Peking University, Beijing
Zhizheng Wu	Apple
Silamu Wushouer	Xinjiang University, Urumchi
Yunqing Xia	Tsinghua University, Beijing
Lei Xie	Northwestern Polytechnical University, Xi'an
Xiang Xie	Beijing Institute of Technology, Beijing
Jieping Xu	School of Information, Renmin University of China, Beijing
Mingxing Xu	Tsinghua University, Beijing
Yufang Yang	Institute of Psychology of the Chinese Academy of Sciences, Beijing
Hongzhi Yu	Northwest Minzu University, Lanzhou
Kai Yu	Shanghai Jiao Tong University, Shanghai
Dong Yu	Tencent AI Lab, Seattle
Baozong Yuan	Beijing Jiaotong University, Beijing
Aibaidula Yusupu	Xinjiang Normal University, Urumchi
Jinsong Zhang	Beijing Language and Culture University, Beijing
Shuwu Zhang	Institute of Automation, Chinese Academy of Sciences, Beijing
Weiqliang Zhang	Tsinghua University, Beijing
Thomas Fang Zheng	Tsinghua University, Beijing
Weibin Zhu	Beijing Jiaotong University, Beijing
Chengqing Zong	Institute of Automation, Chinese Academy of Sciences, Beijing
Yiqing Zu	Iflytek Co. Ltd., Hefei

Contents

Uyghur Word Stemming Based on Stem and Affix Features.	1
<i>Hankiz Yilahun, Sediyeğyl Enwer, and Askar Hamdulla</i>	
GLEU-Guided Multi-resolution Network for Short Text Conversation	13
<i>Xuan Liu and Kai Yu</i>	
Applying Functional Partition in the Investigation of Lexical Tonal-Pattern Categories in an Under-Resourced Chinese Dialect	24
<i>Junru Wu, Yiya Chen, Vincent J. van Heuven, and Niels O. Schiller</i>	
Typology of Convergences and Divergences of English Monophthongs by EFL Learners from Guanhua Regions	36
<i>Yuan Jia, Yu Wang, Aijun Li, Dawei Song, and Liang Xu</i>	
Perception of English Phonemes by Chinese College Students	47
<i>Yanqin Feng, Hao Yan, and Liangkai Zhai</i>	
Collaborative Learning for Language and Speaker Recognition.	58
<i>Lantian Li, Zhiyuan Tang, Dong Wang, Andrew Abel, Yang Feng, and Shiyue Zhang</i>	
HelloNPU: A Corpus for Small-Footprint Wake-Up Word Detection Research	70
<i>Senmao Wang, Jingyong Hou, Lei Xie, and Yufeng Hao</i>	
Multi-task Learning in Prediction and Correction for Low Resource Speech Recognition	80
<i>Danish Bukhari, Jiangyan Yi, Zhengqi Wen, Bin Liu, and Jianhua Tao</i>	
Acoustic Model Compression with Knowledge Transfer.	89
<i>Jiangyan Yi, Jianhua Tao, Zhengqi Wen, Ya Li, and Hao Ni</i>	
Mongolian Text-to-Speech System Based on Deep Neural Network.	99
<i>Rui Liu, Feilong Bao, Guanglai Gao, and Yonghe Wang</i>	
Using Mandarin Training Corpus to Realize a Mandarin-Tibetan Cross-Lingual Emotional Speech Synthesis.	109
<i>Peiwen Wu, Hongwu Yang, and Zhenye Gan</i>	

Emotion Recognition Using Support Vector Machine
and Deep Neural Network 122
Ruinian Chen, Ying Zhou, and Yanmin Qian

Distance-Dependent Modeling of Head-Related Transfer Functions
Based on Spherical Fourier-Bessel Transform 132
Xiaoke Qi and Jianhua Tao

Author Index 143