Lecture Notes in Artificial Intelligence

10077

Subseries of Lecture Notes in Computer Science

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/1244>

Alberto Abad · Alfonso Ortega
António Teixeira · Carmen García Mateo
Carlos D. Martínez Hinarejos · Fernando Perdigão
Fernando Batista · Nuno Mamede (Eds.)

Advances in Speech and Language Technologies for Iberian Languages

Third International Conference, IberSPEECH 2016
Lisbon, Portugal, November 23–25, 2016
Proceedings

Editors

Alberto Abad
INESC-ID/IST
Universidade de Lisboa
Lisbon
Portugal

Alfonso Ortega
I3A/University of Zaragoza
Zaragoza
Spain

António Teixeira
DETI/IEETA
University of Aveiro
Aveiro
Portugal

Carmen García Mateo
AtlantTIC Research Center
Universidad de Vigo
Vigo
Spain

Carlos D. Martínez Hinarejos
Universitat Politècnica de València
Valencia
Spain

Fernando Perdigão
University of Coimbra
Coimbra
Portugal

Fernando Batista
INESC-ID/ISCTE-IUL
Lisbon
Portugal

Nuno Mamede
INESC-ID/IST
Universidade de Lisboa
Lisbon
Portugal

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-49168-4 ISBN 978-3-319-49169-1 (eBook)
DOI 10.1007/978-3-319-49169-1

Library of Congress Control Number: 2016956496

LNCS Sublibrary: SL7 – Artificial Intelligence

© Springer International Publishing AG 2016

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.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The Spanish Thematic Network on Speech Technology (RTTH) and the International Speech Communication Association (ISCA) Special Interest Group on Iberian Languages (SIG-IL) are pleased to present a selection of papers from the IberSPEECH 2016 conference—*IX Jornadas en Tecnologías del Habla* and the V Iberian SLTech Workshop—that was held in Lisbon, Portugal, during November 23–25 2016.

The IberSPEECH series of conferences have become one of the most relevant scientific events for the community working in the field of speech and language processing of Iberian languages. This is demonstrated by the increased interest and success of previous editions, starting in Vigo 2010 with FALA and continuing in Madrid 2012 and Las Palmas de Gran Canaria 2014 with the new denomination: IberSPEECH. Prior to these past joint editions of the *Jornadas en Tecnologías del Habla* and the Iberian SLTech Workshop, separate events were held. The former ones have been organized by the RTTH (<http://www.rthabla.es>) since 2000. This network was created in 1999 and currently includes over 200 researchers and 30 research groups in speech technology all over Spain. The latter (Iberian SLTech) have been organized by the SIG-IL (<http://www.il-sig.org/>) of the ISCA since 2009. Overall, these events have attracted the attention of many researchers over the years, mainly from Spain, Portugal, and from other Iberian-speaking countries in Latin America, but also from several other research groups from all around the world. These previous efforts have contributed to the consolidation of a strong and active community addressing the problems of speech and language automatic processing of Iberian languages.

The IberSPEECH 2016 conference –the third of its kind using this name– brought together the *IX Jornadas en Tecnologías del Habla* and the V Iberian SLTech Workshop events, and it was co-organized by INESC-ID Lisboa, the RTTH, and the ISCA SIG-IL. Maintaining the identity of previous editions, this new event represented not only a step forward for the support of researchers in Iberian languages, but also a new challenge for the community, since it was the first edition to be held outside Spain. In order to promote interaction and discussion among all the members of the community, the Organizing Committee planned a three-day event with a wide variety of activities, such as technical paper presentations, keynote lectures, evaluation challenges, presentation of demos and research projects, and recent PhD thesis.

This *Lecture Notes in Artificial Intelligence* volume contains only a selection of the full regular articles presented during the IberSPEECH 2016 conference. To ensure the quality of all the contributions, each submitted article was reviewed by three members of the Scientific Review Committee, who provided feedback to improve the final article version, in addition to acceptance/rejection recommendations for the Technical Program Committee. Authors of the accepted papers had time to address the comments before submitting the camera-ready versions. We had 48 regular full papers submitted and only 27 of them were selected for publication in this volume. This selection was

based on the scores and comments provided by our Scientific Review Committee, which includes over 66 researchers from different institutions mainly from Spain, Portugal, and Latin America, but also from France, Germany, Hungary, Italy, Norway, Sweden, UK, and the USA, to whom we also would like to express our deepest gratitude. The selected articles in this volume are organized into four different topics:

- Speech Production, Analysis, Coding and Synthesis
- Automatic Speech Recognition
- Paralinguistic Speaker Trait Characterization
- Speech and Language Technologies in Different Application Fields

The Organizing Committee of IberSPEECH is extremely proud of the contents of this volume. We trust that we have achieved and delivered the scientific quality standards that the researchers in the field of speech and language technologies for Iberian languages value.

As mentioned previously, besides the excellent research articles included in this volume, the conference included a wide variety of scientific activities that are worth mentioning in more detail, particularly, the ALBAYZIN Technology Competitive Evaluations, the Special Session contributions, and the keynote lectures.

The ALBAYZIN Technology Competitive Evaluations have been organized alongside the conference since 2006, promoting the fair and transparent comparison of technology in different fields related to speech and language technology. In this edition we had two different challenge evaluations: Speaker Diarization and Search on Speech. The organization of each one of these evaluations requires the preparation of development and test data, which is released along with a clear set of rules to the participants, and gathering and comparing results from participants. This organization was carried out by different groups of researchers with the support of the ALBAYZIN Committee. Although article contributions and results from the evaluations are not included in this volume, we would like to express our gratitude to the organizers and also to the participants of these evaluation challenges.

In addition to full articles, a subset of which are represented here, and to system description papers of the ALBAYZIN participants, Special Session papers were also included in the conference program. These were intended to describe either progress in current or recent research and development projects, demonstration systems, or PhD Thesis extended abstracts to compete in the PhD Award.

Moreover, we had the pleasure of hosting three extraordinary keynote speakers: Prof. Elmar Nöth (University of Erlangen-Nuremberg, Germany), Dr. Bhuvana Ramabhadran (IBM T.J. Watson Research Center, USA) and Prof. Steve Renals (University of Edinburgh, UK). We would like to acknowledge their extremely valuable participation that helped to elevate the quality of the overall event.

Finally, we would like to thank all those whose effort made possible this conference, including the members of the Organizing Committee, the Local Organizing Committee, the ALBAYZIN Committee, the Scientific Reviewer Committee, the authors, the conference attendees, the supporting institutions, and so many people who gave their

best to achieve a successful conference. We would also like to thank Springer for the possibility of publishing another volume of selected papers from an IberSPEECH conference.

November 2016

Alberto Abad
Alfonso Ortega
António Teixeira
Carmen García Mateo
Carlos D. Martínez Hinarejos
Fernando Perdigão
Fernando Batista
Nuno Mamede

Organization

General Chair

Alberto Abad	INESC-ID/IST, University of Lisbon, Portugal
Alfonso Ortega	Universidad de Zaragoza, Spain
António Teixeira	Universidade de Aveiro, Portugal

Technical Program Chair

Carmen García Mateo	Universidad de Vigo, Spain
Fernando Perdigão	Universidade de Coimbra, Portugal
C.D. Martínez-Hinarejos	Universitat Politècnica de València, Spain

Publication Chair

Nuno Mamede	INESC-ID/IST, University of Lisbon, Portugal
Fernando Batista	INESC-ID/ISCTE-IUL, Portugal

Special Session and Awards Chair

Juan Luis Navarro Mesa	Universidad de Las Palmas de Gran Canaria, Spain
Doroteo Torre Toledano	Universidad Autónoma de Madrid, Spain
Xavier Anguera	ELSA Corp., USA

Plenary Talks Chair

Isabel Trancoso	INESC-ID/IST, University of Lisbon, Portugal
-----------------	--

Evaluations Chair

Luis J. Rodríguez Fuentes	Universidad del País Vasco, Spain
Rubén San-Segundo	Universidad Politécnica de Madrid, Spain

Publicity and Sponsorship Chair

Helena Moniz	INESC-ID/FL, University of Lisbon, Portugal
--------------	---

Albayzin Committee

Luis J. Rodríguez Fuentes	Universidad del País Vasco, Spain
Rubén San-Segundo	Universidad Politécnica de Madrid, Spain

Alberto Abad	INESC-ID/IST, University of Lisbon, Portugal
Alfonso Ortega	Universidad de Zaragoza, Spain
António Teixeira	Universidade de Aveiro, Portugal

Local Organizing Committee

Alberto Abad	INESC-ID/University of Lisbon, Portugal
Fernando Batista	INESC-ID/ISCTE-IUL, Portugal
Nuno Mamede	INESC-ID/IST, University of Lisbon, Portugal
David Martins de Matos	INESC-ID/IST, University of Lisbon, Portugal
Helena Moniz	INESC-ID/FL, University of Lisbon, Portugal
Rubén Solera-Ureña	INESC-ID, Portugal
Isabel Trancoso	INESC-ID/IST, University of Lisbon, Portugal

Scientific Review Committee

Alberto Abad	INESC-ID/IST, University of Lisbon, Portugal
Plinio Barbosa	University of Campinas, Brazil
Jorge Baptista	INESC-ID/University of Algarve, Portugal
Fernando Batista	INESC-ID/ISCTE-IUL, Portugal
José Miguel Benedí Ruiz	Universitat Politècnica de València, Spain
Carmen Benítez	Universidad de Granada, Spain
Antonio Bonafonte	Universitat Politècnica de Catalunya, Spain
German Bordel	University of the Basque Country UPV/EHU, Spain
Alessio Brutti	FBK, Italy
Paula Carvalho	INESC-ID/Universidade Europeia, Portugal
Diamantino Caseiro	Google Inc., USA
Maria Jose Castro-Bleda	Universitat Politècnica de Valencia, Spain
Ricardo Cordoba	Grupo de Tecnología del Habla, Spain
Conceicao Cunha	IPS Munich, Germany
Carme De-La-Mota	Universitat Autònoma de Barcelona, Spain
Laura Docío Fernández	University of Vigo, Spain
Daniel Erro	University of the Basque Country UPV/EHU, Spain
David Escudero	University of Valladolid, Spain
Ruben Fernandez	Universidad Politécnica de Madrid, Spain
Javier Ferreiros	GTH, Universidad Politécnica de Madrid, Spain
Julian Fierrez	Universidad Autónoma de Madrid, Spain
Ascension Gallardo	Universidad Carlos III de Madrid, Spain
Carmen Garcia Mateo	Universidade de Vigo, Spain
Juan I. Godino Llorente	Universidad Politécnica de Madrid, Spain
Javier Hernando	Universitat Politècnica de Catalunya, Spain
Lluís Felip Hurtado Oliver	Universitat Politècnica de Valencia, Spain
Irina Illina	LORIA, France
Eduardo Lleida	University of Zaragoza, Spain
José David Lopes	KTH, Sweden
Paula López Otero	Universidade de Vigo, Spain

Ranniery Maia	Toshiba Research Europe Ltd., UK
C.D. Martínez Hinarejos	Universitat Politècnica de València, Spain
Helena Moniz	INESC-ID/FL, University of Lisbon, Portugal
Nicolas Morales Mombiola	Nuance Communications GmbH, Germany
Asuncion Moreno	Universitat Politècnica de Catalunya, Spain
Antonio Moreno-Sandoval	Universidad Autonoma Madrid, Spain
Climent Nadeu	Universitat Politècnica de Catalunya, Spain
Juan L. Navarro-Mesa	Universidad de Las Palmas de Gran Canaria, Spain
Eva Navas Cordón	University of the Basque Country, Spain
Géza Németh	University of Technology and Economics, Hungary
Nelson Neto	Universidade Federal do Pará, Brazil
Alfonso Ortega	University of Zaragoza, Spain
Thomas Pellegrini	Université de Toulouse; IRIT, France
Carmen Peláez-Moreno	University Carlos III Madrid, Spain
Mikel Penagarikano	University of the Basque Country, Spain
Fernando Perdigao	IT/Universidade de Coimbra, Portugal
Ferran Pla	DSIC, Universitat Politècnica de València, Spain
José L. Pérez-Córdoba	University of Granada, Spain
Paulo Quaresma	Universidade de Evora, Portugal
Andreia Rauber	University of Tübingen, Germany
Luis J. Rodriguez-Fuentes	University of the Basque Country UPV/EHU, Spain
Eduardo Rodriguez Banga	University of Vigo, Spain
José A.R. Fonollosa	Universitat Politècnica de Catalunya, Spain
Rubén San-Segundo	Universidad Politécnica de Madrid, Spain
Emilio Sanchis	Universidad Politécnica Valencia, Spain
Diana Santos	University of Oslo, Norway
Encarna Segarra	DSIC, Universidad Politécnica de Valencia, Spain
Alberto Simões	ESEIG, Instituto Politécnico do Porto, Portugal
Rubén Solera-Ureña	INESC-ID, Portugal
Joan Andreu Sanchez	Universitat Politècnica de València, Spain
António Teixeira	University of Aveiro, Portugal
Javier Tejedor	GEINTRA/Universidad de Alcala, Spain
Doroteo Toledano	Universidad Autónoma de Madrid, Spain
Pedro Torres-Carrasquillo	MIT Lincoln Laboratory, USA
Isabel Trancoso	INESC-ID/IST, University of Lisbon, Portugal
Amparo Varona	University of the Basque Country, Spain
Aline Villavicencio	Universidade Federal do Rio Grande do Sul, Brazil

Organizing Institutions

INESC-ID Lisboa

Spanish Thematic Network on Speech Technology (RTTH)

ISCA Special Interest Group on Iberian Languages (SIG-IL)

Support and Partner Institutions

Instituto Superior Técnico, University of Lisbon

Ministerio de Economía y Competitividad, Gobierno de España

FCT, Fundação para a Ciência e a Tecnologia

Câmara Municipal de Lisboa

TAP Portugal

Contents

Speech Production, Analysis, Coding and Synthesis

Study of the Effect of Reducing Training Data in Speech Synthesis Adaptation Based on Frequency Warping.	3
<i>Agustin Alonso, Daniel Erro, Eva Navas, and Inma Hernaez</i>	
A Dynamic FEC for Improved Robustness of CELP-Based Codec	14
<i>Nadir Benamirouche, Bachir Boudraa, Angel M. Gomez, José L. Pérez-Córdoba, and Iván López-Espejo</i>	
Objective Comparison of Four GMM-Based Methods for PMA-to-Speech Conversion	24
<i>Daniel Erro, Inma Hernaez, Luis Serrano, Ibon Saratxaga, and Eva Navas</i>	
Adding Singing Capabilities to Unit Selection TTS Through HNM-Based Conversion	33
<i>Marc Freixes, Joan Claudi Socoró, and Francesc Alías</i>	
A Novel Error Mitigation Scheme Based on Replacement Vectors and FEC Codes for Speech Recovery in Loss-Prone Channels	44
<i>Domingo López-Oller, Angel M. Gomez, and José L. Pérez-Córdoba</i>	
Language-Independent Acoustic Cloning of HTS Voices: An Objective Evaluation	54
<i>Carmen Magariños, Daniel Erro, Paula Lopez-Otero, and Eduardo R. Banga</i>	
Prosodic Break Prediction with RNNs	64
<i>Santiago Pascual and Antonio Bonafonte</i>	
Surgery of Speech Synthesis Models to Overcome the Scarcity of Training Data	73
<i>Arnaud Pierard, D. Erro, I. Hernaez, E. Navas, and Thierry Dutoit</i>	

Automatic Speech Recognition

An Analysis of Deep Neural Networks in Broad Phonetic Classes for Noisy Speech Recognition	87
<i>F. de-la-Calle-Silos, A. Gallardo-Antolín, and C. Peláez-Moreno</i>	

Automatic Speech Recognition with Deep Neural Networks for Impaired Speech	97
<i>Cristina España-Bonet and José A.R. Fonollosa</i>	
Detection of Publicity Mentions in Broadcast Radio: Preliminary Results . . .	108
<i>Maria Pilar Fernández-Gallego, Álvaro Mesa-Castellanos, Alicia Lozano-Díez, and Doroteo T. Toledano</i>	
Deep Neural Network-Based Noise Estimation for Robust ASR in Dual-Microphone Smartphones	117
<i>Iván López-Espejo, Antonio M. Peinado, Angel M. Gomez, and Juan M. Martín-Doñas</i>	
Better Phoneme Recognisers Lead to Better Phoneme Posteriorgrams for Search on Speech? An Experimental Analysis	128
<i>Paula Lopez-Otero, Laura Docio-Fernandez, and Carmen Garcia-Mateo</i>	
Crowdsourced Video Subtitling with Adaptation Based on User-Corrected Lattices	138
<i>João Miranda, Ramón F. Astudillo, Ângela Costa, André Silva, Hugo Silva, João Graça, and Bhiksha Raj</i>	
Paralinguistic Speaker Trait Characterization	
Acoustic Analysis of Anomalous Use of Prosodic Features in a Corpus of People with Intellectual Disability	151
<i>Mario Corrales-Astorgano, David Escudero-Mancebo, and César González-Ferreras</i>	
Detecting Psychological Distress in Adults Through Transcriptions of Clinical Interviews	162
<i>Joana Correia, Isabel Trancoso, and Bhiksha Raj</i>	
Automatic Annotation of Disfluent Speech in Children’s Reading Tasks	172
<i>Jorge Proença, Dirce Celorico, Carla Lopes, Sara Candeias, and Fernando Perdigão</i>	
Automatic Detection of Hyperarticulated Speech	182
<i>Eugénio Ribeiro, Fernando Batista, Isabel Trancoso, Ricardo Ribeiro, and David Martins de Matos</i>	
Acoustic-Prosodic Automatic Personality Trait Assessment for Adults and Children.	192
<i>Rubén Solera-Ureña, Helena Moniz, Fernando Batista, Ramón F. Astudillo, Joana Campos, Ana Paiva, and Isabel Trancoso</i>	

Speech and Language Technologies in Different Application Fields

Evaluating Different Non-native Pronunciation Scoring Metrics with the Japanese Speakers of the SAMPLE Corpus	205
<i>Vandria Álvarez Álvarez, David Escudero Mancebo, César González Ferreras, and Valentín Cardeñoso Payo</i>	
Global Analysis of Entrainment in Dialogues	215
<i>Vera Cabarrão, Isabel Trancoso, Ana Isabel Mata, Helena Moniz, and Fernando Batista</i>	
A Train-on-Target Strategy for Multilingual Spoken Language Understanding	224
<i>Fernando García-Granada, Encarna Segarra, Carlos Millán, Emilio Sanchis, and Lluís-F. Hurtado</i>	
Collaborator Effort Optimisation in Multimodal Crowdsourcing for Transcribing Historical Manuscripts	234
<i>Emilio Granell and Carlos-D. Martínez-Hinarejos</i>	
Assessing User Expertise in Spoken Dialog System Interactions	245
<i>Eugénio Ribeiro, Fernando Batista, Isabel Trancoso, José Lopes, Ricardo Ribeiro, and David Martins de Matos</i>	
Automatic Speech Feature Learning for Continuous Prediction of Customer Satisfaction in Contact Center Phone Calls	255
<i>Carlos Segura, Daniel Balcells, Martí Umbert, Javier Arias, and Jordi Luque</i>	
Reversible Speech De-identification Using Parametric Transformations and Watermarking	266
<i>Aitor Valdivielso, Daniel Erro, and Inma Hernaez</i>	
Bottleneck Based Front-End for Diarization Systems	276
<i>Ignacio Viñals, Jesús Villalba, Alfonso Ortega, Antonio Miguel, and Eduardo Lleida</i>	
Author Index	287