

Mohamed Chetouani Amir Hussain
Bruno Gas Maurice Milgram
Jean-Luc Zarader (Eds.)

Advances in Nonlinear Speech Processing

International Conference
on Nonlinear Speech Processing, NOLISP 2007
Paris, France, May 22-25, 2007
Revised Selected Papers

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Mohamed Chetouani

Bruno Gas

Maurice Milgram

Jean-Luc Zarader

Université Pierre et Marie Curie

Institut des Systèmes Intelligents et de Robotique (ISIR)

3 rue Galilée, 94200 Ivry-sur-Seine, France

E-mail: {mohamed.chetouani, bruno.gas, maurice.milgram, jean-luc-zarader}@upmc.fr

Amir Hussain

University of Stirling

Department of Computing Science & Mathematics

Stirling FK9 4LA, Scotland, UK

E-mail: ahu@cs.stir.ac.uk

Library of Congress Control Number: 2007941810

CR Subject Classification (1998): I.2.7, J.5, C.3

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743

ISBN-10 3-540-77346-0 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-77346-7 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper SPIN: 12206032 06/3180 5 4 3 2 1 0

Preface

We present in this volume a collection of revised selected papers from the ISCA Tutorial and Research Workshop on Nonlinear Speech Processing (NOLISP 2007) held in Paris, France, 22–25 May, 2007. NOLISP 2007 was organized by the University Pierre and Marie Curie (UPMC) with the generous support of ISCA (International Speech Communication Association), EURASIP and the IEEE. NOLISP 2007 was the first follow-on workshop to a series of three earlier events related to nonlinear speech processing, that were organized within the framework of the European COST action 277 Nonlinear speech processing (2001–2005). The financial support of ISCA enabled the attendance of leading researchers from various parts of the world.

The exciting field of speech processing has witnessed tremendous development over the past 20 years or so, thanks to both technological progress and to the growing focus of research on a number of key application areas. However, some specificities of the speech signal are still not well addressed by the currently available models. Hence, new nonconventional models and processing techniques need to be investigated in order to foster and/or accompany future progress, even if they do not match immediately the level of performance and understanding of the current state-of-the-art approaches.

The purpose of NOLISP is to present and discuss novel ideas, work and results related to such alternative techniques for speech processing, which depart from mainstream approaches. It declared its intent to be an interdisciplinary forum, intertwining research in different fields of speech processing with its growing applications in everyday practice.

One of the special characteristics of the NOLISP volumes is that the authors usually propose new improvements for speech processing by drawing inspiration from other exciting fields including, amongst others, statistical signal processing, pattern classification, multi-modal processing, perceptual-based criteria, auditory processing and machine learning-based approaches.

This volume contains a collection of revised and selected papers presented at NOLISP 2007. After a thorough review process, 24 papers were accepted for publication, including the contribution of three invited speakers. A total of 8 sessions containing 32 papers were accepted for presentation, covering specific aspects such as speaker recognition, speech analysis, voice pathologies, speech recognition, speech enhancement, audio and visual speech processing and applications.

In this book, the papers have been divided into the following sections:

- Nonlinear and Nonconventional Techniques
- Speech Synthesis
- Speaker Recognition
- Speech Recognition

- Speech Analysis
- Exploitation of nonlinear techniques

Finally, we would like to express our gratitude to members of the NOLISP Organizing Committee, and to all the people who participated in the event (delegates, invited speakers, scientific committee members). The editors would also like to address a special mention to the people who helped in the peer-review process as special or additional reviewers.

October 2007

Mohamed Chetouani
Amir Hussain
Bruno Gas
Maurice Milgram
Jean-Luc Zarader

Organization

NOLISP 2007 was organized by the Institute of Intelligent Systems and Robotics (CNRS FRE2507), University Pierre and Marie Curie (Paris-6).

Scientific Committee

Frédéric Bimbot	IRISA, Rennes (France)
Mohamed Chetouani	UPMC, Paris (France)
Gérard Chollet	ENST, Paris (France)
Tariq Durrani	University of Strathclyde, Glasgow (UK)
Marcos Faundez-Zanuy	EUPMt, Barcelona (Spain)
Bruno Gas	UPMC, Paris (France)
Hynek Hermansky	OGI, Portland (USA)
Amir Hussain	University of Stirling, Scotland (UK)
Eric Keller	University of Lausanne (Switzerland)
Bastiaan Kleijn	KTH, Stockholm (Sweden)
Gernot Kubin	TUG, Graz (Austria)
Petros Maragos	Nat. Tech. Univ. of Athens (Greece)
Stephen McLaughlin	University of Edinburgh (UK)
Kuldip Paliwal	University of Brisbane (Australia)
Bojan Petek	University of Ljubljana (Slovenia)
Jean Rouat	University of Sherbrooke (Canada)
Jean Schoentgen	Univ. Libre Bruxelles (Belgium)
Isabel Trancoso	INESC (Portugal)

Organizing Committee

Mohamed Chetouani	UPMC, Paris (France)
Bruno Gas	UPMC, Paris (France)
Amir Hussain	University of Stirling, Scotland (UK)
Maurice Milgram	UPMC, Paris (France)
Michèle Vie	UPMC, Paris (France)
Jean-Luc Zarader	UPMC, Paris (France)

Sponsoring Institutions

University Pierre and Marie Curie (Paris-6)
International Speech Communication Association (ISCA)
European Association For Signal, Speech And Image Processing (EURASIP)
IEEE UKRI Industry Applications Society Chapter

Table of Contents

Non-Linear and Non-Conventional Techniques

Phase-Based Methods for Voice Source Analysis	1
<i>Christophe d’Alessandro, Baris Bozkurt, Boris Doval, Thierry Dutoit, Nathalie Henrich, Vu Ngoc Tuan, and Nicolas Sturmel</i>	
Some Experiments in Audio-Visual Speech Processing	28
<i>G. Chollet, R. Landais, T. Hueber, H. Bredin, C. Mokbel, P. Perrot, and L. Zouari</i>	
Exploiting Nonlinearity in Adaptive Signal Processing	57
<i>Phebe Vayanos, Mo Chen, Beth Jelfs, and Danilo P. Mandic</i>	

Speech Synthesis

Mixing HMM-Based Spanish Speech Synthesis with a CBR for Prosody Estimation	78
<i>Xavi Gonzalvo, Ignasi Iriondo, Joan Claudi Socoró, Francesc Alías, and Carlos Monzo</i>	
Objective and Subjective Evaluation of an Expressive Speech Corpus . . .	86
<i>Ignasi Iriondo, Santiago Planet, Joan-Claudi Socoró, and Francesc Alías</i>	

Speaker Recognition

On the Usefulness of Linear and Nonlinear Prediction Residual Signals for Speaker Recognition	95
<i>Marcos Faundez-Zanuy</i>	
Multi Filter Bank Approach for Speaker Verification Based on Genetic Algorithm	105
<i>Christophe Charbuillet, Bruno Gas, Mohamed Chetouani, and Jean Luc Zarader</i>	
Speaker Recognition Via Nonlinear Phonetic- and Speaker-Discriminative Features	114
<i>Lara Stoll, Joe Frankel, and Nikki Mirghafori</i>	
Perceptron-Based Class Verification	124
<i>Michael Gerber, Tobias Kaufmann, and Beat Pfister</i>	

Speech Recognition

Manifold Learning-Based Feature Transformation for Phone Classification	132
<i>Andrew Errity, John McKenna, and Barry Kirkpatrick</i>	
Word Recognition with a Hierarchical Neural Network	142
<i>Xavier Domont, Martin Heckmann, Heiko Wersing, Frank Joublin, Stefan Menzel, Bernhard Sendhoff, and Christian Goerick</i>	
Hybrid Models for Automatic Speech Recognition: A Comparison of Classical ANN and Kernel Based Methods	152
<i>Ana I. García-Moral, Rubén Solera-Ureña, Carmen Peláez-Moreno, and Fernando Díaz-de-María</i>	
Towards Phonetically-Driven Hidden Markov Models: Can We Incorporate Phonetic Landmarks in HMM-Based ASR?	161
<i>Guillaume Gravier and Daniel Moraru</i>	
A Hybrid Genetic-Neural Front-End Extension for Robust Speech Recognition over Telephone Lines	169
<i>Sid-Ahmed Selouani, Habib Hamam, and Douglas O'Shaughnessy</i>	
Efficient Viterbi Algorithms for Lexical Tree Based Models	179
<i>Salvador España-Boquera, Maria Jose Castro-Bleda, Francisco Zamora-Martínez, and Jorge Gorbe-Moya</i>	

Speech Analysis

Non-stationary Self-consistent Acoustic Objects as Atoms of Voiced Speech	188
<i>Friedhelm R. Dreyer</i>	
The Hartley Phase Cepstrum as a Tool for Signal Analysis	204
<i>Ioannis Paraskevas and Maria Rangoussi</i>	
Voiced Speech Analysis by Empirical Mode Decomposition	213
<i>Aïcha Bouzid and Noureddine Ellouze</i>	
Estimation of Glottal Closure Instances from Speech Signals by Weighted Nonlinear Prediction	221
<i>Karl Schnell</i>	
Quantitative Perceptual Separation of Two Kinds of Degradation in Speech Denoising Applications	230
<i>Anis Ben Aïcha and Sofia Ben Jebara</i>	

Exploitation of non-linear techniques

An Efficient VAD Based on a Generalized Gaussian PDF	246
<i>Oscar Pernía, Juan M. Górriz, Javier Ramírez, Carlos G. Puntonet, and Ignacio Turias</i>	
Estimating the Dispersion of the Biometric Glottal Signature in Continuous Speech	255
<i>Pedro Gómez, Agustín Álvarez, Luis Miguel Mazaira, Roberto Fernández, Victoria Rodellar, Rafael Martínez, and Cristina Muñoz</i>	
Trajectory Mixture Density Networks with Multiple Mixtures for Acoustic-Articulatory Inversion	263
<i>Korin Richmond</i>	
Application of Feature Subset Selection Based on Evolutionary Algorithms for Automatic Emotion Recognition in Speech	273
<i>Aitor Álvarez, Idoia Cearreta, Juan Miguel López, Andoni Arruti, Elena Lazkano, Basilio Sierra, and Nestor Garay</i>	
Author Index	283