

Contents

Invited Talks

Low-Resource Speech Recognition and Keyword-Spotting	3
<i>Mark J.F. Gales, Kate M. Knill, and Anton Ragni</i>	
Big Data, Deep Learning – At the Edge of X-Ray Speaker Analysis	20
<i>Björn W. Schuller</i>	

Conference Papers

A Comparison of Covariance Matrix and i-vector Based Speaker Recognition	37
<i>Nikša Jakovljević, Ivan Jokić, Slobodan Jošić, and Vlado Delić</i>	
A Trainable Method for the Phonetic Similarity Search in German Proper Names	46
<i>Oliver Jokisch and Horst-Udo Hain</i>	
Acoustic and Perceptual Correlates of Vowel Articulation in Parkinson’s Disease With and Without Mild Cognitive Impairment: A Pilot Study	56
<i>Michaela Strinzel, Vasilisa Verkhodanova, Fedor Jalvingh, Roel Jonkers, and Matt Coler</i>	
Acoustic Cues for the Perceptual Assessment of Surround Sound	65
<i>Ingo Siebert, Oliver Jokisch, Alicia Flores Lotz, Franziska Trojahn, Martin Meszaros, and Michael Maruschke</i>	
Acoustic Modeling in the STC Keyword Search System for OpenKWS 2016 Evaluation	76
<i>Ivan Medennikov, Aleksei Romanenko, Alexey Prudnikov, Valentin Mendelev, Yuri Khokhlov, Maxim Korenevsky, Natalia Tomashenko, and Alexander Zatvornitskiy</i>	
Adaptation Approaches for Pronunciation Scoring with Sparse Training Data	87
<i>Federico Landini, Luciana Ferrer, and Horacio Franco</i>	
An Algorithm for Detection of Breath Sounds in Spontaneous Speech with Application to Speaker Recognition.	98
<i>Sri Harsha Dumpala and K.N.R.K. Raju Alluri</i>	

An Alternative Approach to Exploring a Video.	109
<i>Fahim A. Salim, Fasih Haider, Owen Conlan, and Saturnino Luz</i>	
An Analysis of the RNN-Based Spoken Term Detection Training.	119
<i>Jan Švec, Luboš Šmídl, and Josef V. Psutka</i>	
Analysis of Interaction Parameter Levels in Interaction Quality Modelling for Human-Human Conversation.	130
<i>Anastasiia Spirina, Olesia Vaskovskaia, Tatiana Karaseva, Alina Skorokhod, Iana Polonskaia, and Maxim Sidorov</i>	
Annotation Error Detection: Anomaly Detection vs. Classification.	141
<i>Jindřich Matoušek and Daniel Tihelka</i>	
Are You Addressing Me? Multimodal Addressee Detection in Human-Human-Computer Conversations	152
<i>Oleg Akhtiamov, Dmitrii Ubskii, Evgeniia Feldina, Aleksei Pugachev, Alexey Karpov, and Wolfgang Minker</i>	
Assessing Spoken Dialog Services from the End-User Perspective: Usability and Experience	162
<i>Otilia Kocsis, Basilis Kladis, Anastasios Tsopanoglou, and Nikos Fakotakis</i>	
Audio-Replay Attack Detection Countermeasures	171
<i>Galina Lavrentyeva, Sergey Novoselov, Egor Malykh, Alexander Kozlov, Oleg Kudashev, and Vadim Shchemelinin</i>	
Automatic Estimation of Presentation Skills Using Speech, Slides and Gestures.	182
<i>Abualsoud Hanani, Mohammad Al-Amleh, Waseem Bazbus, and Saleem Salameh</i>	
Automatic Phonetic Transcription for Russian: Speech Variability Modeling	192
<i>Vera Evdokimova, Pavel Skrelin, and Tatiana Chukaeva</i>	
Automatic Smoker Detection from Telephone Speech Signals.	200
<i>Amir Hossein Poorjam, Soheila Hesaraki, Saeid Safavi, Hugo van Hamme, and Mohamad Hasan Bahari</i>	
Bimodal Anti-Spoofing System for Mobile Security	211
<i>Eugene Luckyanets, Aleksandr Melnikov, Oleg Kudashev, Sergey Novoselov, and Galina Lavrentyeva</i>	
Canadian English Word Stress: A Corpora-Based Study of National Identity in a Multilingual Community	221
<i>Tatiana Shevchenko and Daria Pozdeeva</i>	

Classification of Formal and Informal Dialogues Based on Turn-Taking and Intonation Using Deep Neural Networks	233
<i>István Szekrényes and György Kovács</i>	
Clustering Target Speaker on a Set of Telephone Dialogs	244
<i>Andrey Shulipa, Aleksey Sholohov, and Yuri Matveev</i>	
Cognitive Entropy in the Perceptual-Auditory Evaluation of Emotional Modal States of Foreign Language Communication Partner	253
<i>Rodmonga Potapova and Vsevolod Potapov</i>	
Correlation Normalization of Syllables and Comparative Evaluation of Pronunciation Quality in Speech Rehabilitation	262
<i>Evgeny Kostyuchenko, Roman Meshcheryakov, Dariya Ignatieva, Alexander Pyatkov, Evgeny Choyazonov, and Lidiya Balatskaya</i>	
CRF-Based Phrase Boundary Detection Trained on Large-Scale TTS Speech Corpora	272
<i>Markéta Jůzová</i>	
Deep Recurrent Neural Networks in Speech Synthesis Using a Continuous Vocoder	282
<i>Mohammed Salah Al-Radhi, Tamás Gábor Csapó, and Géza Németh</i>	
Design of Online Echo Canceller in Duplex Mode	292
<i>Andrey Barabanov and Evgenij Vikulov</i>	
Detection of Stance and Sentiment Modifiers in Political Blogs	302
<i>Maria Skeppstedt, Vasiliki Simaki, Carita Paradis, and Andreas Kerren</i>	
Digits to Words Converter for Slavic Languages in Systems of Automatic Speech Recognition	312
<i>Josef Chaloupka</i>	
Discriminating Speakers by Their Voices — A Fusion Based Approach	322
<i>Halim Sayoud, Siham Ouamour, and Zohra Hamadache</i>	
Emotional Poetry Generation	332
<i>Aitzol Astigarraga, José María Martínez-Otzeta, Igor Rodriguez, Basilio Sierra, and Elena Lazkano</i>	
End-to-End Large Vocabulary Speech Recognition for the Serbian Language	343
<i>Branislav Popović, Edvin Pakoci, and Darko Pekar</i>	

Examining the Impact of Feature Selection on Sentiment Analysis for the Greek Language	353
<i>Nikolaos Spatiotis, Michael Paraskevas, Isidoros Perikos, and Iosif Mporas</i>	
Experimenting with Hybrid TDNN/HMM Acoustic Models for Russian Speech Recognition	362
<i>Irina Kipyatkova</i>	
Exploring Multiparty Casual Talk for Social Human-Machine Dialogue	370
<i>Emer Gilmartin, Benjamin R. Cowan, Carl Vogel, and Nick Campbell</i>	
First Experiments to Detect Anomaly Using Personality Traits vs. Prosodic Features	379
<i>Cedric Fayet, Arnaud Delhay, Damien Lolive, and Pierre-François Marteau</i>	
Fusion of a Novel Volterra-Wiener Filter Based Nonlinear Residual Phase and MFCC for Speaker Verification	389
<i>Purvi Agrawal and Hemant A. Patil</i>	
Hesitations in Spontaneous Speech: Acoustic Analysis and Detection	398
<i>Vasilisa Verkhodanova, Vladimir Shapranov, and Irina Kipyatkova</i>	
Human as Acmeologic Entity in Social Network Discourse (Multidimensional Approach)	407
<i>Rodmonga Potapova and Vsevolod Potapov</i>	
Improved Speaker Adaptation by Combining I-vector and fMLLR with Deep Bottleneck Networks	417
<i>Thai Son Nguyen, Kevin Kilgour, Matthias Sperber, and Alex Waibel</i>	
Improving of LVCSR for Causal Czech Using Publicly Available Language Resources	427
<i>Petr Mizera and Petr Pollak</i>	
Improving Performance of Speaker Identification Systems Using Score Level Fusion of Two Modes of Operation	438
<i>Saeid Safavi and Iosif Mporas</i>	
Improving Speech-Based Emotion Recognition by Using Psychoacoustic Modeling and Analysis-by-Synthesis	445
<i>Ingo Siebert, Alicia Flores Lotz, Olga Egorow, and Andreas Wendemuth</i>	
In Search of Sentence Boundaries in Spontaneous Speech	456
<i>Natalia Bogdanova-Beglarian</i>	

Investigating Acoustic Correlates of Broad and Narrow Focus Perception by Japanese Learners of English	464
<i>Gábor Pintér, Oliver Jokisch, and Shinobu Mizuguchi</i>	
Language Adaptive Multilingual CTC Speech Recognition.	473
<i>Markus Müller, Sebastian Stüker, and Alex Waibel</i>	
Language Model Optimization for a Deep Neural Network Based Speech Recognition System for Serbian	483
<i>Edvin Pakoci, Branislav Popović, and Darko Pekar</i>	
Lexico-Semantical Indices of “Deprivation – Aggression” Modality Correlation in Social Network Discourse	493
<i>Rodmonga Potapova and Liliya Komalova</i>	
Linguistic Features and Sociolinguistic Variability in Everyday Spoken Russian	503
<i>Natalia Bogdanova-Beglarian, Tatiana Sherstinova, Olga Blinova, and Gregory Martynenko</i>	
Medical Speech Recognition: Reaching Parity with Humans.	512
<i>Erik Edwards, Wael Salloum, Greg P. Finley, James Fone, Greg Cardiff, Mark Miller, and David Suendermann-Oeft</i>	
Microphone Array Post-filter in Frequency Domain for Speech Recognition Using Short-Time Log-Spectral Amplitude Estimator and Spectral Harmonic/Noise Classifier	525
<i>Sergey Salishev, Ilya Klotchkov, and Andrey Barabanov</i>	
Multimodal Keyword Search for Multilingual and Mixlingual Speech Corpus.	535
<i>Abhimanyu Popli and Arun Kumar</i>	
Neural Network Doc2vec in Automated Sentiment Analysis for Short Informal Texts	546
<i>Natalia Maslova and Vsevolod Potapov</i>	
Neural Network Speaker Descriptor in Speaker Diarization of Telephone Speech	555
<i>Zbyněk Zajíc, Jan Zelinka, and Luděk Müller</i>	
Novel Linear Prediction Temporal Phase Based Features for Speaker Recognition	564
<i>Ami Gandhi and Hemant A. Patil</i>	
Novel Phase Encoded Mel Cepstral Features for Speaker Verification	572
<i>Apeksha J. Naik, Rishabh Tak, and Hemant A. Patil</i>	

On a Way to the Computer Aided Speech Intonation Training	582
<i>Boris Lobanov, Yelena Karnevskaia, and Vladimir Zhitko</i>	
On Residual CNN in Text-Dependent Speaker Verification Task.	593
<i>Egor Malykh, Sergey Novoselov, and Oleg Kudashev</i>	
Perception and Acoustic Features of Speech of Children with Autism Spectrum Disorders	602
<i>Elena Lyakso, Olga Frolova, and Aleksey Grigorev</i>	
Phase Analysis and Labeling Strategies in a CNN-Based Speaker Change Detection System	613
<i>Marek Hruš and Petr Salajka</i>	
Preparing Audio Recordings of Everyday Speech for Prosody Research: The Case of the ORD Corpus.	623
<i>Tatiana Sherstinova</i>	
Recognizing Emotionally Coloured Dialogue Speech Using Speaker-Adapted DNN-CNN Bottleneck Features	632
<i>Kohei Mukaihara, Sakriani Sakti, and Satoshi Nakamura</i>	
Relationship Between Perception of Cuteness in Female Voices and Their Durations.	642
<i>Ryohei Ohno, Masanori Morise, and Tetsuro Kitahara</i>	
Retaining Expression on De-identified Faces	651
<i>Li Meng and Aruna Shenoy</i>	
Semi-automatic Facial Key-Point Dataset Creation	662
<i>Miroslav Hlaváč, Ivan Gruber, Miloš Železný, and Alexey Karpov</i>	
Song Emotion Recognition Using Music Genre Information	669
<i>Athanasios Koutras</i>	
Spanish Corpus for Sentiment Analysis Towards Brands	680
<i>Maria Navas-Loro, Víctor Rodríguez-Doncel, Idafen Santana-Perez, and Alberto Sánchez</i>	
Speech Enhancement for Speaker Recognition Using Deep Recurrent Neural Networks.	690
<i>Maxim Tkachenko, Alexander Yamshinin, Nikolay Lyubimov, Mikhail Kotov, and Marina Nastasenko</i>	
Stance Classification in Texts from Blogs on the 2016 British Referendum	700
<i>Vasiliki Simaki, Carita Paradis, and Andreas Kerren</i>	

The “Retrospective Commenting” Method for Longitudinal Recordings of Everyday Speech.	710
<i>Arto Mustajoki and Tatiana Sherstinova</i>	
The 2016 RWTH Keyword Search System for Low-Resource Languages. . . .	719
<i>Pavel Golik, Zoltán Tüske, Kazuki Irie, Eugen Beck, Ralf Schlüter, and Hermann Ney</i>	
The Effect of Morphological Factors on Sentence Boundaries in Russian Spontaneous Speech	731
<i>Anton Stepikhov and Anastassia Loukina</i>	
The Pausing Method Based on Brown Clustering and Word Embedding	741
<i>Arman Kaliyev, Sergey V. Rybin, and Yuri Matveev</i>	
Unsupervised Document Classification and Topic Detection	748
<i>Jaromír Novotný and Pavel Ircing</i>	
Using a High-Speed Video Camera for Robust Audio-Visual Speech Recognition in Acoustically Noisy Conditions	757
<i>Denis Ivanko, Alexey Karpov, Dmitry Ryumin, Irina Kipyatkova, Anton Saveliev, Victor Budkov, Dmitriy Ivanko, and Miloš Železný</i>	
Utilizing Lipreading in Large Vocabulary Continuous Speech Recognition. . .	767
<i>Karel Paleček</i>	
Vocal Emotion Conversion Using WSOLA and Linear Prediction.	777
<i>Susmitha Vekkot and Shikha Tripathi</i>	
Voice Conversion for TTS Systems with Tuning on the Target Speaker Based on GMM	788
<i>Vadim Zahariev, Elias Azarov, and Alexander Petrovsky</i>	
VoiScan: Telephone Voice Analysis for Health and Biometric Applications.	799
<i>Ladan Baghai-Ravary and Steve W. Beet</i>	
Web Queries Classification Based on the Syntactical Patterns of Search Types	809
<i>Alaa Mohasseb, Mohamed Bader-El-Den, Andreas Kanavos, and Mihaela Cocea</i>	
What Speech Recognition Accuracy is Needed for Video Transcripts to be a Useful Search Interface?	820
<i>Yang Chao and Marie-Luce Bourguet</i>	
Author Index	829

Speech and Computer

19th International Conference, SPECOM 2017, Hatfield,

UK, September 12-16, 2017, Proceedings

Karpov, A.; Potapova, R.; Mporas, I. (Eds.)

2017, XV, 831 p. 255 illus., Softcover

ISBN: 978-3-319-66428-6