

Contents

Invited Talks

Automatic Speech Recognition Based on Neural Networks.	3
<i>Ralf Schlüter, Patrick Doetsch, Pavel Golik, Markus Kitza, Tobias Menne, Kazuki Irie, Zoltán Tüske, and Albert Zeyer</i>	
Machine Processing of Dialogue States; Speculations on Conversational Entropy	18
<i>Nick Campbell</i>	
Speech Recognition Challenges in the Car Navigation Industry.	26
<i>Attila Vékony</i>	

Conference Papers

A Comparison of Acoustic Features of Speech of Typically Developing Children and Children with Autism Spectrum Disorders.	43
<i>Elena Lyakso, Olga Frolova, and Aleksey Grigorev</i>	
A Deep Neural Networks (DNN) Based Models for a Computer Aided Pronunciation Learning System.	51
<i>Mohamed S. Elaraby, Mustafa Abdallah, Sherif Abdou, and Mohsen Rashwan</i>	
A Linguistic Interpretation of the Atom Decomposition of Fundamental Frequency Contour for American English.	59
<i>Tijana Delić, Branislav Gerazov, Branislav Popović, and Milan Sečujski</i>	
A Phonetic Segmentation Procedure Based on Hidden Markov Models	67
<i>Edvin Pakoci, Branislav Popović, Nikša Jakovljević, Darko Pekar, and Fathy Yassa</i>	
A Preliminary Exploration of Group Social Engagement Level Recognition in Multiparty Casual Conversation	75
<i>Yuyun Huang, Emer Gilmartin, Benjamin R. Cowan, and Nick Campbell</i>	
An Agonist-Antagonist Pitch Production Model	84
<i>Branislav Gerazov and Philip N. Garner</i>	
An Algorithm for Phase Manipulation in a Speech Signal	92
<i>Darko Pekar, Siniša Suzić, Robert Mak, Meir Friedlander, and Milan Sečujski</i>	

An Exploratory Study on Sociolinguistic Variation of Russian Everyday Speech.	100
<i>Natalia Bogdanova-Beglarian, Tatiana Sherstinova, Olga Blinova, and Gregory Martynenko</i>	
Adaptation of DNN Acoustic Models Using KL-divergence Regularization and Multi-task Training	108
<i>László Tóth and Gábor Gosztolya</i>	
Advances in STC Russian Spontaneous Speech Recognition System	116
<i>Ivan Medennikov and Alexey Prudnikov</i>	
Approaches for Out-of-Domain Adaptation to Improve Speaker Recognition Performance	124
<i>Andrey Shulipa, Sergey Novoselov, and Aleksandr Melnikov</i>	
Assessment of the Relation Between Low-Frequency Features and Velum Opening by Using Real Articulatory Data	131
<i>Alexander Sepulveda-Sepulveda and German Castellanos-Dominguez</i>	
Automatic Summarization of Highly Spontaneous Speech	140
<i>András Beke and György Szaszák</i>	
Backchanneling via Twitter Data for Conversational Dialogue Systems	148
<i>Michimasa Inaba and Kenichi Takahashi</i>	
Bio-Inspired Sparse Representation of Speech and Audio Using Psychoacoustic Adaptive Matching Pursuit.	156
<i>Alexey Petrovsky, Vadzim Herasimovich, and Alexander Petrovsky</i>	
Combining Atom Decomposition of the F0 Track and HMM-based Phonological Phrase Modelling for Robust Stress Detection in Speech.	165
<i>György Szaszák, Máté Ákos Tündik, Branislav Gerazov, and Aleksandar Gjoreski</i>	
Comparative Analysis of Classifiers for Automatic Language Recognition in Spontaneous Speech	174
<i>Konstantin Simonchik, Sergey Novoselov, and Galina Lavrentyeva</i>	
Comparison of Retrieval Approaches and Blind Relevance Feedback Methods Within the Czech Speech Information Retrieval	182
<i>Lucie Skorkovská</i>	
Convolutional Neural Network in the Task of Speaker Change Detection.	191
<i>Marek Hríz and Marie Kunešová</i>	

Design of a Speech Corpus for Research on Cross-Lingual Prosody Transfer	199
<i>Milan Sečujiški, Branislav Gerazov, Tamás Gábor Csapó, Vlado Delić, Philip N. Garner, Aleksandar Gjoreski, David Guennec, Zoran Ivanovski, Aleksandar Melov, Géza Németh, Ana Stojković, and György Szaszák</i>	
Designing High-Coverage Multi-level Text Corpus for Non-professional-voice Conservation	207
<i>Markéta Jůzová, Daniel Tihelka, and Jindřich Matoušek</i>	
Designing Syllable Models for an HMM Based Speech Recognition System . . .	216
<i>Kseniya Proença, Kris Demuynck, and Dirk Van Compernelle</i>	
Detecting Filled Pauses and Lengthenings in Russian Spontaneous Speech Using SVM	224
<i>Vasilisa Verkhodanova and Vladimir Shapranov</i>	
Detecting Laughter and Filler Events by Time Series Smoothing with Genetic Algorithms	232
<i>Gábor Gosztolya</i>	
Detecting State of Aggression in Sentences Using CNN.	240
<i>Denis Gordeev</i>	
DNN-Based Acoustic Modeling for Russian Speech Recognition Using Kaldi	246
<i>Irina Kipyatkova and Alexey Karpov</i>	
DNN-Based Duration Modeling for Synthesizing Short Sentences.	254
<i>Péter Nagy and Géza Németh</i>	
Emotional Speech of 3-Years Old Children: Norm-Risk-Deprivation	262
<i>Olga Frolova and Elena Lyakso</i>	
Ensemble Deep Neural Network Based Waveform-Driven Stress Model for Speech Synthesis	271
<i>Bálint Pál Tóth, Kornél István Kis, György Szaszák, and Géza Németh</i>	
Evaluation of Response Times on a Touch Screen Using Stereo Panned Speech Command Auditory Feedback	279
<i>Hunor Nagy and György Wersényi</i>	
Evaluation of the Speech Quality During Rehabilitation After Surgical Treatment of the Cancer of Oral Cavity and Oropharynx Based on a Comparison of the Fourier Spectra.	287
<i>Evgeny Kostyuchenko, Mescheryakov Roman, Dariya Ignatieva, Alexander Pyatkov, Evgeny Choyazonov, and Lidiya Balatskaya</i>	

Experiments with One-Class Classifier as a Predictor of Spectral Discontinuities in Unit Concatenation	296
<i>Daniel Tihelka, Martin Grüber, and Markéta Jůzová</i>	
Exploring GMM-derived Features for Unsupervised Adaptation of Deep Neural Network Acoustic Models	304
<i>Natalia Tomashenko, Yuri Khokhlov, Anthony Larcher, and Yannick Estève</i>	
Feature Space VTS with Phase Term Modeling	312
<i>Maxim Korenevsky and Aleksei Romanenko</i>	
Finding Speaker Position Under Difficult Acoustic Conditions	321
<i>Evgeniy Shuranov, Aleksandr Lavrentyev, Alexey Kozlyayev, Galina Lavrentyeva, and Valeriya Volkovaya</i>	
Fusing Various Audio Feature Sets for Detection of Parkinson's Disease from Sustained Voice and Speech Recordings.	328
<i>Evaldas Vaiciukynas, Antanas Verikas, Adas Gelzinis, Marija Bacauskiene, Kestutis Vaskevicius, Virgilijus Uloza, Evaldas Padervinskis, and Jolita Ciceliene</i>	
HAVRUS Corpus: High-Speed Recordings of Audio-Visual Russian Speech.	338
<i>Vasilisa Verkhodanova, Alexander Ronzhin, Irina Kipyatkova, Denis Ivanko, Alexey Karpov, and Miloš Železný</i>	
Human-Smartphone Interaction for Dangerous Situation Detection and Recommendation Generation While Driving.	346
<i>Alexander Smirnov, Alexey Kashevnik, and Igor Lashkov</i>	
Improving Automatic Speech Recognition Containing Additive Noise Using Deep Denoising Autoencoders of LSTM Networks	354
<i>Marvin Coto-Jiménez, John Goddard-Close, and Fabiola Martínez-Licona</i>	
Improving the Quality of Automatic Speech Recognition in Trucks.	362
<i>Maxim Korenevsky, Ivan Medennikov, and Vadim Shchemelinin</i>	
Improving Recognition of Dysarthric Speech Using Severity Based Tempo Adaptation	370
<i>Chitralkha Bhat, Bhavik Vachhani, and Sunil Kopparapu</i>	
Improving Robustness of Speaker Verification by Fusion of Prompted Text-Dependent and Text-Independent Operation Modalities.	378
<i>Iosif Mporas, Saeid Safavi, and Reza Sotudeh</i>	

Improvements to Prosodic Variation in Long Short-Term Memory Based Intonation Models Using Random Forest	386
<i>Bálint Pál Tóth, Balázs Szórádi, and Géza Németh</i>	
In-Document Adaptation for a Human Guided Automatic Transcription Service	395
<i>André Mansikkaniemi, Mikko Kurimo, and Krister Lindén</i>	
Interaction Quality as a Human-Human Task-Oriented Conversation Performance	403
<i>Anastasiia Spirina, Olesia Vaskovskaia, Maxim Sidorov, and Alexander Schmitt</i>	
Investigation of Segmentation in i-Vector Based Speaker Diarization of Telephone Speech	411
<i>Zbyněk Zajíc, Marie Kunešová, and Vlasta Radová</i>	
Investigation of Speech Signal Parameters Reflecting the Truth of Transmitted Information	419
<i>Victor Budkov, Irina Vatamaniuk, Vladimir Basov, and Daniyar Volf</i>	
Investigating Signal Correlation as Continuity Metric in a Syllable Based Unit Selection Synthesis System	427
<i>Sai Sirisha Rallabandi, Sai Krishna Rallabandi, Naina Teertha, Kumaraswamy R., and Suryakanth V. Gangashetty</i>	
Knowledge Transfer for Utterance Classification in Low-Resource Languages	435
<i>Andrei Smirnov and Valentin Mendelev</i>	
Language Identification Using Time Delay Neural Network D-Vector on Short Utterances	443
<i>Maxim Tkachenko, Alexander Yamshinin, Nikolay Lyubimov, Mikhail Kotov, and Marina Nastasenko</i>	
Lexical Stress in Punjabi and Its Representation in PLS	450
<i>Swaran Lata, Swati Arora, and Simerjeet Kaur</i>	
Low Inter-Annotator Agreement in Sentence Boundary Detection and Annotator Personality	461
<i>Anton Stepikhov and Anastassia Loukina</i>	
LSTM-Based Language Models for Spontaneous Speech Recognition	469
<i>Ivan Medennikov and Anna Bulusheva</i>	

Measuring Prosodic Entrainment in Italian Collaborative Game-Based Dialogues.	476
<i>Michelina Savino, Loredana Lapertosa, Alessandro Caffò, and Mario Refice</i>	
Microphone Array Directivity Improvement in Low-Frequency Band for Speech Processing	484
<i>Mikhail Stolbov and Sergei Aleinik</i>	
Modeling Imperative Utterances in Russian Spoken Dialogue: Verb-Central Quantitative Approach	491
<i>Olga Blinova</i>	
Multimodal Perception of Aggressive Behavior.	499
<i>Rodmonga Potapova and Liliya Komalova</i>	
On Individual Polyinformativity of Speech and Voice Regarding Speakers Auditive Attribution (Forensic Phonetic Aspect)	507
<i>Rodmonga Potapova and Vsevolod Potapov</i>	
Online Biometric Identification with Face Analysis in Web Applications	515
<i>Gerasimos Arvanitis, Konstantinos Moustakas, and Nikos Fakotakis</i>	
Optimization of Zelinski Post-filtering Calculation.	523
<i>Sergei Aleinik</i>	
Phonetic Aspects of High Level of Naturalness in Speech Synthesis	531
<i>Vera Evdokimova, Pavel Skrelin, Andrey Barabanov, and Karina Evgrafova</i>	
Polybasic Attribution of Social Network Discourse	539
<i>Rodmonga Potapova and Vsevolod Potapov</i>	
Precise Estimation of Harmonic Parameter Trend and Modification of a Speech Signal	547
<i>Andrey Barabanov, Valentin Magerkin, and Evgenij Vikulov</i>	
Profiling a Set of Personality Traits of a Text's Author: A Corpus-Based Approach.	555
<i>Tatiana Litvinova, Olga Zagorovskaya, Olga Litvinova, and Pavel Seredin</i>	
Prosody Analysis of Malay Language Storytelling Corpus	563
<i>Izzad Ramli, Noraini Seman, Norizah Ardi, and Nursuriati Jamil</i>	
Quality Assessment of Two Fullband Audio Codecs Supporting Real-Time Communication.	571
<i>M. Maruschke, O. Jokisch, M. Meszaros, F. Trojahn, and M. Hoffmann</i>	

Robust Speech Analysis Based on Source-Filter Model Using Multivariate Empirical Mode Decomposition in Noisy Environments	580
<i>Surasak Boonkla, Masashi Unoki, and Stanislav S. Makhanov</i>	
Scenarios of Multimodal Information Navigation Services for Users in Cyberphysical Environment	588
<i>Irina Vatamaniuk, Dmitriy Levonevskiy, Anton Saveliev, and Alexander Denisov</i>	
Scores Calibration in Speaker Recognition Systems	596
<i>Andrey Shulipa, Sergey Novoselov, and Yuri Matveev</i>	
Selecting Keypoint Detector and Descriptor Combination for Augmented Reality Application	604
<i>Lukáš Bureš and Luděk Müller</i>	
Semi-automatic Speaker Verification System Based on Analysis of Formant, Durational and Pitch Characteristics	613
<i>Elena Bulgakova and Aleksey Sholohov</i>	
Speaker-Dependent Bottleneck Features for Egyptian Arabic Speech Recognition	620
<i>Aleksei Romanenko and Valentin Mendelev</i>	
Speech Acts Annotation of Everyday Conversations in the ORD Corpus of Spoken Russian	627
<i>Tatiana Sherstinova</i>	
Speech Enhancement with Microphone Array Using a Multi Beam Adaptive Noise Suppressor	636
<i>Mikhail Stolbov and Alexander Lavrentyev</i>	
Speech Features Evaluation for Small Set Automatic Speaker Verification Using GMM-UBM System	645
<i>Ivan Rakhmanenko and Roman Meshcheryakov</i>	
Speech Recognition Combining MFCCs and Image Features	651
<i>Stamatis Karlos, Nikos Fazakis, Katerina Karanikola, Sotiris Kotsiantis, and Kyriakos Sgarbas</i>	
Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech. . . .	659
<i>Natalia Bogdanova-Beglarian, Tatiana Sherstinova, Olga Blinova, Olga Ermolova, Ekaterina Baeva, Gregory Martynenko, and Anastasia Ryko</i>	
Statistical Analysis of Acoustical Parameters in the Voice of Children with Juvenile Dysphonia	667
<i>Miklós Gábor Tulics, Ferenc Kazinczi, and Klára Vicsi</i>	

Stress, Arousal, and Stress Detector Trained on Acted Speech Database.	675
<i>Róbert Sabo, Milan Rusko, Andrej Ridzik, and Jakub Rajčáni</i>	
Study on the Improvement of Intelligibility for Elderly Speech Using Formant Frequency Shift Method	683
<i>Yuto Tanaka, Mitsunori Mizumachi, and Yoshihisa Nakatoh</i>	
Text Classification in the Domain of Applied Linguistics as Part of a Pre-editing Module for Machine Translation Systems	691
<i>Ksenia Oskina</i>	
Tonal Specification of Perceptually Prominent Non-nuclear Pitch Accents in Russian	699
<i>Nina Volskaya and Tatiana Kachkovskaia</i>	
Toward Sign Language Motion Capture Dataset Building	706
<i>Zdeněk Krňoul, Pavel Jedlička, Jakub Kanis, and Miloš Železný</i>	
Trade-Off Between Speed and Accuracy for Noise Variance Minimization (NVM) Pitch Estimation Algorithm	714
<i>Andrey Barabanov and Aleksandr Melnikov</i>	
Unsupervised Trained Functional Discourse Parser for e-Learning Materials Scaffolding.	722
<i>Varvara Krayvanova and Svetlana Duka</i>	
Author Index	729

Speech and Computer

18th International Conference, SPECOM 2016,

Budapest, Hungary, August 23-27, 2016, Proceedings

Ronzhin, A.; Potapova, R.; Németh, G. (Eds.)

2016, XVIII, 731 p. 197 illus., Softcover

ISBN: 978-3-319-43957-0