

**23rd International Conference on Speech and Computer
6th International Conference on Interactive Collaborative Robotics**

SPECOM & ICR 2021

www.specom.nw.ru/2021

Joint Conference Program

27-30 September 2021

St. Petersburg, Russia => Online



SPECOM 2021

ICR 2021

Organizers



**SPC
RAS**



TUM

General Sponsor



HUAWEI

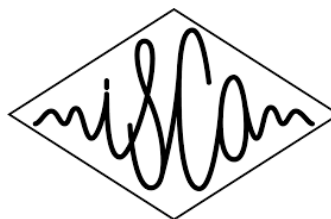
Sponsors



ASM Solutions



Supporters



Service Agency



SPECOM and ICR History

- 23rd SPECOM-2021 and 6th ICR-2021, Online
- 22nd SPECOM-2020 and 5th ICR-2020, Online
- 21st SPECOM-2019 and 4th ICR-2019, Istanbul, Turkey
- 20th SPECOM-2018 and 3rd ICR-2018, Leipzig, Germany
- 19th SPECOM-2017 and 2nd ICR-2017, Hatfield, United Kingdom
- 18th SPECOM-2016 and 1st ICR-2016, Budapest, Hungary
- 17th SPECOM-2015, Athens, Greece
- 16th SPECOM-2014, Novi Sad, Serbia
- 15th SPECOM-2013, Pilsen, Czech Republic
- 14th SPECOM-2011, Kazan, Russia
- 13th SPECOM-2009, St. Petersburg, Russia
- 12th SPECOM-2007, Moscow, Russia
- 11th SPECOM-2006, St. Petersburg
- 10th SPECOM-2005, Patras, Greece
- 9th SPECOM-2004, St. Petersburg
- 8th SPECOM-2003, Moscow
- 7th SPECOM-2002, St. Petersburg
- 6th SPECOM-2001, Moscow
- 5th SPECOM-2000, St. Petersburg
- 4th SPECOM-1999, Moscow
- 3rd SPECOM-1998, St. Petersburg
- 2nd SPECOM-1997, Cluj-Napoca
- 1st SPECOM-1996, St. Petersburg

Organizers

The SPECOM and ICR conferences are organized by

- St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS) of the St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg, Russia

In a cooperation with

- Moscow State Linguistic University (MSLU, Moscow, Russia)
- Technical University of Munich (TUM, Munich, Germany)

Committees

General Chairs (SPECOM)

Alexey Karpov, SPC RAS, St. Petersburg
Rodmonga Potapova, MSLU, Moscow

General Chairs (ICR)

Andrey Ronzhin, SPC RAS, St. Petersburg
Gerhard Rigoll, TUM, Munich, Germany
Roman Meshcheryakov, ICS RAS, Moscow

Program Committee (SPECOM)

Shyam Agrawal, India
Tanel Alumäe, Estonia
Elias Azarov, Belarus
Peter Birkholz, Germany
Milana Bojanic, Serbia
Nick Campbell, Ireland
Eric Castelli, France
Josef Chaloupka, Czech Republic
Khalid Choukri, France
Vladimir Chuchupal, Russia
Vlado Delić, Serbia
Olivier Deroo, France
Febe De Wet, South Africa
Anna Esposito, Italy
Yannick Estève, France
Vera Evdokimova, Russia
Mauro Falcone, Italy
Vasiliki Foufi, Switzerland
Todor Ganchev, Bulgaria
Philip Garner, Switzerland
Gábor Gosztolya, Hungary
Tunga Gungör, Turkey
Abualseoud Hanani, Palestine
Ruediger Hoffmann, Germany
Marek Hruz, Czech Republic
Kristiina Jokinen, Japan
Oliver Jokisch, Germany
Denis Juvet, France
Tatiana Kachkovskaia, Russia
Alexey Karpov, Russia
Heysem Kaya, The Netherlands
Tomi Kinnunen, Finland
Irina Kipyatkova, Russia
Daniil Kocharov, Russia

Liliya Komalova, Russia
Evgeny Kostyuchenko, Russia
Mikko Kurimo, Finland
Benjamin Lecouteux, France
Natalia Loukachevitch, Russia
Elena Lyakso, Russia
Joseph Mariani, France
Konstantin Markov, Japan
Jindřich Matoušek, Czech Republic
Yuri Matveev, Russia
Ivan Medennikov, Russia
Peter Mihajlik, Hungary
Nobuaki Minematsu, Japan
Wolfgang Minker, Germany
Iosif Mporas, UK
Ludek Muller, Czech Republic
Bernd Möbius, Germany
Sebastian Möller, Germany
Satoshi Nakamura, Japan
Géza Németh, Hungary
Oliver Niebuhr, Denmark
Thomas Niesler, South Africa
Sergey Novoselov, Russia
Stavros Ntalampiras, Italy
Hemant Patil, India
Alexey Petrovsky, Russia
Nickolay Petrovsky, Belarus
Dimitar Popov, Bulgaria
Branislav Popović, Serbia
Rodmonga Potapova, Russia
Vsevolod Potapov, Russia
Fabien Ringeval, France
Gerhard Rigoll, Germany
Paolo Rosso, Spain

Leon Rothkrantz, The Netherlands
Milan Rusko, Slovakia
Sergey Rybin, Russia
Sakriani Sakti, Japan
Albert Ali Salah, The Netherlands
Andrey Savchenko, Russia
Maximilian Schmitt, Germany
Tanja Schultz, Germany
Friedhelm Schwenker, Germany
Milan Sečujski, Serbia
Tatiana Sherstinova, Russia
Tatiana Shevchenko, Russia
Anastasia Shimorina, France
Nickolay Shmyrev, Russia
Vered Silber-Varod, Israel
Vasiliki Simaki, Sweden
Pavel Skrelin, Russia
Sebastian Stüker, Germany
György Szaszák, Hungary
Ivan Tashev, USA
António Teixeira, Portugal
Natalia Tomashenko, France
Laszlo Toth, Hungary
Isabel Trancoso, Portugal
Jan Trmal, USA
Liliya Tsurulnik, USA
Dirk Van Compernelle, Belgium
Vasilisa Verkhodanova, The Netherlands
Matthias Wolff, Germany
Junichi Yamagishi, Japan
Zeynep Yucel, Japan
Miloš Železný, Czech Republic

Program Committee (ICR)

Andrey Ronzhin, Russia (co-chair)
Roman Meshcheryakov,
Russia (co-chair)
Gerhard Rigoll, Germany (co-chair)
Christos Antonopoulos, Greece
Branislav Borovac, Serbia
Sara Chaychian, UK
Ivan Ermolov, Russia
Oliver Jokisch, Germany

Dimitrios Kalles, Greece
Igor Kalyaev, Russia
Alexey Kashevnik, Russia
Dongheui Lee, Germany
Viktor Glazunov, Russia
Mehmet Guzey, Turkey
Evgeni Magid, Russia
Iosif Mporas, UK
Viacheslav Pshikhopov, Russia

Mirko Rakovic, Serbia
Yulia Sandamirskaya, Switzerland
Jesus Savage, Mexico
Hooman Samani, UK
Evgeny Shandarov, Russia
Lev Stankevich, Russia
Tilo Strutz, Germany
Sergey Yatsun, Russia

Additional Reviewers (SPECOM)

Alexandr Axyonov, Russia
Denis Dresvyanskiy, Germany
Stefan Hillmann, Germany
Denis Ivanko, Russia
Nikša Jakovljević, Serbia
Ildar Kagirow, Russia
Andrey Kuznetsov, Russia
Ilya Makarov, Russia
Olesia Makhnytkina, Russia

Danila Mamontov, Germany
Maxim Markitantov, Russia
Dragisa Miskovic, Serbia
Tijana Nosek, Serbia
Lukáš Pícek, Czech Republic
Ana P. Rocha, Portugal
Dmitry Ryumin, Russia
Petr Salajka, Czech Republic
Samuel Silva, Portugal

Nikola Simić, Serbia
Lubos Smidl, Czech Republic
Siniša Suzić, Serbia
Ana Rita Valente, Portugal
Alena Velichko, Russia
Oxana Verkholyak, Russia
Alisa Vinogradova, Russia
Zbyněk Zajíc, Czech Republic

Local Organising Committee

Alexey Karpov (Chair)
Andrey Ronzhin
Rodmonda Potapova
Daniil Kocharov
Anton Saveliev
Irina Kipyatkova

Dmitry Ryumin
Dmitriy Levonevskiy
Natalia Kashina
Anastasiya Molotilova
Irina Novikova
Natalia Dormidontova

Irina Vatamanyuk
Polina Chernousova
Irina Podnozova
Denis Ivanko

Invited Speaker

Prof. Junichi Yamagishi

National Institute of Informatics, Tokyo, Japan



Junichi Yamagishi was awarded a Ph.D. by Tokyo Institute of Technology, Japan in 2006. From 2007-2013 he was a research fellow in the Centre for Speech Technology Research (CSTR) at the University of Edinburgh, UK. He was appointed Associate Professor at National Institute of Informatics, Japan in 2013. He is currently a Professor at NII, Japan. His research topics include speech processing, machine learning, signal processing, biometrics, digital media cloning and media forensics. He served previously as co-organizer for the bi-annual ASVspoof challenge and the bi-annual Voice conversion challenge. He also served as a member of the IEEE Speech and Language Technical Committee (2013-2019), an Associate Editor of the IEEE/ACM Transactions on Audio Speech and Language Processing (2014-2017), and a chairperson of ISCA SynSIG (2017- 2021). He is currently a PI of JST-CREST and ANR supported VoicePersona project and a Senior Area Editor of the IEEE/ACM TASLP.

Keynote Lecture

Detection of Generated Voices and Faces – ASVspoof and Deepfake Detection

In this talk, we introduce our recent projects on automatic detection of generated voices or faces. For detection of computer-generated voices, we have constructed a large dataset that contains various types of synthesized and converted speech and organized the ASVspoof challenge where over 50 teams competed for the best anti-spoofing performance. Based on analysis of the challenge results, we show important lessons learned from both the attacker and defender perspectives. In the latter part of the talk, we describe another project on deepfake detection and show new countermeasure models using capsule networks or multi-task models that indicate manipulated facial regions for improving explainability.

Invited Speaker

Dr. Mac Schwager

Stanford University, Stanford, USA



Mac Schwager is an Assistant Professor of Aeronautics and Astronautics at Stanford University. He directs the Multi-robot Systems Lab (MSL) where he studies distributed algorithms for control, perception, and learning in groups of robots and autonomous systems. He is interested in a range of applications including cooperative surveillance with teams of UAVs, autonomous driving in cluttered, dynamic environments, cooperative robotic manipulation, distributed SLAM, distributed trajectory planning, and autonomous drone and car racing. He obtained his BS degree from Stanford, and his MS and PhD degrees from MIT. He was a postdoctoral researcher at the University of Pennsylvania and at MIT. Prior to joining Stanford, he was an assistant professor at Boston University from 2012 to 2015. He received the NSF CAREER award in 2014, the DARPA YFA in 2018, and has received numerous best paper awards in conferences and journals including the IEEE Transactions on Robotics best paper award in 2016, the Best Paper Award in Robot Manipulation in ICRA 2018, and the Best Paper Award in Multi-Robot Systems in ICRA 2020.

Keynote Lecture

Enabling Robots to Cooperate and Compete: Distributed Optimization and Game Theoretic Methods for Multiple Interacting Robots

For robots to effectively operate in our world, they must master the skills of dynamic interaction. Autonomous cars must safely negotiate their trajectories with other vehicles and pedestrians as they drive to their destinations. UAVs must avoid collisions with other aircraft, as well as dynamic obstacles on the ground. Disaster response robots must coordinate to explore and map new disaster sites. In this talk I will describe recent work in my lab using distributed optimization to obtain algorithms for robots to cooperate, and game theoretic methods to obtain algorithms for robots to compete. I will describe a general and flexible method, called SOVA, for deriving distributed optimization algorithms for a variety of multi-robot collaborative problems. I will present an algorithm for fleets of autonomous cars to cooperatively track a large number of vehicles and pedestrians in a city, an algorithm for multiple robots to manipulate an object to a goal while avoiding collisions, and a distributed multi-robot SLAM algorithm, all derived using the SOVA method. I will also discuss algorithms based on the theory of dynamic games, in which each actor has its own objective and constraints. I will describe examples in autonomous drone racing, car racing, and autonomous driving that use game theoretic principles to solve for Nash equilibrium trajectories in real-time, in a receding horizon fashion. Throughout the talk, I will show results from hardware experiments with ground robots, autonomous cars, and quadrotor UAVs collaborating and competing in the scenarios above.

Keynote Lecture by General Sponsor



AI enabled Speech Processing for Smart Devices

Speakers

D.Sc. Alexey Petrovsky

Huawei, Russian Research Center, Moscow Media Algorithm Laboratory, Russia

and

Dr. Evgeniy Shuranov

Huawei, Russian Research Center, St. Petersburg CBG AI Laboratory, Russia

Program at a Glance

Day 0: 27 September 2021, Monday (Room 1)	
13:00-18:00	Zoom setup and test connections
Day 1: 28 September 2021, Tuesday (Rooms 1 & 2)	
09:50-10:00	Zoom connection
10:00-10:20	Opening ceremony (Room 1)
10:20-11:20	Keynote lecture by Prof. JUNICHI YAMAGISHI (Japan): Automatic Detection of Generated Voices and Faces: ASVspoof and Deepfake Detection
11:20-11:30	Group photo in Zoom
11:30-12:30	Keynote lecture by HUAWEI (Russian Research Center): AI enabled Speech Processing for Smart Devices
12:30-13:30	Lunch break & Online entertainment
13:30-15:30	SPECOM Session 1 (6 full presentations, 20 min) Speaker Recognition
	ICR Session 1 (6 full presentations, 20 min) Collaborative Robots
15:30-15:50	Short break
15:50-17:30	SPECOM Session 2 (10 short presentations, 10 min) Speech and Audio Signal Processing
	ICR Session 2 (5 full presentations, 20 min) Robot Control Systems
17:30-17:40	Short break
17:40-18:40	Keynote lecture by Prof. MAC SCHWAGER (USA): Enabling Robots to Cooperate and Compete: Distributed Optimization and Game Theoretic Methods for Multiple Interacting Robots (Room 1)
Day 2: 29 September 2021, Wednesday (Rooms 1 & 2)	
09:50-10:00	Zoom connection
10:00-12:00	SPECOM Session 3 (6 full presentations, 20 min) Automatic Speech Recognition
	ICR Session 3 (8 short presentations, 15 min) Robotic Systems
12:00-12:20	Short break
12:20-13:50	SPECOM Session 4 (9 short presentations, 10 min) Speech Recognition and Understanding
13:50-15:00	Lunch break & Online entertainment
15:00-17:00	SPECOM Session 5 (6 full presentations, 20 min) Speech Synthesis
17:00-17:20	Short break
17:20-19:00	SPECOM Session 6 (5 full presentations, 20 min) Computational Paralinguistics
	SPECOM Session 7 (5 full presentations, 20 min) Multimodal Analysis and Synthesis
Day 3: 30 September 2021, Thursday (Room 2)	
09:50-10:00	Zoom connection
10:00-12:00	SPECOM Session 8 (6 full presentations, 20 min) Spoken Language Processing
12:00-12:20	Short break
12:20-13:50	SPECOM Session 9 (9 short presentations, 10 min) Spoken Language Analysis
13:50-15:00	Lunch break
15:00-17:00	SPECOM Session 10 (6 full presentations, 20 min) Natural Language Processing
17:00-17:20	Short break
17:20-18:20	SPECOM Session 11 (6 short presentations, 10 min) Speech and Audio Signal Processing
18:20-18:30	Closing ceremony

Detailed Scientific Program

Day 0: 27 September 2021, Monday (Room 1)		
13:00-18:00	Zoom setup and test connections	
Day 1: 28 September 2021, Tuesday (Rooms 1 & 2)		
09:50-10:00	Zoom connection	
10:00-10:20	Opening ceremony (Room 1)	
10:20-11:20	Keynote lecture by Prof. JUNICHI YAMAGISHI (Japan): “Automatic Detection of Generated Voices and Faces: ASVspoof and Deepfake Detection” <i>Chair: Alexey Karpov (Room 1)</i>	
11:20-11:30	Group photo in Zoom	
11:30-12:30	Keynote lecture by HUAWEI (Russian Research Center): “AI enabled Speech Processing for Smart Devices” <i>Chair: Alexey Karpov (Room 1)</i>	
12:30-13:30	Lunch break & Online entertainment (Room 1)	
13:30-15:30	SPECOM Session 1 (6 full presentations, 20 min) Speaker Recognition (Room 1) <i>Chair: Iosif Mporas</i>	ICR Session 1 (6 full presentations, 20 min) Collaborative Robots (Room 2) <i>Chair: Konstantin Yakovlev</i>
13:30-13:50	Influence of Speaker Pre-training on Character Voice Representation <i>Mathias Quillot, Jarod Duret, Richard Dufour, Mickaël Rouvier and Jean-Francois Bonastre (France)</i>	Step Path Simulation for Quadruped Walking Robot <i>Dmitry Dobrynin and Yulia Zhiteneva (Russia)</i>
13:50-14:10	Spectral Root Features for Replay Spoof Detection in Voice Assistants <i>Ankur T. Patil, Harsh Kotta, Rajul Acharya and Hemant A. Patil (India)</i>	Prioritized SIPP Algorithm for Multi-Agent Path Finding Problem with Kinematic Constraints <i>Zain Alabedeen Ali and Konstantin Yakovlev (Russia)</i>
14:10-14:30	Predicting Biometric Error Behaviour from Speaker Embeddings and a Fast Score Normalization Scheme <i>İsmail Rasim Ülgen, Mustafa Erden and Levent M. Arslan (Turkey)</i>	Analysis of Kinematic Diagrams and Design Solutions of Mobile Agricultural Robots <i>Iuliia Vasiunina, Konstantin Krestovnikov, Aleksandr Bykov and Aleksei Erashov (Russia)</i>
14:30-14:50	Evaluating X-vector-based Speaker Anonymization Under White-box Assessment <i>Pierre Champion, Denis Jouvet and Anthony Larcher (France)</i>	DronePort: Smart Drone Battery Management System <i>Zdeněk Bouček, Petr Neduchal and Miroslav Flidr (Czechia)</i>
14:50-15:10	Applying EEND Diarization to Telephone Recordings from a Call Center <i>Zbynek Zajic, Marie Kunesova and Ludek Muller (Czechia)</i>	Classification of Aerial Manipulation Systems and Algorithms for Controlling the Movement of Onboard Manipulator <i>Vinh Nguyen, Tien Ngo, Quyen Vu and Andrey Ronzhin (Russia, Vietnam)</i>
15:10-15:30	Speaker Authorization for Air Traffic Control Security <i>Marian Trnka, Sakhia Darjaa, Milan Rusko, Meilin Schaper and Tim H. Stelkens-Kobsch (Slovakia, Germany)</i>	Adaptive Event Triggered Control of Nonholonomic Mobile Robots <i>Hacı Mehmet Güzey (Turkey)</i>
15:30-15:50	Short break	

15:50-17:30	SPECOM Session 2 (10 short presentations, 10 min) Speech and Audio Signal Processing (Room 1) Chair: Ilya Chizhov	ICR Session 2 (5 full presentations, 20 min) Robot Control Systems (Room 2) Chair: Dmitriy Levonevskiy
15:50-16:00	Assessing Speaker-Independent Character Information for Acted Voices <i>Mathias Quillot, Richard Dufour and Jean-Francois Bonastre (France)</i>	Collision Avoidance for Mobile Robots using Proximity Sensors <i>Yucheng Tang, Ilshat Mamaev, Hosam Alagi, Bengt Abel and Bjorn Hein (Germany)</i>
16:00-16:10	Language Adaptation for Speaker Recognition Systems using Contrastive Learning <i>Vincent Brignatz, Jarod Duret, Driss Matrouf and Mickael Rouvier (France)</i>	
16:10-16:20	Modified Group Delay Function using Different Spectral Smoothing Techniques for Voice Liveness Detection <i>Shrishti Singh, Kuldeep Khorja and Hemant A. Patil (India) – VIDEO</i>	Visual Data Processing Framework for a Skin-Based Human Detection <i>Valery Myrzin, Tatyana Tsoy, Yang Bai, Mikhail Svinin and Evgeni Magid (Russia, Japan)</i>
16:20-16:30	Kurdish Spoken Dialect Recognition Using X-vector Speaker Embedding <i>Arash Amani, Mohammad Mohammadamini and Hadi Veisi (Iran, France)</i>	
16:30-16:40	Improvement of Speaker Number Estimation by Applying an Overlapped Speech Detector <i>Elena Timofeeva, Elena Evseeva, Valeriia Zaluskaia, Vlada Kapranova, Sergei Astapov and Vladimir Kabarov (Russia)</i>	Prioritizing Tasks within a Robotic Transportation System for a Smart Hospital Environment <i>Ruslan Safin, Roman Lavrenov, Tatyana Tsoy, Evgeni Magid, Mikhail Svinin, Sumantra Dutta Roy and Subir Kumar Saha (Russia, Japan, India)</i>
16:40-16:50	A Mobile Application for Detection of Amyotrophic Lateral Sclerosis via Voice Analysis <i>Denis Likhachov, Maxim Vashkevich, Elias Azarov, Katsiaryna Malhina and Yuliya Rushkevich (Belarus)</i>	
16:50-17:00	Prosodic Changes with Age: a Longitudinal Study on a Famous European Portuguese Native Speaker <i>Ana Rita Valente, Catarina Oliveira, Luciana Albuquerque, António Teixeira and Plínio A. Barbosa (Portugal, Brazil)</i>	Development of Matrix of Combined Force and Proximity Sensors for use in Robotics <i>Konstantin Krestovnikov, Aleksei Erashov and Aleksandr Bykov (Russia)</i>
17:00-17:10	Assessing Velar Gestures Timing in European Portuguese Nasal Vowels with RT-MRI Data <i>Nuno Almeida, Conceição Cunha, Samuel Silva and António Teixeira (Portugal)</i>	
17:10-17:20	Text-Independent Speaker Verification Employing CNN-LSTM-TDNN Hybrid Networks <i>Jahangir Alam, Abderrahim Fathan and Woohyun Kang (Canada)</i>	A Framework to Process Natural Speech in Service Robots using a Combination of a Speech Recognition System, Universal Dependencies Extracted by means of the Stanford Parser and an Expert System <i>Alessandro Presacco, Jesus Savage, Julio Cruz and Stephany Ortuño (Mexico)</i>
17:20-17:30	End-to-End Voice Spoofing Detection Employing Time Delay Neural Networks and Higher Order Statistics <i>Jahangir Alam, Abderrahim Fathan and Woohyun Kang (Canada)</i>	
17:30-17:40	Short break	
17:40-18:40	Keynote lecture by Prof. MAC SCHWAGER (USA): Enabling Robots to Cooperate and Compete: “Distributed Optimization and Game Theoretic Methods for Multiple Interacting Robots” Chair: Anton Saveliev (Room 1)	

Day 2: 29 September 2021, Wednesday (Rooms 1 & 2)	
09:50-10:00	Zoom connection
10:00-12:00	<p>SPECOM Session 3 (6 full presentations, 20 min) Automatic Speech Recognition (Room 1) Chair: Gerhard Rigoll</p> <p>ICR Session 3 (8 short presentations, 15 min) Robotic Systems (Room 2) Chair: Roman Lavrenov</p>
10:00-10:20	<p>An Equal Data Setting for Attention-based Encoder-Decoder and HMM/DNN Models: a Case Study in Finnish ASR <i>Aku Rouhe, Astrid Van Camp, Mittul Singh, Hugo Van Hamme and Mikko Kurimo (Finland, Belgium)</i></p> <p>A Technique to Provide an Efficient System Recovery in the Fog- and Edge-environments of Robotic Systems <i>Anna Klimenko and Igor Kalyaev (Russia)</i></p>
10:20-10:40	<p>Induced Local Attention for Transformer Models in Speech Recognition <i>Tobias Watzel, Ludwig Kürzinger, Lujun Li and Gerhard Rigoll (Germany)</i></p> <p>Dedicated Payload Stabilization System in a Service Robot for Catering <i>Petr Smirnov and Artem Kovalev (Russia)</i></p>
10:40-11:00	<p>Text and Synthetic Data for Domain Adaptation in End-to-End Speech Recognition <i>Juan Hussain, Christian Huber, Sebastian Stüker and Alexander Waibel (Germany, USA)</i></p> <p>The Efficiency Improvement of Robots Group Operation by Means of Workload Relocation <i>Eduard Melnik, Irina Safronenkova and Sergey Kapustyan (Russia)</i></p>
11:00-11:20	<p>Study on Acoustic Model Personalization in a Context of Collaborative Learning Constrained by Privacy Preservation <i>Salima Mdhaffar, Marc Tommasi and Yannick Estève (France)</i></p> <p>Algorithm for Radio Survey of the Cyber-Physical Systems Operating Areas Using Unmanned Aerial Vehicles <i>Alexander Denisov and Dmitriy Levonevskiy (Russia)</i></p>
11:20-11:40	<p>Recognition of Heavily Accented and Emotional Speech of English and Czech Holocaust Survivors using Various DNN Architectures <i>Josef V. Psutka, Aleš Pražák and Jan Vaněk (Czechia)</i></p> <p>The Problem Statement of Cognitive Modeling in Social Robotic Systems <i>Galina Gorelova, Eduard Melnik and Irina Safronenkova (Russia)</i></p>
11:40-12:00	<p>Optimised Code-Switched Language Model Data Augmentation in Four Under-Resourced South African Languages <i>Joshua Jansen van Vueren and Thomas Niesler (South Africa)</i></p> <p>Semantic Segmentation in the Task of Long-Term Visual Localization <i>Lukáš Bureš and Luděk Müller (Czechia)</i></p>
12:00-12:20	<p>Development of a Multi-sensor Emotional Response System for Social Robots <i>Estilla Hefter, Charlotte Perry, Nicholas Coiro, Harry Parsons, Shuo Zhu and Chunxu Li (UK, China) – VIDEO</i></p> <p>"MEOW" - A Social Robot Platform for School <i>Egor Polyntsev, Evgeny Shandarov and Vladimir Zhelonkin (Russia)</i></p>
12:00-12:20	Short break

12:20-13:50	SPECOM Session 4 (9 short presentations, 10 min) Speech Recognition and Understanding (Room 1) Chair: Yannick Esteve
12:20-12:30	Regularized Forward-Backward Decoder for Attention Models <i>Tobias Watzel, Ludwig Kürzinger, Lujun Li and Gerhard Rigoll (Germany)</i>
12:30-12:40	Methods for Using Class Based N-gram Language Models in the Kaldi Toolkit <i>Edvin Pakoci and Branislav Popović (Serbia)</i>
12:40-12:50	Multilingual Training Set Selection for ASR in Under-Resourced Malian Languages <i>Ewald van der Westhuizen, Trideba Padhi and Thomas Niesler (South Africa)</i>
12:50-13:00	A Study of Multilingual End-to-End Speech Recognition for Kazakh, Russian, and English <i>Saida Mussakhoyeva, Yerbolat Khassanov and Huseyin Atakan Varol (Kazakhstan)</i>
13:00-13:10	USC: An Open-Source Uzbek Speech Corpus and Initial Speech Recognition Experiments <i>Muhammadjon Musaev, Saida Mussakhoyeva, Ilyos Khujayorov, Yerbolat Khassanov, Mannon Ochilov and Huseyin Atakan Varol (Uzbekistan, Kazakhstan)</i>
13:10-13:20	Where are We in Semantic Concept Extraction for Spoken Language Understanding? <i>Sahar Ghannay, Antoine Caubrière, Salima Mdhaffar, Gaëlle Laperrière, Bassam Jabaian and Yannick Estève (France)</i>
13:20-13:30	Initial Experiments on Question Answering from the Intrinsic Structure of Oral History Archives <i>Adam Chýlek, Jan Švec and Luboš Šmíd (Czechia)</i>
13:30-13:40	Speaker-invariant Speech-To-Intent Classification for Low-Resource Languages <i>Anosha Ignatius and Uthayasanker Thayasivam (Sri Lanka)</i>
13:40-13:50	Experimental Analysis of Expert and Quantitative Estimates of Syllable Recordings in the Process of Speech Rehabilitation <i>Dariya Novokhrestova, Evgeny Kostyuchenko, Ilya Hodashinsky and Lidiya Balatskaya (Russia)</i>
13:50-15:00	Lunch break & Online entertainment (Room 1)
15:00-17:00	SPECOM Session 5 (6 full presentations, 20 min) Speech Synthesis (Room 1) Chair: Oliver Jokisch
15:00-15:20	Word-level Style Control for Expressive, Non-attentive Speech Synthesis <i>Konstantinos Klapsas, Nikolaos Ellinas, June Sig Sung, Hyounghmin Park and Spyros Raptis (Greece, South Korea)</i>
15:20-15:40	Improved Prosodic Clustering for Multispeaker and Speaker-Independent Phoneme-Level Prosody Control - <i>Myrsini Christidou, Alexandra Vioni, Nikolaos Ellinas, Georgios Vamvoukakis, Konstantinos Markopoulos, Panos Kakoulidis, June Sig Sung, Hyounghmin Park, Aimilios Chalamandaris and Pirros Tsiakoulis (Greece, South Korea)</i>
15:40-16:00	Assessing Speaker Interpolation in Neural Text-to-Speech <i>Roman Korostik, Javier Latorre, Sivanand Achanta and Yannis Stylianou (Russia, UK, USA)</i>
16:00-16:20	Speaker Adaptation with Continuous Vocoder-based DNN-TTS <i>Ali Raheem Mandeel, Mohammed Salah Al-Radhi and Tamás Gábor Csapó (Hungary)</i>
16:20-16:40	Explicit Control of the Level of Expressiveness in DNN-based Speech Synthesis by Embedding Interpolation <i>Tijana Nosek, Siniša Suzić, Mia Vujović, Darko Pekar, Milan Sečujski and Vlado Delic (Serbia)</i>
16:40-17:00	Synthesis Speech based Data Augmentation for Low Resource Children ASR <i>Virender Kadyan, Hemant Kathania, Prajval Govil and Mikko Kurimo (India, Finland)</i>
17:00-17:20	Short break

17:20-19:00	SPECOM Session 6 (5 full presentations, 20 min) Computational Paralinguistics (Room 1) Chair: Heysem Kaya	SPECOM Session 7 (5 full presentations, 20 min) Multimodal Analysis and Synthesis (Room 2) Chair: Wolfgang Minker
17:20-17:40	Deep Learning based Engagement Recognition in Highly Imbalanced Data <i>Denis Dresvyanskiy, Wolfgang Minker and Alexey Karpov (Germany, Russia)</i>	Designing and Deploying an Interaction Modality for Articulatory-Based Audiovisual Speech Synthesis <i>Nuno Almeida, Diogo Cunha, Samuel Silva and António Teixeira (Portugal)</i>
17:40-18:00	Dialog Speech Sentiment Classification for Imbalanced Datasets <i>Sergis Nicolaou, Lambros Mavrides, Georgina Tryfou, Kyriakos Tolia, Konstantinos Panousis, Sotirios Chatzis and Sergios Theodoridis (Cyprus, Denmark)</i>	Multimodal Corpus Analysis of Autoblog 2020: Lecture Videos in Machine Learning <i>Abner Hernandez and Seung Hee Yang (Germany)</i>
18:00-18:20	Automatic Selection of the Most Characterizing Features for Detecting COPD in Speech <i>Loes van Bommel, Wieke Harmsen, Catia Cucchiari and Helmer Strik (The Netherlands)</i>	X-Bridge: Image-to-Image Translation with Reconstruction Capabilities <i>Ivan Gruber, Marek Hruz, Miloš Železný and Alexey Karpov (Czechia, Russia)</i>
18:20-18:40	Increasing the Precision of Dysarthric Speech Intelligibility and Severity Level Estimate <i>Mohammad Soleymanpour, Michael T. Johnson and Jeffrey Berry (USA)</i>	OCR Improvements for Images of Multi-Page Historical Documents <i>Ivan Gruber, Marek Hruz, Pavel Ircing, Petr Neduchal, Tomáš Zítka, Miroslav Hlaváč, Zbyněk Zajíc, Jan Švec and Martin Bulín (Czechia)</i>
18:40-19:00	An Ensemble Approach for the Diagnosis of COVID-19 from Speech and Cough Sounds <i>Abderrahim Fathan, Jahangir Alam and Woohyun Kang (Canada)</i>	Imagined, Intended, and Spoken Speech Envelope Synthesis from Neuromagnetic Signals <i>Debadatta Dash, Paul Ferrari, Karinne Berstis and Jun Wang (USA)</i>

Day 3: 30 September 2021, Thursday (Room 2)	
09:50-10:00	Zoom connection
10:00-12:00	SPECOM Session 8 (6 full presentations, 20 min) Spoken Language Processing (Room 2) Chair: Milos Zelezny
10:00-10:20	Learning Mizo Tones from F0 Contours using 1D-CNN <i>Parismita Gogoi, Sishir Kalita, Wendy Lalhminghlui, Priyankoo Sarmah and S. R. M. Prasanna (India)</i>
10:20-10:40	Articulation During Voice Disguise: a Pilot Study <i>Lauri Tavi, Tomi Kinnunen, Einar Meister, Rosa González Hautamäki and Anton Malmi (Finland, Estonia, Singapore)</i>
10:40-11:00	Velocity Differences between Velum Raising and Lowering Movements <i>Peter Birkholz and Christian Kleiner (Germany)</i>
11:00-11:20	Is it a Filler or a Pause? A Quantitative Analysis of Filled Pauses in Hebrew <i>Vered Silber-Varod, Mária Gósy and Anat Lerner (Israel, Hungary)</i>
11:20-11:40	Human and Transformer-Based Prosodic Phrasing in Two Speech Genres <i>Jan Volín, Markéta Řezáčková and Jindřich Matoušek (Czechia)</i>
11:40-12:00	An ASR-based Tutor for Learning to Read: How to Optimize Feedback to First Graders <i>Yu Bai, Cristian Tejedor-García, Ferdy Hubers, Catia Cucchiari and Helmer Strik (The Netherlands)</i>
12:00-12:20	Short break
12:20-13:50	SPECOM Session 9 (9 short presentations, 10 min) Spoken Language Analysis (Room 2) Chair: Elena Lyakso
12:20-12:30	Child's Emotional Speech Classification by Human across Two Languages: Russian & Tamil <i>Elena Lyakso, Olga Frolova, Nersisson Ruban and A. Mary Mekala (Russia, India)</i>
12:30-12:40	Automatic Recognition of the Psychoneurological State of Children: Autism Spectrum Disorders, Down Syndrome, Typical Development <i>Yuri Matveev, Anton Matveev, Olga Frolova and Elena Lyakso (Russia)</i>
12:40-12:50	Complex Rhythm Adjustments in Multilingual Code-Switching across Mandarin, English and Russian <i>Tatiana Sokoreva, Tatiana Shevchenko and Mariya Chyrvonaya (Russia)</i>
12:50-13:00	Intraspeaker Variability of a Professional Lecturer: Ageing, Genre, Pragmatics vs. Voice Acting (Case Study) <i>Anna Dunashova (Russia)</i>
13:00-13:10	Pragmatic Markers of Russian Everyday Speech: Invariants in Dialogue and Monologue <i>Natalia Bogdanova-Beglarian, Olga Blinova, Tatiana Sherstinova and Tatiana Sulimova (Russia)</i>
13:10-13:20	Can Your Eyes Tell Us Why You Hesitate? Comparing Reading Aloud in Russian as L1 and Japanese as L2 <i>Valeriya Prokaeva, Elena Riekhakaynen and Vladislav Zubov (Russia)</i>
13:20-13:30	What Causes Phonetic Reduction in Russian Speech: New Evidence from Machine Learning Algorithms <i>Maria Dayter and Elena Riekhakaynen (Russia)</i>
13:30-13:40	Media Content vs Nature Stimuli Influence on Human Brain Activity <i>Rodmonga Potapova, Vsevolod Potapov, Natalia Lebedeva, Ekaterina Karimova and Nikolay Bobrov (Russia)</i>
13:40-13:50	Influence of the Aggressive Internet Environment on Cognitive Personality Disorders (in Relation to the Russian Young Generation of Users) <i>Rodmonga Potapova, Tatiana Agibalova, Vsevolod Potapov and Olga Tuchina (Russia)</i>
13:50-15:00	Lunch break

15:00-17:00	SPECOM Session 10 (6 full presentations, 20 min) Natural Language Processing (Room 2) Chair: Daniil Kocharov
15:00-15:20	Who is Selling to Whom - Feature Evaluation for Multi-block Classification in Invoice Information Extraction <i>Hien Thi Ha and Ales Horak (Czechia)</i>
15:20-15:40	Opinion Classification via Word and Emoji Embedding Models with LSTM <i>Ilyos Rabbimov, Sami Kobilov and Iosif Mporas (Uzbekistan, UK)</i>
15:40-16:00	Mind Your Tweet: Abusive Tweet Detection <i>Paras Tiwari and Sawan Rai (India)</i>
16:00-16:20	Perceiving Speech Aggression with and without Textual Context on Twitter Social Network Site <i>Liliya Komalova and Diana Kulagina (Russia)</i>
16:20-16:40	Toxic Comment Classification Service in Social Network <i>Mikhail Dolgushin, Dayana Ismakova, Yuliya Bidulya, Igor Krupkin, Galina Barskaya and Anastasiya Lesiv (Russia)</i>
16:40-17:00	Neural Network Recognition of Russian Noun and Adjective Cases in the Google Books Ngram Corpus <i>Andrey Savinkov, Vladimir Bochkarev, Anna Shevlyakova and Stanislav Khristoforov (Russia)</i>
17:00-17:20	Short break
17:20-18:20	SPECOM Session 11 (6 short presentations, 10 min) Speech and Audio Signal Processing (Room 2) Chair: Irina Kipyatkova
17:20-17:30	Speaker-aware Training of Speech Emotion Classifier with Speaker Recognition <i>Lyudmila Savchenko and Andrey Savchenko (Russia)</i>
17:30-17:40	Analysis of Dialogues of Typically Developing Children, Children with Down Syndrome and ASD using Machine Learning Methods <i>Olesia Makhnytkina, Aleksey Grigorev and Aleksander Nikolaev (Russia)</i>
17:40-17:50	Acoustic Characteristics of Speech Entrainment in Dialogues in Similar Phonetic Sequences <i>Svetlana Zimina and Vera Evdokimova (Russia)</i>
17:50-18:00	Learning Efficient Representations for Keyword Spotting with Triplet Loss <i>Roman Vygon and Nikolay Mikhaylovskiy (Russia)</i>
18:00-18:10	End-to-End Russian Speech Recognition Models with Multi-Head Attention <i>Irina Kipyatkova (Russia)</i>
18:10-18:20	Speaker-Dependent Visual Command Recognition in Vehicle Cabin: Methodology and Evaluation <i>Denis Ivanko, Dmitry Ryumin, Alexandr Axyonov and Alexey Kashevnik (Russia)</i>
18:20-18:30	Closing ceremony (Room 2)

Zoom online connection

Room 1: SPECOM Sessions (Days 1 & 2) and Keynotes

<https://us02web.zoom.us/j/86525238435?pwd=eFQvM29kV3RyQkcrUTNHUXpVdGJBUT09>

ID: 865 2523 8435

Password: 611009

Room 2: ICR Sessions and SPECOM Session 7 & Day 3

<https://us06web.zoom.us/j/82776250333?pwd=SWxydzU3b2IPODMxTnBVeWsrNEsrUT09>

ID: 827 7625 0333

Password: 118713

Proceedings

Free access to the electronic proceedings is available till end-October 2021 via the conference web page: <http://www.specom.nw.ru/2021/program/>

- 23rd International Conference SPECOM 2021 (**Springer LNCS/LNAI, vol. 12997**)
- 6th International Conference ICR 2021 (**Springer LNCS/LNAI, vol. 12998**)