ICON 2021

The 18th International Conference on Natural Language Processing

Proceedings of the Workshop on Speech and Music Processing 2021 (SMP2021)

December 16, 2021

Preface

Speech and music are two prominent research areas in the domain of audio signal processing. With recent advancements in speech and music technology, the area has grown manifolds, bringing together the interdisciplinary researchers of computer science, musicology and speech analysis. The languages we speak propagate as sound wave through various media and allow communication or entertainment for us, humans. The music we hear or create can be perceived in different aspects as rhythm, melody, harmony, timbre, or mood. The multifaceted nature of speech or music information requires algorithms, systems using sophisticated signal processing, and machine learning techniques to better extract useful information. This workshop will provide both profound technological knowledge and a comprehensive treatment of essential topics in speech and music processing.

Recent computational advancement has opened up several avenues to explore further the domain of speech and music. A deep understanding of both speech and music in terms of perception, emotion, mood, gesture and cognition is in the forefront, and many researchers are working in these domains. In this digital age, overwhelming data have been generated across the world that requires efficient processing for better maintenance, retrieval, indexing and querying. Machine learning and artificial intelligence are most suited for these computational tasks.

The SMP-2021 workshop was organized with the following objects: (i) to bring researchers and developers together who work on speech and music domain. (ii) to provide a platform for researchers to discuss speech prosody, Indian as well as western music and (iii) to encourage researchers to collaborate and create more annotated resources.

Technical Session:

The SMP-2021 workshop received nine submissions by authors from India, China, Canada, and Ireland. Each paper was reviewed by 2-3 experts. Based on reviewers' comments, six papers were accepted for presentation at the workshop. However, five papers were presented during the workshop session.

The accepted and presented papers include a variety of topics from both speech and music processing domains. Two papers covered speech emotion recognition in multimodal context and speech prosody in Hindi language. While three papers covered music, out of which two papers presented language, artist and melody identification from Indian classical music and one paper discussed about Dorabella Cipher as western music context.

Acknowledgement:

Organizers would like to thank everyone who supported us to organize the Workshop on Speech and Music Processing 2021 (SMP2021). Specifically, the ICON-2021organizers and the technical program committee members need for the workshop facilitation and support in reviewing papers, respectively.

SMP-2021 Organizers:

- Dr. Anupam Biswas, Department of CSE, National Institute of Technology Silchar, India
- Dr. Rabul H. Laskar, Department of ECE, National Institute of Technology Silchar, India
- Dr. Pinki Roy, Department of CSE, National Institute of Technology Silchar, India

Biography of Organizers:

Anupam Biswas received his Ph.D. degree in computer science and engineering from Indian Institute of Technology (BHU), Varanasi, India in 2017. He has received his M. Tech. and B. E. Degree in computer science and engineering from Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India in 2013 and Jorhat Engineering College, Jorhat, Assam in 2011 respectively. He is currently working as an Assistant Professor in the Department of Computer Science & Engineering, National Institute of Technology Silchar, Assam. He has published several research papers in reputed international journals, conference and book chapters. His research interests include Machine learning, Social Networks, Computational music, Information retrieval, and Evolutionary computation. He is Principal Investigator of two on-going DST-SERB sponsored research projects in the domain of machine learning and evolutionary computation. He has served as Program Chair of International Conference on Big Data, Machine Learning and Applications (BigDML 2019) and currently serving as Publicity Chair of BigDML 2021. He has served as General Chair of 25th International Symposium Frontiers of Research in Speech and Music (FRSM 2020) and co-edited the proceedings of FRSM 2020 published as book volume in Springer AISC Series. He edited two books titled "Health Informatics: A Computational Perspective in Healthcare" and "Principles of Social Networking: The New Horizon and Emerging Challenges" published by Springer. Also edited a book "Principles of Big Graph: In-depth Insight" with Advances in Computers book Series of Elsevier, currently in press.

Rabul Hussain Laskar is an Associate Professor in the Department of Electronics and Communication Engineering at National Institute of Technology Silchar. He completed his Bachelor of Engineering (B.E.) in Electronics and Communication Engineering in 1998, Masters in Signal processing (M.Tech) in 2009, and Doctorate in speech signal processing in 2013. He has been involved in teaching and research for the past 21 years. He is leading the signal processing research group in ECE department. His area of specialization includes digital signal processing, machine learning, computer vision, pattern recognition, analog and digital communication. He has been teaching subjects like signals and systems, linear algebra and random process, communication system engineering, digital speech processing, digital image processing, etc., among the student and research community. He has published 80+ research articles in various referred journals, 70+ papers in international/national conferences, 10+ chapters in different book series. He has been an active reviewer in various international journals and conferences. He has supervised 12 Ph.D. scholars, 10 post-graduate, 40 undergraduates in different areas of speech, image, video and biomedical signal processing. There are 08 Ph.D. students working under his supervision to complete their thesis. He has been involved in various R&D projects sponsored by DIT, SERB, MCIT, DST, BARC- BRNS as principal investigator. He is the senior member of IEEE, IETE fellow, IEI fellow.

Pinki Roy is working in NIT-Silchar for past 17 years. Her research interests include machine intelligence, language identification, Speech Processing, Health informatics and Cloud Computing. She has published more than 24 SCI/SCIE/Scopus Journals with highest impact factor of more than 7. She has more than 30 publications in various IEEE conferences/other reputed conferences and in Book Chapters. She had travelled to international destinations like Singapore, Sydney (Australia) and London to attend world top level conferences where her research work was highly appreciated. Already 3 PhD scholars have attained PhD degrees under her guidance in the area of cloud computing and speech processing. At present 4 scholars are working under her guidance. She was the recipient of Distinguished Lady Alumnus Award",

from Dr. Babasaheb Ambedkar technological university, Lonere, Maharashtra, 2014,"Young Scientist Award", Venus International foundation, Chennai 2015. Awarded for major contribution in research during PhD, "Rastriya Gaurav Award", India International Friendship Society, New Delhi 2015. She was also recipient of "Bharat Excellence Award", "Best Golden personalities Award", "Global Award for Education" by the Friendship Forum, New Delhi in the year 2016.She has organized workshops on Machine learning where people from the Academics and Industry were invited to bridge the gap between both the organizations. She has chaired several sessions in many renowned International Conferences and delivered expert talks in various workshops from time to time. She has served as TPC member and had reviewed many papers in International conferences and Journals.

Program Committee:

- Dr. Alicja Wieczorkowska, Polish-Japanese Academy of Information Technology, Poland
- Dr. Archi Banerjee, IIT Kharagpur, India
- Dr. Kaustav Kanti Ganguly, New York University, Abu Dhabi
- Prof. Debasish Samanta, IIT Kharagpur, India
- Prof. Kaushik Roy, West Bengal State University Barasat, India
- Dr. Suyel Namsudra, NIT Patna, India
- Dr. Vijay Bhaskar Semwal, MANIT Bhopal, India
- Dr. Mridula Verma, IDRBT Hyderabad, India
- Dr. Prabu Mohandas, NIT Trichy, India
- Dr. Durgesh Singh, IIITDM, Jabalpur, India
- Dr. Thoudam Doren Singh, NIT Silchar, India
- Dr. Rahul Chandra Kushwaha, RGU, Arunachal Pradesh, India
- Dr. Pravin Kumar, Allahabad University, Prayagraj, India

Table of Contents

Conference Program

16 December 2021

Classifying Emotional Utterances by Employing Multi-modal Speech Emotion Recognition
Dipankar Das

Prosody Labelled Dataset for Hindi
Esha Banerjee, Atul Kr. Ojha and Girish Jha

Multitask Learning based Deep Learning Model for Music Artist and Language Recognition
Yeshwant Singh and Anupam Biswas

Comparative Analysis of Melodia and Time-Domain Adaptive Filtering based Model for Melody Extraction from Polyphonic Music
Ranjeet Kumar, Anupam Biswas, Pinki Roy and Yeshwant Singh

Dorabella Cipher as Musical Inspiration
Bradley Hauer, Colin Choi, Abram Hindle, Scott Smallwood and Grzegorz Kondrak