



Preface: Special Section: Special Issue on Next Generation Acoustic Signal Processing (Articles 1–9)

Amy Neustein¹ · Parikshit N. Mahalle² · Mohd. Shafi Pathan³ · Nilanjan Dey⁴

Published online: 9 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Acoustics is the branch of physics that deals with the study of all mechanical waves in gases, liquids, and solids including topics such as vibration, sound, ultrasound and infrasound. The acoustic interface for future multimedia and communication terminals should be hands-free and as natural as possible. Signal processing algorithms enable acousticians to separate signals from noise, to perform automatic speech recognition or to compress information for more efficient storage or transmission. Artificial Intelligence [AI] is emerging as a most prevailing technology in the speech recognition process. To develop robust Next Generation Acoustic Signal.

This special issue focuses on the recent research works that introduces new methods and techniques to implement Acoustic Signal Processing for Next-Generation Human/Machine Interfaces. This special issue includes 09 articles.

In the first article, Sumita Nainan et al. performed comparative analysis of accuracies obtained in ASR with employment of classical Gaussian mixture model (GMM), support vector machine (SVM) which is the machine learning algorithm and the state of art 1-D CNN as classifiers is presented. Authors propose considering dynamic voice

features along with static features as relevant speaker information in them lead to substantial improvement in the accuracy for ASR. As concatenation of features leads to the redundancy and increased computation complexity, Fisher score algorithm was employed to select the best contributing features resulting in improvement in accuracy.

In the second article, Lakshmi et al. presents an extensive survey of reversible watermarking techniques based on DE algorithms for their characteristics, considerations, performance with respect to complexity, embedding capacity, perceptual impact, and fragility, to find the applicability of those algorithms. The paper also outlines the open research problems and provides some future directions for the researchers working in the related domain.

In the third article, Yosra et al. described and evaluated a work on Part-of-speech (POS) tagging for the Khasi language by using the Conditional Random Field (CRF) method. The main significance of this work, is to experiment with the CRF model for PoS tagging in the Khasi language. This method produces a reliable agreement on the features of the language. POS tagging for Khasi is essential for creating lemmatizers which are used to lessen a word to its root structure and the POS corpus or dataset can be used in other NLP applications.

In the fourth article, Hsein Kew presented a method to generate an audio output based on spectroscopy data in order to discriminate two classes of data, based on the features of our spectral dataset. To do this, we first perform spectral pre-processing, and then extract features, followed by machine learning, for dimensionality reduction. The features are then mapped to the parameters of a sound synthesiser, as part of the audio processing, so as to generate audio samples in order to compute statistical results and identify important descriptors for the classification of the dataset. The results of classification accuracy using FM synthesis as the sound synthesiser and PCA as the dimensionality reduction method yields a mean classification accuracies of 93.81%

✉ Mohd. Shafi Pathan
shafipathan@gmail.com

Parikshit N. Mahalle
aalborg.pnm@gmail.com

Nilanjan Dey
neelanjan.dey@gmail.com

¹ Linguistic Technology Systems, Inc., Fort Lee, NJ, USA

² Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune, India

³ Department of Computer Engineering, Vidyavardhini's College of Engineering & Technology, Vasai, Maharashtra, India

⁴ Department of Computer Science and Engineering, JIS University, Kolkata, India

and 88.57% for the coffee dataset and the fruit puree dataset respectively.

In the fifth article, Nabanita Das et al. provides a state-of-the-art summary and present approaches for using the widely used machine learning and deep learning methods to detect the challenges along with future research directions of speech enhancement systems. With the advancement of machine learning and deep learning, classification of speech has become more significant. Methods of speech enhancement consist of different stages, such as feature extraction of the input speech signal, feature selection, feature selection followed by classification. Deep learning techniques are also an emerging field in the classification domain, which is discussed in this review.

In the sixth article, Debajyoty Banik proposed a Method based on the syntactic phrase-level information, phrase table is updated with the weights estimated using probability distribution. Evaluation on English–Hindi technical domain corpora shows that our proposed method is more effective in producing better output in terms of BLEU, RIBES and NIST metrics. We shows that the proposed methods works well for other language pairs like Hindi–Konkani and Bengali–Hindi. Finally, we realised that this minor probabilistic change can improve the accuracy of the machine translation system a lot.

In the seventh article, Jyotismita Chaki provides the survey of the state-of art for understanding ASC's general research scope, including different types of audio; representation of audio like acoustic, spectrogram; audio feature extraction techniques like physical, perceptual, static, dynamic; audio pattern matching approaches like pattern matching, acoustic phonetic, artificial intelligence; classification, and clustering techniques. The aim of this state-of-art paper is to produce a summary and guidelines for using the broadly used methods, to identify the challenges as well as future research directions of acoustic signal processing.

In the eighth article, Zaid Albataineh addresses the problem of scalability to higher dimensional space in blind source separation (BSS), where the number of sources is greater than two. Herein, we propose two schemes of pairwise non-parametric independent component analysis (ICA) algorithms based on Convex Cauchy–Schwarz Divergence (CCS–DIV) for high dimensional problem in BSS. Author extends the pairwise method to the scenario of more than two sources; two improved ICA algorithms are developed. Moreover and we employ adaptive sampling technique that samples the signal into small time blocks to evaluate the integration of the CCS–DIV and reduce the computational

complexity. The two presented methods enable fast and efficient demixing of sources in real-world high dimensional source applications.

In the ninth article, Soumi Majumder et al. provided a summary of the challenges and research directions in Human resource. With the invention of chatbots, which is a significant domain in AI and natural language processing, the organizations have become more AI centric. A bot is considered to be an effective communication system which can be used among the employees as well as customers for performing some communication-oriented works within an organization without any human intervention. In this work, an overview of the chatbots in HRM, often called as a HR-bot, is studied to emphasise the usefulness in real time considering the relevant challenges such as cost factor, complex business domains, limited responsiveness etc.

In this issue, the guest editors selected 09 research articles and confirmed they will be effective and valuable for multitude of readers/researchers. Note that technical standard and quality of published content are based on the strength of the submitted articles. We are grateful to the authors for their imperative research contributions to this issue and their patience during the revision stages. We take this opportunity to give our special thanks to the Editor-in-chief Amy Neustein, for all the support, and competence rendered to this special issue.

Guest editors

Amy Neustein, *Editor in Chief of the International Journal of Speech Technology*.

Parikshit N. Mahalle, *Professor & Head of Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Information Technology, Pune. India aalborg.pnm@gmail.com*.

Mohd. Shafi Pathan, *Professor in Computer Engineering, Vidyavardhini's College of Engineering & Technology, Vasai, Maharashtra, India. shafipathan@gmail.com*

Nilanjan Dey, *Associate Professor in Department of Computer Science and Engineering, JIS University, Kolkata, India. neelanjan.dey@gmail.com*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.