Congratulations to Mr.SREEKUMAR G on the successful completion and awarding of PhD

Thesis title: Design, Implementation and Evaluation of High Resolution DoA Estimation Algorithms in Phased Arrays using Chirp Signals
Supervisor: Dr. Leena Mary
Date of Defense: 2 March 2020
University: Mahatma Gandhi University

Thesis Summary: The thesis reports the techniques developed for the detection and DoA estimation of sources transmitting a variety of chirp signals, using phased array processing techniques. The algorithms developed as part of the thesis have been demonstrated on the chirp signals (LFM and QFM), which include simulated as well as real data obtained from a practical sonar array. The problem of estimating the DoAs of multiple far-field moving contacts, producing broad-banded chirp signals, in both active and passive mode of operation is addressed. For the DoA estimation of sources stemming from a variable class of chirp signals, the conventional and subspace beam-forming algorithms are re-cast in the thesis, in the FrFT/GTFT domains. For the recasting, the challenge lies in the derivation of the array steering vector in the new domain. Using the newly derived steering vectors, the DoA estimation techniques are evaluated through elaborate simulation. The results of simulation experiments reported in the present work show that FrFT based approach enjoys better estimation over the conventional frequency domain beam-forming for chirp signals in terms of accuracy, resolution and computational efficiency. The pronounced performance is manifested even with low SNR, spatially coherent signals and with limited snapshots and sensors as well. The performance metrics used for the study are RMSE, 3 dB beam width and CPU time for bearing accuracy, resolution of closely located contacts and computation efficiency, respectively. It is seen from computer simulations that FrFT-MUSIC has better accuracy and resolution in general, whereas FrFT-MVDR performs better in multipath environments. Experimental verification of the QFM chirp detection employing GTFT has performed admirably well in underwater environment, when the method was tested using real data acquired from a practical sonar array. It is observed that even though the transmitted chirp is linear, the reflected response from the underwater medium contact is found non-linear due to the absorption of high frequency terms. Therefore GTFT based method stands one step better over conventional FrFT, in respect of the performance of practical underwater object contacts. Also the work reaps the advantage of FrFT/GTFT methods, extending to non-uniform sparse arrays such as nested and co-prime arrays. Such modified geometries are capable of detecting more number of sources, than the number of sensors by exploiting the difference co-array structure based on the correlation of the observations.

Congratulations to Ms.DEEKSHITHA.G on the successful completion and awarding of PhD

Thesis Title: Multistage Spoken Term Detection for Searching Speech Databases
Supervisor: Dr. Leena Mary
Date of Defence: 22 july 2021
University: APJ Abdul Kalam Technological University

Thesis Summary: The availability of high-speed, low-cost Internet enables the use of multimedia files for a variety of applications in our day-to-day activities. A considerable volume of audio archives is available on different websites. Such vast audio resources are useful only if the needed file can be retrieved accurately and efficiently. Audio search refers to the search and retrieval of a particular audio file from a large audio database. Since most of the audio archives are not well indexed or labelled, it is still a challenging task. Spoken Term Detection (STD) refers to the process of locating the occurrences of spoken queries in a large speech database. Generally, two methods have been adopted for STD: an Automatic Speech Recognition (ASR) based label sequence matching or feature-based template matching. ASR-based techniques utilize phoneme models of a language, which require a considerable amount of labelled training data in the selected language. Hence such techniques are considered as language-dependent, and it is not feasible to develop ASR for each language. The feature-based template matching techniques address this task in a language-independent manner, but they are computationally complex. This work combines the positive aspects of both the methods by introducing a multistage architecture to address the task of STD for low-resourced languages. Two different approaches have been proposed for Language-Dependent (LD) and Language-Independent (LI) STD.

Congratulations to Ms.STARLET BEN ALEX on the successful completion and awarding of PhD

Thesis Title: Multilevel Prosodic Features for Automatic Emotion and Speaker Recognition from Speech using Deep Learning Techniques
Supervisor: Dr. Leena Mary
Date of Defense: 9 November 2021
University: APJ Abdul Kalam Technological University

Thesis Summary: Speech is the primary mode of communication for human beings. Apart from the intended message, the speech signal contains several implicit characteristics, including that of the speaker and the conveyed emotion. Speech processing tries to extract the desired information from the speech signal to facilitate better human-machine interactions. Human speech signal carries many attributes specific to the underlying emotion or speaker identity, occurring at multiple levels. However, it is difficult to isolate the attributes related to a speaker or emotion. The proper choice of these attributes would improve recognition accuracy. Prosodic characteristics are important in human speech communication and lend naturalness and intelligibility to speech. Prosodic features such as duration, energy and fundamental frequency (F0) vary among speakers/emotions. This thesis is an attempt to identify, extract and characterize prosodic features at multiple levels of speech signal for emotion and speaker recognition. Another goal of this thesis is to develop an efficient classification scheme for making prediction or classification decisions. Recently, the focus of research in various speech processing areas has changed to classifiers incorporating deep learning. This thesis, therefore, also examines the applicability of various deep learning methods to model the emotional/speaker-specific information present in the speech signal.

Congratulations to Mr.THOMSON DAVIS on the successful completion and awarding of PhD

Thesis Title: Auto-reconfigurable and Hardware-efficient Filter bank Structures for Auditory Compensation in Digital Hearing Aids
Supervisor: Dr. Manju Manuel
Date of Defense: 8 April 2022
University: APJ Abdul Kalam Technological University

Thesis Summary: Speaking and listening constitute the most important means of human communication. Several factors like aging, trauma, drugs, noise etc. contribute to hearing impairment. The use of a hearing aid provides improved communication for a hearing impaired person. The major impediment in the wide usage of hearing aid is its high cost. Hence, this thesis attempts to design low-complex filter bank structures for the auditory compensation of digital hearing aid. A digital filter bank is the back bone of the auditory compensation system of a hearing aid. A 17-band fixed filter bank is developed for the audiogram matching. The coefficients of the FIR filter are fractionally and octave interpolated to generate the filter bank. Though the design and gain adjustment are simple, the usage of hardware is not optimal while compensating flat audiograms. Hence a more flexible reconfigurable filter bank design is proposed. A reconfigurable hearing aid can be used for different impairments without modifying the basic hardware. A three-level reconfigurable filter bank structure is proposed for audiogram matching with 4, 8 and 16 bands. The control switches in the structure are adjusted to vary the configuration. The variations in hearing deficiency are considered for selecting the suitable level of filter bank for matching various audiograms. A better utilization of hardware is possible here, compared to the fixed filter bank method. An important highlight of this thesis is the concept of auto-reconfigurability of hearing aids. This feature is introduced in two types of region-based filter banks, namely, four-region reconfigurable filter bank and three-region reconfigurable filter bank. Auto-reconfigurability implies the automated selection of the optimal frequency bands in the various regions. The gradient of the hearing deficiency obtained from the audiogram is utilized for the optimum band allocation in various regions of the hearing profile. The region-based reconfigurable filter bank permits the optimal usage of hardware. In the proposed four-region reconfigurable filter bank, the overall hearing spectrum is divided into four regions. Here, three levels of filter bank design are employed with 1, 2 and 4 bands in each region. The computational complexity of the approach is less than that of the existing methods. Only 26 coefficient multipliers are required for the design of the filter bank structure. Another region-based approach proposed in this thesis is the three-region filter bank structure in which the hearing spectrum is equally divided into three regions. Besides, four levels of filter bank design are utilized for sub-band decomposition in each region. The main highlight of this method is the further reduction in hardware components compared to the four-region approach. The three-region reconfigurable filter bank structure demands only 18 multipliers for auditory compensation. In order to evaluate the effectiveness of the auto-reconfigurable filter bank structures, hardware realization is performed using Xilinx Kintex-7 FPGA. The claims about the hardware-efficient design are confirmed by the device and power utilization results of the four-region and three-region approaches. The realization of a low-complex filter bank design leads to the development of a reconfigurable cost-effective hearing aid for wide use.