K Venkata Vijay Girish

PhD (Completed), MILE Lab, Department of Electrical Engineering,

Indian Institute of Science, Bangalore- 560012, INDIA

Email: vijay.girish@gmail.com **Phone:** +91 9480515318

Web: https://sites.google.com/site/kvvijaygirish/

Date of Birth

23-10-1985

Research Interests

I am interested in designing innovative mathematical and machine learning models to analyze, process and extract useful information from various signals and data. I am inclined towards mathematical, analytical and out of the box thinking to solve any problem. My broad areas of interest are:

- Machine Listening: Audio Classification and Analysis pertaining to Speaker, Background Noise and Language Classification, Source Separation, Audio Signal Analysis, Audio Segmentation, Speech Enhancement and Multilingual Speech Recognition
- Machine Learning, Artificial Intelligence, Deep learning, Data Analytics and Signal Processing
- Sparse Signal Processing, Dictionary Learning and Adaptation, Image and Video Processing

Education

PhD, Systems and Signal Processing, August 2010- April 2017, September 2017 (Thesis submitted, Thesis Defense)
Department of Electrical Engineering, Indian Institute of Science, Bangalore, GPA: 6.0/8.0

- *Research Advisors: Prof. A. G. Ramakrishnan, Department of Electrical Engineering, Indian Institute of Science and Dr. T. V. Ananthapadmanabha, Voice and Speech Systems, Bangalore*

B.Tech., Electrical and Electronics Engineering, August 2004- May 2008
National Institute of Technology Karnataka, Surathkal, GPA: 7.75/10.00

Class- 10+2, CBSE, May 2003
Kendriya Vidyalaya No.-2, Kharagpur, Percentage: 82.0

Class- 10, CBSE, May 2001
Kendriya Vidyalaya No.-2, Kharagpur, Percentage: 75.8

PhD Thesis Summary

Title: Speech and noise analysis using sparse representation and acoustic phonetics knowledge.

This thesis addresses different aspects of speech and noise analysis using two different approaches, namely (1) A supervised and adaptive sparse representation based approach for identifying the type of background noise and the speaker and separating the speech and background noise, and (2) An unsupervised acoustic-phonetics knowledge based approach for detecting transitions between broad phonetic classes and significant excitation instants called as glottal closure instants (GCIs) in a speech signal, for applications like speech segmentation, recognition and modification.

In the supervised sparse representation based approach, a dictionary learning based noise classification algorithm is proposed using a cosine similarity measure for learning atoms of the dictionary. We have used the Active Set Newton Algorithm (ASNA) and supervised non-negative matrix factorization for source recovery in the testing phase. Based on the objective measure of signal to distortion ratio (SDR), we get frame-wise noise classification accuracy of 97.8% for fifteen different noise sources taken from NOISEX database. For speaker classification on clean speech, using high energy subsets of test frames and dictionary atoms, sum of weights measure on concatenated dictionaries give good accuracy. We have then dealt with noisy speech signals assuming a single speaker speaking in a noisy environment. The speaker and/or the noise belonging to unknown sources is handled by adaptation. Adaptive noise dictionary gives an improvement of about 18% in speaker classification accuracy and 4 dB in SDR over an out-of-set dictionary, after enhancement of noisy speech at an SNR of 0 dB. We have also addressed the classification of speakers and subsequent separation of speakers in overlapped speech, obtaining a mean speaker classification accuracy of 84% for the speaker 1 to speaker 2 ratio (S1S2R) of 0 dB.

In the unsupervised acoustic-phonetics knowledge based approach, we detect transitions between broad phonetic classes in a speech signal which has applications such as landmark detection and segmentation. A rule-based approach using relative thresholds learnt from a small development set is devised to detect transitions of silence to non-silence, sonorant to non-sonorant and vice-versa. This approach does not require significant training data for determining the parameters of the proposed approach. When tested on the entire TIMIT database for clean speech, 93.6% of the detected transitions are within a tolerance of 20 ms from the hand labeled boundaries. The proposed method is also tested on the test set of the TIMIT database for robustness with respect to white, babble and Schroeder noise, and about 90% of the detected transitions are within a tolerance of 20 ms at a SNR of 5 dB.

We have also proposed subband analysis of linear prediction residual (LPR) to estimate the GCIs from voiced speech segments. The GCI detection performance of the proposed algorithm is quantified using the following measures: identification rate (IDR), miss rate (MR), false alarm rate (FAR), standard deviation of error (SDE) and accuracy to 0.25 ms. It is evaluated using 6 different databases and compared with 3 state-of-the-art LPR based methods. The proposed method is comparable to the best of the LPR based techniques for clean and noisy speech.

Industry experience

- *Senior Technical Leader, Huawei Technologies India,* April 2017- present
  - Analysis of universities and preparing proposals for collaboration and setting up Innovation lab
  - Analysis and review of machine learning services on various cloud platforms
  - Brainstorming and proposed patents on noise and audio analysis for cloud platform
  - AI/ML resource analysis and gave technical talks on machine learning and cloud services
  - Created proposals for getting new project on audio detection/analysis, speaker classification and voice assistant on Edge and Cloud platform
  - Submitted research papers on distributed audio classification and anomaly prediction
- *Senior Engineer, Relay and Integrated Solutions, Larsen and Toubro Limited, Mumbai,* August 2008- July 2010
  - Development of new product: Intelligent Motor Protection Relay, MCOMP
  - Indepth testing, troubleshooting and analysis of new product
  - New product documentation and management
  - Development of Data Concentrator Systems
- *Industrial Training: Inplant Practical Training at the Bharat Heavy Electricals Limited, Electronics Division, Bangalore,* May 2006-June 2006

Research Publications

Conference publications

Published:

- Sayan Ghosh, K V Vijay Girish, T.V. Sreenivas, *Relationship between Indian Languages Using Long Distance Bigram Language Models*, Proc. International Conference on Natural Language Processing (ICON 2011), Dec 16-19, 2011, Chennai, India, pp. 104-113
- Vikram Ramesh Lakkavalli, K V Vijay Girish, A G Ramakrishnan, *Sub-band Envelope Approach to Obtain Instants of Significant Excitation in Speech*, Proc. National Conference on Communications (NCC 2012), Feb 3-5, 2012, Kharagpur, India, pp. 19
- Vikram R L, K V Vijay Girish, Harshavardhan S, A G Ramakrishnan, T V Ananthapadmanabha, *Sub-band Analysis of Linear Prediction Residual for the Estimation of Glottal Closure Instants*, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014), May 4-9, 2014, Florence, Italy
- K V Vijay Girish, A G Ramakrishnan and T V Ananthapadmanabha, *Hierarchical classification of speaker and background noise and estimation of SNR using sparse representation*, Interspeech 2016, September 8-12, 2016, San Francisco
- K V Vijay Girish, Veena Vijai and A G Ramakrishnan, *Relationship between spoken Indian languages by clustering of long distance bigram features of speech*, INDICON 2016, IISc Bangalore
- K V Vijay Girish, T V Ananthapadmanabha and A G Ramakrishnan, *Cosine similarity based dictionary learning and source recovery for classification of diverse audio sources*, INDICON 2016, IISc Bangalore
- K V Vijay Girish and A G Ramakrishnan, *Enhancement of noisy Tamil speech for improved quality of perception for the hearing impaired*, 16th Tamil Internet Conference 2017, Toronto
- K V Vijay Girish, A G Ramakrishnan and Neeraj Kumar, *A system for distributed audio classification using sparse representation over cloud for IOT*, accepted in COMSNETS 2018, Bangalore

Journal publications

Under review:

- T V Ananthapadmanabha, K V Vijay Girish and A G Ramakrishnan, *Relative occurrences and difference of extrema for detection of transitions between broad phonetic classes*, submitted to Sadhana

Technical Reports

- T V Ananthapadmanabha, K V Vijay Girish, A G Ramakrishnan, *Detection of transitions between broad phonetic classes in a speech signal*, arXiv:1411.0370 [cs.SD]
- K V Vijay Girish, A G Ramakrishnan and T V Ananthapadmanabha, *Adaptive dictionary based approach for background noise and speaker classification and subsequent source separation*, arXiv:1609.09764 [cs.SD]

Technical Skills

- Relevant Subjects:
  During B.Tech *August 2004 - May 2008:*
  Computer Programming, Introduction to Algorithms and Data Structures, Numerical Methods, Digital Signal Processing, Digital System Design, Microprocessors, Computer Organization and Architecture
  During PhD *August 2010- current:*
  Matrix Theory, Linear and Non Linear Optimization, Probability Theory, Convex Optimization, Advanced Digital Signal Processing, Pattern Recognition and Neural Networks, Machine Learning, Data Mining, Compressive Sensing and Sparse Signal Processing, Time Frequency Analysis, Speech Information Processing, Automatic Speech Recognition Algorithms, Digital Image Processing
- Programming languages known: *C, Python, Matlab, VHDL, Latex*
- Softwares used: *Maxwell, PSpice, Xilinx, Modelsim, Matlab, Simulink and Praat*

- Operating system: *Worked on Windows XP and Linux*

Project Mentorship

- Veena Vijai, from Birla Institute of Technology and Science, Pilani - K. K. Birla Goa Campus on Relationship between spoken Indian languages by clustering of long distance bigram features of speech, May-July, 2016
- B Shubashree, from SSN College of Engineering, Chennai on Unsupervised background noise change identification using dictionary learning, June- August, 2016

Project experience

- **Project Associate** DRDO project on Speaker and background change detection from August 2016 to April 2017
  - Understanding and delivering the project requirements
  - Regular meeting with DRDO scientists for understanding their requirements and presenting work progress
  - Visited DRDO CAIR Lab for testing algorithms on the data provided by DRDO

Teaching Experience

Linear and Nonlinear Optimization offered by Prof. Muthuvel Arigovindan at Indian Institute of Science, Bangalore during August-December, 2013: Responsibilities include conducting tutorial classes

Speech Information Processing offered by Prof. A G Ramakrishnan at Indian Institute of Science, Bangalore during January-May, 2014: Responsibilities include preparing assignments and projects

Matrix Theory offered by Prof. A G Ramakrishnan at Indian Institute of Science, Bangalore during August-December, 2014: Responsibilities include preparing assignments, clearing doubts, evaluation and grading of students.

Academic Achievements

- Got through AIEEE 2004 with an All India Rank of 4525 and State Rank (West Bengal) of 53
- Secured an All India Rank of 5439 in IIT-JEE 2004
- Got through GATE 2010 (Electrical Engineering) with 98.8 percentile
- Has been among top 3 students of the school consistently
- Distinctive performance in 2nd National Science Olympiad in 2000

Talks and Poster Presentations

- Assisted Prof. A. G. Ramakrishnan in conducting a tutorial on Insights into Signal Processing,Transforms and Linear Algebra at International Conference on Biomedical Engineering, 2011 held at MIT Manipal, during 8-9 December, 2011
- Gave a talk on my accepted paper in National Conference on Communications, held at IIT Kharagpur, during 3-5 February, 2012
- Presented poster on my accepted paper in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014) at Florence, Italy during 4-9 May, 2014
- Presented poster on my accepted paper in INTERSPEECH-2016 at Hyatt Regency, San Francisco, USA during 8-12 September, 2016
- Gave a talk and poster presentation on "Analysis of audio intercepts: Can we identify and locate the speaker? " at EECS Research Students Symposium - 2016, held at IISc Bangalore during 28-29 April, 2016

- Presented poster on “Adaptive and supervised sparse representation based approach for noisy speech analysis” in IISconnect: Industry Interaction Day at IISc Bangalore on 3rd October, 2016
- Gave an IEEE talk on “Throwing light on sound” in NIT Goa and BITS Goa, on 12th February, 2017

Hobbies and Extra-curricular activities

- Participated in Google Code Jam in the past five years
- Bagged 2nd prize in Foxhunt in the TechFest Engineer 2006 conducted by NITK Surathkal
- Secured 5th position in Simplicity, an International Online Matlab Programming contest conducted during Engineer 2008, a Technical Festival conducted by NITK Surathkal
- Bagged 3rd prize in Science Slam in Pravega Sci-Tech 2014 conducted by Swissnex at IISc Bangalore
- Volunteer in the organization team of IISc Electrical Sciences Divisional Symposium, 2013
- Active Member of Spicmacay, Voice club, IEEE and Management Forum at NITK Surathkal
- Active member of Mess committee at IISc Bangalore from July 2013 - May 2015
- Playing badminton, guitar, photography, running, swimming, biking, triathlon, programming, technology, reading, writing blogs

Languages Known

- English, Hindi and Telugu

References

Prof. A G Ramakrishnan
Chairman, Department of Electrical Engineering,
Indian Institute of Science Bangalore, India.
E-mail: ramkiag@ee.iisc.ernet.in

Dr. T V Ananthapadmanabha
Founder and Chief innovator,
Voice and Speech Systems, Bangalore, India
E-mail: tva.blr@gmail.com

Prof. T V Sreenivas
Professor, Department of Electrical Communication Engineering,
Indian Institute of Science Bangalore, India.
E-mail: tvsreenivas@gmail.com

December 5, 2017