K. Sreenivasa Rao

Curriculum Vitae

AFFILIATION:

Professor
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
Kharagpur – 721 302, India
Phone: 91-3222-282336
Fax: 91-3222-282206
Email: ksrao@iitkgp.ac.in
Website: https://cse.iitkgp.ac.in/~ksrao/

AGE AND DATE OF BIRTH: 52 Years; July 25, 1969

ACADEMIC DETAILS:

Ph.D. (2005) : Thesis on "*Acquisition and Incorporation of Prosodic Knowledge for Speech Systems in Indian Languages*", Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, India.

M.E (1993): Specialization in Communication Systems, Department of Electronics and Communication Engineering, PSG Tech., Coimbatore, India.

B.Tech (1990): Department of Electronics and Communication Engineering, Nagarjuna University, Guntur, India.

AREAS OF RESEARCH INTEREST:

Signal processing, Speech processing, Audio, Music and Multimedia, Machine learning and Pattern Recognition and Big-data analytics.

APPOINTMENTS HELD:

Feb 2018 to till date: Professor, Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India.

Dec 2015 to Jan 2018: Associate Professor, Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India.

Feb 2013 to Nov 2015: Associate Professor, School of Information Technology, Indian Institute of Technology Kharagpur, India.

May 2007 to Jan 2013: Assistant Professor, School of Information Technology, Indian Institute of Technology Kharagpur, India.

Oct 2005 to Apr 2007: Assistant Professor, Department of Electronics and Communication Engineering, Indian Institute of Technology Guwahati, India.

Feb 2005 to Sept 2005: Project Officer, Department of Computer Science and Engineering, Indian Institute of Technology Madras, India.

Jan 2001 to Jan 2005: Ph.D. Scholar, Department of Computer Science and Engineering, Indian Institute of Technology Madras, India.

Oct 1998 to Dec 2000: Assistant Professor, Department of Electronics and Communication Engineering, Bapatla Engineering College, India.

July 1993 to Sept 1998: Lecturer, Department of Electronics and Communication Engineering, Bapatla Engineering College, India.

INSTITUTIONAL RESPONSIBILITIES

July 2019 to June 2021: Faculty Adviser, M.Tech (CSE)

May 2016 to July 2017: Chairman: GATE-2017 and JAM-2017.

May 2015 to April 2016: Vice-Chairman: GATE-2016 and JAM-2016.

May 2014 to April 2015: Vice-Chairman: GATE-2015 and JAM-2015.

July 2015 to June 2017: Faculty Adviser, M.Tech (CSE) and M.Tech (IT).

July 2009 to June 2014: Faculty Adviser, M.Tech (IT).

Oct 2009 to Sept 2011: Associate Warden, LLR Hall.

AWARDS & RECOGNITION

1. Featuring in the Stanford University's list of Globally Top 2% Researcher List for the year 2020 in the research area of AI & Image Processing.

PATENTS

1. System and Method for Synchronizing Acoustic Signal of Vopiced Speech and its Corresponding Electroglottography Signal : Filed (Ref : 805/KOL/2014)

2. Method and apparatus to detect voice activity using Harmonics of Phase of Zero Frequency Filtered Speech Signal : Filed (Ref : 1237/KOL/2015)

GOOGLE SCHOLAR CITATIONS

Number of citations	=	4949
H-index	=	37
i10-index	=	121

SCOPUS INDEX CITATIONS

Scopus Index	=	35093227600
Number of citations	=	2893
H-index	=	28

PUBLICATIONS

Books

1. K. Sreenivasa Rao (2012), Predicting Prosody from Text for Text-to-Speech Synthesis, Springer, **ISBN** 978-1-4614-1337-0

2. K. Sreenivasa Rao and Shashidhar G. Koolagudi (2012), Emotion Recognition using Speech Features, Springer, **ISBN** 978-1-4614-5142-6

3. K. Sreenivasa Rao and Shashidhar G. Koolagudi (2013), Robust Emotion Recognition using Spectral and Prosodic Features, Springer, **ISBN** 978-1-4614-6359-7

4. K. Sreenivasa Rao and Anil Kumar Vuppala (2014), Speech Processing in Mobile Environments, Springer, **ISBN** 978-3-319-03116-3 (130 pages).

5. K. Sreenivasa Rao and Sourjya Sarkar (2014), Robust Speaker Recognition in Noisy Environments, Springer, **ISBN** 978-3-319-07129-9.

6. K. Sreenivasa Rao, V. Ramu Reddy and Sudhamay Maity (2015), Language Identification using Spectral and Prosodic Features, Springer, **ISBN** 978-3-319-17162-3.

7. K. Sreenivasa Rao and Dipanjan Nandi (2015), Language Identification using Excitation Source Features, Springer, **ISBN** 978-3-319-17724-3.

8. K. Sreenivasa Rao and Manjunath K E (2017), Speech Recognition using Articulatory and Excitation Source Features, Springer, ISBN 978-3-319-49219-3.

9. K. Sreenivasa Rao and Narendra N P (2019), Source Modeling Techniques for Quality Enhancement in Statistical Parametric Speech Synthesis, Springer, ISBN 978-3-030-02758-2.

Refereed Journals

1. Kumud Tripathi and K. Sreenivasa Rao, "CycleGAN based Speech Mode Transformation Model for Robust Multilingual ASR", *Circuits Systems & Signal Processing*, Springer, 2022.

2. Hareesh Mandalapu, Aravinda Reddy P N, Raghavendra ramachandra, K. Sreenivasa Rao, Pabitra Mitra, S. R. M. Prasanna and Christoph Busch, "Multilingual Audio-Visual Smartphone Dataset and Evaluation", in *IEEE Access*, 2022.

3. Kishore Kumar R and K. Sreenivasa Rao, “Phoneme Segmentation based Unsupervised Pattern Discovery and Clustering of Speech Signals”, *Circuits Systems & Signal Processing*, Springer, 2022.

4. Kishore Kumar R and K. Sreenivasa Rao, “A Novel approach to Unsupervised Pattern Discovery in Speech using Convolutional Neural Network”, *Computer Speech and Language*, Elsevier, vol. 71, pp. 1–13, 2022.

5. Nirmalya Sen, Md Sahidullah, Hemant Patil, Shyamal Kumar das Mandal, K. Sreenivasa Rao and Tapan Kumar Basu, “Utterance partitioning for speaker recognition: an experimental review and analysis with new findings under GMM-SVM framework”, *International Journal of Speech Technology (Springer)*, 2021.

6. Hareesh Mandalapu, Aravinda Reddy P N, Raghavendra ramachandra, K. Sreenivasa Rao, Pabitra Mitra, S. R. M. Prasanna and Christoph Busch, "Audio-Visual Biometric Recognition and Presentation Attack Detection: A Comprehensive Survey", in *IEEE Access*, vol. 9, pp. 37431-37455, 2021.

7. Kumud Tripathi and K. Sreenivasa Rao, VOP Detection for Read and Conversation Speech using CWT Coefficients and Phone Boundaries, *Journal of Ambient Intelligence and Humanized Computing (Springer)*, 2021.

8. Kumud Tripathi and K. Sreenivasa Rao, Robust Vowel Region Detection Method for Multimode Speech, *Multimedia Tools and Applications (Springer)*, 2021.

9. Tanumay Mandal, K. Sreenivasa Rao and Sanjay K. Gupta, Identification of glottal instants using electroglottographic signal for vulnerable cases of voicing, *IET Healthcare Technology Letters*, 2021.

10. Manjunath K E, Dinesh Babu Jayagopi, K. Sreenivasa Rao, Srinivasa Raghavan, and V. Ramasubramanian,. Approaches for Multilingual Phone Recognition in Code-Switched and Non-Code-Switched Scenarios using Indian Languages, T*ransactions on Asian and Low-Resource Language Information Processing (TALLIP)*, ACM, 2021.

11. Pradeep Rengaswamy, K. Sreenivasa Rao and Pallab Dasgupta, SongF0: A Spectrum based Fundamental Frequency Estimation for Monophonic Songs, *Circuits, Systems and Signal Processing (CSSP)*, Springer, 2020.

12. Debopriyo Banerjee, Krothapall Sreenivas Rao, Shamik Sural, and Niloy Ganguly BOXREC: Recommending a Box of Preferred Outfits in Online Shopping ACM Transactions on Intelligent Systems and Technology, 2020

13. M. Kiran Reddy, Pavu Alku and K. Sreenivasa Rao, "Detection of Specific Language Impairment in Children Using Glottal Source Features," in *IEEE Access*, vol. 8, pp. 15273-15279, 2020.

14. Pradeep Rengaswamy, M. Gurunath Reddy, K. Sreenivasa Rao and Pallab Dasgupta, hf0: A hybrid pitch extraction method for multimodal voice, *Circuits, Systems and Signal Processing (CSSP)*, Springer, 2020.

15. Manjunath K E, Dinesh Babu Jayagopi, K. Sreenivasa Rao and V. Ramasubramanian, Articulatory feature based methods for performance improvement of multilingual phone recognition systems using Indian languages, *SADHANA*, Academy Proceedings in Engineering Sciences, Indian Academy of Sciences, Springer, 2020.

16. Kumud Tripathi, M. Kiran Reddy and K. Sreenivasa Rao, Multilingual and multimode phone recognition for Indian languages, *Speech Communication*, Elsevier, 2020.

17. Kumud Tripathi and K. Sreenivasa Rao, VEP detection for Read, Extempore and Conversation speech, *IETE Journal of Research*, Taylor & Francis, 2020.

18. R Pradeep, M. Kiran Reddy, K. Sreenivasa Rao and Pallab Dasgupta, Robust f0 extraction from monophonic signals using adaptive sub-band filtering, Speech Communication, Elsevier, Vol. 116, pp. 77-85, 2020.

19. M. Kiran Reddy and K. Sreenivasa Rao, Excitation modelling using epoch features for statistical parametric speech synthesis, Computer Speech & Language., Vol. 60, 2019.

20. M. Kiran Reddy and K. Sreenivasa Rao, DNN-based cross-lingual voice conversion using Bottleneck Features, Neural Processing Letters, 2019.

21. R Pradeep, M. Kiran Reddy and K. Sreenivasa Rao, LSTM-based robust voicing decision applied to DNN-based speech synthesis, *Automatic Control and Computer Sciences*, Vol. 53, No. 4, PP. 328-332, 2019.

22. Saikat Biswas, Pabitra Mitra, and K Sreenivasa Rao, “Relation Prediction of Co-morbid Diseases Using Knowledge Graph Completion”, *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB),* 2019.

23. Y. Madhu Keerthana, M. Kiran Reddy , and K. Sreenivasa Rao, "CWT-Based Approach for Epoch Extraction From Telephone Quality Speech", *IEEE Signal Processing Letters*, Vol. 26 (8), pp. 1107-1111, Aug. 2019.

24. Hari Krishna D M and K. Sreenivasa Rao, "Children Story Classification in Indian Languages using Linguistic and Keyword based Features", ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 2019.

25. R. Pradeep and K. Sreenivasa Rao, Incorporation of Manner of Articulation Constraint in LSTM for Speech Recognition, *Circuits, Systems and Signal Processing (Springer),* 2019.

26. Kishore Kumar Ravi, Lokendra Birla and K Sreenivasa Rao, A Robust Unsupervised Pattern Discovery and Clustering of Speech Signals, *Pattern Recognition Letters (Elsevier)*, Vol. 116, pp. 254-261, 2018.

27. Gurunath Reddy M and K. Sreenivasa Rao, Predominant Melody Extraction from Vocal Polyphonic Music Signal by Time Domain Adaptive Filtering Based Method, *Circuits, Systems and Signal Processing (Springer),* Vol. 37, No. 7, pp. 2911-2933, 2018.

28. Manjunatha K E and K. Sreenivasa Rao, Improvement of Phone Recognition Accuracy using Articulatory Features, *Circuits, Systems and Signal Processing (Springer),* Vol. 37, No. 2, pp. 704-722, 2018.

29. M. Kiran Reddy and K. Sreenivasa Rao, Inverse filter based excitation model for HMM-based speech synthesis system, *IET Signal Processing*, Vol. 12, pp. 544-548, 2018.

30. Jainath Yadav, Md. Shah Fahad and K. Sreenivasa Rao, Epoch Detection from Emotional Speech Signal using Zero Time Windowing, Speech Communication, Elsevier, Vol. 96, pp. 142-149, 2018.

31. Jainath Yadav and K. Sreenivasa Rao, Neural Network and GMM based Feature Mappings for Consonant-Vowel Recognition in Emotional Environment, International Journal of Speech Technology (Springer), Vol. 21, pp. 421-433, 2018.

32. Manjunath K E, Dinesh Babu Jayagopi, K. Sreenivasa Rao and V. Ramasubramanian, "Development and Analysis of Multilingual Phone Recognition Systems using Indian Languages", International Journal of Speech Technology (Springer), 2018.

33. Prasenjit Dhara and K. Sreenivasa Rao, Automatic Note Transcription System for Hindustani

Classical Music, International Journal of Speech Technology (Springer), 2018.

34. Kumud Tripathi and K. Sreenivasa Rao, Improvement of phone recognition accuracy using speech mode classification, International Journal of Speech Technology (Springer), Vol. 21, pp. 489-500, 2018.

35. Arup Kumar Datta and K. Sreenivasa Rao, Language identification using phase information, International Journal of Speech Technology (Springer), Vol. 21, pp. 509-519, 2018.

36. M. Kiran Reddy and K. Sreenivasa Rao, Robust pitch extraction method for HMM-based speech synthesis system, *IEEE Signal Processing Letters*, Vol. 24 (8), pp. 1133-1137, 2017.

37. S. Samsekai Manjabhat, Shashidhar G. Koolagudi, K. Sreenivasa Rao and Pravin Bhaskar Ramteke, Raga and Tonic Identification in Carnatic Music, *Journal of New Music Research*, Vol. 46, No. 3, pp. 229-245, 2017.

38. Sourjya Sarkar and K. Sreenivasa Rao, Supervector-based approaches in discriminative freamework for speaker verification in noisy environments, *International Journal of Speech Technology (Springer)*, Vol. 20 (2), pp. 387-416, 2017.

39. Sunil Kumar S. B., Tanumay Mandal and K. Sreenivasa Rao, Robust Glottal Activity Detection using the Phase of an Electroglottographic Signal, *Biomedical Signal Processing & Control*, (Elsevier), Vol. 36, pp. 27-38, 2017.

40. Arijul Haque and K. Sreenivasa Rao, Modification of energy spectra, epoch parameters and prosody for emotion conversion in speech, *International Journal of Speech Technology (Springer)*, Vol. 20 (1), pp. 15-25, 2017.

41. Narendra N. P. and K. Sreenivasa Rao, Parameterization of excitation signal for improving the quality of HMM-based speech synthesis system, Circuits, Systems and Signal Processing (Springer), Vol. 36, No. 9, pp. 3650-3673, 2017.

42. N. P. Narendra and K. Sreenivasa Rao, Generation of creaky voice for improving the quality of HMM-based speech synthesis, *Computer Speech and Language*, Elsevier, Vol. 42, pp. 38-58, 2017.

43. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, "Implicit processing of LP residual for language identification", *Computer Speech and Language*, Elsevier, Vol. 41, pp. 68-87, 2017.

44. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, "Parametric representation of excitation source information for language identification", *Computer Speech and Language*, Elsevier, Vol. 41, pp. 88-115, 2017.

45. V. Ramu Reddy and K. Sreenivasa Rao, "Prosody modeling for syllable based text-to-speech synthesis using feedforward neural networks" *Neurocomputing*, Vol. 171, pp. 1323-1334, 2016.

46. K. E. Manjunath and K. Sreenivasa Rao, "Articulatory and Excitation Source Features for Speech

Recognition in Read, Extempore and Conversation Modes" *International Journal of Speech Technology (Springer)*, Vol. 19, pp. 121-134, 2016.

47. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, "Implicit excitation source features for robust language identification" *International Journal of Speech Technology (Springer)*, Vol. 18, pp. 459-477, 2015.

48. N. P. Narendra and K. Sreenivasa Rao, "Time-Domain Deterministic Plus Noise Model based Hybrid Source Modeling for HMM-Based Speech Synthesis" *Speech Communication (Elsevier)*, Vol. 77, pp. 65-83, 2016.

49. S. B. Sunil Kumar and K. Sreenivasa Rao, "Voice/Non-voice Detection Using Phase of Zero Frequency Filtered Speech Signal" *Speech Communication (Elsevier)*, Vol. 81, pp. 90-103, 2016.

50. Jainath Yadav and K. Sreenivasa Rao, Prosodic Mapping using Neural Networks for Emotion Conversion in Hindi Language, *Circuits, Systems and Signal Processing (Springer)*, Vol. 35, pp. 139-162, 2016.

51. Narendra N. P. and K. Sreenivasa Rao, Robust Voicing Detection and F0 Estimation for HMM-Based Speech Synthesis *by* N. P. *Circuits, Systems and Signal Processing (Springer)*, Vol. 34, pp. 2597-2619, 2015.

52. K. Sreenivasa Rao and Ketan Pachpande, Segmentation, indexing and retrieval of TV broadcast news bulletins using Gaussian mixture models and vector quantization codebooks, *International Journal of Speech Technology*, Springer, Vol. 17, pp. 259-269, 2014.

53. Sourjya Sarkar and K. Sreenivasa Rao, Stochastic feature compensation methods for speaker verification in noisy environments, *Applied Soft Computing, Elsevier,* Vol. 19, pp. 198-214, 2014.

54. K. E. Manjunath and K. Sreenivasa Rao, Source and System Features for Phone Recognition *International Journal of Speech Technology (Springer)*, Vol. 18, pp. 257-270, 2015.

55. K. Sreenivasa Rao and Shashidhar G Koolagudi, Recognition of emotions from video using acoustic and facial features, *Signal Image and Video Processing*, Springer Vol-9, No.-5, pp-1029-1045, July. 2015.

56. Avinash Kumar Singh, Jayanta Mukhopadhyay, K. Sreenivasa Rao and Kapinaiah Viswanath, Classification of Infant Cries Using Dynamics of Epoch Features, *Journal of Intelligent Systems*, Vol. 22, No. 3, pp. 253-267, July 2013.

57. K. Sreenivasa Rao, Dipanjan Nandi and Shashidhar G Koolagudi, Film segmentation and indexing using autoassociative neural networks, *International Journal of Speech Technology,* Springer, Vol. 17, pp. 65-74 (2014).

58. V. Ramu Reddy, Sudhamay Maity and K. Sreenivasa Rao, Recognition of Indian languages using multi-level spectral and prosodic features, *International Journal of Speech Technology,* Springer,

Vol. 16, No. 4, pp. 489-510, 2013.

59. K. Sreenivasa Rao, Sudhamay Maity and V. Ramu Reddy, Pitch Synchronous and Glottal Closure based Speech Analysis for Language Recognition, *International Journal of Speech Technology,* Springer Vol. 16, No. 4, pp. 413-430, 2013.

60. Jainath Yadav and K. Sreenivasa Rao, Detection of Vowel Offset Point from Speech Signal, *IEEE Signal Processing Letters*, Vol. 20, No. 4, pp. 299-302, 2013.

61. V. Ramu Reddy and K. Sreenivasa Rao, "Two-Stage Intonation Modeling using Feedforward Neural Networks for syllable based Text-to-Speech Synthesis", *Computer Speech and Language*, Elsevier, Vol. 27, pp. 1105-1126, 2013.

62. Anil Kumar Vuppala and K. Sreenivasa Rao, Vowel Onset Point Detection for Noisy Speech using Spectral Energy at Formant Frequencies, *International Journal of Speech Technology,* Springer, Vol. 16, No. 2, pp. 229-235, 2013.

63. Anil Kumar Vuppala and K. Sreenivasa Rao, "Speaker Identification under Background Noise using Features Extracted from Steady Vowel Regions", *International Journal of Adaptive control and Signal processing (Wiley),* Vol. 27, No. 9, pp. 781-792, September 2013.

64. K. Sreenivasa Rao and Anil Kumar Vuppala, Non-Uniform Time Scale Modification Using Instants of Significant Excitation and Vowel Onset Points, *Speech communication,* Elsevier, Vol. 55, No. 6, pp. 745-756, July 2013.

65. K. Sreenivasa Rao, Shashidhar G Koolagudi and Ramu Reddy Vempada, Emotion Recognition from Speech using global and local prosodic features, *International Journal of Speech Technology,* Springer, Vol. 16, No.2, pp. 143-160, June 2013.

66. K. Sreenivasa Rao and Shashidhar G Koolagudi, Characterization and recognition of emotions from speech using excitation source information, *International Journal of Speech Technology,* Springer, Vol. 16, No.2, pp. 181-201, June 2013.

67. N. P Narendra and K. Sreenivasa Rao, "Optimal weight tuning method for unit selection cost functions in syllable based text-to-speech synthesis", *Applied Soft Computing, Elsevier,* Vol. 13, pp. 773-781, 2013.

68. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, "Improved Speaker Identification in Wireless Environment", *International Journal of Signal and Imaging Systems Engineering, Inderscience,* Vol. 6, No. 3, pp. 130-137, 2013.

69. K. Sreenivasa Rao, Unconstrained pitch contour modification using instants of significant excitation, *Circuits, Systems & Signal Processing*, Springer, Vol. 31, No. 6, pp. 2133-2152, December 2012.

70. Anil Kumar Vuppala, Jainath Yadav, Saswat Chakrabarti and K. Sreenivasa Rao, Vowel Onset Point Detection for Low Bit Rate Coded Speech, *IEEE Transactions on Audio, Speech and*

Language Processing, Vol. 20, No. 6, pp. 1894-1903, Aug. 2012.

71. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, “Improved Vowel Onset Point Detection using Epoch Intervals", *International Journal of Electronics and Communications, Elsevier, Vol. 66, No. 8, pp. 697-700, Aug. 2012.*

72. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, “Spotting and Recognition of Consonant-Vowel Units from Continuous Speech using Accurate Vowel Onset Points", *Circuits, Systems & Signal Processing, Springer Vol. 31, No. 4, pp.1459-1474, Aug. 2012.*

73. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, “Improved consonant– vowel recognition for low bit-rate coded speech", *International Journal of Adaptive control and Signal processing (Wiley), Vol 26, Issue 4, pp. 333-349, April 2012.*

74. K. Sreenivasa Rao, Jainath Yadav, Sourjya Sarkar, Shashidhar G. Koolagudi and Anil Kumar Vuppala, Neural network based feature transformation for emotion independent speaker identification, *International Journal of Speech Technology,* Springer, Vol. 15, No.3, pp. 335-349, 2012.

75. Shashidhar G Koolagudi and K. Sreenivasa Rao, Emotion Recognition from Speech using sub-syllabic and pitch synchronous spectral features, *International Journal of Speech Technology,* Springer, Vol. 15, No.3, pp. 495-511, Dec. 2012.

76. Rabul Hussain Laskar, Kalyan Banerjee, Fazal Ahmed Talukdar and K. Sreenivasa Rao, A pitch synchronous approach to design voice conversion system using source-filter correlation, *International Journal of Speech Technology,* Springer, Vol. 15, No.3, pp. 419-431, 2012.

77. R. H. Laskar, D. Chakrabarty, F. A. Talukdar, K. Sreenivasa Rao and K. Banerjee, Comparing ANN and GMM in a voice conversion framework, *Applied Soft Computing,* Elsevier, Vol. 12, No. 11, pp. 3332-3342, Nov. 2012.

78. N. P Narendra and K. Sreenivasa Rao, “Syllable Specific Unit Selection Cost Functions for Text-to-Speech Syntheisis”, ACM *Transactions on speech and language processing,* Vol. 9, No. 3, Nov. 2012.

79. Shashidhar G Koolagudi and K. Sreenivasa Rao, Emotion Recognition from Speech : A Review, *International Journal of Speech Technology,* Springer, Vol. 15, No.3, pp. 99-117, 2012.

80. Shashidhar G Koolagudi and K. Sreenivasa Rao, Emotion Recognition from Speech using Source, System and Prosodic features, *International Journal of Speech Technology,* Springer, Vol. 15, No.3, pp. 265-289, 2012.

81. K. Sreenivasa Rao, Application of Prosody models for Developing Speech systems in Indian languages, *International Journal of Speech Technology,* Springer, Vol. 14, pp. 19-33, 2011.

82. Shashidhar G Koolagudi and K. Sreenivasa Rao, Two Stage Emotion Recognition Based on Speaking Rate, *International Journal of Speech Technology,* Springer, Vol. 14, pp. 35-48, 2011.

83. K. Sreenivasa Rao, “Role of Neural network models for developing speech systems”, *SADHANA*, Academy Proceedings in Engineering Sciences, Indian Academy of Sciences, Vol. 36, Part-5, pp. 783-836, Springer, Oct. 2011.

84. K. Sreenivasa Rao, V. K. Saroj, Sudhamay Maity and Shashidhar G Koolagudi, Recognition of emotions from video using neural network models, *Expert systems and applications,* Elsevier, Vol. 38, No.10, pp. 13181-13185, Sep. 2011.

85. N. P. Narendra, K. Sreenivasa Rao, Krishnendu Ghosh, Ramu Reddy Vempada and Sudhamay Maity “Development of Syllable-based Text-to-Speech Synthesis System in Bengali", *International Journal of Speech Technology,* Springer, Vol. 14, No.3, pp. 167-181, 2011.

86. K. Sreenivasa Rao and Shashidhar G Koolagudi, Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech, *Journal of Systems, Cybernetics and Informatics,* Vol. 9, No. 4, pp. 24-33, 2011.

87. Anil Kumar Vuppala, K. Sreenivasa Rao, Saswat Chakrabarti, P Krishnamoorthy, and S R M Prasanna “Recognition of Consonant-Vowel (CV) Units under Background Noise using Combined Temporal and Spectral Preprocessing", *International Journal of Speech Technology*, Springer, Vol. 14, No.3, pp. 259-272, 2011.

88. K. Sreenivasa Rao, Voice Conversion by Mapping the Speaker-specific features using Pitch Synchronous Approach, *Computer Speech and Language,* Elsevier, Vol. 24, pp. 474-494, July 2010.

89. K. Sreenivasa Rao, Real time prosody modification, *Journal of Signal and Information Processing*, Vol. 1, pp. 50-62, Nov. 2010.

90. K. Sreenivasa Rao and Shashidhar G Koolagudi, Selection of suitable features for modeling the durations of syllables, *Journal of Software Engineering and Applications,* Vol. 3, pp. 1107-1117, Dec. 2010.

91. K. Sreenivasa Rao and B. Yegnanarayana, Intonation modeling for Indian languages, *Computer Speech and Language,* Elsevier, Vol. 23, pp. 240-256, Apr. 2009.

92. K. Sreenivasa Rao and B. Yegnanarayana, Duration modification using Glottal Closure Instants and Vowel Onset Points, *Speech communication,* Elsevier, Vol. 51, pp. 1263-1269, Dec. 2009

93. K. Sreenivasa Rao and B. Yegnanarayana, “Modeling durations of syllables using neural networks”, *Computer Speech and Language,* Elsevier, Vol. 21, pp. 282-295, Apr. 2007.

94. K. Sreenivasa Rao, S. R. M. Prasanna and B. Yegnanarayana, “Determination of instants of significant excitation in speech using Hilbert envelope and group delay function”, *IEEE Signal Processing Letters*, vol. 14, No. 10, pp. 762-765, Oct. 2007.

95. K.Sreenivasa Rao and B. Yegnanarayana, “Prosody modification using instants of significant excitation,”*IEEE Trans. Speech and Audio Processing, vol. 14*, pp. 972-980, May 2006.

Book Chapters

1. Multilingual Phone Recognition: Comparison of Traditional versus Common Multilingual Phone-Set Approaches and Applications in Code-Switching, by K. E. Manjunath, K. M. Srinivasa Raghavan, K. Sreenivasa Rao, Dinesh Babu Jayagopi and V. Ramasubramanian, *in Advances in Signal Processing and Intelligent Recognition Systems*, Springer Nature Singapore, 2020.

2. Infant Cry Recognition using Source, System, Prosody and Epoch features, K. Sreenivasa Rao, A. K. Singh, J. Mukhopadhyay, Siva Ayyappa Kumar, Sunil Kumar S B and Ramu Raddy Vempada in Acoustic analysis of Infant Cries, Toddler Vocalizations, and Yound Adult Dysarthria, Speech Technology in Medicine and Health Care, published by de Gruyter, 2019.

3. Excitation modeling method based on inverse filtering for HMM-based speech synthesis" by M. Kiran Reddy and K. Sreenivasa Rao, in *Machine Intelligence and Signal Processing (AISC series of Springer),* 2017.

4. Hybrid source modeling method utilizing optimal residual frames for HMM-based speech synthesis by N. P. Narendra and K. Sreenivasa Rao in Mining Intelligence and Knowledge Exploration, (Published by: Springer, 2015) Reference.

5. Indexing and Retrieval of Speech Documents by Piyush Kumar P. Singh, K. E. Manjunath, R. Ravi Kiran, Jainath Yadav and K. Sreenivasa Rao in Advanced Computing, Networking and Informatics - Volume 1, (Published by: Springer, 2014) Reference.

6. Importance of Utterance Partitioning in SVM Classifier with GMM Supervectors for Text-Independent Speaker Verification, Nirmalya Sen, Hemant.A. Patil, Shyamal Kr. Das Mandal and K. Sreenivasa Rao, Mining Intelligence and Knowledge Exploration (LNCS), Vol. 8284, pp. 780-789, Springer, 2013.

7. Corpus Based Emotional Speech Synthesis in Hindi, Ravi Kalyan Bhakat, N.P. Narendra, and K. Sreenivasa Rao, Pattern Recognition and Machine Intelligence (LNCS), Vol. 8251, pp. 390-395, Springer, 2013.

8. Duration Modeling Using Multi-model Based on Positional Information, Ramu Reddy Vempada, and K. Sreenivasa Rao, Pattern Recognition and Machine Intelligence (LNCS), Vol. 8251, pp. 404-409, Springer, 2013.

9. Data-driven Phase break prediction for Bengali Text-to-Speech system, Krishnendu Ghosh and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 306, pp. 118-129, Springer, 2012.

10. Intensity modeling for Syllable based Text-to-Speech synthesis, V. Ramu Reddy and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary

Computing, Vol. 306, pp. 106-117, Springer, 2012.

11. Speaker Recognition in Emotional Environment, Shashaidhar G. Koolagudi, Kritika Sharma and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS) : Echo-friendly computing and communication systems, Vol. 305, pp 117-124, Springer, 2012.

12. Emotion Recognition from Semi Natural Speech using Artificial Neural Networks and Excitation Source Features, Shashaidhar Koolagudi, Swati Devliyal, Anurag Barthwal and K.Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 306, pp. 273-282, Springer, 2012.

13. Spoken language identification using spectral features, Shashaidhar G. Koolagudi, Deepika Rastogi and K.Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 306, pp. 496-497, Springer, 2012.

14. Vowel Recognition from Telephonic Speech using MFCC features and Gaussian Mixture Models, Shashidhar G. Koolagudi, Sujata Negi Thakur, Anurag Barthwal, Manoj Kumar Singh, Ramesh S. Rawat and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS) : Echo-friendly computing and communication systems, Vol. 305, pp 170-177, Springer, 2012.

15. Real Life Emotion Classification from speech using Gaussian Mixture Models, Shashidhar G. Koolagudi, Anurag Barthwal, Swati Devliyal, and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 306, pp. 250-261, Springer, 2012.

16. Text independent emotion recognition using spectral features, Rahul Chauhan, Jainath Yadav, S. G. Koolagudi and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 168, Part-2, pp. 359-370, Springer, 2011.

17. Segment Specific Concatenation Cost for Syllable Based Bengali TTS, N. P. Narendra and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 168, Part-2, pp. 371-382, Springer, 2011.

18. Effect of Noise on Recognition of Consonant-Vowel (CV) Units, Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 168, Part-2, pp. 191-200, Springer, 2011.

19. Effect of Noise on Vowel Onset Point Detection, Anil Kumar Vuppala, Jainath Yadav, K. Sreenivasa Rao and Saswat Chakrabarti, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 168, Part-2, pp. 201-211, Springer, 2011.

20. Robust speaker recognition in noisy environments: Using dynamics of speaker-specific prosody, Shashidhar G Koolkagudi, K. Sreenivasa Rao, Ramu Reddy, Anil Kumar Vuppala and Saswat Chakrabarti, Chapter 8, pp: 183-204, in Springer book "Forensic speaker recognition", 2011.

21. Emotion Classification Based on Speaking Rate, Shashidhar G. Koolagudi, Sudhin Ray and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 94, pp. 316-327, Springer, 2010.

22. Effect of speech coding on recognition of Consonant-Vowel (CV) units, Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, Communications in Computer and Information Science (CCIS): Contemporary Computing, Vol. 94, pp. 284-294, Springer, 2011.

23. Unit selection using linguistic, prosodic and spectral distance for developing text-to-speech system in Hindi, K. Sreenivasa Rao, Shashidhar G Koolagudi, Sudhamay Maity and Amol Taru, Pattern Recognition and Machine Intelligence (LNCS), Vol. 5909, pp. 531-536, Springer, 2009

24. Exploring Speech Features for Classifying Emotions along Valence Dimension, Shashidhar G. Koolagudi and K. Sreenivasa Rao, Pattern Recognition and Machine Intelligence (LNCS), Vol. 5909, pp. 537-542, Springer, 2009.

25. IITKGP-SESC : Speech database for emotion analysis, Shashidhar G. Koolagudi, Vuppala Anil Kumar, Saswat Chakrabarti, Sudhamay Maity and K. Sreenivasa Rao, Communications in Computer and Information Science (CCIS) : Contemporary Computing, Vol. 40, pp. 485-492, Springer, 2009.

26. Modeling supra-segmental features of syllables using neural networks, K. Sreenivasa Rao in Speech, Audio, Image and Biomedical signal processing using neural networks, pp. 71-95. Springer, 2008.

27. Voice Transformation by Mapping the Features at Syllable Level, K. Sreenivasa Rao, R. H. Laskar and Shashidhar G. Koolagudi in Lecture Notes in Computer Science: Pattern Recognition and Machine Intelligence, Vol. 4815, pp. 479-486, Springer, 2007.

28. Two-stage duration model for Indian languages using neural networks, K. Sreenivasa Rao and B.Yegnanarayana, in Lecture Notes in Computer Science : Neural Information Processing, Vol. 3316, pp. 1179-1185, Springer, 2004.

Refereed Conferences

1. Soumava Paul, Gurunath Reddy M, K Sreenivasa Rao and Partha Pratim Das, “Knowledge Distillation for Singing Voice Detection”, INTERSPEECH, BRNO, CZECHIA, 2021.

2. Gurunath Reddy M, K. Sreenivasa Rao and Partha Pratim Das, “Glottal Closure Instants Detection from EGG Signal by Classification Approach”, Interspeech-2020, Shanghai, China, Oct 25 – 29, 2020.

3. Kumud Tripathi and K. Sreenivasa Rao, “Multi-Lingual Speech Mode Classification Model for Indian Languages”, National Conference on Communications (NCC-2020) (IEEE Explore), IIT Kharagpur, Kharagpur, India., Feb. 21-23, 2020.

4. Gurunath Reddy M and K. Sreenivasa Rao, “Glottal Closure Instants Detection from Speech Signal by Deep Features Extracted from Raw Speech and Linear Prediction Residual”, Interspeech-2019, Graz, Austria, Sept 15-19, 2019.

5. Kiran Reddy M and K. Sreenivasa Rao, “Mel-scaled Wavelet-based Features for Spoofing Speech Detection”, International Conference on Electrical, Control and Computer Engineering, University of Malaysia, Pahang, Malaysia, July 2019.

6. Manjunath K E, Srinivasa Raghavan K. M, K. Sreenivasa Rao, Dinesh Babu Jayagopi and V Ramasubramanian, “Comparison of Common Multilingual Phone-set Based and LID-switched Monolingual Approaches for Multilingual Phone Recognition using Indian Languages”, *International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCT)*, IIIT Bangalore, Bangalore, India, 2019.

7. Gurunath Reddy M, Tanumay Mandal and K. Sreenivasa Rao, “Glottal Closure Instants Detection From Pathological Acoustic Speech Signal Using Deep Learning”, NIPS-2018 ML4H, Montreal, Canada, 2018.

8. Kishore Kumar R, Sandipan Sarkar, Pradeep Rengaswamy and K. Sreenivasa Rao, “Audio Mining: Unsupervised Spoken Term Detection over an Audio Database”, In *7th International Conference on Advances in Computing, Communications and Informatics (ICACCI-2018),* Bangalore, India. 2018.

9. Kumud Tripathi and K. Sreenivasa Rao, “Discriminative sparse representation for speech mode classification”, In *7th International Conference on Advances in Computing, Communications and Informatics (ICACCI-2018)*, Bangalore, India. 2018.

10. M. Kiran Reddy and K. Sreenivasa Rao, “DNN-based Bilingual (Telugu-Hindi) Polyglot Speech Synthesis,” in International Conference on Advances in Computing, Communications and Informatics (ICACCI-2018), Bangalore, India, 2018.

11. Tanumay Mandal, K. Sreenivasa Rao and Sanjay Kumar Gupta, "Classification of disorders in vocal folds using Electroglottographic Signal," In INTERSPEECH, Hyderabad, India, 2018.

12. Gurunath Reddy M, K. Sreenivasa Rao, Partha Pratim Das, “Harmonic-Percussive Source Separation of Polyphonic Music by Suppressing Impulsive Noise Events,” in INTERSPEECH, Hyderabad, India, 2018.

13. Kumud Tripathi and K. Sreenivasa Rao, "Analysis of sparse representation based feature on speech mode classification," in INTERSPEECH, Hyderabad, India, 2018.

14. Manjunath K E, K. Sreenivasa Rao, Dinesh Babu Jayagopi and V Ramasubramanian, “Indian languages ASR: A multilingual phone recognition framework with IPA based common phone-set, predicted articulatory features and feature fusion”, In INTERSPEECH-2018 (ISCA) , Hyderabad, India. 2018.

15. R. Pradeep and K. Sreenivasa Rao, “Modifying LSTM Posteriors with Manner of Articulation Knowledge to Improve Speech Recognition Performance”, in *IEEE 17th International Conference on Machine Learning and Applications (ICMLA 2018),* Orlando, Florida, USA, 2018.

16. Suma S M, Shashidhar G. Koolagudi, Pravin B. Ramteke and K. Sreenivasa Rao, “Note Transcription from Carnatic Music”, International Conference on Advanced Computing Networking and Informatics (ICACNI-2018), Springer, NIT Silchar, Assam – 788010, India.

17. Debopriyo Banerjee, Niloy Ganguly, Shamik Sural and K. Sreenivasa Rao, “One for the Road: Recommendign Male Street Attire”, 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-2018), Melbourne, Australia, June 2018.

18. Tanumay Mandal and K. Sreenivasa Rao, “Robust Detection of Glottal Activity using Unwrapped Phase Electroglottographic Signal”, International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2018) (IEEE Explore), Calgary, Alberta, Canada, April 2018.

19. R. Pradeep and K. Sreenivasa Rao, “Manner of Articulation Based Split Lattices for Phoneme Recognition”, National Conference on Communications (NCC-2018) (IEEE Explore), IIT Hyderabad, Hyderabad, India., Feb. 25-28, 2018.

20. Gurunath Reddy, M. and K. Sreenivasa Rao, "Predominant vocal melody extraction from enhanced partial harmonic content", EUSIPCO, Kos, Greece, 2017.

21. Gurunath Reddy, M. and K. Sreenivasa Rao, "Automatic Evaluation of Hindustani Learner’s SARGAM Practice", EUSIPCO, Kos, Greece, 2017.

22. Gurunath Reddy, M. and K. Sreenivasa Rao, "Neutral to Joyous Happy Emotion Conversion", 14th IEEE India Council International Conference (INDICON), IIT Roorkee, December, 2017.

23. M. Kiran Reddy and K. Sreenivasa Rao, "Excitation modeling method based on inverse

filtering for HMM-based speech synthesis" in Machine Intelligence and Signal Processing (AISC series of Springer), IIT Indore, Indore, December 2017.

24. Pradeep R and K. Sreenivasa Rao, "Split Acoustic Modeling in Decoder for Phoneme Recognition", 14th IEEE India Council International Conference (INDICON), IIT Roorkee, December, 2017.

25. S.B. Sunil Kumar, K.Sreenivasa Rao, Tanumay Mandal," Accurate Synchronization of Speech and EGG signal using Phase Information", INTERSPEECH, Stockholm, Sweden, 2017

26. Kumud Tripathi, Parakrant Sarkar and K. Sreenivasa Rao, “Sentence Based Discourse Classification for Hindi Story Text-to-Speech (TTS) System”. *In 13th International Conference on Natural Language Processing (ICON)*, Dec 17-20, IIT (BHU), Varanasi, 2016

27. Prasenjit Dhara, Pradeep Rengaswamy, K. Sreenivasa Rao, “Designing Automatic Note Transcription System for Hindustani Classical Music”, I*EEE 5th International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sept. 21-24, Jaipur, 2016.

28. Jainath Yadav, Md. Shah Fahad, Ranjeet Kumar and K. Sreenivasa Rao, "Speaker Identification in Emotional Environment using Trajectory-based Stochastic Feature Mapping", *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2016)*, Dec. 2016, Jaipur, India.

29. Manish Rai, Jainath Yadav, K. Sreenivasa Rao, Neetish Kumar and Md. Shah Fahad, "Language Identification Using PLDA Based on I-Vector in Noisy Environment", *International Conference on Advances in Computing, Communications and Informatics (ICACCI-2016)*, Sept. 2016.

30. Pradeep R and K S Rao, “Deep Neural Networks for Kannada Phoneme recognition”, *IEEE 9 th International Conference on Contemporary Computing (IC3)*, pp. 67-72, New Delhi, August 2016.

31. Pradeep R, Gurunath Reddy M, K S Rao and P Dasgupta, “A Robust Non-Parametric and Filtering Based Approach for Glottal Closure Instant Detection”, *INTERSPEECH*, San Francisco, USA, 2016.

32. Gurunath Reddy M and K. Sreenivasa Rao, "Enhanced Harmonic Content and Vocal Note Based Predominant Melody Extraction from Vocal Polyphonic Music Signals", *INTERSPEECH*, San Francisco, USA, 2016.

33. Parakrant Sarkar and K. Sreenivasa Rao, Data-driven pause prediction for synthesis of storytelling style speech based on discourse modes, International Conference on Electronics, Computing and Communication Technologies (CONECCT-2015) (IEEE Explore), IIIT Bangalore, Bangalore, India., July 10-11 (2015)

34. Harikrishna D M and K. Sreenivasa Rao, Children story classification based on structure of the

story, International Conference on Advances in Computing, Communications and Informatics (ICACCI-2015) (IEEE Explore), SCMS Aluva, Kochi, India., August 10-13 (2015)

35. Parakrant Sarkar and K. Sreenivasa Rao, Modeling pauses for synthesis of storytelling style speech using unsupervised word features, International Conference on Advances in Computing, Communications and Informatics (ICACCI-2015) (IEEE Explore), SCMS Aluva, Kochi, India., August 10-13 (2015)

36. R. Pradeep, Prasenjit Dhara, K. Sreenivasa Rao and Pallab Dasgupta, Raga identification based on normalized note histogram features, International Conference on Advances in Computing, Communications and Informatics (ICACCI-2015) (IEEE Explore), SCMS Aluva, Kochi, India., August 10-13 (2015)

37. Arijul Haque and K. Sreenivasa Rao, Analysis and modification of spectral energy for neutral to sad emotion conversion, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

38. Arup Kumar Dutta and K. Sreenivasa Rao, Robust language identification using power normalized cepstral coefficients, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

39. Gurunath Reddy M and K. Sreenivasa Rao, Neutral to happy emotion conversion by blending prosody and laughter, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

40. Harikrishna D M and K. Sreenivasa Rao, Multi-stage children story speech synthesis for Hindi, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

41. Parakrant Sarkar and K. Sreenivasa Rao, Analysis and modeling pauses for synthesis of storytelling speech based on discourse modes, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

42. Randheer Bagi, Jainath Yadav and K. Sreenivasa Rao, Improved recognition rate of language identification system in noisy environment, International Conference on Contemporary Computing (IC3-2015) (IEEE Explore), JIIT Noida, Noida, India., August 20-22 (2015)

43. N. P. Narendra and K. Sreenivasa Rao, Automatic detection of creaky voice using epoch parameters, Interspeech-2015, Dresden, Germany., September 6-10 (2015)

44. Arijul Haque and K. Sreenivasa Rao, Modification and incorporation of excitation source features for emotion conversion, International Conference on Computer, Communication and Control (IC4-2015) (IEEE Explore), MGI Indore, Indore, India., September 10-12 (2015)

45. Arup Kumar Dutta and K. Sreenivasa Rao, Analysis of linear prediction residual signal, its magnitude and phase for language identification on NIST LRE (2003) database, International

Conference on Computer, Communication and Control (IC4-2015) (IEEE Explore), MGI Indore, Indore, India., September 10-12 (2015)

46. Gurunath Reddy M, Procheta Sen, Manjunath K E, Arup Dutta, Arijul Haque, Parakrant Sarkar and K. Sreenivasa Rao, Automatic pitch accent contour transcription for Indian languages, International Conference on Computer, Communication and Control (IC4-2015) (IEEE Explore), MGI Indore, Indore, India., September 10-12 (2015)

47. Harikrishna D M and K. Sreenivasa Rao, Classification of children stories in Hindi using keywords and POS density, International Conference on Computer, Communication and control (IC4-2015) (IEEE Explore), MGI Indore, Indore, India., September 10-12 (2015)

48. Tanumay Mandal and K. Sreenivasa Rao, Analysis of perturbation in pitch period and contact quotient for classifying age groups, International Conference on Computer, Communication and Control (IC4-2015) (IEEE Explore), MGI Indore, Indore, India., September 10-12 (2015)

49. N. P. Narendra and K. Sreenivasa Rao, A deterministic plus noise model of excitation signal using principal component analysis for parametric speech synthesis, International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2016) (IEEE Explore), Shanghai, China., March 20-25 (2016)

50. Gurunath Reddy M and K. Sreenivasa Rao, Predominant melody extraction from vocal polyphonic music signal, International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2016) (IEEE Explore), Shanghai, China., March 20-25 (2016)

51. N. P. Narendra, M. Kiran Reddy and K. Sreenivasa Rao, Excitation modeling for HMM-based speech synthesis based on principal component analysis, National Conference on Communications (NCC-2016) (IEEE Explore), IIT Guwahati, Guwahati, India., March 4-6 (2016)

52. Harikrishna D M and K. Sreenivasa Rao, Emotion-specific features for classifying emotions in story text, National Conference on Communications (NCC-2016) (IEEE Explore), IIT Guwahati, Guwahati, India., March 4-6 (2016)

53. Parakrant Sarkar, Arijul Haque, Arup Kumar Dutta, Gurunath Reddy M., Harikrishna D. M., Prasenjit Dhara, Rashmi Verma, N. P. Narendra, Sunil Kr. S.B., Jainath Yadav and K. Sreenivasa Rao, Designing prosody rule-set for converting neutral TTS speech to storytelling style speech for Indian languages Bengali, Hindi and Telugu, 7th International Conference on Contemporary Computing (IC3), IEEE Explore, JIIT, Noida, India, August 20-24 (2014)

54. Ramu Reddy Vempada, K. Sreenivasa Rao and Parakrant Sarkar, Duration Modeling by Multi-Models based on Vowel Production characteristics, 11th International Conference on Natural Language Processing, (ICON 2014), Goa University, Goa, India., December 18-21 (2014)

55. Rashmi Verma, Parakrant Sarkar and K. Sreenivasa Rao, Conversion of Neutral speech to Storytelling Style Speech, Proceedings of the Eighth IEEE International Conference on

Advances in Pattern Recognition (ICAPR 2015) (IEEE Explore), ISI, Kolkata, India., January 4-7 (2015)

56. Parakrant Sarkar and K. Sreenivasa Rao, Data-Driven Pause Prediction for Speech Synthesis in Storytelling Style Speech, 21st National Conference on Communication (NCC-2015) (IEEE Explore), IIT Bombay, Mumbai, India., Feb 27 - March 1, 2015 (2015)

57. K. E. Manjunath, K. Sreenivasa Rao and M. Gurunath Reddy, Two-stage phone recognition system using articulatory and spectral features, International Conference on Signal Processing and Communication Engineering Systems (SPACES-2015) (IEEE Explore), Guntur, India, January 2-3 (2015)

58. K. E. Manjunath, K. Sreenivasa Rao and M. Gurunath Reddy, Improvement of phone recognition accuracy using source and system features, International Conference on Signal Processing and Communication Engineering Systems (SPACES-2015) (IEEE Explore), Guntur, India, January 2-3 (2015)

59. M. Gurunath Reddy, D. M. Harikrishna, K. Sreenivasa Rao and K. E. Manjunath, Telugu emotional story speech synthesis using SABLE markup language, International Conference on Signal Processing and Communication Engineering Systems (SPACES-2015) (IEEE Explore), Guntur, India, January 2-3 (2015)

60. M. Bhaykar, K. Sreenivasa Rao and K. E. Manjunath, Speaker Independent Recognition from Speech using Combination of Different Classification Models, 18th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI-2014), Florida, USA., July (2014)

61. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, Sub-segmental, Segmental and Supra-segmental Analysis of Linear Prediction Residual Signal for Language Identification, IEEE International Conference on Signal Processing and Communications (SPCOM) (IEEE Explore), IISC Bangaluru, Bangaluru, India., July (2014)

62. Dipanjan Nandi, Arup Kumar Dutta and K. Sreenivasa Rao, Significance of CV transition and steady vowel regions for language identification, IEEE International Conference on Contemporary Computing (IC3) (IEEE Explore), JIIT, Noida, India., August 20-24 (2014).

63. N. P. Narendra and K Sreenivasa Rao, Optimal residual frame based source modeling for HMM-based speech synthesis, Proceedings of the Eighth IEEE International Conference on Advances in Pattern Recognition (ICAPR 2015) (IEEE Explore), ISI, Kolkata, India., January 4-7 (2015)

64. Shashidhar, Koolagudi G., B. Shivakranthi, K. Sreenivasa Rao, and Pravin B. Ramteke, Contribution of Telugu vowels in identifying emotions, Proceedings of the Eighth IEEE International Conference on Advances in Pattern Recognition (ICAPR 2015) (IEEE Explore), ISI, Kolkata, India., January 4-7 (2015)

65. Jainath Yadav, Anshu Kumari and K. Sreenivasa Rao, Emotion Recognition using LP Residual

at Sub-segmental, Segmental and Supra-segmental levels, International Conference on Communication, Information & Computing Technology (ICCICT), Central University of Bihar, Patna, Indi, January (2015)

66. Jainath Yadav and K. Sreenivasa Rao, Generation of emotional speech by prosody imposition on Sentence, Word and Syllable level fragments of neutral speech, International Conference on Cognitive Computing and Information Processing (CCIP) (IEEE Explore), Noida, India., March 3-4 (2015)

67. Sourjya Sarkar and K.Sreenivasa Rao, A Novel Boosting Algorithm for Improved i-Vector based Speaker Verification in Noisy Environments, 15th Annual Conference on the International Speech Communication Association (INTERSPEECH-2014), Singapore, September 14-18 (2014)

68. Manjunath K. E. and K. Sreenivasa Rao, Automatic Phonetic Transcription for Read, Extempore and Conversation Speech for an Indian Language: Bengali, NCC-2014, IIT Kanpur, Kanpur, India, Feb. 2014.

69. Jainath Yadav and K. Sreenivasa Rao, Emotional-speech synthesis from neutral-speech using prosody imposition, *International Conference on Recent Trends in Computer Science and Engineering (ICRTCSE-2014)*, Central University of Bihar, Patna, India, February 8-9, 2014.

70. V. Ramu Reddy and K. Sreenivasa Rao, High quality text-to-speech synthesis system with efficient duration models developed using coding schemes based on vowel production characteristics, *13th International Conference on Intelligent Systems Design and Applications (ISDA-13)*, Universiti Putra Malaysia, Malaysia, December 2013.

71. R Ravi Kiran, Sunil Kumar, Manjunath K E, Apoorv Chaturvedi, Biswajit Satapathy, Debadatta Pati and K. Sreenivasa Rao, Automatic Phonetic and Prosodic Transcription for Speech Documents in Indian languages: Bengali and Oriya, *International Conference on Natural Language Processing (ICON-2013)*, CDAC, Noida, India, December 2013.

72. Bellamkonda Bhaskar, Dipanjan Nandi and K. Sreenivasa Rao, Analysis of Language Identification Performance based on Gender and Hierarchial Grouping Approaches, *International Conference on Natural Language Processing (ICON-2013)*, CDAC, Noida, India, December 2013.

73. Avinash Kumar Singh, Jayanta Mukhopadyay, Sunil Kumar S B and K. Sreenivasa Rao, Infant Cry Recognition using Excitation Source Features, *IEEE INDICON*, Mumbai, India, Dec. 2013.

74. Sourjya Sarkar, K. Sreenivasa Rao and Dipanjan Nandi, Multilingual Speaker Recognition on Indian Languages, *IEEE INDICON*, Mumbai, India, Dec. 2013.

75. Manjunath K E, Sunil Kumar. S. B, Debadatta Pati, Biswajit Satapathy and K. Sreenivasa Rao, Development of Consonant-Vowel Recognition Systems for Indian Languages: Bengali and Oriya, *IEEE INDICON*, Mumbai, India, Dec. 2013.

76. Sunil Kumar S B, K. Sreenivasa Rao and Debadatta Pati, Phonetic and Prosodically Rich Transcribed Speech Corpus in Indian Languages: Bengali and Oriya, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

77. Manjunath K E, K. Sreenivasa Rao and Debadatta Pati, Development of Phonetic Engine for Indian Languages: Bengali and Oriya, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

78. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, Robustness of Excitation Source Information for Language Independent Speaker Recognition, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

79. Sourjya Sarkar and K. Sreenivasa Rao, Significance of Utterance Partitioning in GMM-SVM based Speaker Verification in Varying Background Environment, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

80. Dipanjan Nandi, Debadatta Pati and K. Sreenivasa Rao, Language Identification using Hilbert Envelope and Phase Information of Linear Prediction Residual, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

81. Hemant A Patil, Tanvina B Patel, Nirmesh J Shah, Hardik B Sailor, Raghava Krishnan, G R Kasthuri T Nagarajan, Lilly Christina, Naresh Kumar, Veera Raghavendra, S P Kishore, S R M Prasanna, Nagaraj Adiga, Sanasam Ranbir Singh, Konjengbam Anand, Pranaw Kumar, Bira Chandra Singh,S L Binil Kumar, T G Bhadran, T Sajini, Arup Saha, Tulika Basu, K Sreenivasa Rao, N P Narendra, Anil Kumar Sao, Rakesh Kumar, Pranhari Talukdar, Purnendu Acharyaa, Somnath Chandra, Swaran Lata and Hema A Murthy, *A Syllable-Based Framework for Unit Selection Synthesis in 13 Indian Languages*, 16th International Oriental COCOSDA Conference, Gurgoan, India, Nov. 2013.

82. Jainath Yadav and K. Sreenivasa Rao, Analysis of Detection of Vowel Offset Point for Coded Speech, 6th International Conference on Contemporary Computing (IC3), pp. 485-490, IEEE Explore, August 2013.

83. Avinash Kumar Singh, Jayanta Mukhopadhyay and K. Sreenivasa Rao, “Classification of Infant Cries using Source, System and Supra-segmental Features”, Indian Conference in Medical Informatics and Telemedicine (ICMIT-2013), IIT Kharagpur, India, March 2013.

84. Avinash Kumar Singh, Jayanta Mukhopadhyay, Sreenivasa Rao K. and Viswanath K, Classification of Infant Cries Using Dynamics of Epoch Features, International Conference on Communications, VLSI and Signal Processing, SIT, Tumkur, India, Feb 2013.

85. Sourjya Sarkar and K. Sreenivasa Rao, Speaker verification in noisy environment using GMM Supervectors, NCC-2013, IIT Delhi, Delhi, India, Feb. 2013.

86. Manav Bhaykar and K. Sreenivasa Rao, Speaker Dependent, Speaker Independent and Cross

Language Emotion Recognition From Speech Using GMM and HMM, NCC-2013, IIT Delhi, Delhi, India, Feb. 2013.

87. Avinash Kumar Singh and K. Sreenivasa Rao, Classification of Infant Cries Using Epoch and Spectral Features, NCC-2013, IIT Delhi, Delhi, India, Feb. 2013.

88. K. Sreenivasa Rao, Ketan Pachpande, V. Ramu Reddy and Sudhamay Maity, Segmentation of TV Broadcast News Using Speaker Specific Information, NCC-2012, IIT Kharagpur, Kharagpur, India, Feb. 2012.

89. Ramu Reddy Vempada and K. Sreenivasa Rao, Modeling the Intensity of Syllables Using Classification and Regression Trees, NCC-2012, IIT Kharagpur, Kharagpur, India, Feb. 2012.

90. Sudhamay Maity, Anil Kumar Vuppala, K. Sreenivasa Rao and Dipanjan Nandi, IITKGP-MLILSC Speech Database for Language Identification, NCC-2012, IIT Kharagpur, Kharagpur, India, Feb. 2012.

91. Krishnendu Ghosh and K. Sreenivasa Rao, Subword based Approach for Grapheme-to-Phoneme Conversion in Bengali Text-to-Speech Synthesis System, NCC-2012, IIT Kharagpur, Kharagpur, India, Feb. 2012.

92. Ramu Reddy Vempada, Siva Ayyappa Kumar B and K. Sreenivasa Rao, Characterization of Infant Cries using Spectral and Prosodic features, NCC-2012, IIT Kharagpur, Kharagpur, India, Feb. 2012.

93. Ramu Reddy Vempada and K. Sreenivasa Rao, Intonation modeling using linguistic, production and prosodic constraints for syllable-based TTS system, International Conference on Modeling Optimization and Computing (ICMOC-2012), Vol. 38, pp. 2772-2783, India, April 2012.

94. Shashidhar G. Koolagudi, Anurag Barthwal, Swati Devliyal, and K. Sreenivasa Rao, “Real Life Emotion Classification using Spectral Features and Gaussian Mixture Models.”, International Conference on Modeling Optimization and Computing (ICMOC-2012), India, April 2012.

95. Shashidhar G. Koolagudi, Swati Devliyal, Bhavna Chavla, Anurag Barthwal and K. Sreenivasa Rao, “Recognition of Emotions from Speech using Excitation Source Features” , International Conference on Modeling Optimization and Computing (ICMOC-2012), India, April 2012.

96. Shashidhar G. Koolagudi, Deepika Rastogi and K. Sreenivasa Rao , “Identification of Languages using Mel Frequency Cepstral Coefficients”, International Conference on Modeling Optimization and Computing (ICMOC-2012), India, April 2012.

97. V. Ramu Reddy and K. Sreenivasa Rao, “Better Human Computer Interaction by enhancing the quality of Text-to-Speech synthesis, 4th International Conference on Intelligent Human Computer Interaction (IHCI-2012), IIT Kharagpur, India, December 2012.

98. Anil Kumar Vuppala, Jainath Yadav, K. Sreenivasa Rao and Saswat Chakrabarti, Effect of Low Bit Rate Speech Coding on Epoch Extraction, ICDECOM, BIT-Mesra, Ranchi, Feb. 2011.

99. Shashidhar G. Koolagudi, Ramu Reddy, Jainath Yadav and K. Sreenivasa Rao, IITKGP-SEHSC: Hindi speech corpus for emotion analysis, ICDECOM, BIT-Mesra, Ranchi, Feb. 2011.

100. Shashidhar G. Koolagudi, Nitin Kumar and K. Sreenivasa Rao, Speech emotion recognition and verification using segmental level prosodic analysis, ICDECOM, BIT-Mesra, Ranchi, Feb. 2011.

101. N. P. Narendra and K. Sreenivasa Rao, Syllable specific target cost formulation for syllable based text-to-speech synthesis in Bengali, 2nd Int. Conf. Computer and Communication Technology (ICCCT), pp. 180-184, Allahabad, India, September 2011.

102. Ramu Reddy Vempada and K. Sreenivasa Rao, Intonation modeling using FFNN for syllable based Bengali text to speech synthesis, 2nd Int. Conf. Computer and Communication Technology (ICCCT), pp. 334-339, Allahabad, India, September 2011.

103. Krishnendu Ghosh and K. Sreenivasa Rao, Memory-based Data-driven Approach for Grapheme to Phoneme. Conversion in Bengali Text to Speech Synthesis System, INDICON-2011, Hyderabad, India, Dec. 2011.

104. Narendra N. P., K. Sreenivasa Rao, Ramu Reddy Vempada, Krishnendu Ghosh and Sudhamay Maity, Development of Bengali screen reader using Festival speech synthesizer, INDICON-2011, Hyderabad, India, Dec. 2011.

105. Krishnendu Ghosh, Ramu Reddy Vempada and K. Sreenivasa Rao, *Phrase Break Prediction for Bengali Text to Speech Synthesis System*, International Conference on Natural Language Processing (ICON), Chennai, India, Dec. 2011.

106. Krishnendu Ghosh, Shashidhar G Koolagudi, Sudhamay Maity, Ramu Redday, Narendra and K. Sreenivasa Rao, Grapheme to phoneme conversion in Bengali for Festival based TTS framework, *International Conference on Natural Language Processing (ICON)*, IIT Kharagpur, Kharagpur, Dec. 2010.

107. Shashidhar G.Koolagudi and K. Sreenivasa Rao, Real life emotion classification using VOP and pitch based spectral features, *IEEE INDICON*, Kolkata, Dec. 2010.

108. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti," Effect of speech coding on speaker identification", *IEEE INDICON*, Kolkata, Dec 2010

109. Anil Kumar Vuppala, K. Sreenivasa Rao and Saswat Chakrabarti, Two-stage isolated consonant-vowel (CV) unit recognition in Indian languages, *ICCCD*, IIT Kharagpur, Kharagpur, Dec. 2010.

110. Hema A Murthy, Ashwin Bellur, Vinodh Viswanath, Badri Narayanan, Anila Susan, G

Kasturi, K. Sreenivasa Rao, Sudhamay Maity, N. P. Narendra, Ramu Reddy, Krishnendu Ghosh, K. G. Sulochana, E. L. Abhilash, T. Sajini, M. Sasikumar, Bira Chandra Singh, Pranav Kumar, P. Vijayaditya, E. Veera Raghavendra and Kishore Prahllad, Building Unit Selection Speech Synthesis in Indian languages: An Initiative by an Indian Consortium, *COCOSDA 2010*, Kathmandu, Nepal, Nov 2010.

111. K. Sreenivasa Rao, Anil Kumar Vuppala, Saswat Chakrabarti and Leena Dutta, "Robust Speaker Recognition on Mobile Devices", *IEEE International Conference on Signal Processing and Communication (SPCOM)*, IISC Bangalore, India, July 2010.

112. Shashidhar G. Koolagudi, Ramu Reddy and K. Sreenivasa Rao, "Emotion Recognition from Speech Signal using Epoch Parameters", *IEEE International Conference on Signal Processing and Communication (SPCOM)*, IISC Bangalore, India, July 2010.

113. K. Sreenivasa Rao, Saurav Nandy and Shashidhar G Koolagudi, Identification of Hindi Dialects using Speech, *14th World Multiconference on Systemics, Cybernetics and Informatics (WMSCI 2010)*, Orlando, USA, June 2010

114. K. Sreenivasa Rao, V K Saroj, Sudhamay Maity and Shashidhar G Koolagudi, Recognition of Emotions from Video, *14th World Multiconference on Systemics, Cybernetics and Informatics (WMSCI 2010)*, Orlando, USA, June 2010

115. K Sreenivasa Rao, Ramu Reddy, Sudhamay Maity and Shashidhar G Koolagudi, Characterization of Emotions using the Dynamics of Prosodic Features, *Int. Conf. on Speech Prosody*, Chicago, USA, May 2010.

116. S. R. M. Prasanna, D. Govind, K. Sreenivasa Rao and B. Yegnanarayana, Fast Prosody Modification using Instants of Significant Excitation, *Int. Conf. on Speech Prosody*, Chicago, USA, May 2010.

117. Anil Kumar Vuppala, Saswat Chakrabarti, and K. Sreenivasa Rao, Feature mapping using neural network models for coded speech recognition, *14th Int. Conf. Cognitive and Neural systems (ICCNS 2010)*, Boston, USA, May 2010.

118. K. Sreenivasa Rao, Anil Kumar Vuppala, Sudhin Ray, and Shashidhar G. Koolagudi, Feature mapping using neural network models for coded speech recognition, *14th Int. Conf. Cognitive and Neural systems (ICCNS 2010)*, Boston, USA, May 2010.

119. K. Sreenivasa Rao, Anil Kumar Vuppala, and Shashidhar G. Koolagudi, Neural network models for emotion recognition using glottal pulse characteristics, *14th Int. Conf. Cognitive and Neural systems (ICCNS 2010)*, Boston, USA, May 2010.

120. K. Sreenivasa Rao, Jainath Yadav, Anil Kumar Vuppala, and Shashidhar G. Koolagudi, Two stage neural network model for recognition of Indian languages from speech, *14th Int. Conf. Cognitive and Neural systems (ICCNS 2010)*, Boston, USA, May 2010.

121. Sabin Kafley, Anil Kumar Vuppala, Arun Chauhan and K. Sreenivasa Rao, Continuous digit

recognition in mobile environment, IEEE TechSym 2010, IIT Kharagpur, Kharagpur, April 2010.

122. Arun Chauhan, Shashidhar G. Koolagudi, Sabin Kafley and K. Sreenivasa Rao, Emotion recognition using LP residual, IEEE TechSym 2010, IIT Kharagpur, Kharagpur, April 2010

123. Anil Kumar Vuppala and K. Sreenivasa Rao, Neural Network Models for Speech Recognition in Mobile Environments, *13th Int. Conf. on Cognitive and Neural systems, Boston*, MA, USA, May 2009.

124. Suparnakanti Das, Sudhamay Maity, K. Sreenivasa Rao and Pabitra Mitra, Strategies for selecting optimal text for Bengali ASR system, *13-th International Conference on Speech and Computer (SPECOM'2009),* St. Petersburg, Russia, June, 2009.

125. K. Sreenivasa Rao, S. R. M. Prasanna and T. Vidya Sagar, Significance of Word and Syllable level Information for Expressive speech processing, submitted to *7th International Conference on Advances in Pattern Recognition*, Kolkata, India, Feb. 2009.

126. Shashidhar G. Koolagudi, Sourav Nandi and K. Sreenivasa Rao, Spectral Features for Emotion Recognition, *IEEE International Advance Computing Conference (IACC 2009)*, Patiala, India, Mar. 2009.

127. Shashidhar G. Koolagudi and K. Sreenivasa Rao, Neural Network Models for Capturing Prosodic Knowledge for Emotion Recognition, *12th Int. Conf. on Cognitive and Neural systems, Boston*, MA, USA, May 2008.

128. K. Sreenivasa Rao and Shashidhar G. Koolagudi, Transformation of speaker characteristics in speech using support vector machines, *15th International Conference on Advanced Computing & Communication (ADCOM-2007),* IIT Guwahati, Guwahati, India, Dec 2007.

129. K. Sreenivasa Rao, S. R. M. Prasanna and T. Vidya Sagar, Emotion Recognition using Multilevel Prosodic Information, *Workshop on Image and Signal Processing (WISP-2007),* IIT Guwahati, Guwahati, India, Dec 2007.

130. T. Vidya Sagar, K. Sreenivasa Rao, S. R. M. Prasanna and S. Dandapat, "Characterization and Incorporation of emotions in speech", *IEEE INDICON*-2006, IIT Delhi, Delhi, India, Sep. 2006.

131. K. Sreenivasa Rao and B.Yegnanarayana, " Voice conversion by prosody and vocal tract modification", *9th IEEE Int. Conf. Information Technology*, Bhubaneswar, Orissa, India, Dec. 2006.

132. K. Sreenivasa Rao and B.Yegnanarayana, "Modeling syllable duration in Indian languages using support vector machines", in *Proc. 2nd Int. Conf. Intelligent Sensing and Information Processing(ICISIP-2005),* Chennai, India, Jan. 2005.

133. L. Mary, K. Sreenivasa Rao and B.Yegnanarayana, "Neural network classifiers for language

identification using syntactic and prosodic features", in *Proc. 2nd Int. Conf. Intelligent Sensing and Information Processing(ICISIP-2005)*, Chennai, India, Jan. 2005.

134. K. Sreenivasa Rao and B.Yegnanarayana, "Impact of constraints on prosody modeling for Indian languiages", in *Proc. 3rd International Conference on Natural Language Processing (ICON-2004)*, Hyderabad, India, Dec. 2004, pp. 225-236.

135. K. Sreenivasa Rao and B.Yegnanarayana, "Neural network models for text-to-speech synthesis", in *Proc. 5th International Conference on Knowledge Based Computer Systems (KBCS-2004)*, Hyderabad, India, Dec. 2004, pp. 520-530.

136. K. Sreenivasa Rao and B.Yegnanarayana, "Intonation modeling for Indian languages", in *Proc. 8th Int. Conf. on Spoken Language Processing (Interspeech-2004)*, Jeju Island, Korea, Oct. 2004, pp. 733-736.

137. L. Mary, K. Sreenivasa Rao, S.V. Gangashetty, and B.Yegnanarayana, ``Neural network models for capturing duration and intonation knowledge for language and speaker identification," in *Proc. 8th Int. Conf. on Cognitive and Neural systems*, Boston, MA, USA, May. 2004.

138. K. Sreenivasa Rao and B.Yegnanarayana, ``Modeling syllable duration in Indian languages using neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, Montreal, Quebec, Canada, May. 2004.

139. K. Sreenivasa Rao, S. V. Gangashetty, and B. Yegnanarayana, ``Duration analysis for Telugu language," in *Proc. Int. Conf. on Natural Language Processing*, Mysore, India, Dec. 2003.

140. S.Rajendran, K. Sreenivasa Rao, B.Yegnanarayana, and K.N. Reddy, ``Syllable duration in broadcast news in Telugu: A preliminary study," in *Proc. National Conf. on Language Technology Tools: Implementation of Telugu/Urdu*, Hyderabad, India, Oct. 2003.

141. S. V. Gangashetty, K. Sreenivasa Rao, A.Nayeemullakhan, C. C. Sekhar, and B.Yegnanarayana, ``Combining evidence from multiple modular networks for recognition of consonant-vowel units of speech," in *Proc. Int. Joint Conf. on Neural Networks*, Portland, Oregon, USA, July 2003.

142. K. Sreenivasa Rao and B.Yegnanarayana, ``Prosodic manipulation using instants of significant excitation," in *Proc. IEEE Int. Conf. Multimedia and Expo*, Baltimore, Maryland, USA, July 2003.

143. K. Sreenivasa Rao, S. V. Gangashetty, and A.Nayeemullakhan, ``Distribution capturing ability of autoassociative neural network models for recognition of consonant-vowel utterances," in *Proc. Int. Conf. on Cognitive and Neural systems*, Boston, MA, USA,May 2003.

144. K. Sreenivasa Rao and B. Yegnanarayana, ``Prosodic Manipulation using Instants of

Significant Excitation," in *Proc. Int. Conf. Acoust., Speech Signal processing*, Apr 2003.

145. K. Kiran Kumar, K. Sreenivasa Rao, and B. Yegnanarayana, ``Duration Knowledge for Text-to-Speech system for Telugu," in *Proc. Int. Conf. Knowledge based computer systems*, Mumbai, India, Dec 2002.

146. B. Yegnanarayana, S.R. Mahadeva Prasanna, and K. Sreenivasa Rao, ``Speech Enhancement using excitation source information," in *Proc. Int. Conf. Acoust., Speech Signal Processing*, Orlando, FL, USA, May 2002.

RESEARCH CONTRIBUTIONS:

Acquisition and incorporation of prosody

- Appropriate models are proposed for capturing the prosodic information, and signal processing methods are developed to incorporate the prosodic information into speech.

Expressive speech processing

- Emotional speech databases were developed in Hindi and Telugu.
- Implicit and explicit excitation source features, pitch synchronous and sub-syllabic spectral features and multi-level global and local prosodic features are proposed for characterizing the emotions.
- Hierarchical models are proposed for improving the accuracy of emotion recognition.
- Acquisition and incorporation of emotion-specific knowledge for developing emotion-aware speech systems
- Robust signal processing methods were proposed for detecting the speech events in expressive speech.
- Signal processing and machine learning methods were proposed for voice as well as expression transformations.

Speech/Speaker/Language Recognition

- Signal processing methods are developed to detect important speech events such as vowel onset and offset points in presence of speech coding and background noise.
- Efficient speech and speaker recognition systems are developed in mobile environment by exploiting the crucial speech events and hybrid recognition models.
- Articulatory and excitation source features are proposed for recognition of speech in read, extempore and conversation modes.

- Detection and Incorporation of Manner of Articulation (MoA) knowledge in LSTM for speech recognition.
- Incorporation of MoA knowledge in decoding graph and lattice rescoring.
- Acquisition and incorporation of MoA knowledge in CTC based End-toEnd ASR system.
- Robust speaker recognition techniques were proposed based on stochastic feature compensation and total variability speaker modeling frameworks.
- Multi-SNR speaker models are proposed for speaker recognition in varying background environments.
- Emotion compensation techniques are proposed for Speaker recognition in emotional environments.
- Robust language recognition systems were developed using (i) spectral features extracted from glottal closure regions, (ii) multi-level prosodic features, (iii) implicit and explicit excitation source features and (iv) phase information.

Text-to-speech synthesis

- Syllable based Bengali text-to-speech system was developed.
- Efficient text analysis and phase break prediction models are developed.
- Accurate prosody models are developed using feedforward neural networks.
- Appropriate syllable-specific features, unit-selection cost functions and weight selection criterion are proposed.
- Bengali screen reader was developed and demonstrated to visually challenged people at NAB Kolkata by conducting workshop for 5 weeks.
- Laughter synthesis and incorporation of appropriate laughter segments for generating the Happy emotion.
- Storyteller style speech synthesis systems were developed in four Indian languages (Hindi, Telugu, Bengali and Malayalam)
- Story-specific prosody models were proposed for enhancing the storyteller style synthesized speech quality.
- Effective source models were proposed and integrated in statistical parametric speech synthesis for generating both modal and creaky voices.

- Customizable polyglot text-to-speech synthesis using cross-lingual voice conversion framework,

Voice conversion

- Autoassociative neural network models are proposed for mapping pitch-synchronous speaker-specific characteristics between source and target speakers.
- High quality cross-lingual voice conversion using bottleneck features from deep neural networks

Analysis and synthesis of vocal folds activity

- Accurate and robust detection of significant instants within a glottal cycle using phase information
- Accurate parameterization of vocal folds activity using phase information
- Analysis and investigation of vocal disorders using the parameters of a glottal cycle
- Simulation of vocal folds activity using EGG and Speech signals.

Analysis of pathological speech

- Dysarthric speech analysis

Music signal processing

- Signal processing methods for extracting the predominant melody from polyphonic music
- Accurate melody extraction from singing voice
- Unified framework for robust and accurate f0 extraction from speech as well as songs.
- Automatic Note transcription
- Automatic generation of lyrics and melody using deep architectures.

Unsupervised Pattern Discovery for Speech Indexing and Retrieval Applications

- Unsupervised pattern discovery methods at frame level and phoneme level
- Unsupervised pattern discovery using image processing, nearest depth first search (NDFS) traversal and machine learning techniques.
- Clustering of speech documents based on the knowledge acquired from the pattern discovery techniques.
- Indexing of clusters using unique keywords and linking the related speech documents to indexed keywords.

- Development of an efficient speech retrieval system using the indexed speech repository.
- Development of the web-interface for captiuring the speech queries and retieving the relevant speech documents in response to the given user query.

PROTO-TYPE SYSTEMS DEVELOPED:

1. Text-to-Speech synthesis system for an Indian language Bengali
2. Bi-lingual (English and Bengali) screen reader for visually challenged people
3. Multi-stage storyteller style speech synthesizers in Hindi, Telugu, Bengali and Malayalam.
4. Polyglot speech synthesis for development of multilingual TTS
5. Vocalfold activity synthesizer is developed using phase information of EGG
6. Online Hindustani music tutoring system (basic SARGAM)
7. Automatic Tanpura tuner system
8. Speech interface for filling Aadhar and Train reservation forms in English and Hindi.
9. Development of speech indexing and retrieval system using unsupervised pattern discovery techniques.

PROJECTS GUIDED

(M. Tech Level): Completed: 44, On-going: 7

1. Real time Prosody modification
2. Voice conversion by mapping spectral and prosodic features
3. Emotion analysis on stressed speech database (SUSE)
4. Issues in the development of Text-to-speech (TTS) system for Hindi
5. Exploration of excitation source information for the analysis of expressive speech
6. Emotion recognition from video using facial features
7. Analysis of Hindi dialects using speech features
8. Emotion recognition from video using acoustic features
9. Continuous digit recognition in noisy environments
10. Emotion recognition using speaking rate features
11. Robust speaker recognition using PCA and ICA
12. Emotion recognition using LP residual signal of speech
13. Accurate vowel onset point detection from speech
14. Emotion recognition under text and speaker variabilities
15. Emotion recognition using prosodic features from sentence, word and syllable levels
16. Infant cry recognition using spectral and prosodic features
17. Automatic news bulletin segmentation using news reader speaking characteristics
18. Indexing and retrieval of speech documents
19. Robust emotion recognition using multiple classifiers
20. Infant cry recognition using source and system features
21. Corpus based emotion synthesis in Hindi
22. Language identification from speech using GMM-SVM framework

23. Characterization of speaker emotion from speech using GMM-SVM framework
24. Conversion of neutral speech to story-teller speech
25. Language identification using multi-stage discriminative approaches
26. Indexing and retrieval of video lectures using speech
27. Speaker recognition using phase information from EGG signal
28. Language identification using phonotactic features
29. Audio retrieval using GMM-posterior features
30. Identification of North Indian language dialects using source, system and prosodic features
31. Automatic generation of companion chords based on melody
32. Unsupervised Keyword Recognition in Speech and clustering of Speech Documents (Vineeth Chinmay Karra, M.Tech (Dual Degree)
33. Multi Column Artificial Neural Network For Vocal and Non Vocal Classification of Bollywood Music (M.Tech, Ravindra Singh Koranga)
34. End-to-End Automatic Speech Recognition Model for Noisy Speech, Vikash Kumar Chaurasia (17CS60R51)
35. Exploiting Machine Learning Techniques for Unsupervised Clustering of Speech Utterances, Aniket Deroy (17CS60R19).
36. Indexing and Retrieval of Video Lectures, Narayan Kunal (18CS60R56).
37. Unsupervised Pattern Discovery Technique for Speech Information Retrieval, Harshita Chouhan (18CS60R40).
38. Context-Aware Outfit Compatibility Prediction & FITB, Vaibhav Mishra (18CS60R34).
39. Detection of Copy-Move Tampering in Digital Audio, Vivek Sharma, (19CS60R38).
40. Semantic segmentation of lecture videos using deep learning techniques for efficient indexing, Hemlata Ramesh Chandewar (19CS60R41)
41. Exploring features for detection of repeated patterns in multi-speaker speech, Shirish Kumar Shukla (19CS60R54).
42. Semi-supervised clustering of speech signals using manner of articulation, Damera Ajay (19CS60R63)
43. Audio Visual Biometric Verification System using Siamese Network, Saurav Likhar (16CS30048).
44. Speech Emotion Recognition Using Convolutional Neural Networks With Data Augmentation And Transfer Learning, B Kushal (16CS30008).

(B.Tech Level): Completed: 9, On-going: 6

1. Incorporation of emotion knowledge into Speech
2. Emotion recognition using multi-level information
3. Person-specific music retrieval using adaptive GMMs
4. Development of portable book reading device
5. Evaluation of SARGAM tutoring system
6. Incorporation of emotions into neutral speech
7. A music identification and recommender system using audio fingerprinting
8. Evaluation of Learner's SARGAM practice (Manikya Singh-14CS10032, K. Jaya Chandra, 14CS30045)
9. Unsupervised keyword spotting using GM-posteriogram features (Alla Sairam, 13CS10007)

THESIS SUPERVISION:

(PhD Level) Completed: 11; On-going: 12

1. Emotion recognition from speech using source, system and prosodic features
2. Vowel onset point detection for speech processing in mobile environment
3. Enhancement of Speaker Recognition Performance for Short Test Segments using GMM-SVM and Polynomial Classifiers
4. Source modeling for improving the quality of HMM-based speech syntheis.
5. Emotion Transformation using Significant Events of Speech
6. Incorporation of manner of articulation knowledge to improve the performance of automatic speech recognition system.
7. Source modeling and cross-lingual voice conversion for high quality polyglot parametric speech synthesis.
8. Speech recognition in multilingual and multimode environment.
9. Modeling Aesthetics in Fashion Recommendation using Machine Learning
10. An Unified Framework for Accurate Pitch Estimation from Human Voice
11. Unsupervised Pattern Discovery and Clustering of Speech Utterances for Information Retrieval

(MS Level) Completed: 14

1. Identification of Indian languages using spectral and prosodic features
2. Text analysis for Bengali text-to-speech synthesis system
3. Prosody modeling for Bengali text-to-speech synthesis system
4. Optimal unit selection criterion for Bengali text-to-speech synthesis system
5. Robust Speaker Recognition in Mobile Environment
6. Analysis and Synthesis of Vocal Fold Activity using an Electroglottographic Signal
7. Articulatory and Excitation Source Features for Phone Recognition
8. Language identification using excitation source features
9. Predominant Melody Extraction from Vocal Polyphonic Music Signals
10. Multi-stage Children Story Speech Synthesis
11. Prosody Modeling for Storytelling Style Speech Synthesis
12. Emotion Conversion in Speech using Source, System and Prosody Modifications
13. Robust language identification using magnitude and phase information
14. Automatic Transcription of Hindustani Classical Music

COURSES TAUGHT:

IIT Kharagpur:

1. Image Processing
2. Programming and Data Structures
3. Computer Architecture and Operating Systems
4. Advanced topics in speech processing
5. Speech processing technology
6. Communication systems and networking
7. Internet and web technology
8. Introduction to Internet
9. Fundamentals of Computing Systems
10. Programming and data structures lab
11. Computing systems lab
12. Internet technologies lab
13. Computer Networks Lab

IIT Guwahati

1. Communication Networks
2. Analog circuits

Bapatla Engineering College

1. Digital signal processing
2. Advanced communication systems
3. Communication systems
4. Analog communication
5. Digital communication
6. Principles of communication
7. Control systems
8. Applied electronics
9. TV engineering
10. Microprocessors and applications
11. Advanced microprocessors
12. Digital electronics
13. Object oriented programming
14. Operating systems
15. Computer networks

WORKSHOP / CONFERENCE / INVITED SEMINAR / SHORT-TERM COURSE ORGANIZED:

Type	Name	Year	# participants
Workshop	Anti-spoofing strategies for unseen audio-visual attacks	2019	40
Workshop	Winter School on Speech and Audio Processing (WISSAP) at College of Engineering Trivendrum	2019	150
Conference	INTERSPEECH – 2018 (i) Area Chair for Speech Synthesis and Spoken Language Generation track and (ii) Member in Finance Committee	2018	1500
Workshop	Winter School on Speech and Audio Processing (WISSAP) at IIT Guwahati	2018	175
Workshop	Design and Development of Speech Interfaces for Form-Filling Applications	2017	20
Workshop	Acquisition, representation, analysis and classification of bio-signals	2017	20
Seminar	Excitation source features for the analysis of distant speech	2015	25
Workshop	Prosodically Guided Phonetic Engine for Searching Speech Databases in Indian Languages	2014	36
Conference (Publication chair)	International conference on Human Computer Interaction (IHCI-2012)	2012	250
Conference (Publication chair)	Indian conference on Medical Informatics and Telemedicine (ICMIT-2013)	2013	75
Short-term course	Advanced Networking	2008	15
Short-term course	Image and video processing	2009	60
Short-term course	Wireless adhoc and sensor networks	2010	50
Short-term course	Basic computers for visually challenged	2011	25
Short-term course	Recent trends in speech processing	2011	30
Short-term course	Recent trends in speech processing	2012	30
Workshop	Basics of speech processing	2009	24
Workshop	Speech enhancement using spectral and temporal methods	2009	25
Workshop	Development of Speaker verification systems	2010	30
Workshop	Signal processing techniques for various speech applications	2010	30
Seminar	Processing of multispeaker speech	2009	25
Seminar	Speech enhancement by temporal processing	2009	25
Seminar	Speech enhancement by spectral processing	2009	25

Seminar	Speaker recognition under degraded conditions	2009	25
Seminar	Signal processing issues in speech processing	2010	30
Seminar	A computational framework for exploring the role of speech production in speech processing/recognition	2011	30
Seminar	Multi-level excitation source information for speaker verification	2012	35

INVITED TALKS:

1. Detection of Vocal Folds Disorders using EGG Signals, Online Short Term Course (STC) on "Deep Learning-Based Speech Processing Techniques for Smart Health and Education Systems-Concepts, Recent Trends and Key Challenges" (DLSP-2021), (September 20-24, 2021) at IIT Patna, Patna, Bihar, India.
2. Deep Learning Techniques for Speech and Music Processing, On-line Faculty Development Program on "Artificial Intelligence, Machine Learning and Robotics", organized by Dept of Computer Science and Engineering, University of Science and Technology, Meghalaya, Inmdia. September 6-10, 2021.
3. Deep Learning Techniques for Audio Indexing and Retrieval Applications, Workshop on Machine Learning for Speech Processing at NIT Sikkim, South Sikkim,India, 2021
4. Unsupervised Pattern Discovery using Convolution Neural Networks, AICTE sponsored Faculty Development Program on "Machine Learning in Image Processing Applications", at PSCMR College of Engineering & Technology, Vijayawada, India. (2021)
5. Incorporation of MOA knowledge in End-to-End ASR system, AICTE sponsored Faculty Development Program on "Machine Learning and its Applications", at PSCMR College of Engineering & Technology, Vijayawada, India. (2021)
6. Application of Deep Learning Techniques for Development of Speech and Audio Systems, AICTE Training And Learning (ATAL) Academy Sponsored Faculty Development Programme on Speech Processing using Deep Learning, IIT Dharwad. (2020)
7. Introduction to speech processing and its applications
8. Speech and audio applications on Internet.
9. Expressive speech processing and its applications
10. Prosody modeling and its application to various speech tasks
11. Acquisition and incorporation of prosody for developing speech systems
12. Speech interface to mobile phones
13. Speaker recognition in varying background environments
14. Mapping of speech parameters for voice conversion and emotion transformation
15. Development text-to-speech synthesis systems in Indian languages

SPONSORED PROJECTS:

PROJECT – 1	
Project Title	Development of Anti-spoofing Strategies for Unseen Audio-Visual Attacks in Handheld Devices
Funding Agency	MHRD
Project Value	59.24 Lakhs
Duration	2 years (March 2019 to March 2021)
Status	In progress
Role	Principal Investigator
PROJECT – 2	
Project Title	Development of Speech Interface for Form-Filling Application (SIFA) in Five Indian Languages
Funding Agency	MHRD & MCIT
Project Value	103.20 Lakhs
Duration	2 years (Sept. 2017 to Aug. 2019)
Status	In progress
Role	Principal Investigator
PROJECT – 3	
Project Title	Center of Excellence in Robotics
Funding Agency	531.75 Lakhs
Project Value	SRIC, IIT Kharagpur
Duration	5 years (April. 2014 to March 2019)
Status	In progress
Role	Co-Principal Investigator
PROJECT – 4	
Project Title	Decoding and exploring ancient classification of Indian music through machine learning method and audience response
Funding Agency	150 Lakhs
Project Value	MHRD
Duration	4 years (Jan. 2014 to Jan. 2018)
Status	In progress
Role	Co-Principal Investigator
PROJECT – 5	

Project Title	Speech Based Access of Agricultural Commodity Prices and Weather Information in 12 Indian Lnaguages/Dialects, ASR Phase-II
Funding Agency	54.62 Lakhs
Project Value	MCIT
Duration	4 years (Aug. 2014 to July 2018)
Status	Completed
Role	Principal Investigator
PROJECT – 6	
Project Title	Robot as a social assistance and robotic framework for exploring rugged and potentially hazardous terrain
Funding Agency	SRIC, IIT Kharagpur
Project Value	6.99 Lakhs
Duration	3.5 years (May 2014 to Oct 2017)
Status	Completed
Role	Co-Principal Investigator
PROJECT – 7	
Project Title	Accurate analysis of vocal folds activity for speech and biomedical applications
Funding Agency	MHRD
Project Value	33 Lakhs
Duration	4 years (Apr. 2014 to March 2018)
Status	Completed
Role	Principal Investigator
PROJECT – 8	
Project Title	Development of Text-to-Speech Synthesis System for Indian Languages (Phase II)
Funding Agency	MCIT
Project Value	97.94 Lakhs
Duration	5 years (Feb. 2012 to June 2017)
Status	Completed
Role	Principal Investigator
PROJECT – 9	
Project Title	Development Prosodically guided phonetic engine for searching speeh databases in Indian languages
Funding Agency	MCIT
Project Value	60.38 Lakhs

Duration	3.5 years (Jan 2012 to June 2015)
Status	Completed
Role	Principal Investigator
PROJECT – 10	
Project Title	Development of Text-to-Speech Synthesis System for Indian Languages (Phase I)
Funding Agency	MCIT
Project Value	35.66 Lakhs
Duration	2 yrs 9 months (Apr. 2009 to Dec 2011)
Status	Completed
Role	Principal Investigator
PROJECT – 11	
Project Title	Speaker recognition system for handheld devices in varying background environments
Funding Agency	DST
Project Value	23.03 Lakhs
Duration	4 years (Oct 2008 to Sept. 2012)
Status	Completed
Role	Principal Investigator
PROJECT – 12	
Project Title	Shruti: A vernacular speech recognition system
Funding Agency	Media Lab Asia
Project Value	20 Lakhs
Duration	3 yrs (Mar 2009 to Feb 2012)
Status	Completed
Role	Co-Principal Investigator
PROJECT – 13	
Project Title	Characterization and Incorporation of Emotions in Speech
Funding Agency	ISIRD (IITKGP)
Project Value	3 Lakhs
Duration	4 yrs (Oct 2007 to Sept. 2011)
Status	Completed
Role	Principal Investigator

OTHER ACTIVITIES:

Membership:

Vice Chairman, IEEE Signal Processing Society Chapter, IIT Kharagpur, 2020-2022.
IEEE Senior Member
IEEE Signal Processing Society Member
ISCA Member
ACM Member
Life Member for Indian Society for Technological Education (ISTE)

Editorial Board Membership:

Editorial Board Member to International Journal of Biosciences and Technology.
Editorial Board Member to The Open Signal Processing Journal.
Editorial Board Member to International Journal of UWB Communications and Systems.

Reviewer:

Reviewer for “IEEE Transactions on Audio Speech and Language Processing”
Reviewer for IEEE Signal Processing Letters, IEEE Access
Reviewer for “Journal of Acoustic Society of America (JASA)”, JASA Express letters.
Reviewer for “Speech Communication”, Elsevier Publishers.
Reviewer for “Computer Speech and Language”, Elsevier Publishers.
Reviewer for “Computers and Electrical Engineering” Elsevier Publishers.
Reviewer for “ERASIP Journal on Audio, Speech and Music Processing”, Elsevier Publishers.
Reviewer for “ACM Transactions on Speech and Language Processing”
Reviewer for “Journal of Circuits, Systems and Signal Processing”, Springer Publishers.
Reviewer for “International Journal of Speech Technology” Springer Publishers
Reviewer for “Biomedical Signal Processing and Control” Elsevier Publishers.
Reviewer for “Applied Soft Computing” Elsevier Publishers.
Reviewer for “Neurocomputing” Elsevier Publishers.
Reviewer for “International Journal of Adaptive Control and Signal Processing” Wiley Publishers
Reviewer for “Journal of Multimedia”, Academy Publisher
Reviewer for many reputed international conferences such as ICASSP, Interspeech, ICSLP, SPCOM and ICON.

BOS Member:

(1) NIT Suratkal
(2) JNTU Kakinada
(3) Vignan University
(4) RVR College of Engineering, Guntur, AP, India.
(5) QIS College of Engineering and Technology, Ongole, AP, India.
(6) Bapatla Engineering College, Bapatla, Guntur, AP.

PRSG Member:

(1) DeITy sponsored project “Development of Robust Person Authentication system using online handwriting” by RGU, Arunachal Pradesh.
(2) DeITy sponsored project “Audio Visual Spoof Resistance and Robust PID” by NIT Nagaland
(3) DeITy sponsored project “Robust Speaker Recognition” by RGU (Arunachal Pradesh)
(4) Project Review Committee Member, Technology Innovation Hub on Autonomous Navigation Foundation, IIT Hyderabad

BOG Nominee:

For NITK Suratkal during selection committee meeting from Feb 10-11, 2018 for recruitment of Faculty members in various departments.

Expert Member:

Faculty selection comittee member for NIT Agartala, IIIT Dharwad, IIIT Sricity and JNTU Anantapur