

Samik Sadhu

Curriculum Vitae

Center for Language and Speech Processing
Johns Hopkins University
Hackerman 226, 3400 North Charles Street, Maryland, USA
☎ +1 443 214 8199
✉ sadhusamik@gmail.com, samiksadhu@jhu.edu
🌐 samiksadhu.weebly.com/

Educational Background

- 2017–Present **Doctor of Philosophy**, *Johns Hopkins University*, Electrical and Computer Engineering Department, Center for Language and Speech Processing, Baltimore, USA, Present CGPA: 4/4.
Ph.D. advisor: Hynek Hermansky
- 2015–2017 **Master of Engineering**, *Indian Institute of Science*, Electrical Engineering Department, Bangalore, India, CGPA: 6.9/8.
Thesis advisor: Prasanta Kumar Ghosh
- 2011–2015 **Bachelor of Engineering**, *Jadavpur University*, Electrical Engineering Department, Kolkata, India, CGPA: 8.3/10.

Relevant Courses

- Signal Processing Speech Information Processing, Audio Signal Processing, Digital Image Processing, Time-frequency Analysis, Speech and Auditory Processing by Humans and Machines, Wavelets and Filter Banks, Signal Quantization and Compression
- Statistics and Applied Mathematics Matrix Analysis and Linear Algebra, Data Analytics, Random Processes, Detection and Estimation Theory
- Optimization and Machine Learning Linear and Nonlinear Optimization, Nonlinear Optimization II, Machine Learning

Research Interests

- Robust Automatic Speech Recognition
- Representation Learning
- Unsupervised and Semi-supervised Learning
- Multi-stream and Multi-modal Speech Recognition
- Continual Learning

Research Projects

- Modulation-based features for Automatic Speech Recognition**, *Hand-crafted as well as data-driven robust front-end feature extraction capturing temporal modulations of speech in different frequency sub-bands [3,4,7,9,11].*
work done at Johns Hopkins University

Representation learning, *Using self-supervised learning models like Auto-regressive Predictive Coding(APC), Contrastive Predictive Coding (CPC), Wav2vec etc. to extract acoustic representations for Automatic Speech Recognition (ASR) [6]* .

work done at Johns Hopkins University & Amazon Alexa ASR

Continual learning in Automatic Speech Recognition, *Gradual training of speech recognition systems with a continuous stream of data [5,8]*.

work done at Johns Hopkins University

Confidence scores in Speech Recognition, *Generating confidence scores during ASR inference and out-of-domain detection [1,5,8]*.

work done at Johns Hopkins University

Keyword spotting in continuous speech, *Using light-weight (non-ASR) systems to detect keywords in continuous speech [12,13]*.

work done at Indian Institute of Science, Bangalore

Scholarships and Awards

Research Award Masters Project Support Grant from Robert Bosch Centre for Cyber Physical Systems(RBCCPS) for Keyword Spotting in Continuous Speech

Internship Experience

Jun-Aug 2019 **Amazon Alexa ASR**, *Applied Scientist Intern, Seattle, USA*, Manager: Roland Maas.
Project: Representation Learning in ASR using Contrastive Predictive Coding(CPC) and Autoregressive Predictive Coding(APC)

Jun-Aug 2020 **Amazon Alexa ASR**, *Applied Scientist Intern, Seattle, USA*, Manager: Che-Wei Huang.
Project: Representation Learning in ASR using wav2vec 2.0 and wav2vec-C [6]

Programming Experience

ASR and Machine Learning Kaldi, ESPNet, PyTorch-Kaldi, PyTorch, Tensorflow 2.0

Programming Languages Python, Bash, MATLAB

Publications (Google Scholar)

- 1 Sustek, Martin, Samik Sadhu, Lukas Burget, Hynek Hermansky, Jesus Villalba, Laureano Moro-Velazquez, and Najim Dehak. "Stabilized training of joint energy-based models and their practical applications." arXiv e-prints (2023): arXiv-2303.
- 2 Sadhu, Samik, and Hynek Hermansky. "Blind Signal Dereverberation for Machine Speech Recognition." arXiv preprint arXiv:2210.00117 (2022).
- 3 Sadhu, Samik, and Hynek Hermansky. "Importance of Different Temporal Modulations of Speech: A Tale of Two Perspectives." arXiv e-prints (2022): arXiv-2204.(**Accepted at ICASSP 2023**)
- 4 Sadhu, S., Hermansky, H. (2022) Complex Frequency Domain Linear Prediction: A Tool to Compute Modulation Spectrum of Speech. Proc. Interspeech 2022, 3208-3212, doi: 10.21437/Interspeech.2022-11095

- 5 Sustek, M., Sadhu, S., Hermansky, H. (2022) Dealing with Unknowns in Continual Learning for End-to-end Automatic Speech Recognition. Proc. Interspeech 2022, 1046-1050, doi: 10.21437/Interspeech.2022-11139
- 6 Sadhu, S., He, D., Huang, C.-W., Mallidi, S.H., Wu, M., Rastrow, A., Stolcke, A., Droppo, J., Maas, R. (2021) wav2vec-C: A Self-Supervised Model for Speech Representation Learning. Proc. Interspeech 2021, 711-715, doi: 10.21437/Interspeech.2021-717
- 7 Sadhu, S., Hermansky, H. (2021) Radically Old Way of Computing Spectra: Applications in End-to-End ASR. Proc. Interspeech 2021, 1424-1428, doi: 10.21437/Interspeech.2021-643
- 8 Sadhu, S., Hermansky, H. (2020) Continual Learning in Automatic Speech Recognition. Proc. Interspeech 2020, 1246-1250, doi: 10.21437/Interspeech.2020-2962
- 9 Sadhu, Samik, and Hynek Hermansky. "Modulation Vectors as Robust Feature Representation for ASR in Domain Mismatched Conditions." Proc. Interspeech 2019 (2019): 3441-3445.
- 10 Wang, Xiaofei, Jinyi Yang, Ruizhi Li, Samik Sadhu, and Hynek Hermansky. "Exploring Methods for the Automatic Detection of Errors in Manual Transcription." arXiv preprint arXiv:1904.04294 (2019)
- 11 Sadhu, Samik, Ruizhi Li, and Hynek Hermansky. "M-vectors: Sub-band Based Energy Modulation Features for Multi-stream Automatic Speech Recognition." ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.
- 12 Sadhu, Samik, and Prasanta Kumar Ghosh. "Low resource point process models for keyword spotting using unsupervised online learning." 2017 25th European Signal Processing Conference (EUSIPCO). IEEE, 2017.
- 13 Masters Thesis - Sadhu, Samik , "Non-ASR Based Keyword Spotting in Continuous Speech", (**Thesis**)