

K.NAVEEN KUMAR

Address

3077 Los Prados St.,
San Mateo, CA 94403

knaveen87@gmail.com

Cell : (213) 545-1684

<https://algoseer.github.io>

Interests

- Machine learning for speech, multimedia and multimodal processing

Current

- **Sony Interactive Entertainment America** Sept 2016-present
Sr. Software Engineer,
Affective Artificial Intelligence,
PlayStation R&D, San Mateo, USA

Education

- **University of Southern California, Los Angeles, CA** Aug 2009- Aug 2016
Ph.D. in Electrical Engineering
Signal Analysis and Interpretation Laboratory
Thesis: *Exploiting latent reliability information for classification tasks*
GPA: 3.91/4.0
- **Indian Institute of Technology, Kharagpur, India** May 2009
B.Tech. in Electrical Engineering
GPA: 8.5/10.0

Experience at SIE R&D

- **Affective computing**
Emotion recognition from speech and video input using various deep learning (DL) methods. Trained different multimodal architectures for reliable fusion of information from multiple channels. Helped develop cloud-based DL solutions for integrating different components in a realtime system.
- **Text-to-speech synthesis**
Trained traditional clustergen and DL based voices. Explored both parametric and concatenative voices to attain good speed and quality tradeoff.
- **Speech-to-text recognition**
Trained KALDI acoustic models for noisy and far field speech recognition using hybrid DNN and CTC based architectures.
- **Facial animation**
DL models for controlling talking face animation rigs using speech and facial video input. Trained related models for audio visual speech recognition.

Projects at SAIL, USC

- **Multimedia Content Analysis (Google/GDI)**
Methods for automatic analysis of character level representation in movies from video and audio modalities. An important challenge is making standard audio/ video algorithms such as diarization robust on movies.
- **Underwater Acoustics (ONR/NSF)**
Classification and detection of objects in underwater sonar images. Also looked at problems of joint optimization of communication and navigation tradeoffs while performing localization tasks underwater.
- **Development of open source speech toolkit (multiple projects)**
Implemented different denoising algorithms for use within the framework of KALDI speech recognition toolkit. I also experimented with fusion of diverse denoising systems to provide robustness to noise conditions. Currently working on an online diarization algorithm for use with KALDI.

- **Dynamics of vocal tract shaping (NIH)**

Image processing techniques for robust segmentation and extraction of landmark points from realtime MR images of the human vocal tract.

- **Be a Scientist: Informal science education (NSF)**

Informal science learning programs are provided by our NGO partner *Iridescent*. I am working on methods for automatic analysis of instruction quality using classroom video in order to provide feedback to the instructors on their teaching.

Peer-Reviewed Journal Papers

1. Krishna Somandepalli, **Naveen Kumar**, Tanaya Guha, Shrikanth S. Narayanan, “*Unsupervised Discovery of Character Dictionaries in Animation Movies*”, IEEE Transactions on Multimedia.
2. **Naveen Kumar**, Fatemeh Fazel, Milica Stojanovic, Shrikanth S. Narayanan, “*Online rate adjustment for adaptive random access compressed sensing of time-varying fields*”, EURASIP Journal on Advances in Signal Processing.
3. **Naveen Kumar**, Urbashi Mitra, Shrikanth S. Narayanan, “*Robust object classification in underwater sidescan sonar images by using reliability aware fusion of shadow features*”, IEEE Journal of Oceanic Engineering.
4. Jangwon Kim, **Naveen Kumar**, Andreas Tsiartas, Ming Li, Shrikanth S. Narayanan, “*Automatic Intelligibility Classification of sentence-level pathological speech*”, Computer Speech and Language : Special Issue on next generation paralinguistics .

Workshops/ Abstracts

1. **Naveen Kumar**, Shrikanth S. Narayanan, “*Detection of Musical Event Drop from Crowdsourced Annotations using a Noisy Channel Model*”, Proc., MediaEval Workshop, 2014
2. **Naveen Kumar**, Rahul Gupta, Tanaya Guha, Colin Vaz, Maarten V. Segbroeck, Jangwon Kim, Shrikanth S. Narayanan “*Affective Feature Design and Predicting Continuous Affective Dimensions from Music*”, Proc., MediaEval Workshop, 2014
3. Sunav Choudhary, D. Kartik, **Naveen Kumar**, Urbashi Mitra, Shrikanth S. Narayanan, “*Active Target Detection with Navigation Costs : A Randomized Benchmark*”, Proc. Allerton 2014.
4. Vikram Ramanarayanan, **Naveen Kumar**, Shrikanth S. Narayanan, “*A framework for unusual event detection in videos of informal classroom settings*”, Workshop on Personalizing Education with Machine Learning, NIPS 2012.

Peer-Reviewed Conference Papers

1. Krishna Somandepalli, Victor Martinez, **Naveen Kumar**, Shri Narayanan, “*Multimodal representation of Advertisements using Segment-level Autoencoders*”, Accepted at ICMI 2018.
2. Haoqi Li, **Naveen Kumar**, Ruxin Chen, Panayiotis Georgiou, “*A deep reinforcement learning framework for identifying funny scenes in movies*”, Accepted at ICASSP 2018.
3. Ankit Goyal, **Naveen Kumar**, Tanaya Guha, Shrikanth S. Narayanan, “*A multimodal mixture-of-experts model for dynamic emotion prediction in movies*”, Accepted at ICASSP 2016.
4. Adarsh Tadimari, **Naveen Kumar**, Tanaya Guha, Shrikanth S. Narayanan, “*Opening big in Box Office? Trailer content can help*”, Accepted at ICASSP 2016.
5. Rahul Gupta, Theodora Chaspari, Jangwon Kim , **Naveen Kumar**, Danny Bone, Shrikanth S. Narayanan, “*Pathological speech processing: state-of-the-art, current challenges, and future directions*”, Accepted at ICASSP 2016.

6. Tanaya Guha, Chewei Huang, **Naveen Kumar**, Zhu Yan, Shrikanth S. Narayanan, “*Gender Representation in Cinematic Content: A Multimodal Approach*”, Accepted at ICMI 2015.
7. **Naveen Kumar**, Shrikanth S. Narayanan, “*A discriminative reliability-aware classification model with applications to intelligibility classification in pathological speech*”, Accepted at Interspeech 2015.
8. Rahul Gupta, **Naveen Kumar**, Shrikanth S. Narayanan, “*Affect prediction in music using boosted ensemble of filters*”, Accepted at EUSIPCO 2015.
9. Tanaya Guha, **Naveen Kumar**, Shrikanth S. Narayanan, Stacy Smith, “*Computationally deconstructing movie narratives*”, Accepted at ICASSP 2015 .
10. Jangwon Kim, **Naveen Kumar**, Sungbok Lee, Shrikanth S. Narayanan, “*Enhanced airway-tissue boundary segmentation for real-time magnetic resonance imaging data*”, Proc. ISSP 2014.
11. **Naveen Kumar**, Shrikanth S. Narayanan, “*Hull Detection Based on Largest Empty Sector Angle with application to analysis of real-time MR images*”, Accepted at ICASSP 2014.
12. **Naveen Kumar**, Maarten Van Segbroeck, Kartik Audhkhasi, Peter Drotar, Shrikanth S. Narayanan, “*Fusion of diverse denoising systems for Robust Automatic Speech Recognition*”, Accepted at ICASSP 2014.
13. Sunav Choudhary, **Naveen Kumar**, Urbashi Mitra, Shrikanth S. Narayanan, “*Active Target Detection with Mobile Agents*”, Accepted at ICASSP 2014.
14. Jangwon Kim, **Naveen Kumar**, Andreas Tsiartas, Ming Li , Shrikanth S. Narayanan, “*Intelligibility Classification of Pathological Speech using Fusion of Multiple Subsystems*”, Proc. Interspeech 2012.
15. **Naveen Kumar**, Andreas Tsiartas, Shrikanth S. Narayanan, “*Features for comparing tune similarity of songs across different languages*”, Proc. MMSP 2012.
16. **Naveen Kumar**, Qun Feng Tan, Shrikanth S. Narayanan, “*Object Classification in Sidescan Sonar Images with Sparse Representation Techniques*”, Proc. ICASSP 2012.
17. **Naveen Kumar**, Adam Lammert, Brendan Englot, Frank S. Hover, Shrikanth S. Narayanan, “*Directional Descriptors using Zernike Moment Phases for Object Orientation Estimation in Underwater Sonar Images*”, Proc. ICASSP 2011.

Graduate Courses	Probability	Wavelets	Pattern Recognition
	Statistics	Speech Recognition	Machine Learning
	Random Processes	Computer Vision	Graphical Models
	Convex Optimization	Natural Language Processing	

Skills

Programming: Python, Java, C/C++, Bash, MATLAB

Tools: SPSS, LaTeX, OpenFST, Git, SVN

Speech Tools: KALDI, HTK, Eesen, Festival

ML Tools: Keras, Tensorflow, PyTorch, sklearn, weka, scipy

OS: Unix, Windows

- Awards**
- Awarded Merit-Cum-Means scholarships for two successive years at IIT Kharagpur.
 - Awarded InfoUSA Summer Research Fellowship 2008 for an internship at USC.
 - Won Interspeech 2012 Computational Paralinguistic Challenge for Intelligibility Classification of Pathological Speech
 - Northern Digital Inc. Excellence award at ISSP 2014, Cologne, Germany for MRI segmentation paper.

**Professional
Activities**

- To organize a workshop at ICME 2018 on Media Analytics for Societal Trends.
- Reviewed papers for SPCOM, ICASSP, Transactions of Affective Computing and IEEE Journal of Oceanic Engineering
- TPC member for Interspeech, ACII
- Member of IEEE, SPS and ISCA