

# Kartik Audhkhasi

IBM T. J. Watson Research Center  
1101 Kitchawan Road  
Yorktown Heights, NY 10598

(914) 945-1332 (Office), (323) 401-3343 (Residence)  
Email: kaudhkha@us.ibm.com

## Occupation

- **Research Staff Member** (July 2014 -)  
Speech Recognition and Machine Learning, IBM Research AI  
IBM T. J. Watson Research Center, Yorktown Heights, New York, USA

## Education

- **Doctor of Philosophy in Electrical Engineering** (Fall 2008 - Summer 2014)  
Thesis: A computational framework for diversity in ensembles for humans and machine systems  
Advisor: Prof. Shrikanth S. Narayanan  
Signal Analysis and Interpretation Lab (SAIL)  
Electrical Engineering Department  
University of Southern California, Los Angeles, California, USA  
GPA: 4.0/4.0
- **Bachelor of Technology in Electrical Engineering** (Fall 2003 - Spring 2008)  
**Master of Technology in Information and Communication Technology**  
Indian Institute of Technology, Delhi, India  
GPA: 9.9/10.0 (Graduate), 8.7/10.0 (Undergraduate)

## Research Interests

- Neural networks, machine learning, automatic speech recognition, natural language processing, statistical signal processing.

## Peer-Reviewed Journal Papers

1. **K. Audhkhasi**, A. Rosenberg, A. Sethy, B. Ramabhadran, B. Kingsbury, “*End-to-end ASR-free keyword search from speech*”, IEEE Journal of Selected Topics in Signal Processing, Special Issue on End-to-End Speech and Language Processing, vol. 11, no. 8, pp. 1351-1359, 2017.
2. A. Hsu, A. K. Gupta, R. Muller, X. Cui, **K. Audhkhasi**, J. Han, “*Representations of speech signals recorded through a dynamic periphery inspired by horseshoe bat biosonar*”, The Journal of the Acoustical Society of America, vol. 142, no. 4, pp. 2706-2707, 2017.
3. **K. Audhkhasi**, A. Rosenberg, G. Saon, A. Sethy, B. Ramabhadran, S. Chen, M. Picheny, “*Recent Progress in End-to-End Models for Spoken Language Processing*”, IBM Journal of Research and Development, vol. 61, no. 4, 2017.
4. R. Gupta, **K. Audhkhasi**, Z. Jacokes, A. Rozga, and S. Narayanan, “*Modeling multiple time series annotations based on ground truth inference and distortion*”, IEEE Transactions on Affective Computing, 2016.
5. **K. Audhkhasi**, O. Osoba, and B. Kosko, “*Noise-enhanced convolutional neural networks*”, Neural Networks, vol. 78, pp. 15-23, 2016.
6. R. Gupta, **K. Audhkhasi**, S. Lee, and S. Narayanan, “*Detecting paralinguistic events in audio stream using context in features and probabilistic decisions*”, Computer Speech & Language, vol. 36, pp. 72-92, 2016.
7. **K. Audhkhasi**, A. Zavou, P. Georgiou, and S. Narayanan, “*Theoretical analysis of diversity in an ensemble of automatic speech recognition systems*”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 3, March 2014.
8. D. Bone, M. S. Goodwin, M. P. Black, C. Lee, **K. Audhkhasi**, and S. Narayanan, “*Applying machine learning to facilitate autism diagnostics: Pitfalls and promises*”, Journal of Autism and Developmental Disorders, 2014.
9. **K. Audhkhasi** and S. Narayanan, “*A globally-variant locally-constant model for fusion of labels from multiple diverse experts without using reference labels*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 4, pp. 769-783, April 2013.
10. **K. Audhkhasi** and A. Kumar, “*Two-scale auditory feature based non-intrusive speech quality evaluation*, IETE Journal of Research, vol. 56, no. 2, pp. 111-118, April 2010.

11. **K. Audhkhasi**, “Automatic evaluation of fluency in spoken language”, IETE Technical Review, vol. 26, no. 2, pp. 108-114, April 2009.

**Peer-Reviewed  
Conference Papers**

1. **K. Audhkhasi**, B. Ramabhadran, G. Saon, M. Picheny, and D. Nahamoo, “Direct Acoustics-to-Word Models for English Conversational Speech Recognition”, Interspeech, 2017.
2. G. Saon, G. Kurata, T. Sercu, **K. Audhkhasi**, S. Thomas, D. Dimitriadis, X. Cui, B. Ramabhadran, M. Picheny, L. Lim, B. Roomi, P. Hall, “English Conversational Telephone Speech Recognition by Humans and Machines”, Interspeech, 2017.
3. **K. Audhkhasi**, A. Rosenberg, A. Sethy, B. Ramabhadran, and B. Kingsbury, “End-to-end ASR-free keyword search from speech”, Proc. ICASSP, 2017.
4. A. Rosenberg, **K. Audhkhasi**, A. Sethy, B. Ramabhadran, M. Picheny, “End-to-end speech recognition and keywords search on low-resource languages”, Proc. ICASSP, 2017.
5. J. Cui, B. Kingsbury, B. Ramabhadran, G. Saon, T. Sercu, **K. Audhkhasi**, A. Sethy, M. Nussbaum-Thom, A. Rosenberg, “Knowledge distillation across ensembles of multilingual models for low-resource languages”, Proc. ICASSP, 2017.
6. D. Serdyuk, **K. Audhkhasi**, P. Brakel, B. Ramabhadran, S. Thomas, Y. Bengio, “Invariant representations for noisy speech recognition”, Proc. NIPS Workshop on End-to-End Speech Recognition, 2016.
7. S. Thomas, **K. Audhkhasi**, J. Cui, B. Kingsbury, and B. Ramabhadran, “Multilingual data selection for low resource speech recognition”, Proc. Interspeech, 2016.
8. **K. Audhkhasi**, A. Sethy, and B. Ramabhadran, “Semantic Word Embedding Neural Network Language Models for Automatic Speech Recognition”, Proc. ICASSP, 2016.
9. J. Chen, L. Wu, **K. Audhkhasi**, B. Kingsbury, and B. Ramabhadran, “Efficient One-Vs-One Kernel Ridge Regression for Speech Recognition”, Proc. ICASSP, 2016.
10. **K. Audhkhasi**, A. Sethy, and B. Ramabhadran, “Diverse embedding neural network language models”, Proc. ICLR, 2015.
11. R. Gupta, **K. Audhkhasi**, and S. Narayanan, “A mixture of experts approach towards intelligibility classification of pathological speech”, Proc. ICASSP, 2015.
12. **K. Audhkhasi**, O. Osoba, and B. Kosko, “Noise benefits in convolutional neural networks”, Proc. Intl. Conf. on Advances in Big Data Analytics, 2014.
13. **K. Audhkhasi**, A. Sethy, B. Ramabhadran, and S. Narayanan, “Semi-supervised term-weighted value rescoring for keyword search”, Proc. ICASSP, 2014.
14. R. Gupta, **K. Audhkhasi**, and S. Narayanan, “Training ensemble of diverse classifiers on feature subsets”, Proc. ICASSP, 2014.
15. N. Kumar, M. V. Segbroeck, **K. Audhkhasi**, P. Drotar, and S. Narayanan, “Fusion of diverse denoising systems for robust automatic speech recognition”, Proc. ICASSP, 2014.
16. **K. Audhkhasi**, A. Zavou, P. Georgiou, and S. Narayanan, “Empirical link between hypothesis diversity and fusion performance in an ensemble of automatic speech recognition systems”, Proc. Interspeech, 2013.
17. **K. Audhkhasi**, O. Osoba, and B. Kosko, “Noise benefits in backpropagation and bidirectional pre-training”, Proc. IJCNN, 2013.
18. **K. Audhkhasi**, O. Osoba, and B. Kosko, “Noisy hidden Markov models for speech recognition”, Proc. IJCNN, 2013.
19. R. Gupta, **K. Audhkhasi**, S. Lee, and S. Narayanan, “Speech paralinguistic event detection using probabilistic time-series smoothing and masking”, Proc. Interspeech (Computational Paralinguistics Challenge), 2013.
20. A. Sethy, S. Chen, E. Arisoy, B. Ramabhadran, **K. Audhkhasi**, S. Narayanan, P. Vozila, “Joint training of interpolated exponential n-gram models”, Proc. ASRU, 2013.
21. D. Bone, T. Chaspari, **K. Audhkhasi**, J. Gibson, A. Tsiartas, S. Lee, and S. Narayanan, “Classifying language-related developmental disorders from speech cues: the promise and the potential confounds”, Proc. Interspeech (Computational Paralinguistics Challenge), 2013.

22. F. Morbini, **K. Audhkhasi**, K. Sagae, R. Artstein, D. Can, P. Georgiou, S. Narayanan, A. Leuski, and D. Traum, “Which ASR should I chose for my dialogue system?”, Proc. SIGDIAL, 2013.
23. **K. Audhkhasi**, A. Sethy, B. Ramabhadran, and S. Narayanan, “Creating ensemble of diverse maximum entropy models”, Proc. ICASSP, 2012.
24. **K. Audhkhasi**, P. Georgiou, and S. Narayanan, “Analyzing quality of crowd-sourced speech transcriptions of noisy audio for acoustic model adaptation”, Proc. ICASSP, 2012.
25. **K. Audhkhasi**, A. Metallinou, M. Li, and S. Narayanan, “Speaker personality classification using systems based on acoustic-lexical cues and an optimal tree-structured Bayesian network”, Proc. Interspeech (Speaker Trait Challenge), 2012.
26. F. Morbini, **K. Audhkhasi**, R. Artstein, M. V. Segbroeck, K. Sagae, P. Georgiou, D. Traum, and S. Narayanan, “A reranking approach for recognition and classification of speech input in conversational dialogue systems”, Proc. SLT, 2012.
27. **K. Audhkhasi**, P. Georgiou, and S. Narayanan, “Reliability-weighted acoustic model adaptation using crowd sourced transcripts”, Proc. Interspeech, 2011.
28. **K. Audhkhasi** and S. Narayanan, “Emotion classification from speech using evaluator reliability-weighted combination of ranked lists”, Proc. ICASSP, 2011.
29. **K. Audhkhasi**, P. Georgiou, and S. Narayanan, “Accurate transcription of broadcast news speech using multiple noisy transcribers and unsupervised reliability metrics”, Proc. ICASSP, 2011.
30. **K. Audhkhasi** and S. Narayanan, “Data-dependent evaluator modeling and its application to emotional valence classification from speech”, Proc. Interspeech, 2010.
31. Q. F. Tan, **K. Audhkhasi**, P. Georgiou, E. Ettelaie, and S. Narayanan, “Automatic speech recognition system channel modeling”, Proc. Interspeech, 2010.
32. **K. Audhkhasi**, P. Georgiou, and S. Narayanan, “Lattice-based lexical cues for word fragment detection in conversational speech”, Proc. ASRU, 2009.
33. **K. Audhkhasi**, K. Kandhway, O. Deshmukh, and A. Verma, “Formant-based technique for automatic filled-pause detection in spontaneous spoken English”, Proc. ICASSP, 2009.
34. O. Deshmukh, K. Kandhway, A. Verma, and **K. Audhkhasi**, “Automatic evaluation of spoken English fluency”, Proc. ICASSP, 2009.
35. **K. Audhkhasi** and A. Verma, “Keyword search using modified minimum edit distance measure”, Proc. ICASSP, 2007.

#### Issued Patents

1. **K. Audhkhasi**, O. Deshmukh, K. Kandhway and A. Verma, “Automatic evaluation of spoken fluency”, US Patent No. 8,457,967, Issued April 6, 2013, Assigned to Nuance Communications, Inc.

#### Teaching Assistance

- USC - EE 562a: Random Processes in Engineering (Fall 2011, Spring 2011, Fall 2010)
- IIT-Delhi - EEP 306: Communication Engineering Lab (Fall 2007), EEL 316: Digital Communications (lab sessions, Spring 2008)

#### Internships

- **IBM T. J. Watson Research Center, Yorktown Heights, USA** (May-August 2011)  
**Summer intern, Human Language Technologies group**  
Developed algorithm for creating diverse ensemble of maximum entropy classifiers.
- **IBM India Research Lab, Delhi, India** (May-July 2008)  
**Summer intern, Speech and Business Intelligence group**  
Worked on filled pause detection using formant trajectories, and automatic fluency evaluation in spontaneous spoken English.
- **IBM India Research Lab, Delhi, India** (May-July 2006)  
**Summer intern, Speech and Business Intelligence group**  
Worked on keyword spotting in spontaneous spoken English using phone lattices.

## Academic Awards and Honors

- Honorable Mention for Best 4-min. Ph.D. Pitch at EE Research Festival, USC (2013)
- Won Interspeech-2013 Computational Paralinguistics Challenge on Social Signals (team led by R. Gupta of SAIL, USC)
- USC EE Best Paper Award (2013)
- USC EE Ming Hsieh Institute PhD Scholar (2012-2013)
- IBM PhD Fellow (2012-2013)
- Best Teaching Assistant in the EE Department, USC (2012)
- Honorable Mention for Best Teaching Assistant in the EE Department, USC (2011)
- IETE M. N. Saha Memorial Award for Best Application Oriented Paper (2010)
- Annenberg Fellow, USC (2008-2012)
- Ranked 1<sup>st</sup> in the EE dual degree 2003 batch at IIT-Delhi at the end of 5 years
- Three Merit Prizes for Academic Excellence, IIT-Delhi
- Ministry of Human Resource Development (Govt. of India) Scholarship (2007-2008)
- Best Undergraduate Summer Intern Project Award, IBM India Research Lab (2006)
- Summer Undergraduate Research Award, IIT-Delhi (2005)
- Ranked 409 (top 0.2%) out of approximately 200,000 candidates from all over India in IIT Joint Entrance Examination (2003)
- Ranked 54 in the Combined Entrance Examination (2003) for the Delhi College of Engg.

## Professional Activities

- Area chair (cross-cutting) for Interspeech 2016.
- Session chair for ICASSP (2016, 2017) and Interspeech 2016.
- Reviewer for ICASSP (2009, 2010, 2013, 2014, 2016, 2017, 2018), Interspeech (2009, 2010, 2011, 2016, 2017), MMSP 2009, SPCOM (2012, 2014), IJCNN (2015), Elsevier Journal of Biomedical Informatics, IEEE Trans. on Communications, IEEE Trans. on Pattern Analysis and Machine Intelligence, Advances in Artificial Intelligence, Fluctuation and Noise Letters.
- Professional Member of IEEE and ISCA.

## Computer Skills

- **Languages and packages:** C/C++, Bash, Python, Lua, Perl, Matlab.
- **Speech/language processing toolkits:** Attila, Kaldi, Sphinx, HTK, OpenFST, Moses.
- **Machine learning toolkits:** Pytorch, Torch, Theano, Keras, Weka.
- **Operating systems:** Linux, Windows, OS X.

## Miscellaneous

- Citizen of India on H1b visa.