Meta Research
Redmond, WA - 98052

Email: anuragkr@ieee.org
Web: anuragkr90.github.io

Anurag Kumar

Current Position

Staff AI Research Lead and Scientist at Meta Research
Research Lead and Scientist working on multimodal and multi-sensory scene understanding and generative AI research.

Research Interests

Artificial Intelligence, Machine Learning, Audio and Speech Signal Processing, Multimodal Learning

Education

Ph.D and M.S, Carnegie Mellon University — Aug. 2013 - Dec. 2018
Language Technologies Institute, School of Computer Science
Dissertation: *Acoustic Intelligence in Machines*
Advisor: Prof. Bhiksha Raj

Bachelors, Indian Institute of Technology Kanpur — 2008 - 2013
BTech - MTech Integrated Dual Degree in Electrical Engineering

Research Experiences

Research Scientist/Lead at Meta Research — Dec 2018 - present

Graduate Research Assistant at Carnegie Mellon University — Aug. 2013 - Sep 2018
Machine Learning and Signal Processing Group
Advisor: Prof. Bhiksha Raj
Dissertation Research: Acoustic Intelligence in Machines

Research Intern at Facebook Research — May 2017 - Aug. 2017
AI Speech Group.
Mentors: Christian Fuegen and Maksim Khadkevich
Research Topic: Large Scale Video Understanding Using Audio Content Analysis

Research Intern at Microsoft Research — July 2015 - Oct. 2015
Multimedia, Interaction, and Communication Group.
Mentor: Dinei Florencio
Research Topic: Speech Enhancement using Deep Learning

Undergrad. Researcher at Indian Institute of Technology Kanpur — 2012 - Jun 2013
Multimodal Information Processing Systems (MiPS) Lab
Advisor: Prof. Rajesh Hegde
Research Topic: Audio Event Recognition and Detection

Visiting Undergraduate Researcher at Carnegie Mellon — May 2011 - July 2011
Machine Learning and Signal Processing Group
Advisor: Prof. Bhiksha Raj and Prof. Richard Stern
Research Topic: Audio Events, Source Separation, Music Information

Publications

Google Scholar for most up to date and complete list
Selected Recent Publications

1. Real Acoustic Fields: An Audio-Visual Room Acoustics Dataset and Benchmark
   Ziyang Chen, Israel D Gebru, Christian Richardt, **Anurag Kumar**, William Laney, Andrew Owens, Alexander Richard
   IEEE/CVF Conference on Computer Vision and Pattern Recognition **(CVPR)**, 2024.
2. A Closer Look at Wav2vec2 Embeddings for On-Device Single-Channel Speech Enhancement
   Ravi Shankar, Ke Tan, Buye Xu, **Anurag Kumar**
   IEEE International Conference on Acoustics, Speech and Signal Processing **(ICASSP)**, 2024.
3. AV-NeRF: Learning Neural Fields for Real-World Audio-Visual Scene Synthesis
   Susan Liang, Chao Huang, Yapeng Tian, **Anurag Kumar**, Chenliang Xu
   Neural Information Processing Systems (**Neurips**), 2023.
4. Egocentric Audio-Visual Object Localization
   Chao Huang, Yapeng Tian, **Anurag Kumar**, Chenliang Xu
   IEEE/CVF Conference on Computer Vision and Pattern Recognition **(CVPR)**, 2023.
5. TorchAudio-Squim: Reference-less Speech Quality and Intelligibility measures in TorchAudio
   Anurag Kumar, Ke Tan, Zhaoheng Ni, Pranay Manocha, Xiaohui Zhang, Ethan Henderson, Buye Xu
   IEEE International Conference on Acoustics, Speech and Signal Processing **(ICASSP)**, 2023.
6. RemixIT: Continual self-training of speech enhancement models via bootstrapped remixing
   Efthymios Tzinis, Yossi Adi, Vamsi Krishna Ithapu, Buye Xu, Paris Smaragdis, **Anurag Kumar**
   IEEE Journal of Selected Topics in Signal Processing **(IEEE JSTSP)**, 2022.
7. Speech Quality Assessment through MOS using Non-Matching References
   Pranay Manocha, **Anurag Kumar**
   Interspeech, 2022.
8. Ego4d: Around the world in 3,000 hours of egocentric video
   Kristen Grauman, Andrew Westbury, · · ·, **Anurag Kumar**, · · ·, Antonio Torralba, Lorenzo Torresani, Mingfei Yan, Jitendra Malik
   IEEE/CVF Conference on Computer Vision and Pattern Recognition **(CVPR)**, 2022.
9. NORESQA: A Framework for Speech Quality Assessment using Non-Matching References
   Pranay Manocha, Buye Xu, **Anurag Kumar**
   Neural Information Processing Systems (**Neurips**), 2021.

10. A Sequential Self Teaching Approach for Improving Generalization in Sound Event Recognition.
Anurag Kumar, Vamsi Krishna Ithapu
37^{th} International Conference on Machine Learning **(ICML)**, 2020

11. Large Scale Audiovisual Learning of Sounds with Weakly Labeled Data.
Haytham Fayek*, **Anurag Kumar*** (* Equal Contribution)
29^{th} International Joint Conference on Artificial Intelligence **(IJCAI)**, 2020

12. Learning Sound Events from Webly Labeled Data
Anurag Kumar , Ankit Shah , Alexander Hauptmann and Bhiksha Raj
28^{th} International Joint Conference on Artificial Intelligence **(IJCAI)**, 2019

Patents

1. Smart Glass Interface for Impaired Users or Users With Disabilities.
Johana Escudero, Scott Selfon, **Anurag Kumar**, Jonathan Lee, Brett Lavalla, Simon Porter, Gregory Sarkis-Kelly, Christi Miller, Yao Ding
US Patent, US-18/163436, 2023.

2. Low Power System for Acoustic Event Detection.
Gael Le Lan, Varun Nagaraja, Ivaylo Enchev, Ivaylo Enchev, **Anurag Kumar**, Sangeeta Srivastava, Yangyang Shi, Ming Sun, Julian Chan, Yun Wang, Xin Lei, Vikas Chandra, Vamsi Krishna Ithapu, Mike Seltzer. US Patent, 2023.

3. Sign Language Detection for Virtual Reality and Augmented Reality Headsets.
Johana Escudero, Scott Selfon, **Anurag Kumar**, Jonathan Lee, Brett Lavalla, Simon Porter, Gregory Sarkis-Kelly, Christi Miller, Yao Ding
US Patent, US-63/348907, 2023.

4. Sign Language Detection for Smart Glasses.
Johana Escudero, Scott Selfon, **Anurag Kumar**, Jonathan Lee, Brett Lavalla, Simon Porter, Gregory Sarkis-Kelly, Christi Miller, Yao Ding.
US Patent, US-18/163420, 2022

5. Audio system for artificial reality applications.
Ashutosh Pandey, Buye Xu, **Anurag Kumar**, Jacob Donley, Paul Calamia, Deliang Wang, Chuming Zhao, Thomas Lunner, Antonio Miller, Neto
US Patent, US-17/714,638, 2022

6. Speech Quality Assessment using Non-Matching References.
Anurag Kumar, Pranay Manocha, Buye Xu
US Patent, US-63/331,502, 2022

7. Neural Radiance Field Systems and Methods for Synthesis Of Audio-Visual Scenes
Susan Liang, Chao Wang, Yapeng Tian, **Anurag Kumar**, Chenliang Xu
US Patent, US-63443258, 2023.

Awards & Honors

- Appointed to IEEE Audio and Acoustic Signal Processing Technical Committee, 2023
- Best Paper Award Finalist at **CVPR 2022**.

- Recognition by AMiner as one of the **most influential scholar** in the area of Multimedia between 2012 and 2021.
- Top Reviewer Recognition at **ICML 2020**
- Finalist **Qualcomm Innovation Fellowship**, 2017
- **Research Fellowship** at Carnegie Mellon University (2013-2018)
- **Gandhian Young Technological Innovation Awards**, 2017
  National Level Award in India for Scientific Innovation
  Mentored and Advised Incoming CMU Graduate Students on Acoustic Intelligence
- **IEEE Signal Processing Society Travel Grant** for ICASSP 2015
- **Best Paper Award** Nomination at National Conference on Communication (NCC), 2014
- **EURASIP Travel Grant** for European Signal Processing Conference, 2013
- **Samsung Innovations Awards 2012**
  For developing Audio Based Event and Context Recognition System
- **Best Speech Technology** project at Carnegie Mellon University Winter School-2010 held at IIIT Hyderabad, India
- **Accepted for fellowship** from Consultancy Development Center Department of Science and Industrial Research, Govt. of India, 2010
- Among Top **0.1%** in IIT Joint Entrance Examination-2008
  Taken by more than 0.5 million students

Recent Talks /Major OSS

- Oct 2023: Torchaudio-Squim Released in PyTorch's Torchaudio Library
- March 2023: Invited Talk in Voice of XR Series at University of Rochester
- March 2023: Invited Talk in ECE and LTI at Carnegie Mellon University
- Sep 2022: **Tutorial** on Learning from Weak Labels at Interspeech 2022.

Professional Activities

- **Journal Editor**

  Associate Editor, IEEE Signal Processing Letters (2024 -)

  Guest Editor: Frontiers in Artificial Intelligence: Special Issue on Audio Event Detection and Recognition.

- **IEEE Audio and Acoustic Signal Processing Technical Committee** (2023 - 2026)
- **Organizer**

  URGENT Challenge on Universal Speech Enhancement, 2024.

  Ego4D Audio-Visual Diarization Benchmark and Challenge, 2022.

  Special Session on On-device AI for Audio and Speech Applications at ICASSP 2021.

- **Program Committee Member/Reviewer (Conferences)**: Regularly for all major AI and Speech Conferences (List Not exhaustive)
  International Conference on Machine Learning (**ICML**), Neural Information Processing Systems (**Neurips**), International Conference on Learning Representations (**ICLR**), AAAI Conference on Artificial Intelligence (**AAAI**), IEEE International Conference on Audio, Speech, and Signal Processing (**IEEE ICASSP**), **Interspeech**.
- **Reviewer (Journals)**: Regularly for several major AI, Audio and Speech Journals (List Not exhaustive)
  Transactions on Machine Learning Research (**TMLR**), IEEE Transactions on Audio Speech and Language Processing (**IEEE TASLP**), IEEE Transactions on Signal Processing (**IEEE TSP**), IEEE Transactions on Multimedia (**IEEE TMM**), Neural Networks, IEEE Signal Processing Letters (**IEEE SPL**), IEEE Transactions on Emerging Topics in Computational Intelligence (**IEEE TETCI**), EURASIP Journal on Audio, Speech, and Music Processing (**EURASIP JASMP**)

Research Advising

PhD Students Co-Advising

- (2021-) Chao Huang, PhD Student at University of Rochester (UoR)
- (2022-) Susan Liang, PhD Student at University of Rochester (UoR)
- (2021-) Joseph Konan, PhD Student at Carnegie Mellon University (CMU)
- (2022-) Muqiao Yang, PhD Student at Carnegie Mellon University (CMU)

PhD Thesis Committee Member

- Muqiao Yang, Carnegie Mellon University
- Chao Huang, University of Rochester
- Ankit Shah, Carnegie Mellon University
- Ravi Shankar, Johns Hopkins University (Graduated 2023)

Interns Advised

- 2023 - Joanna Hong, PhD Student at Korea Advanced Institute of Science and Technology (KAIST)
- 2023 - Jaesung Huh, PhD Student at University of Oxford
- 2022 - Ravi Shankar, PhD Student at Johns Hopkins University (JHU)
- 2022 - Haibin Wu, PhD Student at National Taiwan University
- 2021 - Efthymios Tzinis, PhD Student at Uni. of Urbana-Champaign (UIUC)
- 2021 - Pranay Manocha, PhD Student at Princeton University
- 2020 - Ke Tan, PhD Student at Ohio State University
- 2020 - Raymond Xia, PhD Student Carnegie Mellon University (CMU)
- 2019 - Panagiotis Tzirakis, PhD Student at Imperial College London

Undergraduate and Masters Students Mentored and Advised

- David Bick, MS student at LTI, CMU
- Ankit Shah, MS student at LTI, CMU.
- Rohan Badlani, Undergraduate student at BITS, Pilani.
- Pranay Manocha, Undergraduate student at IIT Guwahati.