

# Kai Zhen

700 N. Woodlawn Ave. Luddy Hall, Bloomington, Indiana, 47404.

<http://kaizhen.us> • zhenk@iu.edu

---

## POSITIONS HELD

### Indiana University

- Research Assistant Spring 2018 --- present
  - Project: efficient end-to-end neural audio coding system
- Teaching Assistant Fall 2015 --- Fall 2017
  - Department of Computer Science
  - Intelligent Systems Engineering Department

### Amazon.com, Inc.

- Applied Scientist Intern Summer 2020
  - Alexa Edge ML team, Pittsburgh, PA
    - Supervisors: Athanasios Mouchtaris, Hieu Duy Nguyen, Feng-Ju (Claire) Chang
    - Project: network compression for on-device ASR solutions
    - Amongst "Top Rated Posters" to be featured at Amazon.Science

### LinkedIn Corporation

- Machine Learning & Relevance Intern Summer 2019
    - Ads-AI Group, Mountain View, CA
      - Supervisors: Sara Smoot, Lijun Peng, Hiroto Udagawa
      - Project: ads response rate prediction in wide-n-deep estimators and BERT
    - Standardization Group, New York City, NY Summer 2018
      - Supervisors: Xiaoqiang Luo, Deirdre Hogan
      - Project: relevance ranking for resume builder with deep neural networks
- 

## EDUCATION

### Ph.D. in Computer Science & Cognitive Science, Double Major (GPA 3.95/4.0)

- Indiana University, Bloomington, United States
- Committee: Minje Kim (advisor), Robert Goldstone, Donald Williamson, Yi Shen
- Dissertation topics: Low-Power Neural Audio Coding, Psychoacoustics

### M.S. in Computer Science (GPA 91.6/100)

2015

- Tsinghua University, Beijing, China

### B.S. in Software Engineering (GPA 91.8/100, Graduated with Honors)

2012

- Xidian University, Xi'an, China
- 

## PROJECT & PUBLICATION

### In Progress

[S002] Kai Zhen, Mi Suk Lee, Jongmo Sung, and Minje Kim, "[Psychoacoustic Calibration of Loss Functions for Efficient End-to-End Neural Audio Coding](#)," submitted to IEEE Signal Processing Letters.

[S001] Kai Zhen and Minje Kim, "Blockwise End-To-End Neural Engine for Efficient And Scalable Speech Coding".

### Peer Reviewed Conference Proceedings

[C003] Kai Zhen, Mi Suk Lee, Jongmo Sung, Seungkwon Beack, and Minje Kim, "[Efficient And Scalable Neural Residual Waveform Coding with Collaborative Quantization](#)," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Barcelona, Spain, May 4-8, 2020.

[C002] Kai Zhen, Mi Suk Lee, Minje Kim. "[A Dual-Stage Context Aggregation Method towards Efficient End-To-End Speech Enhancement](#)," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Barcelona, Spain, May 4-8, 2020.

- [C001] **Kai Zhen**, Jongmo Sung, Mi Suk Lee, Seungkwon Beack, and Minje Kim, "[Cascaded Cross-Module Residual Learning towards Lightweight End-to-End Speech Coding](#)," In Proc. *Annual Conference of the International Speech Communication Association (Interspeech)*, Graz, Austria, September 15-19, 2019.

### Peer Reviewed Workshops & Forums

- [W004] **Kai Zhen**, Hieu Duy Nguyen, Feng-Ju (Claire) Chang, Athanasios Mouchtaris. Network Sparsification for On-Device ASR. *Amazon Machine Learning Conference (AMLC) Workshop on Network Inference Optimization*, 2020.
- [W003] **Kai Zhen**, Aswin Sivaraman, Jongmo Sung, Minje Kim. [On Psychoacoustically Weighted Cost Functions Towards Resource-efficient Deep Neural Networks for Speech Denoising](#). *The 7th Annual Midwest Cognitive Science Conference*, 2018.
- [W002] Peter Miksza, Kevin Watson, **Kai Zhen**, Sanna Wager, Minje Kim. Relationships between experts' subjective ratings of jazz improvisations and computational measures of melodic entropy. *The Improvising Brain III: Cultural Variation and Analytical Techniques Symposium*, Atlanta, GA, in Feb, 2017.
- [W001] **Kai Zhen** and David Crandall. [Finding egocentric image topics through convolutional neural network based representations](#) (extended abstract). In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshop on Egocentric Computer Vision*, 2016.

### Patents

- [P004] Mi Suk Lee, Jongmo Sung, Minje Kim, **Kai Zhen**., "[Audio signal encoding method and audio signal decoding method, and encoder and decoder performing the same](#)," U.S. Patent Application No. 16/543,095
- [P003] Minje Kim, Aswin Sivaraman, **Kai Zhen**, Jongmo Sung, et al, "[Audio signal encoding method and apparatus and audio signal decoding method and apparatus using psychoacoustic-based weighted error function](#)", *US Patent Application*, US 2019 / 0164052 A1.
- [P002] Minje Kim, **Kai Zhen**, Mi Suk Lee, et al, "Apparatus and Method for Speech Processing Using a Densely Connected Hybrid Neural Network," *US Patent Application* (pending), 2019
- [P001] Minje Kim, **Kai Zhen**, Mi Suk Lee, "Scalable and Efficient Neural Waveform Coding with Collaborative Quantization," *US Patent Application* (pending), 2019

---

## PROFESSIONAL ACTIVITIES

### Conference Reviewer

- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) - 2019, 2020
- IEEE International Conference on Data Mining (ICDM), 2020
- Association for the Advancement of Artificial Intelligence (AAAI) - 2017, 2018

### Journal Reviewer

- European Association for Signal Processing (EURASIP) Journal on Audio, Speech, and Music Processing

---

## INVITED TALKS

- [T002] IU Hearing Sciences Seminar, March, 2019
- [T001] IU Grey Matters, Graduate and Post-doc Colloquium, March, 2019

---

## CRAFTSMANSHIP

### Deep Learning/Artificial Intelligence (over 4 years experience)

- TensorFlow, PyTorch, etc;
- recommendation, feature learning, autoregressive modeling, recognition, etc

### Audio Signal Processing (over 4 years experience)

- bitrate efficient and scalable audio/speech coding, speech enhancement;
- subjective/objective audio quality assessment;
- psychoacoustic models and optimization skills;
- end-to-end speech recognition (RNN-Transducer).

**Machine Learning (over 5 years experience)**

- regression (GLMix) and classification (decision trees, SVM);
- dimension reduction (PCA/ICA/NMF/ISOMAP);
- clustering analysis (k-means, GMM);
- topic modeling (LDA).

**Big Data Processing (acquired from 2 summer internships)**

- Hadoop, HDFS, Spark (PySpark).
- 

**TEACHING & TUTORING****Graduate Level**

- "Machine Learning for Signal Processing" (ENGR-E 599, ISE IU), Fall 2017
- "Elements of Artificial Intelligence" (CSCI-B 551, CS IU), Fall 2016
- "Computer Vision" (CSCI-B 657, CS IU), Spring 2016
- "Data Structures" (ENGR-E 599, ISE IU), Fall 2015

**Undergraduate Level**

- "Introduction of Artificial Intelligence" (CSCI-B 351, CS IU), Spring 2017