Kai Zhen

700 N. Woodlawn Ave. Luddy Hall, Bloomington, Indiana, 47404.
812-606-6665 • www.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

LinkedIn Corporation

- Machine Learning & Relevance Intern
  - Ads-AI Group, Mountain View, CA — Summer 2019
    - 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 (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 Submission

[S001] **Kai Zhen**, Mi Suk Lee, Jongmo Sung, Seungkwon Beack, and Minje Kim, "Efficient And Scalable Neural Residual Waveform Coding with Collaborative Quantization" (submitted to ICASSP 2020).

[S002] **Kai Zhen**, Mi Suk Lee, Minje Kim. "A Dual-Staged Context Aggregation Method towards Efficient End-To-End Speech Enhancement" (submitted to ICASSP 2020).

Peer Reviewed Conference Proceedings

[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

[W001] **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.

[W003] **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

[P001] 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

[P003] Minje Kim, **Kai Zhen**, Jongmo Sung, Mi Suk Lee, Seungkwon Beack, et al, "Method and Apparatus of Cascaded Residual Learning Pipeline for Audio Coding," *US Patent Application* (pending), 2019

PROFESSIONAL ACTIVITIES

Conference Reviewer

- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) - 2019, 2020, *reviewer*
- Association for Advances in Artificial Intelligence (AAAI) - 2017, 2018, *sub-reviewer*

Journal Reviewer

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

CRAFTSMANSHIP

Deep Learning/Artificial Intelligence (>2 years experience)

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

Audio Signal Processing (>2 years experience)

- bitrate efficient and scalable audio/speech coding, speech enhancement;
- subjective/objective audio quality assessment;
- psychoacoustic models and optimization skills.

Machine Learning (>2 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).