

Xuankai Chang

☎ 667-229-7626 | ✉ Email | 🏠 HomePage

EDUCATION

Carnegie Mellon University

Ph.D student at Language Technologies Institute. Advised by Shinji Watanabe

Pittsburgh, PA, U.S.

Jan. 2021 – Present

Johns Hopkins University

Ph.D student in Electronic and Computer Engineering. Advised by Shinji Watanabe

Baltimore, MD, U.S.

Aug. 2019 – Dec. 2020

Shanghai Jiao Tong University

Master in Computer Science. Advised by Yanmin Qian

Bachelor in Computer Science.

Shanghai, China

Sep. 2016 – Mar. 2019

Sep. 2012 – Jun. 2016

RESEARCH INTERESTS

Speech Recognition, Multi-speaker Speech Recognition, Speech Separation, Unsupervised Representation Learning.

EXPERIENCE

Research Intern (Multi-Speaker ASR)

Mitsubishi Electric Research Laboratories

- Mentor: *Takaaki Hori*
- End-to-end multi-speaker ASR.

May 2021 – Aug. 2021

Cambridge, MA, U.S.

Research Intern (End-to-End Speaker-attributed ASR)

Microsoft

- Mentor: *Naoyuki Kanda, Takuya Yoshioka*
- End-to-end speaker-attributed ASR model on long-form multi-speaker speech.
- The work was summarized as a paper, accepted in ICASSP 2021.

May 2020 – Aug. 2020

Redmond, WA, U.S.

Visiting Graduate Student (Multi-channel Multi-speaker ASR)

Remote collaboration with Johns Hopkins University

- Mentor: *Shinji Watanabe*
- Built an end-to-end multi-channel multi-speaker ASR model.
- Introduced a multi-speaker neural beamforming into End-to-End ASR models.
- The work was presented in the ASRU 2019 and won the best paper award.

Feb. 2019 – Jun. 2019

Baltimore, MD, U.S.

Visiting Graduate Student (End-to-End Multi-speaker ASR)

CLSP at Johns Hopkins University

- Mentor: *Shinji Watanabe*
- Built a state-of-the-art monaural multi-speaker end-to-end ASR model training from random initialization
- Introduced speaker-parallel-attention for multi-speaker E2E ASR.
- The work was presented in the ICASSP 2019.

Sep. 2018 – Dec. 2018

Baltimore, MD, U.S.

Intern (Low-Resource Keyword Spotting)

AISpeech co.

- Mentor: *Kai Yu*
- Built a CTC-based keyword spotting system with unrestricted vocabulary.
- The work was presented in the ICASSP 2016.

Dec. 2015 – May. 2016

Suzhou, China

Journal Papers

- Wangyou Zhang, **Xuankai Chang**, Yanmin Qian, Shinji Watanabe, “Improving End-to-End Single-Channel Multi-Talker Speech Recognition,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1385-1394.
- Yanmin Qian, Chao Weng, **Xuankai Chang**, Shuai Wang, Dong Yu, “Past review, current progress and challenges ahead on cocktail party problem,” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 40-63, Jan 2018
- Yanmin Qian, **Xuankai Chang**, Dong Yu, “Single-channel multi-talker speech recognition with permutation invariant training,” *Speech Communication*, vol. 104, pp. 1-11, Nov 2018

Conference Papers

- Xuankai Chang**, Naoyuki Kanda, Yashesh Gaur, Xiaofei Wang, Zhong Meng, and Takuya Yoshioka. ”Hypothesis Stitcher for End-to-End Speaker-attributed ASR on Long-form Multi-talker Recordings.” in *Proc. IEEE ICASSP*, 2021
- Pengcheng Guo, Florian Boyer, **Xuankai Chang**, Tomoki Hayashi, Yosuke Higuchi, et al. “Recent Developments on ESPnet Toolkit Boosted by Conformer.” in *Proc. IEEE ICASSP*, 2021, pp. 5874–5878
- Jiatong Shi, Jonathan D Amith, **Xuankai Chang**, Siddharth Dalmia, Brian Yan, and Shinji Watanabe, “Highland Puebla Nahuatl Speech Translation Corpus for Endangered Language Documentation” in *Proc. the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*, 2021, pp. 53–63
- Chenda Li, Jing Shi, Wangyou Zhang, Aswin Shanmugam Subramanian, **Xuankai Chang**, Naoyuki Kamo, Moto Hira, Tomoki Hayashi, Christoph Boeddeker, Zhuo Chen, and others, “ESPnet-SE: end-to-end speech enhancement and separation toolkit designed for ASR integration” in *Proc. IEEE SLT*, 2021, pp. 785–792
- Naoyuki Kanda, **Xuankai Chang**, Yashesh Gaur, Xiaofei Wang, Zhong Meng, Zhuo Chen, Takuya Yoshioka, “Investigation of End-To-End Speaker-Attributed ASR for Continuous Multi-Talker Recordings,” in *Proc. IEEE SLT*, 2021, pp. 809-816
- Jing Shi*, **Xuankai Chang***, Pengcheng Guo*, Shinji Watanabe, Yusuke Fujita, Jiaming Xu, Bo Xu, Lei Xie, “Sequence to Multi-Sequence Learning via Conditional Chain Mapping for Mixture Signals,” in *Proc. NeurIPS*, 2020, pp. 3735–3747
- Xuankai Chang**, Aswin Shanmugam Subramanian, Pengcheng Guo, Shinji Watanabe, Yuya Fujita, Motoi Omachi “End-to-End ASR with Adaptive Span Self-Attention,” in *Proc. ISCA Interspeech*, 2020
- Wangyou Zhang, Aswin Shanmugam Subramanian, **Xuankai Chang**, Shinji Watanabe and Yanmin Qian “End-to-End Far-Field Speech Recognition with Unified Dereverberation and Beamforming,” in *Proc. ISCA Interspeech*, 2020
- Xuankai Chang**, Wangyou Zhang, Yanmin Qian, Jonathan Le Roux, Shinji Watanabe “End-To-End Multi-Speaker Speech Recognition With Transformer,” in *Proc. IEEE ICASSP*, 2020, pp. 6134-6138
- Xuankai Chang**, Wangyou Zhang, Yanmin Qian, Jonathan Le Roux, Shinji Watanabe “MIMO-SPEECH: End-to-end multi-channel multi-speaker speech recognition,” in *Proc. IEEE ASRU*, 2019 (**Best Paper Award**)
- Wangyou Zhang*, **Xuankai Chang***(contributed equally), Yanmin Qian “Knowledge distillation for end-to-end monaural multi-talker ASR system,” in *Proc. ISCA Interspeech*, 2019, pp. 2633-2637
- Xuankai Chang**, Yanmin Qian, Kai Yu, Shinji Watanabe “End-to-end monaural multi-speaker ASR system without pretraining,” in *Proc. IEEE ICASSP*, 2019, pp. 6256-6260
- Xuankai Chang**, Yanmin Qian, Dong Yu, “Monaural multi-talker speech recognition with attention mechanism and gated convolutional networks,” in *Proc. ISCA Interspeech*, 2018, pp. 1586–1590
- Xuankai Chang**, Yanmin Qian, Dong Yu, “Adaptive permutation invariant training with auxiliary information for monaural multi-talker speech recognition,” in *Proc. IEEE ICASSP*, 2018, pp. 5974-5978
- Dong Yu, **Xuankai Chang**, Yanmin Qian, “Recognizing multi-talker speech with permutation invariant training,” in *Proc. ISCA Interspeech*, 2017, pp. 2456-2460
- Yimeng Zhuang, **Xuankai Chang**, Yanmin Qian , Kai Yu, “Unrestricted vocabulary keyword spotting using LSTM-CTC,” in *Proc. ISCA Interspeech*, 2016, pp. 938-942

AWARDS

- 2019, **Best Paper Award**, IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)
- 2019, Outstanding Graduates of Shanghai Jiao Tong University
- 2018, Excellent Master Scholarship – The Education Development Fund of Department of Computer Science and Technology of the Graduates in 1985 & Yuanqing Yang’s Educational Fund
- 2017, The Huawei Scholarship (A scholarship sponsored by Huawei company, for excellent students)
- 2015, National Endeavor Fellowship of Shanghai Jiao Tong University