QIBING BAI

+86 132-7888-2787 ⋄ qibingbai@link.cuhk.edu.cn ⋄ https://p1ping.github.io

EDUCATION

The Chinese University of Hong Kong, Shenzhen *08/2023 – Present*
Ph.D. in Computer Science *Shenzhen, China*
Supervisor: Haizhou Li
Research Interests: Speech Synthesis, Speech Representation Learning, Speaker Recognition

Southern University of Science and Technology *09/2020 – 06/2023*
Master of Engineering *Shenzhen, China*
Supervisor: Yu Zhang, Jimmy Liu, Tom Ko

Central South University *09/2016 – 06/2020*
Bachelor's in Electronic Engineering *Changsha, China*

WORK EXPERIENCE

Direct Speech-to-Speech Translation (Internship at ByteDance) *02/2022 – 03/2023*
- Objective: To bridge the performance gap between end-to-end speech-to-speech translation (S2ST) models and cascade models in terms of translation quality, speech intelligibility, and timbre preservation.
- Outcome: We synthesized target audios for GigaST (a pseudo En-Zh ST dataset) and trained Transformer-based models with various strategies. The best end-to-end model achieved a BLEU score of 36.9, only slightly lower than a cascade model (ST + TTS). The model was deployed for internal testing on Lark. We also experimented with back-translation to train the model with realistic speaker-parallel data.

PUBLICATIONS

Shuai Wang*, **Qibing Bai***, Qi Liu, Jianwei Yu, Zhengyang Chen, Bing Han, Yanmin Qian, and Haizhou Li. *Leveraging In-the-Wild Data for Effective Self-Supervised Pretraining in Speaker Recognition*, Proc. ICASSP, 2024.

Qibing Bai, Tom Ko, and Yu Zhang, *A Study of Modeling Rising Intonation in Cantonese Neural Speech Synthesis*, Proc. Interspeech, 2022.

Rui Wang*, **Qibing Bai***, Junyi Ao, Long Zhou, Zhixiang Xiong, Zhihua Wei, Yu Zhang, Tom Ko, and Haizhou Li. *LightHuBERT: Lightweight and Configurable Speech Representation Learning with Once-for-All Hidden-Unit BERT*, Proc. Interspeech, 2022.

Qianqian Dong, Fengpeng Yue, Tom Ko, Mingxuan Wang, **Qibing Bai** and Yu Zhang, *Leveraging Pseudo-labeled Data to Improve Direct Speech-to-Speech Translation*, Proc. Interspeech, 2022.

Fangfang Zhou, **Bing Bai**, Yitao Wu, Minghui Chen, Zengsheng Zhong, Rongchen Zhu, Yi Chen, Ying Zhao, *FuzzyRadar: Visual Analysis for Understanding Fuzzy Clusters*, Journal of Visualization, 2019.

HONORS & REWARDS

SRIBD PhD Fellowship (Gold Class), The Chinese University of Hong Kong, Shenzhen *2023*
Excellent Teaching Assistant, Southern University of Science and Technology *2022*
Graduation with Distinction, Central South University *2020*
The HollySys Scholarship, Central South University *2019*
The First Prize Scholarship, Central South University *2019*
Outstanding Student, Central South University *2018*

SKILLS

Programming & Toolkits	Python, C/C++, PyTorch, fairseq, ESPnet
Languages	Mandarin (native), English (C1, TOEFL 103), French (≈A1)

SOCIETY MEMBERSHIPS

Student Member, International Speech Communication Association (ISCA)   *2022 – present*
Student Member, IEEE Signal Processing Society   *2023 – present*

EXTRACURRICULAR ACTIVITIES

Class President, Central South University   *2019 – 2020*
Volunteer for Language Day, Southern University of Science and Technology   *2021*