

Rui-Chen Zheng | 郑瑞晨

Phone: (+86)181-8926-6050 | Email: zhengruichen@mail.ustc.edu.cn

Homepage: <https://zhengrachel.github.io/>

Education

PhD Candidate in Information and Communication Engineering

2021.09 – 2026.06 (Expected)

University of Science and Technology of China

Hefei, Anhui, China

- Supervised by Prof. Zhen-Hua Ling
- GPA: 3.9/4.3

Bachelor of Electronic Information Engineering

2017.09 – 2021.06

University of Science and Technology of China

Hefei, Anhui, China

- Thesis: Method and Practice on Text-to-Speech Without Text
- GPA: 3.89/4.3
- Minor in Business Administration

Experiences

Visiting Student

2025.09 – 2026.04 (Expected)

Center for Speech Technology Research (CSTR), University of Edinburgh

Edinburgh, Scotland, UK

- Host by Dr. Korin Richmond.
- Funded by the Speech generation for Indigenous language education (SGILE) project.

Research Intern

2025.03 – 2025.07

Tongyi Speech Group, Alibaba Inc.

Hangzhou, Zhejiang, China

- Advised by Dr. Qian Chen, Mr. Chong Deng and Mr. Qinglin Zhang

Research Interest & Publications

➤ Articulation-to-Speech Synthesis, Articulation-Acoustic Relationship

- [1]. **Rui-Chen Zheng**, Yang Ai, Zhen-Hua Ling, “Speech Reconstruction from Silent Lip and Tongue Articulation by Diffusion Models and Text-Guided Pseudo Target Generation”, in *Proc. ACM Multimedia 2024*, pages 6559-6568.
- [2]. **Rui-Chen Zheng**, Yang Ai, Zhen-Hua Ling, “Incorporating Ultrasound Tongue Images for Audio-Visual Speech Enhancement”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (Volume: 32), pages 1430-1444.
- [3]. **Rui-Chen Zheng**, Yang Ai, Zhen-Hua Ling, “Incorporating Ultrasound Tongue Images for Audio-Visual Speech Enhancement Through Knowledge Distillation”, in *Proc. INTERSPEECH 2023*, pages 844-848.
- [4]. **Rui-Chen Zheng**, Yang Ai, Zhen-Hua Ling, “Speech Reconstruction from Silent Lip and Tongue Articulation by Pseudo Target Generation and Domain Adversarial Training”, in *Proc. ICASSP 2023*, pages 1-5.

➤ Speech Coding, Speech LLMs, Speech Synthesis

- [1]. **Rui-Chen Zheng**, Hui-Peng Du, Xiao-Hang Jiang, Yang Ai, Zhen-Hua Ling, “ERVQ: Enhanced Residual Vector Quantization with Intra-and-Inter-Codebook Optimization for Neural Audio Codecs”, *IEEE Transactions on Audio, Speech, and Language Processing* (Volume: 33), pages 2539-2550.
- [2]. **Rui-Chen Zheng**, Wenrui Liu, Hui-Peng Du, Qinglin Zhang, Chong Deng, Qian Chen, Wen Wang, Yang Ai, Zhen-Hua Ling, “Say More with Less: Variable-Frame-Rate Speech Tokenization via Adaptive Clustering and Implicit Duration Coding”, *Under Review*.
- [3]. **Rui-Chen Zheng**, Hui-Peng Du, Yang Ai, Li-Rong Dai, “Enhancing Noise Robustness for Neural Speech Codecs through Resource-Efficient Progressive Quantization Perturbation Simulation”, *Under Review*.
- [4]. Chao Wang*, **Rui-Chen Zheng***, Yang Ai, Zhen-Hua Ling, “Understanding Textual Capability Degradation in Speech LLMs via Parameter Importance Analysis”, *Under Review*.
- [5]. Hui-Peng Du, Yang Ai, **Rui-Chen Zheng**, Zhen-Hua Ling, “Is GAN Necessary for Mel-Spectrogram-based Neural Vocoder?”, accepted by *IEEE Signal Processing Letters*.
- [6]. Yao Guo, Yang Ai, **Rui-Chen Zheng**, Hui-Peng Du, Xiao-Hang Jiang, Zhen-Hua Ling, “Vision-Integrated High-Quality Neural Speech Coding”, accepted by *INTERSPEECH 2025*.
- [7]. Xiao-Hang Jiang, Yang Ai, **Rui-Chen Zheng**, Zhen-Hua Ling, “A Streamable Neural Audio Codec With Residual Scalar-Vector Quantization for Real-Time Communication”, *IEEE Signal Processing Letters* (Volume: 32), pages 1645-1649.
- [8]. Hui-Peng Du, Yang Ai, **Rui-Chen Zheng**, Zhen-Hua Ling, “APCodec+: A Spectrum-Coding-Based High-Fidelity and High-Compression-Rate Neural Audio Codec with Staged Training Paradigm”, in *Proc. ISCSLP 2024*, pages 676-680.
- [9]. Xiao-Hang Jiang, Yang Ai, **Rui-Chen Zheng**, Hui-Peng Du, Zhen-Hua Ling, “MDCTCodec: A Lightweight MDCT-based Neural Audio Codec towards High Sampling Rate and Low Bitrate Scenarios”, in *Proc. SLT 2024*, pages 540-547.

[10].Fei-Liu, Yang Ai, Hui-Peng Du, Ye-Xin Lu, **Rui-Chen Zheng**, Zhen-Hua Ling, “Stage-Wise and Prior-Aware Neural Speech Phase Prediction”, in *Proc. SLT 2024*, pages 638-644.

[11].Yang Ai, Ye-Xin Lu, Xiao-Hang Jiang, Zheng-Yan Sheng, **Rui-Chen Zheng**, Zhen-Hua Ling, “A Low-Bitrate Neural Audio Codec Framework with Bandwidth Reduction and Recovery for High-Sampling-Rate Waveforms”, in *Proc. Interspeech 2024*, pages 1765-1769.

Selected Honors

- National Graduate Scholarship 2025.10
- JAC-NIO Scholarship 2024.12
- Honor Rank for Top 5% Graduates of USTC. 2021.06
- Huawei Scholarship 2020.12
- USTC Outstanding Student Scholarship, Gold Award 2019.12 & 2018.12
- Top-Notch Program Funding 2019.12 & 2018.12

Service

- Invited Reviewers for *Speech Communication*, *IEEE Signal Processing Letters* and *IEEE/ACM Transactions on Audio, Speech, and Language Processing*.

Teaching Assistant Experience

- Fundamentals of Speech Signal Processing by Prof. Zhen-Hua Ling, USTC 2022 Fall
- Fundamentals of Speech Signal Processing by Prof. Zhen-Hua Ling, USTC 2021 Fall
- Computer Programing Design A by Lecturer. Hu Si, USTC 2020 Fall

Skills

Language – Chinese/English

- TOEFL iBT: 106 (Reading: 28, Listening: 29, Speaking: 26, Writing: 23) 2023.11