

# Yicheng Gu

yichenggu@link.cuhk.edu.cn | github.com/VocodexElysium

## KEY SKILLS

---

**Solid Domain Knowledge:** Extensive knowledge of Deep Learning and Digital Signal Processing; specifically applied to Neural Vocoder, Neural Audio Codec, and Digital Audio Effects.

**Multi-disciplinary:** Experience encompassing Research, Open Source Systems, Dataset Development, Commercial Sound Design, and Music Production.

## EDUCATION

---

**The Chinese University of Hong Kong, Shenzhen**

Shenzhen, China

*Bachelor of Engineering in Computer Science and Engineering*

*Sep. 2022 – Present*

- GPA: 3.965 / 4.0, Ranking: 2 / 308 (Top 1%)

**Aalto University**

Espoo, Finland

*Exchange Student in Computer Science*

*Sep. 2024 – Present*

## EXPERIENCE

---


**Human Language Technology Lab**

Shenzhen, China

*Research Assistant, School of Data Science, CUHK-Shenzhen*

*Oct. 2022 – Present*

Supervisor: Prof. Zhizheng Wu

- Singing Voice Conversion
  - \* Investigated characteristics and the complementary role of different Content-Based Features for the Singing Voice Conversion system.
- Neural Vocoder
  - \* Built a Discriminator based on the Constant-Q Transform (CQT) and Continuous Wavelet Transform (CWT) via Representation Learning and explored their complementary role with the Short-Time Fourier Transform (STFT) to improve the Vocoder's synthesis quality. The methods have been implemented and supported by NVIDIA BigVGAN .

**Shanghai AI Laboratory**

Shanghai, China

*Research Assistant*

*Dec. 2023 – Present*

Supervisor: Dr. Yanhong Zeng

- Video-to-Audio Generation
  - \* Integrated IP-Adapter and Sound Event Detection model to existing Audio Generation pipeline, obtaining both Audio-Visual Synchronization and Text-Controllability.

**Aalto University**

Espoo, Finland

*Visiting Scholar*

*Sep. 2024 – Present*

Supervisor: Prof. Lauri Juvela

- Differentiable Digital Signal Processing

## PROJECTS

---

### Amphion

*An Open-Source Audio, Music and Speech Generation Toolkit*

- Migrate and adapt various well-known, widely used, or SOTA vocoders into our system, including: MelGAN, HiFi-GAN, BigVGAN, NSF-HiFiGAN, APNet, etc.
- Integrate comprehensive Objective Evaluation Metrics to the framework.

### FoleyCrafter

*Bring Silent Videos to Life with Lifelike and Synchronized Sounds*

- Integrate baseline model and develop the temporal adapter for controlling timestamp details.

## PUBLICATIONS

---

**Yicheng Gu**, Xueyao Zhang, Liumeng Xue, Zhizheng Wu, "Multi-Scale Sub-Band Constant-Q Transform Discriminator for High-Fidelity Vocoder," ICASSP 2024.

**Yicheng Gu**, Xueyao Zhang, Liumeng Xue, Haizhou Li, Zhizheng Wu, "An Investigation of Time-Frequency Representation Discriminators for High-Fidelity Vocoder," Submitted to TASLP.

Xueyao Zhang\*, Liumeng Xue\*, **Yicheng Gu\***, Yuancheng Wang\*, et al., "Amphion: An Open-Source Audio, Music and Speech Generation Toolkit," SLT 2024.

Xueyao Zhang, **Yicheng Gu**, et al., "Leveraging Content-based Features from Multiple Acoustic Models for Singing Voice Conversion," Machine Learning for Audio Workshop at NeurIPS 2023.

Yiming Zhang, **Yicheng Gu**, Yanhong Zeng, et al., "FoleyCrafter: Bring Silent Videos to Life with Lifelike and Synchronized Sounds," Submitted to NeurIPS 2024.

Haorui He\*, Zengqiang Shang\*, Chaoren Wang\*, Xuyuan Li\*, **Yicheng Gu**, et al., "Emilia: An Extensive Multilingual and Diverse Speech Dataset for Large-Scale Speech Generation," SLT 2024.

Xueyao Zhang, Zihao Fang, **Yicheng Gu**, et al., "Leveraging Diverse Semantic-based Audio Pretrained Models for Singing Voice Conversion," SLT 2024.

## HONORS AND AWARDS

---

- "LanHuaYing" Scholarship (Top 10 admitted students in Zhejiang Province, 2022)
- The Nobel Class (Top 1, 2024)
- The Academic Performance Scholarship, Class A (Top 1%, 2024)
- The Academic Performance Scholarship, Class B (Top 3%, 2023)
- University Entrance Scholarship (Top 0.7% in Zhejiang Province, 2022)
- Gold Award of the 2019 Shanghai STEED Programming Contest (Top 3, 2019)
- First Prize of the 16th Youth Computer Robot Competition in Provinces (Top 3, 2019)
- Deans List (Top 10%, 2022)
- Bowen Entrance Scholarship (2022)
- Undergraduate Research Award (2023-2024)

## CROSS-DISCIPLINARY ABILITIES

---

**Music Production:** experienced in Composing, Mixing, and Mastering for different genres including Pop, Electronic, etc.

**Sound Design:** experienced in Sound Designing for movie, game, etc.