

XINTONG HU

✉ erics.zju.china@gmail.com ☎ +86 198-8399-2523 🌐 <https://ericsxt.github.io/> 🏠 Erics-zju

EDUCATION

Zhejiang University, HangZhou · China 09/2022 – Present

B.S. in Computer Science, GPA 91/100 Rank 1 / 353 (Top 1%) IELTS: 7.0 (L7.5 R7.0 W6.0 S7.5)

PUBLICATIONS

(* denotes equal contribution)

[1] (**NeurIPS 2025 Poster**) MRSAudio: A Multimodal Recorded Spatial Audio Dataset.

Wenxiang Guo*,Changhao Pan*,Zhiyuan Zhu*,**Xintong Hu***,Yu Zhang*,Li Tang,Rui Yang,Han Wang,Zongbao Zhang,Yuhan Wang,Yixuan Chen,Hankun Xu,Ke Xu,Pengfei Fan,Zhetao Chen,Yanhao Yu,Fei Wu,Zhou Zhao†.
[WebPage] [PDF]

[2] (**NeurIPS 2025 Poster**) Tree of Preferences for Diversified Recommendation.

Hanyang Yuan, Ning Tang, Tongya Zheng, Jiarong Xu, **Xintong Hu**, Renhong Huang, Shunyu Liu, Jiacong Hu, Jiawei Chen, Mingli Song.
[Code] [PDF]

RESEARCH EXPERIENCE

XLANG NLP Lab, The University of Hong Kong Hong Kong

Research Assistant to Prof. Tao Yu Topic: **Embodied AI** 06/2025 – Present

YiWise Lab, Zhe Jiang University Hang Zhou

Research Assistant to Prof. Zhou Zhao Topic: **Spatial Audio** 02/2025 – 06/2025

VIPA Lab, Zhe Jiang University Hang Zhou

Research Assistant to Prof. Mingli Song Topic: **Recommendation** 07/2024 – 02/2025

PROJECTS

VLA-Sim2Real: A project focused on bridging the simulation-to-reality gap 06/2025 – present

- Developed [RealANNO], a multi-view video annotation platform for efficient and precise data labeling.
- Constructed a fine-grained, natural language-annotated video dataset to improve robotic instruction-following capabilities by redefining annotation standards and retraining pre-trained models.

MRSAudio: A Multimodal Recorded Spatial Audio Dataset 02/2025 – 06/2025

- **Dataset**:Establish MRSAudio, a 500-hour multimodal spatial audio dataset with ambisonic audio, synchronized video, motion trajectories, and fine-grained annotations (transcripts, lyrics, scores), covering 4 real-world scenarios (daily life/speech/singing/music).
- **Benchmark**:unified benchmark for 5 spatial audio tasks (spatialization, text-to-speech, singing synthesis, music generation, sound localization) enabling 3D-aware audio modeling.

SELECTED AWARDS AND HONORS

- **National Scholarship** 10/2024
- Zhejiang Province "Shangde Scholar" Award (Single Recipient) 11/2024
- Zhejiang University CS "Campus Star" Honor (Top 10 Students) 05/2025
- Zhejiang University First-Class Scholarship 09/2024