TAO TU

Visiting Researcher, National Tsing Hua University

Mail: ttaoretw@gmail.com
Website: ttaoretw.github.io
Google Scholar: NfbMw74AAAAJ
Phone: +886 903314316
Location: Hsinchu, Taiwan

EDUCATION

National Taiwan University (NTU) — Taipei, Taiwan
M.S. in Computer Science and Information Engineering — Sept 2018 – Nov 2020

- *GPA:* 4.3 / 4.3; *Ranking:* 1 / 137
- *Advisors:* Professor Lin-shan Lee and Professor Hung-yi Lee
- *Thesis:* Semi-supervised Text-to-speech Synthesis Using Sequential Quantized Representation Auto-Encoder

National Tsing Hua University (NTHU) — Hsinchu, Taiwan
B.S. in Electrical Engineering and Computer Science (EECS) — Sept 2014 – June 2018

- *GPA:* 4.13 / 4.3; *Ranking*: 1 / 22

RESEARCH INTEREST

- 3D scene geometry understanding (e.g., static/dynamic scene reconstruction)
- 3D machine perception (e.g., object detection and semantic segmentation)
- Multi-modal learning (mainly for 3D vision, speech, and text)

PUBLICATIONS (Google Scholar)

[1] **Tao Tu**, Shun-Po Chuang, Yu-Lun Liu, Cheng Sun, Ke Zhang, Donna Roy, Cheng-Hao Kuo, Min Sun, "ImGeoNet: Image-induced Geometry-aware Voxel Representation for Multi-view 3D Object Detection", *International Conference on Computer Vision (ICCV),* 2023
- Made the detection model geometry-aware by leveraging scene geometry inferred from multi-view images

[2] **Tao Tu**, Qing Ping, Govind Thai, Gokhan Tur, Prem Natarajan, "Learning Better Visual Dialog Agents with Pretrained Visual-Linguistic Representation", *Conference on Computer Vision and Recognition (CVPR),* 2021
- Leveraged generalization capability of a large-scale pretrained visual-linguistic representation to improve visual dialog understanding

[3] **Tao Tu**, Yuan-jui Chen, Alexander H. Liu, Hung-yi Lee, "Semi-supervised Learning for Multi-speaker Text-to-speech Synthesis Using Discrete Speech Representation", *Interspeech,* 2020
- Use unlabeled audio data and a proposed discrete speech representation to reduce data requirements for our multi-speaker text-to-speech model

[4] **Tao Tu***, Alexander H. Liu***,** Hung-yi Lee, Lin-shan Lee, "Towards Unsupervised Speech Recognition and Synthesis with Quantized Speech Representation Learning", *International Conference on Acoustic, Speech and Signal Processing (ICASSP),* 2020
- Learned a discrete latent speech representation from a large amount of unlabeled data with limited labeled data, which remarkably aligned with a human-defined linguistic system

[5] **Tao Tu***, Yuan-jui Chen*, Cheng-chieh Yeh, Hung-yi Lee, "End-to-end Text-to-speech for Low-resource Languages by Cross-Lingual Transfer Learning", *Interspeech,* 2019
- Proposed cross-lingual transfer learning for low-resource language end-to-end TTS

EXPERIENCES

Visiting Researcher — June 2023 – Present
Vision and Learning Lab led by Prof. Ming-Hsuan Yang**, University of California, Merced** — Merced, CA
- Articulated 3D model reconstruction from a single casual video clip

Research Assistant — Oct 2022 – Present
Vision Science Lab led by Prof. Min Sun, **National Tsing Hua University** — Hsinchu, Taiwan
- Leveraged scene geometry inferred from multi-view images to improve image-based 3D object detection (ICCV'23) for indoor environment
- Open-vocabulary image-based 3D object detection for indoor environment

Algorithm Engineer — Dec 2020 – Oct 2022
Computing AI algorithm Team, MediaTek Inc. — Hsinchu, Taiwan
- Developed voice wakeup engine with neural network design, engine acceleration, and simulator design
- Parallelizing neural network inference based on model architecture and hardware constraints, such as latency and power
- Designed predictive models for latency and power to avoid need for time-consuming cycle-accurate IP models

Applied Scientist Intern — July 2020 – Oct 2020
Alexa AI-Natural Understanding Team, Amazon — Sunnyvale, CA (remote)
- Improved visual-dialog model development through leveraging the generalization ability of a large-scale pretrained representation (CVPR'21)

Teaching Assistant – Linear Algebra, **NTU** — Fall 2018 & Fall 2019

Teaching Assistant – Deep Learning, **NTHU** — Fall 2017

AWARDS & HONORS

Academic Excellence Scholarship (NTHU EECS) — June 2017
- Recognized as outstanding student in academics

Academic Achievement Award × 6 **(NTHU EECS)** 2014 – 2018

- Recognized for being in top 5% of students in academics

MediaTek vAward × 3 2021

- Recognized for remarkable contributions and achievements

TALKS

Deep Learning for Speech Processing March 2020
Dept. of Electrical Engineering, Chung Hua University

End-to-end Text-to-speech Implementation Tutorial Aug 2019
Telecommunication Training Institute, Chunghwa Telecom

A Guide to Neural-based Text-to-speech System July 2019
AI summer school held by Artificial Intelligence Center, National Taiwan University

PROJECTS

Tacotron-pytorch | Author

- An open-source end-to-end text-to-speech system
- Achieved 100+ stars for the PyTorch implementation of the *Tacotron* model, along with speech and text preprocessing

Nerfstudio | Contributor

- An open-source platform for training and testing neural radiance fields (NeRFs)
- My contribution: Enhanced training efficiency with a 20× speedup by resolving PyTorch device-related issues

Stella VSLAM | Contributor

- An open-source monocular, stereo, and RGBD visual SLAM system
- My contribution: Improved model performance by addressing use of outdated features

SKILLS

Programming languages	Python, C, C++, Shell Script
Frameworks & Tools	PyTorch, Tensorflow, Open3D, COLMAP, Scipy, Numpy, Git, CMake
Languages	Mandarin (Advanced), English (Proficient)