WANG ZHENG

(+86)13931805565 • zheng-wa22@mails.tsinghua.edu.cn • cheeseofneo.github.io

EDUCATION

USTEP, Electrical Engineering and Information System, University of Tokyo   Oct 2024 - Mar 2025
Lab: TAKEUCHI LAB

M.S.C., Electrical Engineering, Tsinghua University   Sep 2022 - Jun 2025
Lab: OMIS(Optical Measurement and Imaging Systems) Lab   3.82/4.0 GPA

B.S.E., Electrical Engineering, Tianjin University   Sep 2018 - Jun 2022
Lab:MNMT(Micro-Nano Manufacture and Technology) Lab   3.72/4.0 GPA

HONORS

NVIDIA AI-Agent Group(**2024**), Tsinghua College Scholarship(**2024**), Membership of China Instrument and Control Society(**2022**), ICM Finalist(Top 1%, **2021**), Merit Student(**2021**), Andon Trust(**2020**), Liu Bao Scholarship(**2019**)

PUBLICATIONS

Wang, Z., Ma, R., Zeng, C., Liu, L., Li, X., Wang, X., & He, B. A fast and precise autofocus method using linear array CCD. In Optical Metrology and Inspection for Industrial Applications XI (Vol. 13241, pp. 446-456). SPIE.

Li, Z., Su, Y., Yang, R., Xie, C., **Wang, Z.**, Xie, Z., ... & Yang, H. (2025). Quantization meets reasoning: Exploring llm low-bit quantization degradation for mathematical reasoning. arXiv preprint arXiv:2501.03035.

PROJECTS

InfiAvatar – A Unified Framework for Realistic 3D Talking Avatars   Mar 2025-Now
To be submitted to ICASSP 2026
- Developed a modular pipeline integrating diffusion-based portrait synthesis, multilingual text-to-speech (TTS), 2D-to-3D lifting, and audio-driven animation, achieving 79% efficiency improvement over existing methods.
- Designed a physiologically-grounded gaze-and-blink enhancement module simulating human-like eye dynamics via coordinated saccadic movements, spontaneous blinking (15–26 BPM), and head–eye compensation.
- Conducted extensive experiments on 40 multilingual sequences, demonstrating +59% gaze naturalness, +128% blink realism, and +53% overall user preference, while maintaining real-time performance and cross-lingual compatibility.

Low-bits Denoising Diffusion Models for Masked Images   Nov 2024-Mar 2025
TAKEUCHI Lab, University of Tokyo
- Accelerated the diffusion process by adopting an aggressive quantization scheme (W1A1) to minimize memory footprint, while simultaneously guiding the model to learn more generalizable image representations.
- Investigated the role of prompts under different noise conditions in Bi-Directional Diffusion Models (BiDM). Designed experiments to decompose the latent noise space using PCA and SVD, enabling fine-grained redistribution of basic noise components and a clearer understanding of prompt–noise interaction.
- Exploring whether the representation ability of LDMs correlates with the perceptual quality of synthesized images. Current experiments employ masking-based learning strategies to assess how partial observation affects latent space representation and final image fidelity.

Exploring LLM Low-Bit Quantization Degradation for Mathematical Reasoning   Dec 2024-Feb 2025
Prof.Yang, **HKPolyU**
- Proposed a novel framework for analyzing and mitigating quantization-induced reasoning errors in LLMs.
- Introduced an automated error taxonomy (Conceptual, Method, Execution, Reasoning) and a compact "Silver Bullet" dataset enabling rapid recovery of math reasoning performance.
- Achieved full-precision level accuracy restoration with only 332 samples and 3–5 minutes of training on a single GPU.

A fast and precise autofocus method using linear array CCD   June 2023-Oct 2024
Master's thesis: **laboratory of measurement and computer vision**
- Design an active autofocus microscope using linear array CCD whose accuracy at the nanometer level and speed at the millisecond level. Develop a focus calibration algorithm, which conducts weighted fusion to the results.
- Propose a centroid extraction algorithm at sub-pixel level based on multiscale feature extraction, and fit linear mapping using sliding finite impulse response filter.

Deep Learning on Small Sample Pointsets for 3D Segmentation   Sep. 2021-May 2022

Bachelor's thesis: **State key laboratory of precision measuring technology and instruments**

- Based on a sample-aware data augmention infrastructure, proposed a novel network for point segmentation **PASN)**. Adjusted feature extractor- PointNet for degeneracy to fit our specific segmentation task.
- Propagate segmentation loss to augmentor and regress pointwise and shapewise matrixs for input samples. Experiments proved accuracy of segmentation increase by **5%** and higher robustness than PointNet on our dataset

INTERNSHIP EXPERIENCE

Baidu, Beijing China: AI Research Intern,*Team Apollo*   July 2024- Oct 2024

- For multisensor perception model, propose an infra of model fusion based on shared weights and backbone fusion.. Also conducts a cross training strategy for multi-objection and utilize distillation to compensate accuracy.
- For bev2instance module in bev perception model, conduct novel optimization the self-attention and deformable attention modules based on precision descent and infra compression and prove efficient in training accuracy.

Momenta, Shenzhen China: System Archicture Intern, *Team Middleware of System*   March 2023-July 2023

- Design the timestamp synchronization architecture of ADAS system, and construct simulated signal generator in **signal simulation platform**, which involved validating signal pathways and creating abnormal signals.
- Developed an testbench-**board verification platform** based on SOMEIP about resource schedulng of hardware, the signal accuracy and communication correctness on a high-concurrence scene.

OTHER WORK EXPERIENCE

ICM 2021, Finalist Award(top 1%):   Jan 2020 - Mar 2020

- Proposed **PRE(Page-Rank Entropy) model** vector network of musicians and quantify influence of musicians by weights of nodes at the time-span and genre-span perspectives.
- Construct local and global similarity evaluation model based on cosine simlarity of multi-dimensional eigen vectors extracted by *PCA*. Utlized k-means clustering method to classify significance of eigens between and within genres.
- Adopted *silding window automatic regression model* as time series forecasting on main eigens to predict the long and short-term developing trend of musical genres.