

## Research Interests

3D Computer Vision, Image/Video Processing, View Synthesis, Structure-from-Motion/SLAM, Scene Understanding, Interactive Vision Techniques

## Education

### University of Maryland, College Park

Ph.D. in Computer Science

Maryland, US

Sep. 2022–present

- Advisor: Jia-Bin Huang

### National Taiwan University

M.S. in Computer Science and Information Engineering

Taipei, Taiwan

Sep. 2018–Jun. 2020

- Thesis: “3D Video Stabilization with Depth Estimation by CNN-based Optimization” **[CVPR2021]**  
Committee: Yi-Ping Hung (advisor), Yung-Yu Chuang, Yu-Chiang Frank Wang, Chu-Song Chen, Kuan-Wen Chen
- Cumulative GPA: 4.24/4.3 (rank 7th/132)

### National Chiao Tung University (now National Yang Ming Chiao Tung University)

B.S. in Computer Science (Network and Multimedia Engineering Program)

Hsinchu, Taiwan

Sep. 2014–Jun. 2018

- Cumulative GPA: 4.14/4.3 (rank 1st/50)
- Academic Achievement Award: 4 times (top 5% ranking in 4 semesters)

## Publications

1. **Yao-Chih Lee**, Ji-Ze Genevieve Jang, Yi-Ting Chen, Elizabeth Qiu, Jia-Bin Huang, “Shape-aware Text-driven Layered Video Editing,” **IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**, 2023. [[webpage](#), [arxiv](#)]
2. **Yao-Chih Lee**, Kuan-Wei Tseng, Guan-Sheng Chen, Chu-Song Chen, “Globally Consistent Video Depth and Pose Estimation with Efficiency,” *arXiv:2208.02709*, 2022. [[arxiv](#)]
3. Kuan-Wei Tseng, **Yao-Chih Lee**, Chu-Song Chen, “Artistic Style Novel View Synthesis Based on A Single Image,” **IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops**, 2022. [[pdf](#)]
4. **Yao-Chih Lee**, Kuan-Wei Tseng, Yu-Ta Chen, Chien-Cheng Chen, Chu-Song Chen and Yi-Ping Hung, “3D Video Stabilization with Depth Estimation by CNN-based Optimization,” **IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**, 2021. [[webpage](#), [pdf](#)]
5. Yu-Ta Chen, Kuan-Wei Tseng, **Yao-Chih Lee**, Chun-Yu Chen, Yi-Ping Hung, “PixStabNet: Fast Multi-Scale Deep Online Video Stabilization with Pixel-based Warping,” **IEEE International Conference on Image Processing (ICIP)**, 2021. [[pdf](#)]
6. Hau Chu, Jia-Hong Lee, **Yao-Chih Lee**, Ching-Hsien Hsu, Jia-Da Li, Chu-Song Chen, “Part-aware Measurement for Robust Multi-View Multi-Human 3D Pose Estimation and Tracking,” **IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)**, 2021. [[pdf](#)]
7. Ping-Jung Duh, Yu-Cheng Sung, **Yao-Chih Lee**, Kuan-Wen Chen, Liang-Yu Fan Chiang, “A Design of Vision-based Navigation System for the Visually Impaired,” *the Conference of Taiwan Computer-Human Interaction (TAICHI)*, 2018.

8. Yu-Cheng Sung, **Yao-Chih Lee**, Sarah Wang, Wei-Ting Hu, Kuan-Wen Chen, "An UAV Autopilot System for Sports Player Tracking," *the Conference of Taiwan Computer-Human Interaction (TAICHI)*, 2017.

## Awards and Achievements

- **Reviewer**, CVPR 2023, ICCV 2023
- **Reviewer**, Pattern Recognition
- **Academic Achievement Award** × 4, (Top 5% ranking)  
Fall 2014, Spring 2016, Fall 2016, and Spring 2017
- **Excellence Award**, Undergraduate Project Competition  
An UAV autopilot system for sports player tracking
- **Departmental Core Course Scholarship**  
Top 3 ranking in the course of Operating System

## Research Experiences

### PhD Student

UMD Vision and Learning Lab advised by Prof. Jia-Bin Huang

Maryland, US  
Sep. 2022–present

- Video Editing
  - Developing text-driven video editing method with consistent shape-aware object editing.

### Research Assistant

AI Application and Integration Lab at Academia Sinica advised by Prof. Chu-Song Chen

Taipei, Taiwan  
Sep. 2020–Mar. 2022

- 3D Vision
  - [**arXiv 2022**] Developed globally consistent video dense depth and camera pose estimation, which outperformed the state-of-the-art by 19% improvement with strong efficiency.
  - [**CVPRW 2022**] Led a research team to develop stylized video view synthesis with 3D geometric constraints.
  - [**CVPRW 2021**] Contributed in a multi-view multi-human 3D pose estimation and tracking system with 100 fps.
- Image Processing
  - Solved CT metal artifact reduction in CT-MRI paired images with conditional GAN and contrastive loss.
- Recognition
  - Led a team of Traditional Chinese scene text detection and recognition in self-supervised learning manners; and developing scene text synthesis algorithms with depth estimation and scene text replacement.

### Research Assistant

Interdisciplinary Human-AI Interaction Research Project

Taipei, Taiwan  
Jul. 2020–Aug. 2020

- Human-AI Interaction
  - Advised by Prof. Yihsiu Chen (Communication, NCCU, Taiwan), Prof. Gary Hsieh (Human Centered Design & Engineering, UW, Seattle). and Prof. Chien-Wen Tina Yuan (Library & Information Studies, NTNU, Taiwan).
  - Developed experimental platforms of human-AI collaboration to serve over 700 participants.

### Graduate Research Assistant

Image and Vision Lab at NTU with MediaTek, Inc. advised by Prof. Yi-Ping Hung

Taipei, Taiwan  
Sep. 2018–Jun. 2020

- Video Processing
  - [CVPR2021] Proposed the first 3D learning-based video stabilization algorithm with self-supervised depth and pose estimation. The method outperforms the state-of-the-art methods, especially in challenging videos.
  - [ICIP2021] Contributed an online video stabilization algorithm with a coarse-to-fine approach, which achieved 54.6 fps and surpassed the state-of-the-art by 29% with robust shape preservation.
- 3D Vision
  - Developed self-supervised monocular depth and camera ego-motion estimation for wild videos.
  - Conducted thorough evaluations on the performance of local feature algorithms for visual SLAM systems.

### Undergraduate Research Assistant

Collaborative Vision Lab at NCTU advised by Prof. Kuan-Wen Chen

Hsinchu, Taiwan  
Aug. 2016–Jun. 2018

- Interactive Vision Technique
  - [TAICHI2017] Developed UAV autopilot and visual tracking system with OCR and human detection.
  - [TAICHI2018] Contributed in a navigation system for visually impaired with streaming semantic segmentation.
- 3D Vision
  - Constructed a semi-automatic feature correspondence annotation system to construct a real-world dataset of matching patches for a learning-based viewpoint- and illumination-invariant local feature extraction.
  - Developed semantic segmentation and SLAM system with 3D reconstruction for virtual reality environments.

## Teaching

- **Teaching Assistant**, Introduction to Data Science (CMSC320), UMD Spring
- **Teaching Assistant**, Introduction to Artificial Intelligence (CMSC421), UMD Fall 2022
- **Teaching Assistant**, 3D Computer Vision with Deep Learning Applications (CSIE5429), NTU Spring 2021
- **Teaching Assistant**, Digital Image Processing (CSIE5612), NTU Fall 2019
- **Teaching Assistant**, Probability (CSIE2121), NTU Spring 2019
- **Teaching Assistant**, Computer Vision for UAV Autopilot (DCP1249), NCTU Spring 2018

## Skills

- **Programming Languages:**  
Python, C/C++, L<sup>A</sup>T<sub>E</sub>X, MatLab, Bash, Javascript, PHP, SQL, C#, Swift
- **Development Tools:**  
Unix, PyTorch, OpenCV, Open3D, COLMAP, OpenGL, TensorFlow, Git, Unity
- **Languages:**  
Chinese (native), English (fluent, TOEFL MyBest: 105)