WEN-HSUAN CHU

Email: chuwenhsuan@cmu.edu
Phone: 412-759-9536
Website: https://wenhsuanchu.github.io

RESEARCH INTERESTS

Computer Vision and applications to Robotics/Embodied AI, inferring 3D geometry and motion from videos, 2D/3D point and object tracking, and learning dynamic models through sensor inputs.

EDUCATION

Carnegie Mellon University *2020 - current*
Ph.D. in Robotics (Advisor: Prof. Katerina Fragkiadaki) GPA: 4.0/4.3
Carnegie Mellon University *2018 - 2020*
M.S. in Robotics (Advisor: Prof. Kris Kitani) GPA: 4.08/4.3
National Taiwan University *2014 - 2018*
B.S. in Electrical Engineering GPA: 3.9/4.3

RESEARCH EXPERIENCE

Robotics Institute, CMU 2020 - Present
Ph.D. Student advised by Prof. Katerina Fragkiadaki

- Video-to-4D lifting, and 3D geometry plus motion understanding from monocular videos.
- Temporarally-consistent video amodal completion using diffusion models, with applications to video-to-4D in occluded videos.
- General purpose Open Vocabulary Object Trackers to detect and track any object in the wild.
- Fine grained point tracking for videos through time and occlusions by tracking in a region-aware manner by leveraging region tracklets to guide point tracking.
- Efficient person and vehicle tracking and Re-ID models across multiple views and cameras for use on onboard computers (collaboration with NREC).
- Leveraging learned particle based physics simulators for control from RGB sensor inputs for downstream manipulation tasks, as well as using sparse mixture of experts models to discover and train "experts" for different kinds of motion.

KLab at Carnegie Mellon University 2018-2020
M.S. Student advised by Prof. Kris Kitani

- Combined a Reinforcement Learning based neural batch sampler, an autoencoder, and a predictor to perform semi-supervised pixel-level anomaly detection and segmentation in images using generated loss profiles from the autoencoder.
- Developed a model for single camera worker action recognition in construction sites using RGB, optical flow, and pose information.

Vision and Learning Lab at National Taiwan University 2017-2018
Undergraduate Student Research advised by Prof. Yu-Chiang Wang

- Learned a diverse patch-based data augmentation policy using the Maximum Entropy Reinforcement Learning framework and a recurrent encoder-classifier network architecture for patch based Few-Shot image classification.
- Utilized semantic information from labels using word2vec to learn semantic-guided spacial attention models for Few-Shot image classification.

PUBLICATIONS

In submission:

1. **Wen-Hsuan Chu**, Lei Ke, Jianmeng Liu, Mingxiao Huo, Pavel Tokmakov, Katerina Fragkiadaki. "Robust Multi-Object 4D Generation for In-the-wild Videos". *Under submission*.
2. Adam Harley, Yang You, Yang Zheng, Xinglong Sun, Nikhil Raghuraman, Sheldon Shiqian Liang, Yunqi Gu, **Wen-Hsuan Chu**, Suya You, Achal Dave, Pavel Tokmakov, Rares Andrei Ambrus, Katerina Fragkiadaki, Leonidas Guibas. "TAG: Tracking at Arbitrary Granularity". *Under submission*.
3. **Wen-Hsuan Chu**, Yizhou Zhao, Jie Li, Rareș Ambruș, Adam Harley, Katerina Fragkiadaki. "Tracking Any Point in a Video with Region-Aware Point Trajectory Transformers". *Under submission*.

Conference Papers:

1. **Wen-Hsuan Chu**, Lei Ke, Katerina Fragkiadaki. "DreamScene4D: Dynamic Multi-Object Scene Generation from Monocular Videos". NeurIPS 2024.
2. **Wen-Hsuan Chu**, Adam Harley, Pavel Tokmakov, Achal Dave, Leonidas J. Guibas, Katerina Fragkiadaki. "Zero-Shot Open-Vocabulary Tracking with Large Pretrained Models". ICRA 2024.
3. **Wen-Hsuan Chu**, Kris M. Kitani. "Neural Batch Sampling with Reinforcement Learning for Semi-Supervised Anomaly Detection". *ECCV 2020*.
4. **Wen-Hsuan Chu**, Yu-Jhe Li, Jing-Cheng Chang, Yu-Chiang Frank Wang. "Spot and Learn:A Maximum-Entropy Patch Sampler for Few-Shot Image Classification". *CVPR 2019*, Long Island, USA.
5. **Wen-Hsuan Chu**, Yu-Chiang Wang. "Learning Semantics-Guided Visual Attention for Few-shot Image Classification". *ICIP 2018*, Athens, Greece.
6. Everett Fall, **Wen-Hsuan Chu**, Liang-Gee Chen. "Atlas Architecture: Constructing and Traversing Generalized Graph Manifolds in Feature Space". Poster session presented at *Workshop on Lifelong Learning: A Reinforcement Learning Approach @ICML 2017*, Sydney, Australia.