Joshua Cao

(412) 954-8151 caoyuchen.joshua@gmail.com Github LinkedIn

Skills

Languages	C++, Python, MATLAB, JavaScript, Swift, C#, SQL, CUDA, LaTeX
Libraries	OpenCV, OpenGL, Scikit-learn, Matplotlib, NumPy, Keras, SciPy, React
Frameworks & Tools	Pytorch, ROS, AWS, GCP, W&B, MongoDB, Spark, Hadoop, XGBoost, Docker

Education

Carnegie Mellon University Sep 2021 - May 2023
MS. in Computational Design (Computer Vision Track) *Pittsburgh, PA*

University of Chinese Academy of Sciences Sep 2016 - July 2020
MS. in Computer Science *Shanghai, China*

Experience

APEX(EzPT) July 2022 - Aug 2022
Intern, Computer Vision Engineer & iOS Developer *San Francisco, USA*

- Built a real-time human **Pose Estimation**, **Classification**, and **Rep-counting** system on iOS App with Pytorch, MediaPipe, OpenCV, SwiftUI, Google Cloud Function, and Firebase, which works as a virtual physical therapist.
- Set up a **CI/CD Dataset System** for auto-updating and training by Colab, KNN, GitHub, DVC and GCP API.

Robot Labs, Carnegie Mellon University Sep 2021 - May 2023
Research Assistant, Advisor: Prof. Daniel Cardoso Llach and Prof. Katerina Fragkiadaki *Pittsburgh, PA*

- **Amazon Alexa Prize: SimBot Challenge**
  - Researched and implemented the Alexa housework-robot tasks such as **ASR**, **Transformer-based text2text parser**, **ViT** and **Mask-RCNN** semantic detector, **RGB-D SLAM navigator**, and robot logic.
  - Distributed simulator engine, learning models, robot logics on EC2s by **Flask REST API**. Used Amazon S3, CloudWatch, and DynamoDB to **collect datasets & logs** from users' utterances to train the language model.
- **ReAC: Husky Ground Robot**
  - Built a ground-robot Husky with an onboard **ROS** system for navigation and obstacle avoidance. Developed RGB, 3D Lidar, IMU **multi-sensors fused SLAM** algorithms. Configured ROS local, global **path planner**.
  - **Simulated environments** with Nvidia Isaac Sim and Gazebo for **pedestrian detection** and **RL training**. Used Github Actions, Docker, and DVC for **dataset version control**.

Mobile Perception Lab Sep 2016 - Dec 2020
Software Engineer, Research Assitant, Advisor: Prof. Laurent Kneip *Shanghai, China*

- Built a **SLAM system** with SIFT Feature, 7/8 Points Matching, loop closure with LM optimization in MATLAB. And researched **Relocalization** of robot hijacks by Extended Kalman Filter and MaskRCNN in Pytorch.
- **Synthesized Semantic Dataset** with ground truth and benchmarks by modeling physically accurate camera.
- Combined SfM and VAE to 3D predict and **reconstruct models from partial continuous RGB-D observation**.
- Used RPC to build a **distributed system** of TX1, RPi3 and DJI N3 to orchestrate CV tasks and flight control.

Publications

Incremental Semantic Localization using Hierarchical Clustering of Object Association Sets **ACCV 2022**
*Lan Hu, Zhongwei Luo, Runze Yuan, **Yuchen Cao*** *https://arxiv.org/abs/2208.13210* *Sep 2022*

Representations and Benchmarking of Modern Visual SLAM Systems **Sensors Journal**
***Yuchen Cao**, Lan Hu and Laurent Kneip.* *https://www.mdpi.com/1424-8220/20/9/2572* *Mar 2020*

Dense Object Reconstruction from RGBD Images with Embedded Deep Shape Representations **ACCV Workshop**
*Hu, Lan, **Yuchen Cao**, Peng Wu and Laurent Kneip.* *https://arxiv.org/abs/1810.04891* *Oct 2018*

Selected Projects

NeRF-based 3D Style Transfer / *Computer Vision & Graphics, Deep Learning* April 2022 - Jan 2023

- Implemented 3D style transfer by combining **Instant-ngp** and **Neural Style Transfer** in the latent space to maintain the consistency of styles throughout the entire scene. Put in a VR app for an art exhibition.