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 - Now
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 *Remote, USA*

- Built a real-time human pose estimation, classification, and rep counting system on iOS App with Pytorch, Keras, Google MediaPipe, OpenCV, SwiftUI, and Firebase. It works as a virtual physical therapist in practice.
- Set up a CI/CD for dataset management and generation system with Colab, KNN, GitHub and Google Cloud API.

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

- **Amazon Alexa Prize: SimBot Challenge**
  - Designed and Implemented the Alexa virtual home robot to finish housework tasks with ASR, language parser, semantic detector, RGB-D SLAM navigator, and robot instruction logic. Language model includes Transformers-based profanity, co-ref, text-to-text, and ViT vision-assisted parser.
  - Set up distributed cloud infrastructure for robot platform and learning models by REST API, Amazon S3, CloudWatch, DynamoDB. Developed dataset collector 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, and configured with ROS local, global path planner. Researched pedestrian tracking and following algorithms.
  - Simulated environments with Nvidia Isaac Sim and Gazebo for pedestrian detection and RL training.

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 extractor in the latent space to maintain the consistency of styles from different perspectives. Put in a VR app for an art exhibition.