Department of
Mechanical Engineering
The University of Hong Kong

DEPARTMENT OF
MECHANICAL ENGINEERING
THE UNIVERSITY OF HONG KONG

# SEMINAR

**Date:** 15 April, 2025 (Tuesday)
**Time:** 9:30am - 12:00noon and 2:00pm - 5:00pm (HKT)
**Venue:** Online Seminars

**Join Zoom Meeting:**
**https://hku.zoom.us/j/93040991383?pwd=DuCbyGrxKvb1KzvkSM3LoF0utas02O.1**

**Meeting ID:** 930 4099 1383

**Password:** 023749

**Supervisor:** Prof. F Zhang

| Name | Time | Title | Abstract |
| --- | --- | --- | --- |
| Fangming Cheng | 9:30AM --10:00 AM | UAV design and visualization system using the integrated planning and control framework | When building a drone for autonomous flying in the unknown environment, choosing a reliable planning system is of |

| | | | |
|---|---|---|---|
| | | | paramount importance. The planning system must be able to deal with suddenly appearing objects and generate safe trajectories with low latency. We choose to use the IPC (Integrated Planning and Control framework) as the planning system for our drone. In the context of autonomous flying in an unknown environment, IPC could react quickly for suddenly appearing objects and replan safely. It can generate actual control commands that strictly meet with the safety constraints immediately in this replan process. Inside IPC, the ROG-Map (the robot-centric occupancy grid map) is used to maintain a probabilistic map and continuously update the status of the surrounding environment, which enables IPC to replan within the safe areas. The concept of the sliding map provided by ROG-Map is also used to maintain a robot-centered 3D sliding map, which continuously rendering the surroundings using the received Lidar point clouds. |
| Liuyu Shi | 10:00AM -- 10:30AM | Real-time Bandwidth-efficient Occupancy Grid Map Synchronization for Multi-Robot Systems | Robot swarms are increasingly being applied in various domains. However, due to the inherent limitation imposed by low real-time communication bandwidth, the synchronization of environmental information among multiple robots remains a persistent and challenging problem in practical applications. In response to this challenge, we introduce a comprehensive framework for synchronizing occupancy grid maps (OGMs) in practical multi-robot systems that operate under communication bandwidth constraints. In our research, we elaborately design the data structure of transmitted local OGMs and employ the Hilbert space-filling curve for voxel sorting. By adopting this approach, data redundancy is effectively increased, resulting in lower information entropy |

| | | | |
|---|---|---|---|
| | | | for compression and significantly reducing the volume of communication data. Finally, our framework outperforms the benchmark method by reducing the average and maximum bandwidth usage by more than 10 times in high-resolution scenarios. Moreover, our method has been successfully applied in the multi-UAV autonomous navigation application, demonstrating its real-time and bandwidth-efficient nature, as well as its practical value. |
| Zhengze Xu | 10:30AM -- 11:00AM | Deep Reinforcement Learning for Drone Navigation: Advances, Challenges, and Outlook | The global drone industry has witnessed significant growth in recent years, driven by expanding applications requiring autonomous navigation, from logistics delivery to infrastructure inspection. Although traditional hierarchical planning frameworks remain effective in structured environments, deep reinforcement learning (DRL) presents a promising end-to-end alternative for navigation systems. This report comprehensively examines recent advancements in DRL methodologies for drone obstacle avoidance and navigation. By comparing vision-based and LiDAR-based perception systems, this report analyzes how sensor characteristics, control architecture design, and environmental adaptability jointly influence navigation performance. Key challenges are identified in training efficiency, sim-to-real transfer capability, and safety assurance mechanisms. The analysis concludes by highlighting persistent technical limitations in these areas and proposing possible avenues for future research development. |
| Yuxi Liu | 11:00AM -- 11:30 AM | MARSIM-RL: A Reinforcement Learning Simulator for LiDAR-Enabled UAV Navigation | As unmanned aerial vehicles (UAVs) increasingly operate in complex environments, high-fidelity simulators have become |

| | | | |
|---|---|---|---|
| | | | critical for developing and validating autonomous navigation systems. However, existing simulation platforms face fundamental challenges in balancing computational efficiency with sensor realism, particularly for LiDAR-based perception. To address three key limitations – sampling efficiency, multi-agent training support, and LiDAR fidelity – we present MARSIM-RL, a reinforcement learning (RL)-oriented simulator developed on the MARSIM framework. By integrating Python APIs with ROS interfaces, our system enables automated generation of randomized 50m×50m point cloud environments and supports parallel training for over 50 UAVs, achieving significant reductions in training time. Compared to Gazebo, the proposed simulator demonstrates real-time performance through innovative point cloud publishing strategies and optimized memory management. For platform validation, we introduce an end-to-end RL navigation system. Experimental results show perfect arrival and collision avoidance rates at 5 m/s, while maintaining computational latency under 7ms. Comparative analyses demonstrate our platform's LiDAR depth imaging quality through real-world point cloud mapping, along with higher training efficiency than conventional simulators. |
| Hoada Zhu | 11:30AM -- 12:00 NOON | FAST-LIVO2: Fast, Direct LiDAR-Inertial-Visual Odometry | As the demand for intelligent robots to operate in real-world environments grows, especially in unstructured or featureless settings, it becomes clear that systems relying on a single sensor cannot provide the accurate and robust pose estimation required. To address this challenge, sensor fusion techniques combining commonly used sensors such as LiDAR, cameras, and IMUs are gaining increasing attention. This approach |

| | | | |
|---|---|---|---|
| | | | leverages the strengths of each sensor to deliver improved pose estimation and enables the creation of precise, dense, and colored point cloud maps, even in environments where individual sensors may struggle. |
| Shuang Geng | 2:00PM -- 2:30 PM | A Lightweight LiDAR-Based UAV Exploration Framework for Large-Scale Scenarios | Autonomous exploration is a fundamental problem for various applications of unmanned aerial vehicles (UAVs). Recently, LiDAR-based exploration has gained significant attention due to its ability to generate high-precision point cloud maps of large-scale environments. While the point clouds are inherently informative for navigation, many existing exploration methods still rely on additional, often expensive, environmental representations. This reliance stems from two main reasons: the need for frontier detection or information gain computation, which typically depends on memory-intensive occupancy grid maps, and the high computational complexity of path planning directly on point clouds, primarily due to costly collision checking. To address these limitations, we present EPIC, a lightweight LiDAR-based UAV exploration framework that directly exploits point cloud data to explore large-scale environments. EPIC introduces a novel observation map based on the quality of point clouds, treating the environment as a collection of small surface patches and evaluating their observation quality. It maintains and updates this quality using spatial hashing. By guiding the UAV from well-observed to poorly-observed areas, EPIC eliminates the need for global occupancy grid maps, while ensuring robust exploration and effective performance across diverse environments.We also propose an incremental topological graph construction method operating |

| | | | |
|---|---|---|---|
| | | | directly on point clouds, enabling real-time path planning in large-scale environments. Leveraging these components, we build a hierarchical planning framework that generates agile and energy-efficient trajectories, achieving significantly reduced memory consumption and computation time compared to most existing methods. Extensive simulations and real-world experiments demonstrate that EPIC achieves faster exploration while significantly reducing memory consumption compared to state-of-the-art methods. |
| Yiming Luo | 2:30PM -- 3:00 PM | Flexible Radio Mapping with a Hybrid Propagation Model and Scalable Autonomous Data Collection | Communication is fundamental for multi-robot collaboration, with accurate radio mapping playing a crucial role in predicting signal strength between robots. However, modeling radio signal propagation in large and occluded environments is challenging due to complex interactions between signals and obstacles. Existing methods face two key limitations: they struggle to predict signal strength for transmitter-receiver pairs not present in the training set, while also requiring extensive manual data collection for modeling, making them impractical for large, obstacle-rich scenarios. To overcome these limitations, we propose FERMI, a flexible radio mapping framework. FERMI combines physics-based modeling of direct signal paths with a neural network to capture environmental interactions with radio signals. This hybrid model learns radio signal propagation more efficiently, requiring only sparse training data. Additionally, FERMI introduces a scalable planning method for autonomous data collection using a multi-robot team. By increasing parallelism in data collection and minimizing robot travel costs between regions, overall data collection |

| | | | |
|---|---|---|---|
| | | | efficiency is significantly improved. Experiments in both simulation and real-world scenarios demonstrate that FERMI enables accurate signal prediction and generalizes well to unseen positions in complex environments. It also supports fully autonomous data collection and scales to different team sizes, offering a flexible solution for creating radio maps. |
| Yunfei Wan | 3:00PM -- 3:30 PM | Mesh-Learner: Texturing Mesh with Spherical Harmonics | Nowadays, Neural Radiance Fields (NeRF), 3D Gaussian Splatting (3DGS) and their subsequent variants have significantly improved the rendering quality and efficiency in the novel view synthesis task. However, they ignore a practical issue: inability to directly integrate with traditional rasterization pipelines. To visualize NeRF or 3DGS in rendering engines like Unity, users typically introduce third-party plugins, which cause rendering artifacts. What's more, there is a need to convert 3DGS or NeRF models into mesh format, for many applications (e.g., games), to interact with existing mesh models. These conversions not only introduce additional workload but also inevitably lead to losses in rendering quality and geometric integrity. To fix this gap, we propose to directly learn a texture from the existing meshes, so the learning outcome can be directly applied to a rasterization pipeline without conversion. We present Mesh-Learner, a 3D reconstruction and rendering framework that is natively compatible with traditional rasterization pipelines. It integrates mesh and spherical harmonic (SH) texture (i.e., texture filled with SH coefficients) into the learning process to learn each mesh's view-dependent radiance end-to-end. Images are rendered by interpolating surrounding SH Texels at each pixel's sampling point using a novel interpolation |

| | | | |
|---|---|---|---|
| | | | method. Conversely, gradients from each pixel are back-propagated to the related SH Texels in SH textures. Mesh-Learner exploits graphic features of rasterization pipeline (texture sampling, deferred rendering) to render, which makes Mesh-Learner naturally compatible with tools (e.g., Blender) and tasks (e.g., 3D reconstruction, scene rendering, reinforcement learning for robotics) that are based on rasterization pipelines. Our system can train vast, unlimited scenes because we transfer only the SH textures within the frustum to the GPU for training. At other times, the SH textures are stored in CPU RAM, which results in moderate GPU memory usage. The rendering results on interpolation and extrapolation sequences in the Replica and FAST-LIVO2 datasets achieve state-of-the-art performance compared to existing state-of-the-art methods (e.g., 3D Gaussian Splatting and M2-Mapping). |
| Ziming Wang | 3:30PM -- 4:00 PM | GUNDAM: Giraffe-Inspired Ultra-Scale Arm with LiDAR-Centric Assistance for Safer Heavy-Duty Aerial Work and Maintenance | In high-risk, heavy-duty aerial work scenarios such as power line inspection and structural maintenance, ensuring operational safety and environmental awareness for large articulated robotic arms remains a formidable challenge. We present GUNDAM, a novel embodied perception system inspired by the giraffe's elevated viewpoint and spatial attentiveness. GUNDAM integrates a LiDAR-centric sensor suite with an articulated arm platform to enable real-time 3D mapping, semantic understanding, and spatial localization in both Cartesian and configuration spaces. By fusing multimodal data from LiDAR, vision, and inertial sensors, our system achieves robust scene comprehension and safe motion planning in complex, cluttered, or hazardous |

| | | | |
|---|---|---|---|
| | | | environments.We further demonstrate GUNDAM's capabilities in scenarios involving large-scale infrastructure, highlighting its potential to improve both autonomy and safety for next-generation industrial maintenance robots. |
| Siqi Liang | 4:00PM -- 4:30 PM | Bridging the Sim-to-Real Gap: Learning Residual Policies for Dynamic Robot Control | Training robots in simulation is efficient, but transferring learned behaviors to the real world remains challenging due to dynamics mismatch—differences between simulated and real-world physics. Traditional approaches like System Identification and Domain Randomization often struggle with complex, high-dimensional discrepancies.<br>This seminar explores learning-based methods that directly compensate for sim-to-real gaps, focusing on residual action policies. These models learn corrective actions to align simulated and real-world dynamics, enabling more precise and agile control. We compare these approaches with conventional techniques, highlighting their advantages in reducing tracking errors and improving real-world performance. Finally, we discuss key challenges and future directions in sim-to-real transfer for dynamic robotic systems. |
| Jiaao Chen | 4:30PM -- 5:00 PM | 3D Dynamic Reconstruction | The simulation of 3D environments constitutes a fundamental challenge in computational modeling, particularly in domains requiring high-fidelity digital twin representations such as 3A gaming and 3D cartoon production. Notable exemplars include the photorealistic 3D environment rendering demonstrated in Black Myth: Wukong. Recent advancements have propelled 3D spatial language processing into a critical research frontier, aiming to resolve open-vocabulary spatial queries through multimodal understanding. State-of-the-art |

| | | | |
|---|---|---|---|
| | | | 3D reconstruction methodologies, exemplified by Neural Radiance Fields (NeRF) and 3D Gaussian Splatting, have revolutionized geometric reconstruction through differentiable volumetric rendering and explicit point cloud parameterization. Concurrently, semantic mapping technologies have emerged as pivotal enablers for robotic applications, addressing complex challenges in scene interpretation and task execution. Contemporary approaches integrate these reconstruction techniques with vision-language foundation models (e.g., CLIP for open-vocabulary semantic embedding, SAM for instance segmentation) to construct semantically annotated 3D maps. However, persistent challenges remain in achieving dynamic semantic comprehension—specifically, the capacity to adapt to transient environmental modifications (e.g., seasonal variations, temporary obstructions) and maintain real-time map updating capabilities. |

**ALL INTERESTED ARE WELCOME**
**For further information, please contact Prof. F. Zhang at 3917 7909.**