Research Statement

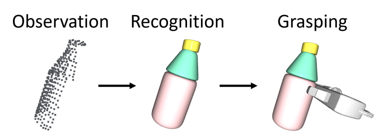
Seungyeon Kim

My research focuses on developing practical solutions for intelligent robots that can operate effectively in real-world scenarios using limited data, while remaining generalizable to unfamiliar environments. Specifically, I aim to design robotic agents that exhibit human-like behaviors and motivations, inspired by the human ability to perform a wide range of manipulation tasks. From my early graduate research on human motor control, a key observation is that the human motor system inherently incorporates certain *inductive biases*, which play a crucial role in enabling humans to rapidly adapt to changing environments. Aligned with this perspective, I am particularly interested in leveraging inductive biases to equip robots with human-level adaptability and performance. In the following sections, I elaborate on my key research contributions and outline future directions.

Research Progress

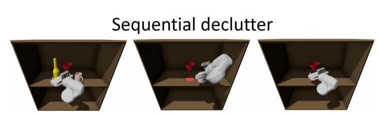
1. Inductive Bias on 3D Recognition System for Object Manipulation

Recognizing and manipulating objects from vision data has emerged as a fundamental aspect of robotics, encompassing tasks such as grasping, pushing, and rearranging objects within complex environments. Specifically, I emphasize the need for an *effective representation* for manipulation – one that captures both object 3D geometry and instance-level information. The first criterion is essential for enabling precise manipulation and avoiding collisions, while the second is critical for target-driven manipulation tasks. To achieve this, my work draws inspiration from the *recognition-by-components theory*, which suggests that humans recognize objects as compositions of simple primitive shapes. Building on this idea, I aim to incorporate inductive biases into object recognition systems to support efficient and effective manipulation.



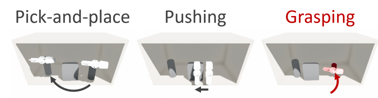
Deformable superquadric shape primitives [2]. As a first step toward this goal, I have proposed *deformable superquadrics* as a set of expressive and efficient shape primitives for *object grasping*. Simple primitives, such as bounding boxes, often fail to capture the complexity of real-world objects, while detailed representations – such as meshes or implicit functions – are computationally expensive and generalize poorly to novel objects, limiting their suitability for real-time planning. Balancing efficiency and expressiveness in shape representation remains a key open challenge.

Deformable superquadrics are parameterized by only eight continuous variables, yet they can represent a wide variety of object parts. Their simple, closed-form surface equations support real-time robot planning and collision checking. I have developed a model that takes visual observations as input and predicts a set of deformable superquadrics that represents the full object shape. Then, I have demonstrated its effectiveness in generating successful collision-free grasp poses in real-time.



Recognizing transparent tableware objects [3]. Recognizing and manipulating *transparent tableware* is particularly challenging due to unreliable depth measurements and limited visibility when objects are placed on shelves or against walls. To address this, I extend the deformable superquadric inductive bias by incorporating superparaboloids and shearing deformations, resulting in a shape representation capable of capturing the concave geometries commonly found in tableware. I also design a novel recognition model that infers the geometry of each object – represented as a union of extended deformable superquadrics – from visual input data. The model outperforms existing state-of-the-art methods in predicting the shapes of transparent objects from partial observations. Furthermore, I demonstrate its practical utility in two downstream tasks: sequential decluttering and object rearrangement for target retrieval.

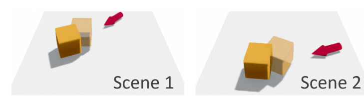
Mechanical search on cluttered shelves [4]. Finding and grasping an *initially invisible target object* on a cluttered shelf presents a significant challenge. In such scenarios, the robot must strategically rearrange surrounding objects to reveal the target's pose and grasp it, all while avoiding collisions with the shelf and nearby objects. To address this, I leverage a superquadric-based object recognition model to effectively reason about potential target poses and assess graspability. I then formulate the manipulation task as a tractable optimal control problem and solve it accordingly. I have validated that this approach enables a robot to successfully find and grasp the target object, even in the presence of real-world vision sensor noise.



2. Equivariant Models for Enhancing Generalizability of Robot Agents

Compared to vision or language data, robotic data is significantly more limited and expensive to collect. One promising research direction I have pursued is the use of *equivariant models* [5, 6]. By designing models that respect the inherent symmetries of tasks or environments, these approaches enable robots to adapt more effectively to novel tasks and settings. I am particularly interested in discovering symmetries within specific tasks and incorporating them into new model designs.

$\mathrm{SE}(2)$-equivariant pushing dynamics learning [5]. In a physics-based model, displacing an object and applying the same relative push should yield a correspondingly displaced motion – reflecting the $\mathrm{SE}(2)$-equivariance of pushing dynamics. This implies that redundant training data across all object poses is unnecessary. Building on this insight, I have developed a data-driven model that explicitly incorporates $\mathrm{SE}(2)$-equivariance. This model achieves state-of-the-art performance in predicting visual pushing dynamics and demonstrates its effectiveness in model-based optimal control across a variety of pushing manipulation tasks.



Ongoing and Future Direction

In summary, my research has focused on two main areas: (1) developing effective shape representations for vision-based object manipulation, and (2) designing generalizable robotic agents that inherently incorporate equivariance. Looking ahead, I am eager to deepen and expand my research to achieve tangible impact in robotics through the following research directions:

- While my previous work has focused on rigid body objects [2, 3, 4], I am now shifting my attention toward developing effective representations for *non-rigid, deformable objects* in the context of vision-based object recognition and manipulation. Deformable objects exhibit a wide range of structural variations – ranging from articulated mechanisms to cloth-like materials – and I believe that effective, task-specific representations can be developed for each category. For example, one of my ongoing research efforts focuses on improving the recognition of articulated objects by incorporating screw theory as an inductive bias.
- One of my major research ambitions is to solve *contact-rich manipulation tasks*, such as robotic assembly. In addition to identifying effective object representations for tasks involving physical contact, I am also interested in the design of robot motion, drawing on my research experience across various domains, including generative policy learning from demonstration [6, 7] and reinforcement learning [8]. In particular, I am passionate about incorporating physical laws and compliant environmental interactions (e.g., admittance control) as inductive biases, enabling robots to perform forceful tasks safely and adaptively.
- When designing adaptable and generalizable robotic agents, I believe the key lies in accounting for the *inherent symmetries* that the model should exhibit [5, 6], as well as the *underlying structure* of the data [9]. In this context, I am also interested in empowering robotic models through a *Riemannian geometric framework*. Moving forward, I aim to explore the types of symmetries present in recognition and dynamics models, and to investigate the underlying geometry of robotic data – with a particular focus on contact-rich manipulation scenarios involving force-torque or tactile sensor data.

References

[1] **Seungyeon Kim**, Jaewoon Kwon, Jin-Min Kim, Frank C. Park, Sang-Hoon Yeo. "On the Encoding Capacity of Human Motor Adaptation". In Journal of Neurophysiology (JNP). 2021.

[2] **Seungyeon Kim***, Taegyun Ahn*, Yonghyeon Lee, Jihwan Kim, Michael Y. Wang, Frank C. Park. "DSQNet: A Deformable Model-Based Supervised Learning Algorithm for Grasping Unknown Occluded Objects". In IEEE Transactions on Automation Science and Engineering (T-ASE). 2022.

[3] Young Hun Kim*, **Seungyeon Kim***, Yonghyeon Lee, Frank C. Park. "T^2SQNet: A Recognition Model for Manipulating Partially Observed Transparent Tableware Objects. In Conference on Robot Learning (CoRL). 2024.

[4] **Seungyeon Kim***, Young Hun Kim*, Yonghyeon Lee, Frank C. Park. "Leveraging 3D Reconstruction for Mechanical Search on Cluttered Shelves". In Conference on Robot Learning (CoRL). 2023.

[5] **Seungyeon Kim**, Byeongdo Lim, Yonghyeon Lee, Frank C. Park. "SE(2)-Equivariant Pushing Dynamics Models for Tabletop Object Manipulations". In Conference on Robot Learning (CoRL). 2022.

[6] Byeongho Lee*, Yonghyeon Lee*, **Seungyeon Kim**, MinJun Son, Frank C. Park. "Equivariant Motion Manifold Primitives". In Conference on Robot Learning (CoRL). 2023.

[7] Yonghyeon Lee, Byeongho Lee, **Seungyeon Kim**, Frank C Park. "Motion Manifold Flow Primitives for Language-Guided Trajectory Generation". arXiv preprint (under review). 2024.

[8] Seokjin Choi*, Yonghyeon Lee*, **Seungyeon Kim**, Che-Sang Park, Himchan Hwang, and Frank C. Park. "Diverse Policy Learning via Random Obstacle Deployment for Zero-Shot Adaptation". In IEEE Robotics and Automation Letters. 2025.

[9] Yonghyeon Lee*, **Seungyeon Kim***, Jinwon Choi, Frank C. Park. "A Statistical Manifold Framework for Point Cloud Data". In International Conference on Machine Learning (ICML). 2022.