Complex Robot Motion Generation by Sequencing Movement Primitives Associated with Multimodal Inputs

Name
Edgar Anarossi
Laboratory’ｓ name
Robot Learning Laboratory
Supervisor’ｓ name
Takamitsu Matsubara

Abstract

We explore robotic motion generation, vital for enabling robots to translate abstract goals into physical actions across industrial automation, household tasks, and service applications. Despite progress, challenges remain in modeling complex, discontinuous motions with abrupt trajectory shifts and integrating multimodal inputs like vision and language. We build on movement primitives, modular frameworks for encoding motion, specifically leveraging Dynamic Movement Primitives (DMPs) as a prominent method in this domain. However, DMPs face limitations in handling sudden changes and diverse sensory data. Our research develops a framework that sequences DMPs to tackle two key issues: accurately generating complex motion such as discontinuous motions and enabling one-shot motion generation for tasks with limited observability, such as those involving occlusions. We propose two approaches: the first leverages deep learning to segment discontinuous motions into multiple DMPs, reducing computational demands and enhancing generalization across trajectories, and the second combines vision and language inputs through advanced reasoning models to dynamically select and sequence DMPs, improving adaptability in unpredictable environments. We validated these methods through experiments on tasks like object cutting, pick-and-place, and navigation in cluttered spaces, using both simulated datasets and real-world trials with physical robots. Evaluations focused on data efficiency, motion accuracy, and adaptability to occlusions, demonstrating significant improvements in generating precise, intent-aligned motions with minimal data. These evaluations confirm the potential of our approaches to advance robotic motion generation for complex real-world scenarios.