FRMD: Fast Robot Motion Diffusion with Consistency-Distilled Movement Primitives for Smooth Action Generation

Xirui Shi[1,2], and Jun Jin[1,2]
[1] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada..
[2] Alberta Machine Intelligence Institute (Amii), Edmonton, Canada..
Email: jjin5@ualberta.ca

INTRODUCTION

Diffusion models have recently emerged as a promising paradigm for robot action generation due to their scalability and ability to capture multimodal distributions [1], [5]. Despite this potential, their direct application to robotics is hindered by two key challenges. First, action-chunking variants generate short-horizon actions efficiently but fail to preserve temporal coherence, often resulting in jerky and unstable trajectories. Second, trajectory-level approaches such as Movement Primitive Diffusion (MPD) [2] leverage Probabilistic Dynamic Movement Primitives (ProDMPs) [3] to ensure smoothness, but require 10–50 iterative denoising steps, leading to prohibitive inference latency for real-time control.

To address this fundamental trade-off between efficiency and expressiveness, we introduce Fast Robot Motion Diffusion (FRMD), a framework that integrates ProDMPs with Consistency Models to enable smooth, structured, and temporally coherent motion generation in a single inference step. FRMD adopts a teacher–student distillation strategy, where a teacher diffusion model generates structured trajectories that are distilled into a student model capable of one-step inference. Experimental evaluations on twelve manipulation tasks from the MetaWorld and ManiSkill benchmarks demonstrate that FRMD achieves state-of-the-art success rates while reducing inference time by an order of magnitude, thereby enabling real-time robotic control.

MATERIALS AND METHODS

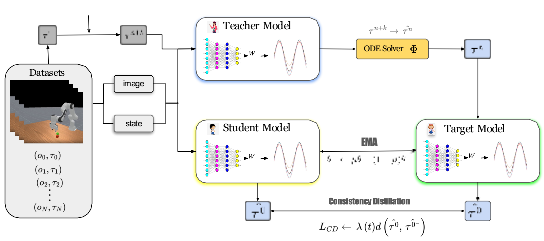

Fig. 1. **Overview of FRMD.** A teacher MPD produces structured ProDMP weights over multiple denoising steps. A student network is distilled to achieve single-step inference, preserving temporal coherence while eliminating latency.

FRMD unifies ProDMP-based trajectory representations with consistency distillation [4] to enable temporally consistent motion generation in a single step. ProDMPs encode trajectories as low-dimensional weight vectors with guaranteed boundary conditions, ensuring smooth execution.

During training, a teacher MPD generates structured trajectories through iterative denoising, while a student model is trained with a self-consistency objective to map noisy trajectories directly to ProDMP parameters in one step. At inference, FRMD samples Gaussian noise, predicts trajectory parameters once, and decodes them into smooth actions, thereby enabling closed-loop control at approximately 30Hz.

RESULTS AND DISCUSSION

FRMD was evaluated against Diffusion Policy (DP) and MPD on twelve tasks drawn from the MetaWorld and ManiSkill benchmarks, spanning easy (4)-medium (5)-difficult (3) levels of difficulty, in a total of 12 tasks. Task success. FRMD achieved the highest average success rate (64.8%), outperforming MPD (64.1%) and DP (50.1%). Notably, on medium-difficulty tasks, FRMD reached 66.3%, indicating that temporal structure was effectively retained through distillation. Inference efficiency. FRMD achieved an average infer- ence latency of 17.2ms, compared with 119.7ms for DP and 168.6ms for MPD, corresponding to a 7–10× improvement. This efficiency enables deployment in time-critical control loops. Motion smoothness. A curvature-based smoothness anal- ysis confirmed that FRMD substantially reduced non-smooth transitions relative to DP. In PlugCharger-v1 (Fig. 2), FRMD produced stable, oscillation-free trajectories, in contrast with the jerky behaviors observed under DP.

TABLE I: AVERAGE PERFORMANCE ACROSS 12 TASKS IN EASY-MEDIUM-DIFFICULT LEVELS.

Method	Success (%)	Inference (ms)
DP	50.1	119.7
MPD	64.1	168.6
FRMD	**64.8**	**17.2**

CONCLUSIONS

This work presents FRMD, a diffusion-based framework that leverages consistency-distilled movement primitives for fast and smooth robot motion generation. By combining trajectory-level representation with one-step inference, FRMD achieves state-of-the-art task performance while dra- matically reducing inference latency. These findings high- light the potential of FRMD as a general-purpose action decoder for real-time embodied AI and robotics.

REFERENCES

[1] C. Chi et al., “Diffusion Policy,” CoRL, 2023. [2] P. Scheikl et al., “Movement Primitive Diffusion,” IEEE RA-L, 2024.
[3] G. Li et al., “ProDMP,” IEEE RA-L, 2023.
[4] Y. Song et al., “Consistency Models,” NeurIPS, 2023.
[5] K. Black et al., “π0: A Multimodal Robot Foundation Model,” arXiv, 2024