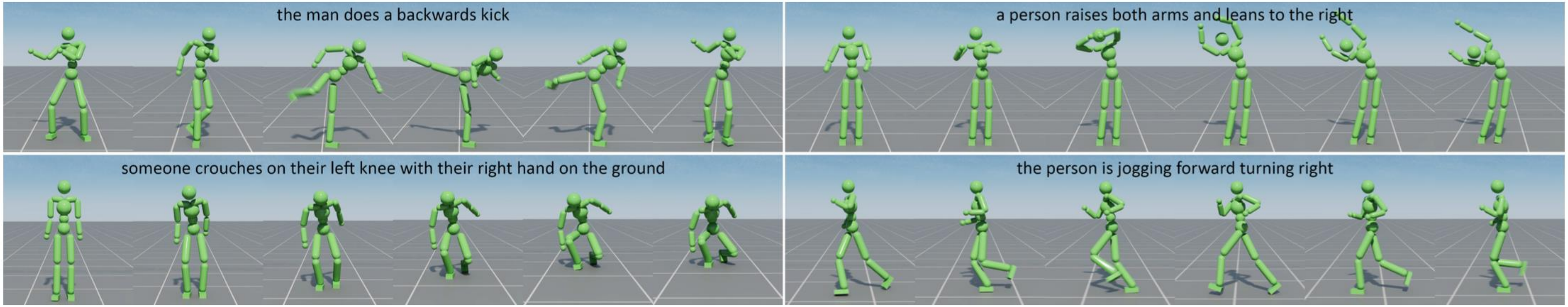


SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation

SIGGRAPH 2024

Jordan, Juravsky Yunrong Guo, Sanja Fidler, Xue Bin Peng
NVIDIA

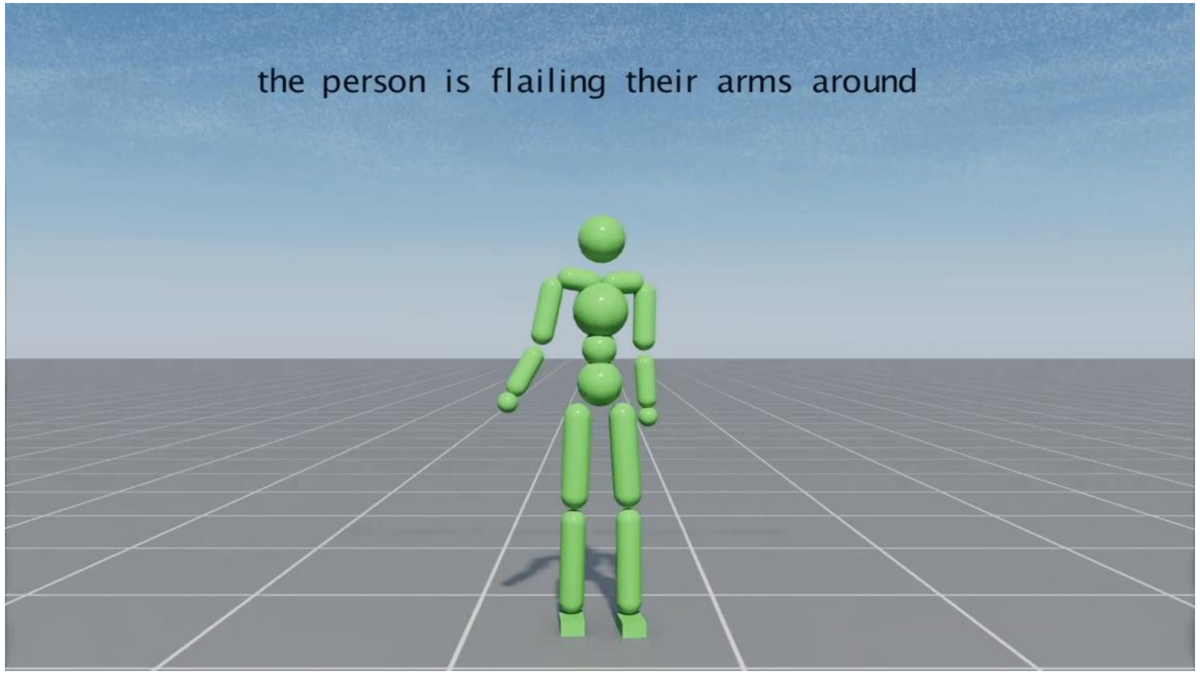
NVIDIA

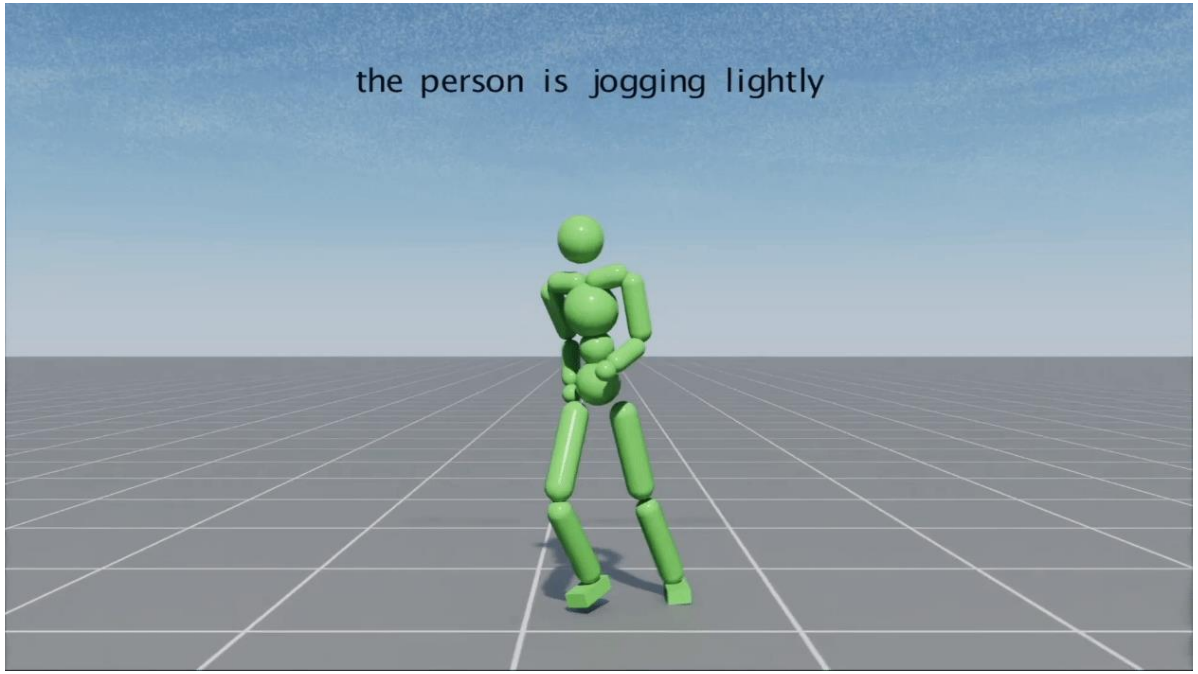
Contents

01 Introduction

Introduction – Goal

Scaling Language-Directed Physics-based Control with Progressive Supervised Distillation



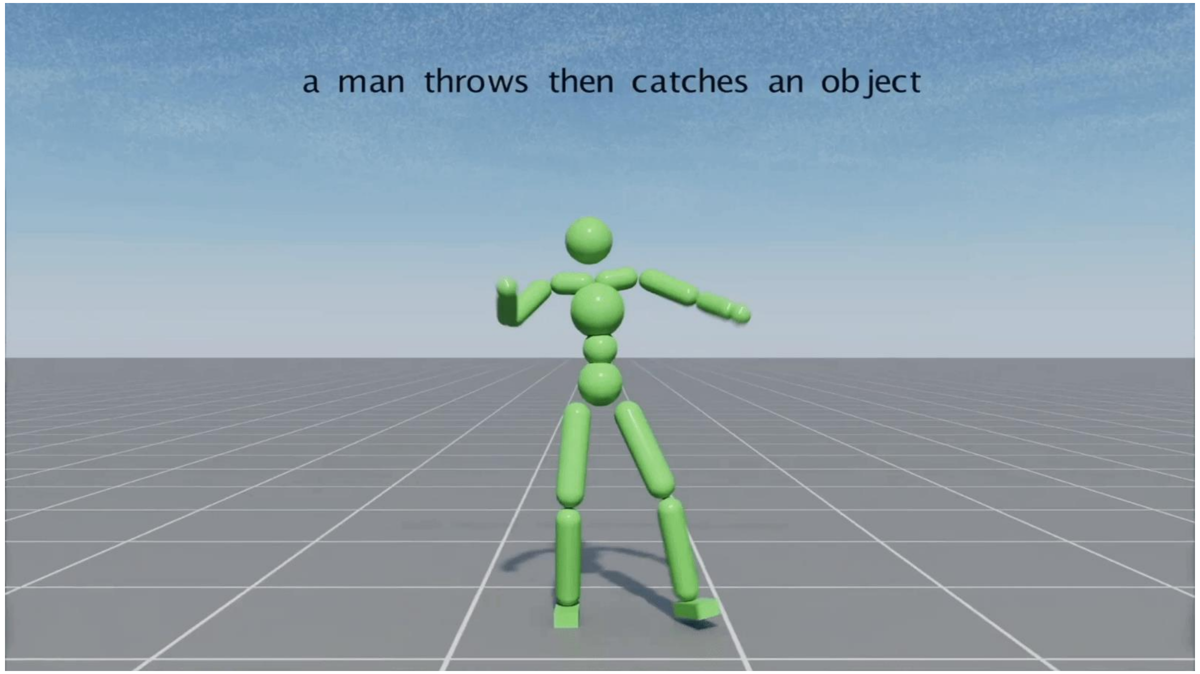


Command : the person is flailing their arms around

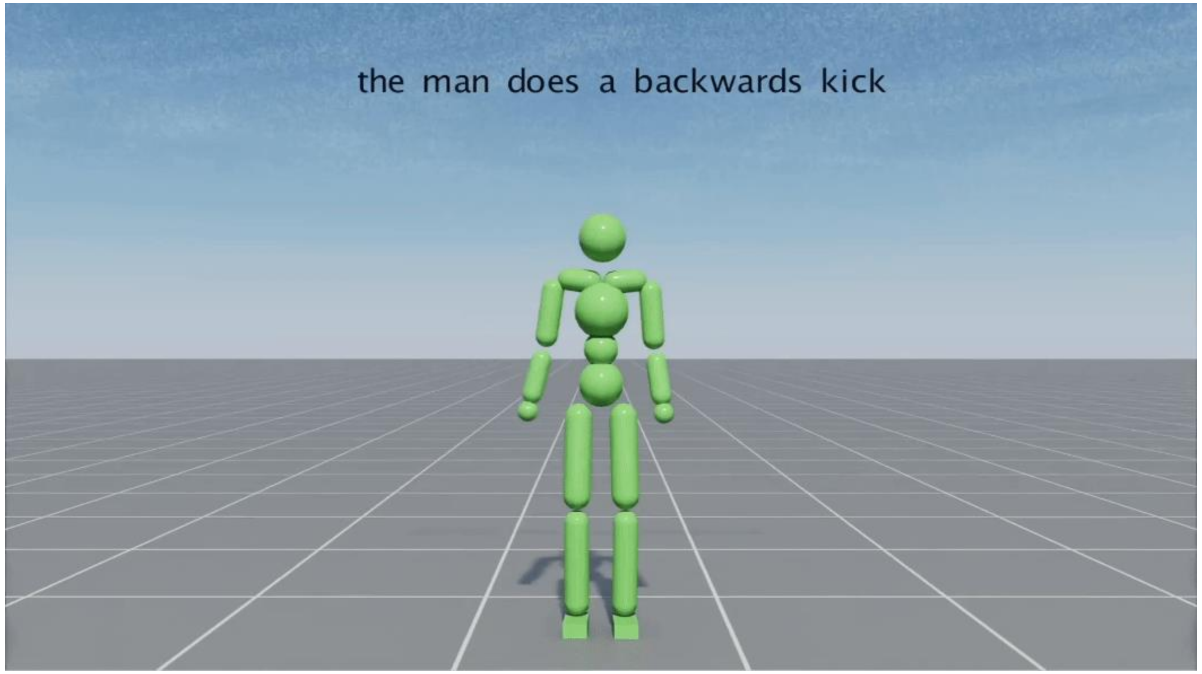
Command : the person is jogging lightly

Introduction – Goal

Scaling Language-Directed Physics-based Control with Progressive Supervised Distillation



Command : a man throws then catches an object



Command : the man does a backwards kick

02 Background

Background – Character animation Model

<Physics-based animation Model>

Advantage

- Realistic movement
- Enable realistic responses to perturbations and environmental variation

Disadvantage

- Can not scale beyond at most several hundred motions

Method

Reinforcement Learning

<Kinematic-based animation Model>

Advantage

- Can scale beyond at most several hundred motions

Disadvantage

- Disable realistic responses to perturbations and environmental variation

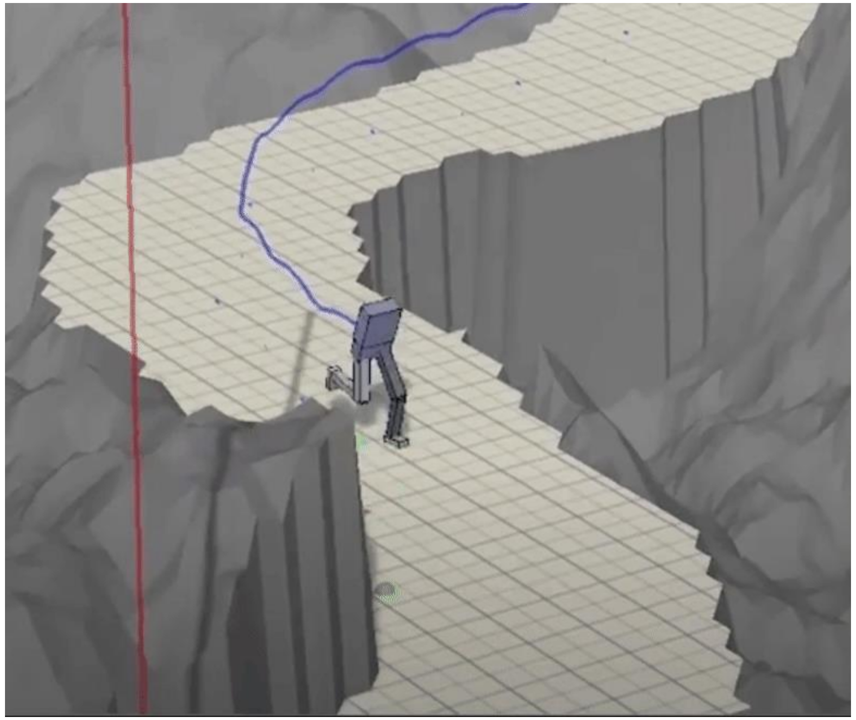
Method

Supervised Learning

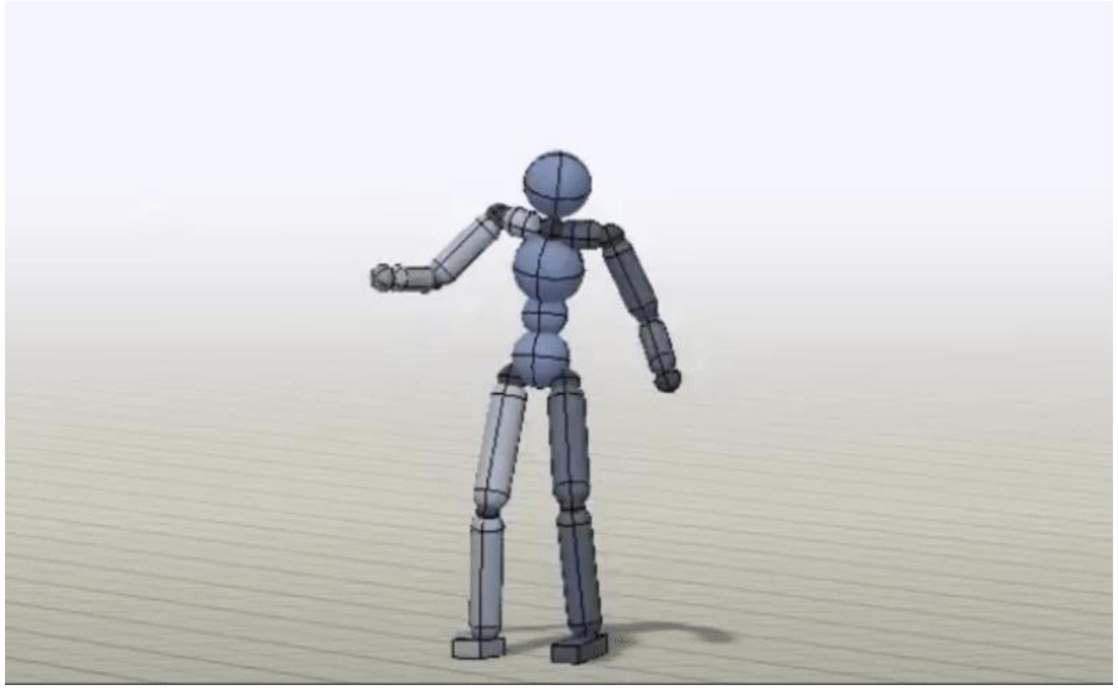
Background – Character animation Model

Physics-based animation model

Method : Deep Reinfocement Learning



<Deeploco, *ACM Transactions on Graphics (TOG)*, 2017>

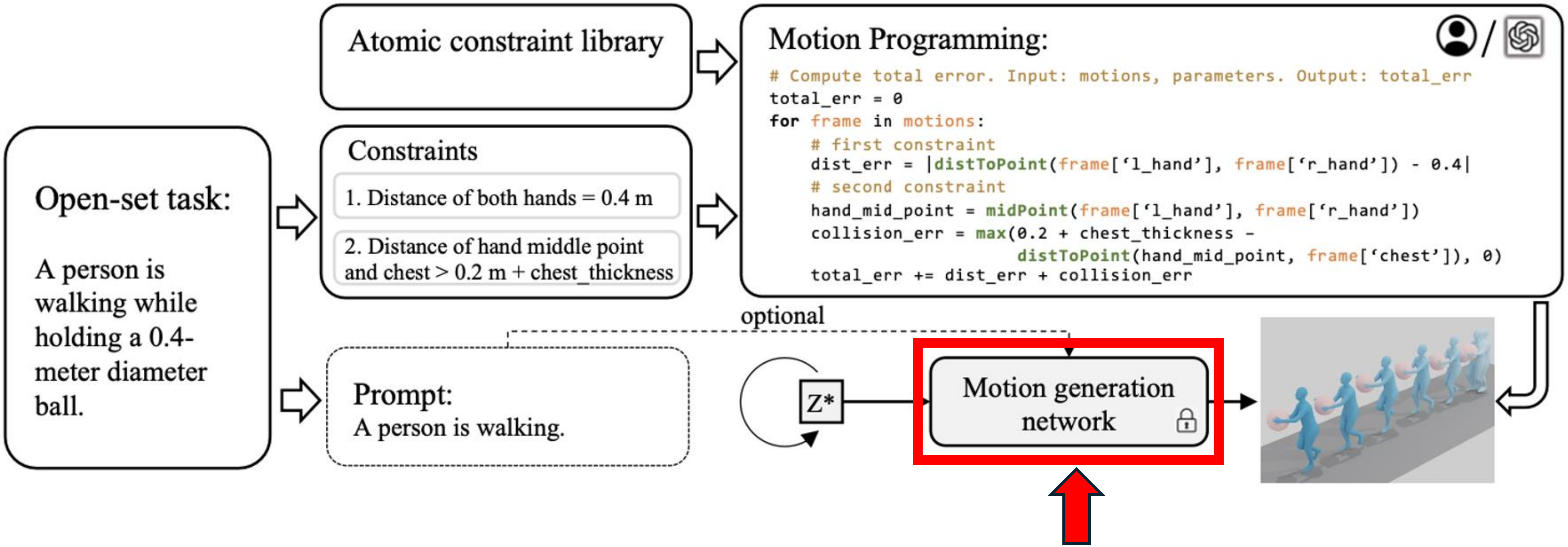


<Deepmimic, *ACM Transactions On Graphics (TOG)*, 2018>

Background – Character animation Model

Kinematic-based animation model

Method : Diffusion Model(Supervised Learning)



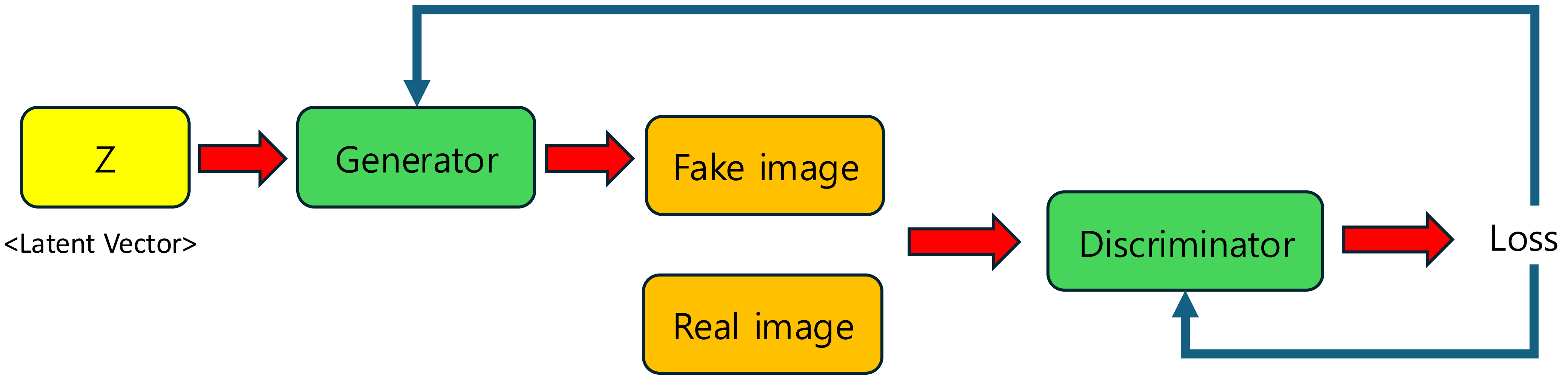
<Programmable Motion Generation for Open-Set Motion Control Tasks, CVPR 2024>

Background – Adversarial Learning

Generative Adversarial Networks(GAN)

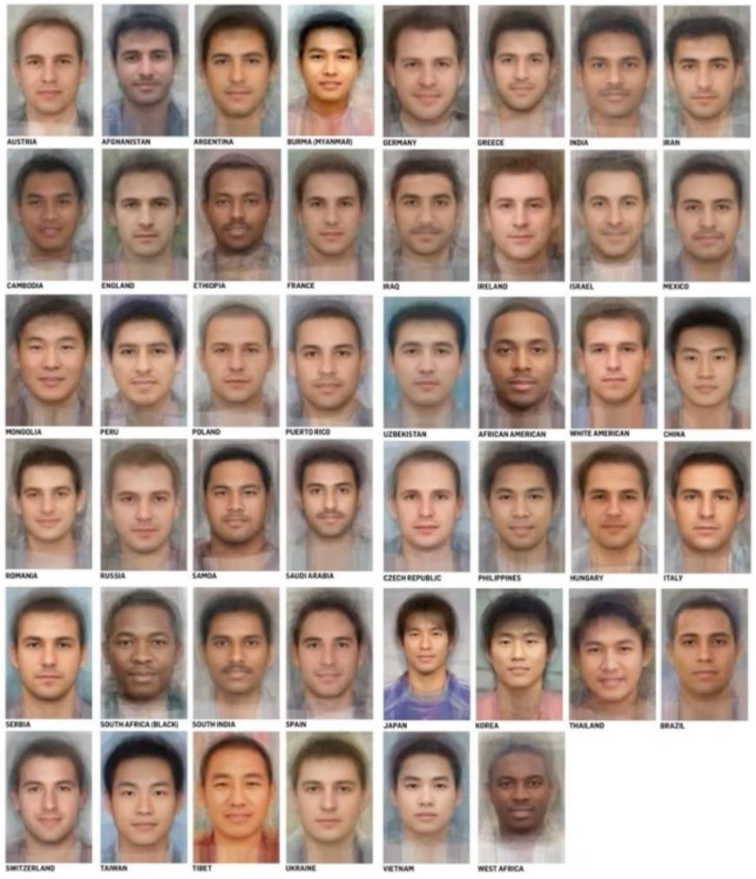
Generator : Generate Image

Discriminator : Probability (Real: 1 ~ Fake: 0)

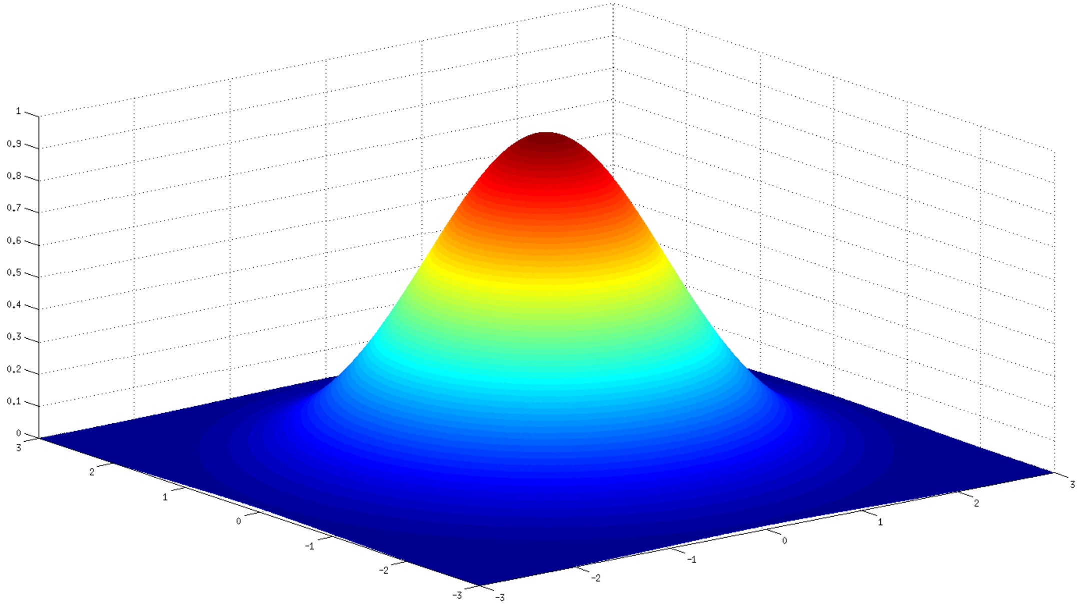


Background – Adversarial Learning

Generative Adversarial Networks(GAN)



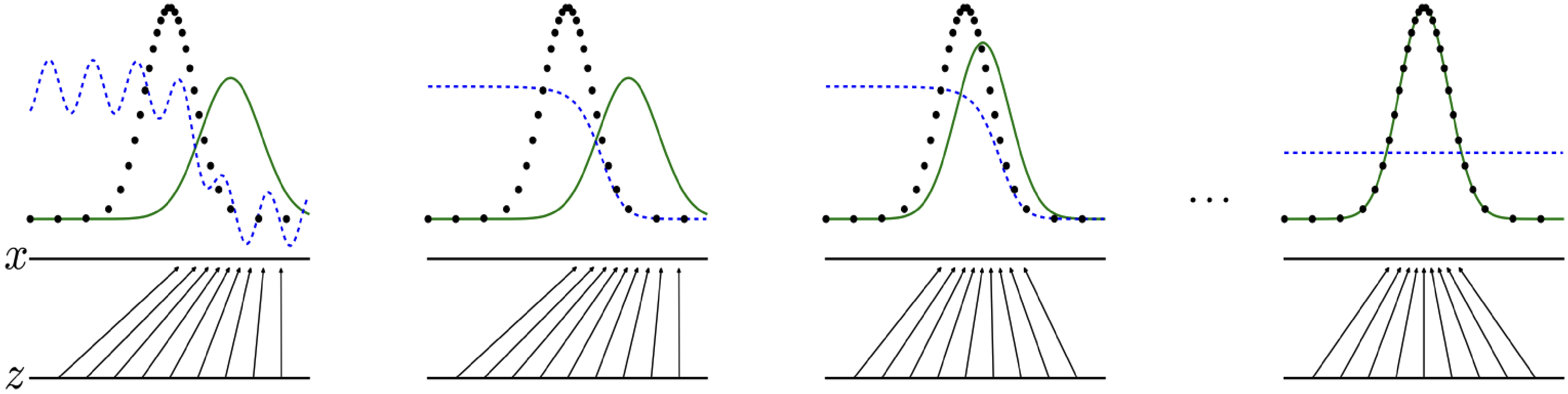
<Image : 64 x 64>



<Distribution : 64 x 64 x 3>

Background – Adversarial Learning

Training of GAN



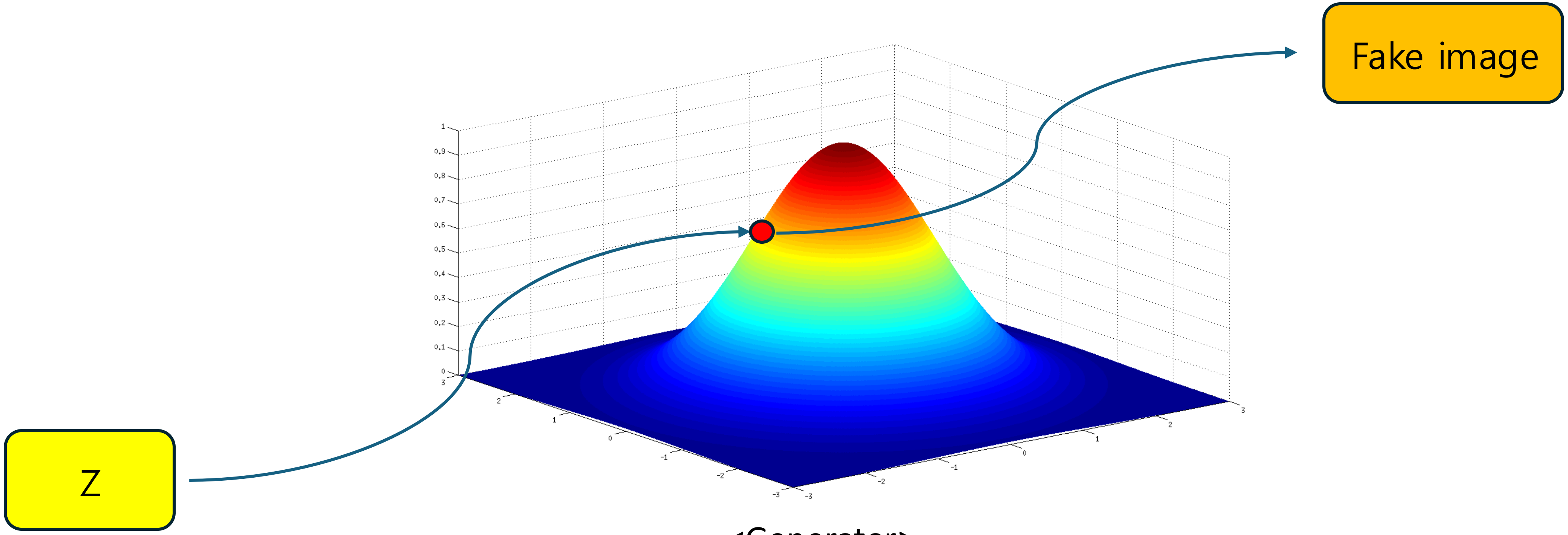
Distribution of Generative model

Distribution of Real data

Distribution of Discriminator

Background – Adversarial Learning

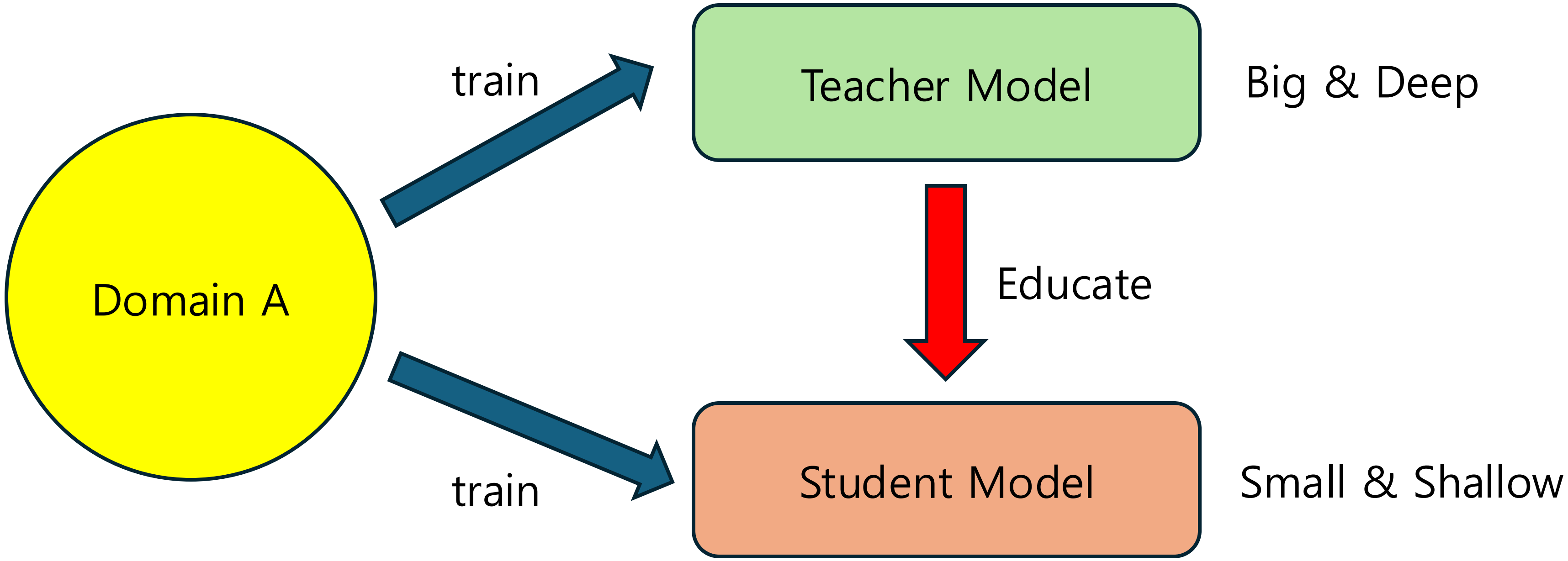
Generative Adversarial Networks(GAN)



<Generator>

Background – Knowledge Distillation

Knowledge Distillation



Advantage

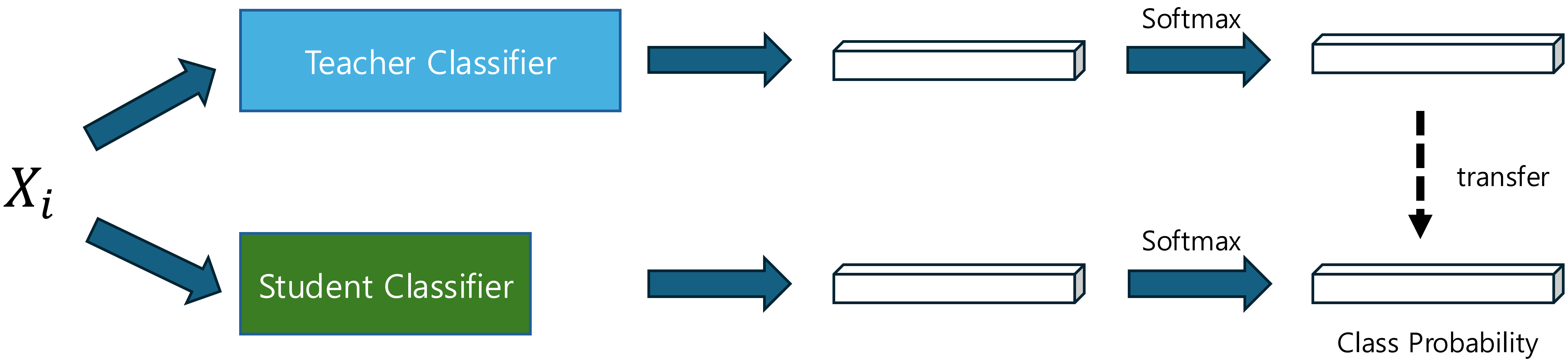
- Model Compression
- Low Computational Complexity
- Boosting Performance

Which Knowledge?

- Hidden Activation(Weight)
- Class Probability

Method: Class Probability

Distilling the Knowledge in a Neural Network, Hinton & Jeff Dean et al, NIPS workshop, 2024

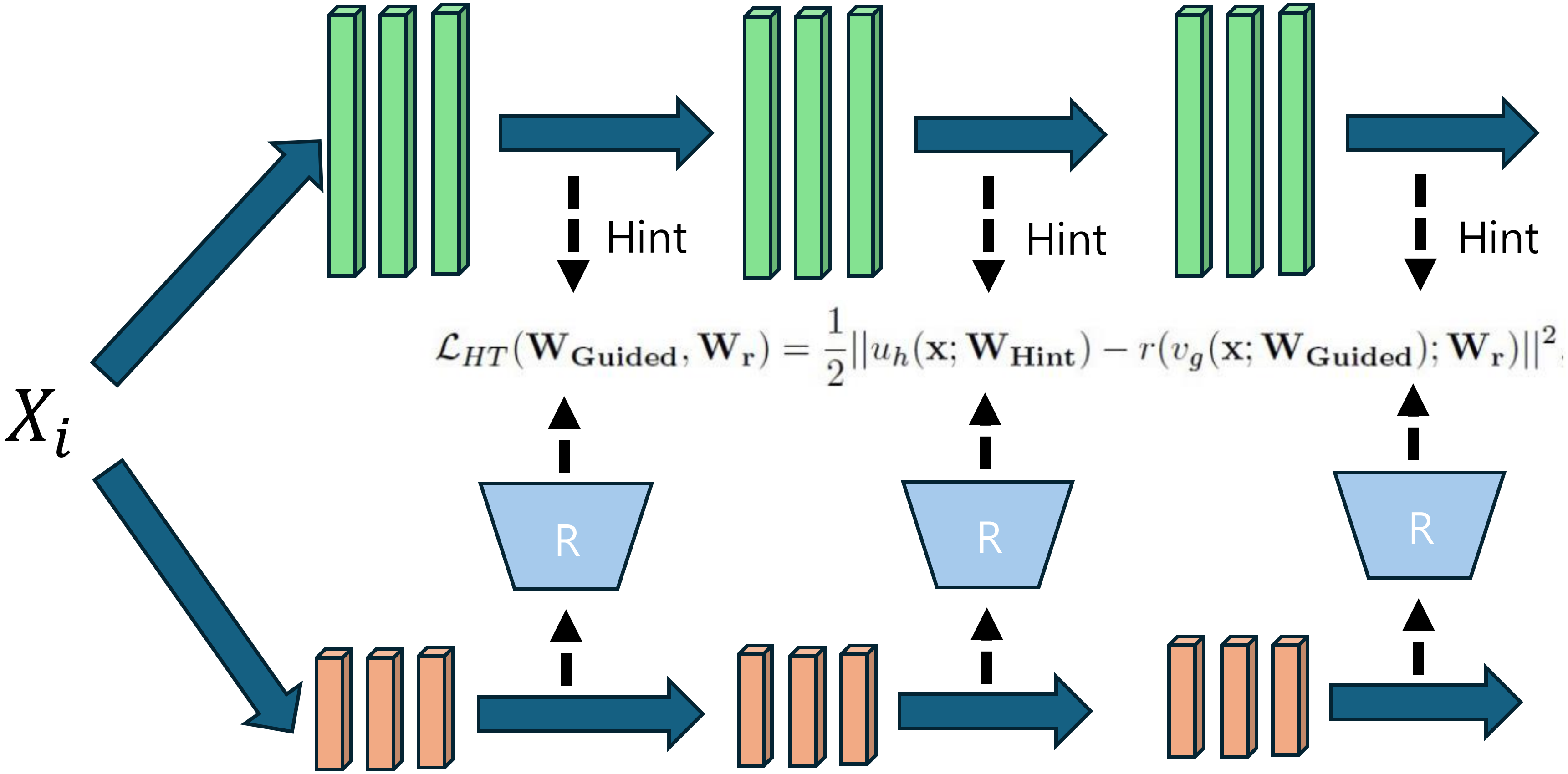


Background – Knowledge Distillation

Method: Hidden Activation(Weight)

FitNets: Hints for Thin Deep Nets, Romero & Bengio et al, ICLR, 2015

<Pre-trained teacher network(Hint)>

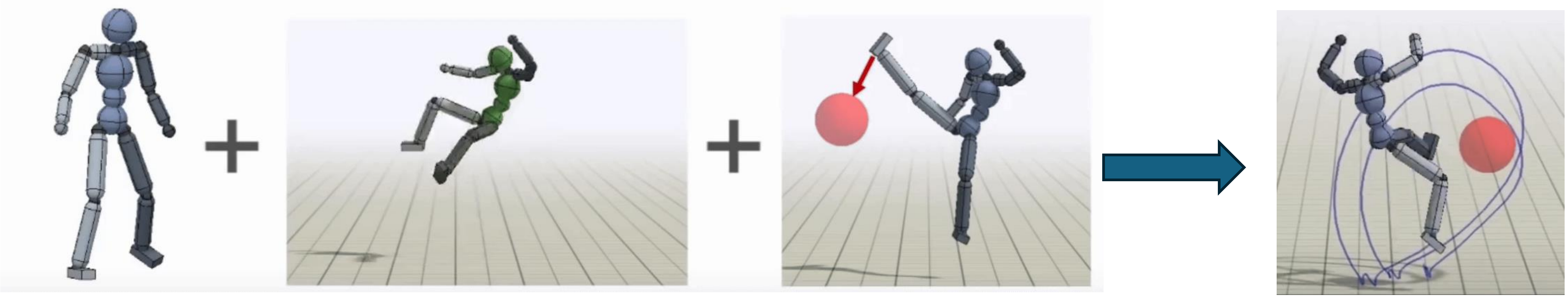


<Student network(Guided)>

Background - DeepMimic

DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills

SIGGRAPH 2018

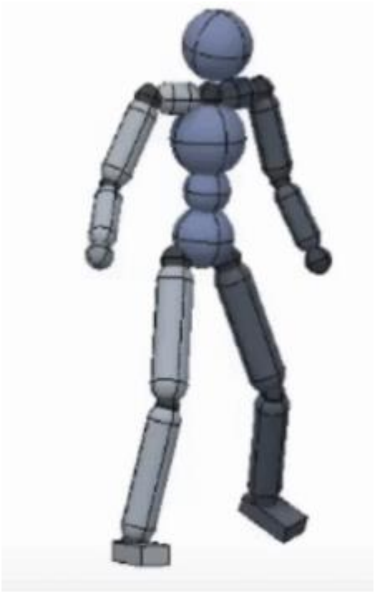


<Character> <Reference motion> <Task: Hit target> <Simulated motion>

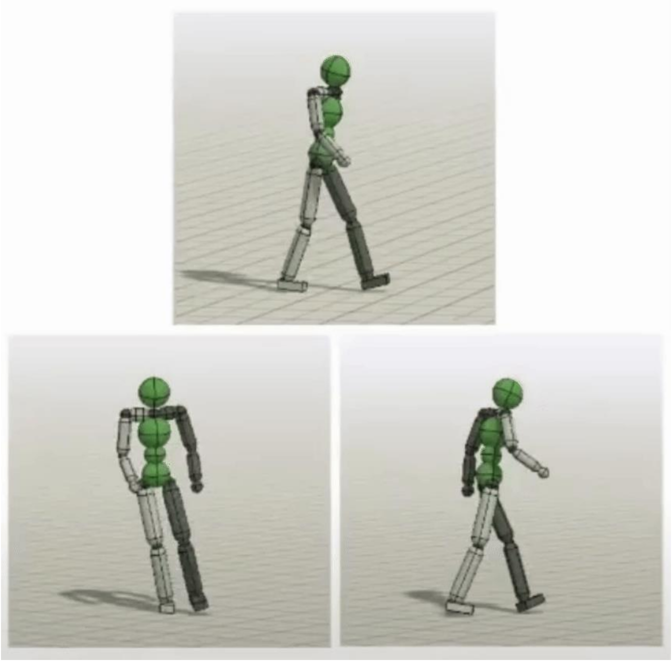
Background - DeepMimic

DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills

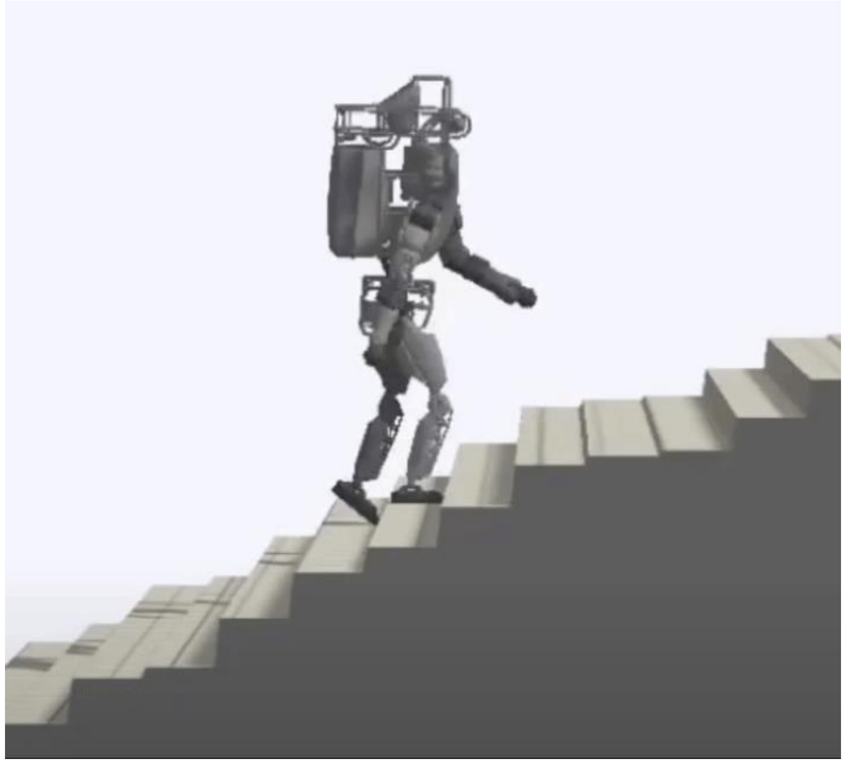
SIGGRAPH 2018



<Character>



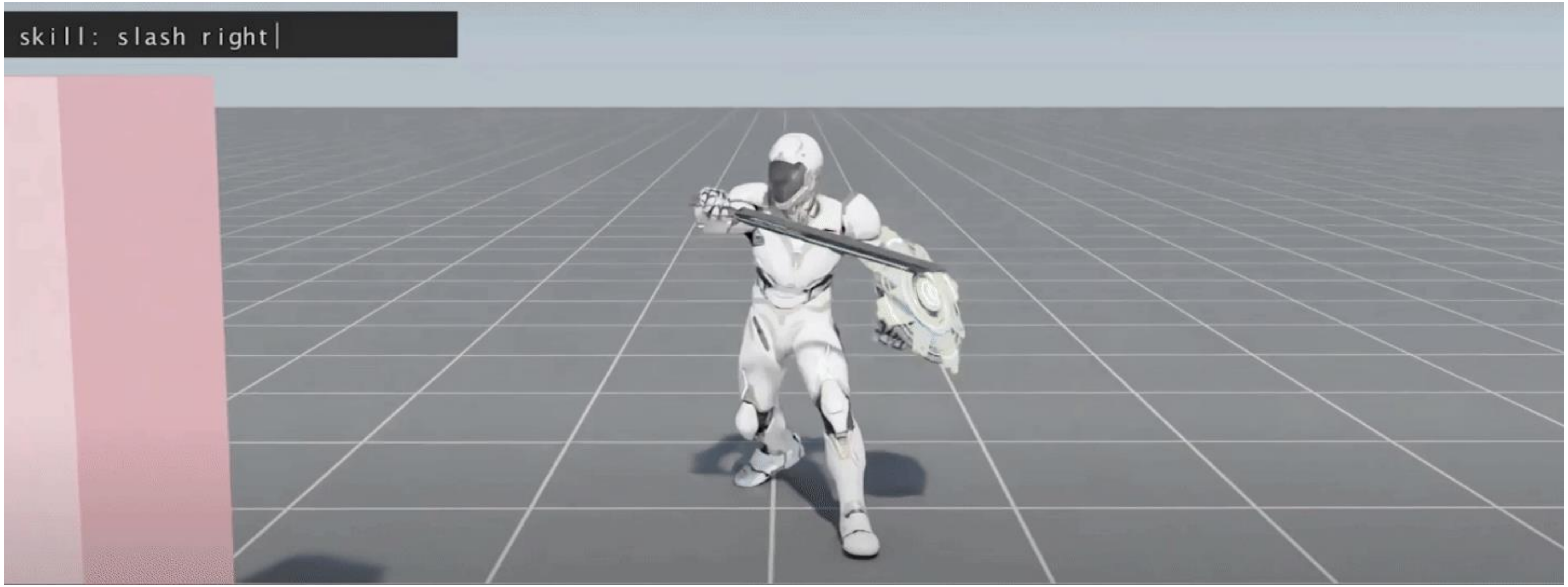
<Reference motion: walking>



<Simulated motion: Physics-based>

Background – PADL

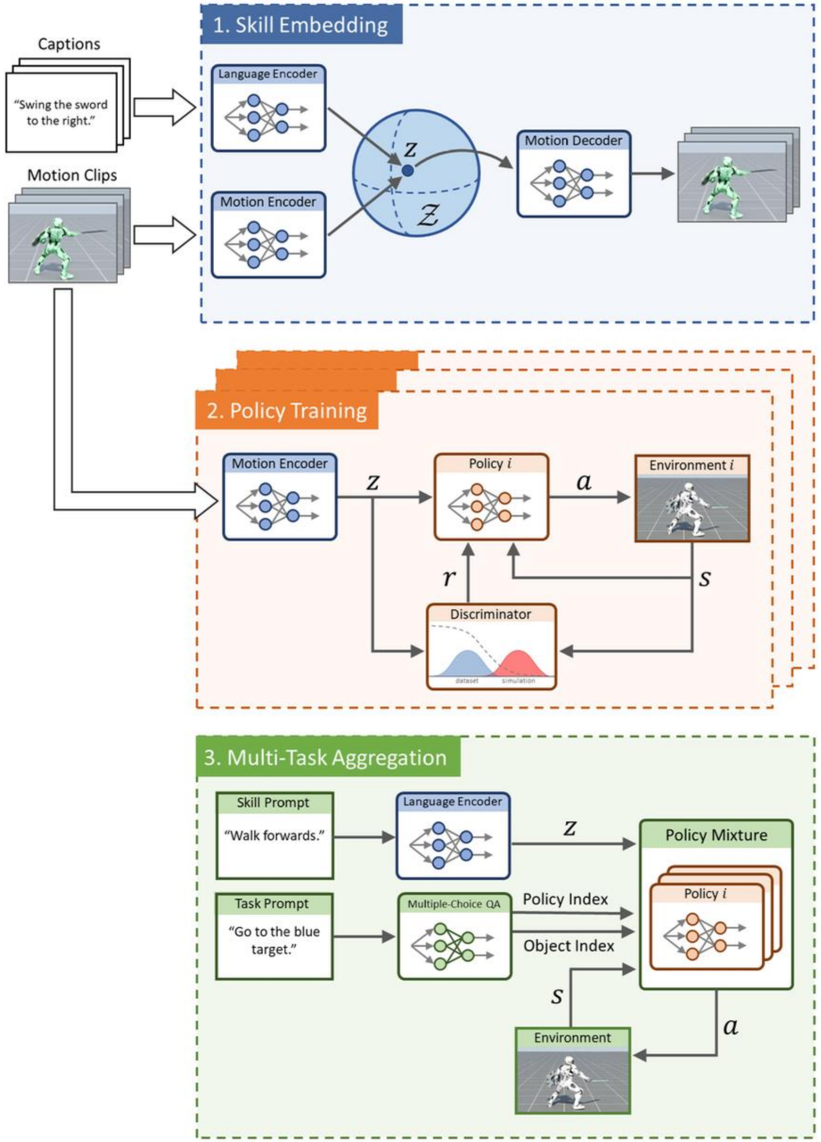
PADL: Language-Directed Physics-Based Character Control, SIGGRAPH Asia 2022 Conference

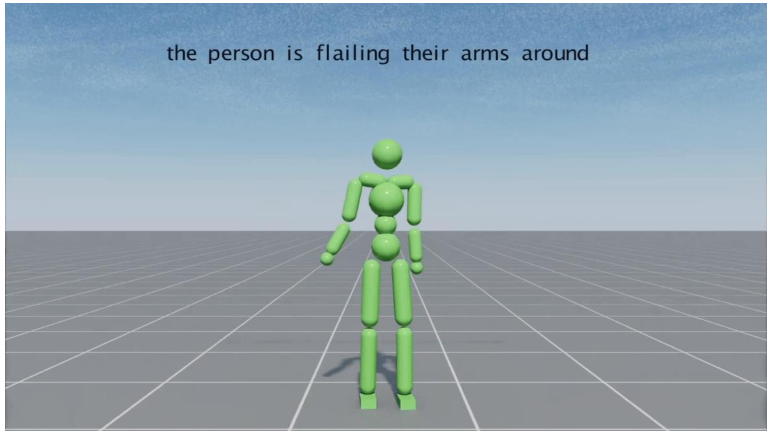


Task : Slash right

Background – PADL

PADL: Language-Directed Physics-Based Character Control, SIGGRAPH Asia 2022 Conference





Data Set : $\{(M_i, C_i)\}$

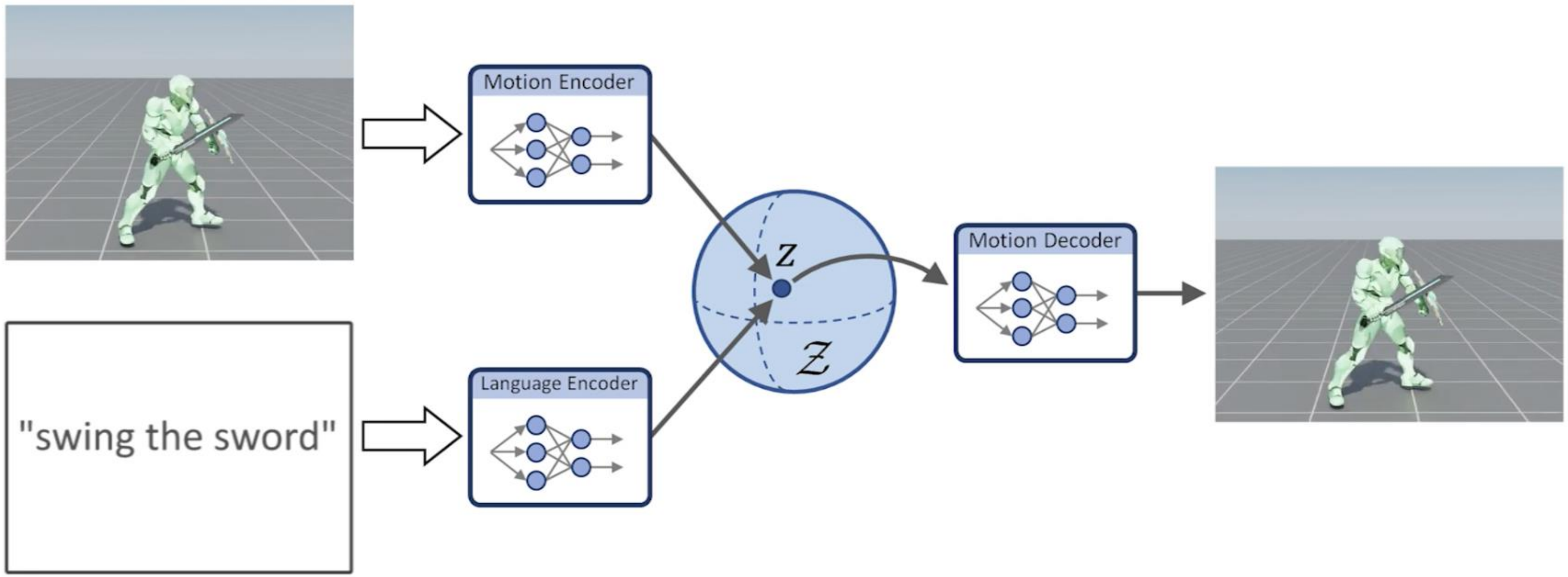
M : Motion, C : Caption

Skill Embedding -> Policy Training -> Multi-Task Aggregation

Background – PADL

PADL: Language-Directed Physics-Based Character Control, SIGGRAPH Asia 2022 Conference

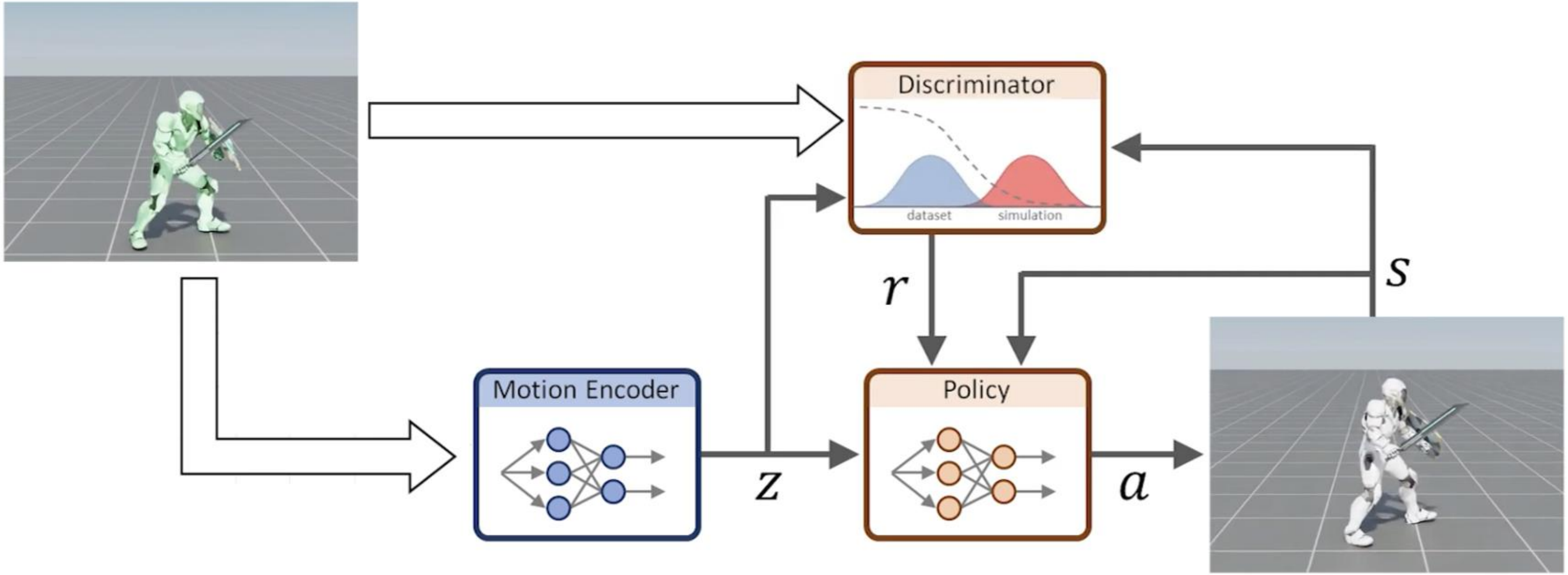
Step1: Skill Embedding(mapping)



Background – PADL

PADL: Language-Directed Physics-Based Character Control, SIGGRAPH Asia 2022 Conference

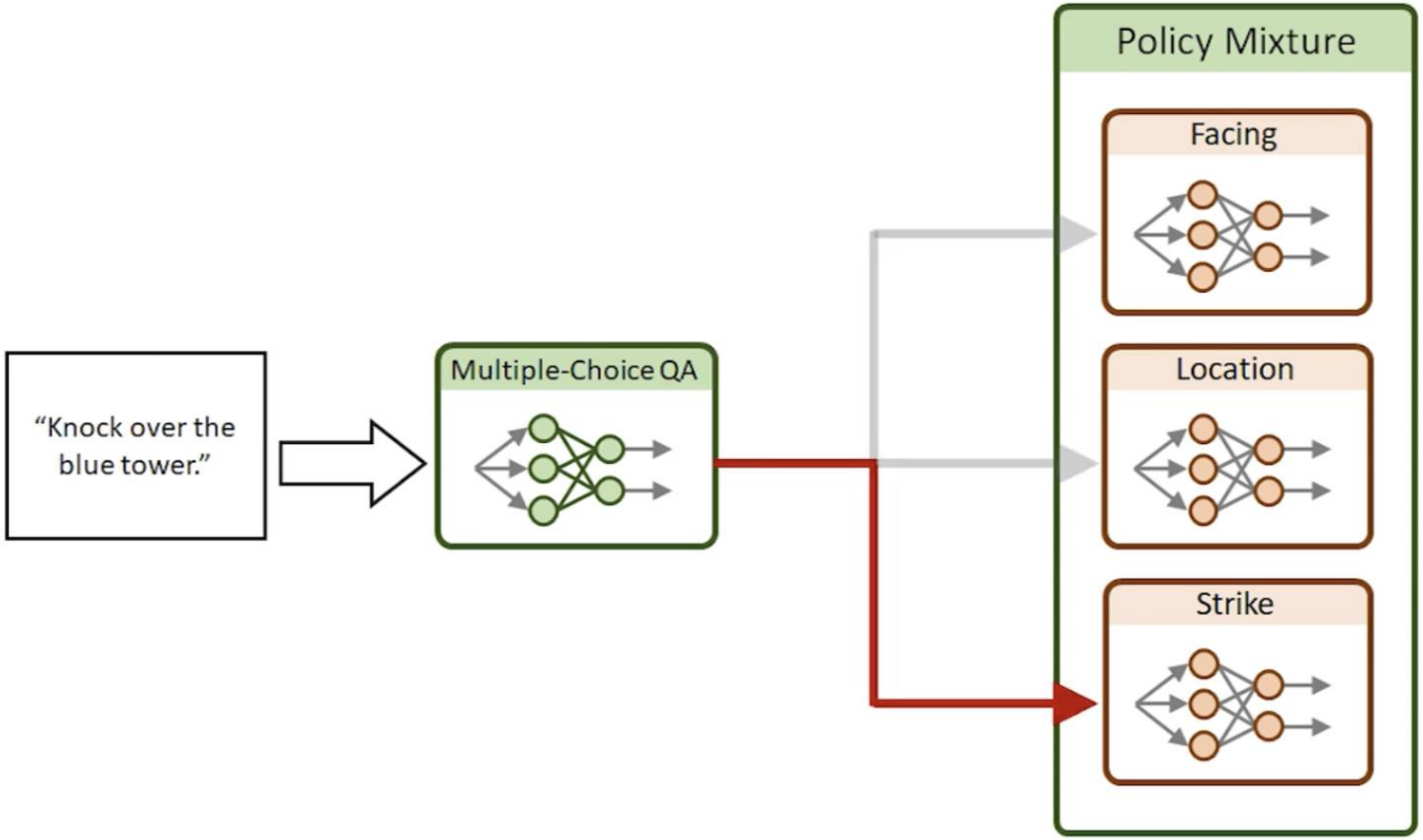
Step2: Policy Training(Adversarial Network Architecture)



Background – PADL

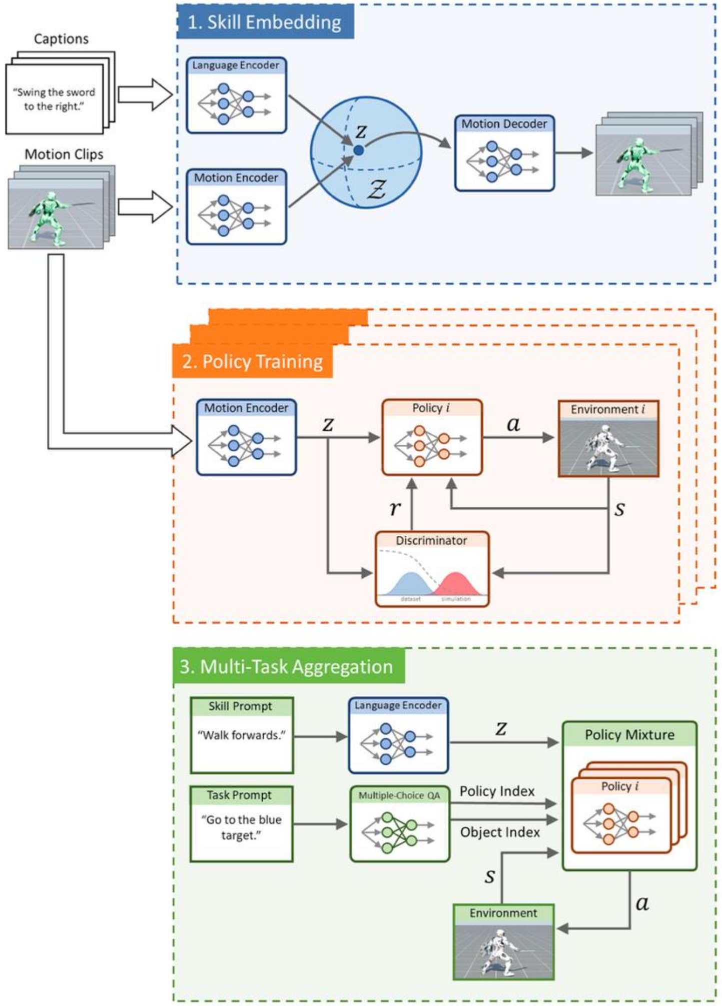
PADL: Language-Directed Physics-Based Character Control, SIGGRAPH Asia 2022 Conference

Step3: Multi-Task Aggregation

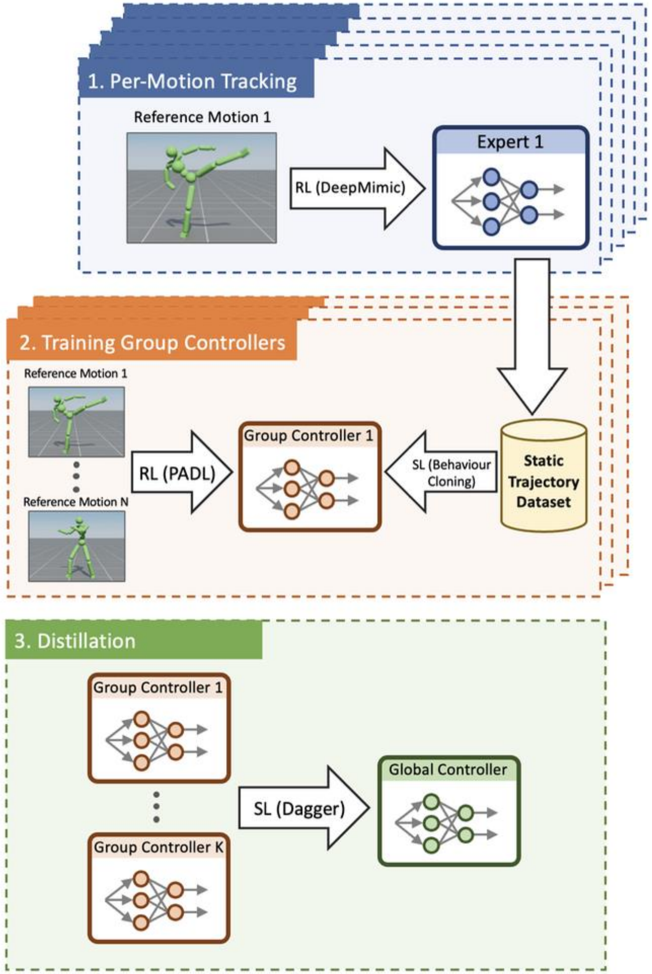


Background – PADL

SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation



PADL

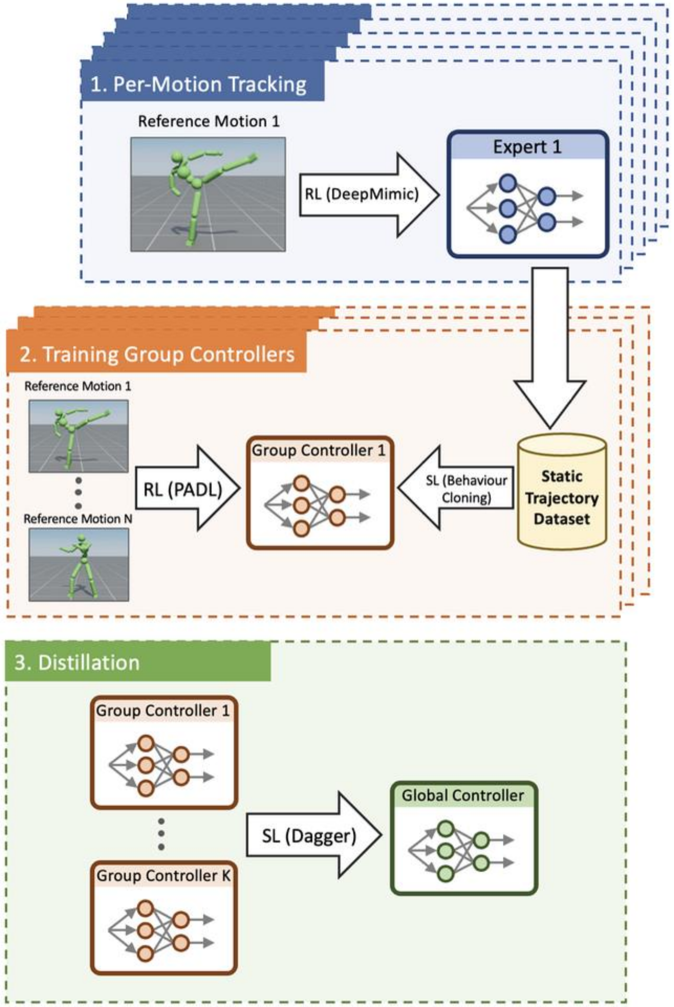


SuperPADL

03 Method

Method - Overview

SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation



Inspired by two observations

RL Method(**Physics-based Model)**

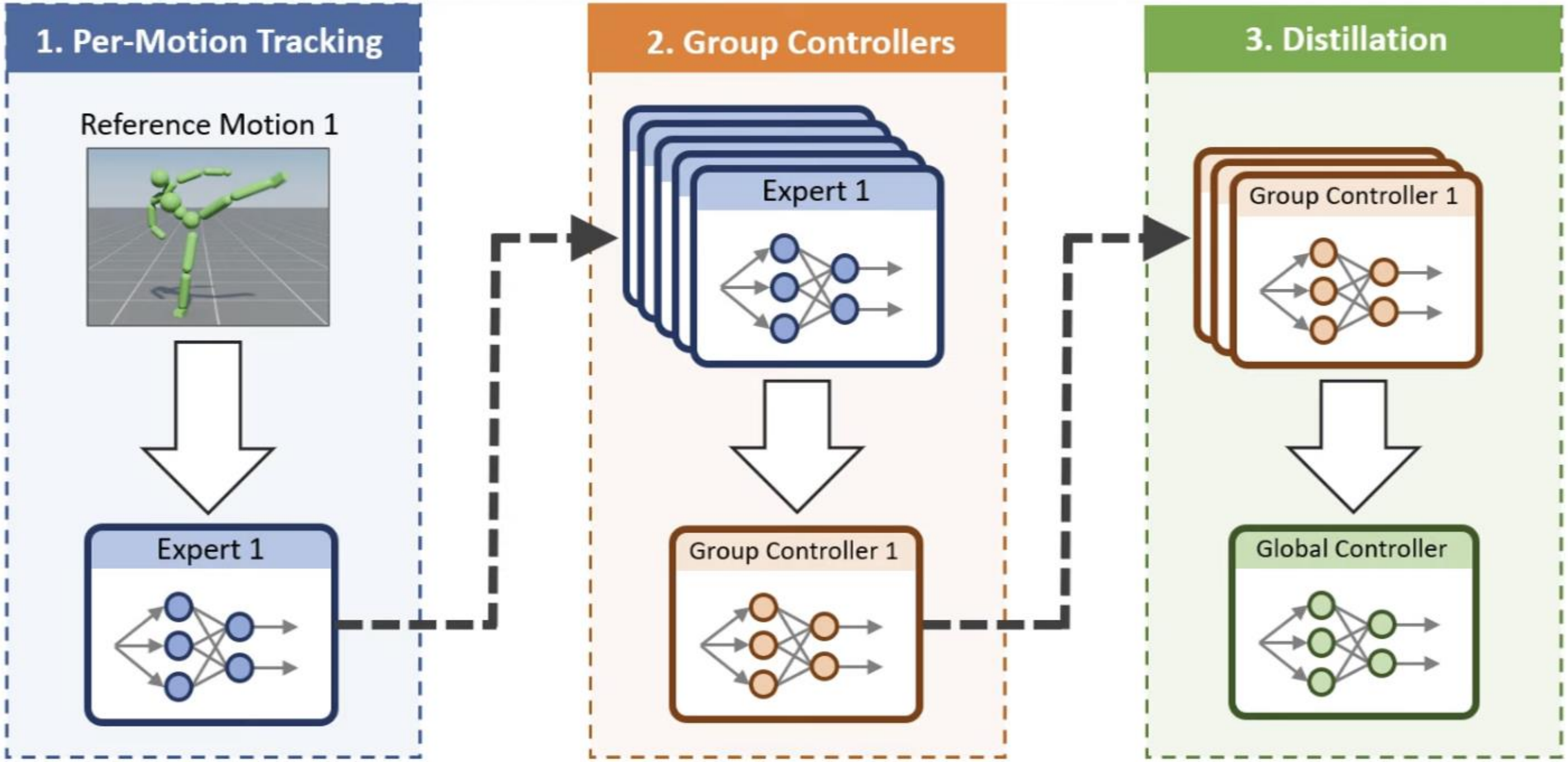
+ able to produce motions with high quality and naturally transition between skills
- can't scale beyond at most several hundred motions

Supervised Method(**Kinematic Model)**

+ can scale to datasets containing thousands of motions using supervised learning
- can't produce motions with high quality and naturally transition between skills

Method - Overview

SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation



RL

RL + Supervised Learning

Supervised Learning

Method – Per motion tracking

Training Per-Motion Expert Tracking Policies

Data Set : D = $\{(M_i,\ C_i)\}$

M : Motion capture sequences, C : Caption

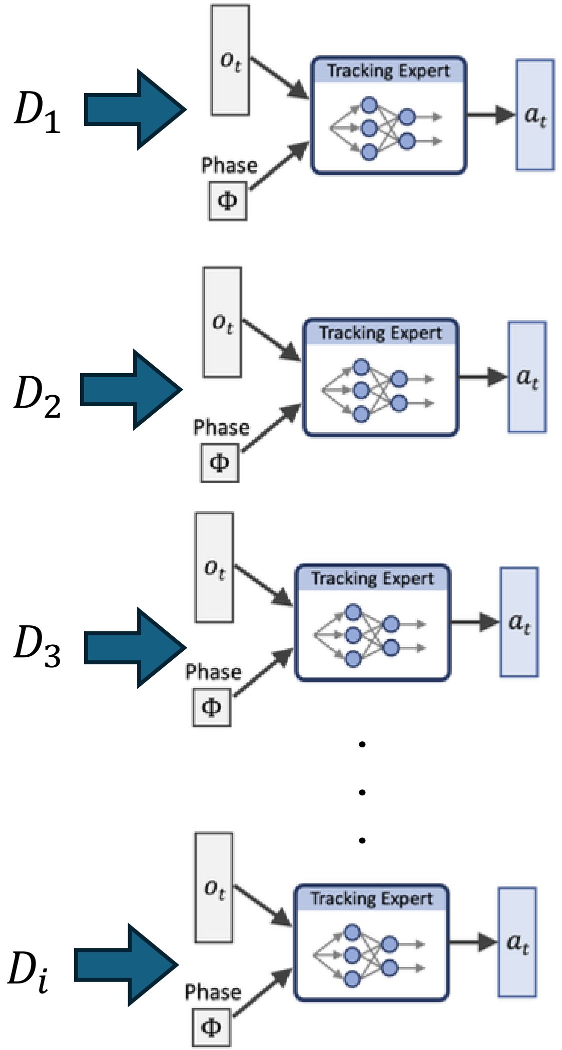
D → DeepMimic → D'

Kinematic-domain → Physics-domain

$$\pi_i^e(\mathbf{a}_t|\mathbf{o}_t, \phi)$$

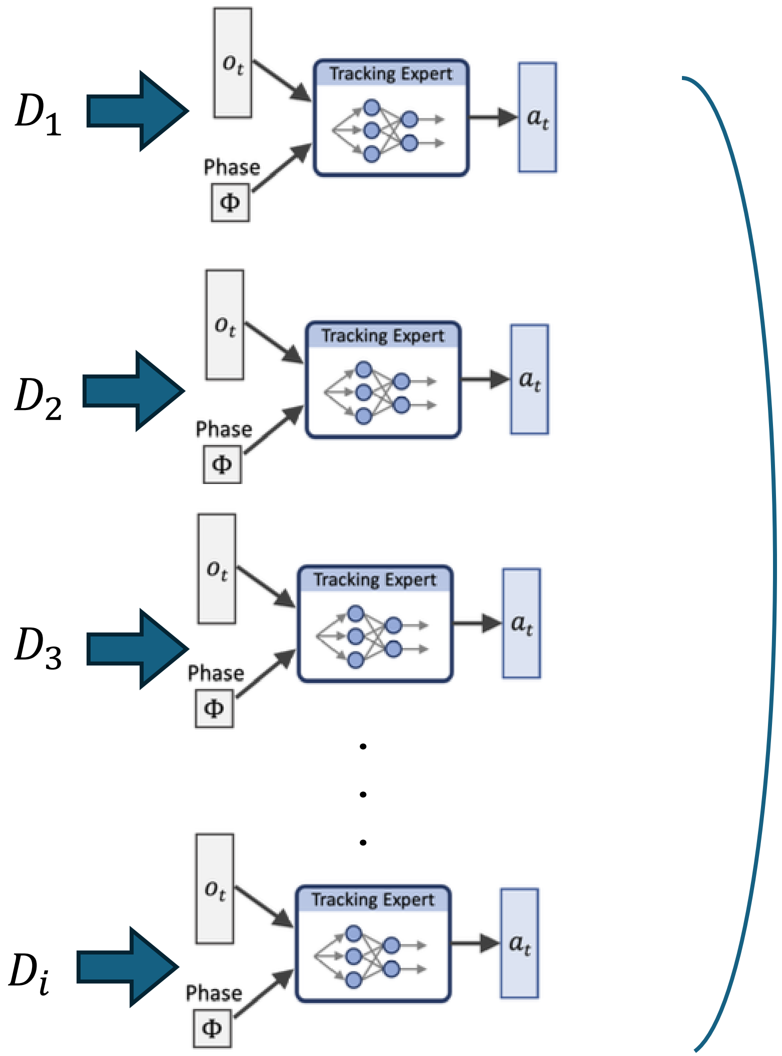
O_t : Current state of character

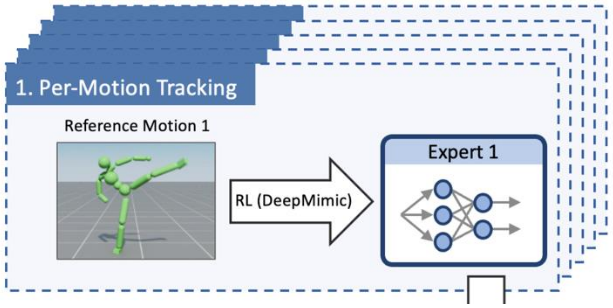
∅ ∈[0, 1] : phase variable that synchronizes the policy to the reference motion



Method – Per motion tracking

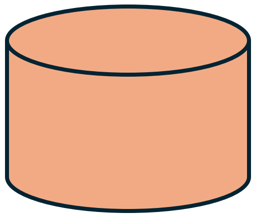
Training Per-Motion Expert Tracking Policies





$$T_i = (\mathbf{o}_1^i, \mathbf{a}_1^i, \mathbf{o}_2^i, \mathbf{a}_2^i, \ldots)$$

Trajectories(each expert)

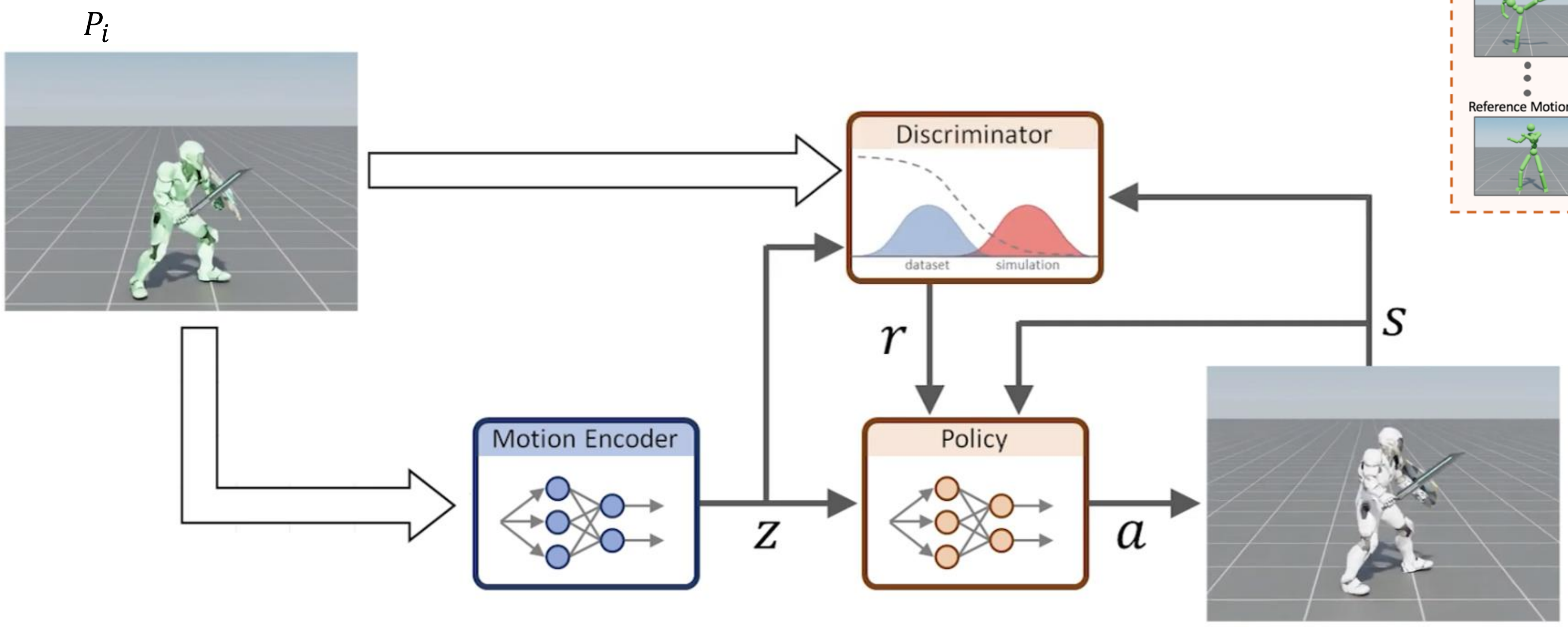


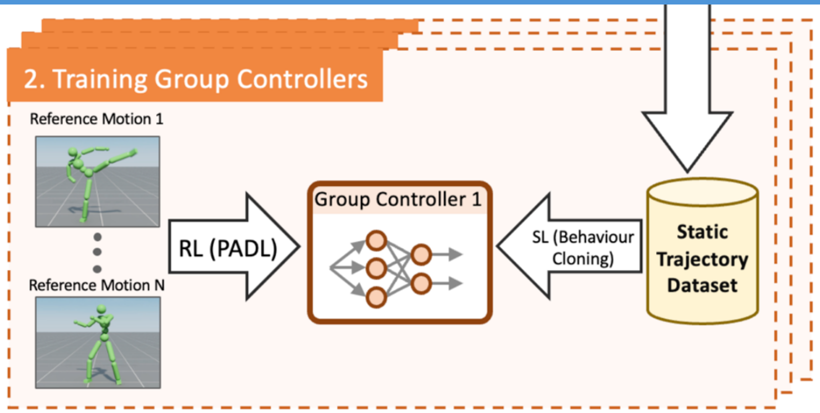
Trajectory Dataset

$$D_T = \{T_i\}$$

Method – Group Controllers

Training Group Controllers with PADL + BC



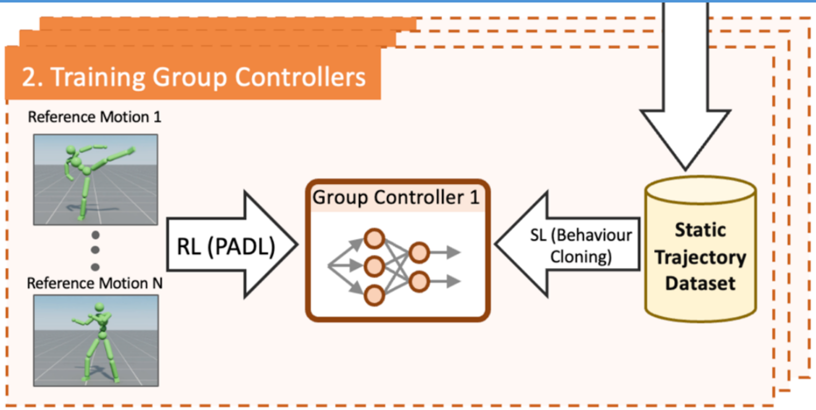


PADL Method

$P_i = \{(m_{20i+1}, C_{20i+1}), (m_{20i+2}, C_{20i+2}), \ldots, (m_{20i+20}, C_{20i+20})\}$: each group(randomly partition the dataset into groups of 20 motions

Method – Group Controllers

Training Group Controllers with PADL + BC



Group 1 → $\pi_1^g(a_t|o_t, I)$

Group 2 → $\pi_2^g(a_t|o_t, I)$

Group 3 → $\pi_3^g(a_t|o_t, I)$

Group 4 → $\pi_4^g(a_t|o_t, I)$

$I \in \{20i + 1, \ldots, 20i + 20\}$

BC: Supervised Learning

Expert

Trajectory Dataset: label

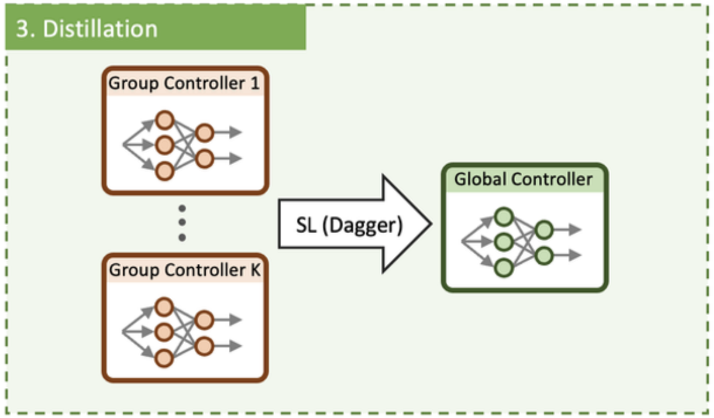
$$\mathcal{L}_{\mathrm{BC}} = \mathop{\mathbb{E}}_{I \sim \{20i+1,\ldots,20i+20\}} \mathop{\mathbb{E}}_{(\mathbf{o},\mathbf{a}) \sim T_I} ||\pi_i^g(\mathbf{o}, I) - \mathbf{a}||_2^2$$

Goal

- Naturally transition between these motions when the input index changes

Method – Distillation

Distilling into a Global Text-Conditioned Policy



$$\pi_1^g(a_t|o_t, I)$$
$$\pi_2^g(a_t|o_t, I)$$
$$\pi_3^g(a_t|o_t, I)$$
$$\cdots$$
$$\pi_i^g(a_t|o_t, I)$$

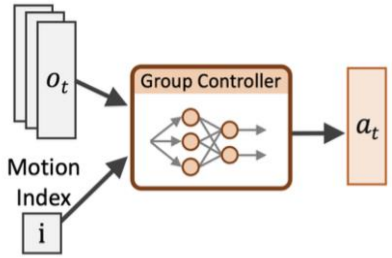
$$\pi^G(a_t|o_t, c)$$

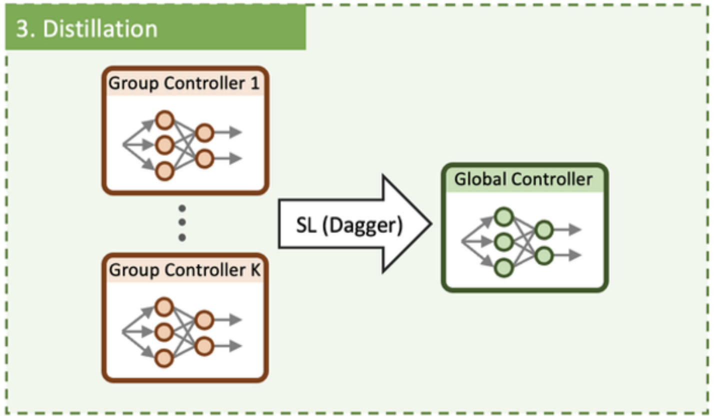
How?

<Teacher Polices> <Student Policy>

Method – Distillation

Distilling into a Global Text-Conditioned Policy

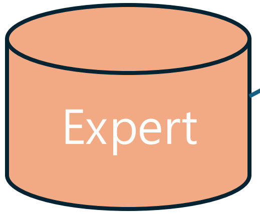




\+ caption(C)

$\pi^G(a_t|o_t, c)$

<Student Policy>



Trajectory Dataset

Supervised Learning

Clip

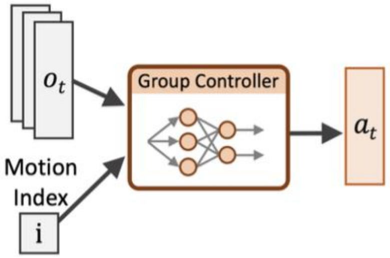
caption, motion -> Latent vector

o_t

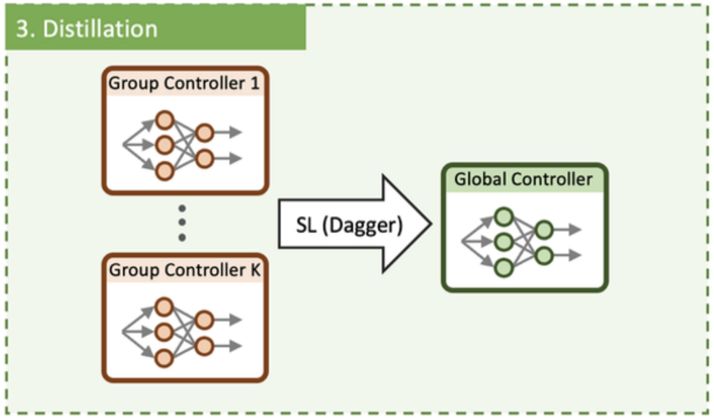
$\pi^G(a_t|o_t, c) \longrightarrow a_t$

Method – Distillation

Distilling into a Global Text-Conditioned Policy



+ caption(C)



$\pi^G(a_t|o_t,c)$

<Student Policy>

Supervised Learning

Expert

Trajectory Dataset

Supervised Learning

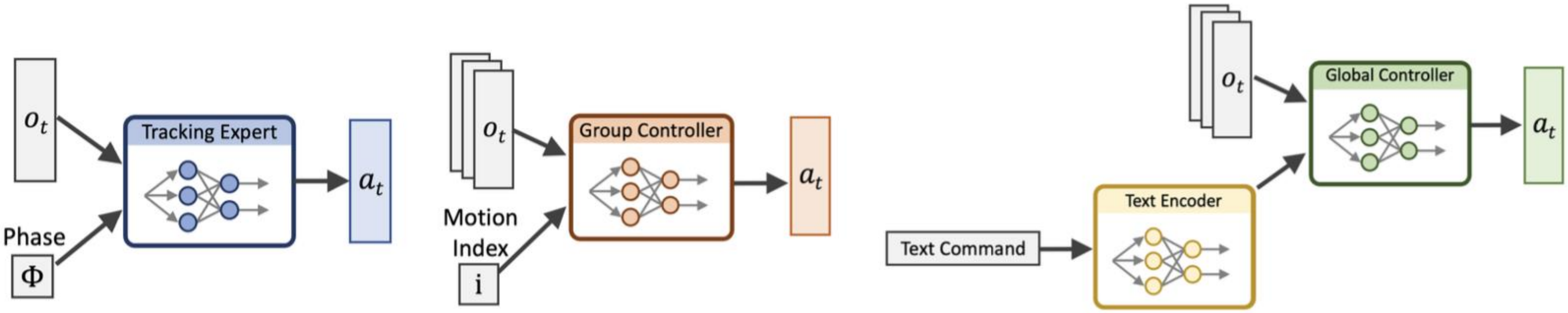
Clip

caption, motion -> Latent vector

o_t

$\pi^G(a_t|o_t,c) \rightarrow a_t$

Method



(a) Expert Controller (b) Group Controller (c) Global Controller

04 Evaluation

Evaluation – Quantitative

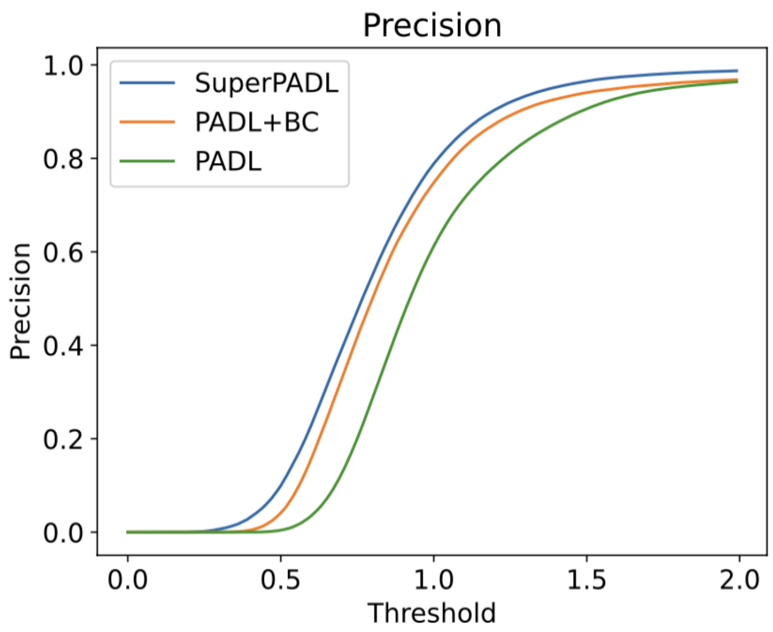
$$\mathrm{Rec}(\tau, \hat{\mathbf{m}}, \epsilon) = \frac{1}{n-9} \sum_{i=0}^{n-10} \mathcal{I}\left(\left(\min_{j \in \{0,\ldots,k-10\}} ||\hat{\mathbf{s}}_{i:i+9} - \mathbf{s}_{j:j+9}||_2\right) \leq \epsilon\right)$$

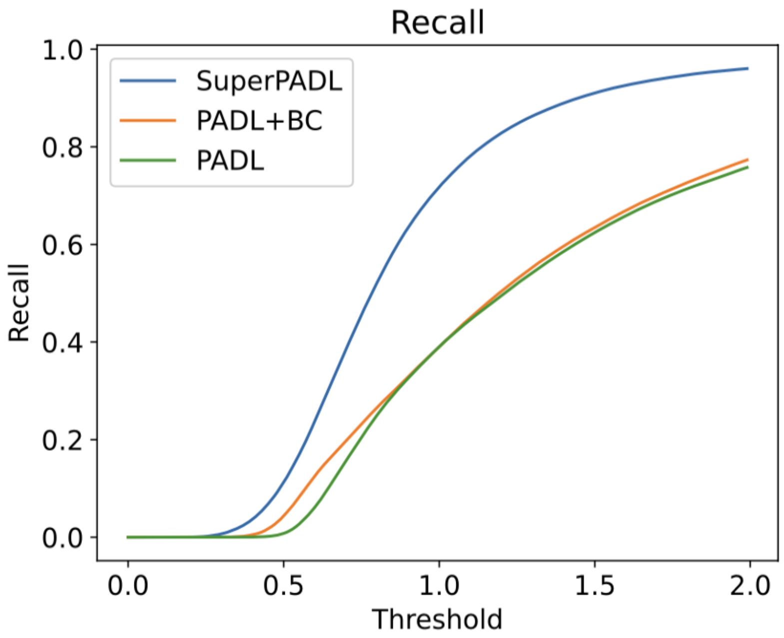
$$\mathrm{Prec}(\tau, \hat{\mathbf{m}}, \epsilon) = \frac{1}{k-9} \sum_{i=0}^{k-10} \mathcal{I}\left(\left(\min_{j \in \{0,\ldots,n-10\}} ||\mathbf{s}_{i:i+9} - \hat{\mathbf{s}}_{j:j+9}||_2\right) \leq \epsilon\right)$$

motion sequence $\hat{\mathbf{m}} = (\hat{\mathbf{s}_0}, \hat{\mathbf{s}_1}, \ldots, \hat{\mathbf{s}_n})$

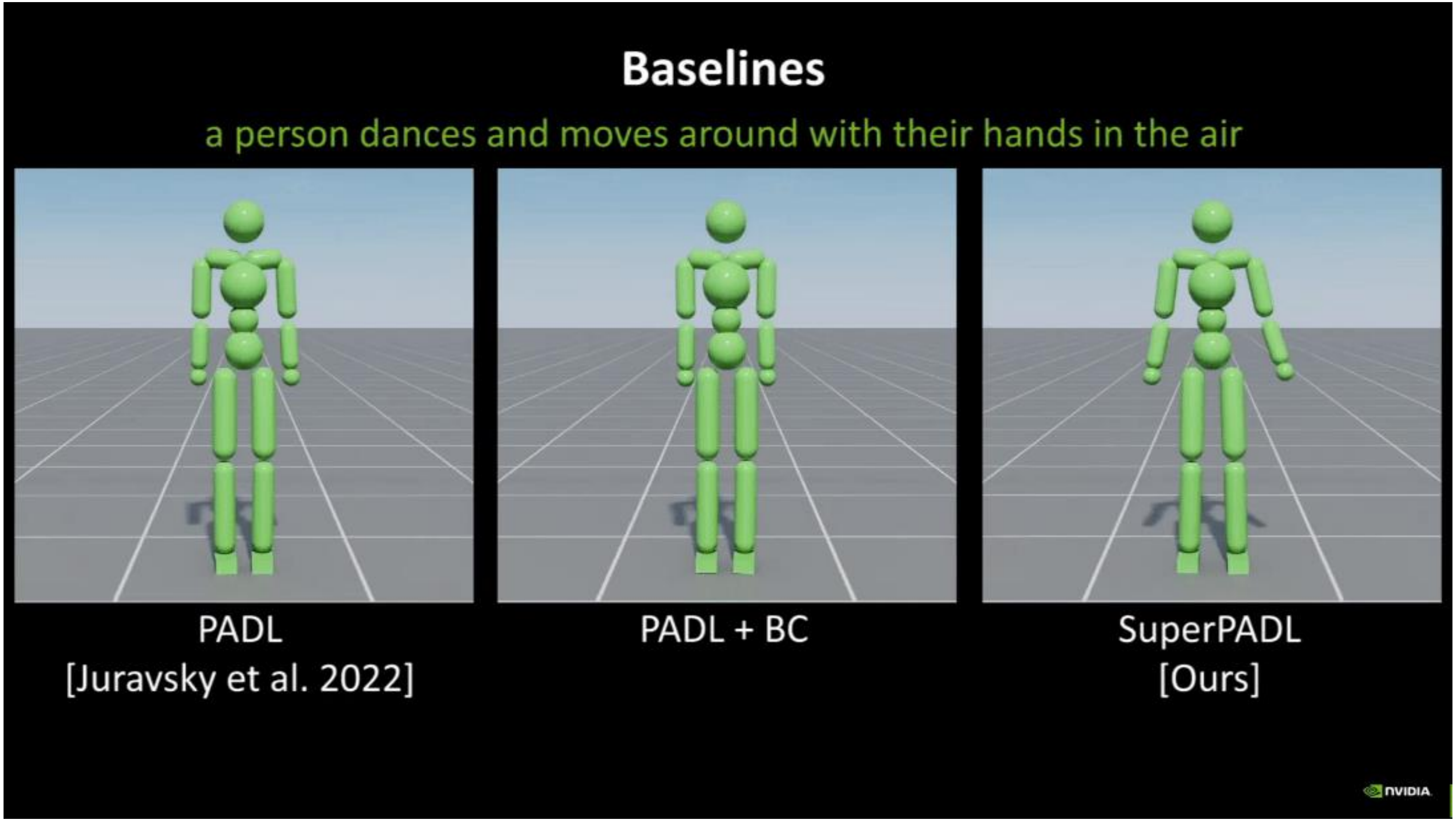
trajectory $\tau = (\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_k)$

I : Indicator variable



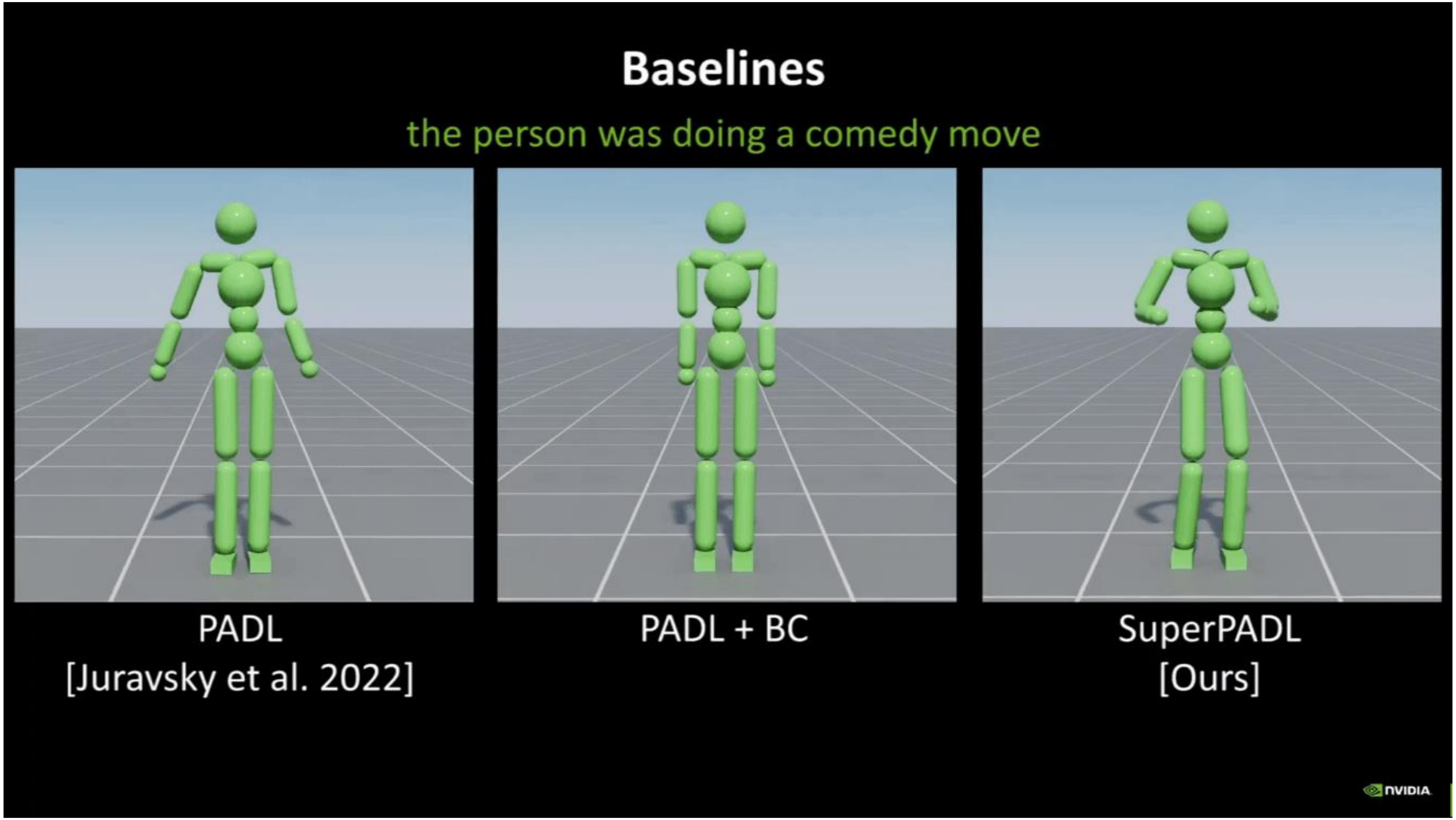


Evaluation - Qualitative



Prompt : a person dances and moves around with their hands in the air

Evaluation – Qualitative



Prompt : the person was doing a comedy move

05 Reference

Reference

- Jordan Juravsky, Yunrong Guo, Sanja Fidler, and Xue Bin Peng. 2024. SuperPADL: Scaling Language-Directed Physics-Based Control with Progressive Supervised Distillation. In ACM SIGGRAPH 2024 Conference Papers (SIGGRAPH '24). Association for Computing Machinery, New York, NY, USA, Article 121, 1–11. https://doi.org/10.1145/3641519.3657492

- Xue Bin Peng, Glen Berseth, Kangkang Yin, and Michiel Van De Panne. 2017. DeepLoco: dynamic locomotion skills using hierarchical deep reinforcement learning. ACM Trans. Graph. 36, 4, Article 41 (August 2017), 13 pages. https://doi.org/10.1145/3072959.3073602

- Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. 2018. DeepMimic: example-guided deep reinforcement learning of physics-based character skills. ACM Trans. Graph. 37, 4, Article 143 (August 2018), 14 pages. https://doi.org/10.1145/3197517.3201311

- Ian J. Goodfellow et al, Generative Adversarial Nets, Part of Advances in Neural Information Processing Systems 27 (NIPS 2014)

- Hanchao Liu, Xiaohang Zhan, Shaoli Huang, Tai-Jiang Mu, and Ying Shan. 2024. Programmable Motion Generation for Open-Set Motion Control Tasks. Retrieved from https://arxiv.org/abs/2405.19283

- Distilling the Knowledge in a Neural Network, Hinton & Jeff Dean et al, NIPS workshop, 2024

- FitNets: Hints for Thin Deep Nets, Romero & Bengio et al, ICLR, 2015

- Jordan Juravsky, Yunrong Guo, Sanja Fidler, and Xue Bin Peng. 2022. PADL: Language-Directed Physics-Based Character Control. In SIGGRAPH Asia 2022 Conference Papers (SA '22). Association for Computing Machinery, New York, NY, USA, Article 19, 1–9. https://doi.org/10.1145/3550469.3555391

- https://www.youtube.com/watch?v=io4dwSbAUW4

- https://www.youtube.com/watch?v=rfOiBqnL_OQ