Towards a Virtual Stuntman

Xue Bin (Jason) Peng

UC Berkeley
Animation
Animation
Computer Animation

[Geijtenbeek et al. 2013]  [Brown et al. 2013]  [Ju et al. 2013]

[Tan et al. 2014]  [Kwon and Hodgins 2017]  [Peng et al. 2018]
Physics-Based Animation

1m/s walk

[Coros et al. 2011]
Deep RL

[Mnih et al. 2015]  [Schulman et al. 2016]  [Chebotar et al. 2017]
Motion Quality

[Schulman et al. 2016] [Heess et al. 2017]
Motivation
Which is Mocap?
Which is Mocap?

Mocap

Simulation
Overview

Character + Reference Motion + Task: Hit Target
Overview
Overview
Overview
Reference Motion
Reference Motion
Reference Motion
Reference Motion
Reference Motion
Reference Motion
Reference Motion

$a_0$  $a_1$  $a_2$  $a_3$  $a_4$
State + Action

State:
• link positions
• link velocities
State + Action

State:
• link positions
• link velocities
State + Action

State:
• link positions
• link velocities

Action:
• PD targets
State + Action

State:
• link positions
• link velocities

Action:
• PD targets
Reward

\[ r_t = \omega^I r^I_t + \omega^G r^G_t \]
Reward

\[ r_t = \omega^I r^I_t + \omega^G r^G_t \]

Imitation Objective
Reward

\[ r_t = \omega^I r^I_t + \omega^G r^G_t \]

Imitation Objective
Reward

\[ r_t = \omega^I r_t^I + \omega^G r_t^G \]

Imitation Objective
Reward

\[ r_t = \omega^I r^I_t + \omega^G r^G_t \]

Imitation Objective
Reward

\[ r_t = \omega^I r^I_t + \omega^G r^G_t \]

Imitation Objective

Task Objective
Proximal Policy Optimization (PPO)

$$\max_{\theta} J(\theta)$$

[Schulman et al. 2017]
Proximal Policy Optimization (PPO)

$$\max_{\theta} \quad J(\theta)$$

s.t. $$\mathbb{E}_{s_t \sim d_\theta(s_t)} \left[ KL \left( \pi_{\theta_{old}}(\cdot | s_t) \| \pi_{\theta}(\cdot | s_t) \right) \right] \leq \delta_{KL}$$

[Schulman et al. 2017]
Proximal Policy Optimization (PPO)

\[
\begin{align*}
\max_{\theta} & \quad J(\theta) \\
\text{s.t.} & \quad \mathbb{E}_{s_t \sim d_\theta(s_t)} \left[ KL \left( \pi_{\theta_{old}}(\cdot|s_t) \big| \pi_{\theta}(\cdot|s_t) \right) \right] \leq \delta_{KL}
\end{align*}
\]

[Schulman et al. 2017]
Proximal Policy Optimization (PPO)

\[
\max_{\theta} \quad J(\theta) \\
\text{s.t.} \quad \mathbb{E}_{s_t \sim d_\theta(s_t)} \left[ KL \left( \pi_{\theta_{old}}(\cdot|s_t) \parallel \pi_\theta(\cdot|s_t) \right) \right] \leq \delta_{KL}
\]

[Schulman et al. 2017]
Proximal Policy Optimization (PPO)

\[
\begin{align*}
\max_{\theta} \quad & J(\theta) \\
\text{s.t.} \quad & \mathbb{E}_{s_t \sim d_\theta(s_t)} \left[ KL \left( \pi_{\theta_{\text{old}}}(\cdot|s_t) \middle| \pi_\theta(\cdot|s_t) \right) \right] \leq \delta_{KL}
\end{align*}
\]

[Schulman et al. 2017]
Humanoid: Walk

Simulation  Reference
Humanoid: Run

Simulation

Reference
Comparison

Ours

[Merel et al. 2017]
No Reference Motion
Locomotion
Humanoid: Cartwheel
Humanoid: Backflip
Humanoid: Frontflip
Humanoid: Roll

Simulation

Reference
Humanoid: Crawl

Simulation

Reference
Humanoid: Dance A

Simulation

Reference
Humanoid: Headspin

Simulation

Reference
Humanoid: Vault 1-Handed
Humanoid: Flare
20+ Skills
Keyframe Animation
T-Rex: Walk

Simulated Character
Dragon: Walk

Simulated Character
Lion: Run

Simulated Character
Tasks

Reference Motion
Tasks

Reference Motion + Task
Humanoid: Spinkick - Strike
Humanoid: Baseball Pitch - Throw
No Reference Motion
Humanoid: Balance Beam
Humanoid: Run – Dense Gaps
Retargeting
Character Retargeting

Reference Motion + Atlas
Atlas: Spinkick
Atlas: Run
Atlas: Getup-Facedown
Atlas: Backflip
Multi-Clip Integration
Multi-Clip Integration

\[ \Pi(a|s) = \sum_{i=1}^{k} p^i(s)\pi^i(a|s) \]
Multi-Clip Integration

Left Cartwheel
Learning from Mocap
Mocap is a Hassle

[Holden 2018]
Skills From Videos
Learning from Videos

Video
Overview

Video → Pose Estimation → Poses → Motion Reconstruction → Reference Motion → Character → Motion Imitation (RL) → π
Overview

Video → Pose Estimation → Poses → Motion Reconstruction

Character → Motion Imitation (RL) → $\pi$
Overview
Overview

Video → Pose Estimation → Poses → Motion Reconstruction

Reference Motion

Character → Motion Imitation (RL) → $\pi$
Pose Estimation
Pose Estimation

Video: Handspring A

Pose Prediction
Human Mesh Recovery (HMR)

[Kanazawa et al., 2018]
Pose Estimation

Video: Backflip A

Pose Prediction
Motion Imitation
Motion Imitation

\[ r_t = \exp \left( -\| - \| \| - \| ^2 \right) \]
Motion Imitation

\[ r_t = \exp \left( - \| \cdot \|^2 \right) \]
Motion Imitation

\[ r_t = \exp \left( -\left\| \cdot \right\|_2^2 \right) \]
Humanoid: Cartwheel B

Video: Cartwheel B  Reference Motion  Policy
Humanoid: Jumping Jack

Video: Jumping Jack

Policy
Humanoid: Backflip B

Video: Backflip B

Policy
Humanoid: Frontflip

Video: Frontflip

Policy
Humanoid: Roll

Video: Roll

Policy
Humanoid: Spin

Video: Spin

Policy
Humanoid: Kip-Up

Video: Kip-Up

Policy
Humanoid: Vault

Video: Vault

Policy
Atlas: Handspring A

Video: Handspring A

Policy
Atlas: Jump

Video: Jump

Policy
Atlas: Vault

Video: Vault

Policy
Atlas: Dance

Video: Dance

Policy
Environment Retargeting
Environment Retargeting
Failure Cases

Video: Gangnam Style  Reference Motion  Simulation
Skills from Videos

Policy
Concluding Remarks

- Simple method can learn a large repertoire of skills
- Minimizing tracking error works (surprisingly) well
Concluding Remarks

- Simple method can learn a large repertoire of skills
- Minimizing tracking error works (surprisingly) well
- A lot of room for improvement for video imitation
  - More end-to-end approach
  - Outdoor sports
  - Multiple actors
Concluding Remarks

• Simple method can learn a large repertoire of skills
• Minimizing tracking error works (surprisingly) well
• A lot of room for improvement for video imitation
  – More end-to-end approach
  – Outdoor sports
  – Multiple actors
• Code: https://github.com/xbpeng/DeepMimic
Collaborators

Pieter Abbeel
Angjoo Kanazawa
Sergey Levine
Jitendra Malik
Michiel van de Panne
Questions?