Presenting Your Research

Huy Ha
huyha@stanford.edu

The story behind an oral presentation

The story behind an oral presentation

"Wow, your student is so passionate about cloth unfolding"

1.Treat Presentations Seriously

1.Treat Presentations Seriously

Start early

Put in the work

Get feedback

Have multiple iterations

1.Treat Presentations Seriously

Start early

Put in the work

Get feedback

Have multiple iterations

Today's Agenda

Treat presentations seriously

A "normal" research presentation (3 minutes)

A guideline to presenting your research (20 minutes)

- Think about your audience
- Make things look pretty

A hopefully improved research presentation! (25 minutes)

Today's Agenda

Treat presentations seriously

A "normal" research presentation (3 minutes)

A guideline to presenting your research (20 minutes)

- Think about your audience
- Make things look pretty

A hopefully improved research presentation! (25 minutes)

UMI on Legs

Making Manipulation Policies Mobile with Manipulation-Centric Whole-body Controllers

Huy Ha*, Yihuai Gao*, Zipeng Fu, Jie Tan, Shuran Song

umi-on-legs.github.io

Problem Motivation

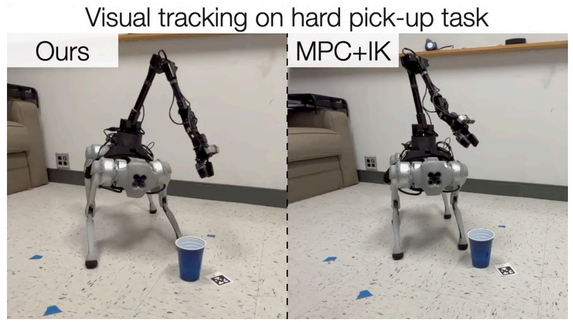
What we want

- Dynamic, autonomous, in-the-wild quadruped manipulation

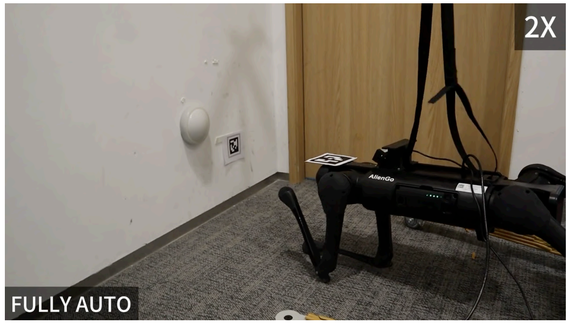
What we need

- Generalizable yet expressive high level policy
- Robust & stable whole body controller for dynamic skills
- Deploy without privileged states (e.g., aruco tags)

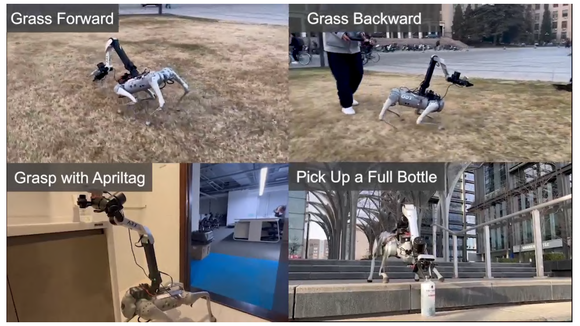
Curiosity-Driven Learning of Joint Locomotion and Manipulation Tasks, *Schwarke et al*



Deep Whole-Body Control, *Fu et al*



Learning Visual Quadrupedal Loco-Manipulation from Demonstrations, *He et al*



RoboDuet, *Pan et al*

Visual Whole-Body Control for Legged Loco-Manipulation, *Liu et al*

Problem Motivation

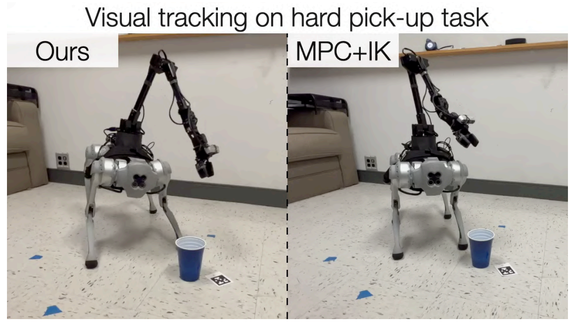
What we want

- Dynamic, autonomous, in-the-wild quadruped manipulation

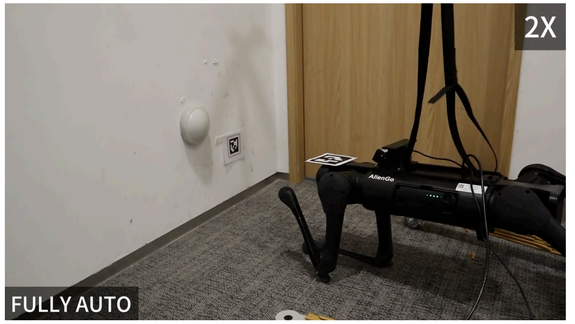
What we need

- Generalizable yet expressive high level policy
- Robust & stable whole body controller for dynamic skills
- Deploy without privileged states (e.g., aruco tags)

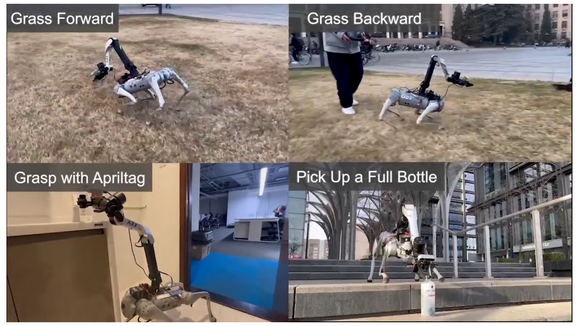
Curiosity-Driven Learning of Joint Locomotion and Manipulation Tasks, *Schwarke et al*



Deep Whole-Body Control, *Fu et al*



Learning Visual Quadrupedal Loco-Manipulation from Demonstrations, *He et al*



RoboDuet, *Pan et al*

Visual Whole-Body Control for Legged Loco-Manipulation, *Liu et al*

Problem Motivation

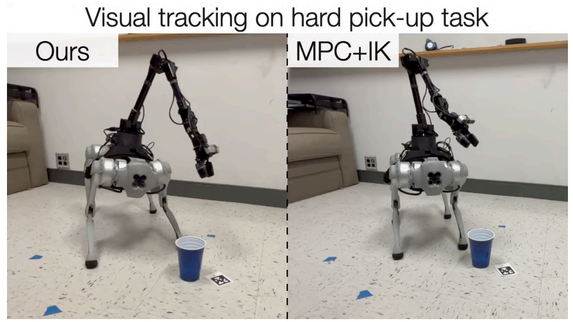
What we want

- Dynamic, autonomous, in-the-wild quadruped manipulation

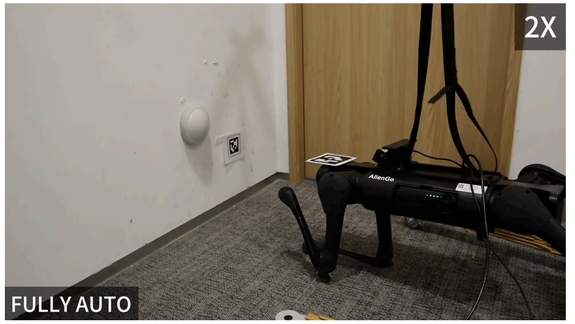
What we need

- Generalizable yet expressive high level policy
- Robust & stable whole body controller for dynamic skills
- Deploy without privileged states (e.g., aruco tags)

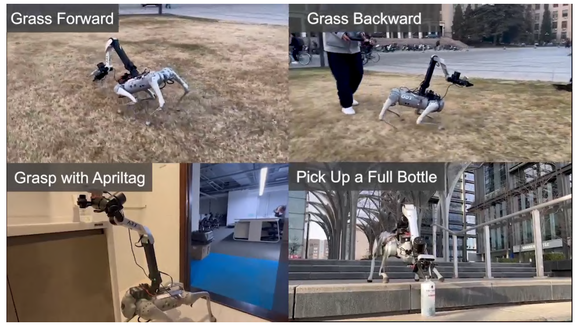
Curiosity-Driven Learning of Joint Locomotion and Manipulation Tasks, *Schwarke et al*



Deep Whole-Body Control, *Fu et al*



Learning Visual Quadrupedal Loco-Manipulation from Demonstrations, *He et al*



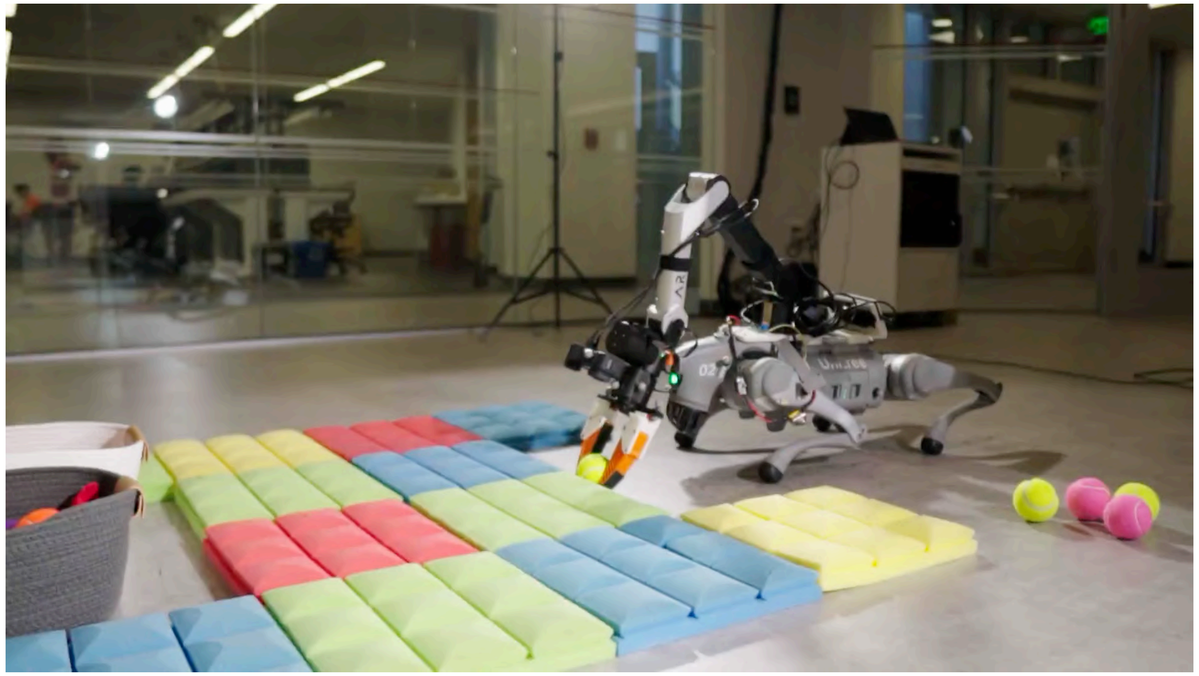
RoboDuet, *Pan et al*

Visual Whole-Body Control for Legged Loco-Manipulation, *Liu et al*

UMI on Legs

Contributions

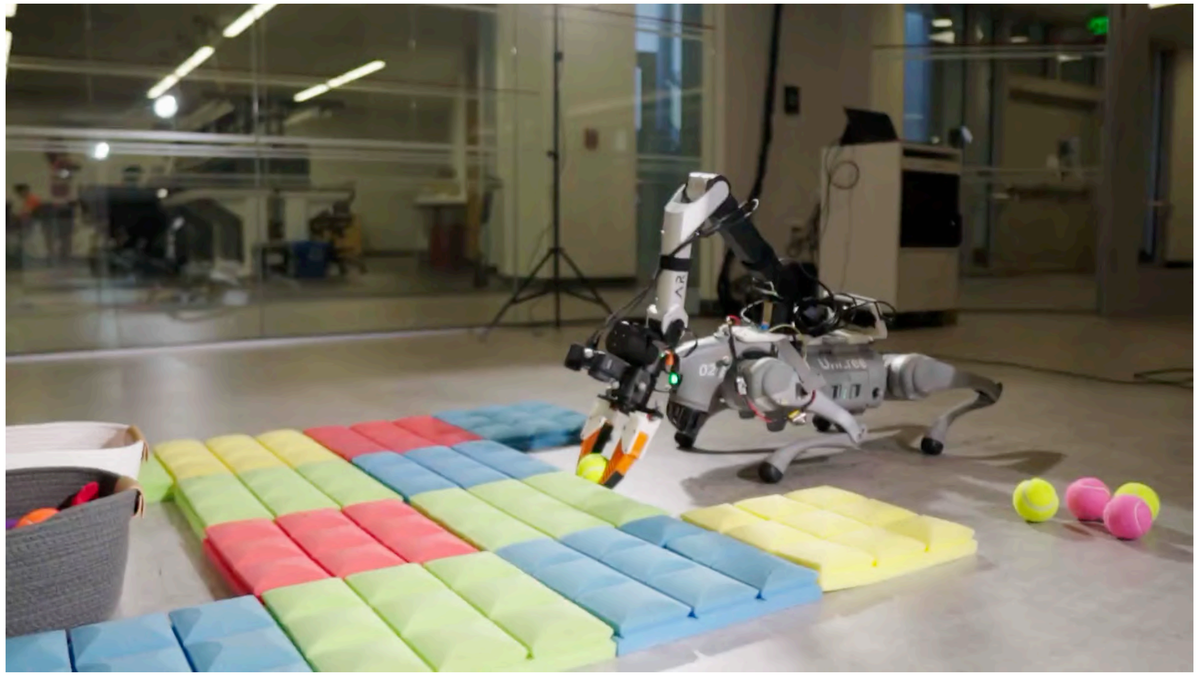
- **Framework** for combining sim and real data for cross-embodiment mobile manipulation systems
- **Manipulation-centric WBC** for tracking complex manipulation trajectories precisely
- **Real-world deployment system** with real-time & robust odometry



UMI on Legs

Contributions

- **Framework** for combining sim and real data for cross-embodiment mobile manipulation systems
- **Manipulation-centric WBC** for tracking complex manipulation trajectories precisely
- **Real-world deployment system** with real-time & robust odometry



Approach

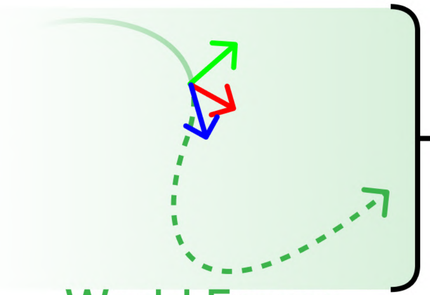
Manipulation Policy

RGB

Diffusion Policy

2Hz

The Interface



World Frame
End-Effector
Trajectory

Whole-body Controller

base state
arm state
leg state
last action

MLP

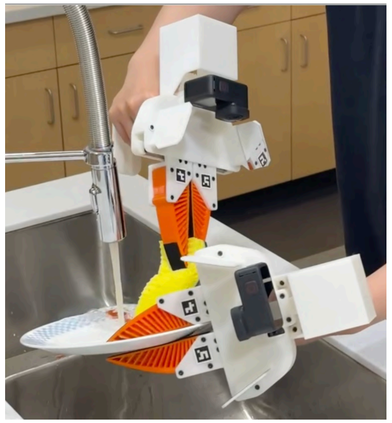
50Hz

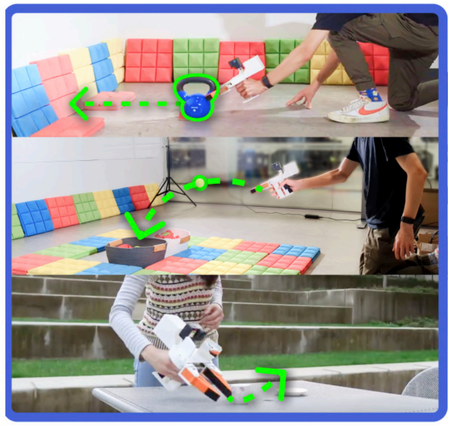
Arm

Leg

Action

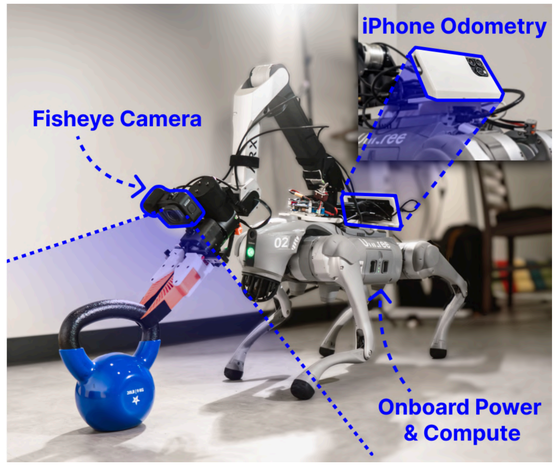
← Behavior-Clone in Real World →





Universal Manipulation Interface, Chi et al 2024

← Reinforcement Learn in Sim →



Quantitative Tossing Results

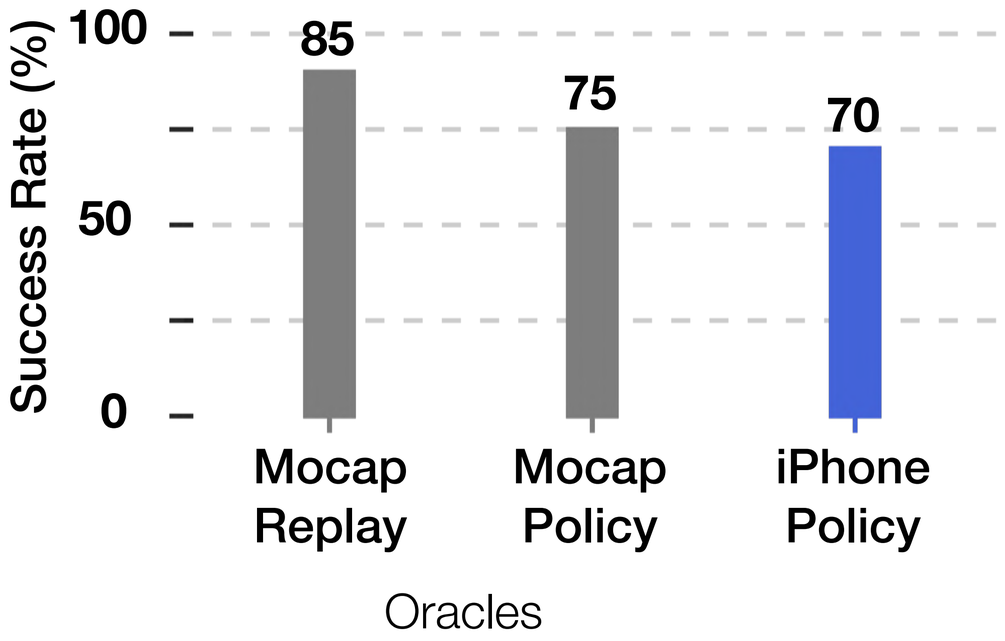
Approach *Units*	Pos Err *cm*↓	Orn Err *deg*↓	Survival *%* ↑	Power *kW*↓
Ours (Unified)	2.18	**3.10**	**99.8**	3.72
Ours (Tossing)	**2.12**	3.35	98.4	3.82
(-) Traj Interface	3.02	4.23	93.0	3.95
(-) Task-space	15.49	15.55	0.0	4.74
(-) UMI Data	2.48	15.67	97.4	**3.69**
DeepWBC	22.2	66.22	0.0	5.92

Controller Ablations

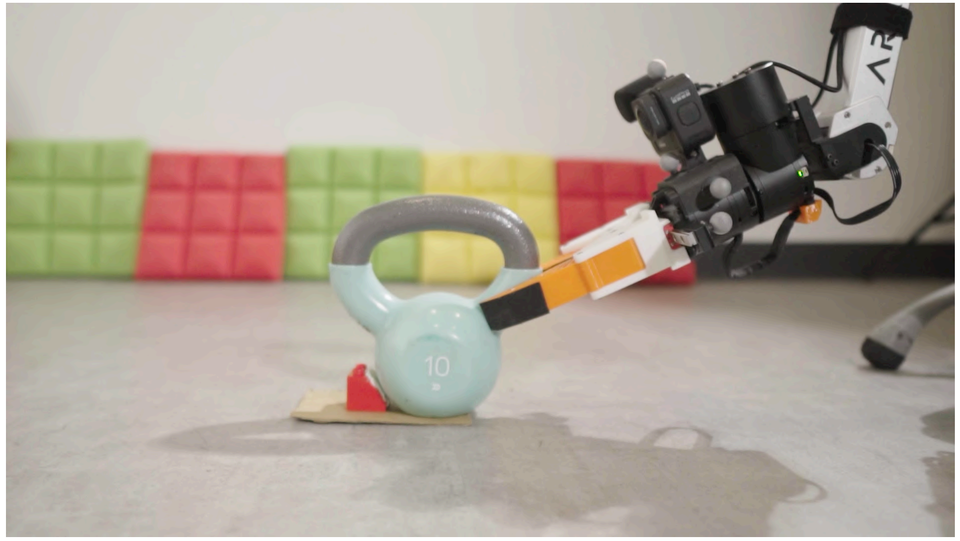
- **More precise and safe** than without trajectory interface
- Outperforms SOTA WBC (DeepWBC)

System Results

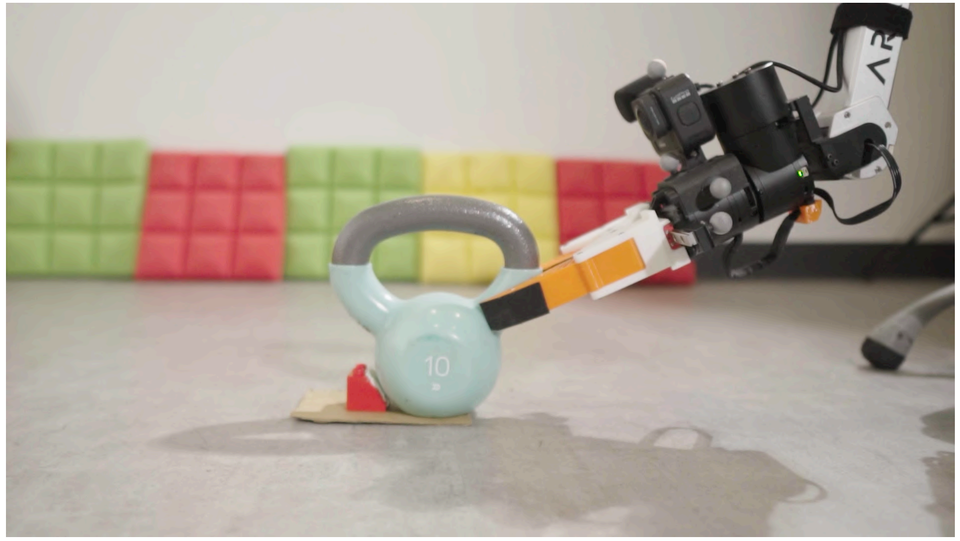
- **iPhone odometry** achieves similar performance to motion capture oracle
- Visuo-motor policy only -10% lower than oracle trajectory replay



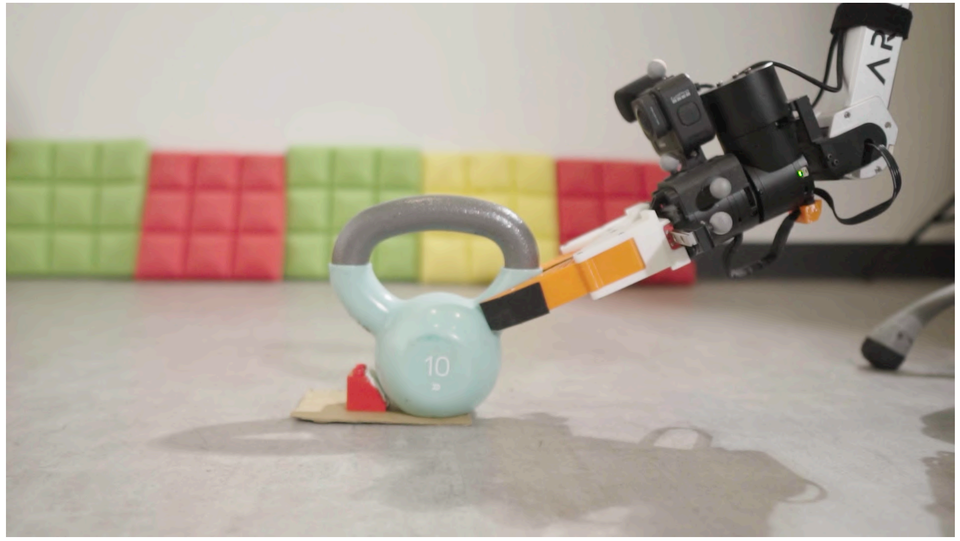
Qualitative Results



Qualitative Results



Qualitative Results



UMI on Legs

Making Manipulation Policies Mobile with Manipulation-Centric Whole-body Controllers

Huy Ha*, Yihuai Gao*, Zipeng Fu, Jie Tan, Shuran Song

umi-on-legs.github.io

Let's roast that presentation 🔥

Let's roast that presentation 🔥

plain, not engaging

just reading the slides and script

hard to follow

overwhelming content in slides, but not talked about

hard to parse what to take away, important points glossed over

feels distant, not approachable, personal

no story, feels uninspired after presentation

Let's roast that presentation 🔥

it's just kinda bad and i'm not sure why..

2. Think about your audience

2. Think about your audience

What do they know?

2. Think about your audience

What do they know?

What do they care about?

2. Think about your audience

What do they know?

What do they care about?

How much effort will they put into understanding?

2. Think about your audience

	What do they know?
	What do they care about?
Problem	How much effort will they put into understanding?
Solution	How can I cater my presentation to this audience?

Who are your audience?

Reviewers

Twitter Bros

Talk Listeners

Conference Attendees

VCs

Who are your audience?

People have limited time and attention.

Who are your audience?

Follow up Works

Reviewers

Twitter Bros

Talk Listeners

Conference Attendees

VCs

Designing an Information Cascade

Designing Information Cascades

What you make

Tweet
Video

Website

Presentation

Paper

Codebase

More effort, More technical details

Designing Information Cascades

What you make

Tweet
Video

Website

Presentation

Paper

Codebase

More effort, More technical details

Who they are for

Twitter Bros
VCs

Conference Attendees

Talk Listeners

Reviewers

Follow Up Works

Designing Information Cascades

What you make

Tweet
Video

Website

Presentation

Paper

Codebase

More effort, More technical details

Who they are for

Twitter Bros
VCs

Conference Attendees

Talk Listeners

Reviewers

Follow Up Works

Information Cascade in Websites

Anatomy of my websites.

TLDR. 1-2 sentences + key takeaway. Plenty of space for emphasis.

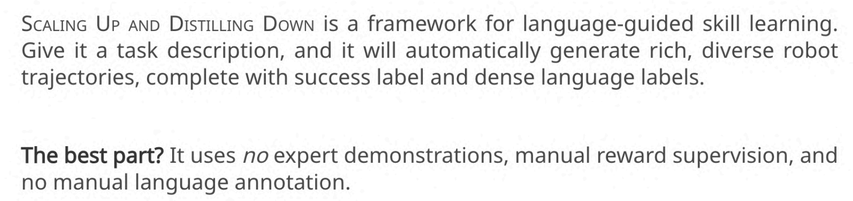
Scaling Up and Distilling Down is a framework for language-guided skill learning. Give it a task description, and it will automatically generate rich, diverse robot trajectories, complete with success label and dense language labels.

The best part? It uses *no* expert demonstrations, manual reward supervision, and no manual language annotation.

Information Cascade in Websites

Anatomy of my websites.

TLDR. 1-2 sentences + key takeaway. Plenty of space for emphasis.

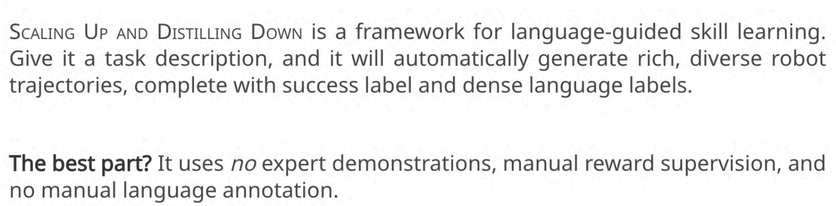


Q&A at end. More information for people who cares. Give candid responses.

Information Cascade in Websites

Anatomy of my websites.

TLDR. 1-2 sentences + key takeaway. Plenty of space for emphasis.



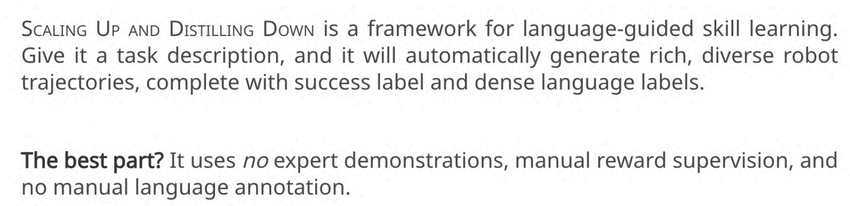
Q&A at end. More information for people who cares. Give candid responses.

More effort, More technical details

Information Cascade in Websites

Anatomy of my websites.

TLDR. 1-2 sentences + key takeaway. Plenty of space for emphasis.



Q&A at end. More information for people who cares. Give candid responses.

It’s not a paper. Stay high level, write only a few sentences in each section. It should motivate interested people to read your paper.

Information Cascade in Websites

Anatomy of my websites.

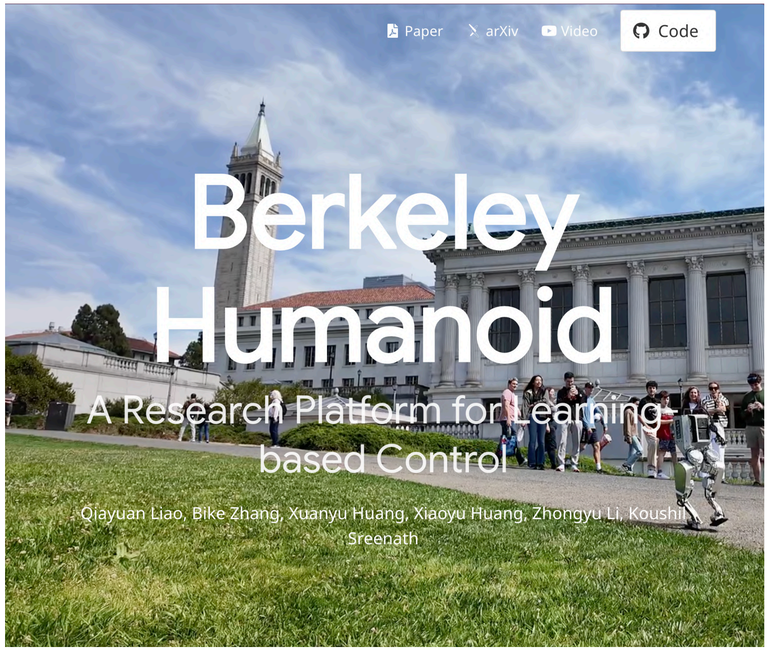
TLDR. 1-2 sentences + key takeaway. Plenty of space for emphasis.

SCALING UP AND DISTILLING DOWN is a framework for language-guided skill learning. Give it a task description, and it will automatically generate rich, diverse robot trajectories, complete with success label and dense language labels.

The best part? It uses *no* expert demonstrations, manual reward supervision, and no manual language annotation.

Q&A at end. More information for people who cares. Give candid responses.

It's not a paper. Stay high level, write only a few sentences in each section. It should motivate interested people to read your paper.



Berkeley Humanoid. Just fork the website! Tweak it to your project :)

Designing Information Cascades

What you make

Tweet
Video

Website

Presentation

Paper

Codebase

More effort, More technical details

Who they are for

Twitter Bros
VCs

Conference Attendees

Talk Listeners

Reviewers

Follow Up Works

Designing Information Cascades

What you make

Tweet
Video

Website

Presentation

Paper

Codebase

More effort, More technical details

Who they are for

Twitter Bros
VCs

Conference Attendees

Talk Listeners

Reviewers

Follow Up Works

Teaching Backward

Teaching Backward

Teaching Forwards

Teaching Backward

Teaching Forwards

Definition, examples

Teaching Backward

Teaching Forwards

Definition, examples

Teaching Backwards

Examples, counter-examples, definitions

Teaching Backward

Teaching Forwards

Definition, examples

Solution, problem/result

Teaching Backwards

Examples, counter-examples, definitions

Teaching Backward

Teaching Forwards

Definition, examples

Solution, problem/result

Teaching Backwards

Examples, counter-examples, definitions

Problem/result, solution

Teaching Backward: Problem, then Solution

Teaching Forwards

We built an iPhone-based odometry system that is real-time and works in the wild. This high frequency odometry is used in determining the body motions, so that the robot can compensate that motion with its arm. This allows the robot to isolate body movements from its gripper.

Teaching Backwards

To isolate sudden body movements from its gripper, the robot must know how its body moves, then react fast counter this movement with its arm. To enable this ability, we built an iPhone-based system, which provides in-the-wild, real-time odometry.

Problem **Solution**

Teaching Backward: Problem, then Solution

Teaching Forwards

We built an iPhone-based odometry system that is real-time and works in the wild. This high frequency odometry is used in determining the body motions, so that the robot can compensate that motion with its arm. This allows the robot to isolate body movements from its gripper.

Teaching Backwards

To isolate sudden body movements from its gripper, the robot must know how its body moves, then react fast counter this movement with its arm. To enable this ability, we built an iPhone-based system, which provides in-the-wild, real-time odometry.

Problem **Solution**

Practical Tips on Catering to Audience

Teach Backwards. Justify everything you say before you say it.

Be Engaging. Tone. Mood. Phrasing. Interactivity. Jokes.

Rehash. In paper writing, presentation, or anywhere, repeat important points multiple times.

Informative Slide Titles. People will zone in and out. Titles can help zoning back in.

Practical Tips on Catering to Audience

Teach Backwards. Justify everything you say before you say it.

Be Engaging. Tone. Mood. Phrasing. Interactivity. Jokes.

Rehash. In paper writing, presentation, or anywhere, repeat important points multiple times.

Informative Slide Titles. People will zone in and out. Titles can help zoning back in.

Practical Tips on Catering to Audience

Teach Backwards. Justify everything you say before you say it.

Be Engaging. Tone. Mood. Phrasing. Interactivity. Jokes.

Rehash. In paper writing, presentation, or anywhere, repeat important points multiple times.

Informative Slide Titles. People will zone in and out. Titles can help zoning back in.

Practical Tips on Catering to Audience

Teach Backwards. Justify everything you say before you say it.

Be Engaging. Tone. Mood. Phrasing. Interactivity. Jokes.

Rehash. In paper writing, presentation, or anywhere, repeat important points multiple times.

Informative Slide Titles. People will zone in and out. Titles can help zoning back in.

Why should I care about people who don't care or aren't paying attention?

Why should I care about people who don't care or aren't paying attention?

Because a good presentation can turn them into people who do.

My View

Research is not just for researchers.

Potential for impacting everything and everyone.

Make it so accessible that everyone can learn about it.

Different resources catering to different people.

My View

Research is not just for researchers.

Potential for impacting everything and everyone.

Make it so accessible that everyone can learn about it.

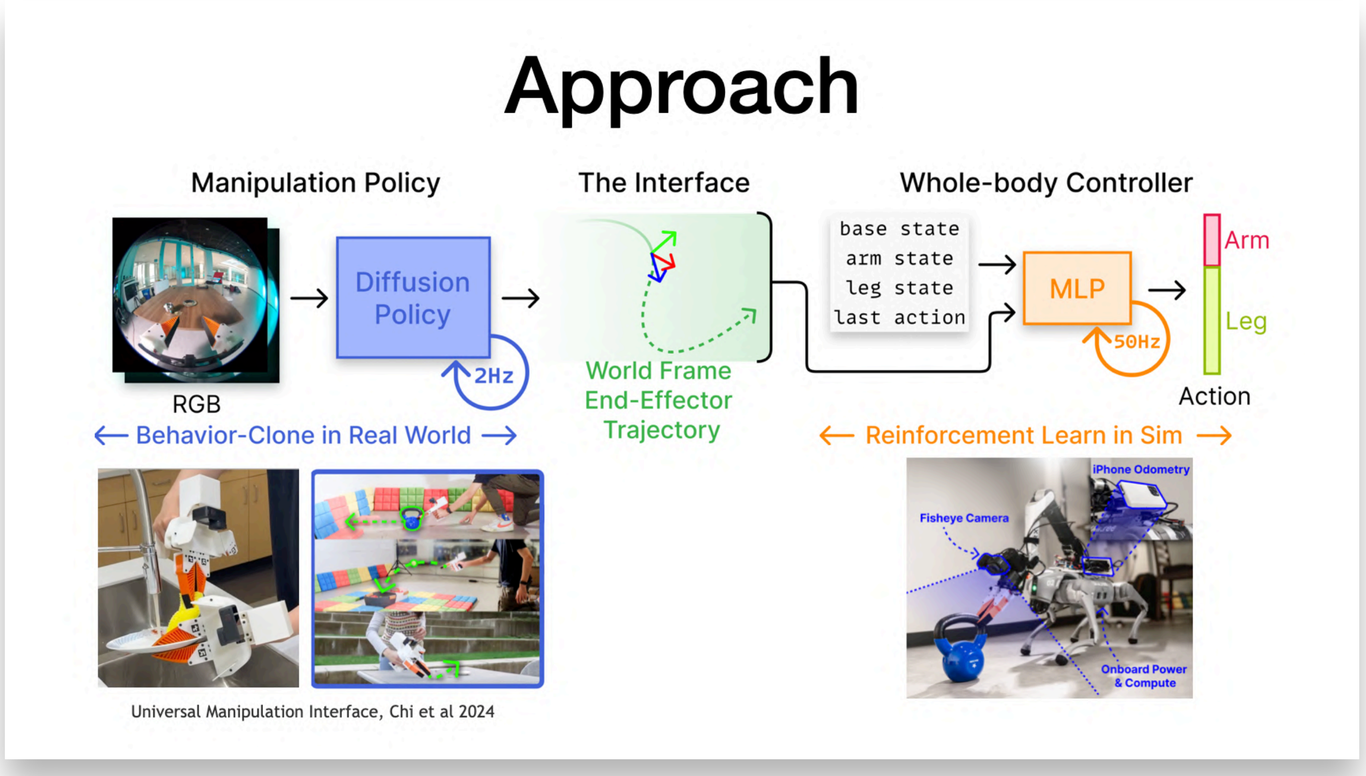
Different resources catering to different people.

3. Control People's Attention

Timing, Pacing

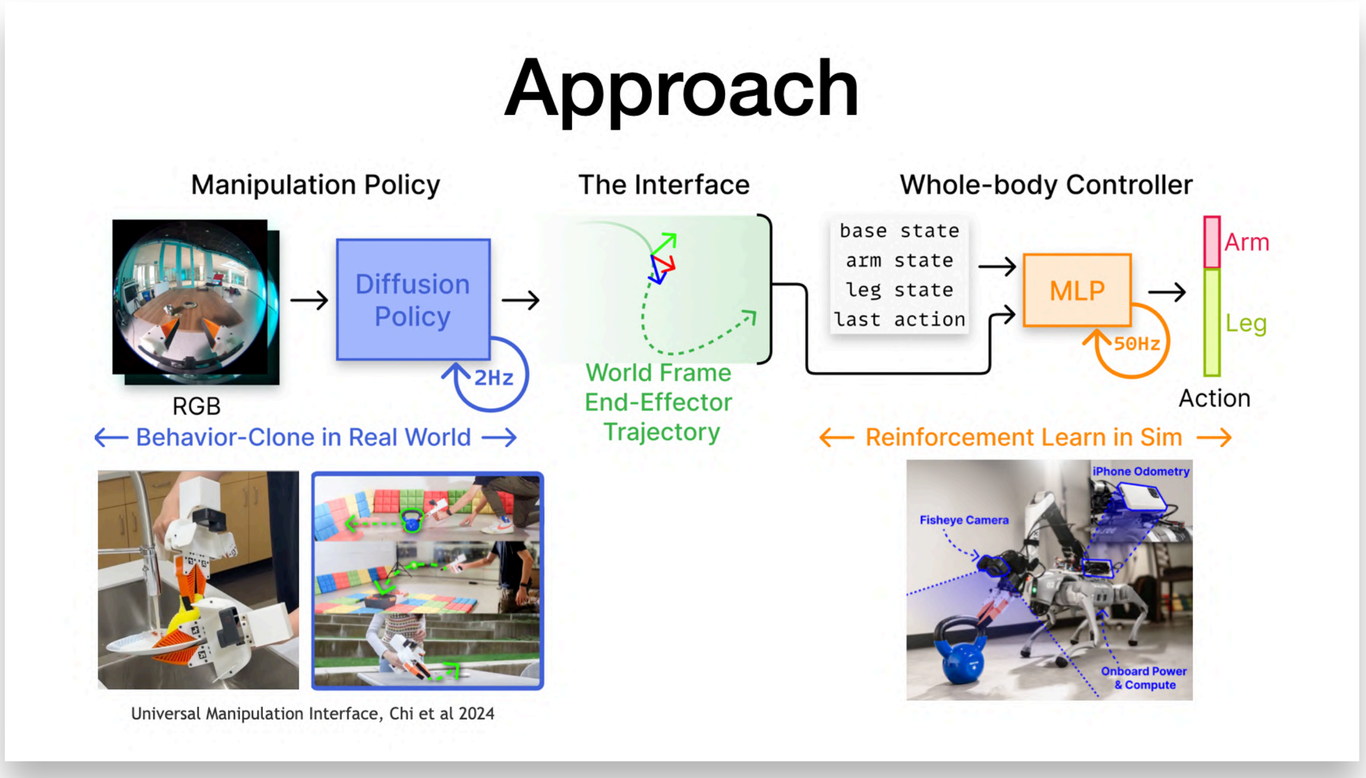
Build up concepts with animation

What not to do



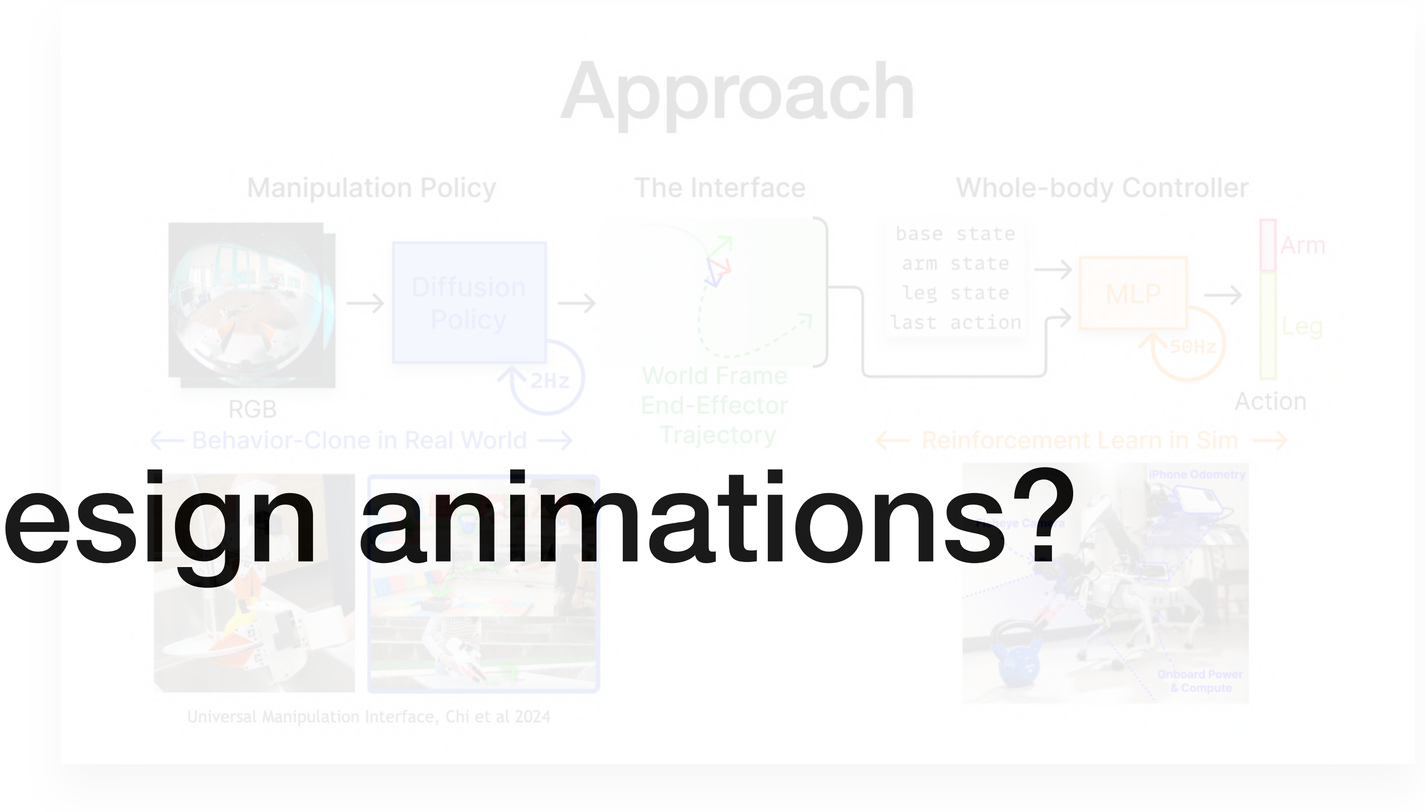
The approach slide from previous presentation

What not to do



The approach slide from previous presentation

What not to do



The approach slide from previous presentation

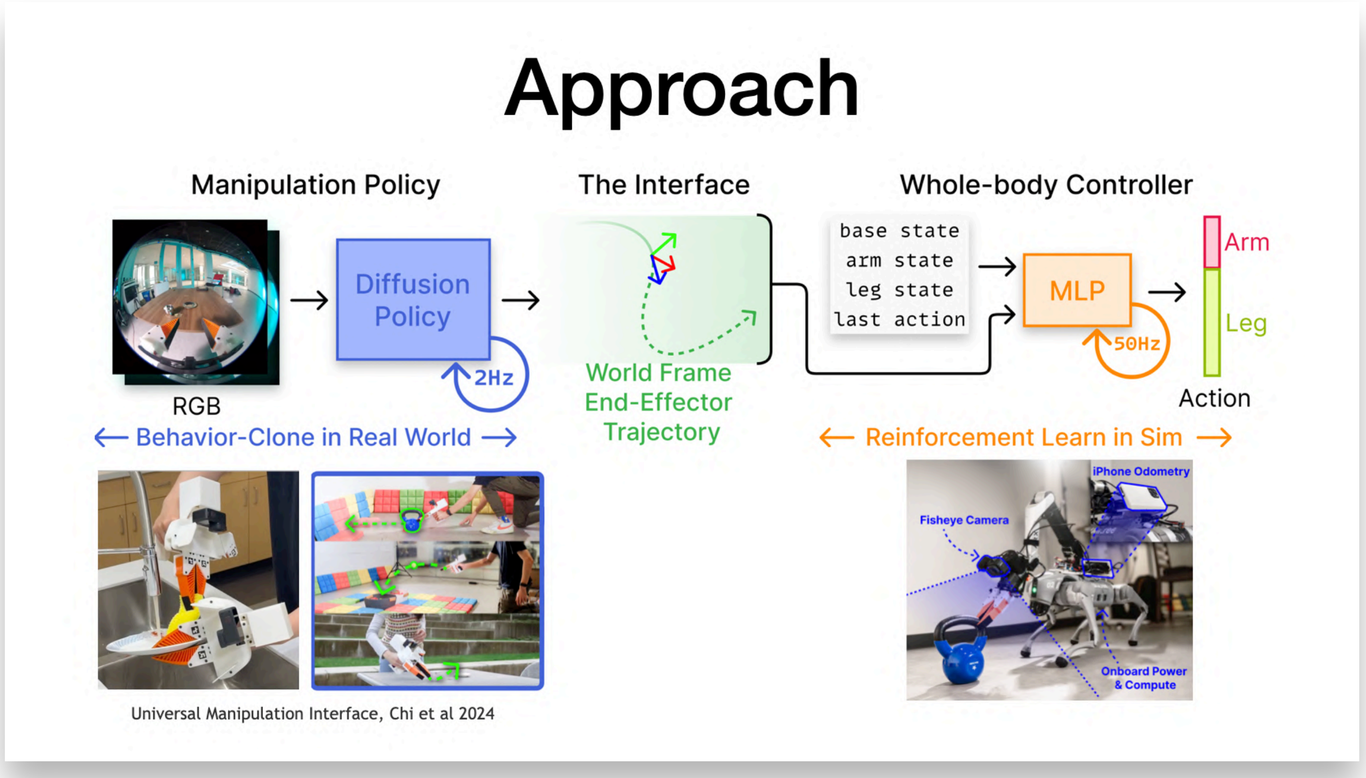
How to design animations?

Script to Animation

"UMI on Legs has two parts. The first part is UMI, which allows ... The second part is the whole body controller learning, which was design to be ... Combined together, the system allows us to achieve... To ensure we can deploy on hardware, we built a custom iOS app..."

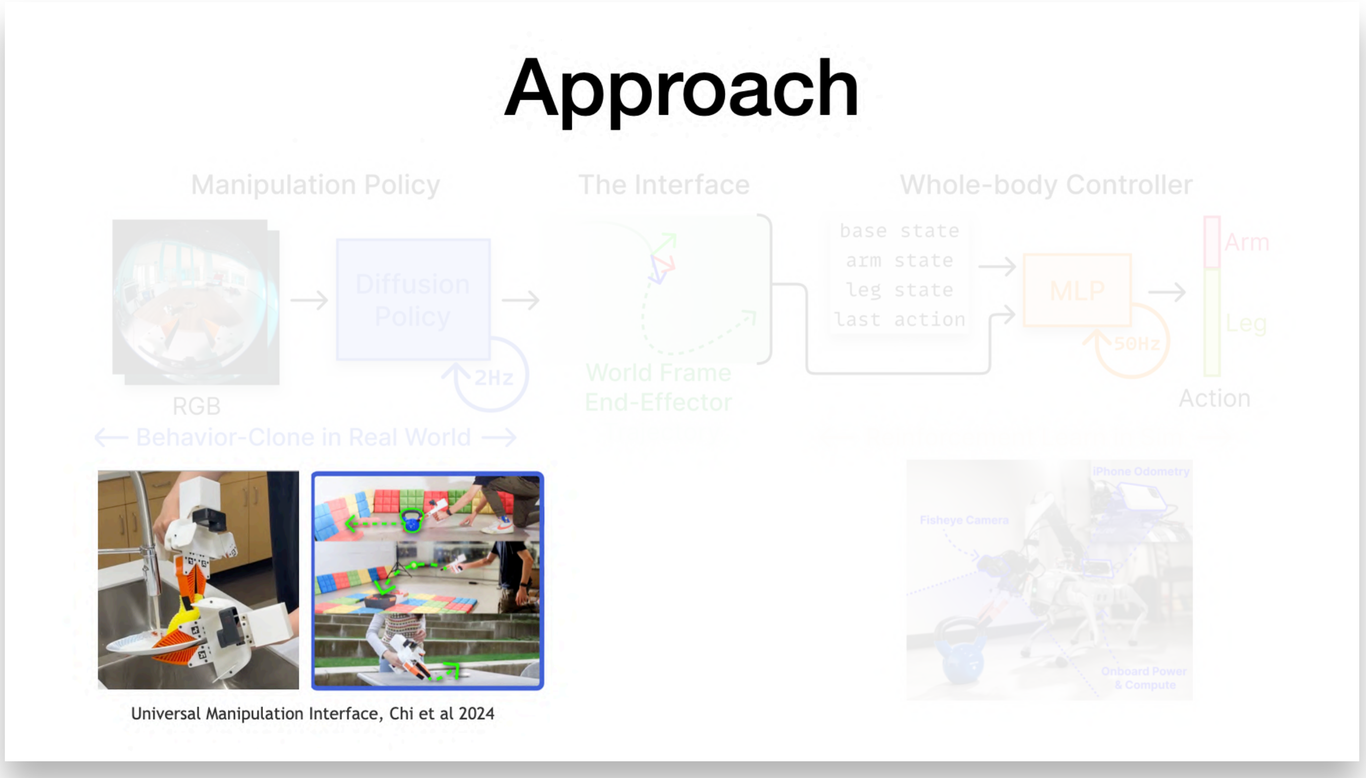
Script to Animation

"UMI on Legs has two parts. The first part is UMI, which allows ... The second part is the whole body controller learning, which was design to be ... Combined together, the system allows us to achieve... To ensure we can deploy on hardware, we built a custom iOS app..."



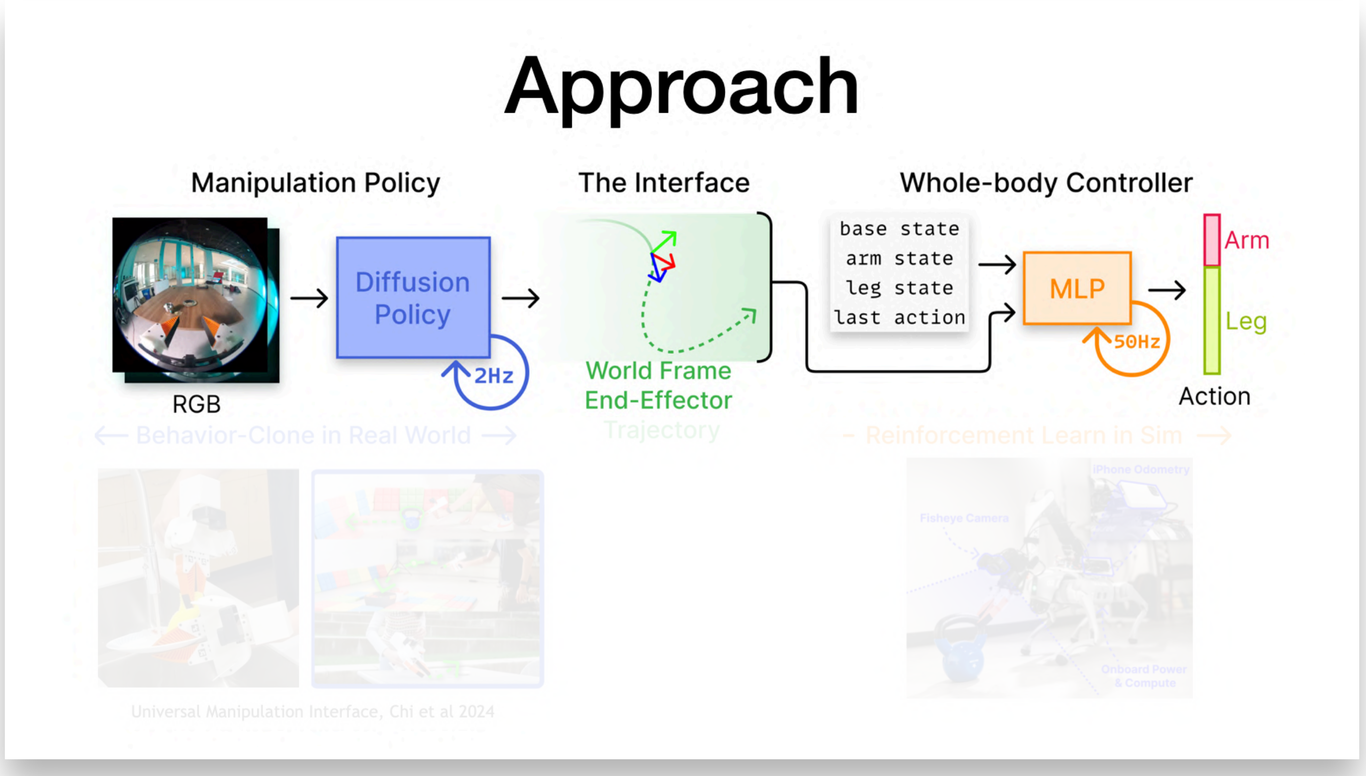
Script to Animation

"UMI on Legs has two parts. The first part is UMI, which allows ... The second part is the whole body controller learning, which was design to be ... Combined together, the system allows us to achieve... To ensure we can deploy on hardware, we built a custom iOS app..."



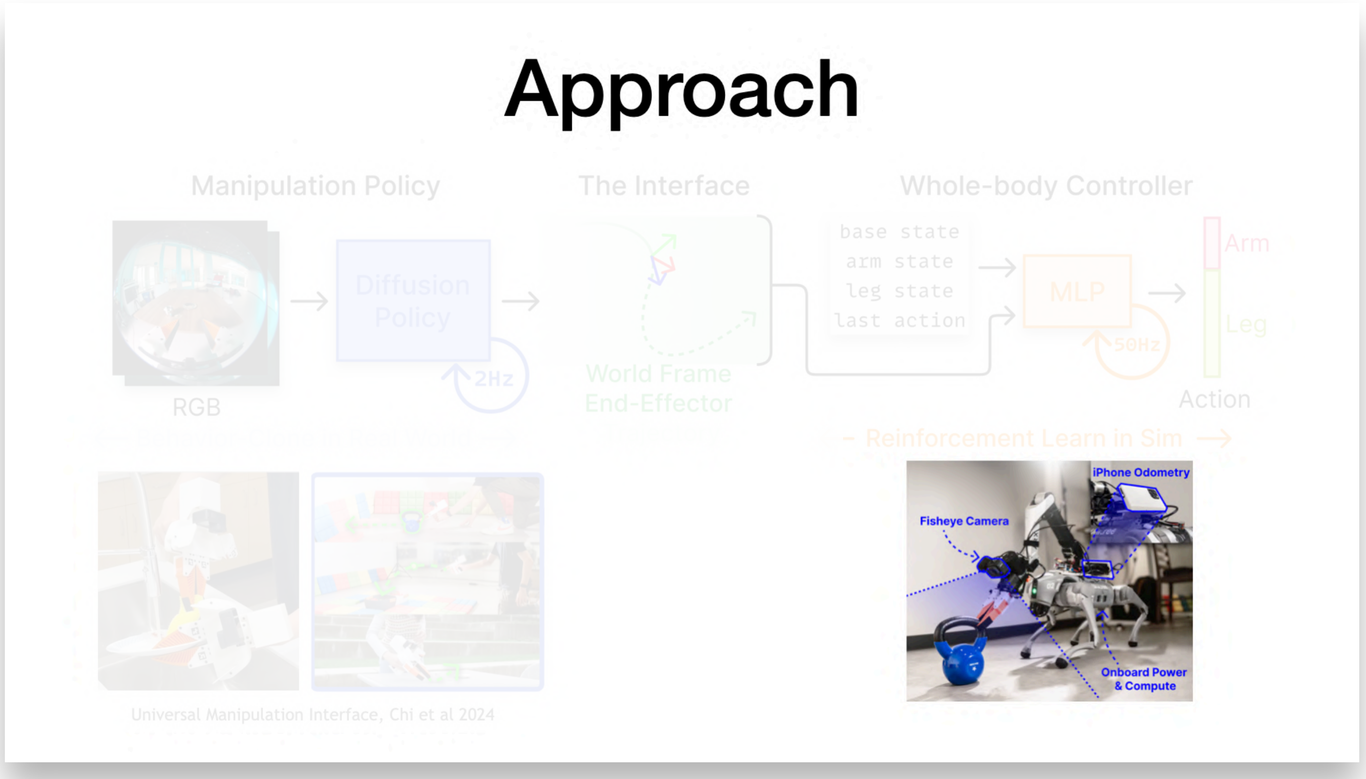
Script to Animation

"UMI on Legs has two parts. The first part is UMI, which allows ... The second part is the whole body controller learning, which was design to be ... Combined together, the system allows us to achieve... To ensure we can deploy on hardware, we built a custom iOS app..."



Script to Animation

"UMI on Legs has two parts. The first part is UMI, which allows ... The second part is the whole body controller learning, which was design to be ... Combined together, the system allows us to achieve... To ensure we can deploy on hardware, we built a custom iOS app..."



With a few seconds of your effort, the viewer can use zero effort to follow along

Practical Tips on Attention Management

Appear/disappear very effective. No fancy magic moves needed, yet most of the gains!

Timing is key. On keywords, not sentence starts/end.

Talk over result videos. No awkward silences.

Practical Tips on Attention Management

Appear/disappear very effective. No fancy magic moves needed, yet most of the gains!

Timing is key. On keywords, not sentence starts/end.

Talk over result videos. No awkward silences.

Practical Tips on Attention Management

Timing is key. On keywords, not sentence starts/end.

On sentence	"[click] Another important point is timing"
On keyword	"Another important point is [click] timing"

Practical Tips on Attention Management

Appear/disappear very effective. No fancy magic moves needed, yet most of the gains!

Timing is key. On keywords, not sentence starts/end.

Talk over result videos. No awkward silences.

4. (Bonus) Make things pretty

Some tips in photography, videography, blender

Videography & Photography

Two Kinds of Visuals

Two different purposes, so keep them separate

Scientific. Objective, fixed view, complete view, end-to-end uncut rollouts, anonymous-safe. Needs to be clear, clean, doesn't have to wow. (belongs to supp material)

Promotional. Sexy, cinematic camera movements, lights, most impressive results. Gets people to click in the first few seconds. (belongs on websites, tweets, etc.)

Two Kinds of Visuals

Two different purposes, so keep them separate

Scientific. Objective, fixed view, complete view, end-to-end uncut rollouts, anonymous-safe. Needs to be clear, clean, doesn't have to wow. (belongs to supp material)

Promotional. Sexy, cinematic camera movements, lights, most impressive results. Gets people to click in the first few seconds. (belongs on websites, tweets, etc.)

You know how to do this.

Two Kinds of Visuals

Two different purposes, so keep them separate

Scientific. Objective, fixed view, complete view, end-to-end uncut rollouts, anonymous-safe. Needs to be clear, clean, doesn't have to wow. (belongs to supp material)

You know how to do this.

Promotional. Sexy, cinematic camera movements, lights, most impressive results. Gets people to click in the first few seconds. (belongs on websites, tweets, etc.)

I'll talk about doing this.

How many shots should I get?

as many as you can, pick later.

How many shots should I get?

Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Meet Espresso and Oat Milk!

Following our lab's caffinated drinks naming tradition, I've decided to name our quadruped Espresso and our new arm Oat Milk. My hope is that, when combined together, Espresso and Oat Milk will be as capable as Latte (our UR5 which has unfolded and folded cloths and washed dishes). Deploying policies from Latte to Espresso and Oat Milk, as we've done, is the first step 🚀.

Espresso - Oat Milk comes with a GoPro on its head 🎩, a 3D printed gripper at the end of its arm 💪, and an iPhone on its butt 🍑. The GoPro streams visual observations through a capture card, serving as the policy observation. Meanwhile, the iPhone runs a custom iOS app we developed, which streams the robot's body pose, allowing world frame stabilization.

NOTE: almost always eye level with robot.

Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Meet Espresso and Oat Milk!

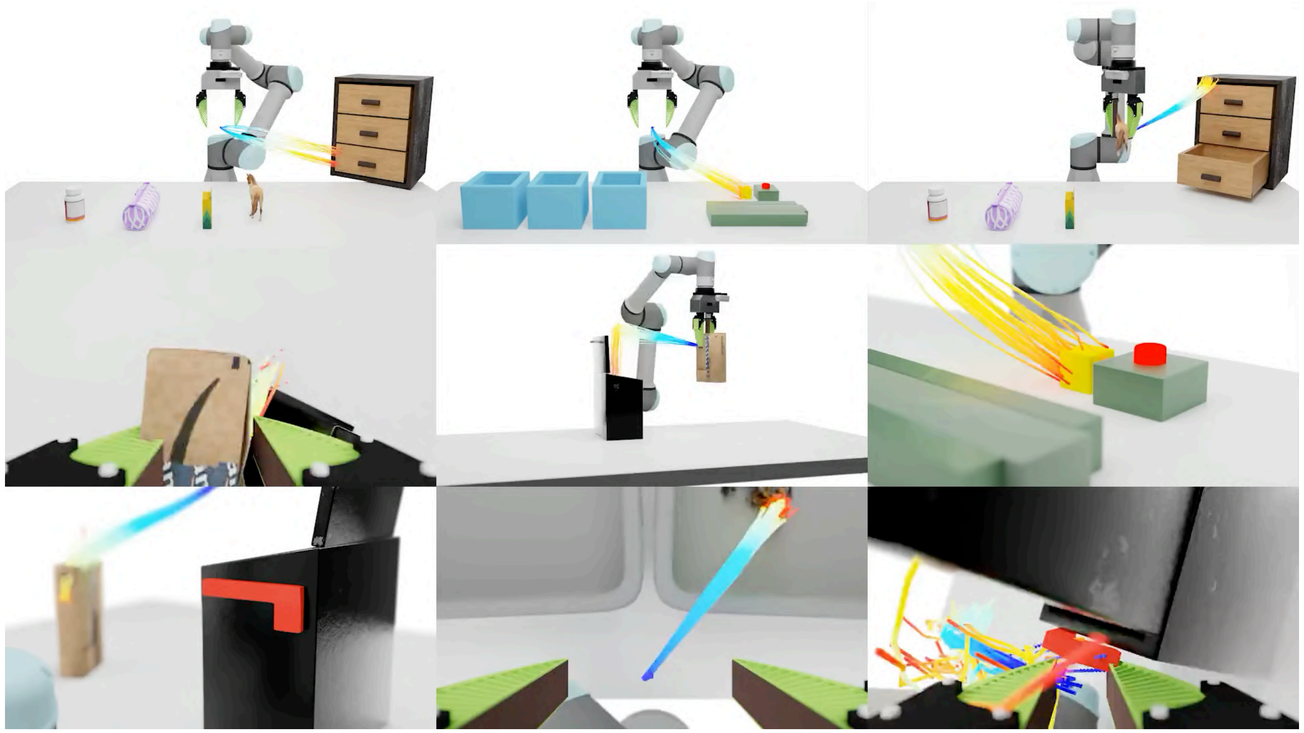
Following our lab's caffinated drinks naming tradition, I've decided to name our quadruped Espresso and our new arm Oat Milk. My hope is that, when combined together, Espresso and Oat Milk will be as capable as Latte (our UR5 which has unfolded and folded cloths and washed dishes). Deploying policies from Latte to Espresso and Oat Milk, as we've done, is the first step 🚀.

Espresso - Oat Milk comes with a GoPro on its head 🎩, a 3D printed gripper at the end of its arm 💪, and an iPhone on its butt 🍑. The GoPro streams visual observations through a capture card, serving as the policy observation. Meanwhile, the iPhone runs a custom iOS app we developed, which streams the robot's body pose, allowing world frame stabilization.

NOTE: almost always eye level with robot.

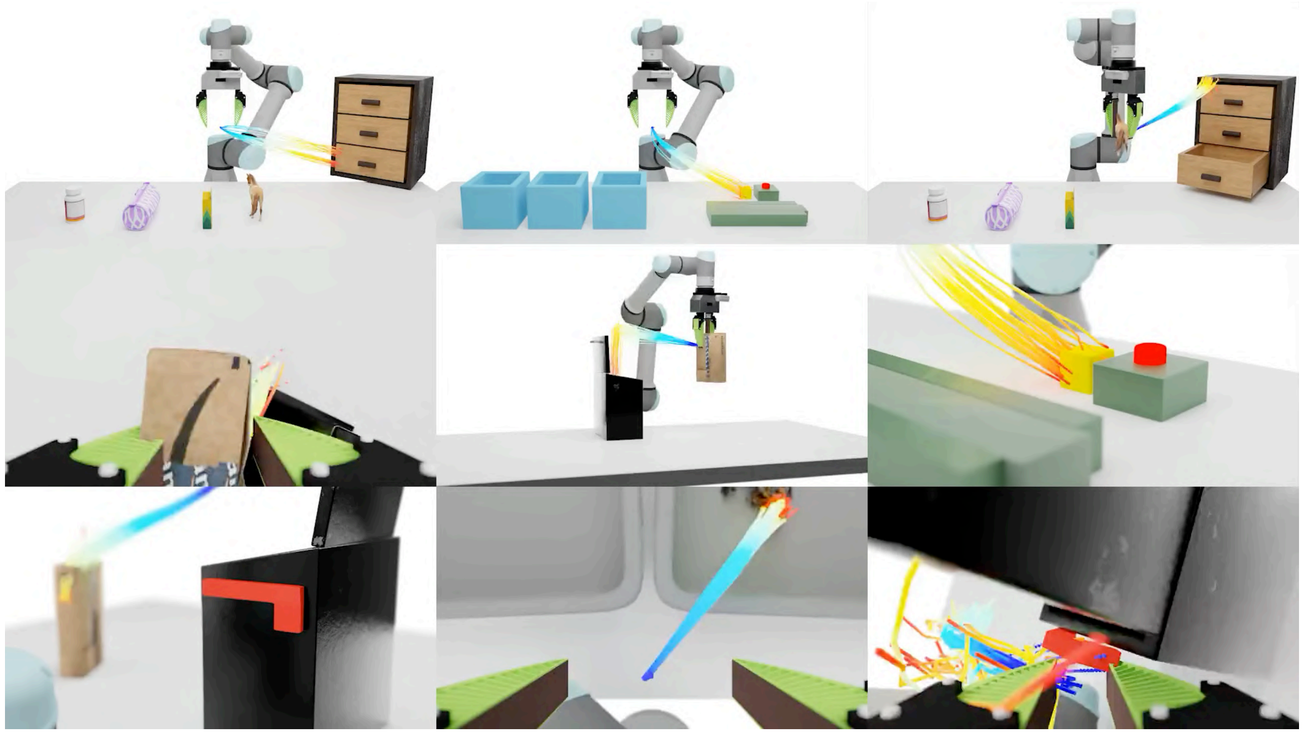
Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.



Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.



Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Movements. Straight lines, or revolving while tracking subject.

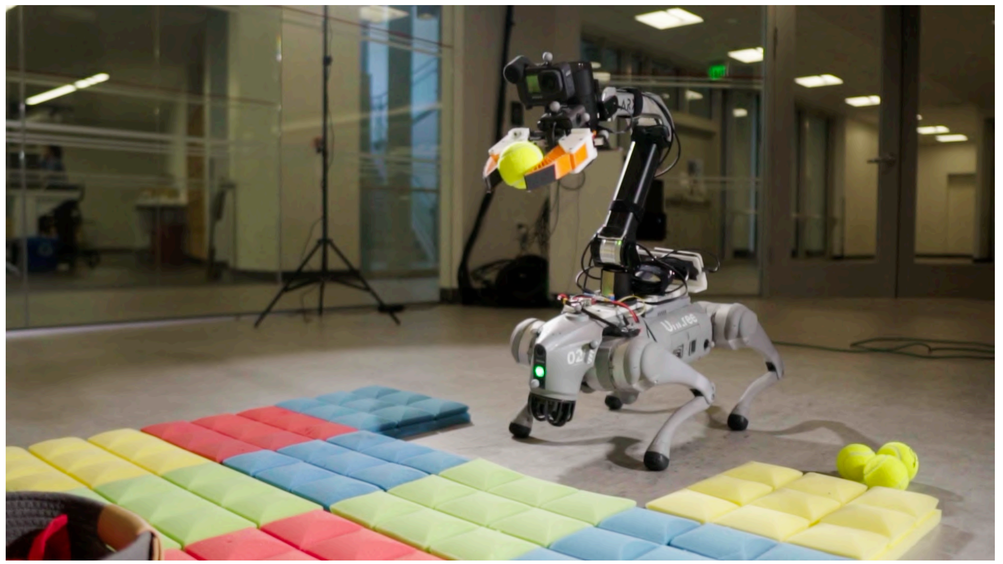
Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Movements. Straight lines, or revolving while tracking subject.

(Bonus) Foreground, Background. Use occlusion/panning for suspense.

Thought process for this shot:
Move backwards, so that first part focuses on robot, but latter part reveals that ball lands into basket.



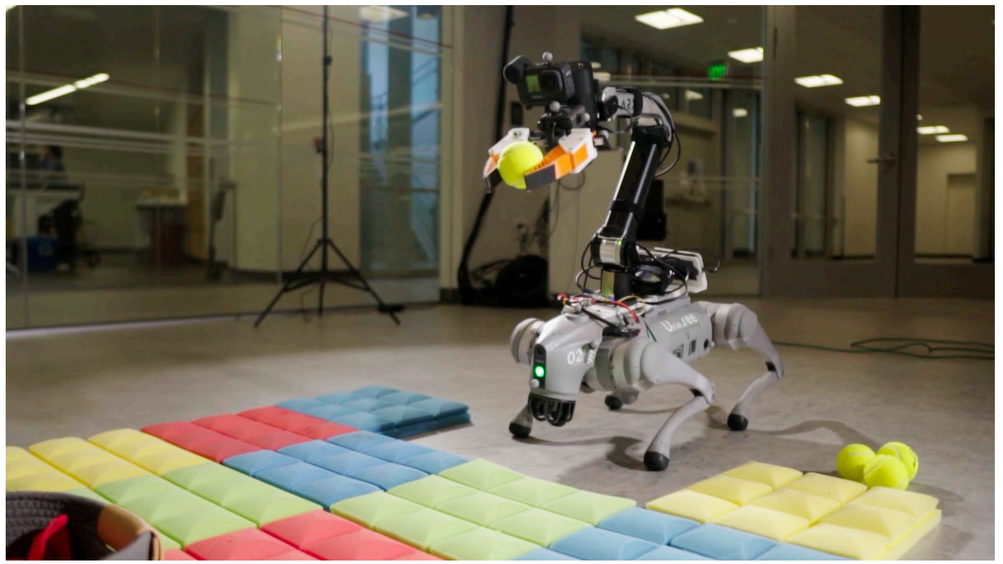
Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Movements. Straight lines, or revolving while tracking subject.

(Bonus) Foreground, Background. Use occlusion/panning for suspense.

Thought process for this shot:
Move backwards, so that first part focuses on robot, but latter part reveals that ball lands into basket.



Characterization of Shots

Focal Lengths. Wider full scene, medium view, close ups, POVs.

Movements. Straight lines, or revolving while tracking subject.

(Bonus) Foreground, Background. Use occlusion/panning for suspense.

Thought process for this shot:
beginning already captures attention, yet it builds up in impressiveness as more information is revealed to viewer. On water => robot arm => robot dog => on a table.

Characterization of Shots

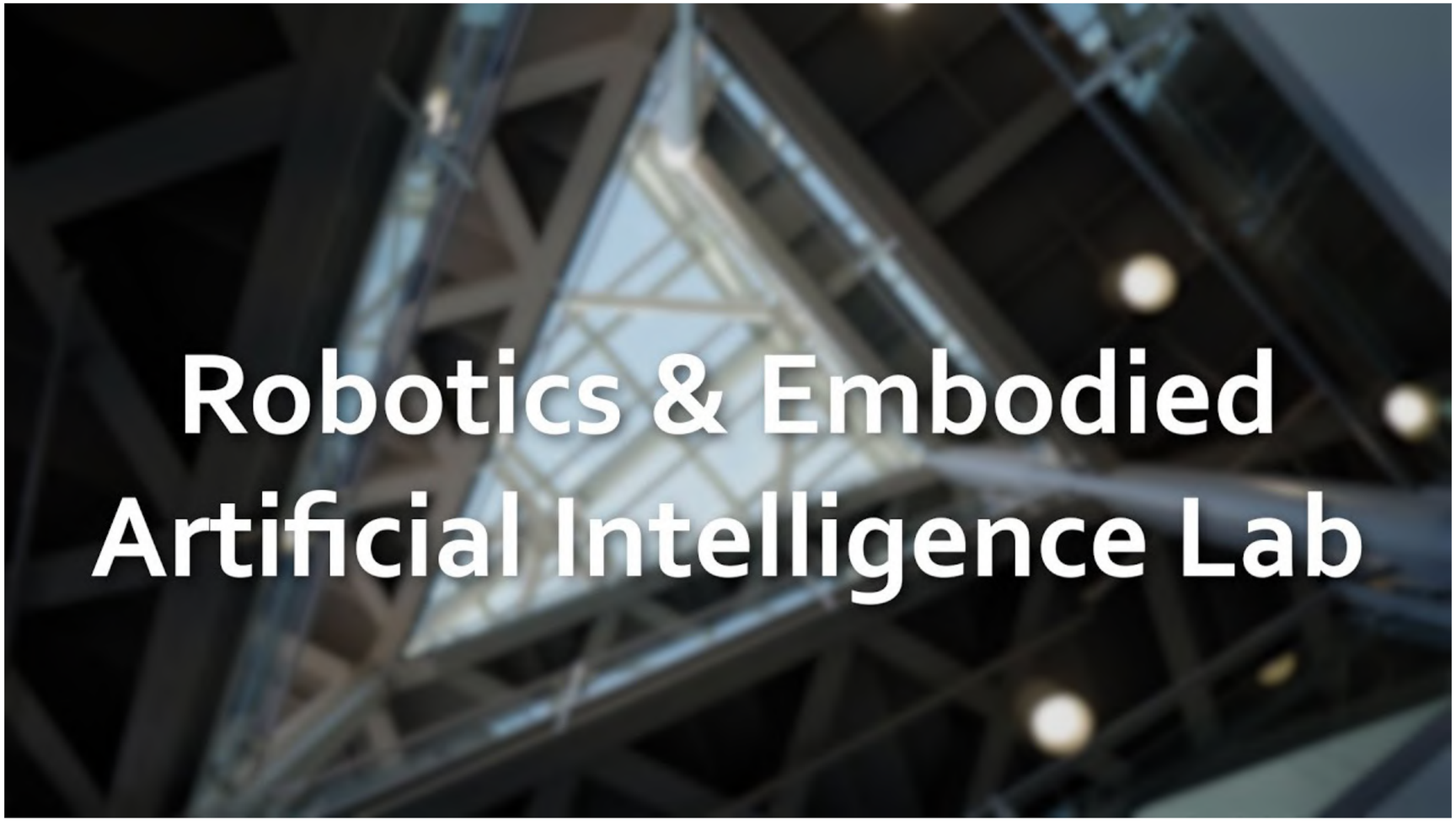
Focal Lengths. Wider full scene, medium view, close ups, POVs.

Movements. Straight lines, or revolving while tracking subject.

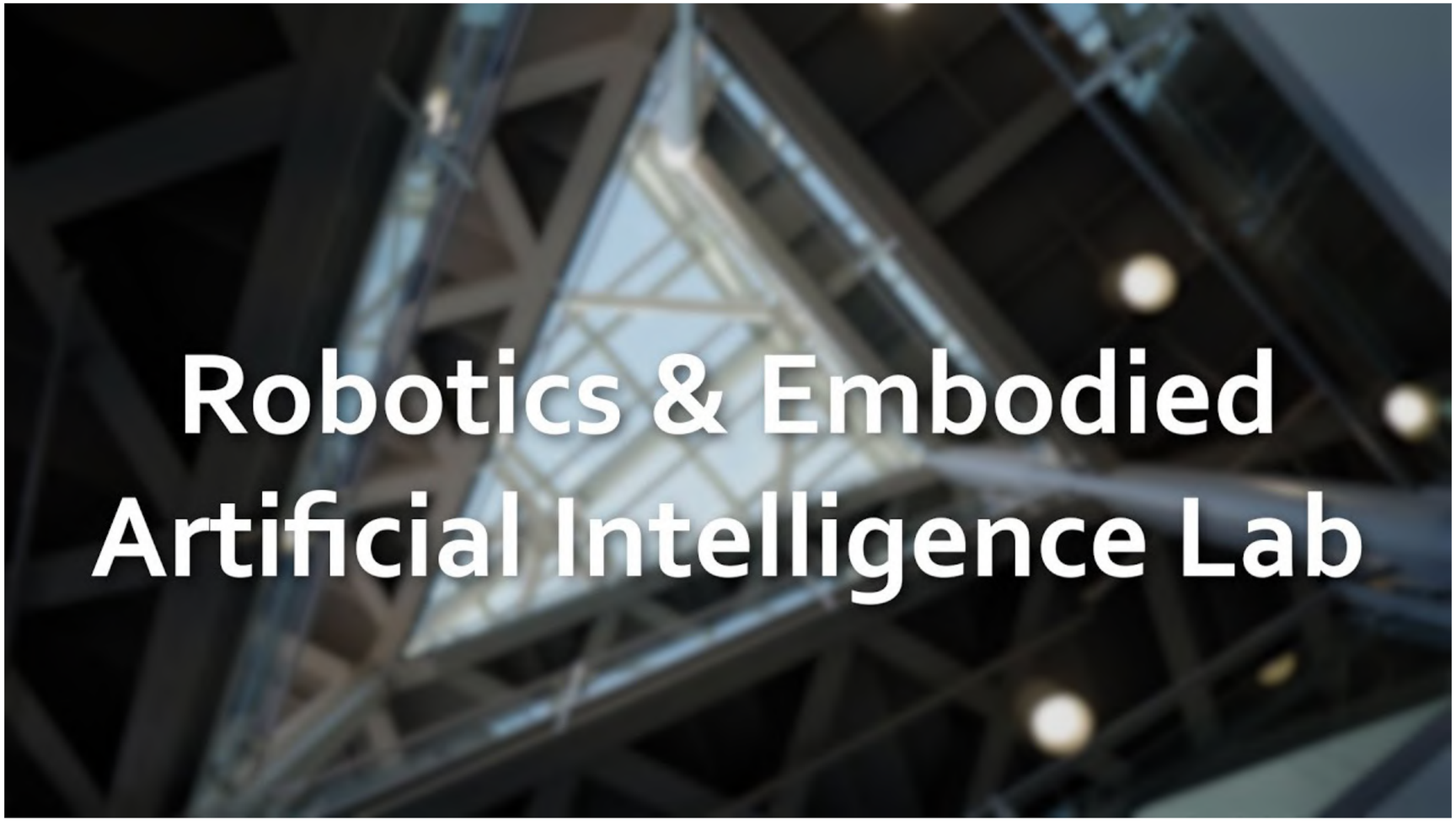
(Bonus) Foreground, Background. Use occlusion/panning for suspense.

Thought process for this shot:
beginning already captures attention, yet it builds up in impressiveness as more information is revealed to viewer. On water => robot arm => robot dog => on a table.

An example of how shots get cut together



An example of how shots get cut together



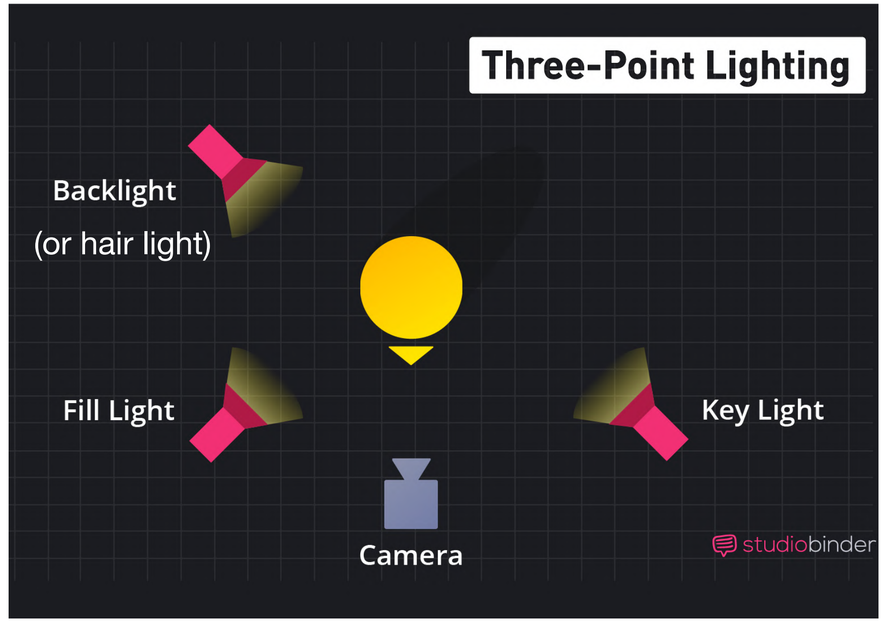
Lights

Much more important than cameras nowadays.

Soft, Diffuse Light. One stand, one video light, one softbox.

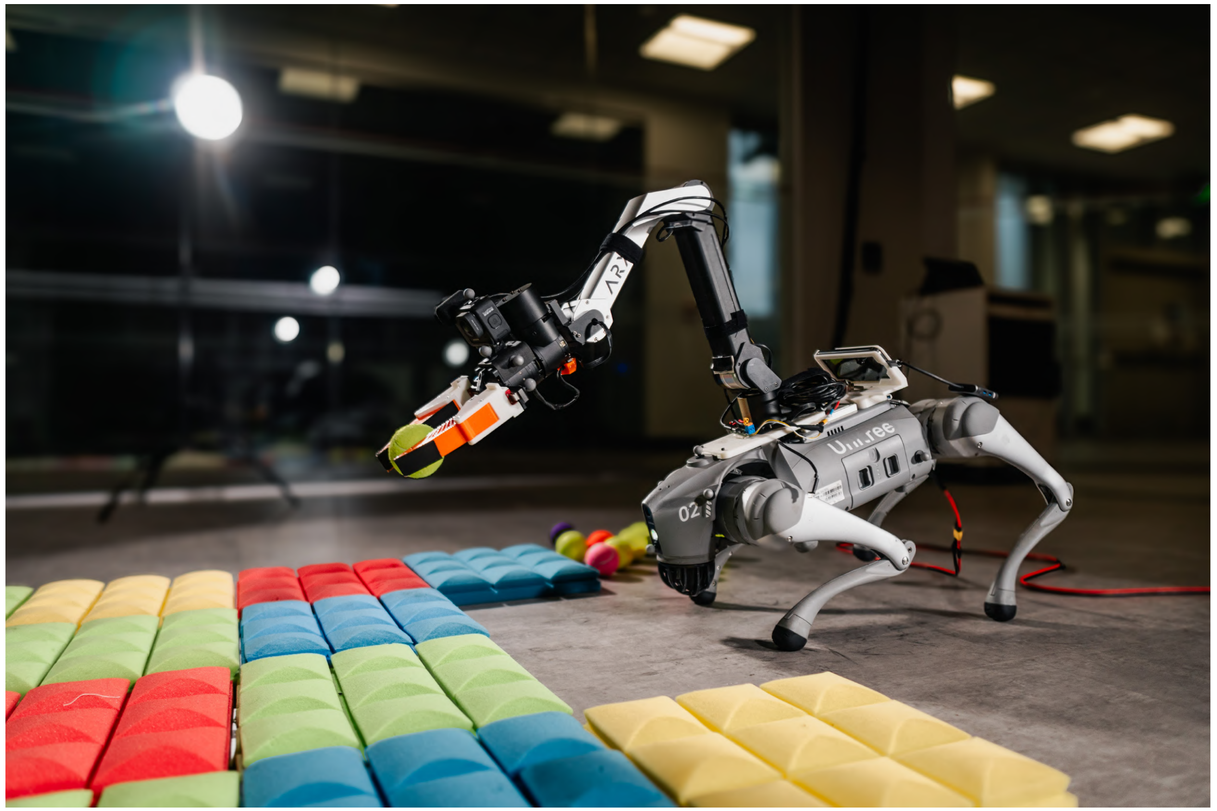
Secondary Fill Light. Finegrained control of shadows.

Hair Light. Helps subject pop!



Lights

I love hair lights



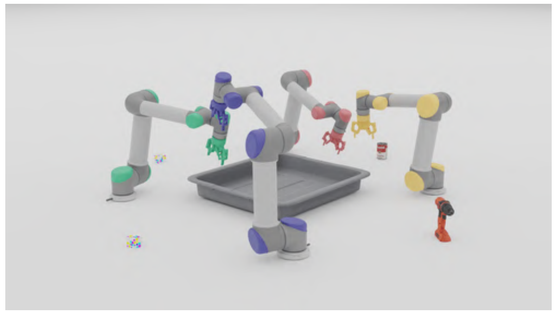

Blender


Blender

“But it’s so hard to use.”

–Robotics PhDs after I tell them they can try using Blender for their next project’s visualizations.

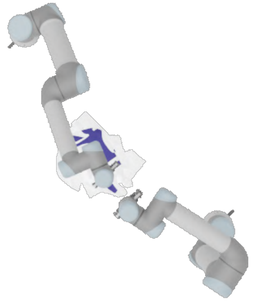
A New Skill Every Project



2020

PyBullet => Blender

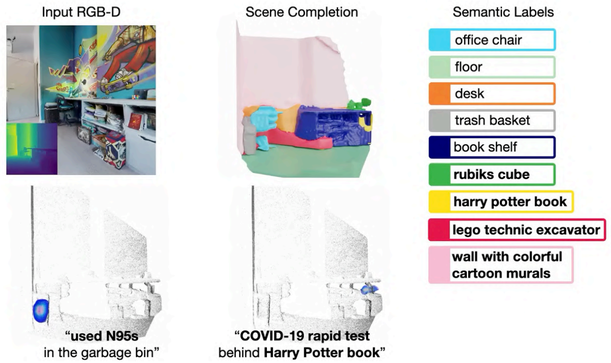
Basic Blender Light, Material, Cam Setup



2021

PyFlex => Blender

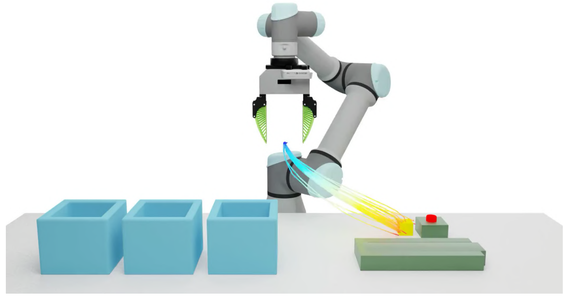
Custom Texturing



2022

AI2 Thor => Blender

Point Cloud Rendering



2023

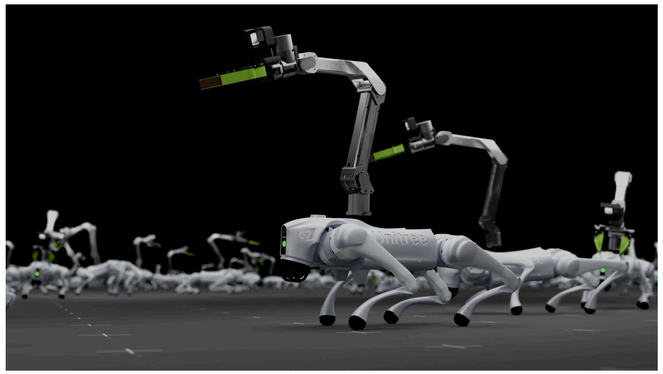
MuJoCo => Blender

Custom Geometry Nodes for Diffusion Policy Inference Traces

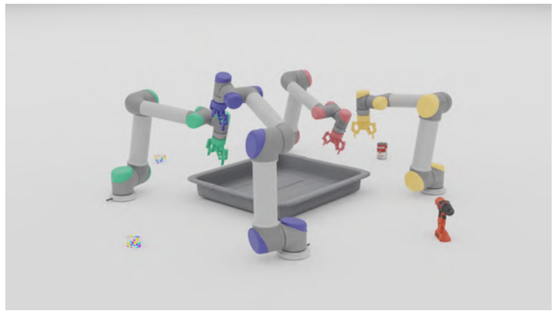
2024

IsaacGym => Blender

Custom Geometry Nodes for large scale animation instancing



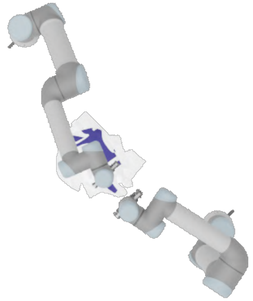
A New Skill Every Project



2020

PyBullet => Blender

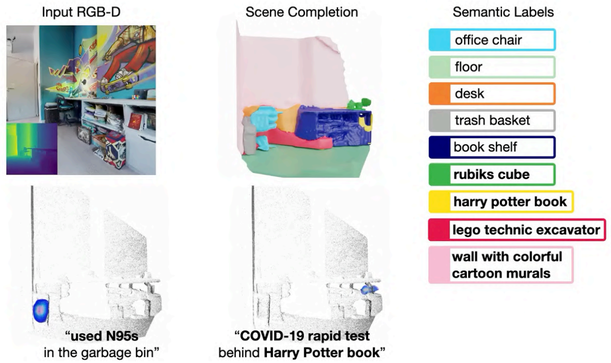
Basic Blender Light, Material, Cam Setup



2021

PyFlex => Blender

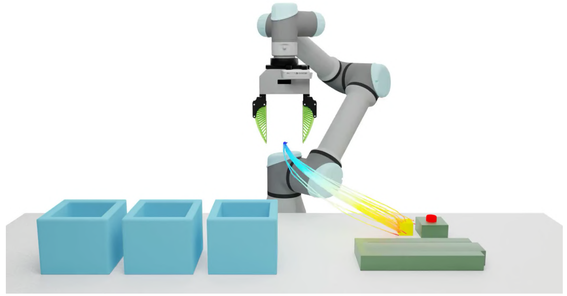
Custom Texturing



2022

AI2 Thor => Blender

Point Cloud Rendering



2023

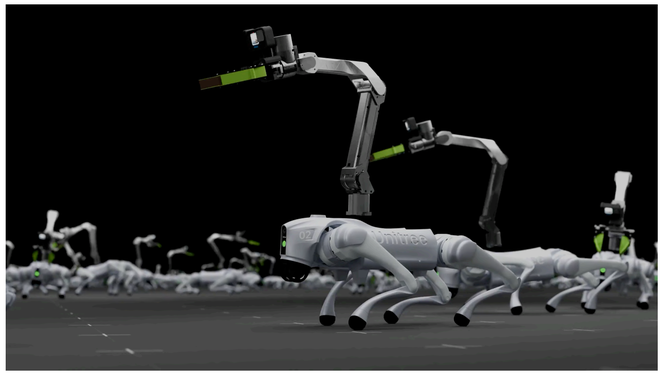
MuJoCo => Blender

Custom Geometry Nodes for Diffusion Policy Inference Traces

2024

IsaacGym => Blender

Custom Geometry Nodes for large scale animation instancing



My Point

start now.

Typical Workflow

Simulation

PyBullet, MuJoCo, …

Setup Blender Scene

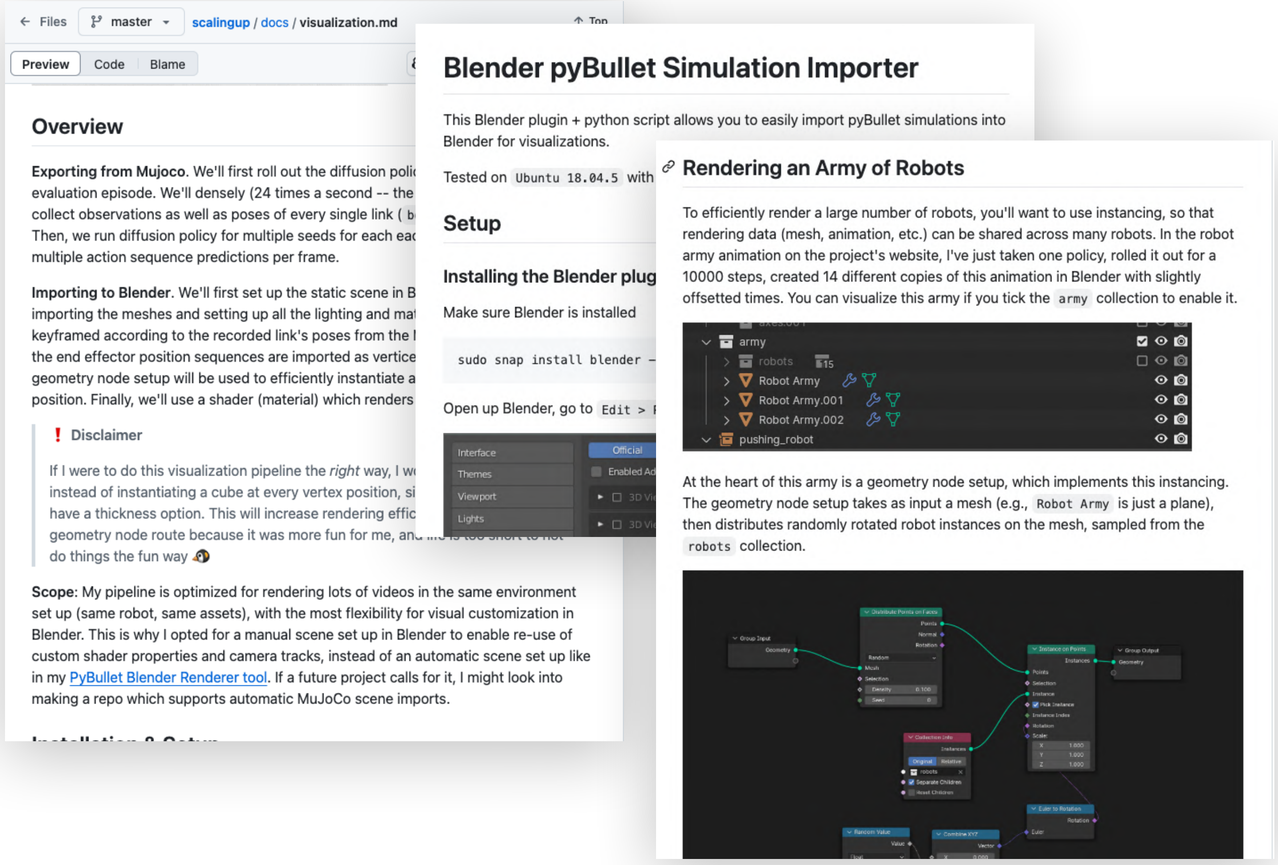
Import robot & obj meshes, materials, cameras, lights

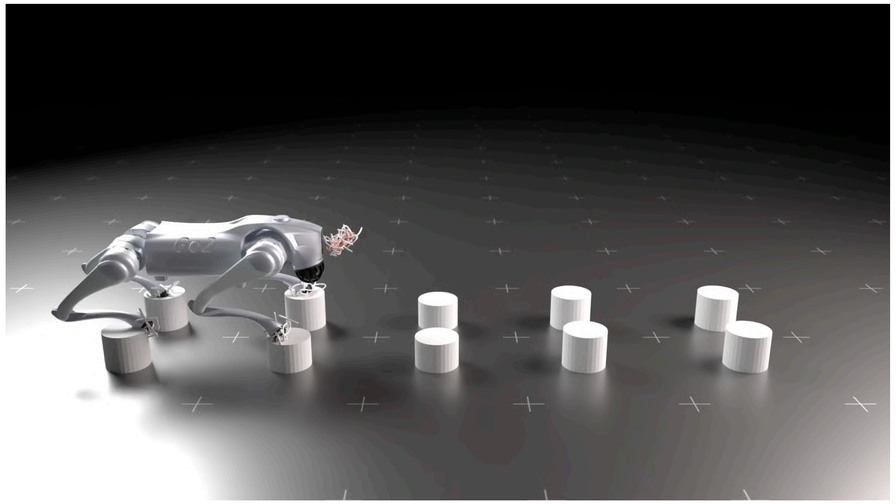
Write importer script

Already open-sourced!

Existing Tools & Resources

Blender Importer scripts are open sourced.

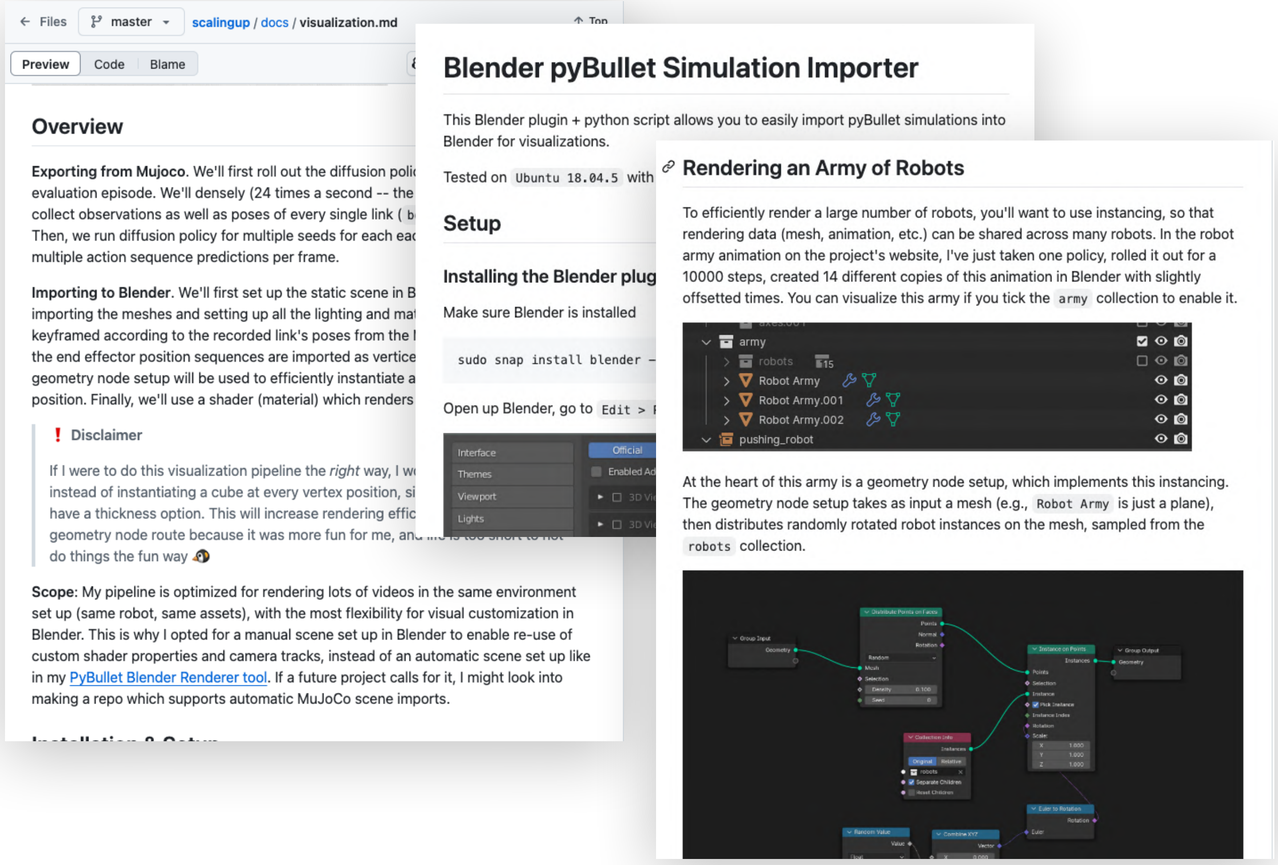


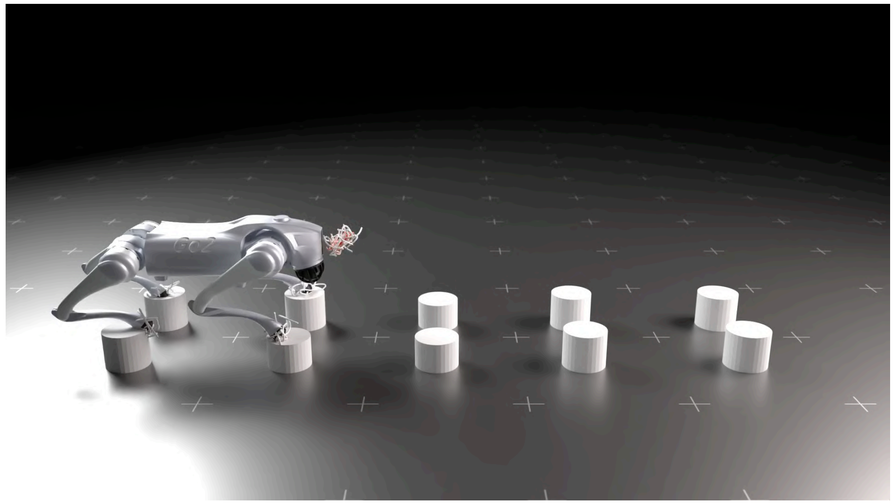


DialMPC from Berkeley.

Existing Tools & Resources

Blender Importer scripts are open sourced.





DialMPC from Berkeley.

Today's Agenda

Treat presentations seriously

A "normal" research presentation

A guideline to presenting your research

- Think about your audience
- Control people's attention
- Make things look pretty

A hopefully improved research presentation! (25 minutes)

Today's Agenda

Treat presentations seriously

A "normal" research presentation

A guideline to presenting your research

- Think about your audience
- Control people's attention
- Make things look pretty

A hopefully improved research presentation! (25 minutes)