EPFL

Expressive Whole-Body Control for Humanoid Robots

Group 12:
Alessio Desogus
Aurélien Laissy
Théo Lemonnier

05.11.2024

EPFL

Introduction

Topic: *Expressive Whole-Body Control in Humanoid Robots*

Main Idea: Introducing ExBody, a framework that enables robots to imitate human-like, expressive movements.

Significance: Aims to improve natural, intuitive human-robot interaction, with applications in assistive and social robotics.

Example Videos

Executive Summary

Robot's type:

- Bipedal humanoid (Unitree H1 platform)
- Human-sized (approx. height and weight similar to a person) with 19 DoFs

Control Methods:

- Position control and goal-conditioned reinforcement learning
- The upper body follows human-like reference movements, while the lower body adapts

Design Methods:

- Reinforcement learning (RL), optimization, and motion retargeting
- Expressive upper-body movements,
- Relax lower body to accommodate hardware constraints and preserve robustness

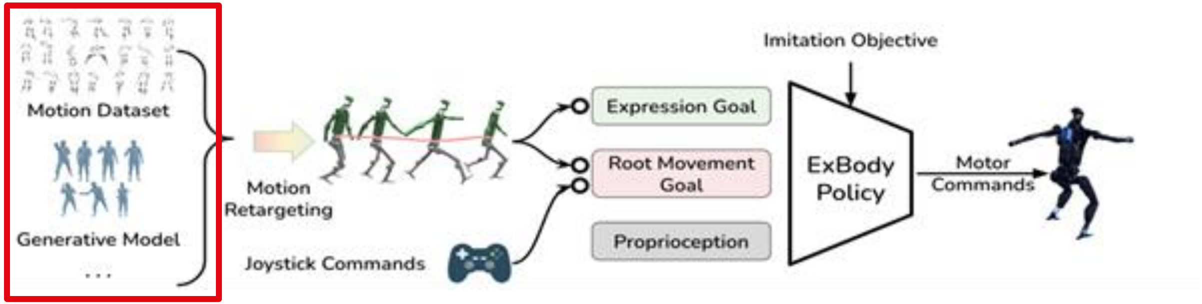
Gait types:

- Walking, Running, Dancing, Social Behaviors and Task-Oriented Movements

Sensors:

- Proprioceptive sensors to track joint angles, angular velocity, and posture.

ExBody Framework



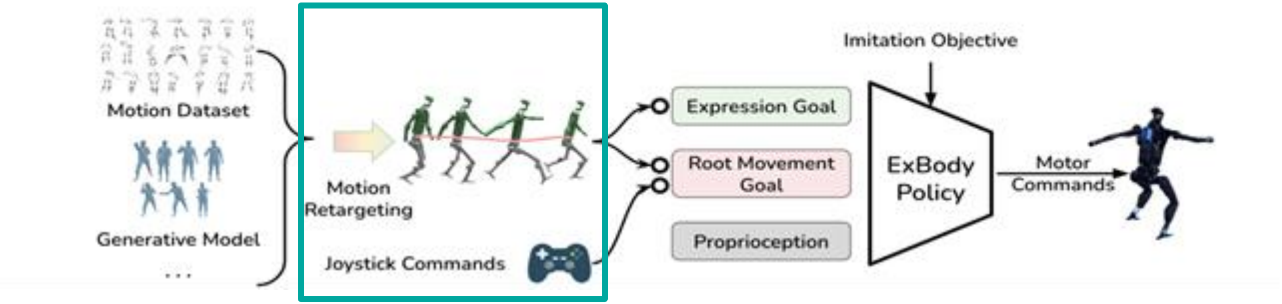
Strategies for Curating Human Behavior Data

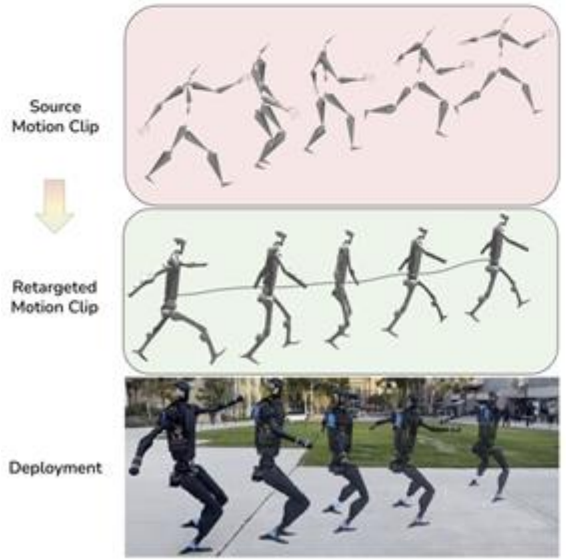
- Training dataset: subset of the **CMU MoCap*** dataset
- Excluding physical interactions with others, heavy objects and rough terrain
  - Total of 780 over 2605 clips
- **Before training the policy:** need to retarget this dataset

	Category	Clips	Length (s)
Training	Walk	546	9076.6
	Dance	78	1552.3
	Basketball	36	766.1
	Punch	20	800.0
	Others	100	1188.0
	Total	780	13383.0

One of the most popular dataset for human motion: http://mocap.cs.cmu.edu/

ExBody Framework

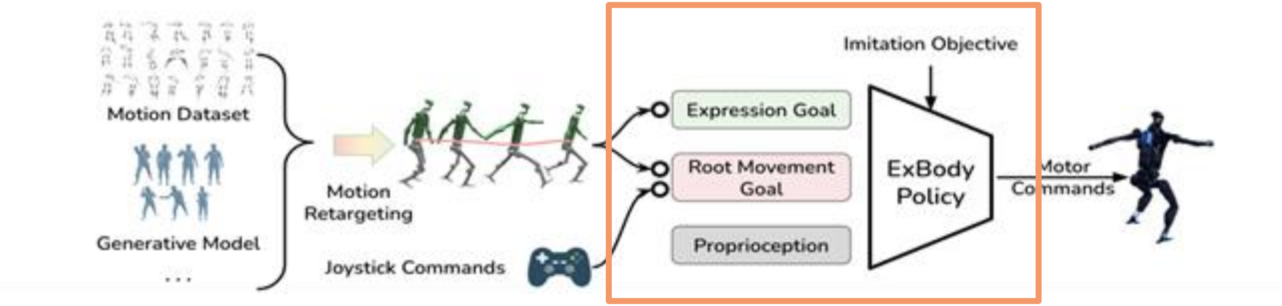




Motion Retargeting to Hardware:

- Adapt the human motion to the robot's framework to account for the **morphological differences**
- 19 DoFs (Unitree H1) vs up-to 69 DoFs (human motion)
- Mapping local joint rotations onto the robot's skeleton
  - Uses a quaternion-based approach

ExBody Framework



Reinforcement Learning Setup:

- Goal-conditioned motor policy **π: G×S→A** where:
  - **G:** the goal space (behavioral objectives)
  - **S:** the observation space
  - **A:** the action space, containing joint positions and torque

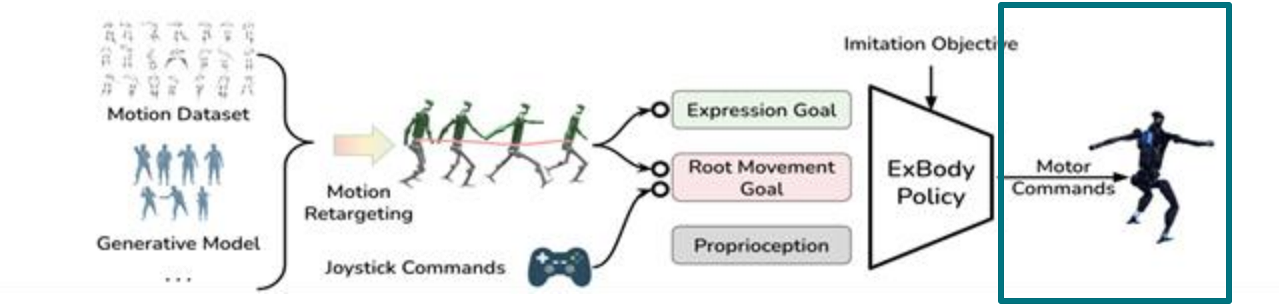
Term	Expression	Weight
Expression Goal G^e		
DoF Position	$\exp(-0.7\lvert \mathbf{q}_{\text{ref}} - \mathbf{q}\rvert)$	3.0
Keypoint Position	$\exp(-\lvert \mathbf{p}_{\text{ref}} - \mathbf{p}\rvert$	2.0
Root Movement Goal G^m		
Linear Velocity	$\exp(-4.0\lvert \mathbf{v}_{\text{ref}} - \mathbf{v}\rvert)$	6.0
Roll & Pitch	$\exp(-\lvert \mathbf{\Omega}^{\phi\theta}_{\text{ref}} - \mathbf{\Omega}^{\phi\theta}\rvert)$	1.0
Yaw	$\exp(-\lvert \Delta y\rvert)$	1.0

Goal Space Design:

- Composite goal space **G = Ge × Gm** where:
  - **Ge** (Expression goal): joint positions of 9 upper body actuators, 3D keypoints of shoulders, elbows, and hands (18 points)
  - **Gm** (Root movement goal): linear velocity ($v \in R^3$), body pose (roll/pitch/yaw) and body height

LEGGED ROBOTS

ExBody Framework

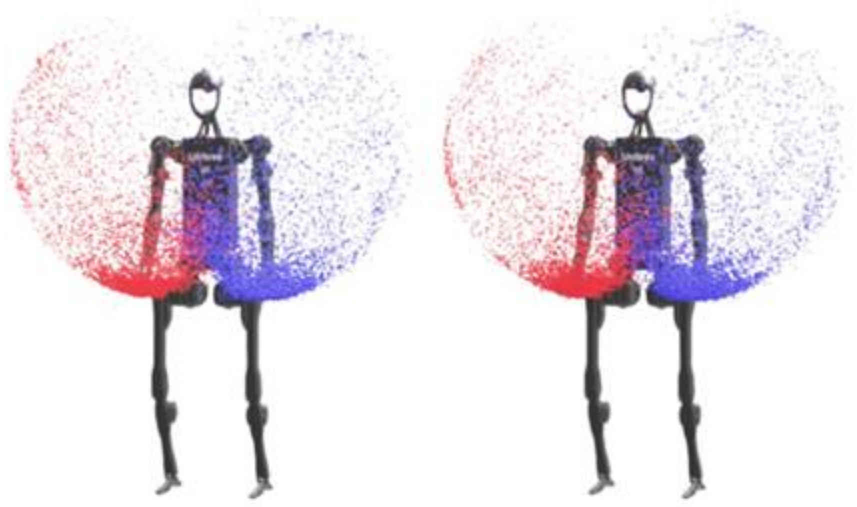


Real-World Deployment:

- During RL training: **develops a robust internal model** of how to move and react to different movement goals
  - Even when these aren't explicitly provided as complete trajectories
- Allows **intuitive control in real-world** settings without needing predefined motion clips
- Joystick inputs mapped to root movement goals (e.g., walking direction and speed)

Results

1. Good tracking of root and upper body



Retargeted motion datasets | Learned ExBody policy rollout

Results

1. Good tracking of root and upper body
2. Good to learn from large motion datasets

Baselines	Motion Sample				Random Sample			
	MEL↑	MELV↑	MERP↑	MEK↑	MEL	MELV	MERP	MEK
→ ExBody (Ours)	16.87	**318.67**	754.92	**659.78**	13.51	**132.14**	523.79	483.67
ExBody + AMP	**17.28**	205.60	**765.85**	635.51	15.59	95.11	583.82	**544.59**
ExBody + AMP NoReg	16.16	87.83	714.74	561.56	15.40	36.76	584.23	515.53
No RSI	0.23	0.63	10.09	7.25	0.22	0.10	7.41	7.15
Random Sample	16.50	181.85	704.73	326.66	16.37	38.51	**586.83**	324.10
Full Body Tracking	13.28	246.11	584.40	397.25	10.76	76.46	407.88	284.69

TABLE IV: Comparisons with baselines. We sample 10,000 trajectories with 4096 environments in simulation and report their mean episode metrics. Motion Sample means we sample g^m from retargeted motions. Random Sample means we uniformly sample g^m in Fig. 3.

Results

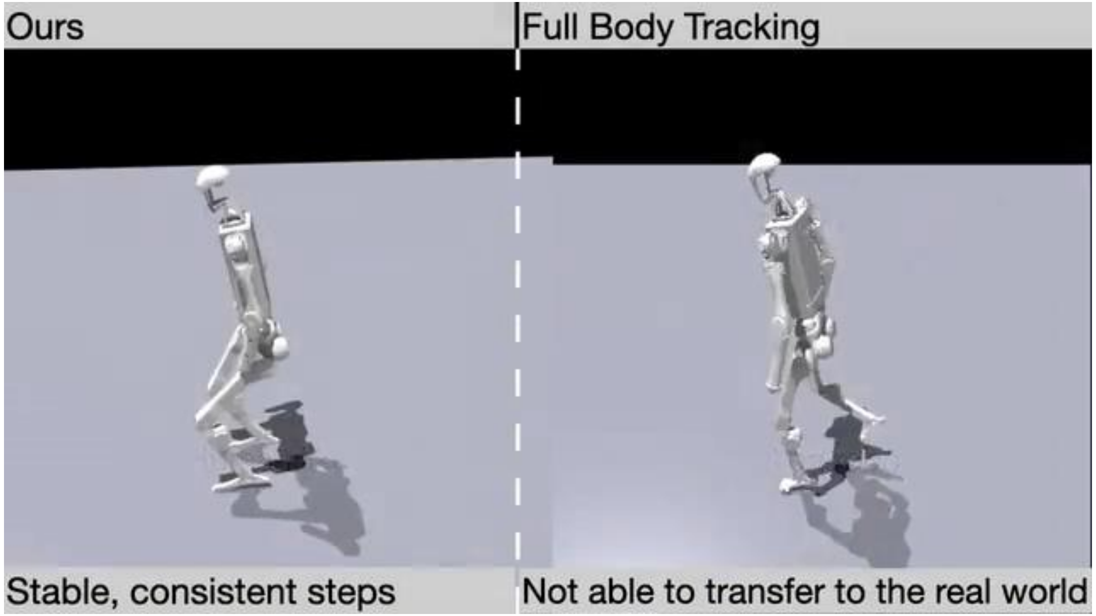
1. Good tracking of root and upper body
2. Good to learn from large motion datasets
3. Better performance with no full DoF tracking

Baselines	Motion Sample				Random Sample			
	MEL↑	MELV↑	MERP↑	MEK↑	MEL	MELV	MERP	MEK
ExBody (Ours)	16.87	**318.67**	754.92	**659.78**	13.51	**132.14**	523.79	483.67
ExBody + AMP	**17.28**	205.60	**765.85**	635.51	15.59	95.11	583.82	**544.59**
ExBody + AMP NoReg	16.16	87.83	714.74	561.56	15.40	36.76	584.23	515.53
No RSI	0.23	0.63	10.09	7.25	0.22	0.10	7.41	7.15
Random Sample	16.50	181.85	704.73	326.66	16.37	38.51	**586.83**	324.10
→ Full Body Tracking	13.28	246.11	584.40	397.25	10.76	76.46	407.88	284.69

TABLE IV: Comparisons with baselines. We sample 10,000 trajectories with 4096 environments in simulation and report their mean episode metrics. Motion Sample means we sample g^m from retargeted motions. Random Sample means we uniformly sample g^m in Fig. 3.

Results

1. Good tracking of root and upper body
2. Good to learn from large motion datasets
3. Better performance with no full DoF tracking



Article Impact

Who's citing this article ?

- Cited 38 times (arXiv)
- Survey, research paper on whole body control, company Nvidia
- US universities (especially CMU); Asian universities (Tsinghua, Shangaï, NTU,...)

Criticism or discussions surrounding the article

- Limited expressiveness (full body motion also rely on lower body articulation)
- Article is not peer reviewed (at least in arXiv)
- Insufficient limitations discussion within the article
- Considered as promising, great potential for RL based methods

Pros and Cons

Pros

- Performs expressive tasks
- RL robust to real-world variations
- Modular training for tasks

Cons

- Lower body not expressive
- Loss of information during retargeting
- Tuning of reward functions is hard

Conclusion

Balanced approach: upper body is expressive, lower is robust

***Future potential*:** paves the way for the development of versatile and reliable social humanoid robots

Exam Questions

1. Why **motion retargeting** is necessary before learning the RL policy?

1. How does the **ExBody** approach balance the need for expressive upper-body movements with the stability requirements of the lower-body in humanoid robots?

Questions & Feedback ?

Group 12:
Alessio Desogus
Aurélien Laissy
Théo Lemonnier

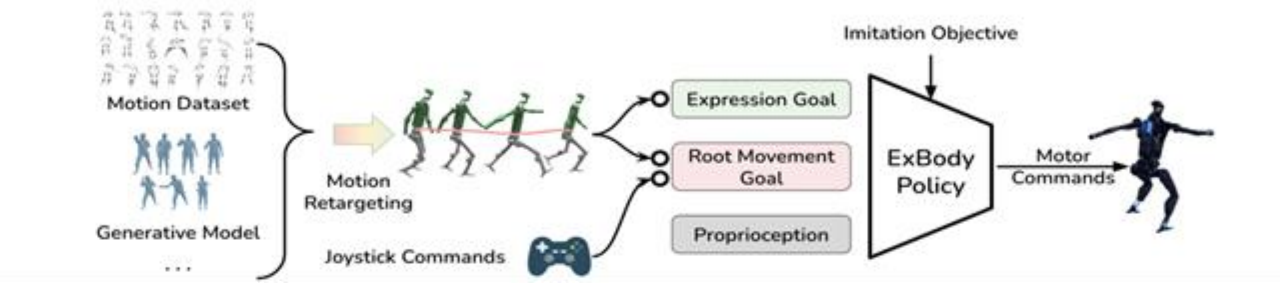
05.11.2024

Exam Questions Answers

1. **Motion retargeting** is necessary to account for the morphological differences between the H1 robot and the human motion data, which has more degrees of freedom (DoFs)

1. **The ExBody framework** achieves this balance by applying reinforcement learning to prioritize expressive movements in the robot's upper body while relaxing control constraints on the lower body. This approach allows the upper body to mimic human-like gestures (e.g., waving, dancing) without requiring exact replication in the legs, which would otherwise risk stability. Instead, the lower body focuses on maintaining robust locomotion.

Appendix - ExBody Framework



Approach Overview:

- Framework designed to enable humanoid robots to generate rich, diverse, and expressive motions while **maintaining stability** in the real world

Key idea:

- Not mimic exactly the same as the reference motion but train a novel controller to take both reference motion and a root movement command as inputs
    - **Upper body:** Imitates expressive human motions
    - **Lower body:** Focuses on robust locomotion without strict imitation constraints

Appendix - Results

- **AMP:** Attention Map Penalization is a regularization technique used in machine learning to encourage models to distribute attention across broader areas of input, reducing over-reliance on small regions and improving generalization.

- **AMP NoReg:** AMP, without regularization, refers to using attention maps directly in model training without any constraints to spread attention, which may lead to over-focusing on specific regions and increase the risk of overfitting.

- **No RSI:** In the context of AMP, RSI (Relevance Score Inference) is a technique used to help the model infer and smooth out the relevance of different areas within attention maps, encouraging a more balanced focus across broader regions. This process aims to reduce the model's tendency to over-concentrate on specific regions, enhancing its ability to generalize and avoid overfitting.

- **Random Sample:** is a sample randomly selected from a dataset, without relying on any specific sequence of movements or actions.

- **Motion sample :** is a specific sample extracted from a sequence of human or robotic movements or actions.