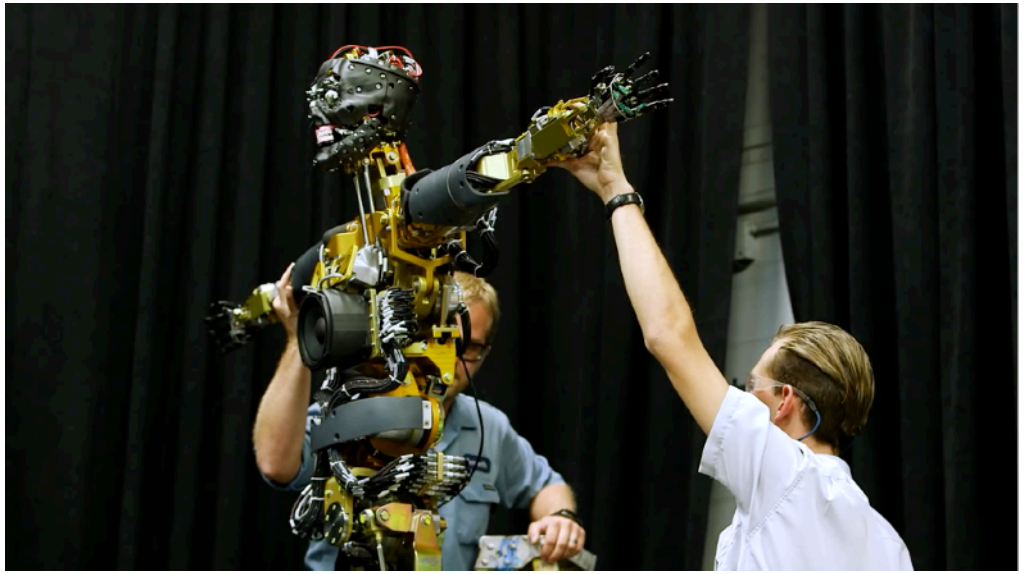
Motion Retargeting

From Optimization-based to Learning-based

Sungjoon Choi, Korea University

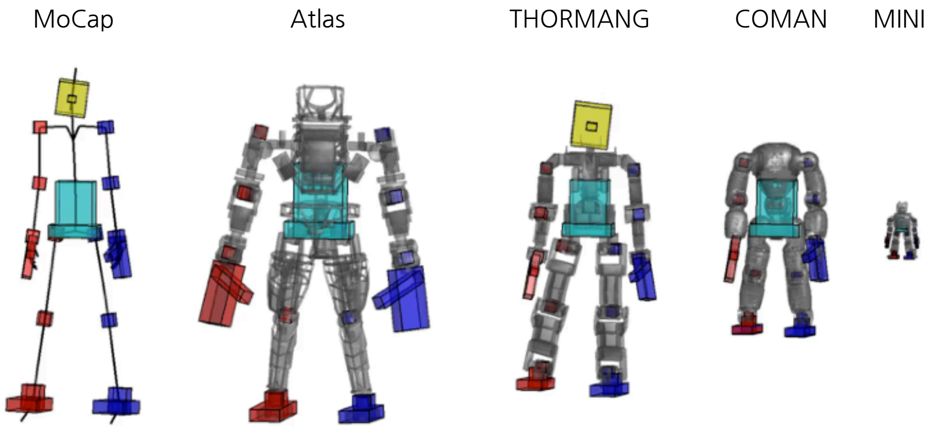
Why Robot Motions?

ROBOT INTELLIGENCE LAB

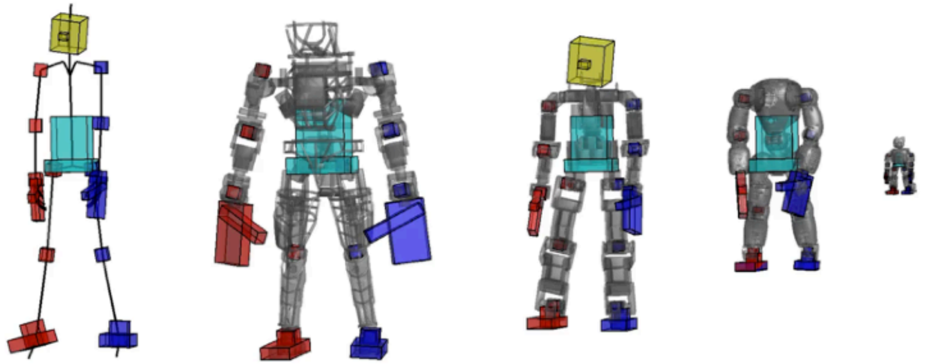


Motion Retargeting

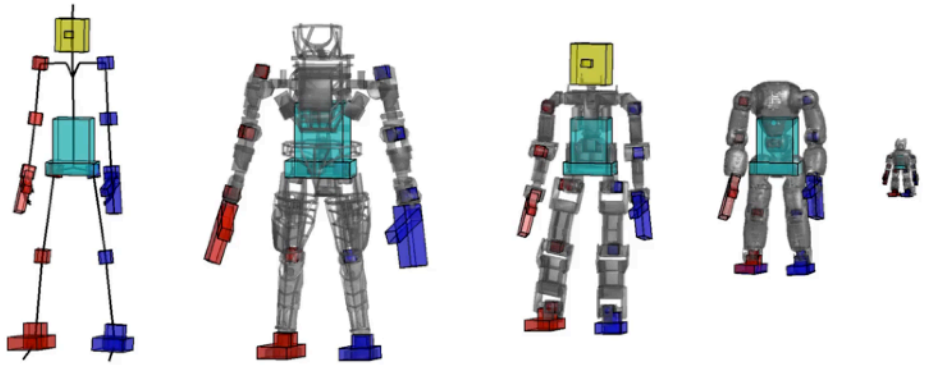
Angry Motion



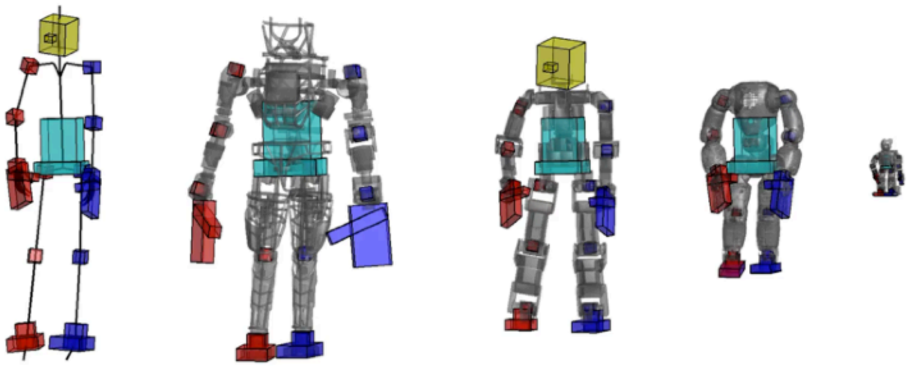
Sad Motion



Happy Motion



Surprise Motion



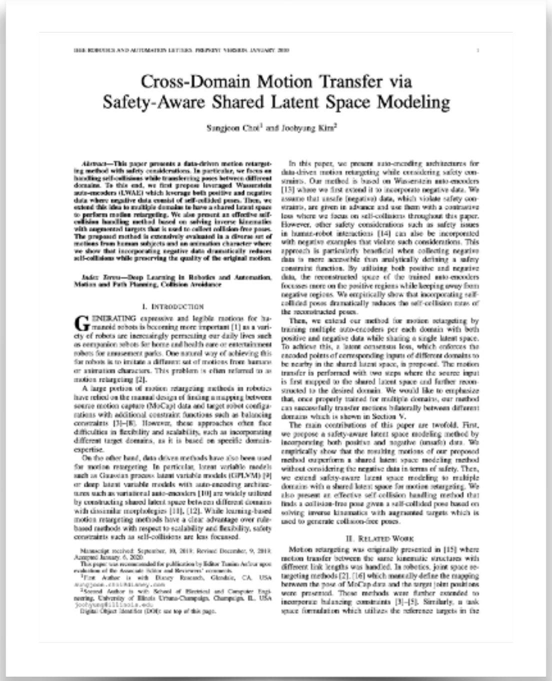
Motion Retargeting Tetralogy

ROBOT INTELLIGENCE LAB


Towards a Natural Motion Generator: a Pipeline to Control a Humanoid based on Motion Data

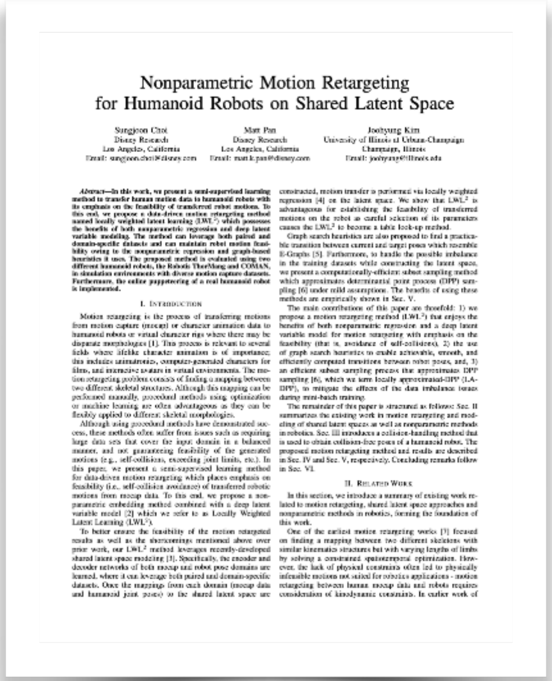
Sungjoon Choi and Joohyung Kim

IROS 2019


Cross-Domain Motion Transfer via Safety-Aware Shared Latent Space Modeling

Sungjoon Choi[1] and Joohyung Kim[2]

ICRA 2020


Nonparametric Motion Retargeting for Humanoid Robots on Shared Latent Space

Sungjoon Choi, Matt Pan, Joohyung Kim

RSS 2020


Self-Supervised Motion Retargeting with Safety Guarantee

Sungjoon Choi[1], Min Jae Song[2], Hyemin Ahn[3], and Joohyung Kim[4]

ICRA 2021

Towards a Natural Motion Generator

"Towards a Natural Motion Generator: a Pipeline to Control a Humanoid based on Motion Data", IROS 2019

Key Question

ROBOT INTELLIGENCE LAB

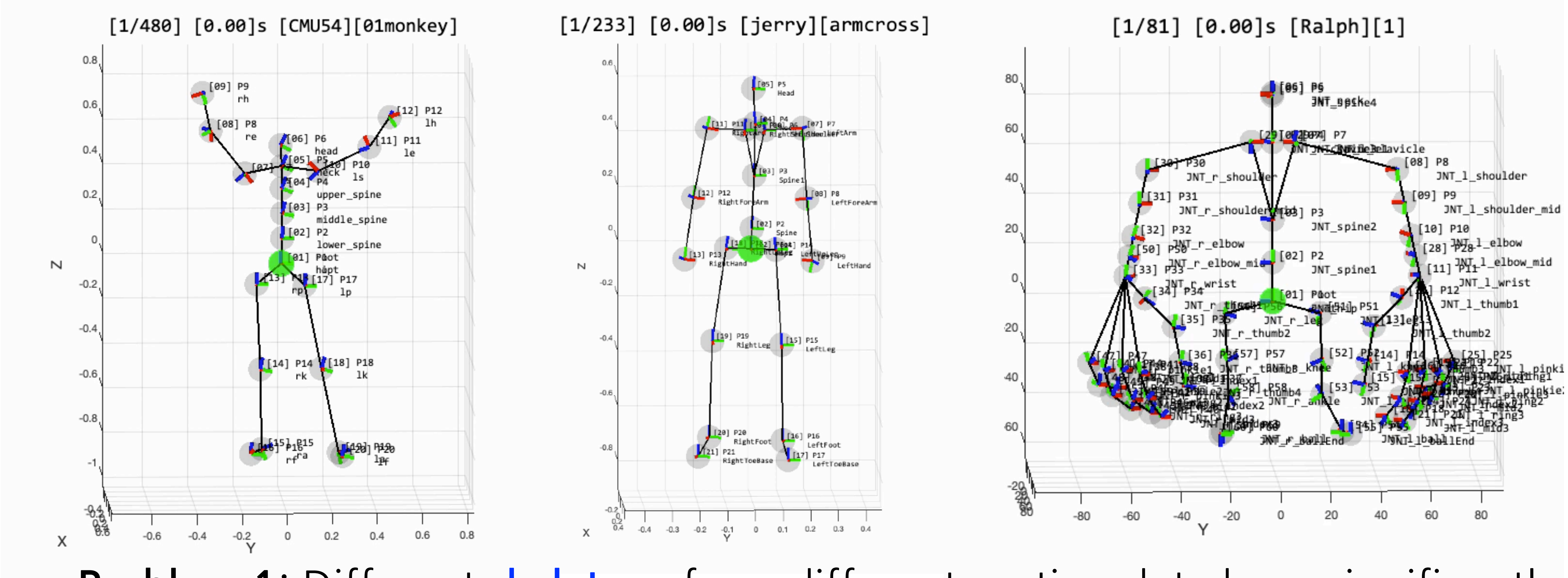
How can we generate **robotics motions** from the **motions of humans or animation characters**?

Problem 1



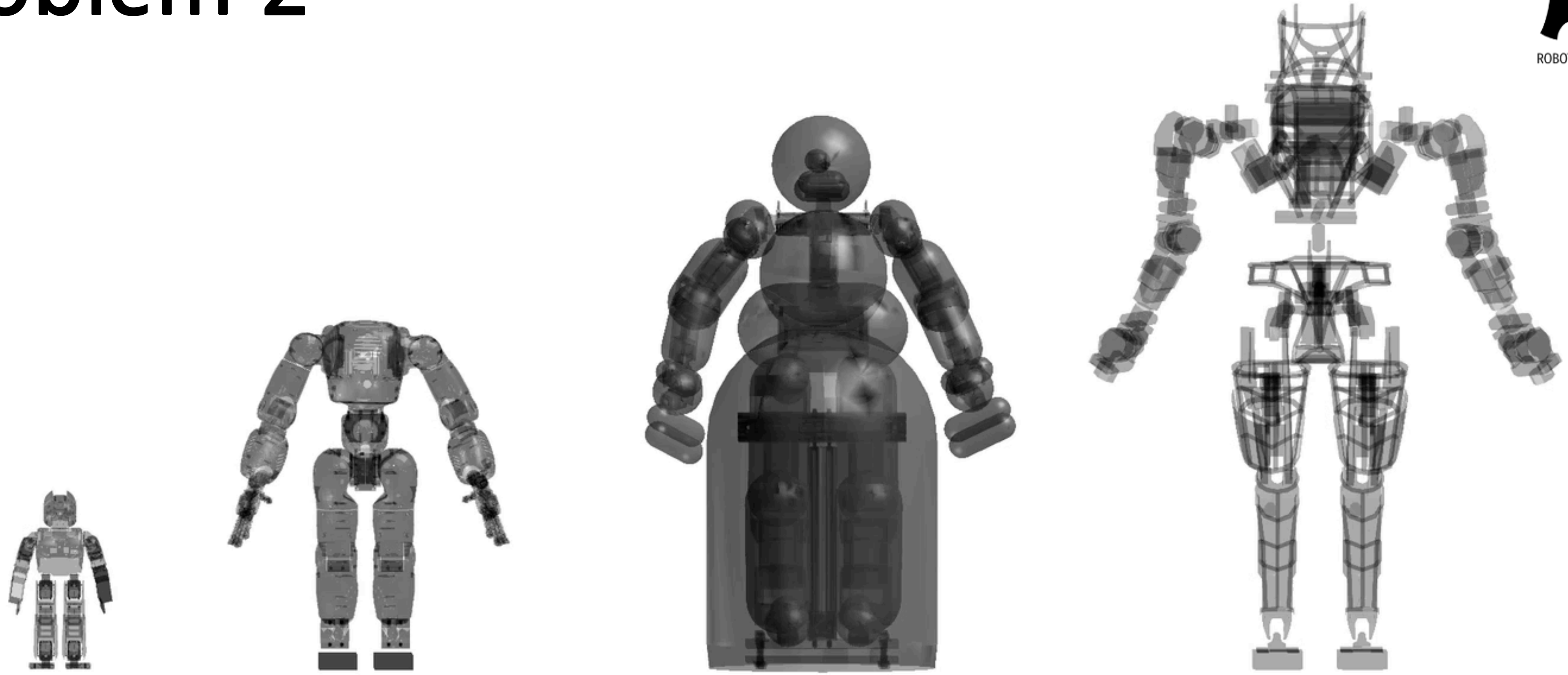
- **<u>Problem 1</u>**: Different skeletons from different motion data have significantly **different shapes** and **morphologies**.

Problem 1



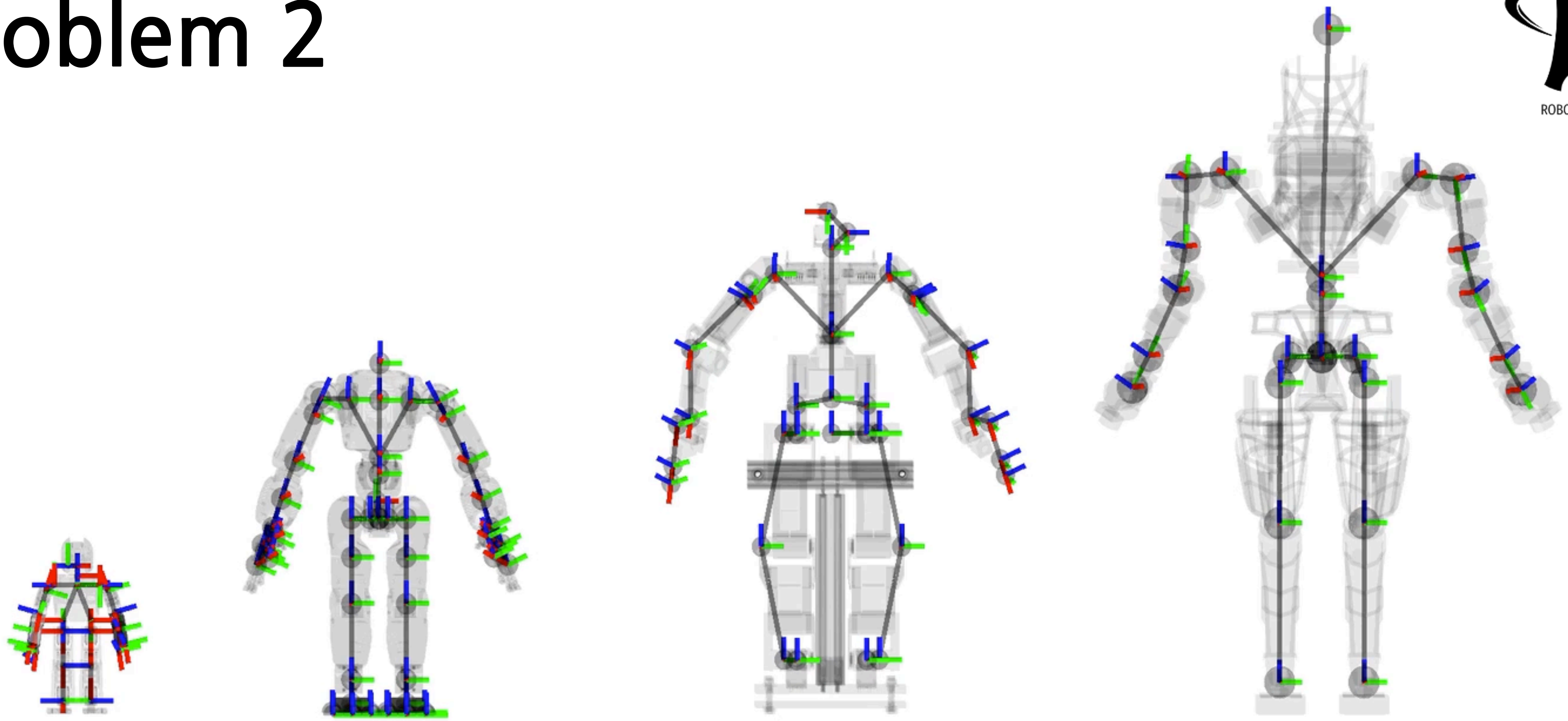
- **Problem 1**: Different skeletons from different motion data have significantly **different shapes** and **morphologies**.

Problem 2



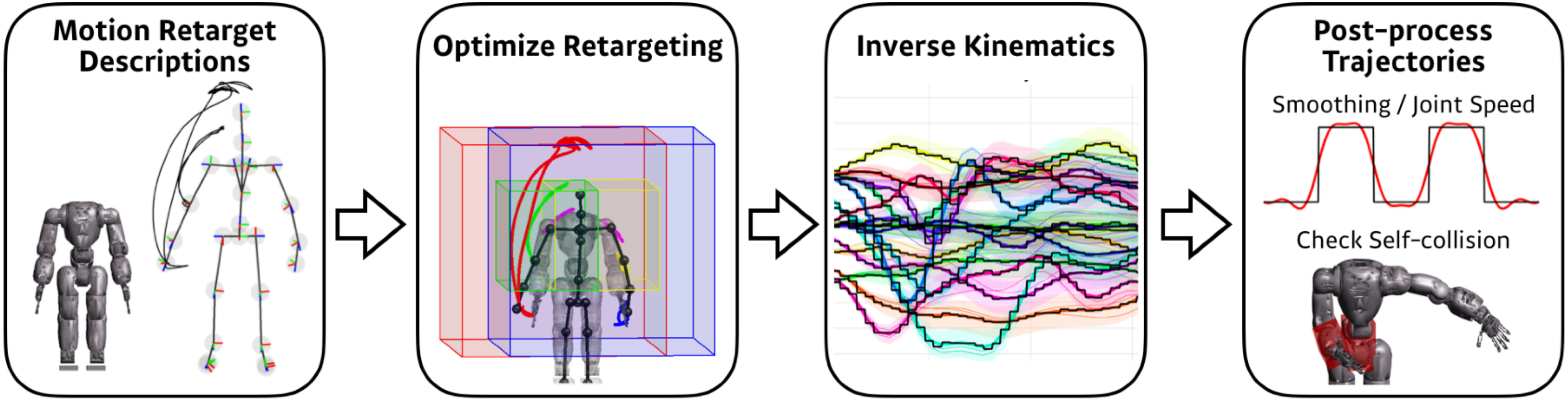
- **<u>Problem 2</u>**: Different **humanoid robots** have significantly **different shapes** and **morphologies**.

Problem 2



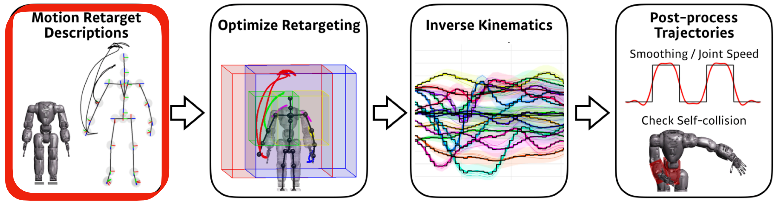
- **Problem 2**: Different humanoid robots have significantly **different shapes** and **morphologies**.

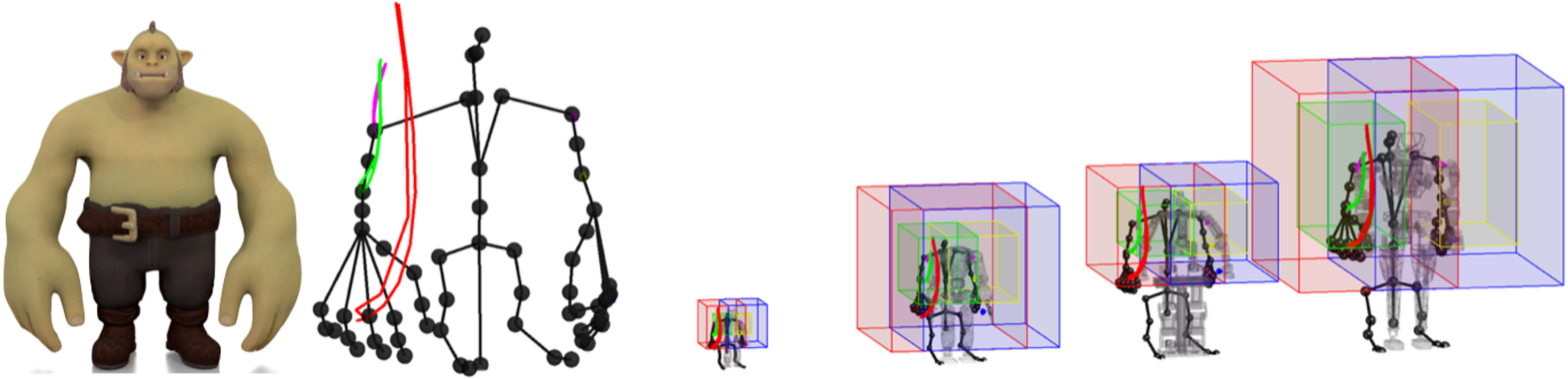
Motion Retargeting



- We present an **optimization-based** motion retargeting method that generates robotic motions from motion capture or animation data.

Motion Retargeting

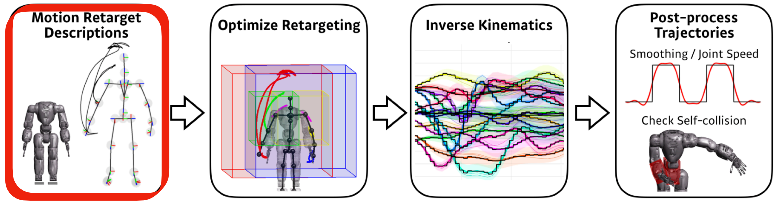


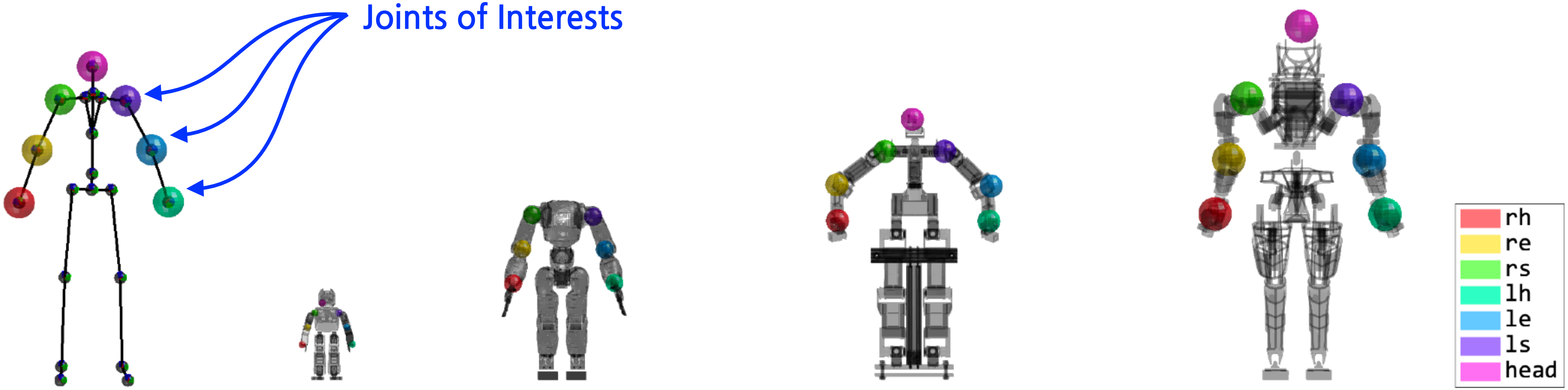


- The cornerstone of the proposed method is a gradient-free optimization-based motion retargeting that adjusts source motion skeleton to fit the target robot.

Joints of Interest

ROBOT INTELLIGENCE LAB

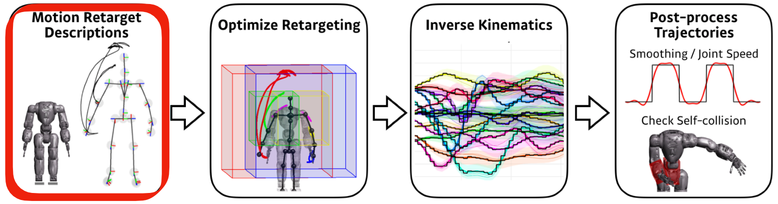


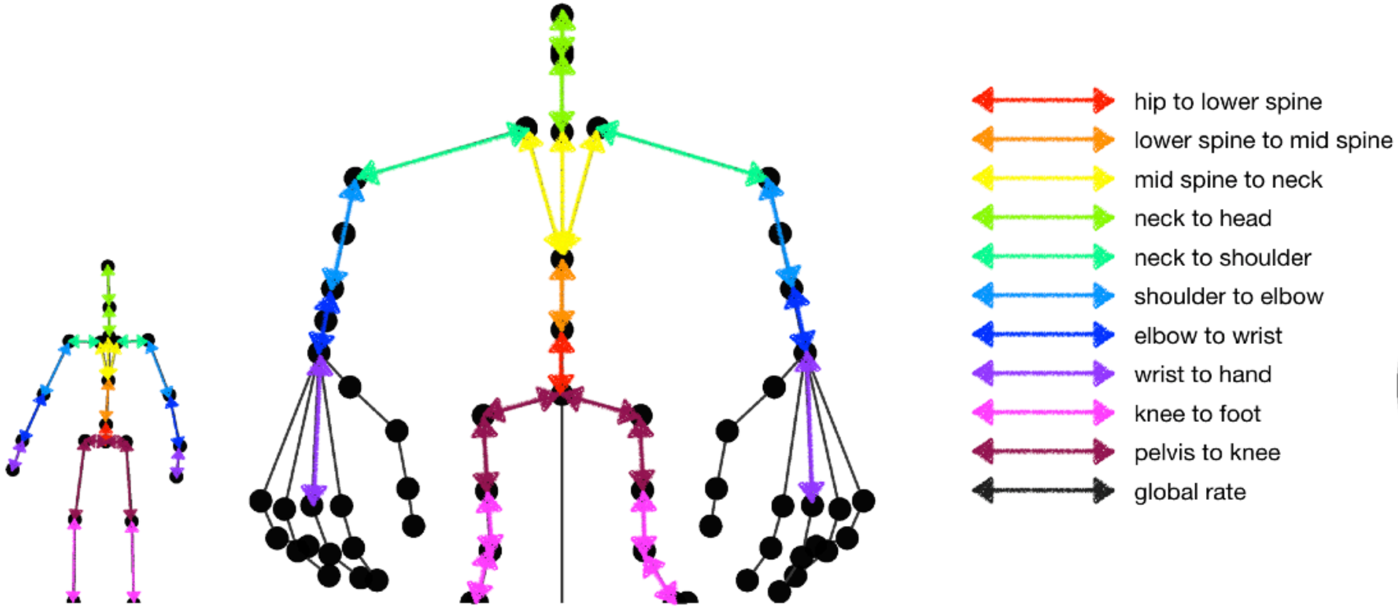


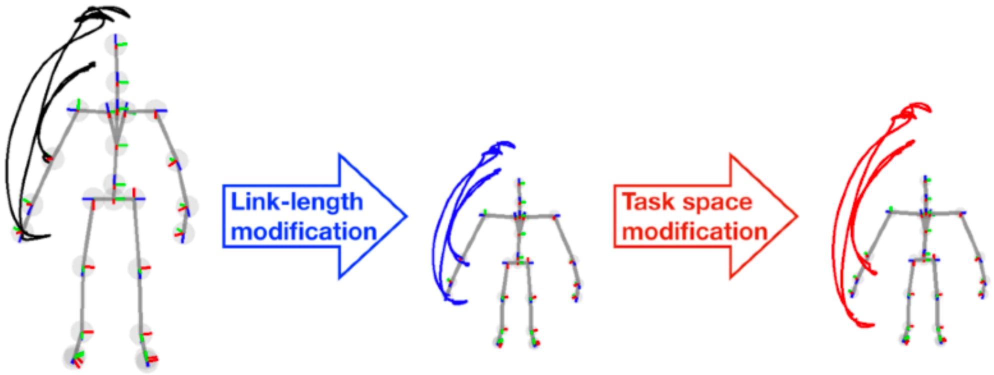
- We first define **joints of interest** of which to transfer from a source skeleton to target humanoid robots.

Motion Transfer Parametrization

ROBOT INTELLIGENCE LAB

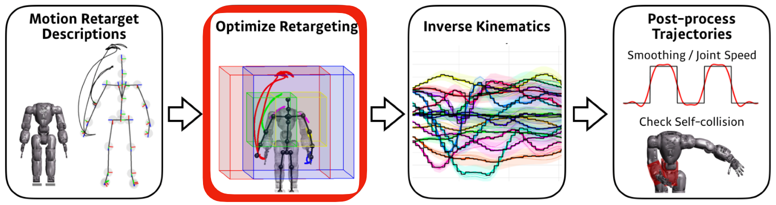


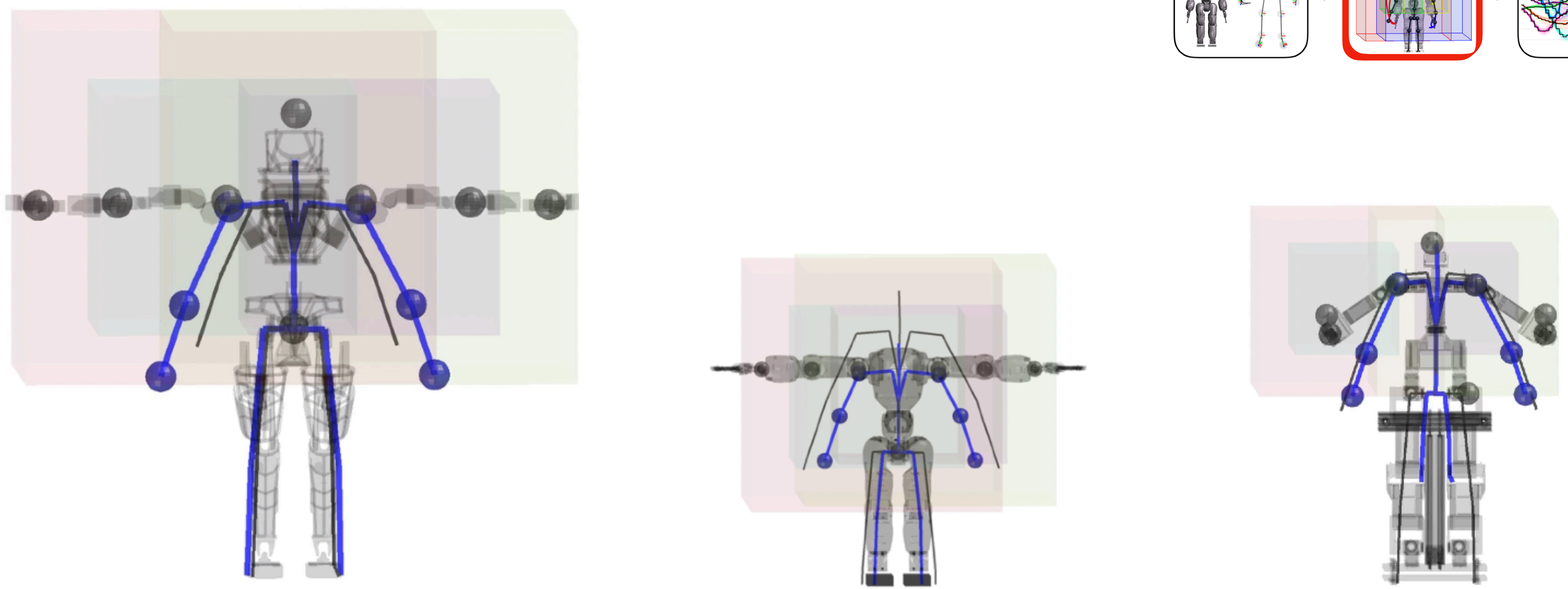




- The motion transfer is parametrized by link-length modification followed by task space modification.

Cost Function

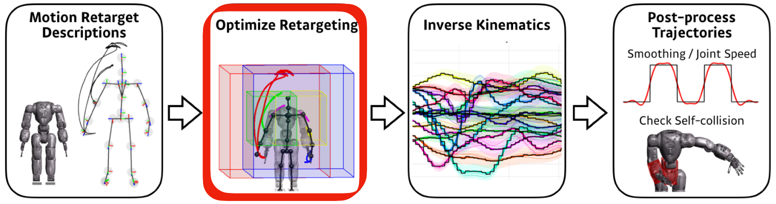




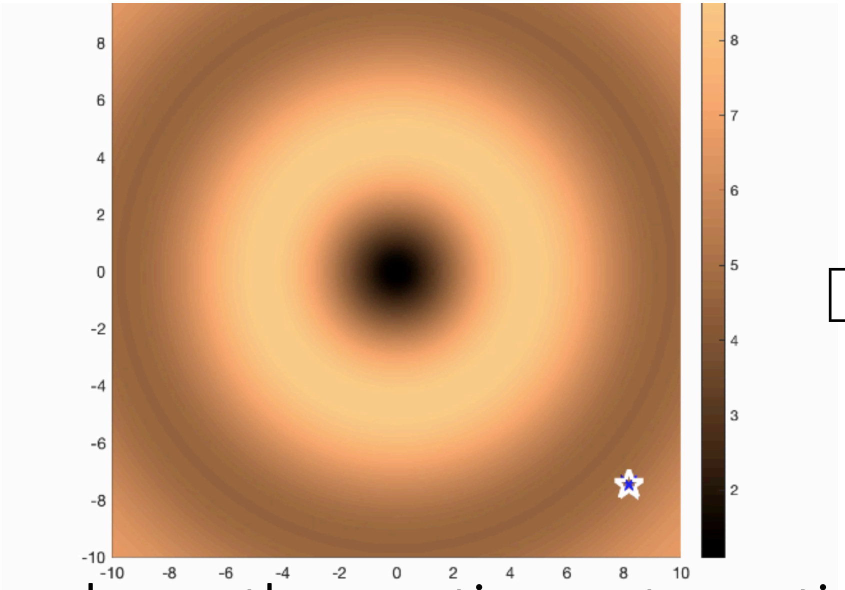
- The **cost function** is composes of four different sub-costs:
  1. Location of joints at the idle pose (A-pose).
  2. Link lengths between joints of interest.
  3. Workspaces of both hands.
  4. Relative trajectories between the original and adjusted source motions.

Black-box Optimization

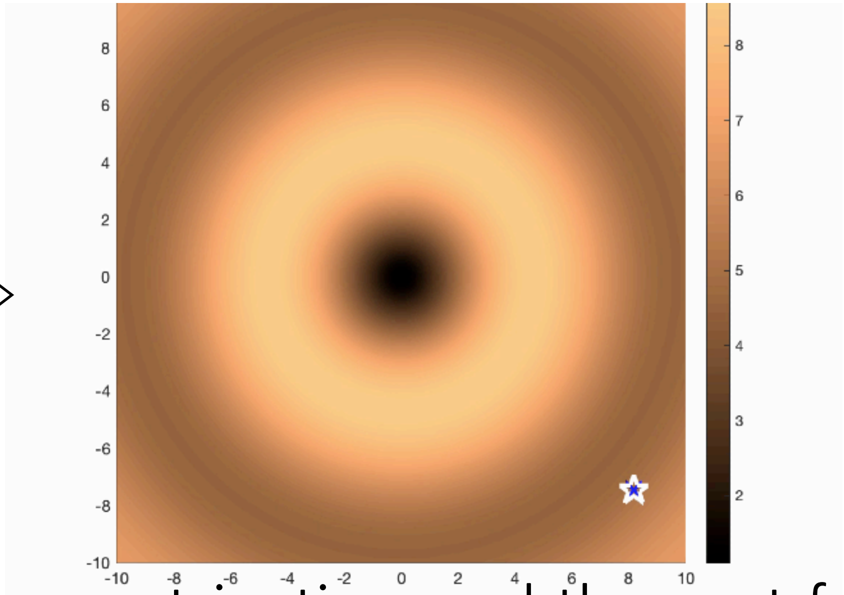
ROBOT INTELLIGENCE LAB



Original Bayesian Optimization

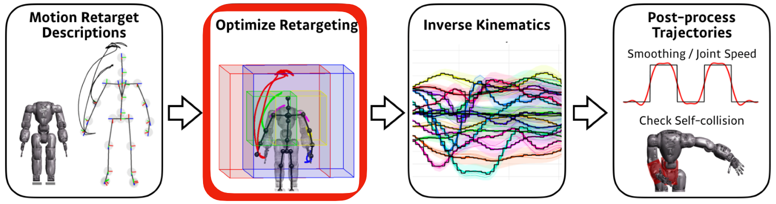


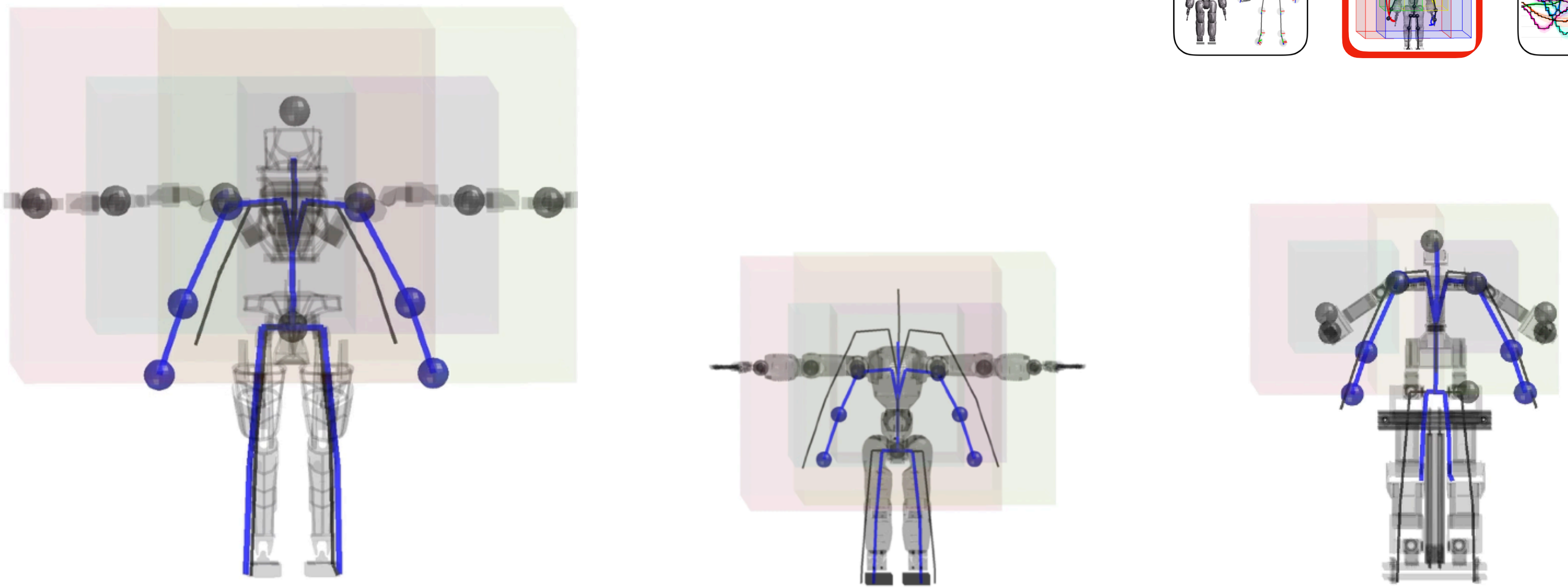
BO + Coordinate Descent



- Once we have the motion retargeting parametrization and the cost function to optimize, a gradient-free optimization method combining **Bayesian optimization with coordinate descent** is used for motion retargeting.

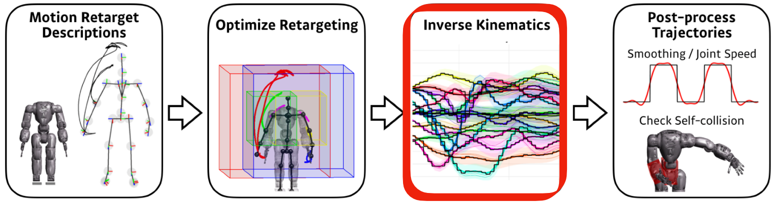
Inverse Kinematics

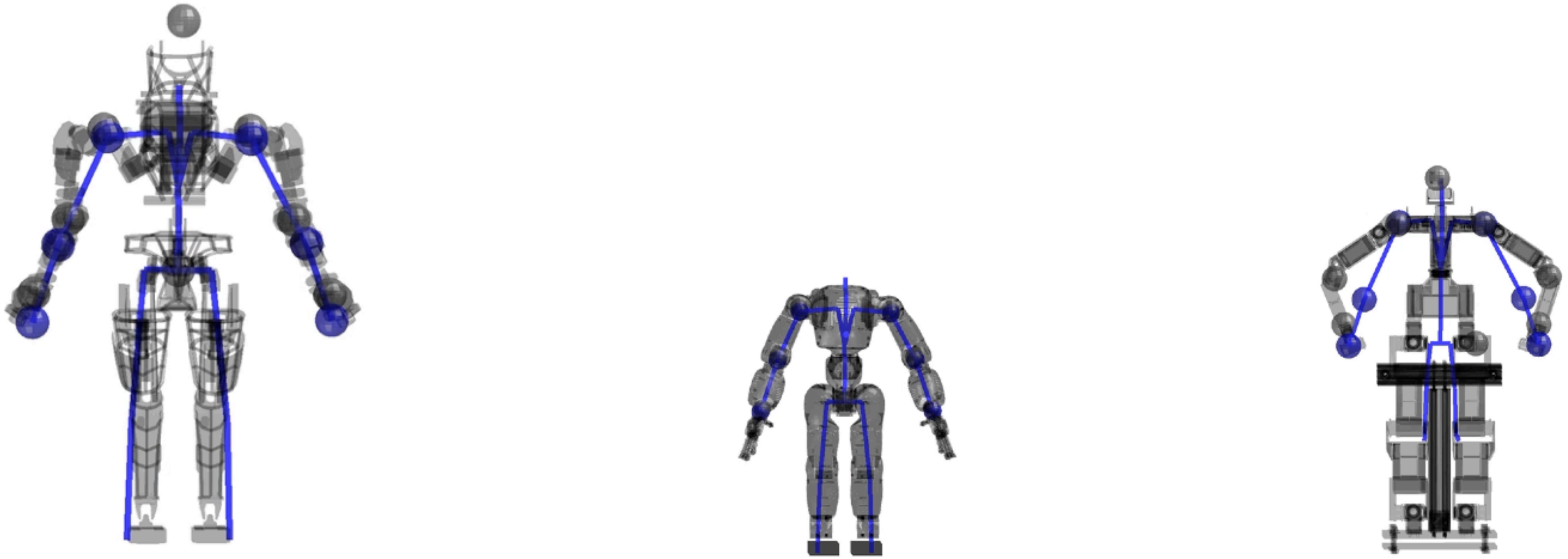




- The joint trajectories of target robots are computed by solving inverse kinematics with multiple target joints using an augmented Jacobian method.

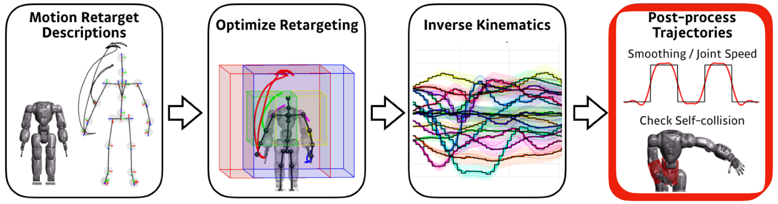
Inverse Kinematics

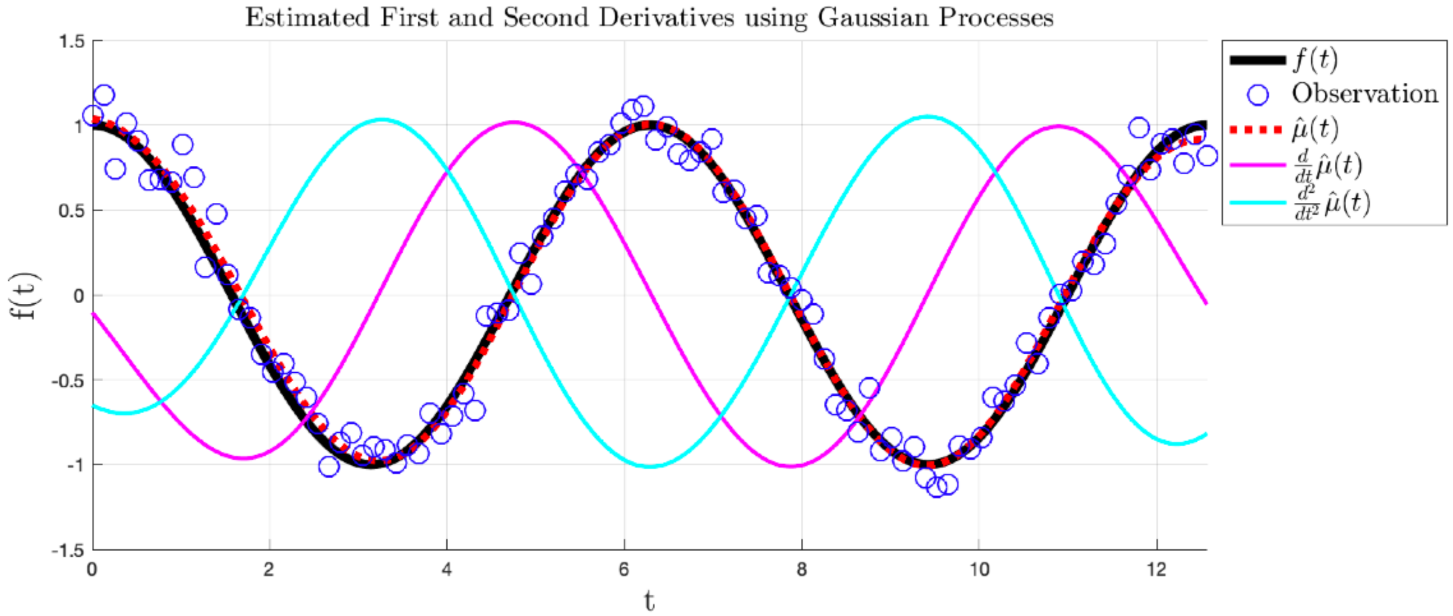




- The joint trajectories of target robots are computed by solving inverse kinematics with multiple target joints using an augmented Jacobian method.

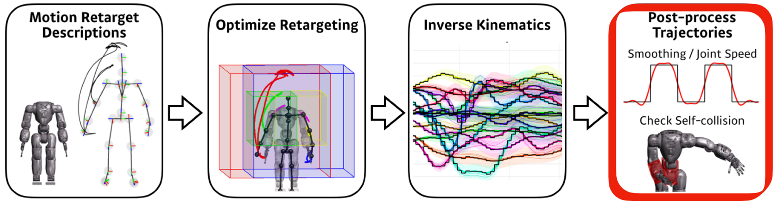
Post Processing



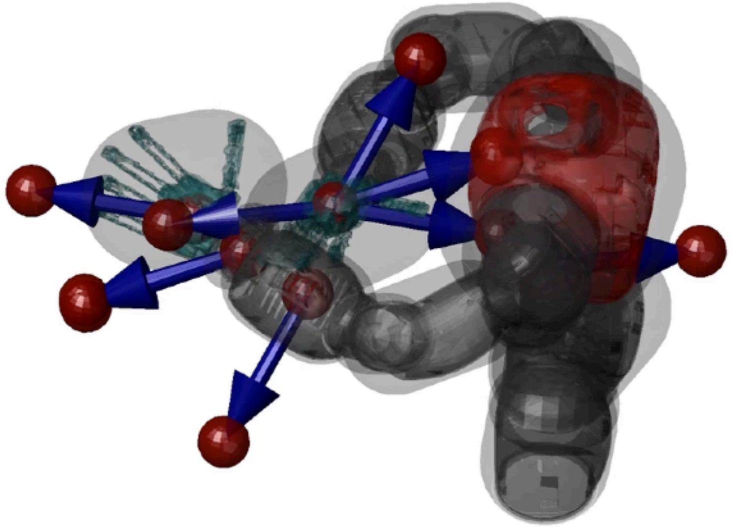


- Additional **trajectory smoothing** using Gaussian process regression as well as **self-collision handling** is processed.

Post Processing

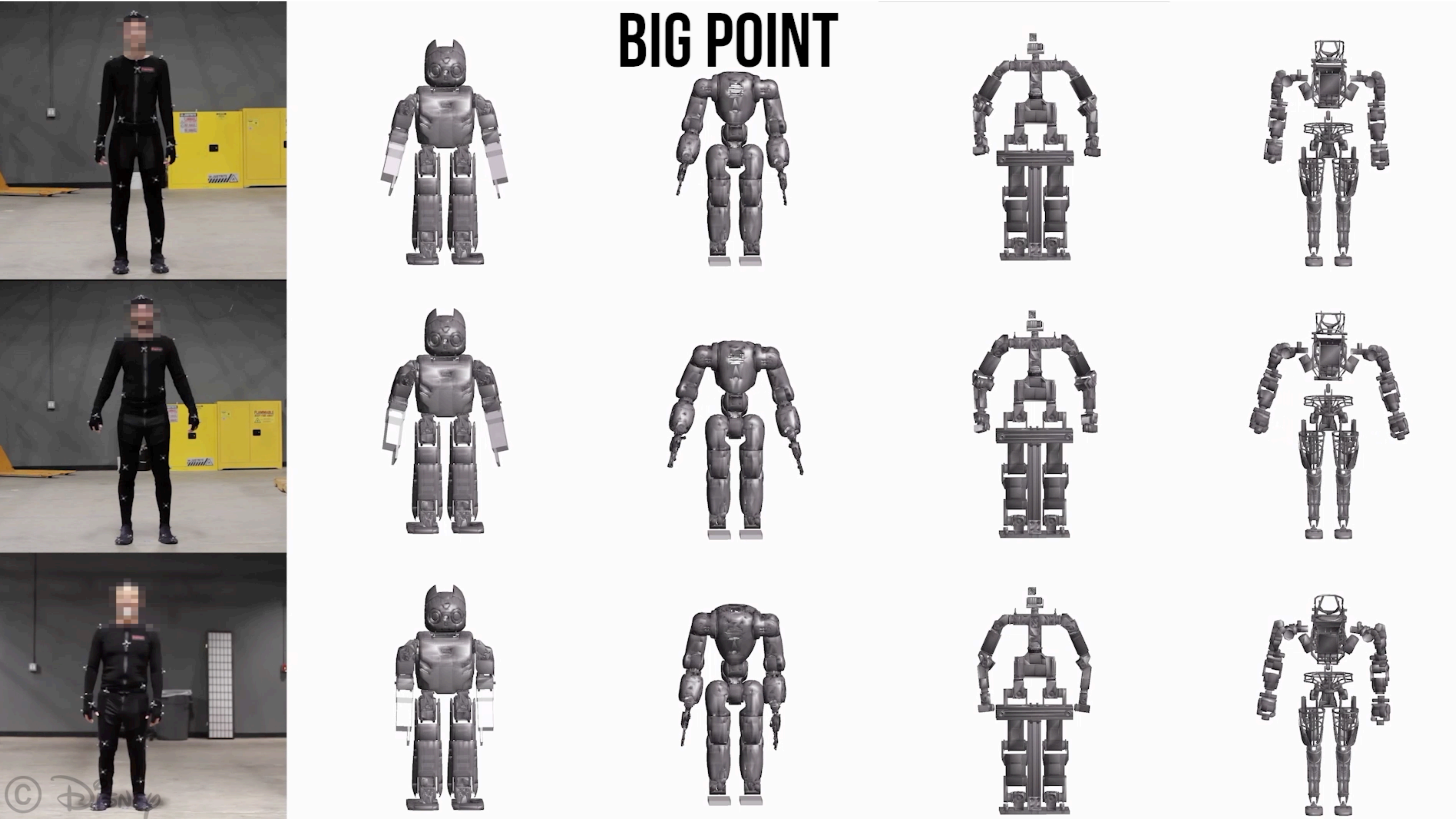


[01] Sum of Signed Distances:[-0.187]

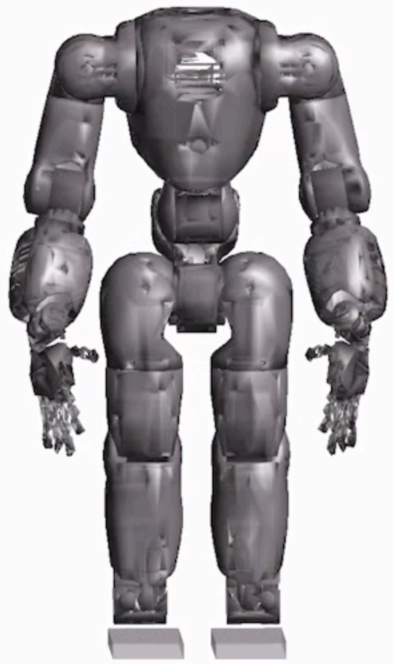


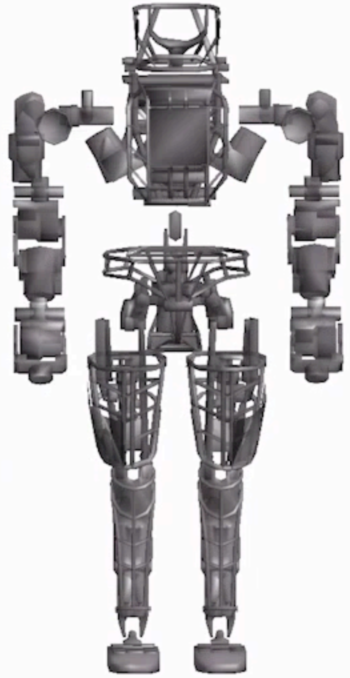
- Additional **trajectory smoothing** using Gaussian process regression as well as **self-collision handling** is processed.

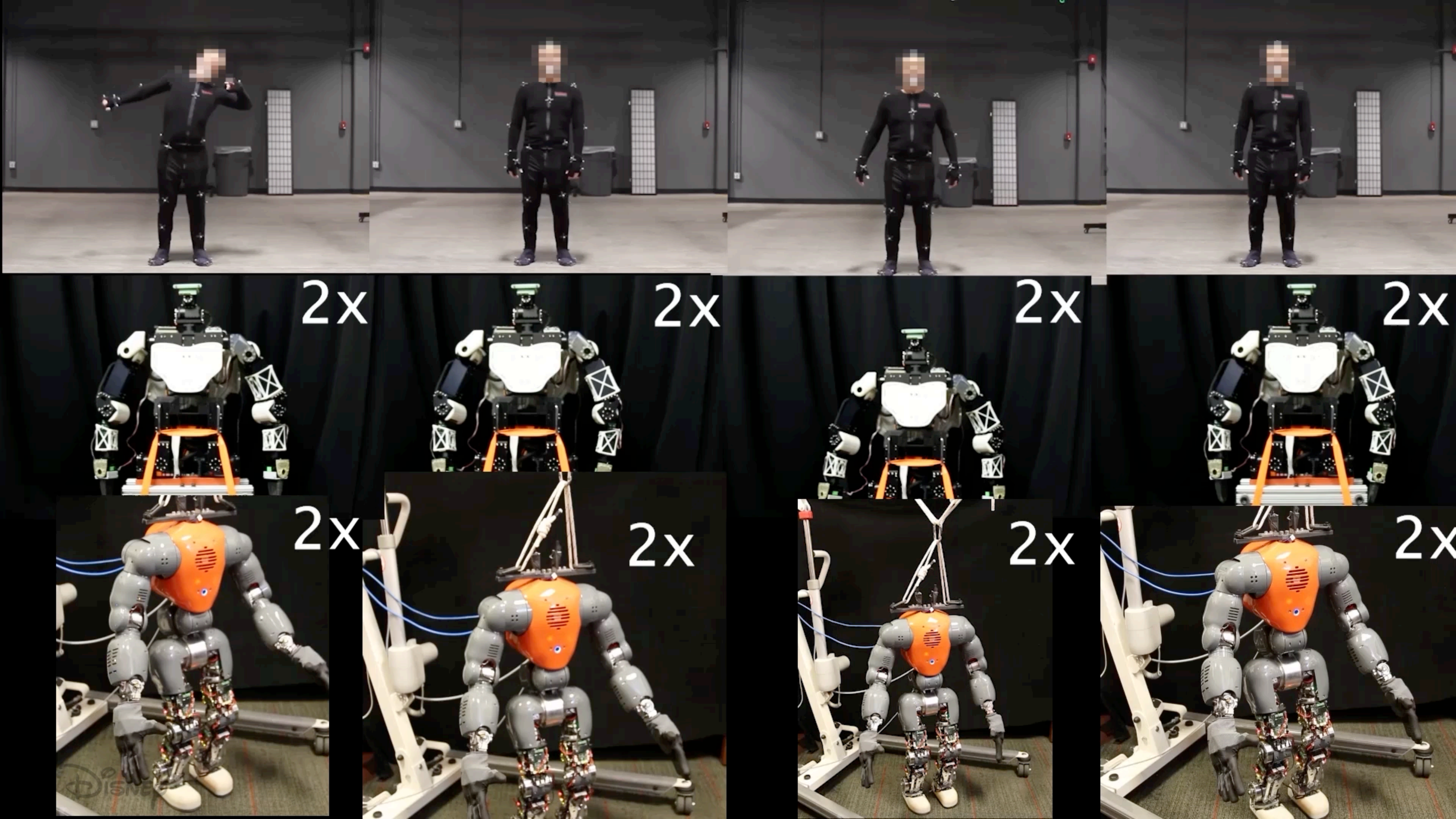
BIG POINT



GO ON






2x
2x
2x
2x
2x
2x
2x
2x

Cross-Domain Motion Transfer

"Cross-Domain Motion Transfer via Safety-Aware Shared Latent Space Modeling", RA-L 2020

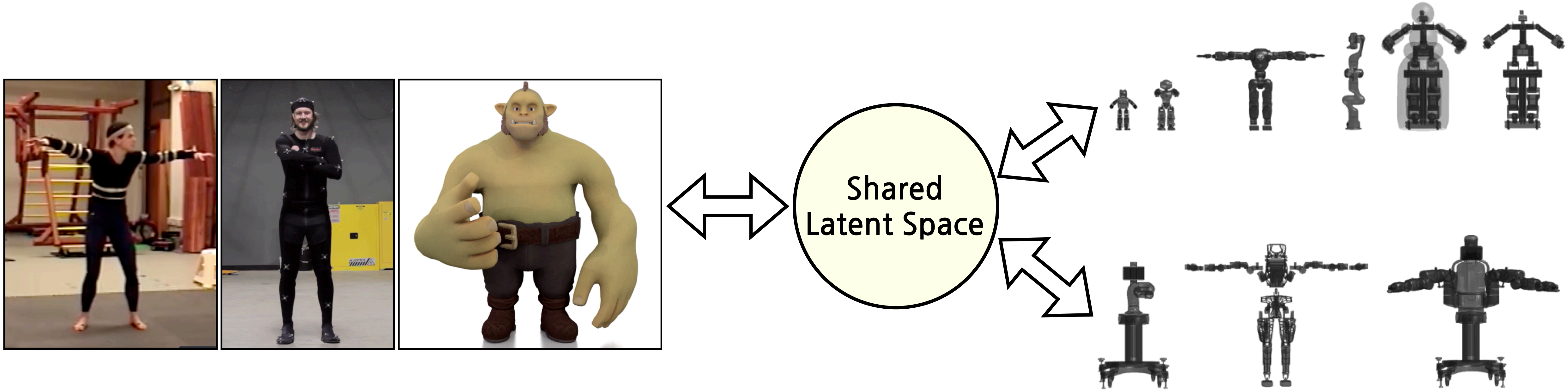
Key Question

ROBOT INTELLIGENCE LAB

How can we **speed-up** the motion retargeting while considering the **feasibility** of the motion?

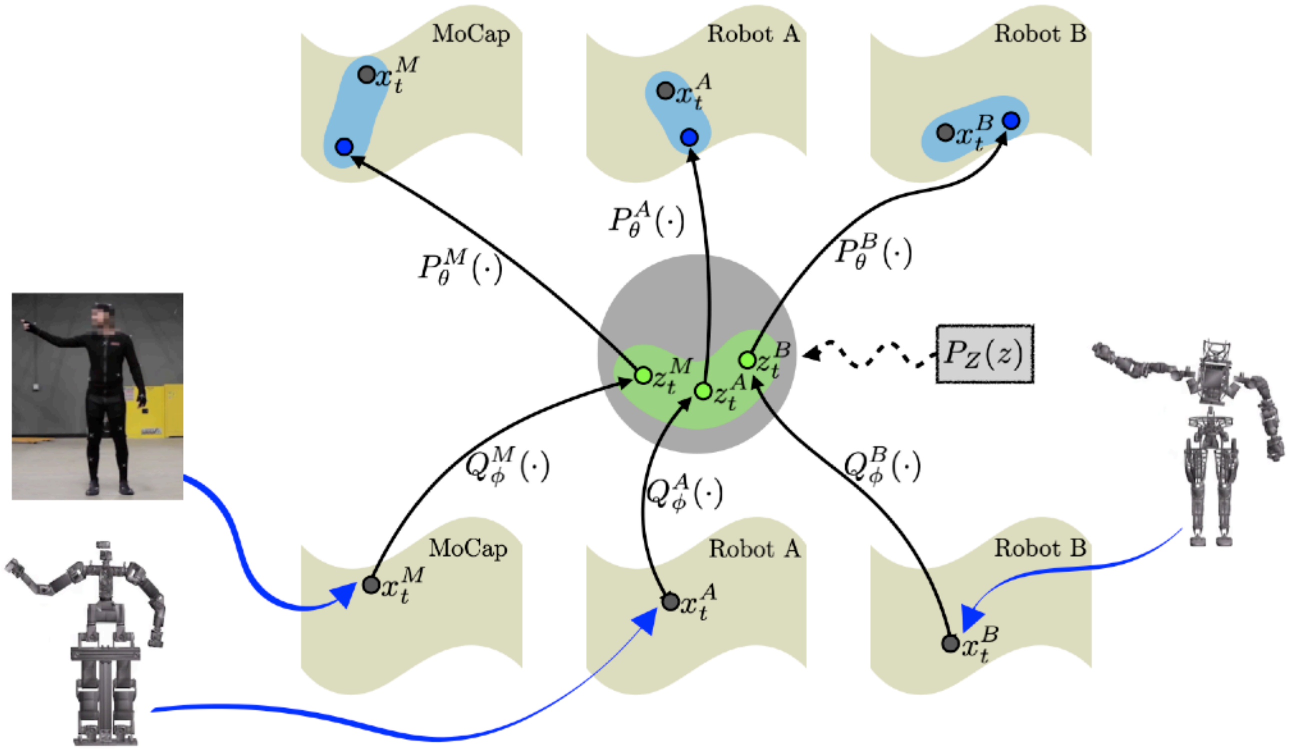
Latent Motion Retargeting

ROBOT INTELLIGENCE LAB



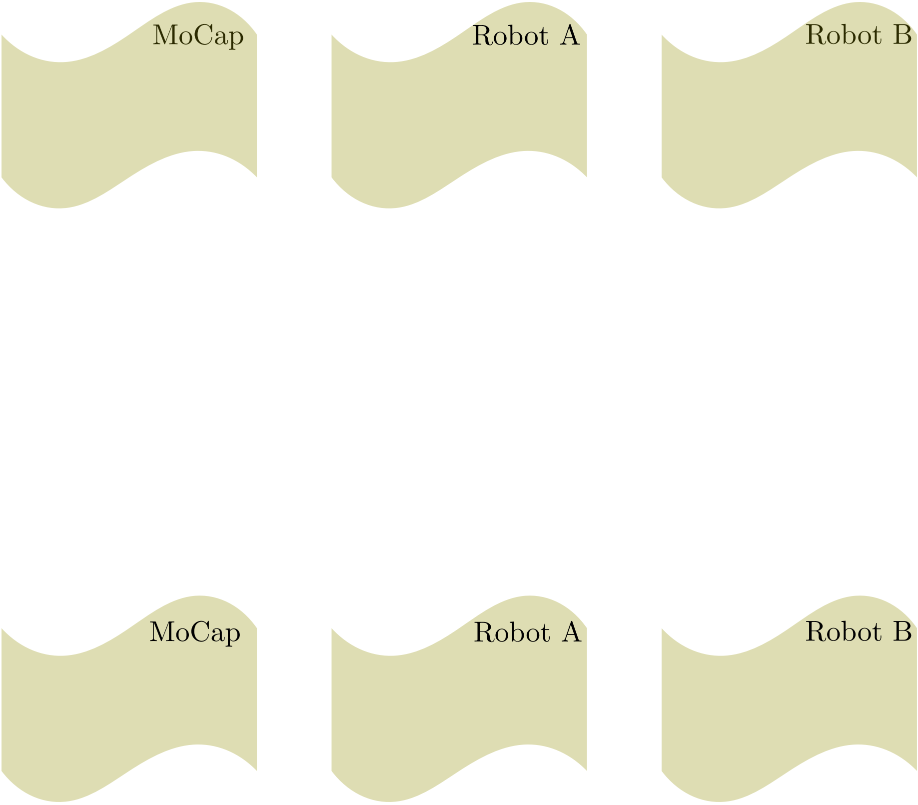
- One way of solving motion retargeting problems is to construct a **shared latent space** between different modalities (different mocap skeletons or different humanoid robot poses).

Shared Latent Space

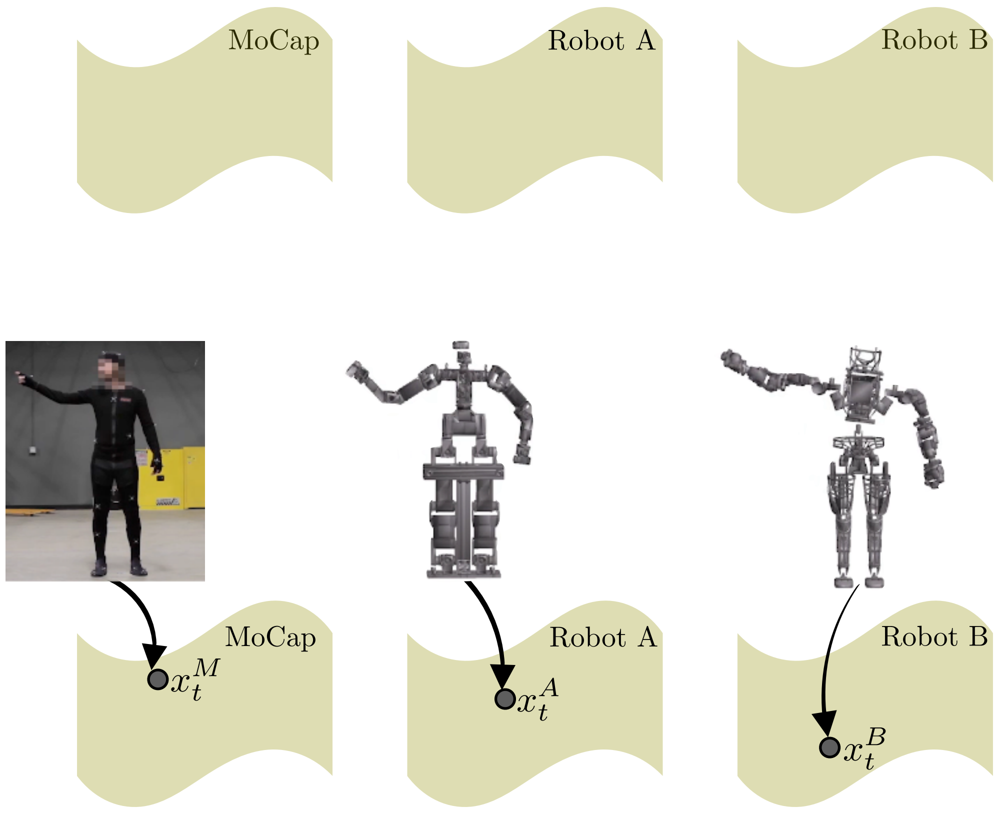


- However, an additional approach is needed to guarantee the feasibility of the transferred results (target humanoid pose).

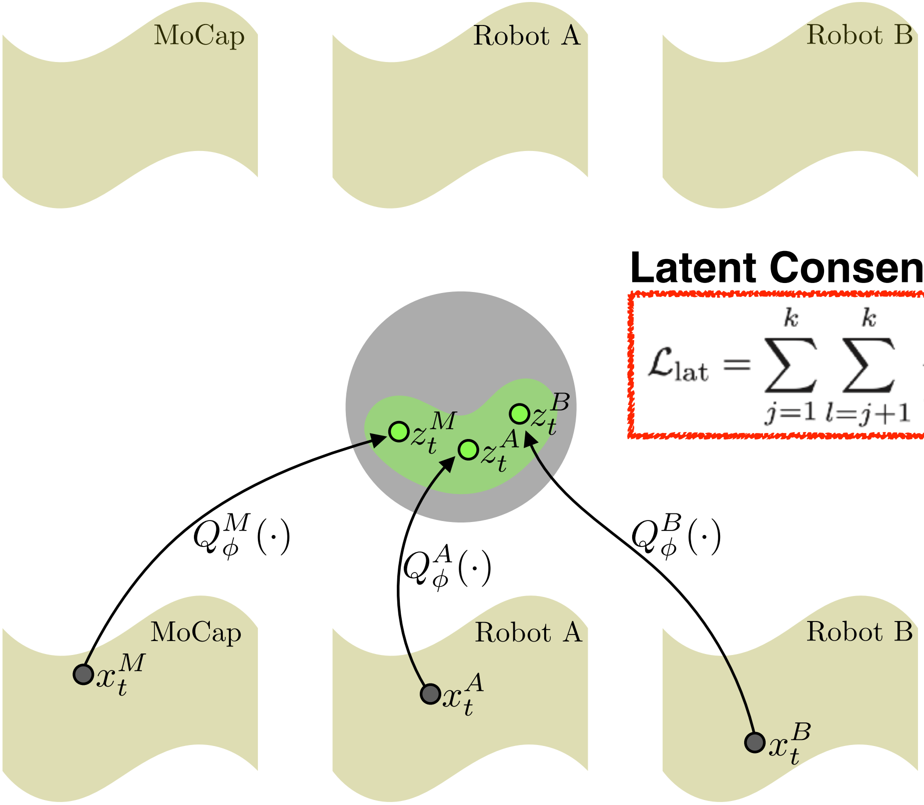
Learning the Shared Latent Space



Learning the Shared Latent Space



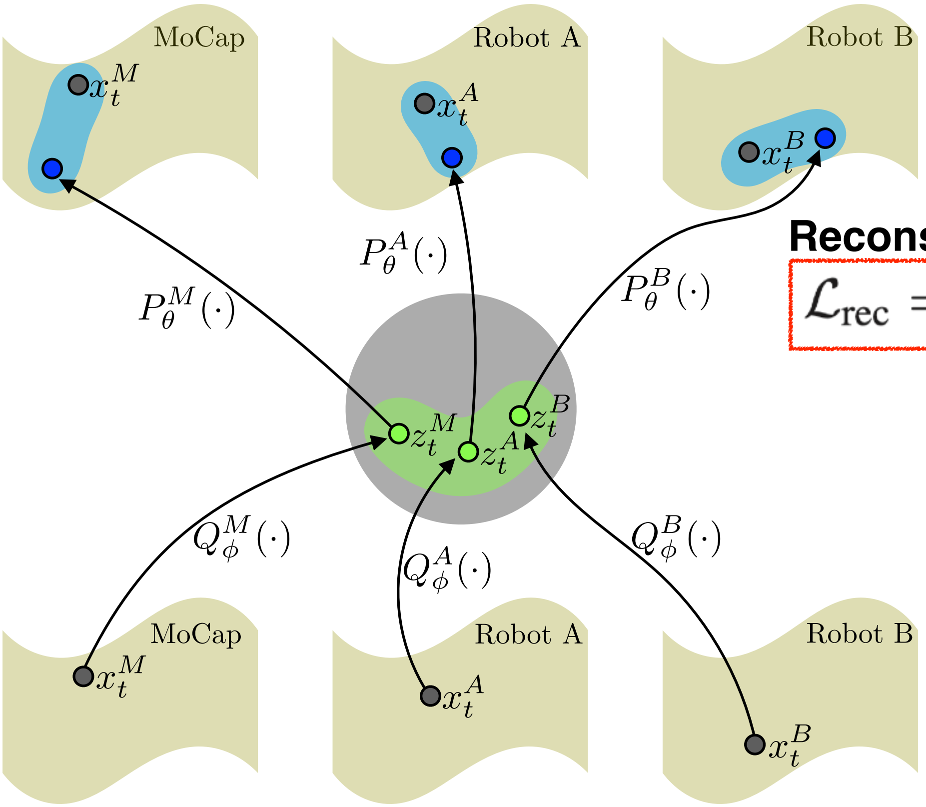
Learning the Shared Latent Space



Latent Consensus Loss

$$\mathcal{L}_{\text{lat}} = \sum_{j=1}^{k} \sum_{l=j+1}^{k} \frac{1}{N} \sum_{i=1}^{N} \| Q_\phi^j(\mathbf{x}_i^j) - Q_\phi^l(\mathbf{x}_i^l) \|_2^2$$

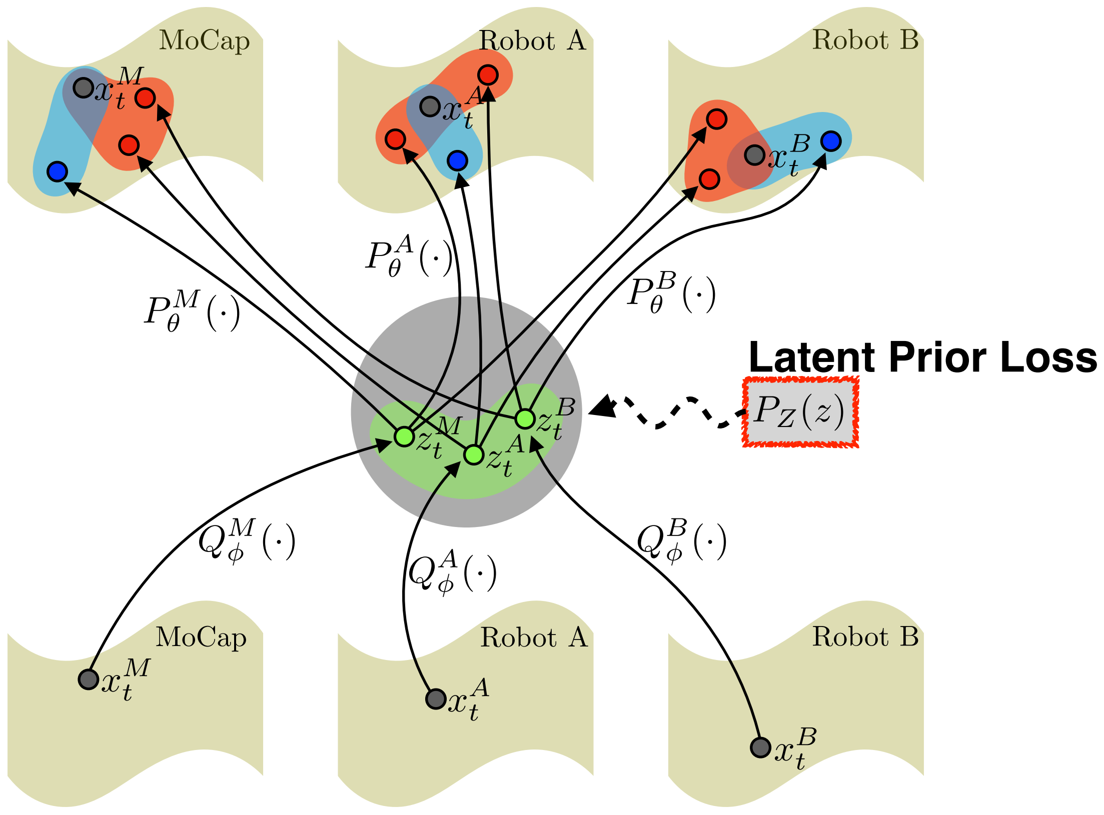
Learning the Shared Latent Space



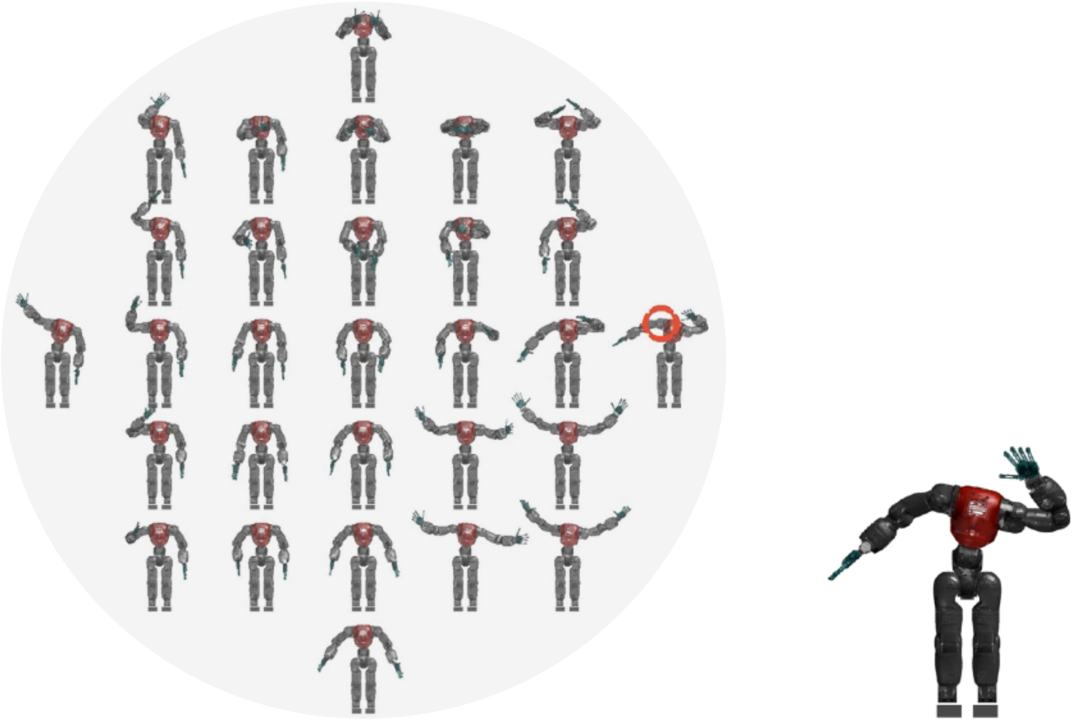
Reconstruction Loss

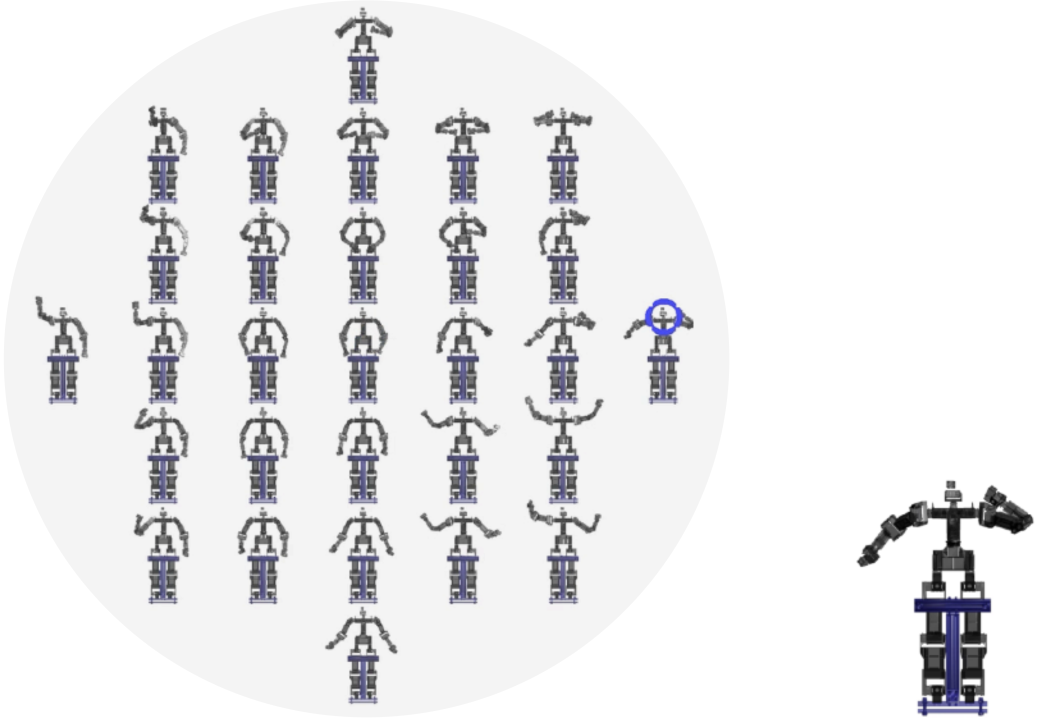
$$\mathcal{L}_{\text{rec}} = \mathbb{E}_{\mathbf{x}\sim p_{\mathbf{x}}}\mathbb{E}_{\mathbf{z}\sim Q_\phi(\mathbf{x})}\left[c(\mathbf{x}, P_\theta(\mathbf{z}))\right]$$

Learning the Shared Latent Space

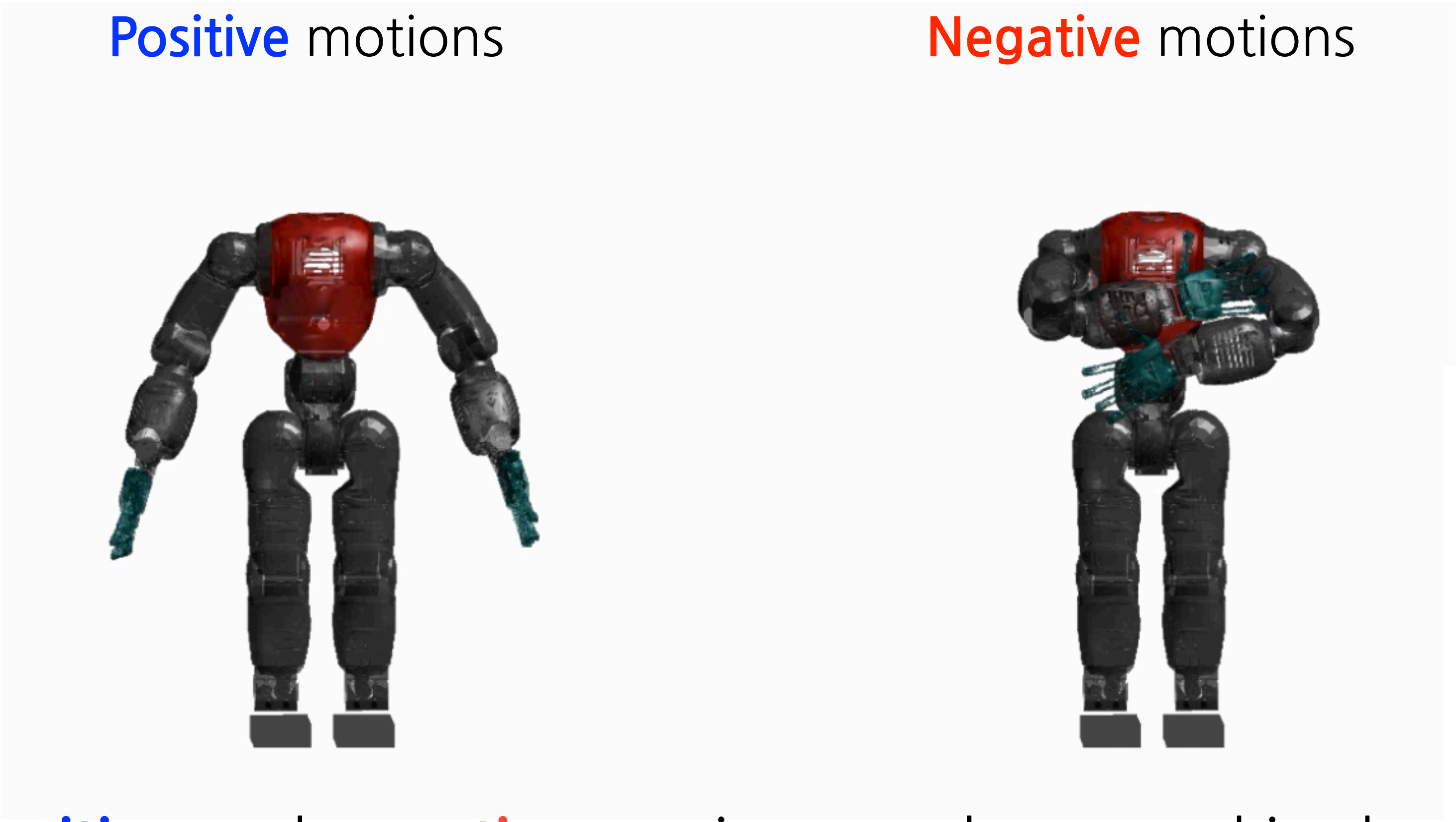


Constructed Shared Latent Space

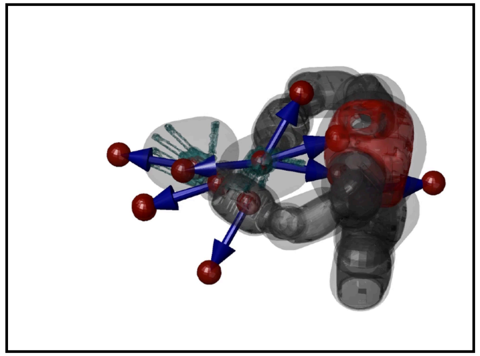




Safety-aware Learning



Self-collision handler



- Both positive and negative motions are leveraged in deep latent space modeling to achieve safety-aware motion retargeting.
- Negative data give explicit knowledge of what not to do.

Results

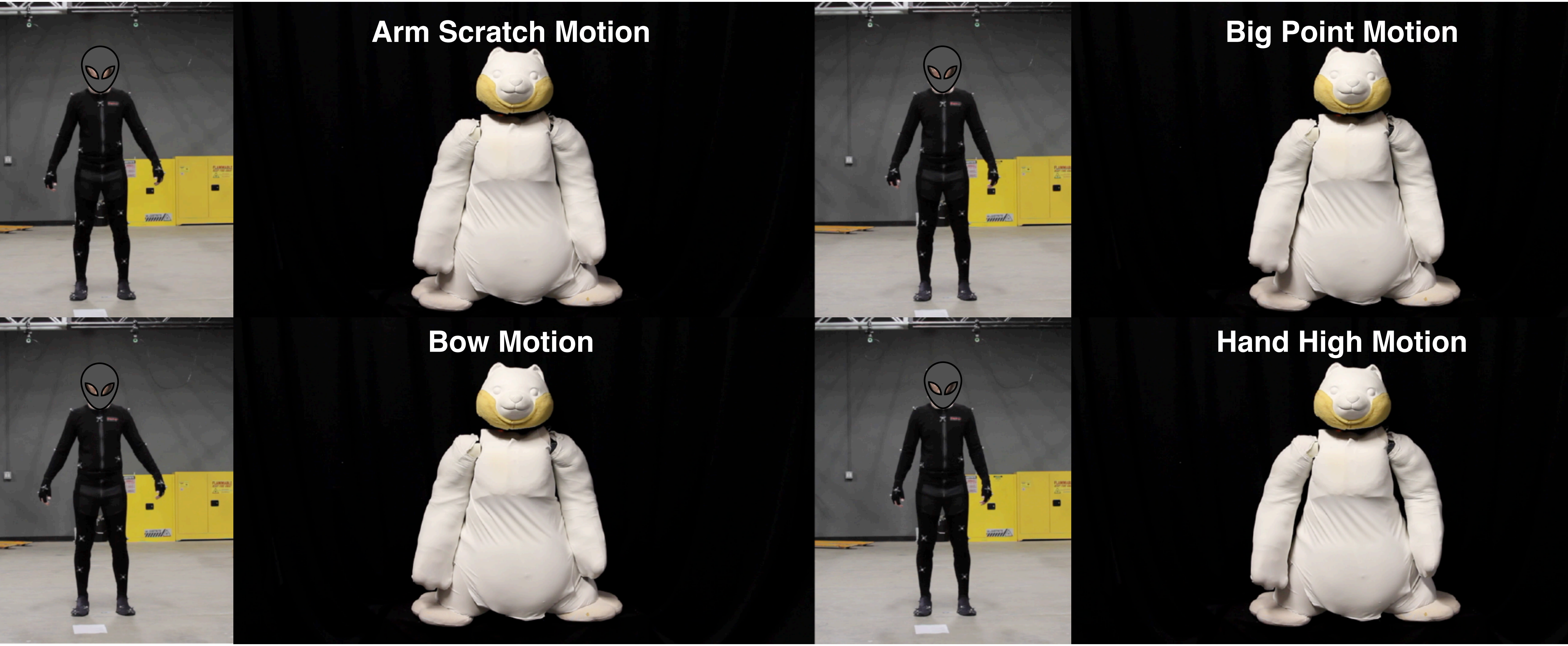
Baseline

Proposed Method

- By utilizing both positive and negative data, the proposed method greatly reduces the self-collision rates.

Results

ROBOT INTELLIGENCE LAB



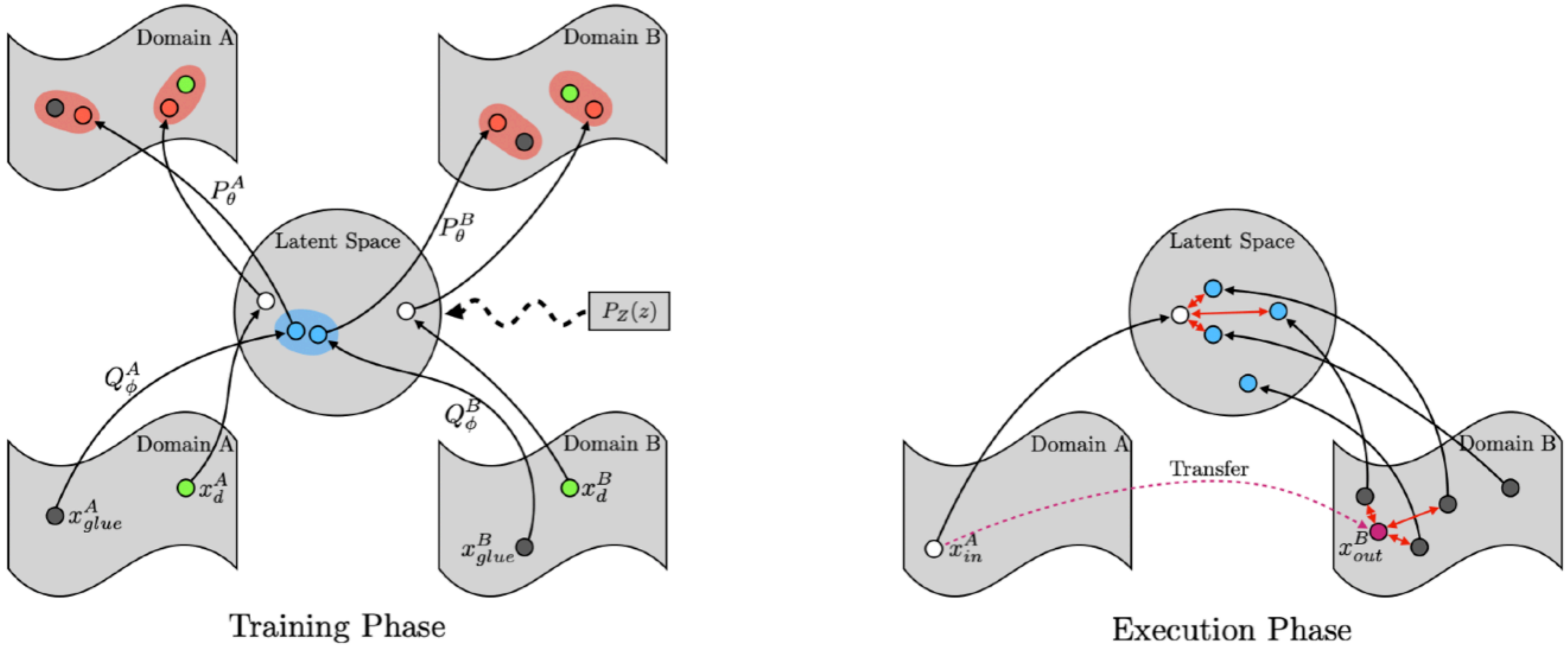
Nonparametric Motion Retargeting

"Nonparametric Motion Retargeting for Humanoid Robots on Shared Latent Space", RSS 2020

Key Question

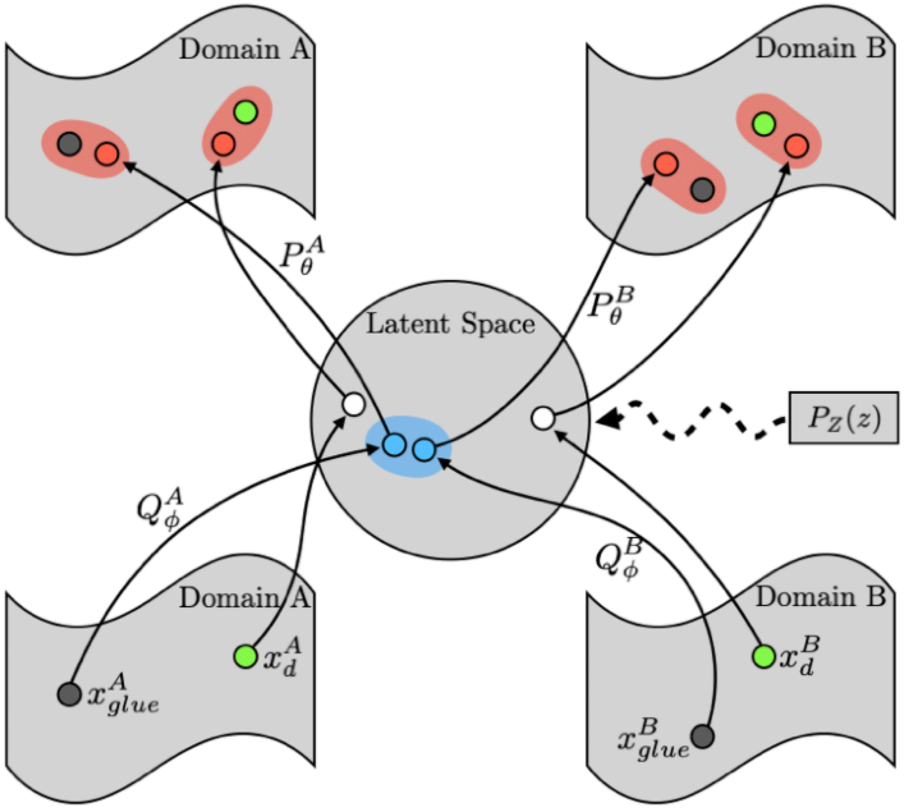
How can we always **guarantee** the feasibility of the retargeted motion?

Nonparametric Motion Retargeting



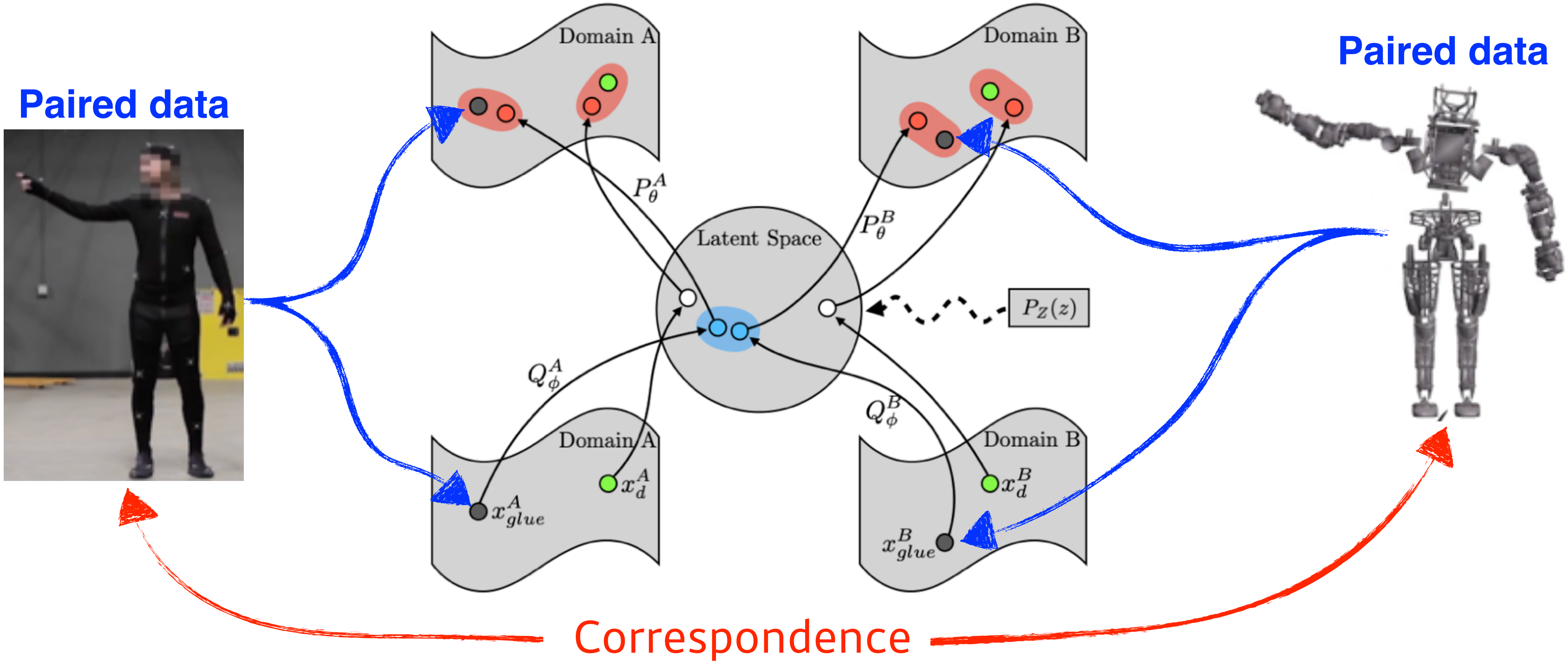
- In general, it is extremely difficult to **guarantee** the feasiblity of the output of deep neural networks.
- In this work, we combine a nonparametric regression method with shared latent space modeling to present **nonparametric motion retargeting**.

Training Phase



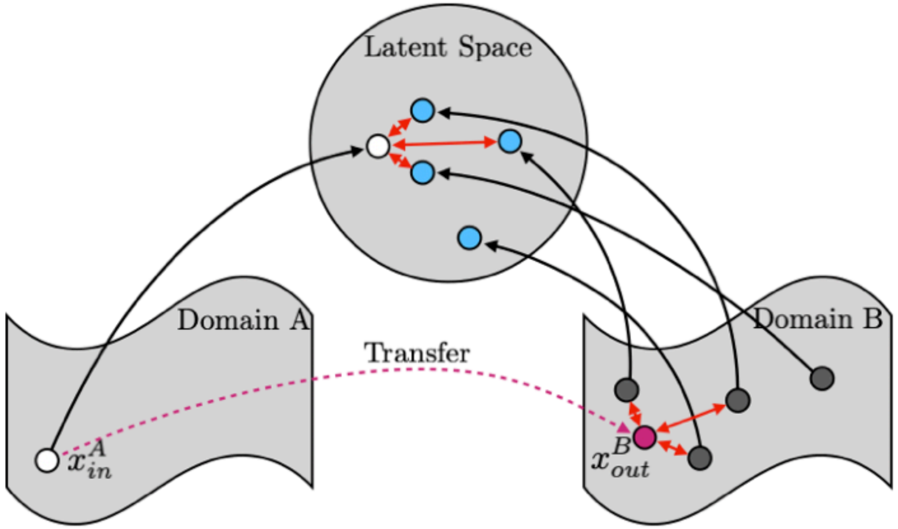
- **Training phase** is similar to our previous method.

Training Phase



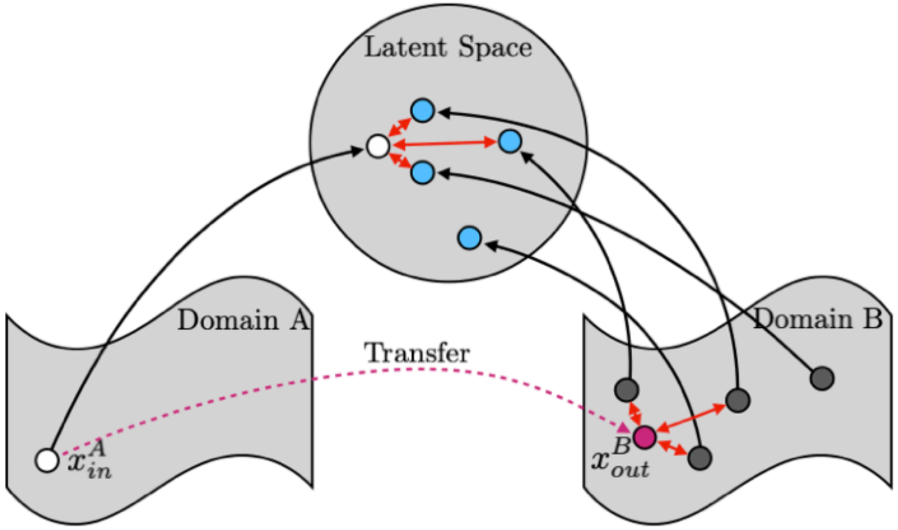
Execution Phase

ROBOT INTELLIGENCE LAB



- On the **Execution phase**, we utilized nonparametric regression (or the nearest neighbor algorithm) on the shared latent space.

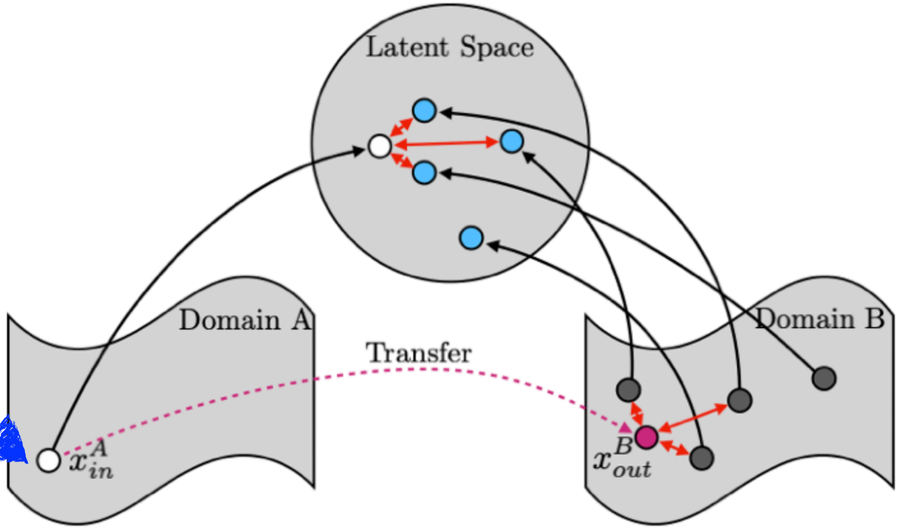
Execution Phase



- On the **Execution phase**, we utilized nonparametric regression (or the nearest neighbor algorithm) on the shared latent space.

Execution Phase

Given input

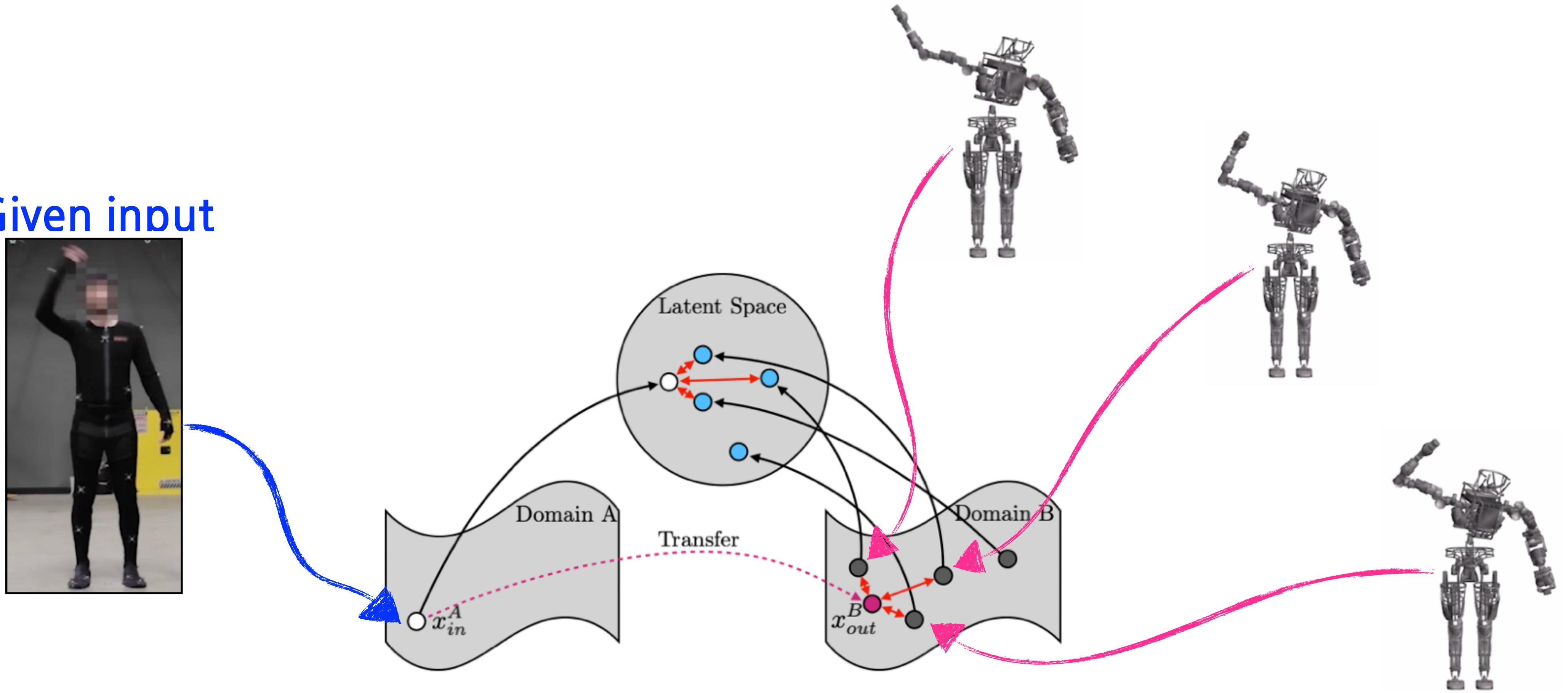


- Once a **mocap pose** is given as an input.

Execution Phase

ROBOT INTELLIGENCE LAB

Given input



- The corresponding robot pose is computed (or selected) by **locally weighed regression** on the shared latent space.

Latent Nonparametric Method

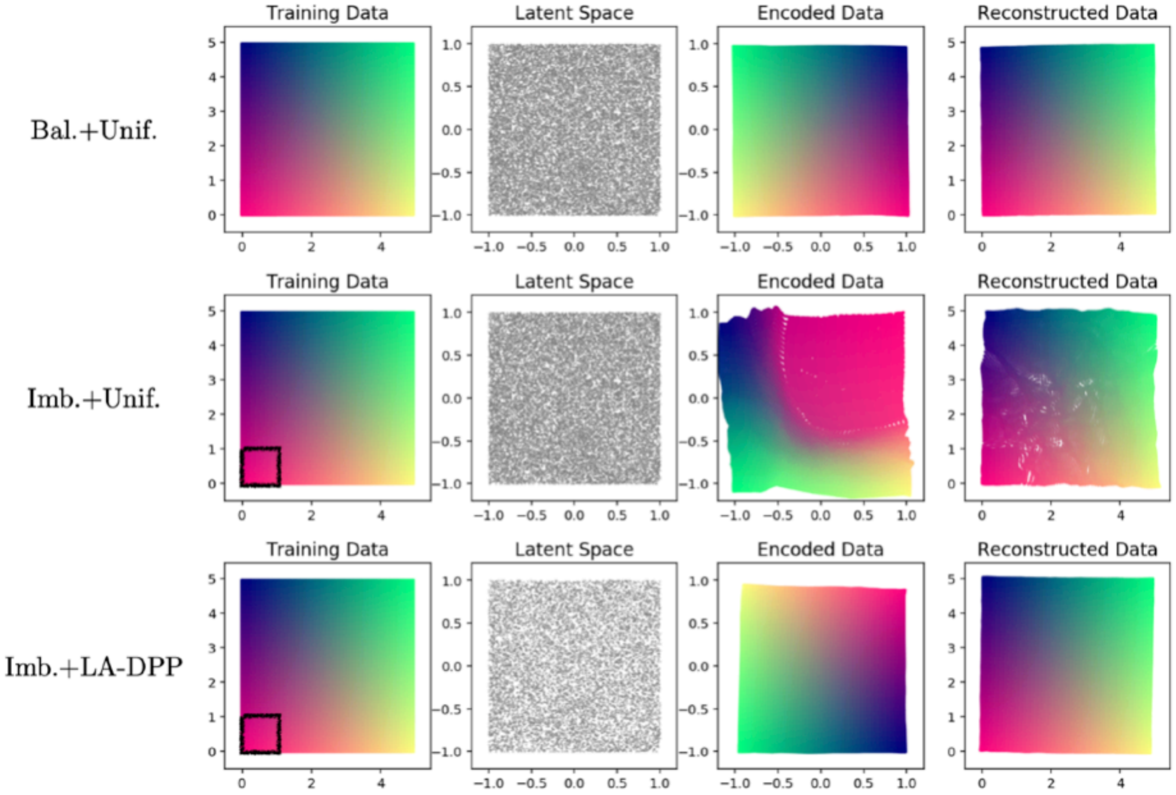
- What is the benefit of leveraging **shared latent space** with **nonparametric methods**?

	Feasibility Guarantee	Paired Data Requirements	Inference Time
Nonparametric Method (e.g., Lookup Table)	YES	YES	$O(N)$
Function approximators (e.g., Multi-layer Perceptron)	NO	YES	$O(1)$
Shared Latent Space (e.g., Cross-Domain Motion Transfer)	NO	NO	$O(1)$
Proposed Method (e.g., Latent Nonparametric Method)	YES	NO	$O(N)$

Efficient Sub-Sampling

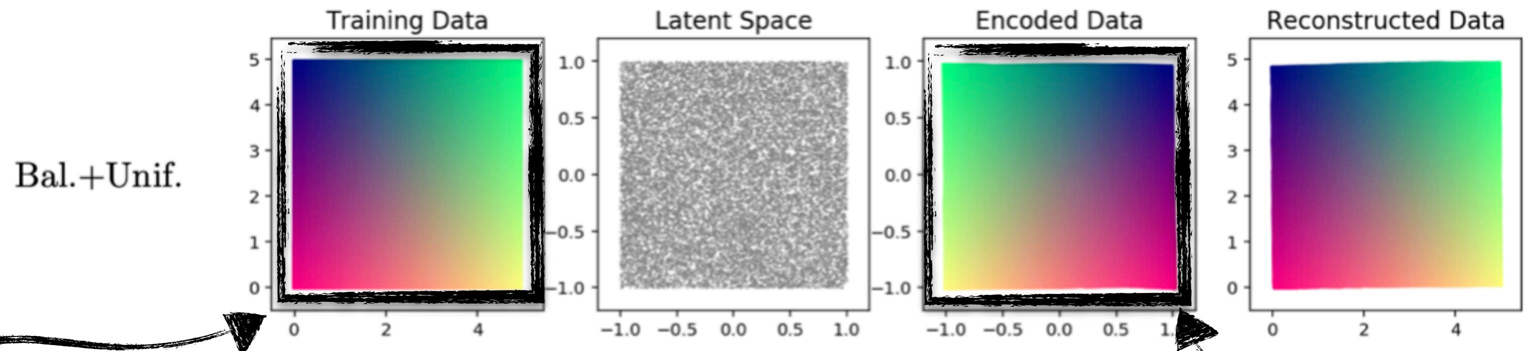
Original DPP

$$
\begin{aligned}
\mathbf{x}_* &= \boxed{\arg\max_{\mathbf{x}\in\mathcal{X}_{\setminus m}} (1-\mathbf{k}^T K^{-1}\mathbf{k}) \det(K)} \\
&\approx \arg\max_{\mathbf{x}\in\mathcal{X}_{\setminus m}} (1-\mathbf{k}^T\mathbf{k}) \\
&= \arg\min_{\mathbf{x}\in\mathcal{X}_{\setminus m}} \|\mathbf{k}\|_2^2 \\
&= \arg\min_{\mathbf{x}\in\mathcal{X}_{\setminus m}} \|\mathbf{k}\|_1 \\
&= \arg\min_{\mathbf{x}\in\mathcal{X}_{\setminus m}} \sum_{i=1}^{m} k(\mathbf{x},\mathbf{x}_i).
\end{aligned}
$$



- An efficient subset sampling method is presented to handle the possible imbalance in the training dataset while constructing the latent space.

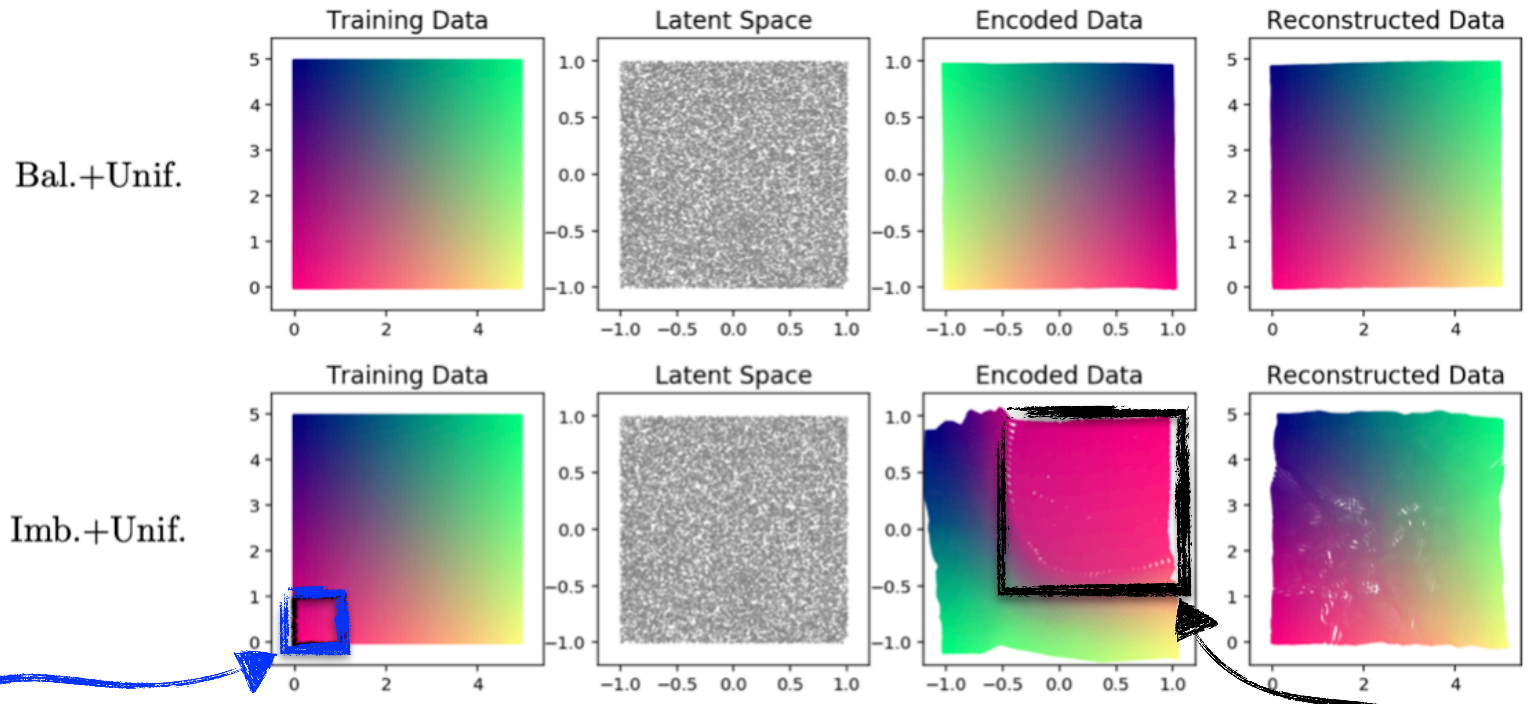
Efficient Sub-Sampling



If the training data are uniformly distributed within the domain (i.e., a **balanced dataset**)

Then, the resulting latent space is also **uniformly covered** with the training data.

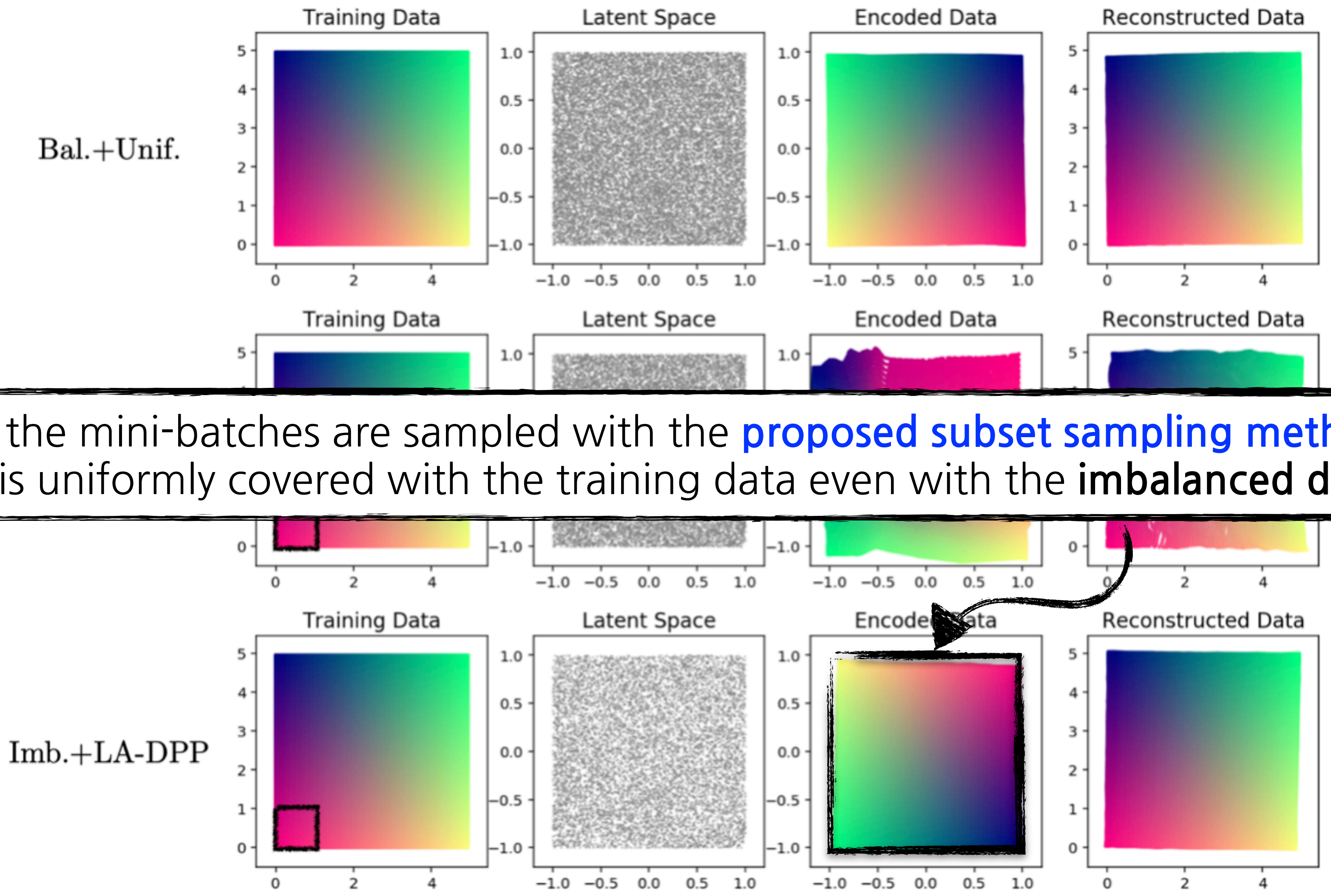
Efficient Sub-Sampling



If the half of training data are located within the small blue box (i.e., an imbalanced dataset)

Then, about the 50% of the latent space is covered with the data from the small imbalanced region in the dataset.

Efficient Sub-Sampling



When the mini-batches are sampled with the proposed subset sampling method, the latent space is uniformly covered with the training data even with the **imbalanced dataset**.

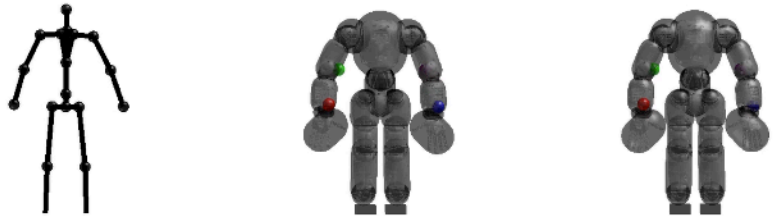
Results

Red color indicates
self-collision is occurred.

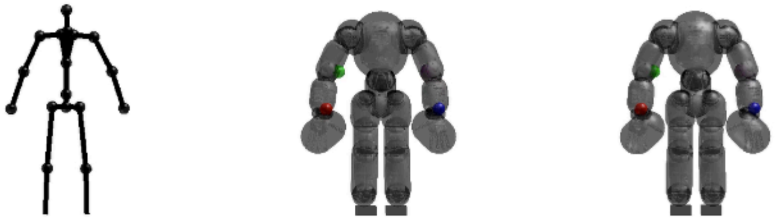
ArmCross

MoCap Proposed Method Baseline

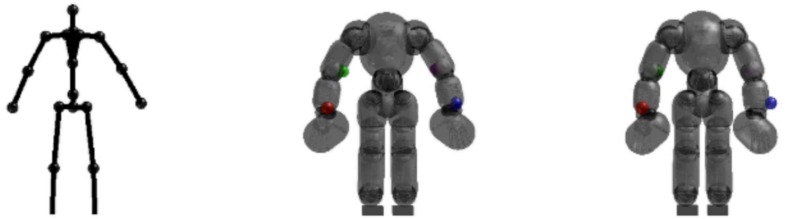
BigPoint



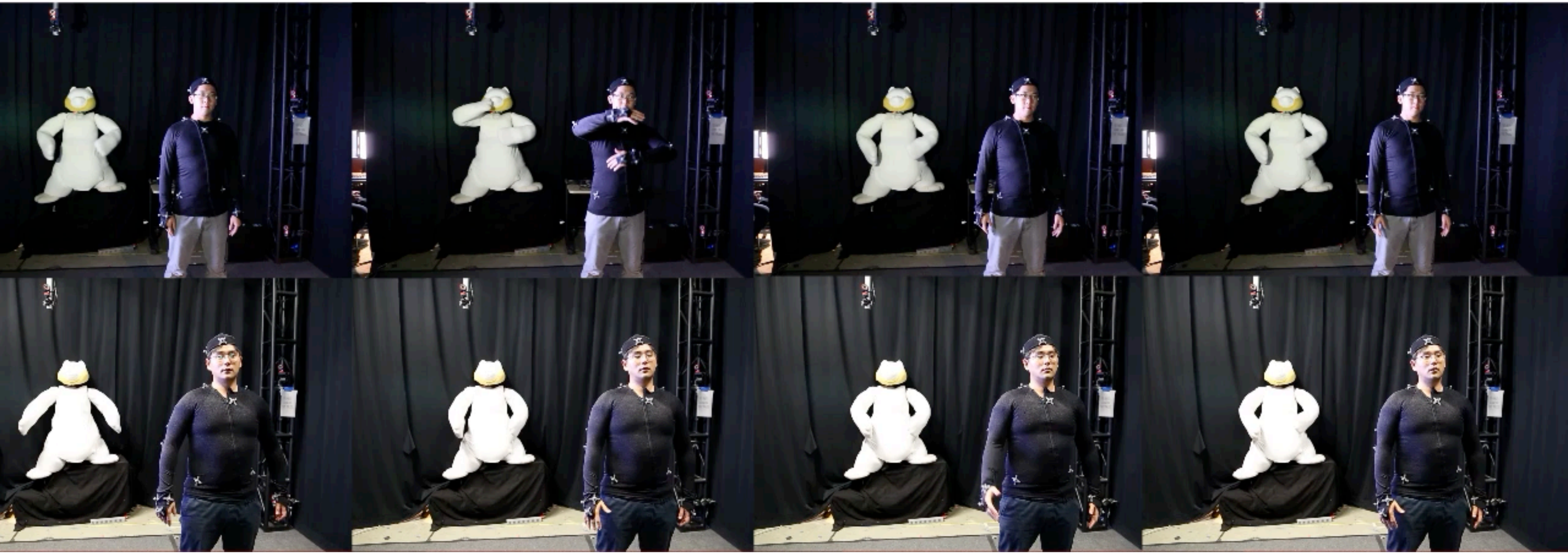
BigWave



BowBeg



Results



- There is no additional self-collision handling routine besides our proposed nonparametric motion retargeting method.

Self-Supervised Motion Retargeting

"Self-Supervised Motion Retargeting with Safety Guarantee", ICRA 2021

Key Questions

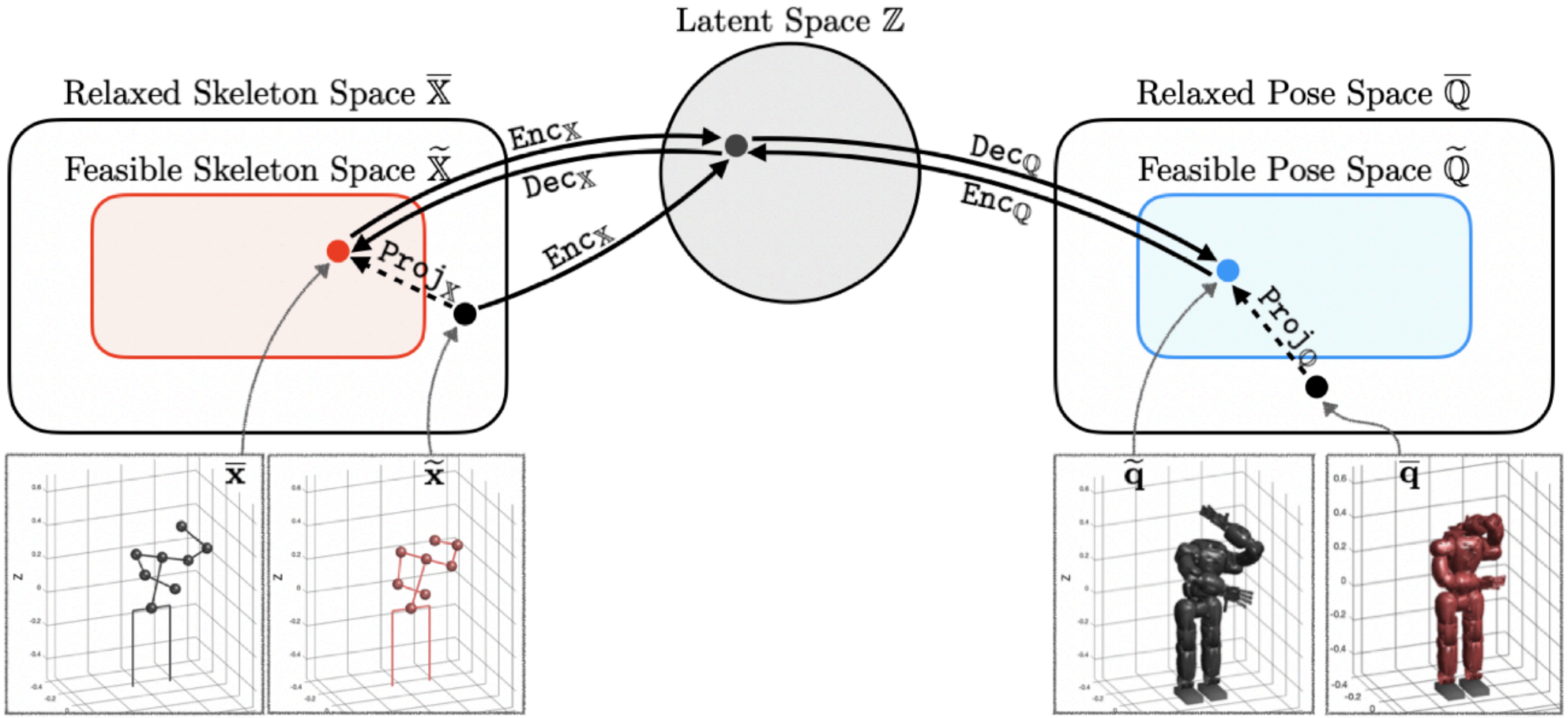
ROBOT INTELLIGENCE LAB

How can we consider **different expressivities** of human skeletons and humanoid robots.

+

Can we train a learning-based motion retargeting method in a **self-supervised manner**?

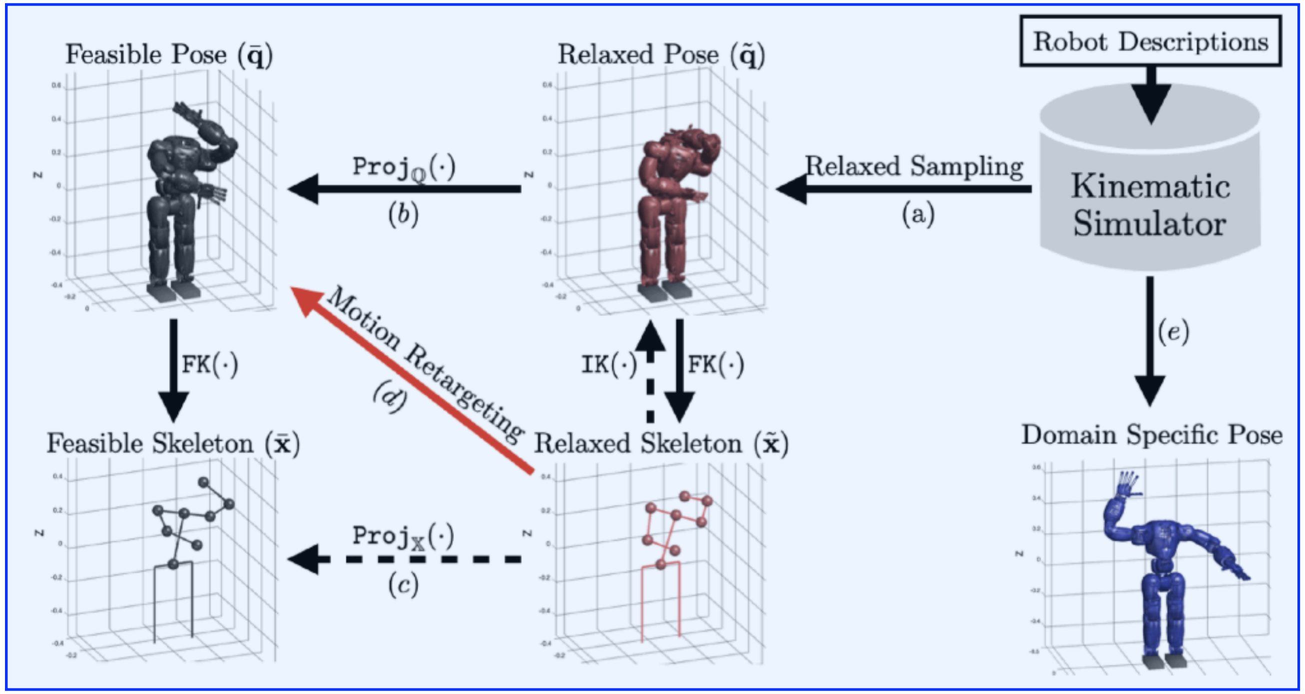
Projection-invariant Mapping

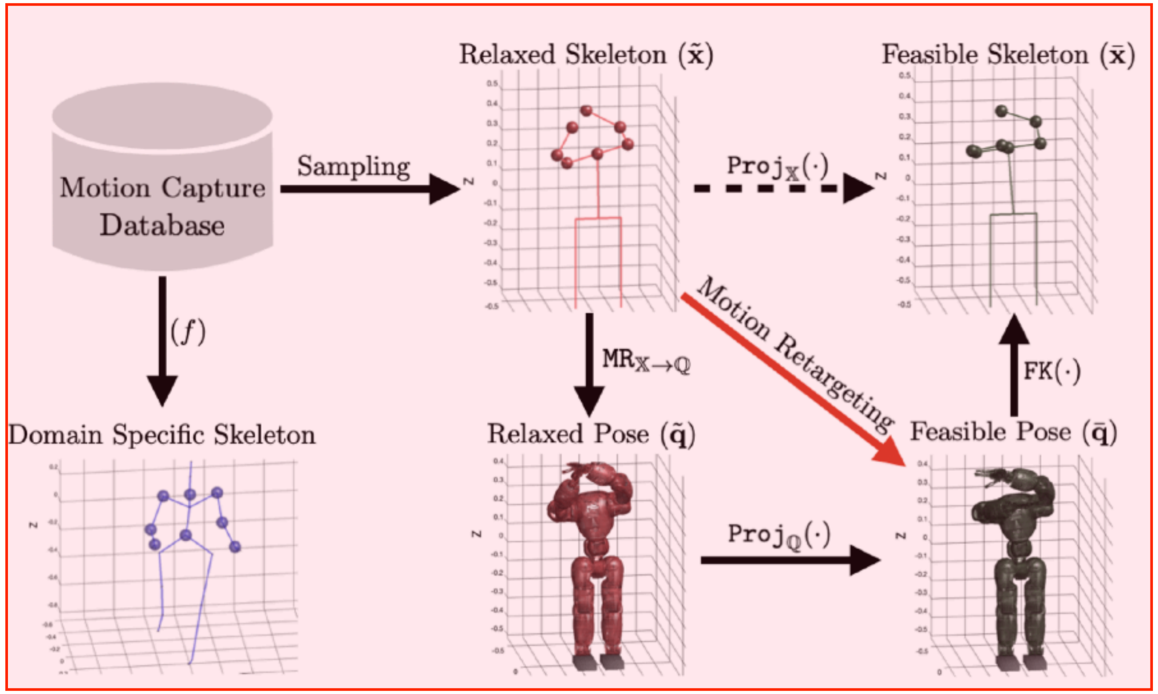


- Due to the limited expressivity of a humanoid robot, **multiple robot poses** may correspond to a **single human pose**. To this end, we propose a self-supervised motion retargeting by constructing a shared latent space between different domains.

Automated Sampling Pipelines

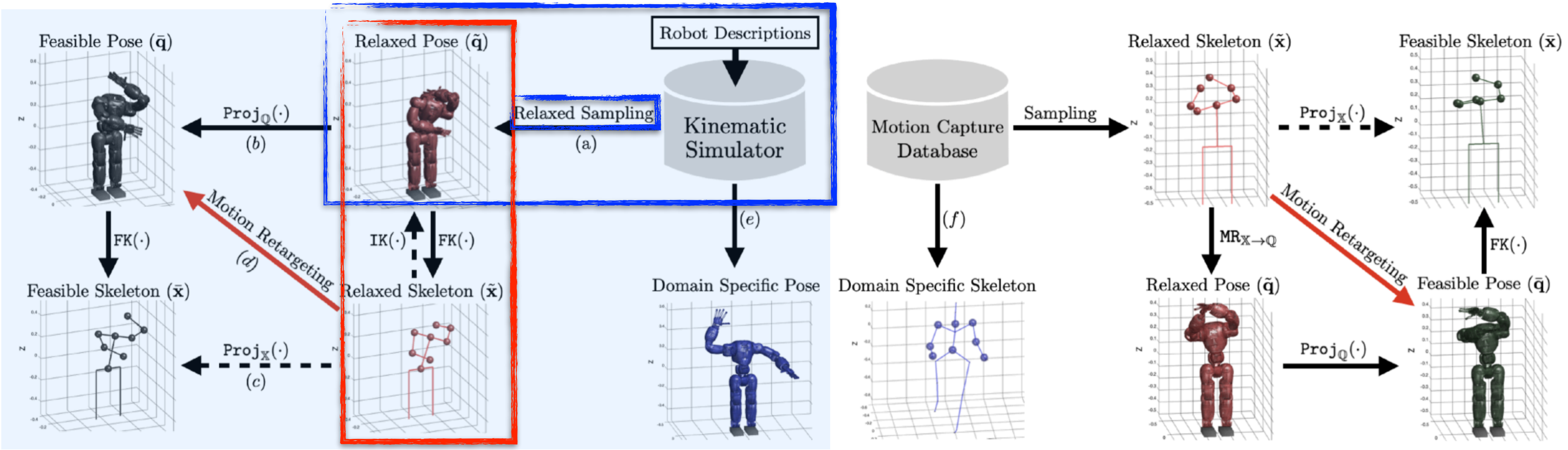
ROBOT INTELLIGENCE LAB





- There are two paths of automated training data sampling pipelines: 1) **from the robot descriptions** and 2) **from the motion capture data**.

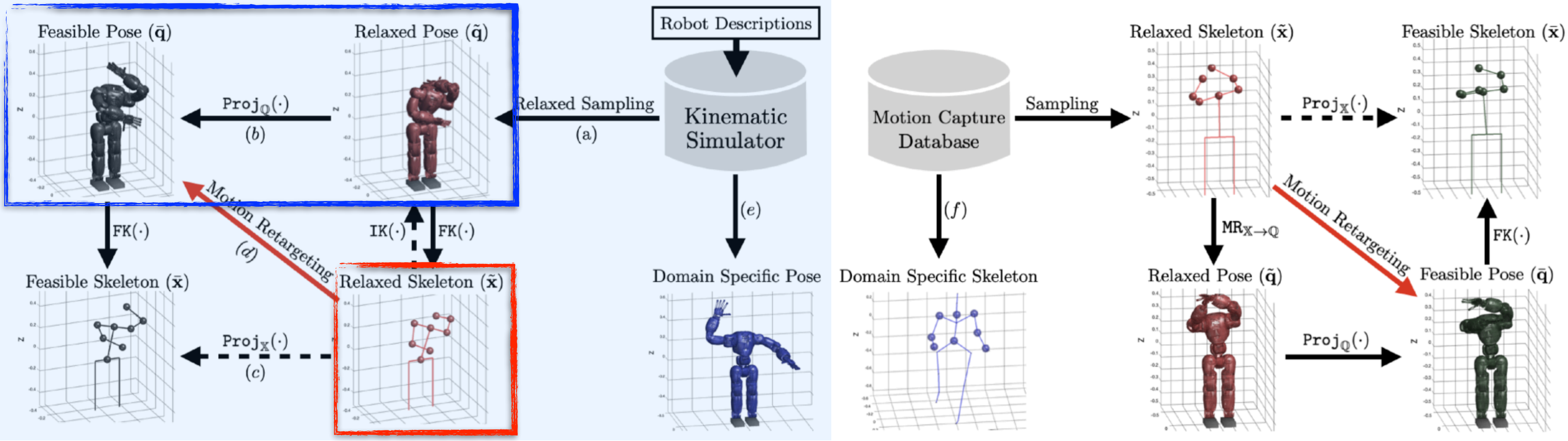
Automated Sampling Pipelines



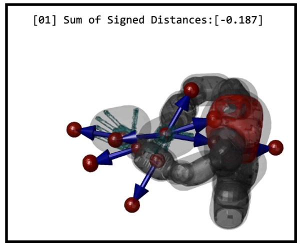
- From the robot descriptions (i.e. URDF), **(a)** we randomly **sample** a relaxed robot pose by sampling joint positions from the **relaxed joint limits** and get the corresponding robot pose by solving **forward kinematics**.

Automated Sampling Pipelines

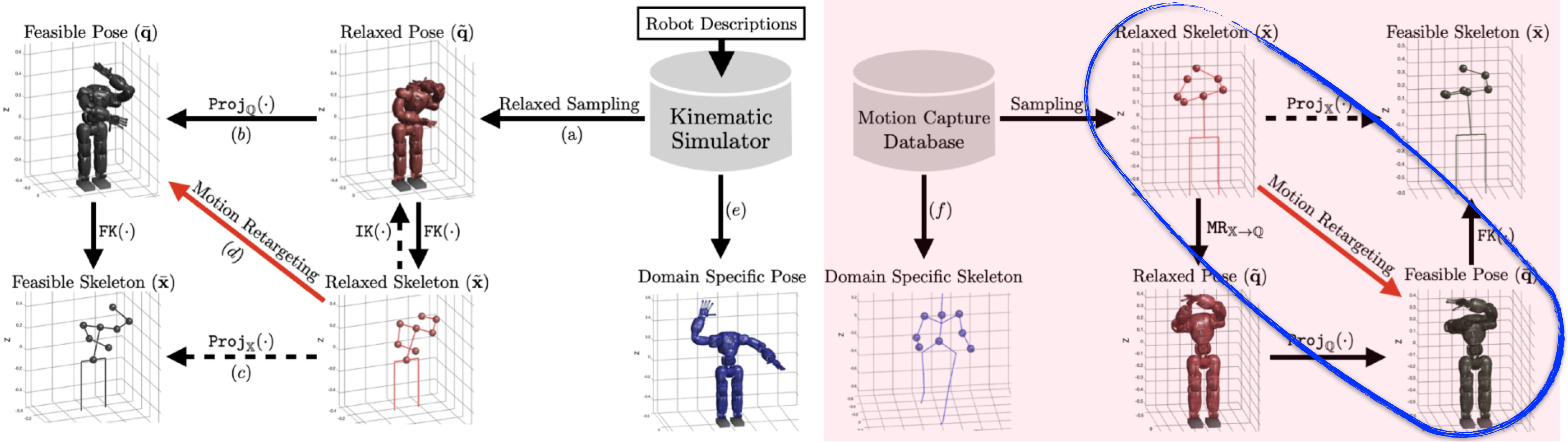
ROBOT INTELLIGENCE LAB



- Then, we find the corresponding **feasible pose** from the relaxed pose using the projection mapping (i.e., handling joint limit violations and self-collisions).



Automated Sampling Pipelines

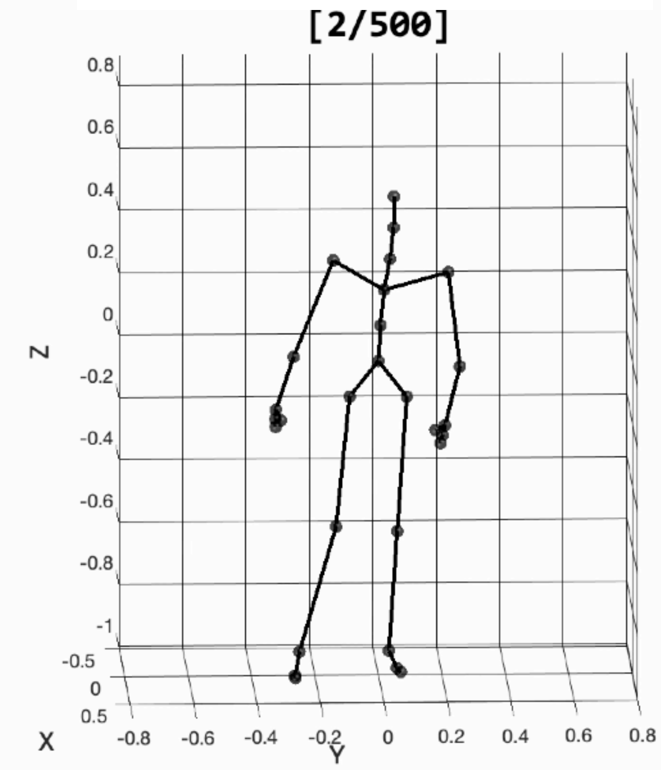


- Similarly, we utilize the **optimization-based motion retargeting** method with the projection mapping the collect training data.

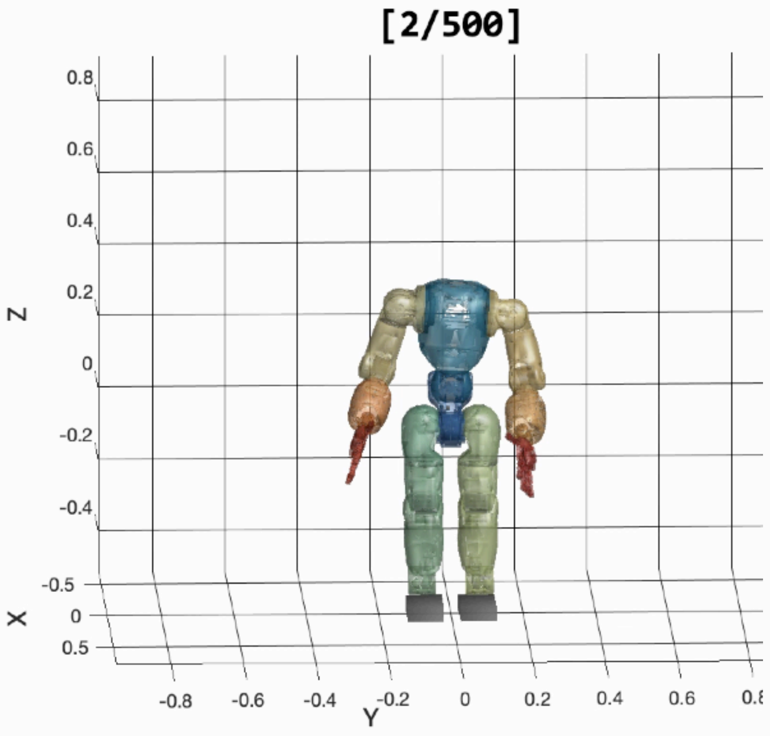
Results

ROBOT INTELLIGENCE LAB

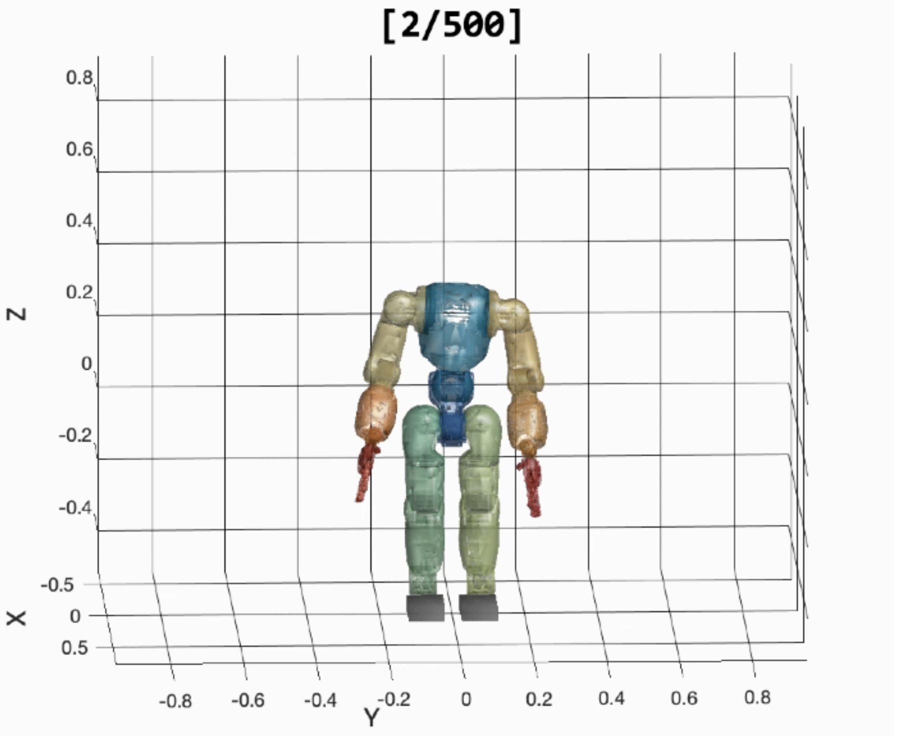
Source Motion



Proposed Method



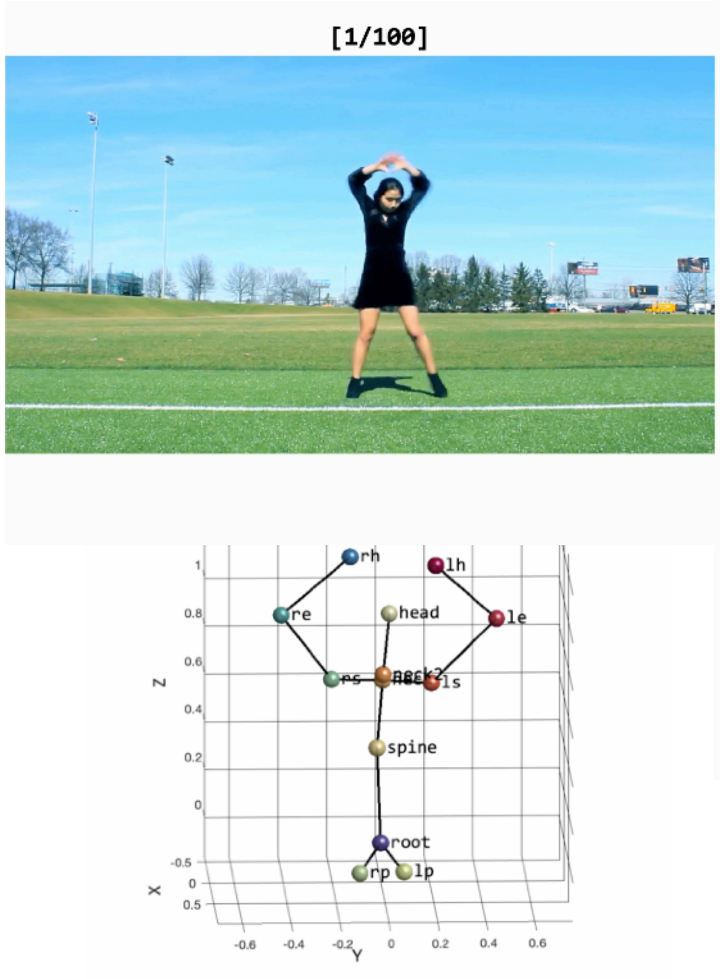
Base Line



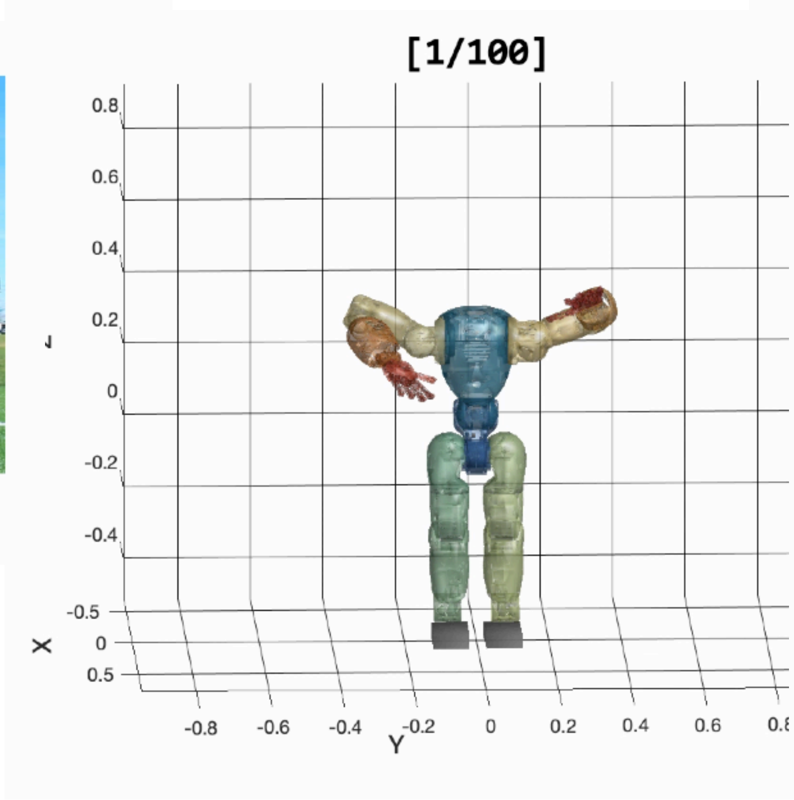
Results

ROBOT INTELLIGENCE LAB

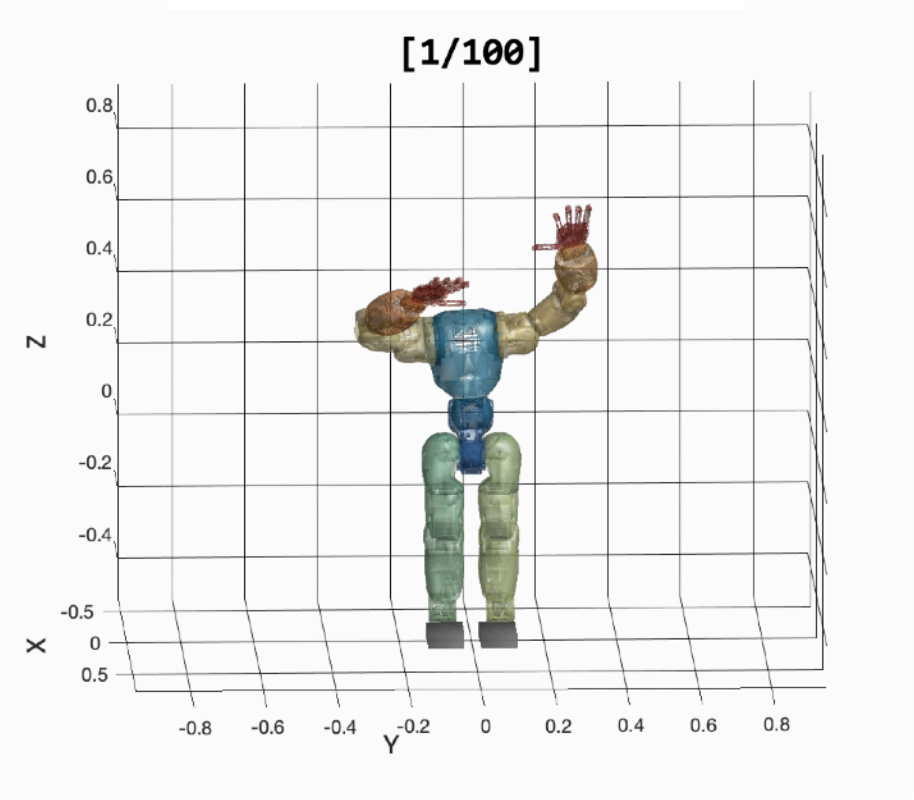
Source Motion



Proposed Method

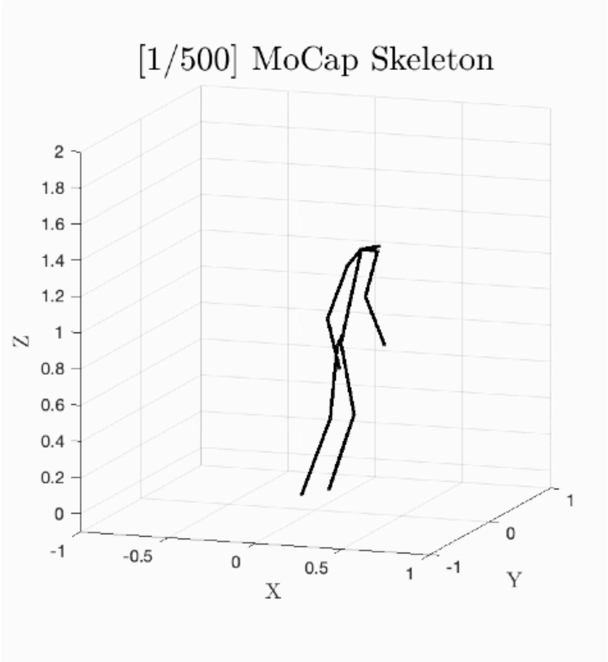


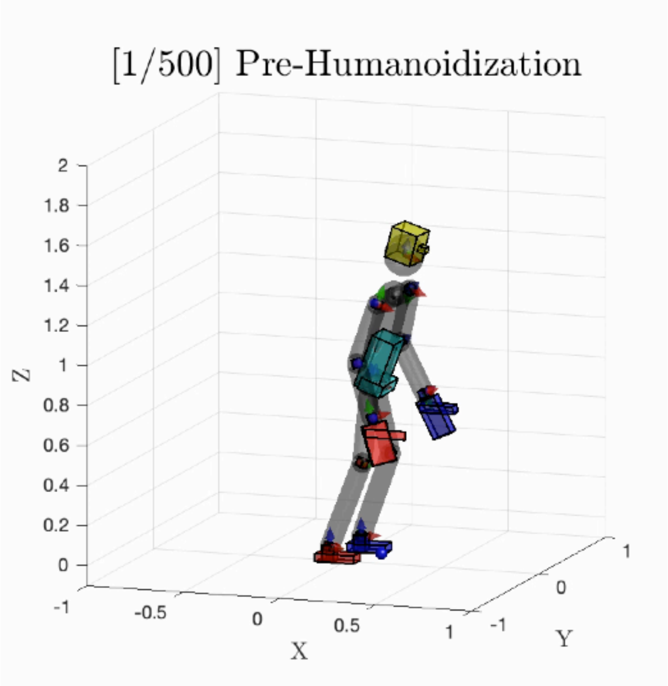
Base Line

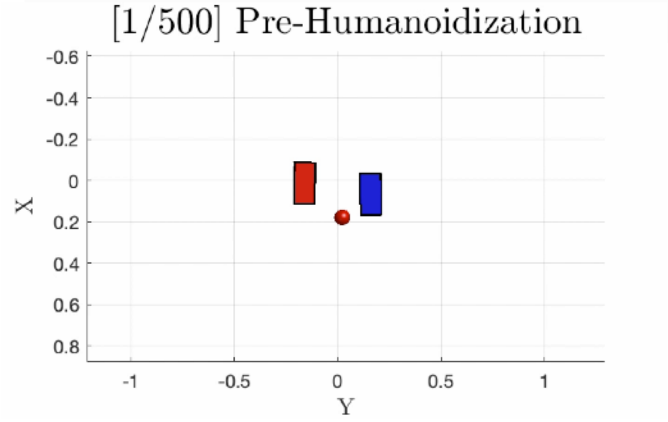


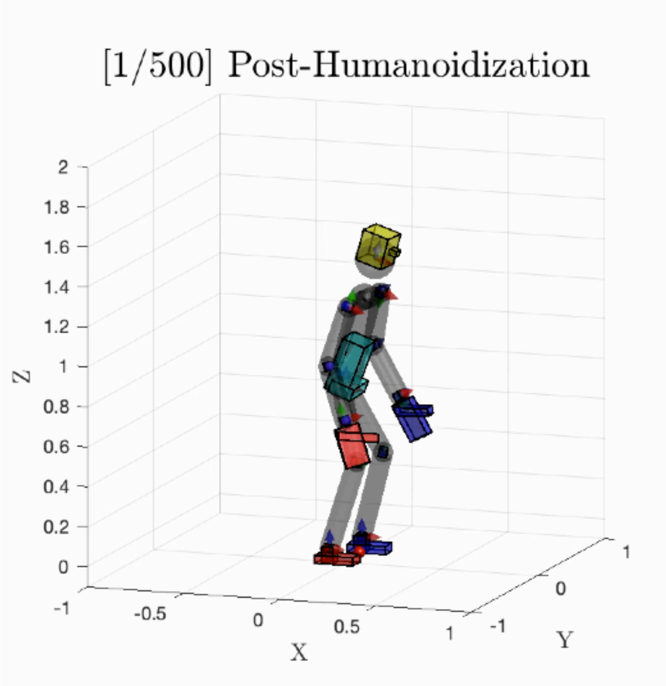
Robust Motion Retargeting

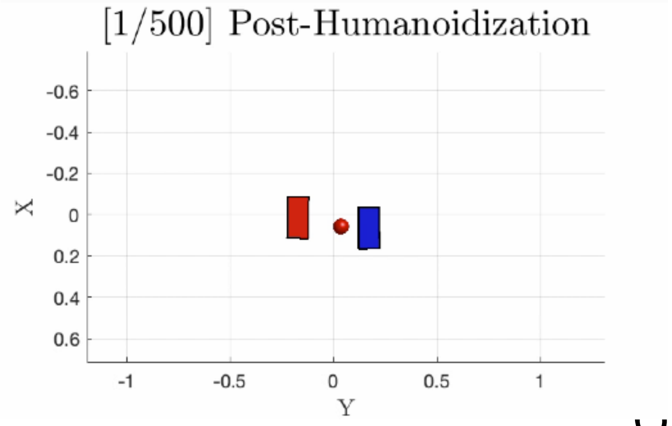
MoCap Handling





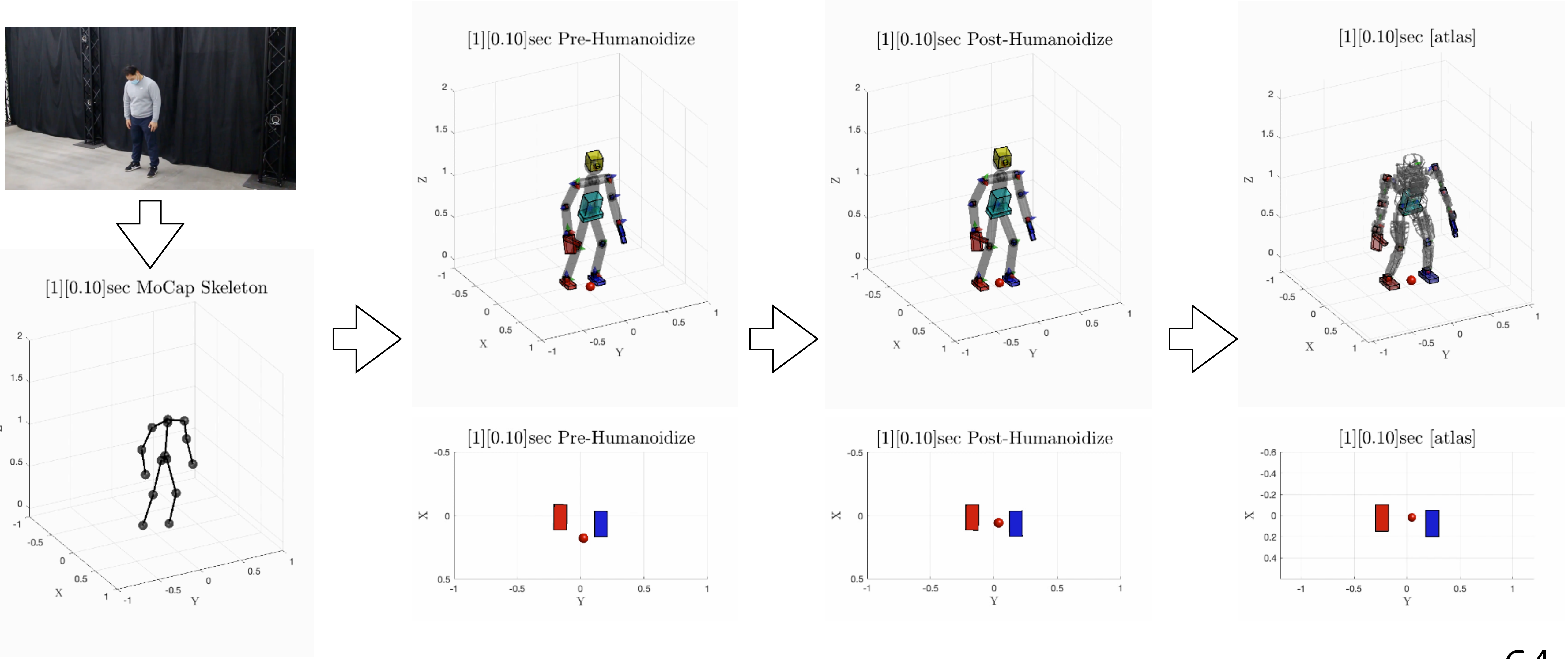






Robust Motion Retargeting

ROBOT INTELLIGENCE LAB



Thank You

ROBOT INTELLIGENCE LAB

ROBOT INTELLIGENCE LAB

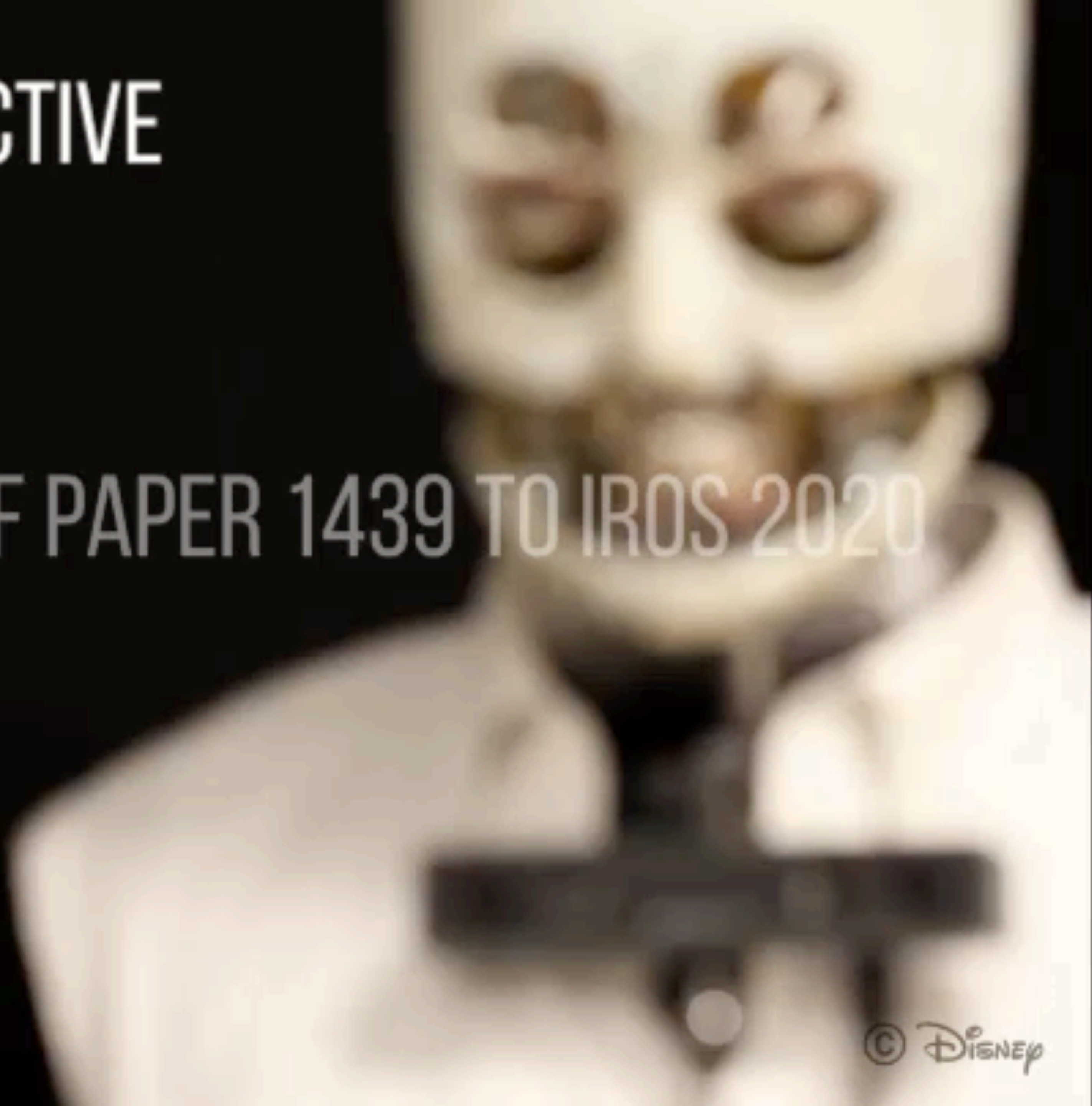
Interactive Gaze

Realistic and Interactive Robot Gaze, IROS 2020

REALISTIC AND INTERACTIVE ROBOT GAZE

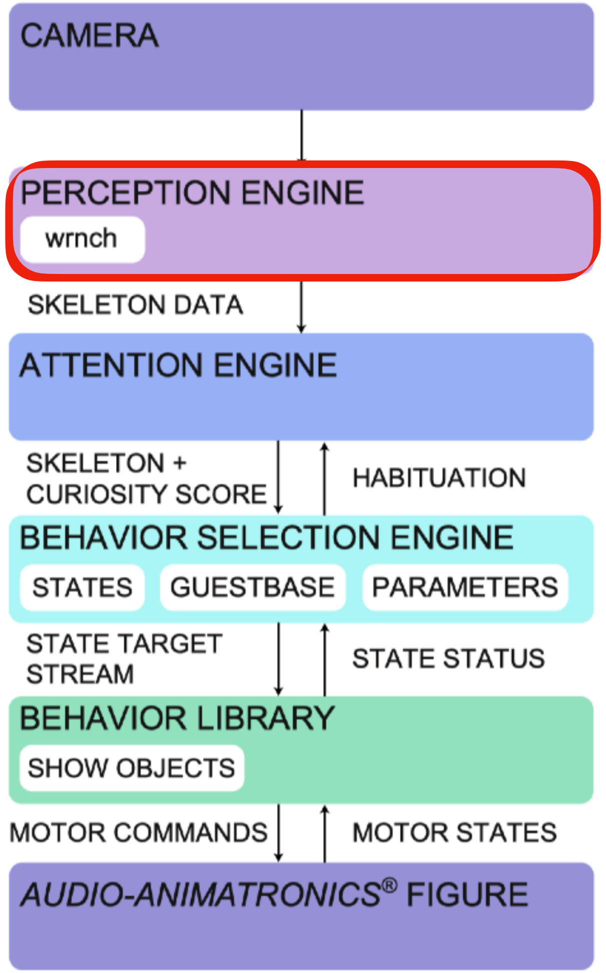
WITH AUDIO

VIDEO SUBMISSION OF PAPER 1439 TO IROS 2020



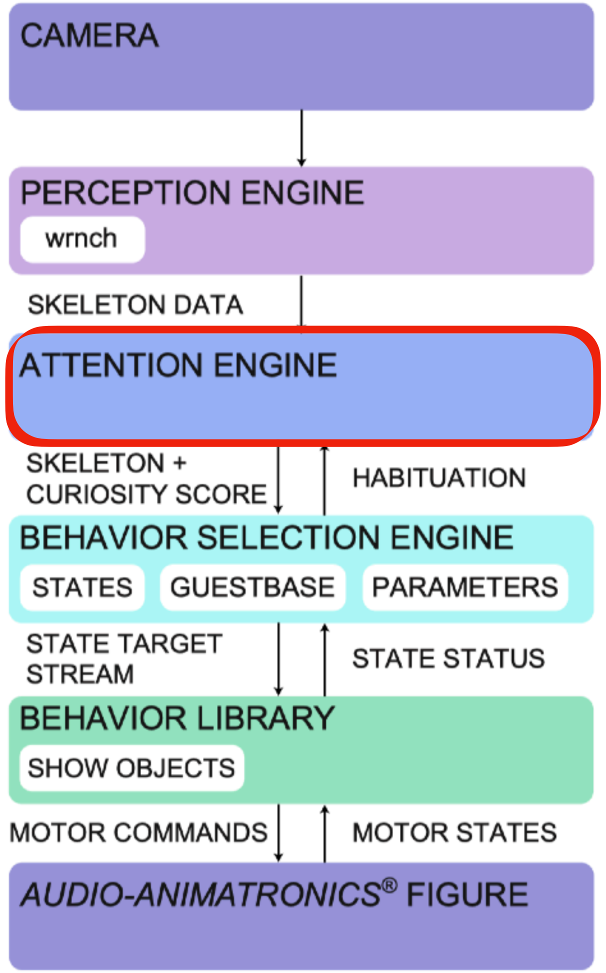
Disney Research

Interactive Gaze



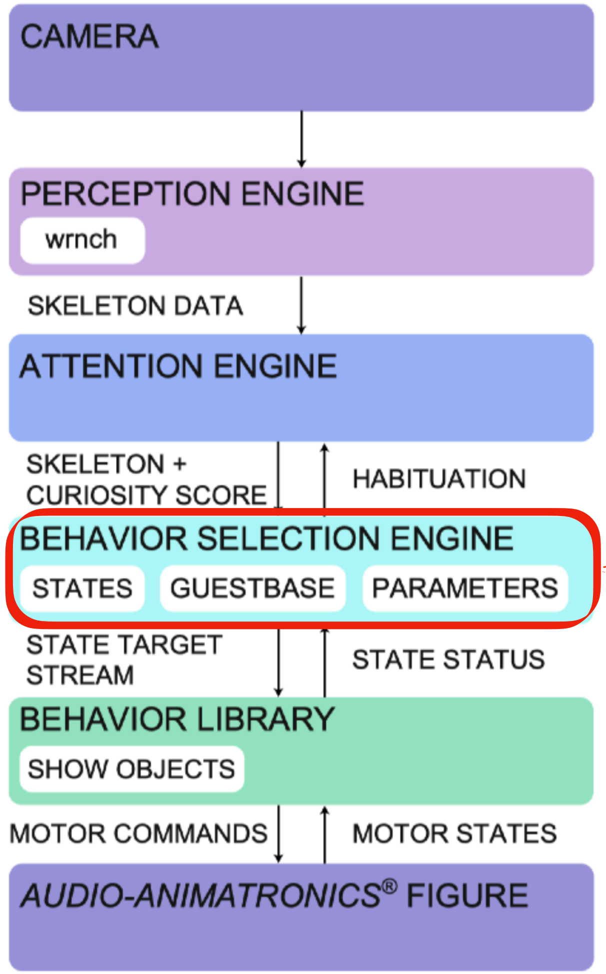
Perception engine infers useful information from visual observation such as skeleton poses on the scene.

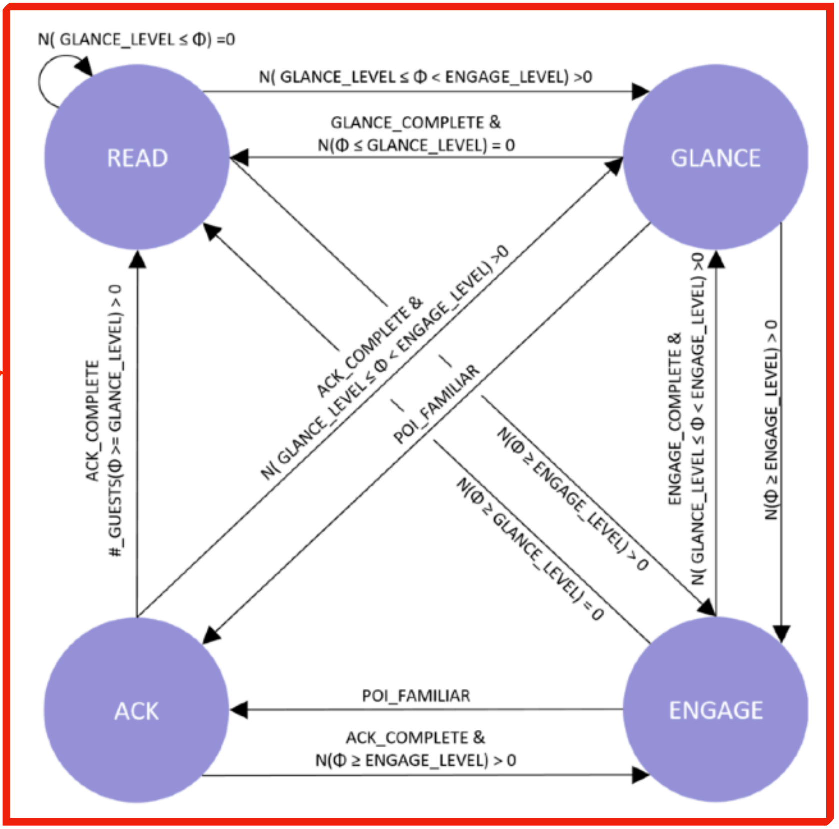
Interactive Gaze



Attention engine computes the **curiosity level** of each person from the sequence of skeleton data.

Interactive Gaze





Behavior selection engine decides which behavior the robot should select based on the curiosity levels of each guest.