Towards Physical Plausible Simulation of Human-Scene Interaction

Bo Dai

bdai[at]hku.hk
bdai[at]feeling-ai.com

Spatial Intelligence

The ability to perceive, understand and interact with the world in three dimensions

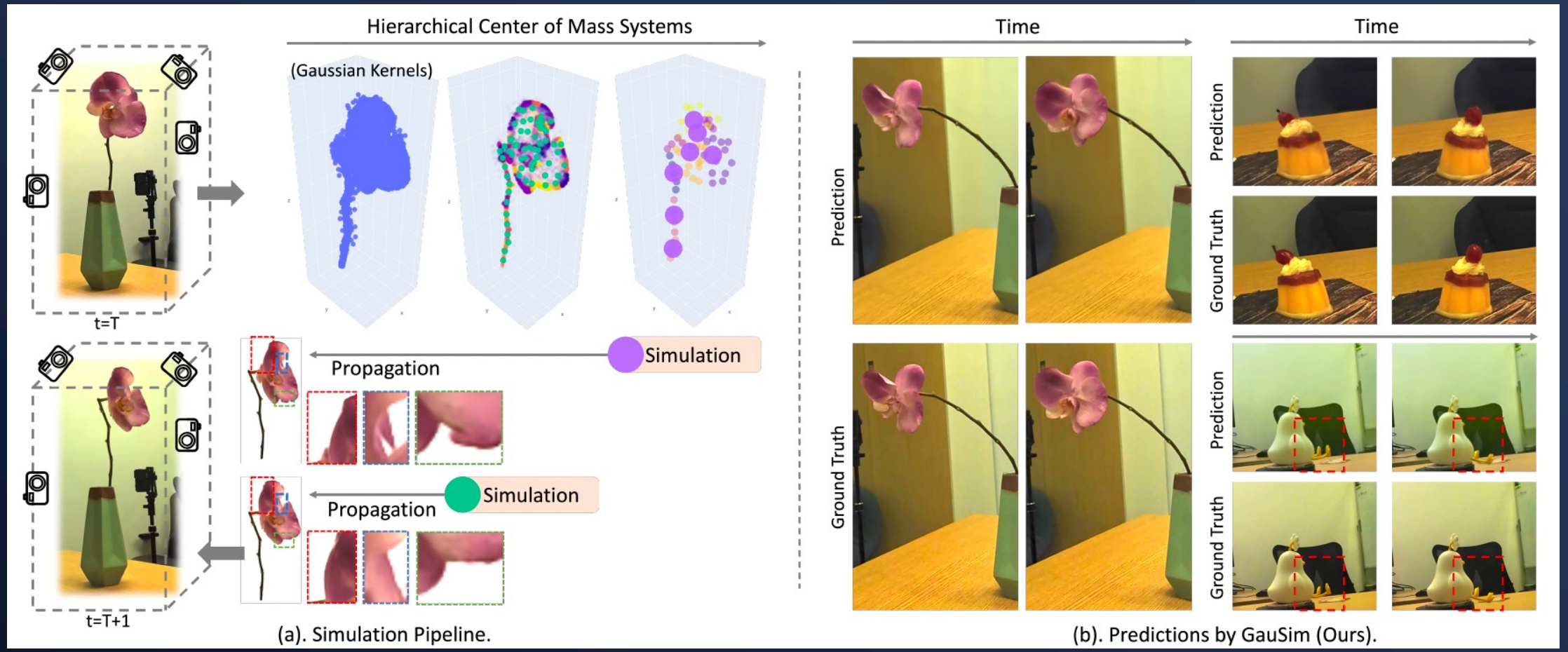
Image Credit: https://www.ainews.com/

Learning of Spatial Intelligence

- 3D virtual worlds that support physical plausible interaction

- General purpose agents that conduct physical plausible interaction

GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects



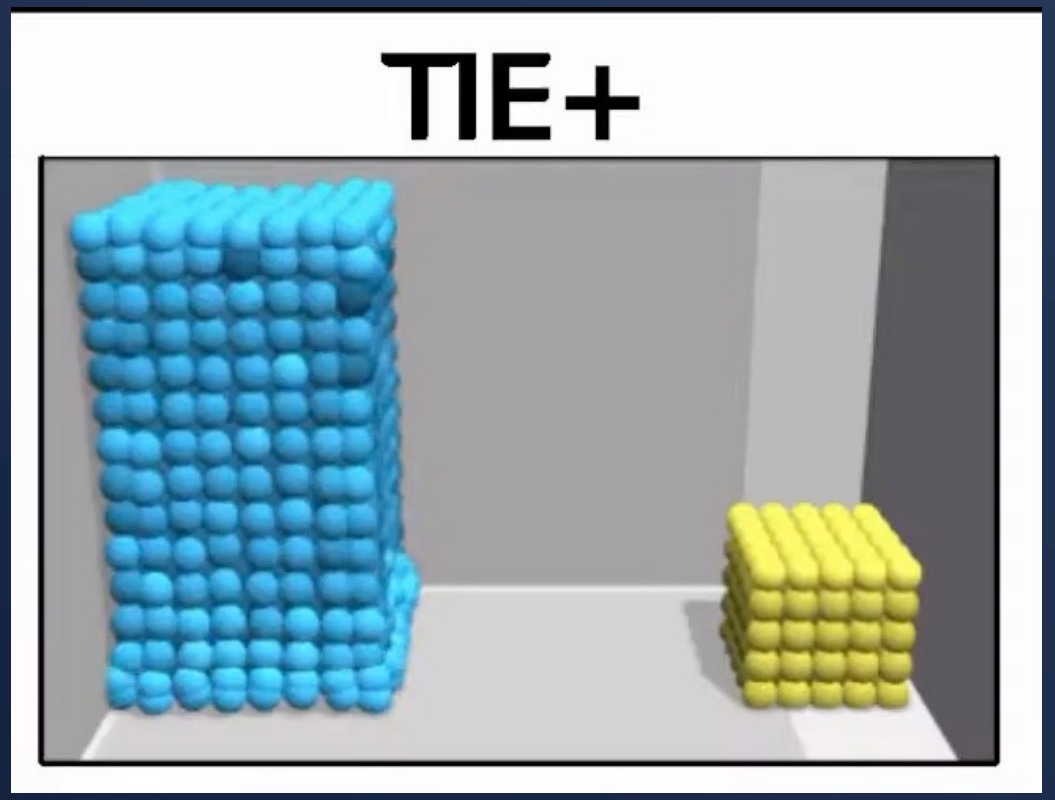
Enabling Physical Plausible Interaction with 3D Gaussian Splatting

GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects

- Combining 3DGS with neural simulation
  - Looks real – perception
  - Acts real – interaction

3DGS (SIGGRAPH 2023)

+



TIE (ECCV 2022)

GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects

- Combining 3DGS with neural simulation
  - Per-kernel deformation matrix F based on interactions of kernels

$$\mathcal{G}_{t+1} = \psi(\mathcal{G}_t, \mathcal{G}_{t-1}, \mathcal{A})$$

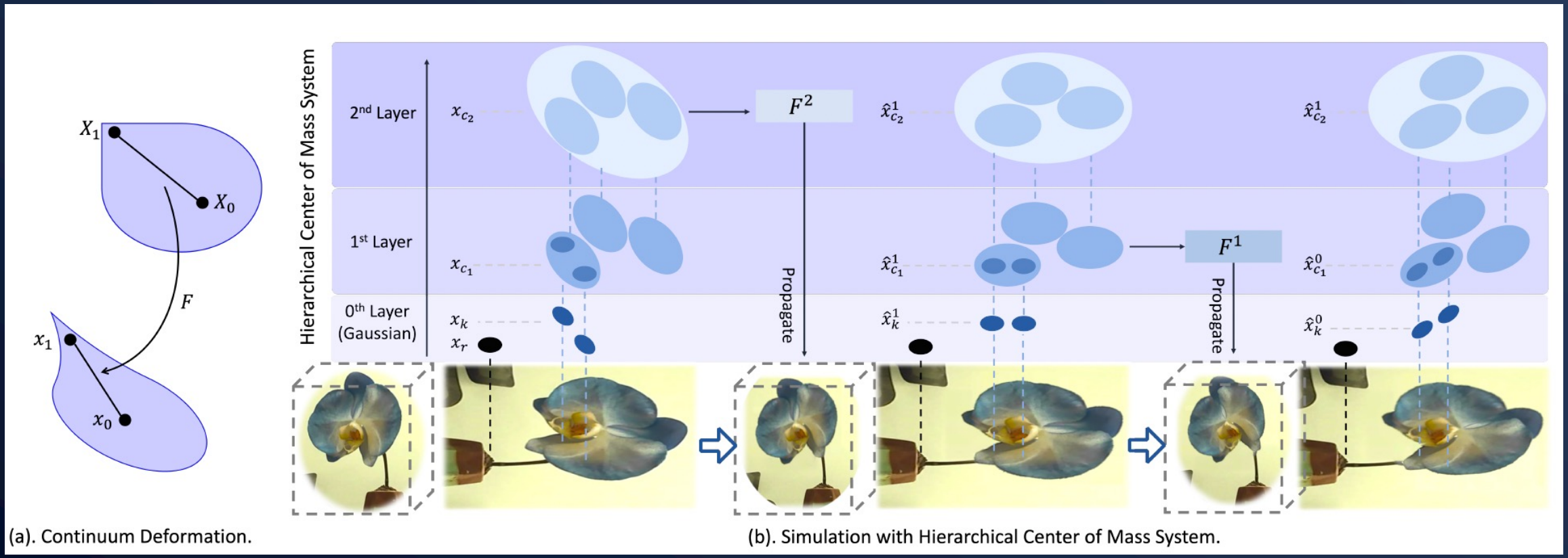
$\mathcal{A} = \{\rho_k, \boldsymbol{a}_k\}_{k\in\mathcal{K}}$ attributes and densities of Gaussian kernels

$\mathcal{G} = \{\boldsymbol{x}_k, \boldsymbol{\sigma}_k, \boldsymbol{c}_k, \alpha_k\}_{k\in\mathcal{K}}$ positions, covariances, colors, opacities of Gaussian kernels

- Instead of treating Gaussian kernels as discrete particles for particle-based simulations, we consider each kernel as a continuous piece of matter
  - Shape of Gaussian kernels are important in 3D-GS

GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects

- Hierarchical Structural Center of Mass Systems



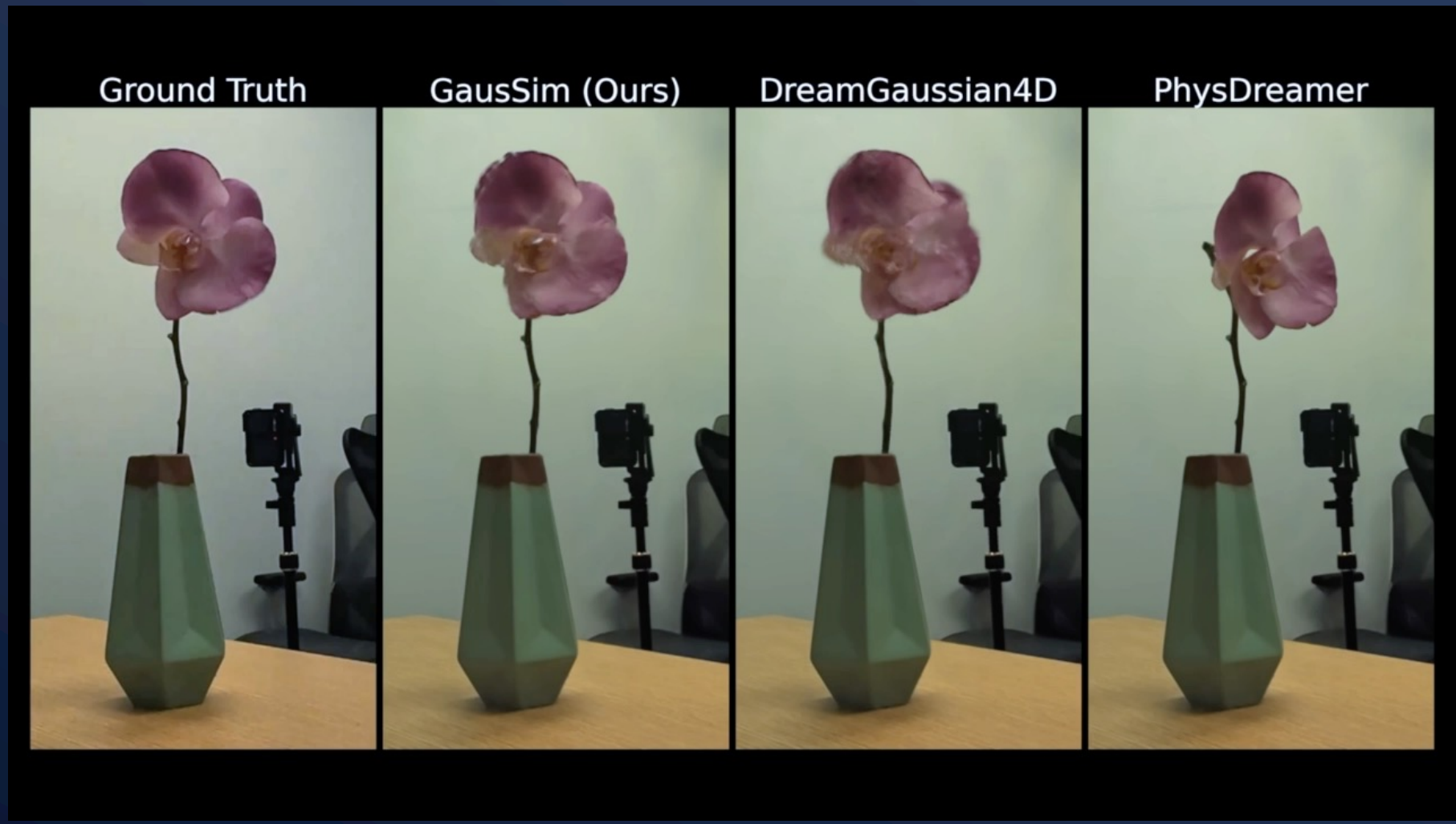
(a). Continuum Deformation.

(b). Simulation with Hierarchical Center of Mass System.

GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects

	Time (s)	GPU (GB)
DG4D [28]	0.14±0.01	7.6
PD [42]	1.67±0.05	4.6
GausSim w/o H	0.36±0.01	3.5
GausSim	**0.13±0.01**	**2.1**

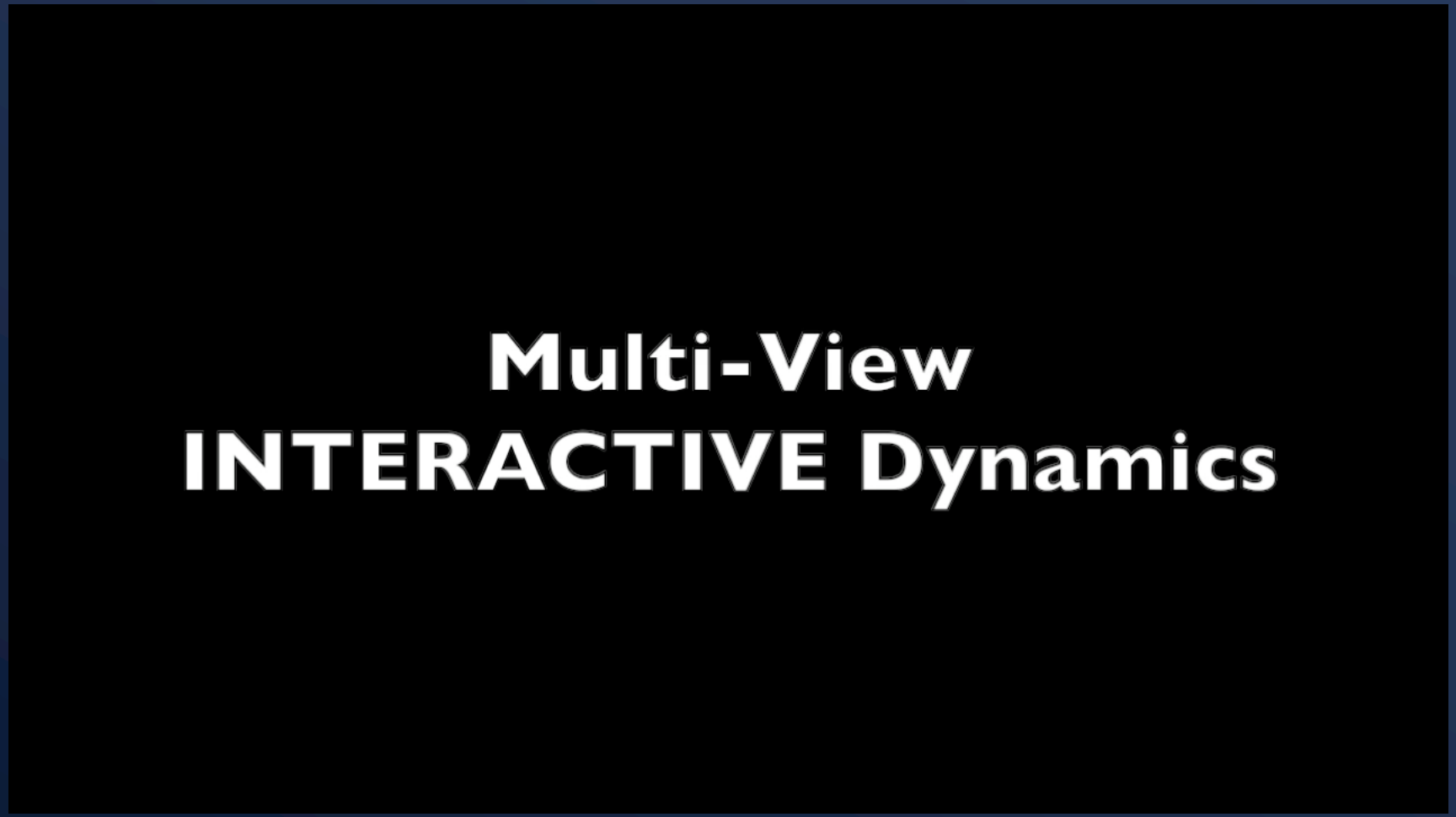
DG4D: DreamGaussian4D
PD: PhysDreamer



GausSim: Foreseeing Reality by Gaussian Simulator for Elastic Objects

	Time (s)	GPU (GB)
DG4D [28]	0.14±0.01	7.6
PD [42]	1.67±0.05	4.6
GausSim w/o H	0.36±0.01	3.5
GausSim	**0.13±0.01**	**2.1**

DG4D: DreamGaussian4D
PD: PhysDreamer

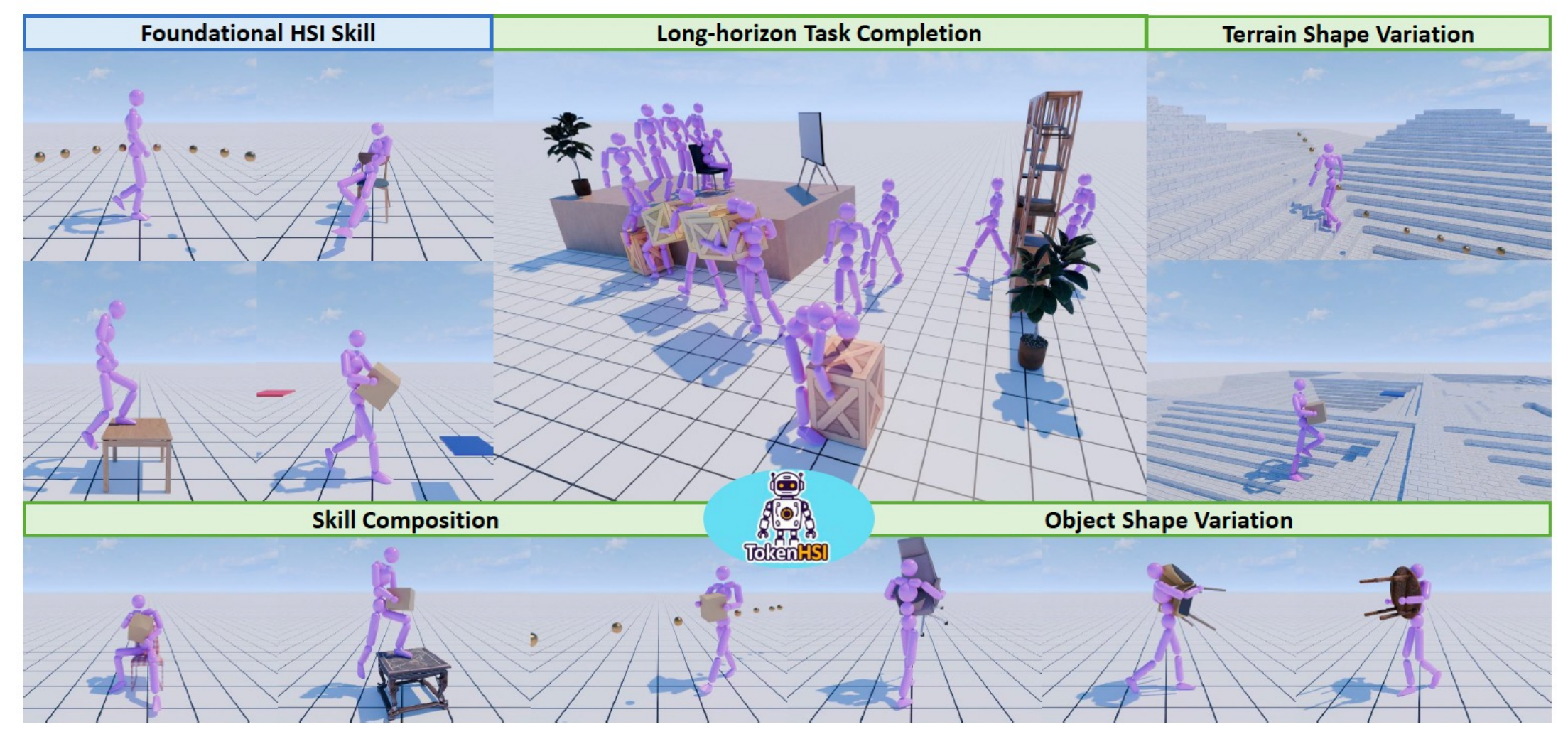


Learning of Spatial Intelligence

- 3D virtual worlds that support physical plausible interaction

- General purpose agents that conduct physical plausible interaction

TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization



Enabling Unified Physical Plausible Interaction of Humanoid Agents

TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization

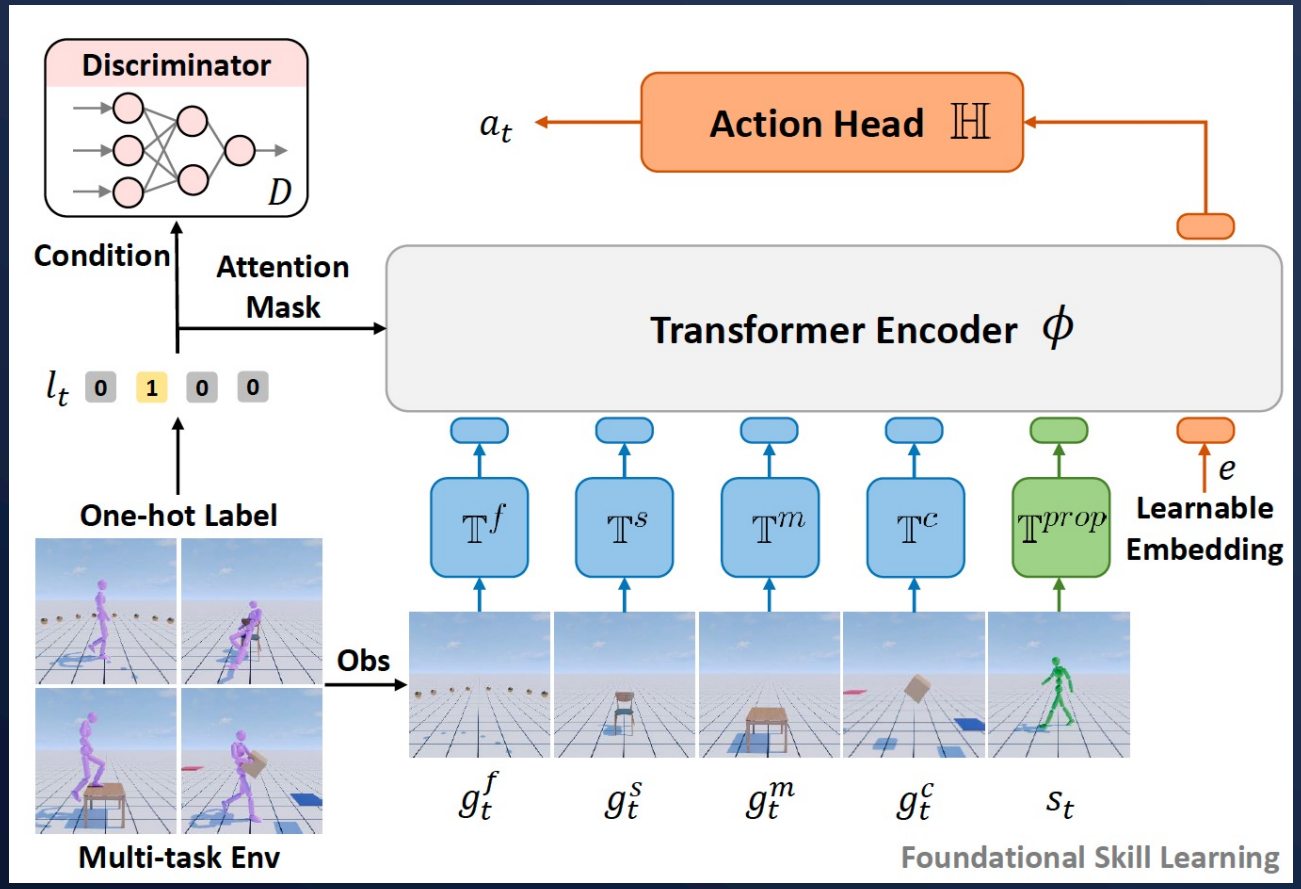
- Unified human-scene interaction
  - Effective joint learning of multiple seen tasks
  - Efficient extension to novel unseen tasks

Task-agnostic humanoid information through a proprioception tokenizer

Task-specific observation information through task tokenizers of same architecture (MLP)

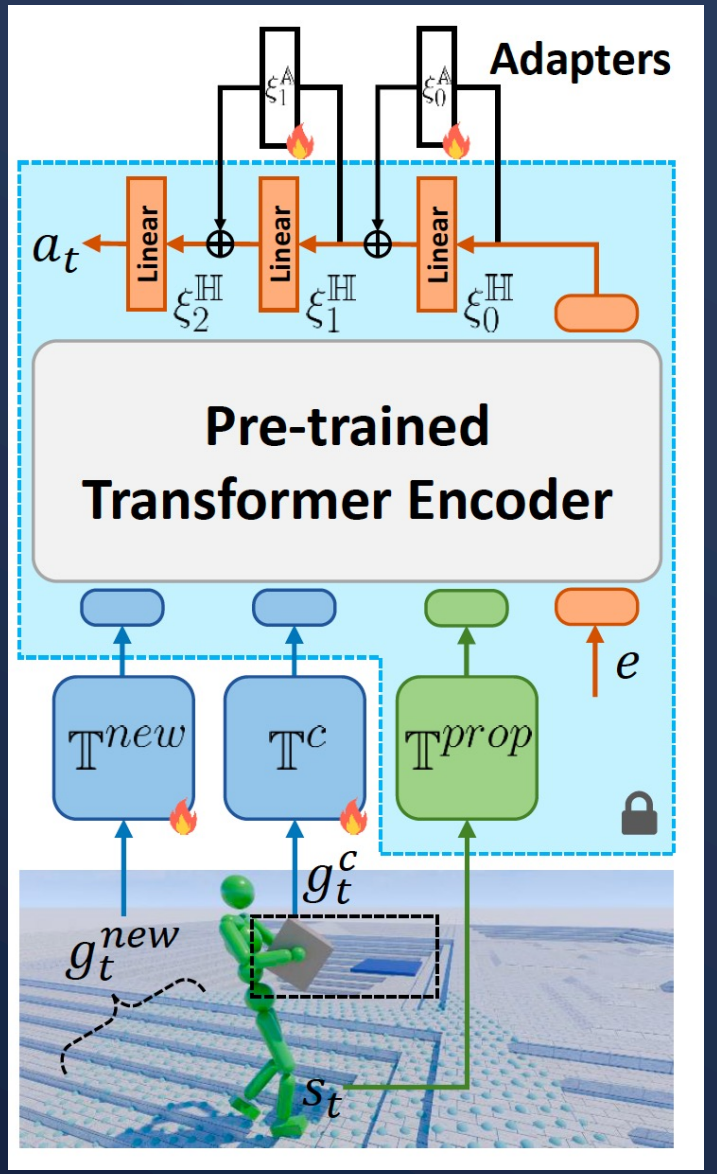
TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization

- Effective joint learning of multiple seen tasks
  - A transformer-based policy network with task-oriented masking

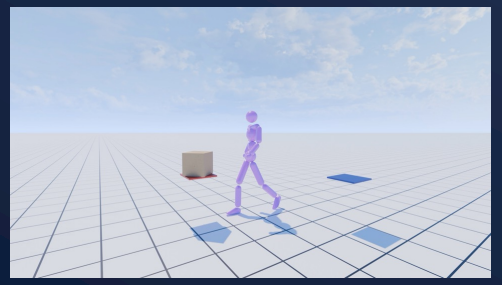


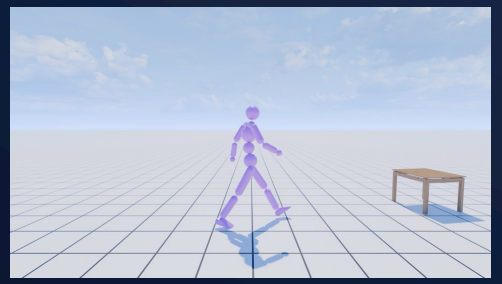
TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization

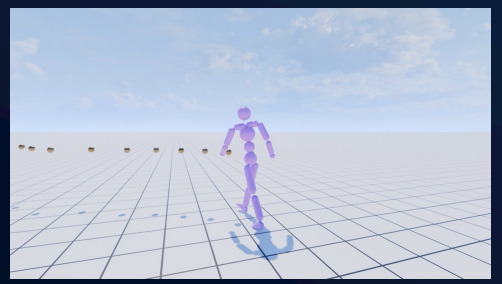
- Efficient extension to novel unseen tasks
  - Task tokenizers for new observation
  - Task LoRAs to the action head



TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization





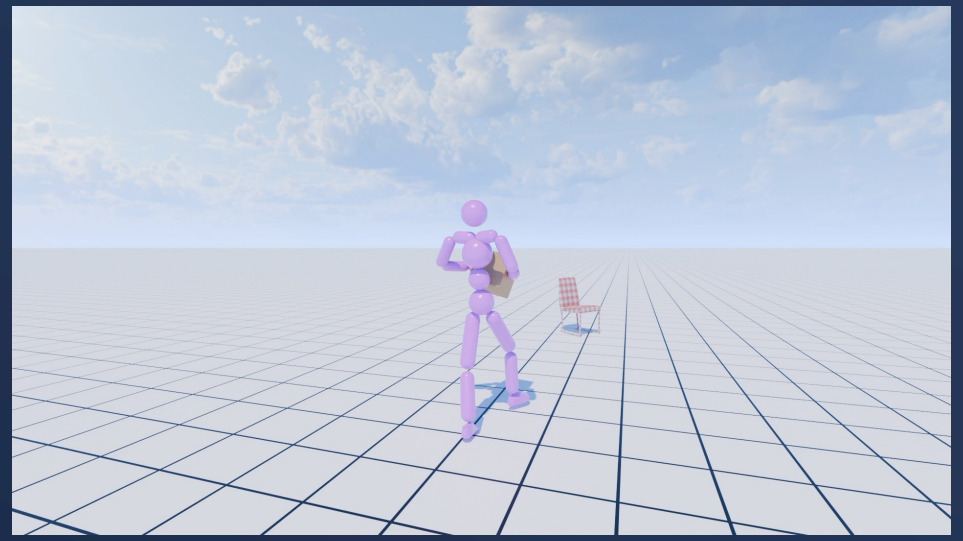


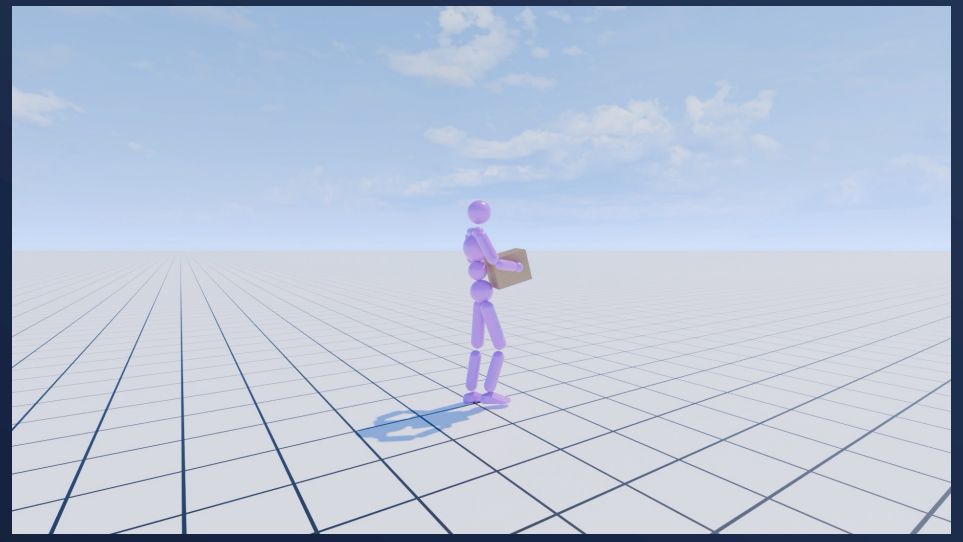
Sit

Carry

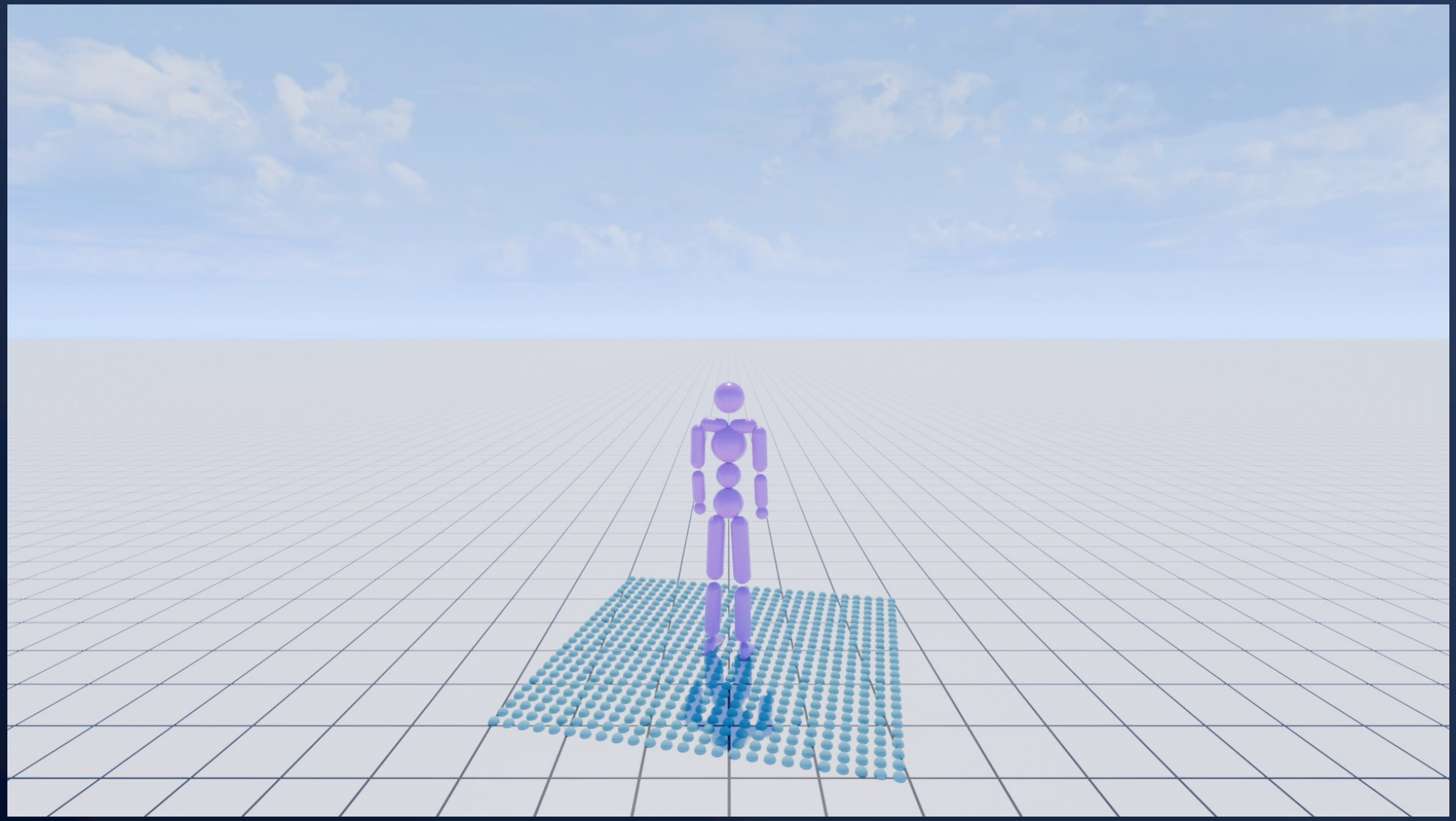
Climb

Follow Traj





TokenHSI: Unified Synthesis of Physical Human-Scene Interactions through Task Tokenization

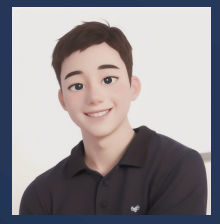
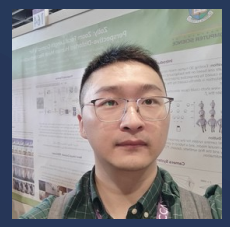
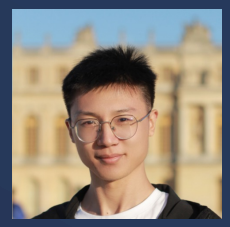


Long-term Compositional Task

- Coauthors





- Project Pages

GausSim

TokenHSI

Poster@China3DV
3-15

- PhD/Postdoc/RA/Researcher/Engineer/Intern Opportunity

bdai[at]hku.hk
bdai[at]feeling-ai.com