PhysX-3D: Physical-Grounded 3D Asset Generation

Ziang Cao[1], Zhaoxi Chen[1],Liang Pan[2], Ziwei Liu[1]

[1]Nanyang Technological University

[2]Shanghai AI Lab

NEURAL INFORMATION PROCESSING SYSTEMS

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

S-LAB FOR ADVANCED INTELLIGENCE

Motivation

- **Absence of Physical-Grounded 3D Datasets.**

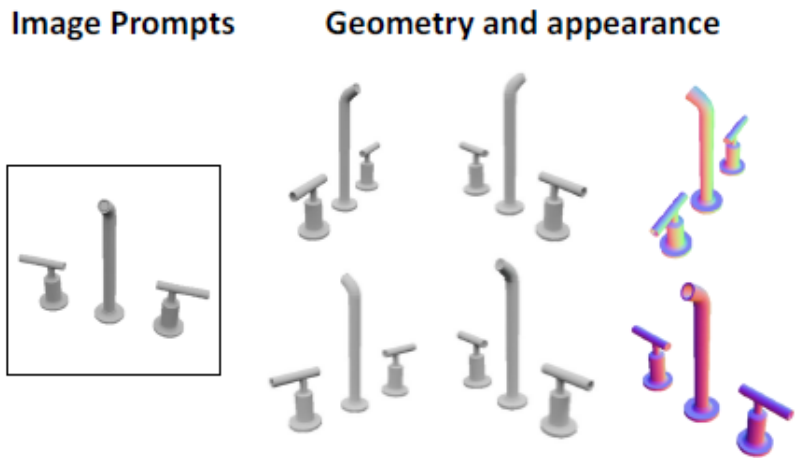
  Most existing 3D datasets primarily provide geometry and texture information, but they rarely include **physical properties** such as **absolute scale**, **material**, **affordance**, **kinematics**, and **function description**, thereby constraining the development of generative models.

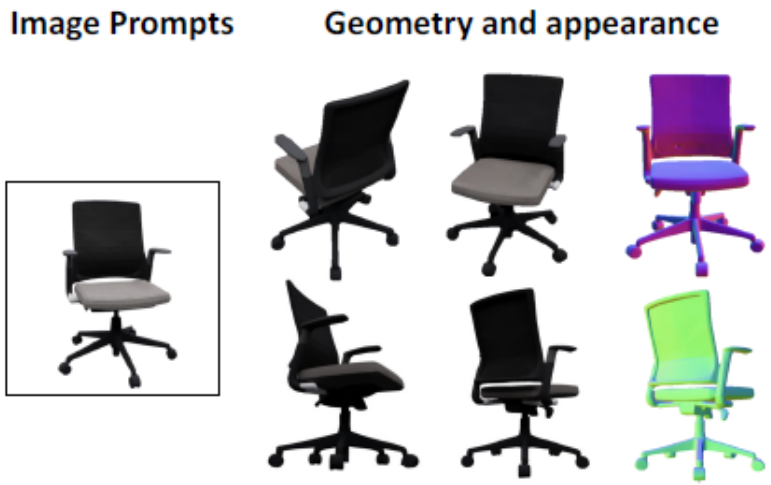
Dataset	# Objs	Part anno	Physical Dim	Material	Affordance	Kinematic	Description	Year
ShapeNet [3]	51K	✗	✗	✗	✗	✗	✗	2015
PartNet [17]	26K	✓	✗	✗	✗	✗	✗	2019
PartNet-Mobility [26]	2.7K	✓	✗	✗	✗	✓	✗	2020
GAPartNet [9]	1.1K	✓	✗	✗	✗	✓	✗	2022
ABO [6]	7.9K	✗	✓	Obj-level	✗	✗	Obj-level	2022
OmniObject3D [25]	6K	✗	✗	✗	✗	✗	✗	2023
Objaverse [8]	818K	✗	✗	✗	✗	✗	✗	2023
PhysXNet (ours)	26K	✓	✓	Part-level	✓	✓	Part-level	2025
PhysXNet-XL (ours)	6M	✓	✓	Part-level	✓	✓	Part-level	2025

Motivation

- **Limitations in current 3D generative methods.**

  Existing 3D generation primarily emphasizes geometries and textures while **neglecting physical-grounded modeling**, hampering their real-world application in physical domains like simulation and embodied AI.





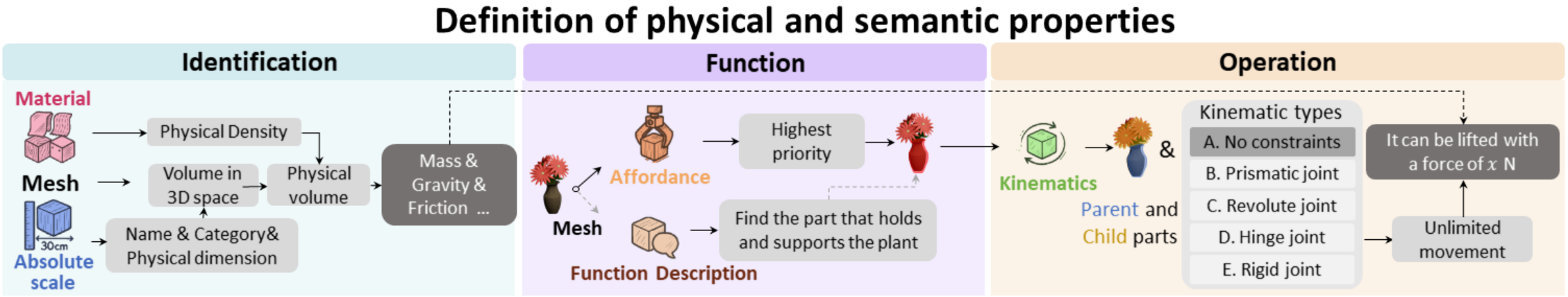
Contribution

- We pioneer the **first end-to-end paradigm** for physical-grounded 3D asset generation, advancing the research frontier in physical-grounded content creation and unlocking new possibilities for downstream applications in simulation.

- We build the **first physical-grounded 3D dataset**, **PhysXNet**, and propose a human-in-the loop annotation pipeline to convert existing geometry-focused datasets into fine-grained physics-annotated 3D datasets efficiently and robustly. In addition, we present an extended version, **PhysXNet-XL**, which includes over 6 million annotated 3D objects generated through procedural methods.

- We design a dual-branch feed-forward framework, **PhysXGen**. It can model the latent interdependencies between structural and physical features to achieve plausible physical predictions while maintaining the native geometry quality.

Definition of physical properties



- **Identification** - determining the basic nature of the object

  absolute scaling and material (material name, Young's modulus, Poisson's ratio, and density)

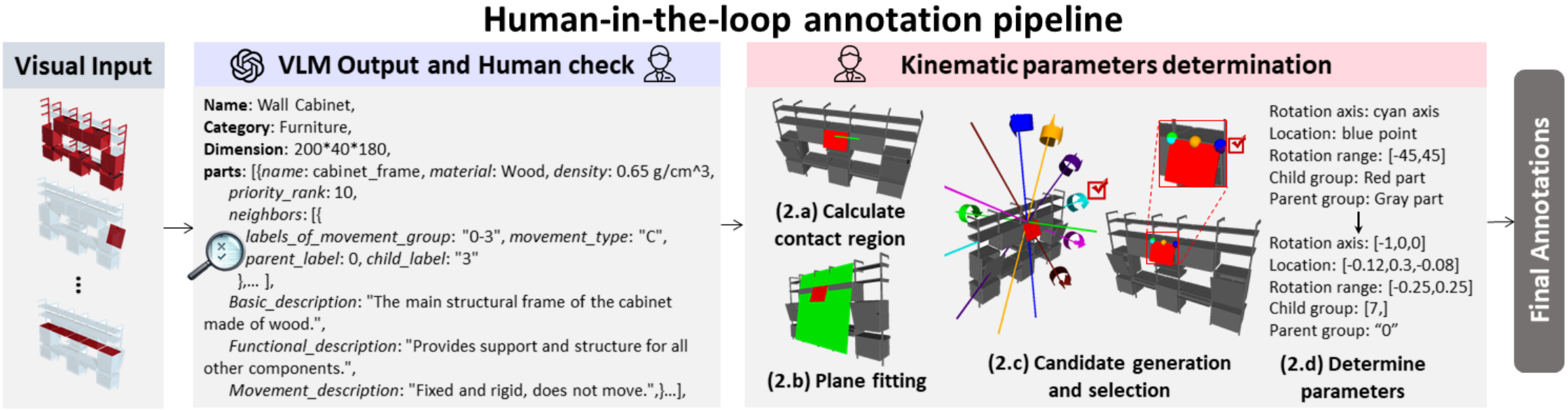
- **Function** - understanding its potential applications

  functional affordance analysis and function descriptions (basic, functional, and kinematic descriptions).

- **Operation** - detailed usage methodologies

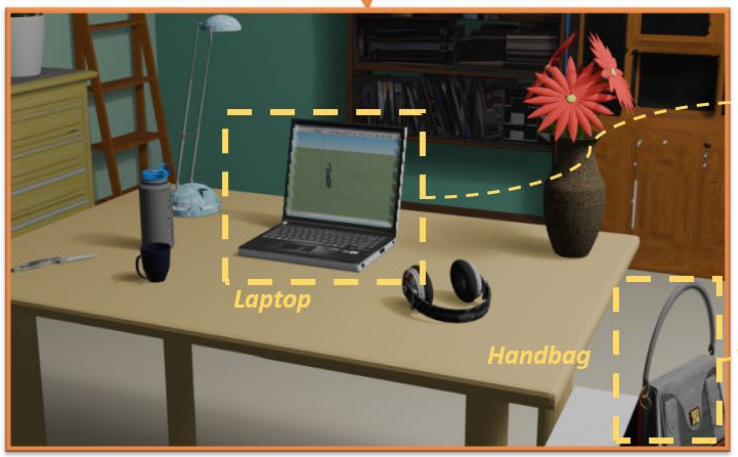
  kinematic parameter

HITL annotation pipeline

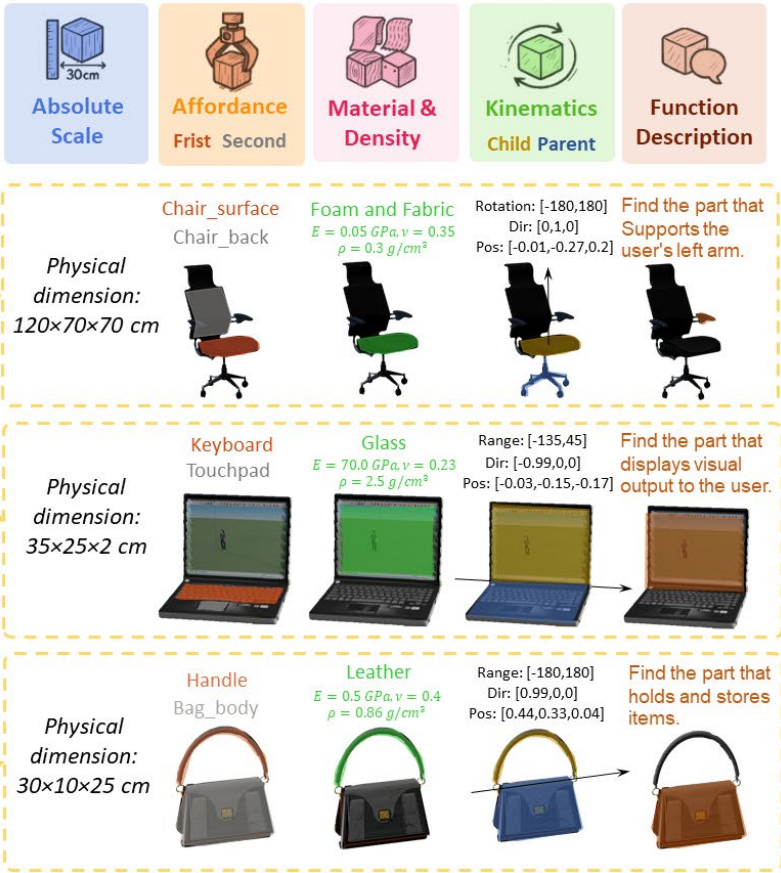


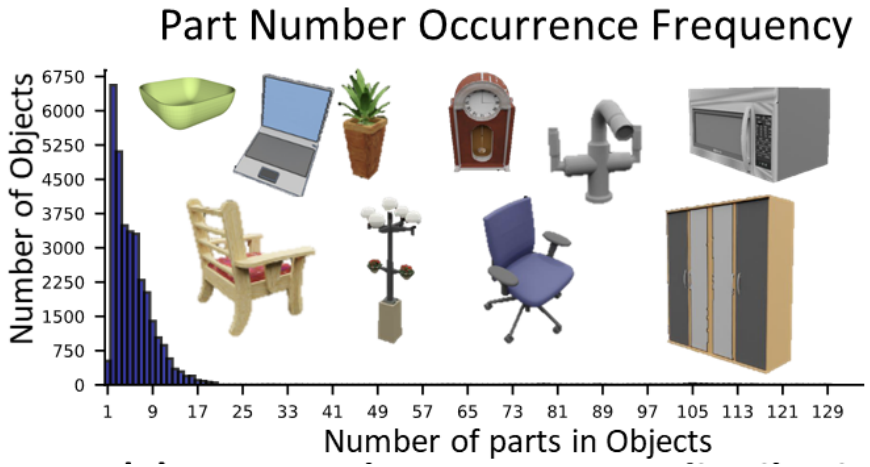
- **Preliminary Data Acquisition：**
  - Utilize GPT-4o to obtain the basic information
  - Human Check
- **Kinematic Parameter Determination**
  - Calculate contact region
  - Plane fitting
  - Candidate generation and selection
  - Determine kinematic parameters by human

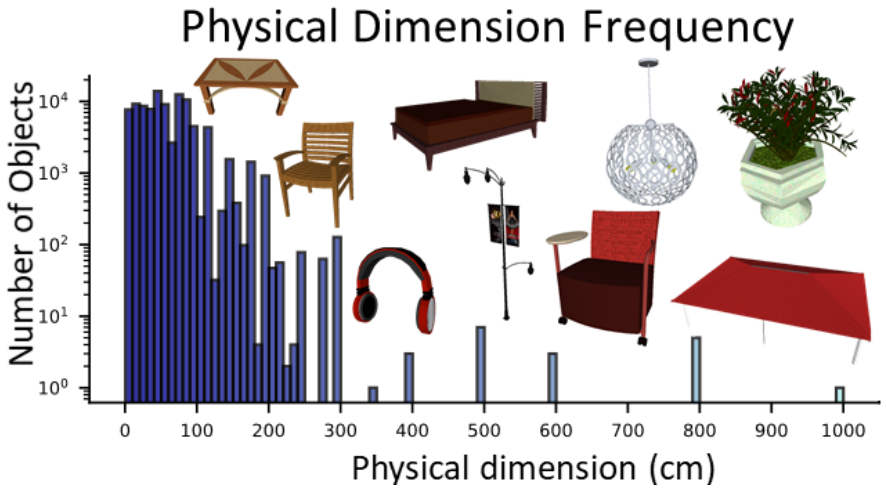
PhysXNet&PhysXNet-XL



Physical properties

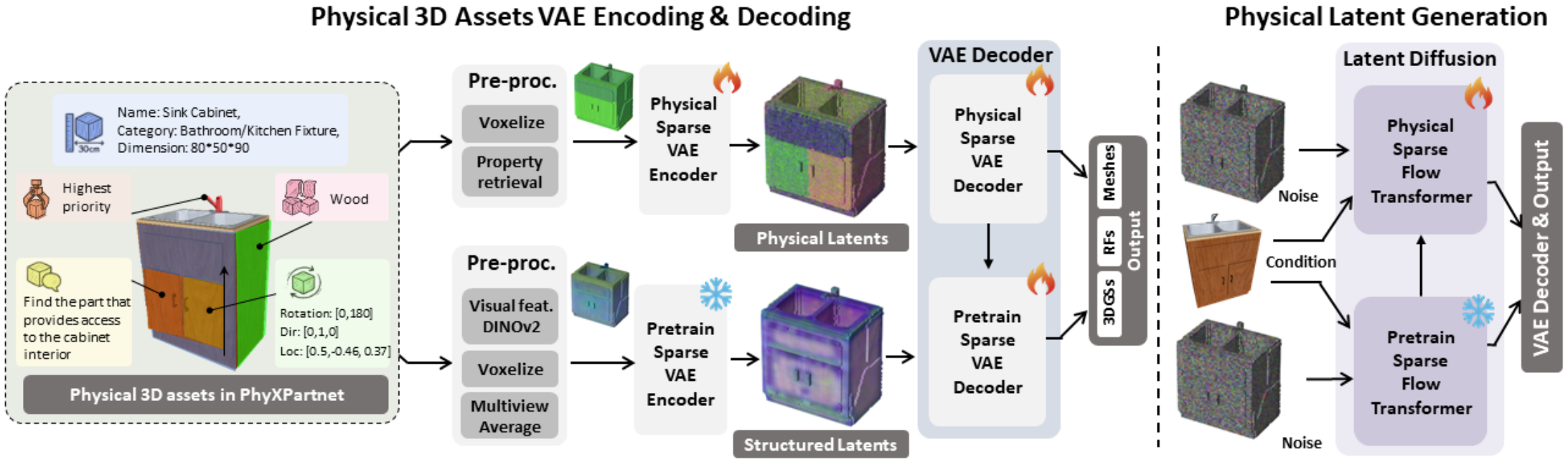






We propose **PhysXNet&PhysXNet-XL** – the first comprehensive physical 3D dataset containing over **26K&6M** richly annotated 3D objects

PhysXGen

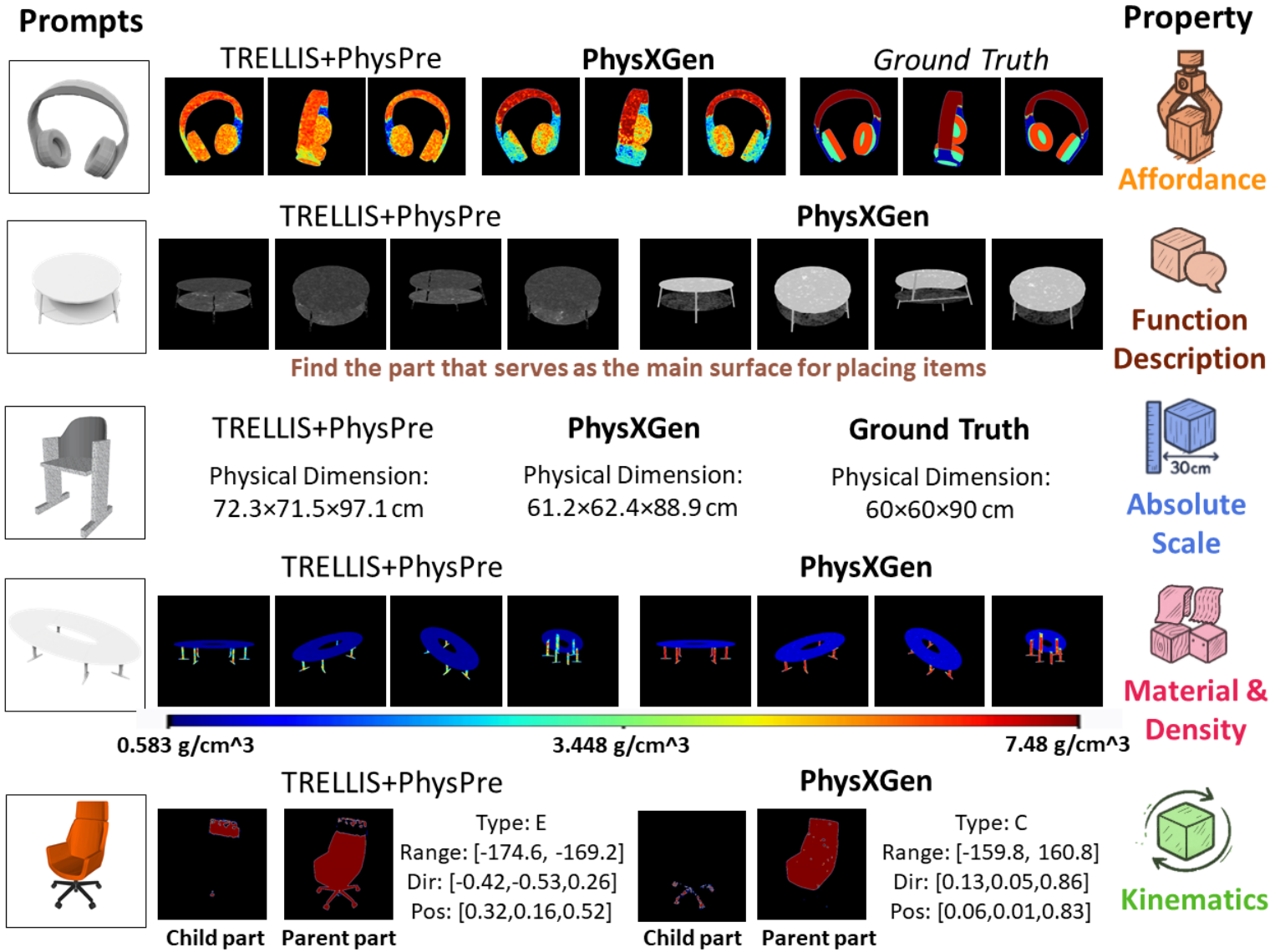


PhysXGen features a two-stage architecture comprising:

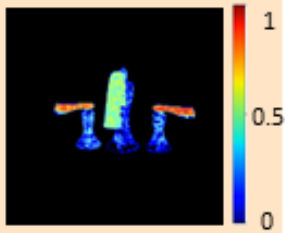
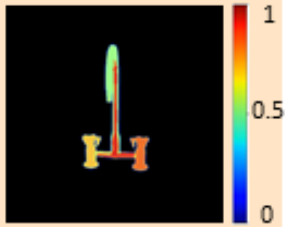
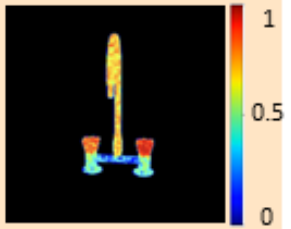
- physical 3D VAE framework for latent space learning
- physics-aware diffusion process for structured latent generation.

Experimental results

Experimental results



Experimental results

Image prompts	Trellis+PartField+GPT-4o			PhysXGen		
	**Absolute Scale** *Physical dimension: 20×15×12 cm*	**Affordance** 	**Material** 	**Absolute Scale** *Physical dimension: 24.31×18.19×14.7 cm*	**Affordance** 	**Material** 
	**Affordance** 	**Function Description** Find the part that delivers water from the valve system to the sink 		**Affordance** 	**Function Description** Find the part that delivers water from the valve system to the sink 	
	**Kinematics** Kinematic type: rotation Range: [-180,180] Dir: [1,0,0] Pos: [0,0,0] **Child part**  **Parent part** 			**Kinematics** Kinematic type: rotation Range: [-87.3,10.2] Dir: [-0.01,0.941,0.05] Pos: [-0.011,0.013,0.001] **Child part**  **Parent part** 		

PhysX-3D: Physical-Grounded 3D Asset Generation

Ziang Cao[1], Zhaoxi Chen[1], Liang Pan[2], Ziwei Liu[1]

[1]Nanyang Technological University, [2]Shanghai AI Lab

Home page:

https://physx-3d.github.io/