

Harmonizing Stochasticity and Determinism:

Scene-responsive Diverse Human Motion Prediction

Paper ID: 4447

(Purple for observation), (Yellow for prediction)



Background

1. Scene-aware Motion Prediction

Most existing diverse human-motion-prediction (HMP) methods focus solely on the stochastic characteristics of human movement, neglecting the external 3D scenes, leading to scene inconsistency in real-world settings.

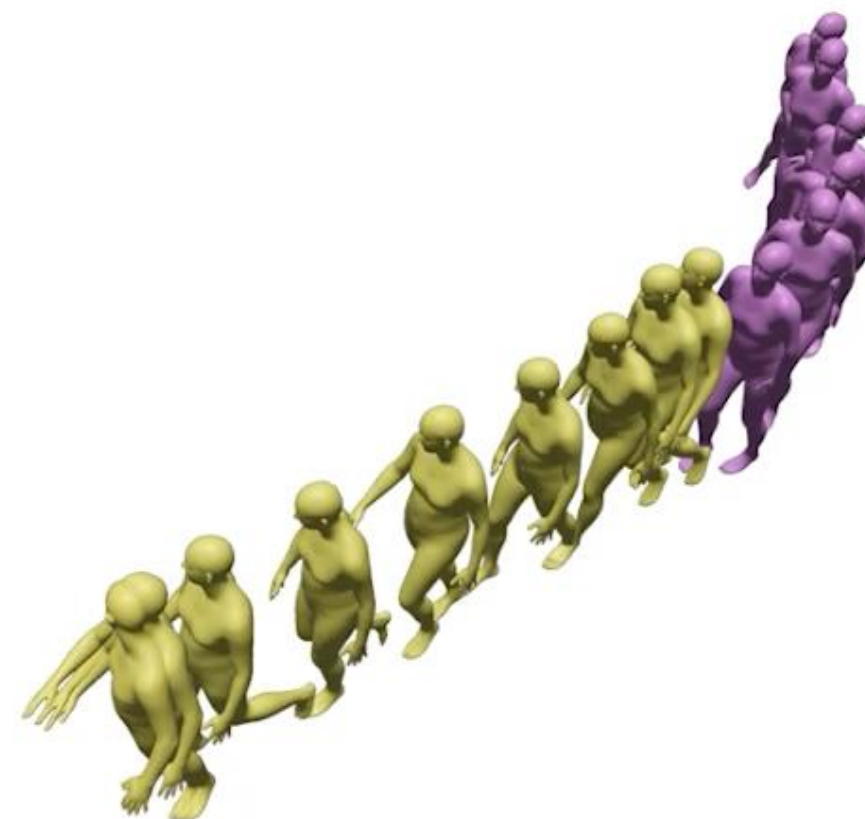
(Purple for observation), (Yellow for prediction)



sit down.



pick up things.



walk forward and maybe
open the door?

Background

1. Scene-aware Motion Prediction

Most existing diverse human-motion-prediction (HMP) methods focus solely on the stochastic characteristics of human movement, neglecting the external 3D scenes, leading to scene inconsistency in real-world settings.

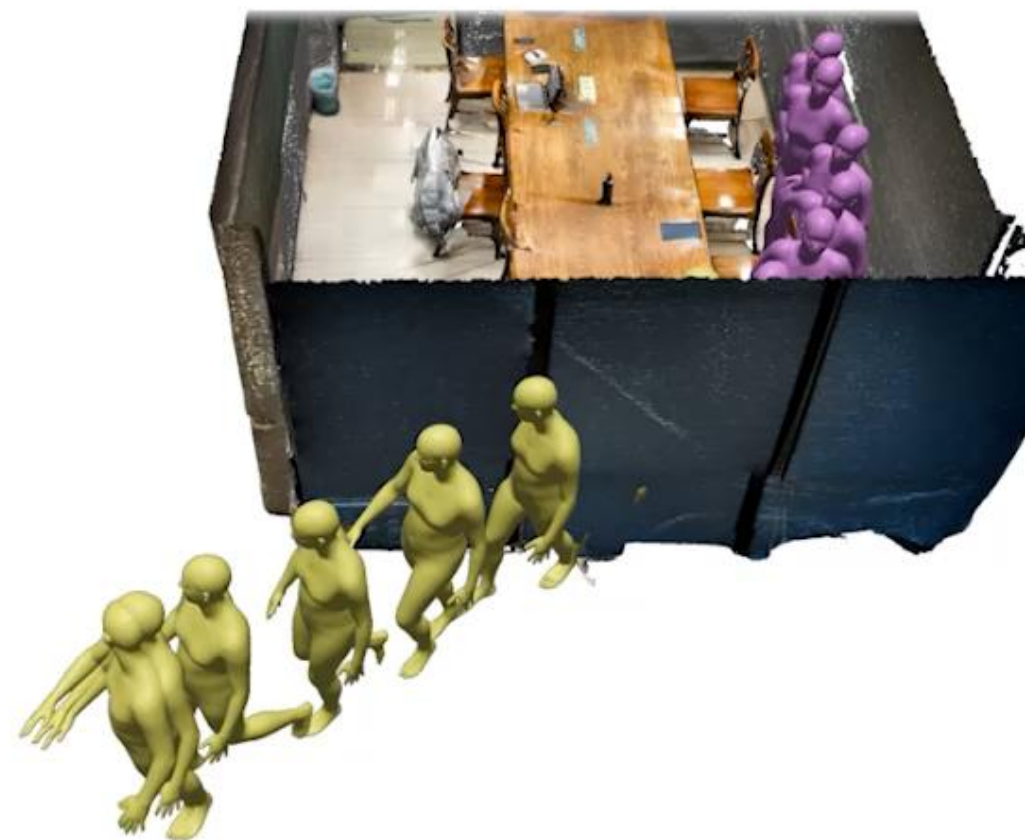
(Purple for observation), (Yellow for prediction)



there's no chair!



where did my things go?

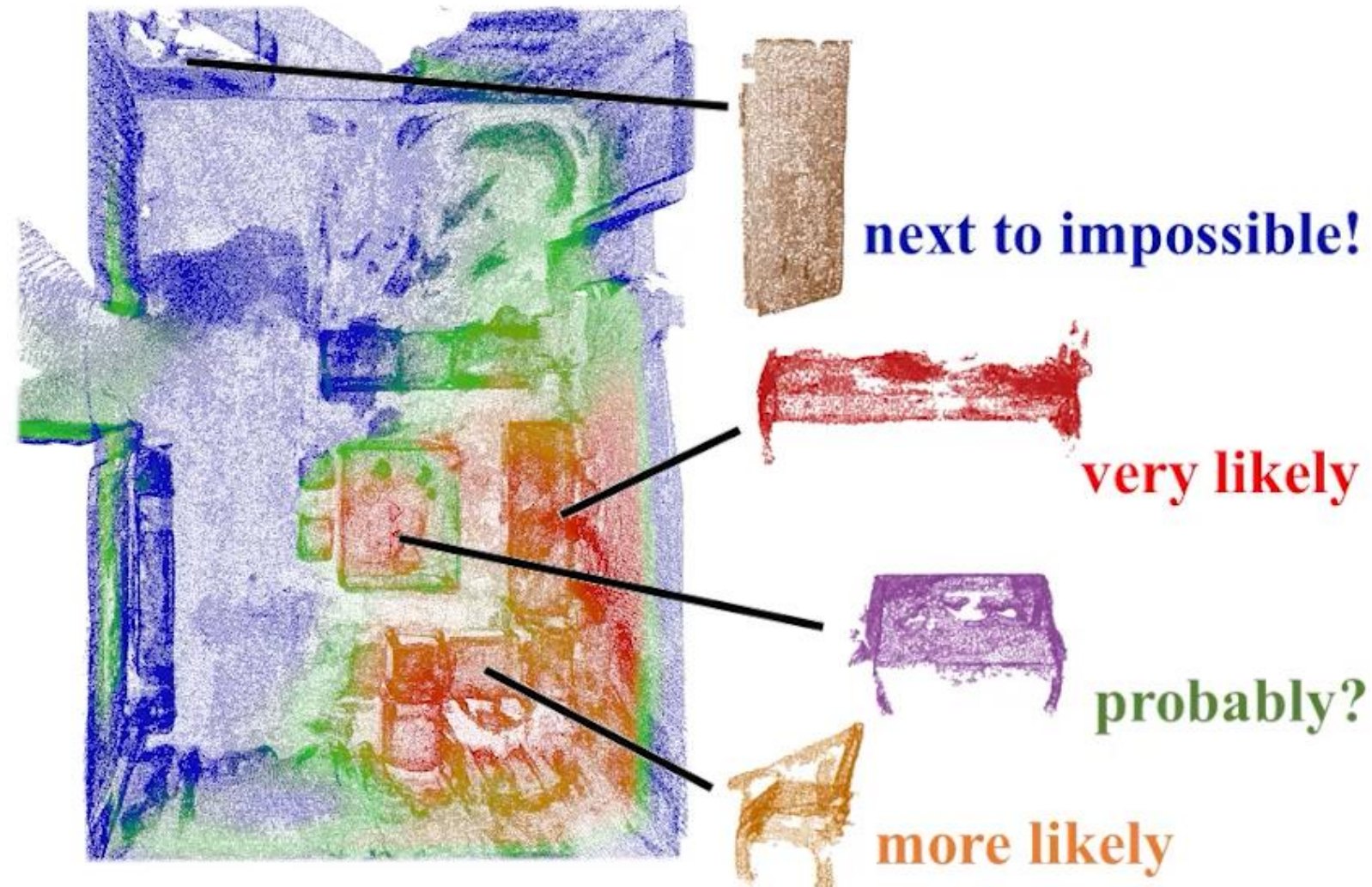


I've magically passed through the wall to exit the room.

Background

2. Cross-modal Motion-Scene Analysis

Unlike conventional scene analysis, scene-aware HMP requires comprehending the interplay between observed motion and 3D scenes. It's essential to discern which areas are viable for interaction based on current observation.



Background

3. Physical Consistency

The predicted sequences must also adhere to physical constraints from the 3D scene environment, including avoiding object penetration and ensuring interactions are consistent with typical human behaviors.



Penetration...



Obstacle free!

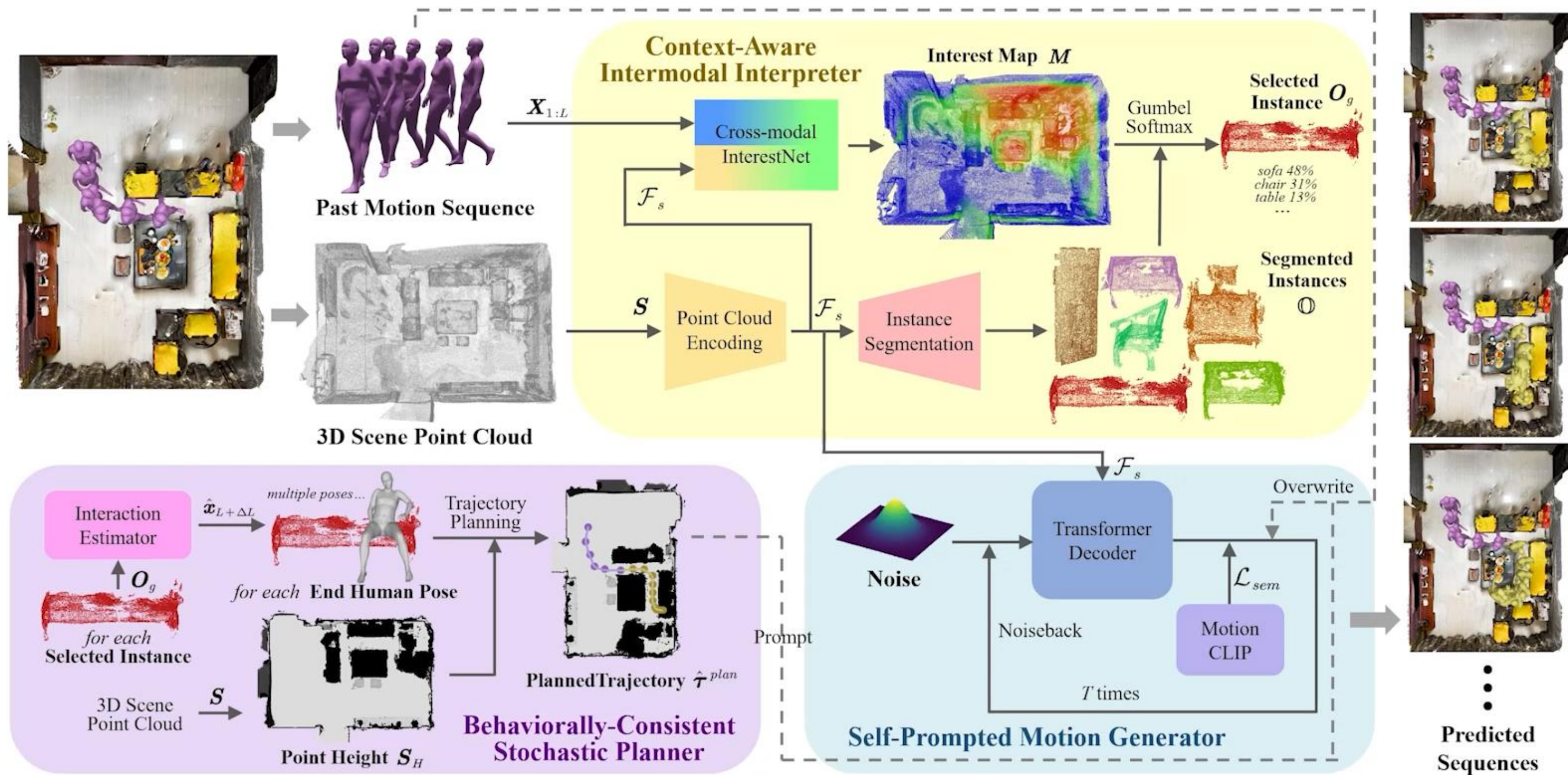


Walk through table?!



Pick things!

Method



Visualizations



BelFusion (ICCV23)
-SoTa baseline-

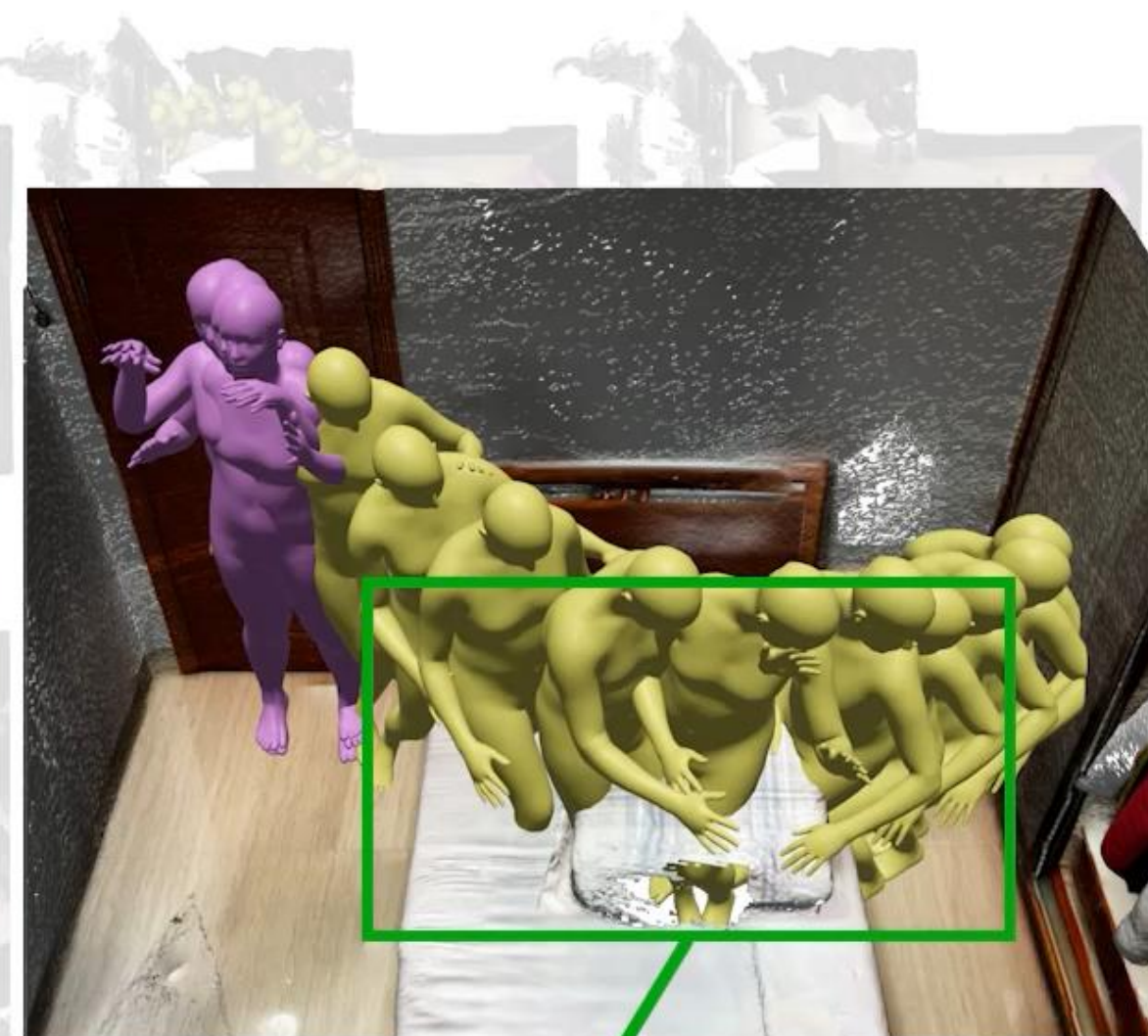
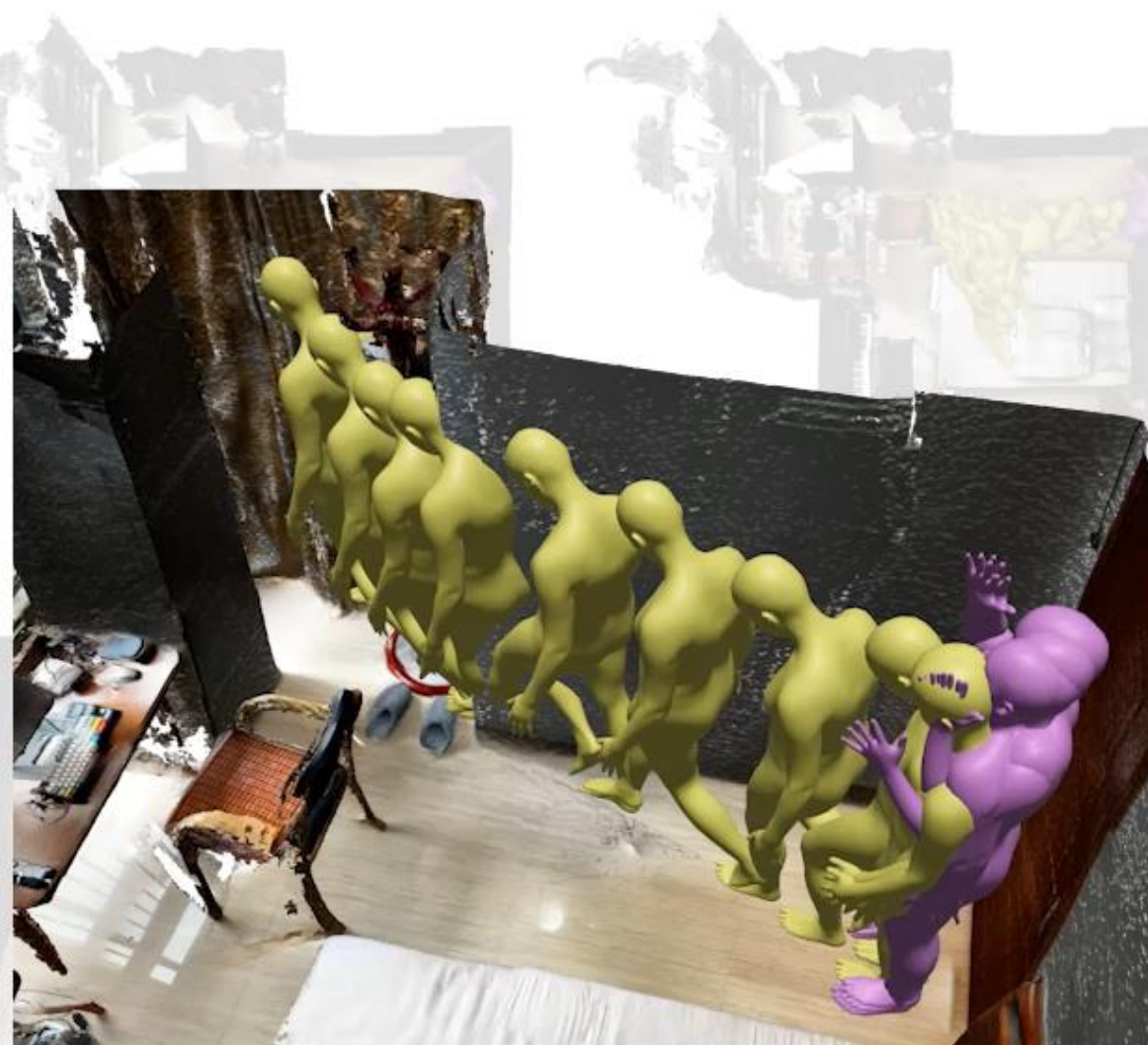


DiMoP3D (ours)

Visualizations



reach for nothing



pick from bed

Visualizations



BelFusion (ICCV23)
-SoTa baseline-



DiMoP3D (ours)

Visualizations



Visualizations



BelFusion (ICCV23)
-SoTa baseline-



DiMoP3D (ours)

Visualizations



Conclusion

- (1) We introduce a novel and challenging task, predicting **diverse** human motion within **real-world 3D scenes**, advancing the scope of the traditional diverse HMP from an idealized open-world context to one that is **more realistic and practical**.
- (2) A novel framework **DiMoP3D** is proposed to tackle this task, harmonizing the **deterministic constraints** of the scene with the stochastic prediction of human motion.
- (3) We show that **crossmodal motion-scene analysis** and **scene-aware motion planning** are essential in predicting high-fidelity diverse motion sequences consisting the 3D scene contexts.