Sitcom-Crafter: A Plot-Driven Human Motion Generation System in 3D Scenes

Jianqi Chen*, Panwen Hu*, Xiaojun Chang, Zhenwei Shi, Michael Kampffmeyer, Xiaodan Liang[+]

Code is available

Motivation and Goal

• Existing works primarily focus on specific type of motion.
• We introduce *Sitcom-Crafter*, a comprehensive system that synthesizes diverse human motions guided by long plot contexts.

Long Plot-Guided Human Motion Generation Results

Amy and Jack are enjoying a relaxing afternoon in a cozy living room. Amy stands by the window, lost in thought, while Jack, wandering near a cabinet, feels the urge to bring some energy into the space. He steps toward Amy and begins guiding a playful dance, bending and stretching his legs in a rhythmic, synchronized way. Amy, intrigued, follows his lead, and soon they are both moving in harmony. As they dance, Jack gets more creative, showing off some playful moves. Amy, amused, begins to imitate him, and soon they are fully immersed in the joy of their spontaneous dance-off.
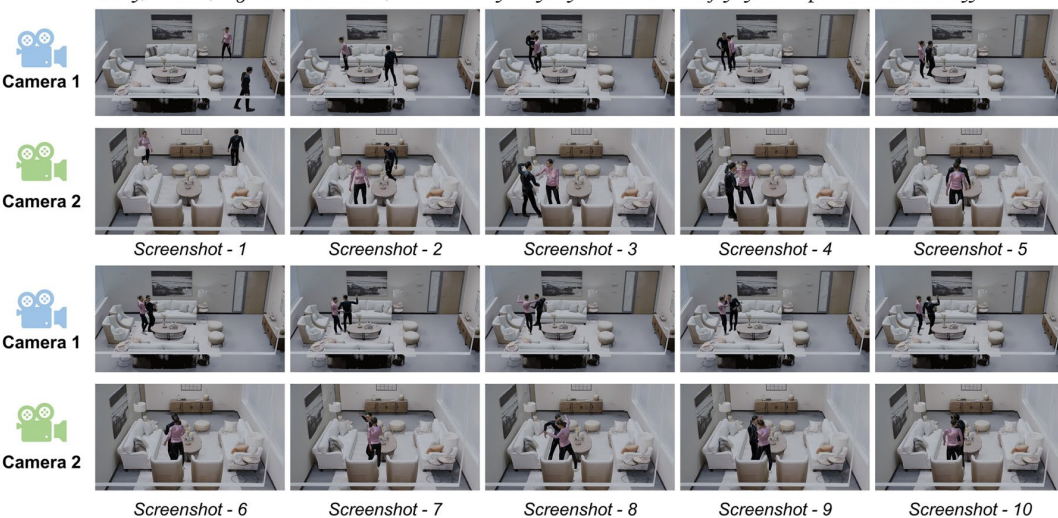

Figure note: The figure displays the final motion generation results of Sitcom-Crafter, guided by long plot context. We display motions under 2-viewpoint for better view.
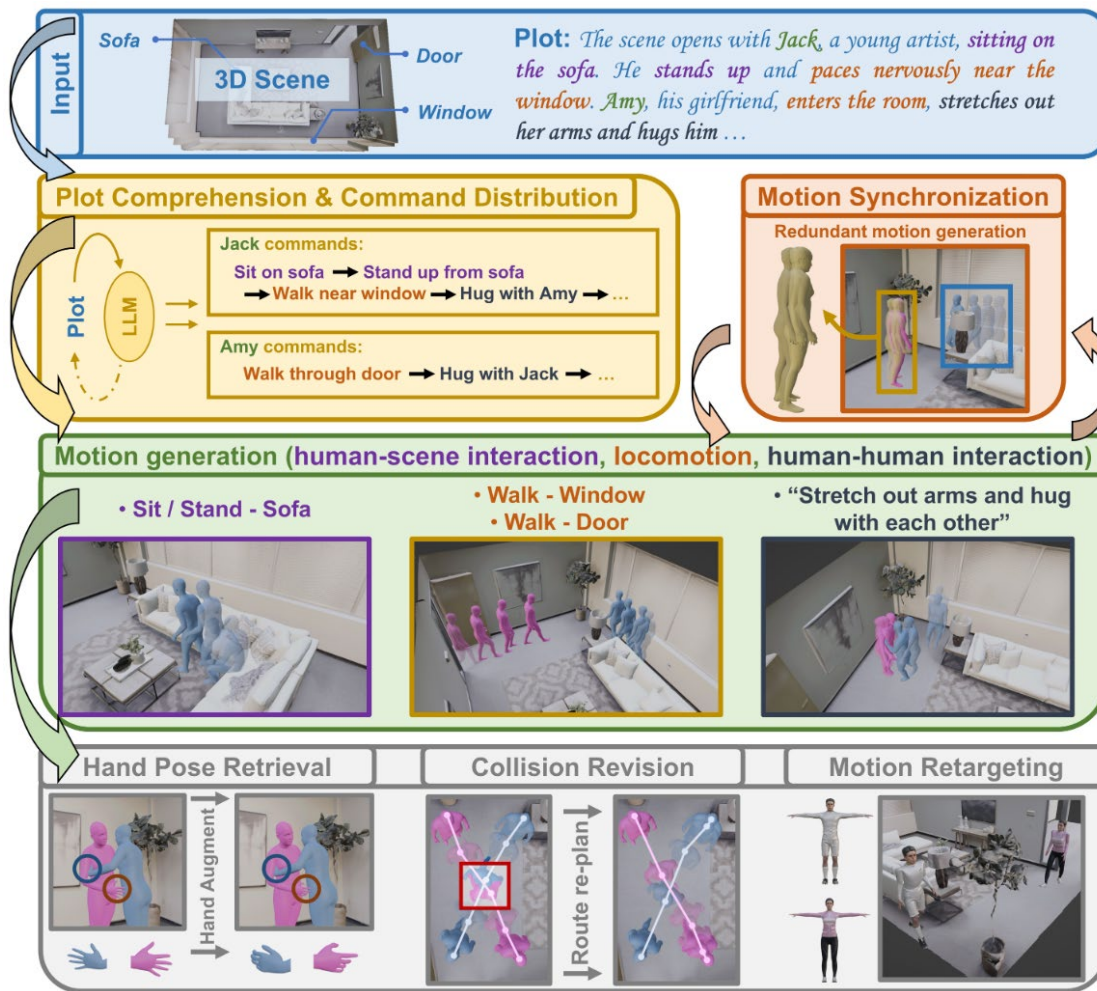
System Workflow

Input:
A 3D scene with information of inside objects, and a plot context.

Process:
• Three fundamental **Generation** modules for diverse types of human motion generation.
• Five peripheral **Augmentation** modules for better motion cohesiveness, quality, and user-friendliness of the system.

Plot: *The scene opens with Jack, a young artist, sitting on the sofa. He stands up and paces nervously near the window. Amy, his girlfriend, enters the room, stretches out her arms and hugs him …*

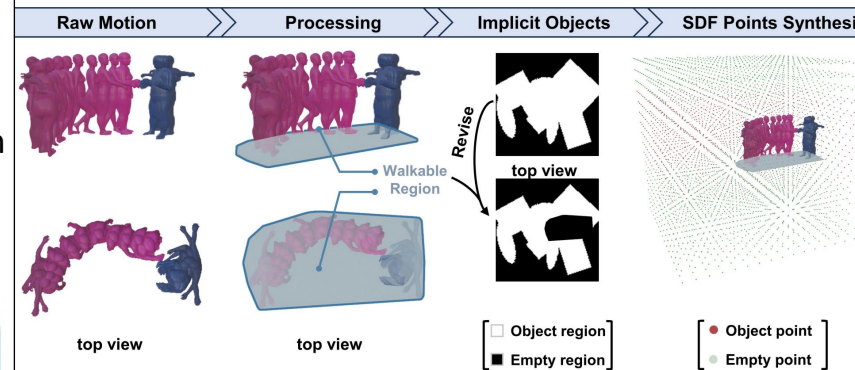## Scene-Aware Human-Human Interaction

Figure note: This figure displays how we synthesize surrounded 3D scene information based on existing human-human interaction motion data.

a) Evaluation on Replica Dataset

Method	FS↓	FP↓	HSP↓	HHP↓
Baseline	0	0.0099	5.5119	0.1991
ComMDM	0.0001	0.0215	10.5532	0.2712
InterGen	0.0001	0.0242	9.6035	0.1774
Ours	**0**	**0.0189**	**5.7529**	**0.1687**

b) Evaluation on InterHuman Dataset

Method	FS↓	FP↓	HSP↓	HHP↓
Real	0	0.0037	0.0043	0.0807
ComMDM	0	0.0076	6.2802	0.1336
InterGen	0	0.0049	6.6408	0.0989
Ours	0	**0.0047**	**1.6950**	**0.0742**

Table note: Comparisons with existing methods on the physical compliance.

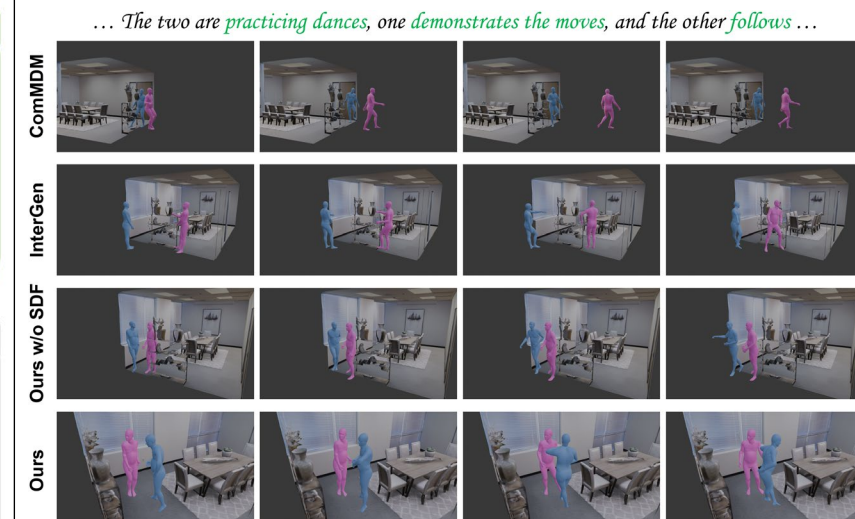*… The two are practicing dances, one demonstrates the moves, and the other follows …*

Figure note: The sentence above is an extracted partial context from the plot.