InstructScene: Instruction-Driven 3D Indoor Scene Synthesis with Semantic Graph Prior

Chenguo Lin, Yadong Mu

Introduction

User Instructions

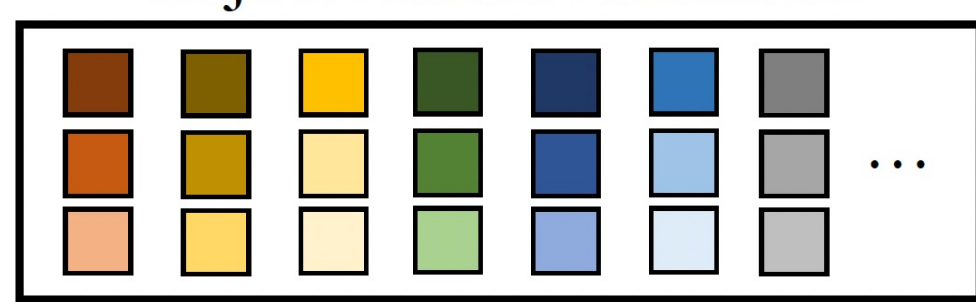Position a grey nightstand to the *left* of a wooden wardrobe with doors. Hang a white ceiling lamp with a wooden circle directly *above* a black double bed.

Design

Semantic Graph Prior

front, front, above, left of, left of, left of, right of, right of, right of, left of
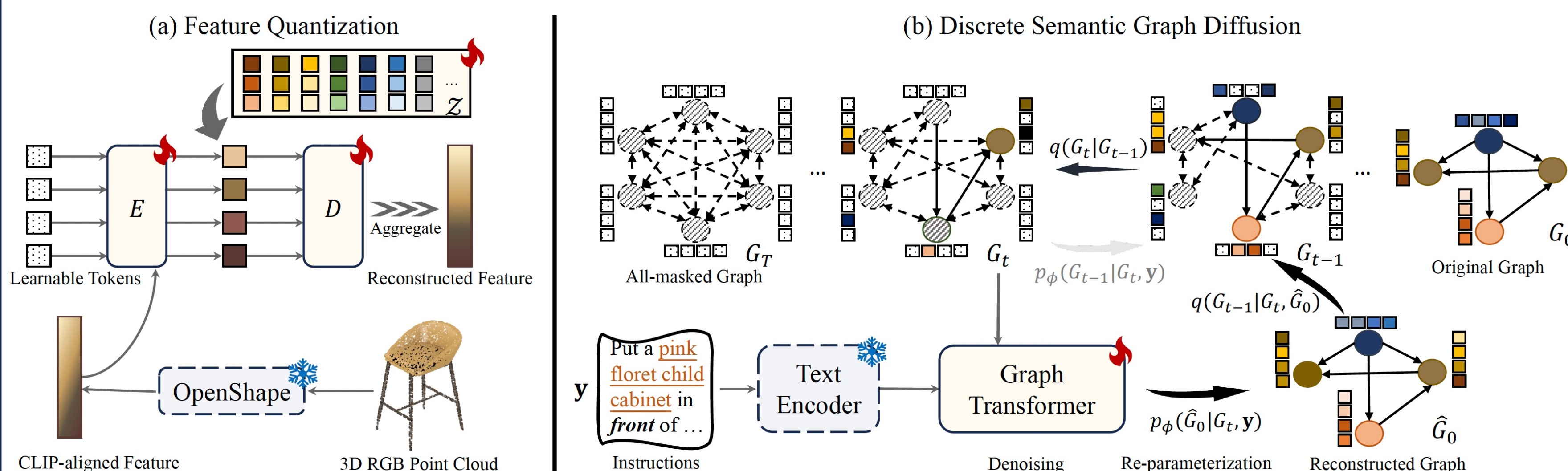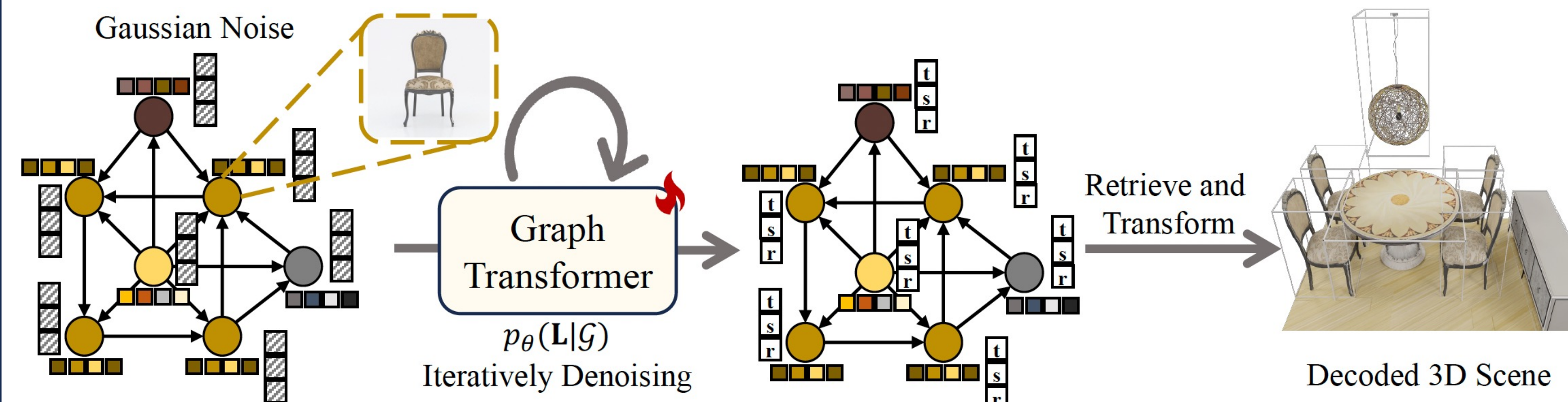
Embody

Decoded 3D Scene
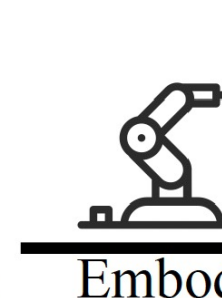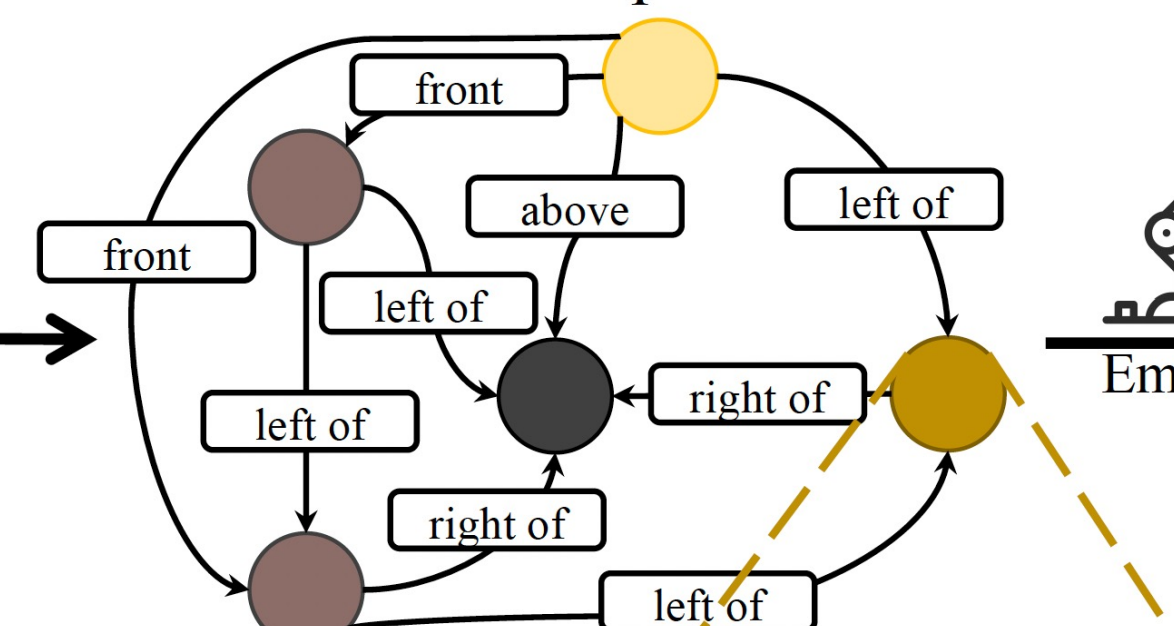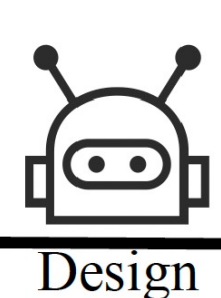
Object Feature Codebook

- **InstructScene** is a generative framework to **synthesize 3D indoor scenes from instructions**.

- It is composed of (1) **a semantic graph prior** and (2) **a layout decoder** to *improve the controllability and fidelity*.

- In the graph, each node is an object with *semantic features*; each edge represents a *spatial relationship* between objects.

- It proceeds to place objects in a scene by decoding precise *7 degrees-of-freedom* attributes.

More details and demos: please scan QR codes above ⬆️

Method Overview

- **Semantic Graph Prior**: Object semantic features are first **quantized**; A graph Transformer is then trained to learn the semantic graph prior by **mask-based discrete diffusion**.

(a) Feature Quantization

Learnable Tokens — E — D — Aggregate — Reconstructed Feature

z

CLIP-aligned Feature — OpenShape — 3D RGB Point Cloud

(b) Discrete Semantic Graph Diffusion

All-masked Graph G_T ... G_t $q(G_t|G_{t-1})$ G_{t-1} ... Original Graph G_0

$p_\phi(G_{t-1}|G_t, \mathbf{y})$ $q(G_{t-1}|G_t, \hat{G}_0)$

y Put a pink floret child cabinet in *front* of ... Instructions — Text Encoder — Graph Transformer — Denoising — $p_\phi(\hat{G}_0|G_t, \mathbf{y})$ — Re-parameterization — $\hat{G}_0$ Reconstructed Graph

- **Layout Decoder**: **Gaussian noises** are attached at each node, and are **iteratively denoised** by layout decoder to generate 7 DoF layout configurations: location, scale and yaw rotation.

Gaussian Noise — Graph Transformer — $p_\theta(\mathbf{L}|\mathcal{G})$ Iteratively Denoising — Retrieve and Transform — Decoded 3D Scene