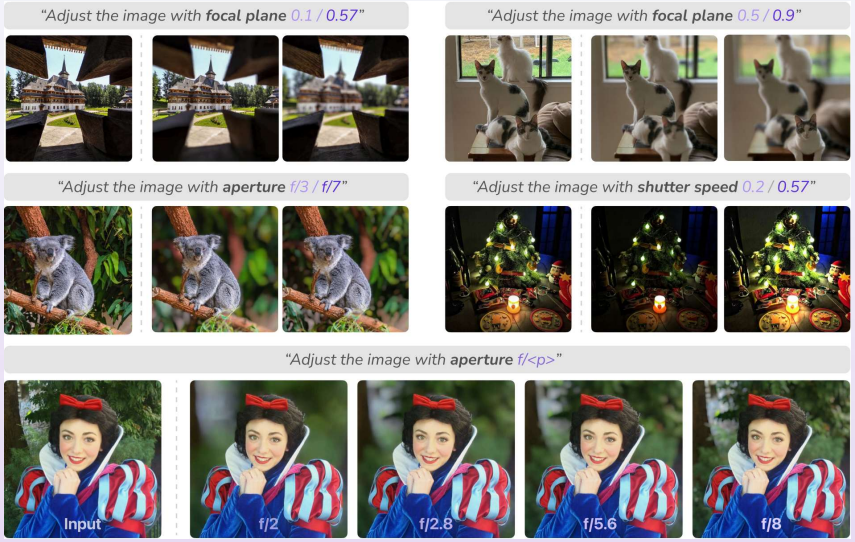
CamEdit: Continuous Camera Parameter Control for Photorealistic Image Editing

Xinran Qin* , Zhixin Wang* , Fan Li, HaoYu Chen, RenJing Pei, WenBo Li, XiaoChun Cao†

Introduction

Existing editing models focus on semantic/stylistic editing, failing to enable precise control over camera parameters.

Challenge: Existing text-driven diffusion model relies on discrete tokens, failing to handle *continuous camera parameters* or maintain *physical consistency* in effects.



CamEdit Enable **continuous, fine-grained, and photorealistic** control over **camera parameters**, such as Aperture, Focal Plane, and Shutter Speed.

Contribution

Existing Editing Methods

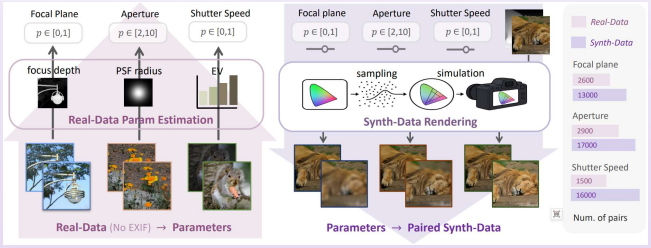
- ☹ Discrete textual tokens limit control precision
- ☹ Fail to capture parameter-specific effects
- ☹ Lack of paired data with camera annotations

CamEdit (Ours)

- ☺ Continuous Parameter Prompting
- ☺ Parameter-Aware Modulation (PAM)
- ☺ CamEdit50K: Real & Synthetic Data

CamEdit, a diffusion-based framework for **photorealistic image editing** that enables continuous and fine-grained control over intrinsic camera parameters

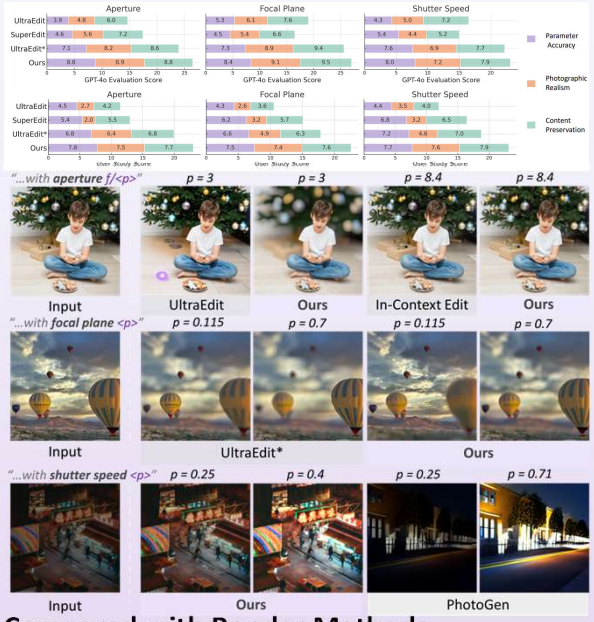
CamEdit50K



Unifying real and synthetic imagery with **explicit ground-truth parameters** for focal plane, aperture, and shutter speed.

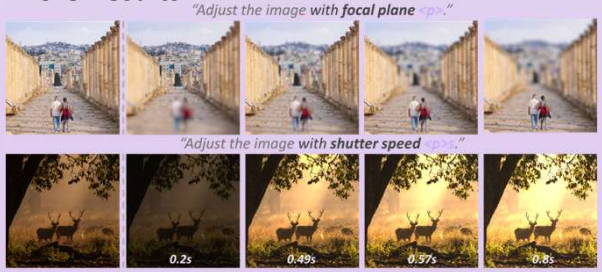
Experiments

Compared with Editing Methods



Compared with Render Methods

More Results

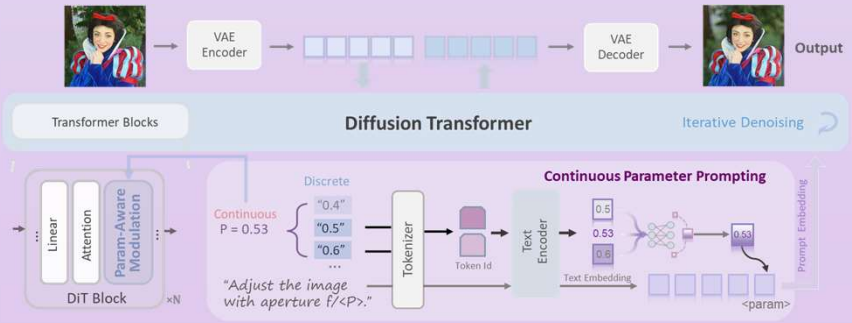


Compared to existing models that lack precise control, our method achieves continuous manipulation across parameters.

Proposed Method

Framework

CamEdit adopts the instruction-driven editing paradigm of IP2P, and equips SD3 with precise continuous control over intrinsic camera parameters.



Continuous Parameter Prompting

By interpolating anchors, we map continuous values into the text space while preserving semantic alignment.

$$e_p = \text{Linear}([e_i, e_{i+1}]) + \text{MLP}(\phi(p))$$

Parameter-Aware Modulation

Spatial: Displace for geometric/depth distortion.

$$\Delta G = \text{MLP}_g(\text{AvgPool}(F), p) \in R^{2\times H\times W}$$

$$F_g = \text{grid\_sample}(F, G_{base} + \Delta G)$$

Channel-wise: Scales for global exposure control.

$$F_p = \gamma(p) \cdot F_g + \beta(p)$$

$$\gamma(p), \beta(p) = \text{MLP}_p(\text{AvgPool}(F), p)$$