

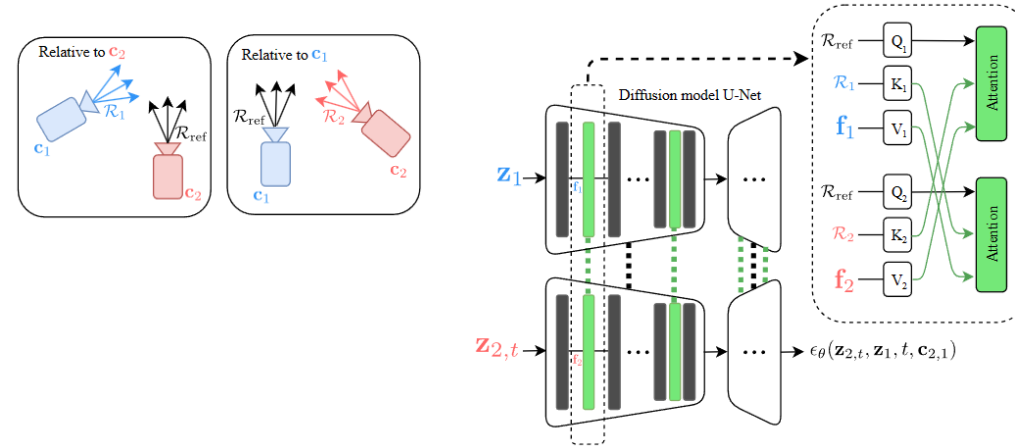
Novel View Synthesis

Ruofeng Yang

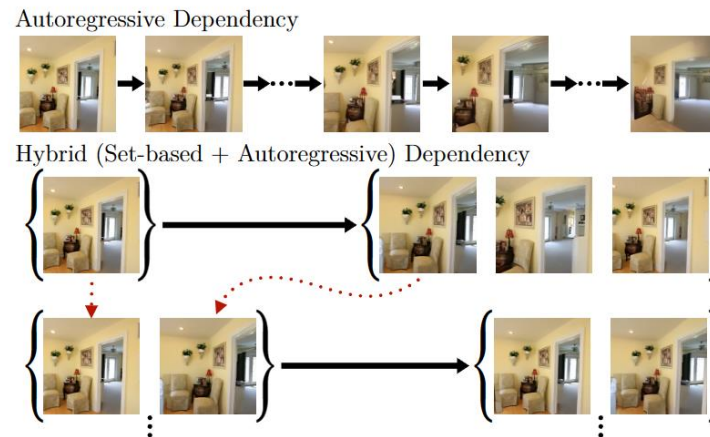
The paradigm of NVS

(A)

1. Latent Diffusion to obtain the latent feature
2. Share feature layer (Reference and target image) and add (relative) camera pose of reference images by using cross-attention layer. [1]



(B) Set to Set (Different frames) [2]



[1] Long-Term Photometric Consistent Novel View Synthesis with Diffusion Models. ICCV 2023.

[2] PolyOculus: Simultaneous Multi-view Image-based Novel View Synthesis. Arxiv 20240228.

The thoughts of NVS

1. Understanding of context information: diffusion transformer
2. Geometric Consistency from different angles:
 - (a) Using some consistency metric such as Thresholded Symmetric Epipolar Distance (TSED) [1] as reward and using RL method such as PPO to fine tuning the pretrained diffusion models [3][4].
 - (b) With some theoretical insight (Diffusion perspective [5][6]; RL perspective [7]).
3. The choice of forward process:

The variance exploding forward process (VESDE) focus on the Negative Log-Likelihood (NLL) metric and some NVS works used VESDE [1][8]

[3] Training Diffusion Models with Reinforcement Learning. ICML 2023 workshop. 20230522

[4] Feedback Efficient Online Fine-Tuning of Diffusion Models. ICML 2023 workshop. 20230522

[5] Theoretical Insights for Diffusion Guidance: A Case Study for Gaussian Mixture Models. 20240304

[6] Reward-directed conditional diffusion: Provable distribution estimation and reward improvement. Neurips 2024

[7] Optimistic policy optimization with bandit feedback. ICML 2020

[8] Generative Novel View Synthesis with 3D-Aware Diffusion Models. ICCV 2023