3D StreetUnveiler with Semantic-aware 2DGS - a simple baseline

Jingwei Xu[1] · Yikai Wang[2] · Yiqun Zhao[3,6] · Yanwei Fu[5] · Shenghua Gao[3,4]

[1] ShanghaiTech University [2] Nanyang Technological University [3] The University of Hong Kong

[4] HKU Shanghai Intelligent Computing Research Center [5] Fudan University [6] Transcengram

(ICLR 2025)

Goal of StreetUnveiler

Remove the cars from the video captured by in-car cameras

Remove from both visual and geometry aspect

Input

3D StreetUnveiler

Clean Street Results

Empty Street Mesh Extraction

Existing Works

[Mirzaei et.al 2023]

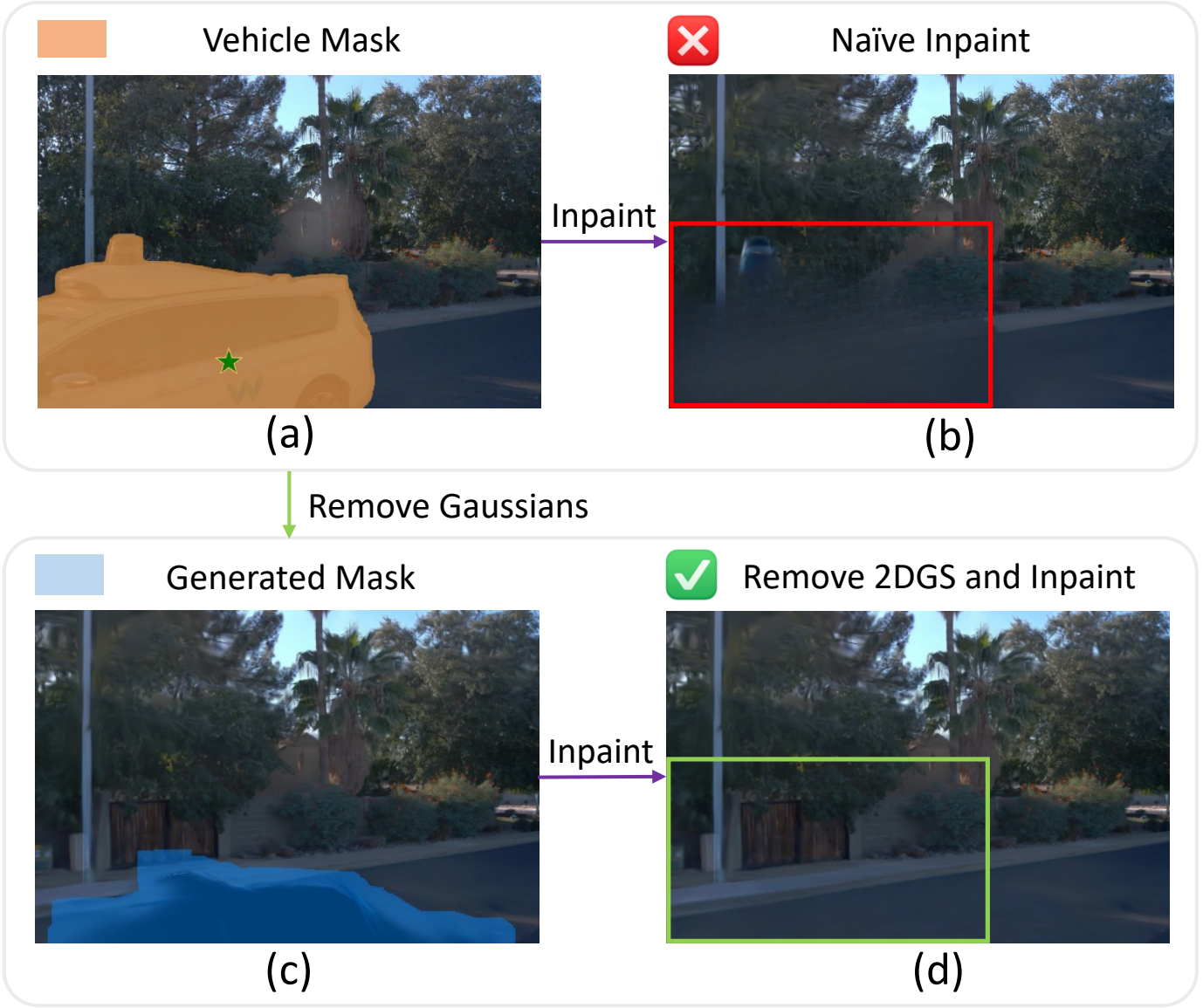
Small scale scenes
equire ground-truth object masks

Unveiling Challenge

- Hard to get a reasonable inpainting mask
- Long trajectory / maintain consistency

Mask Challenge

- How to obtain the blue mask?

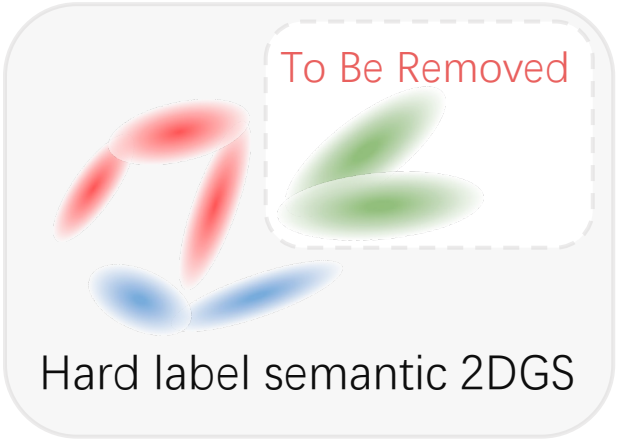


Long Trajectory Consistency Challenge

Inpaint each image seperately

3D Reconstruction

Solution to mask challenge



Render

Get the mask through alpha value

Some more training objectives:

$$\mathcal{L}_{\text{ds}} = \sum_k \mathcal{L}_{\text{d}}^k \qquad \mathcal{L}_\alpha = \frac{1}{N}\sum_p \alpha_p$$

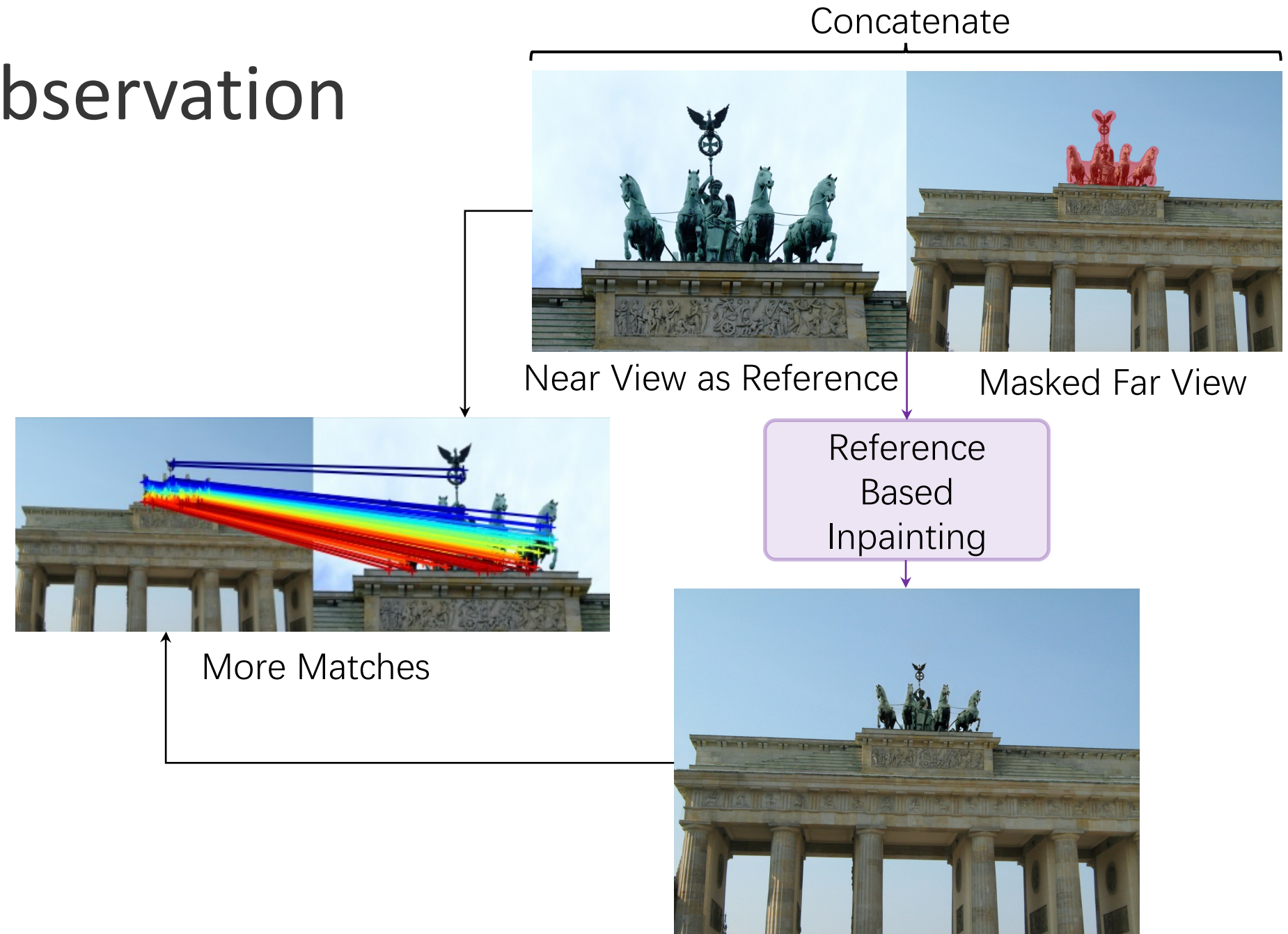
Motivation: Nature of the captured video

Frame T_n

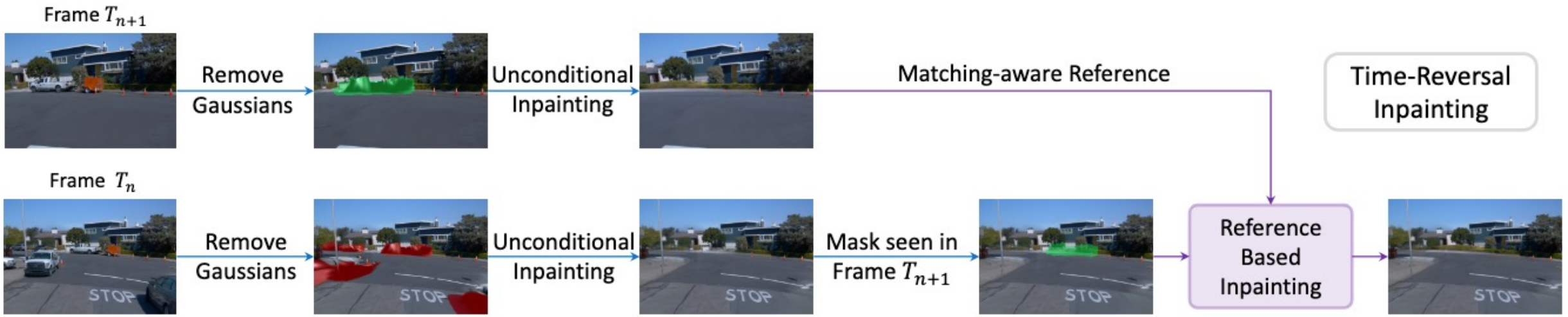
Frame T_{n+1}

Removed object is more close in T_{n+1}

Observation

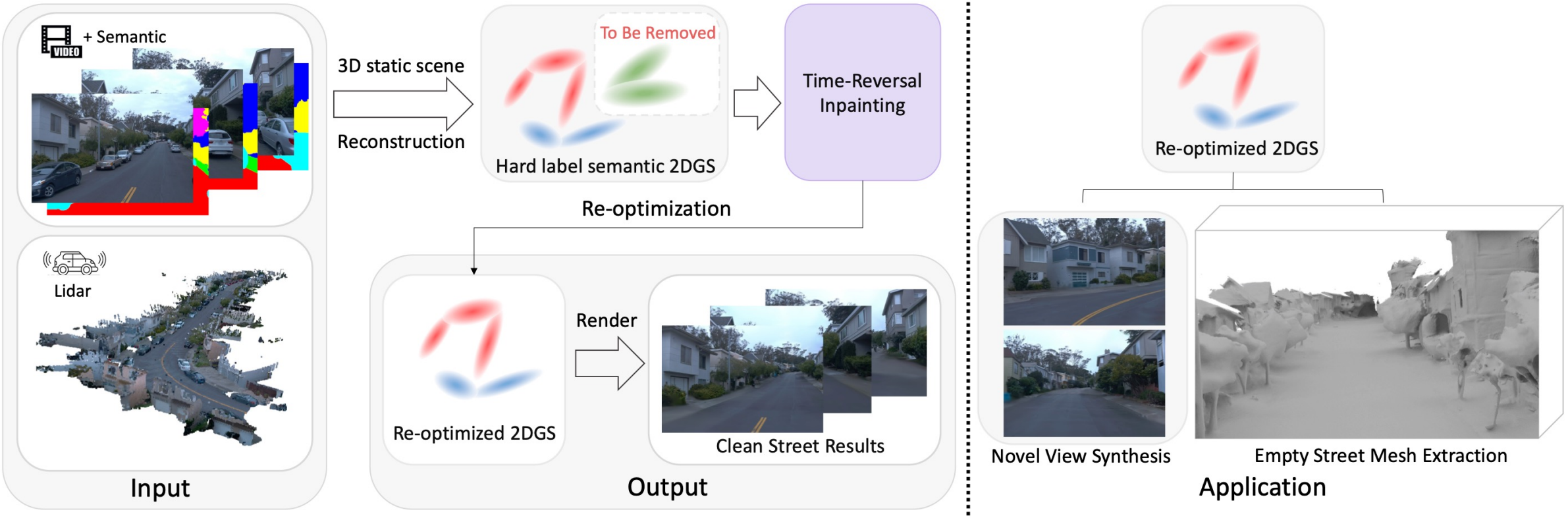


Time Reversal Inpainting



Inpaint T_n with T_{n+1} as reference

Overall Pipeline



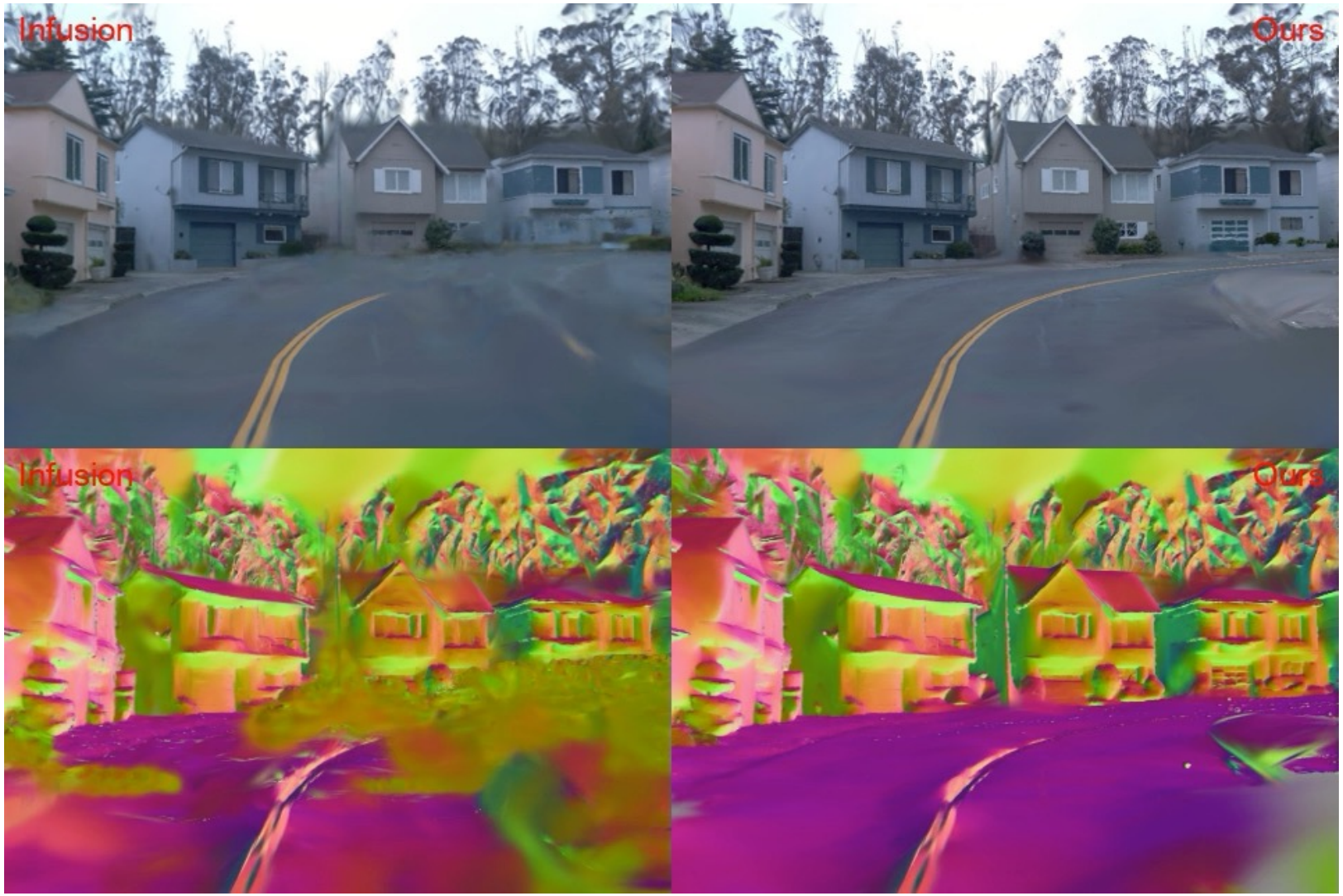
Qualitative Results (with 2D Inpaint)

SDXL [Podell et.al 2023]

Qualitative Results (with Video Inpaint)

ProPainter [Zhou et.al 2023]

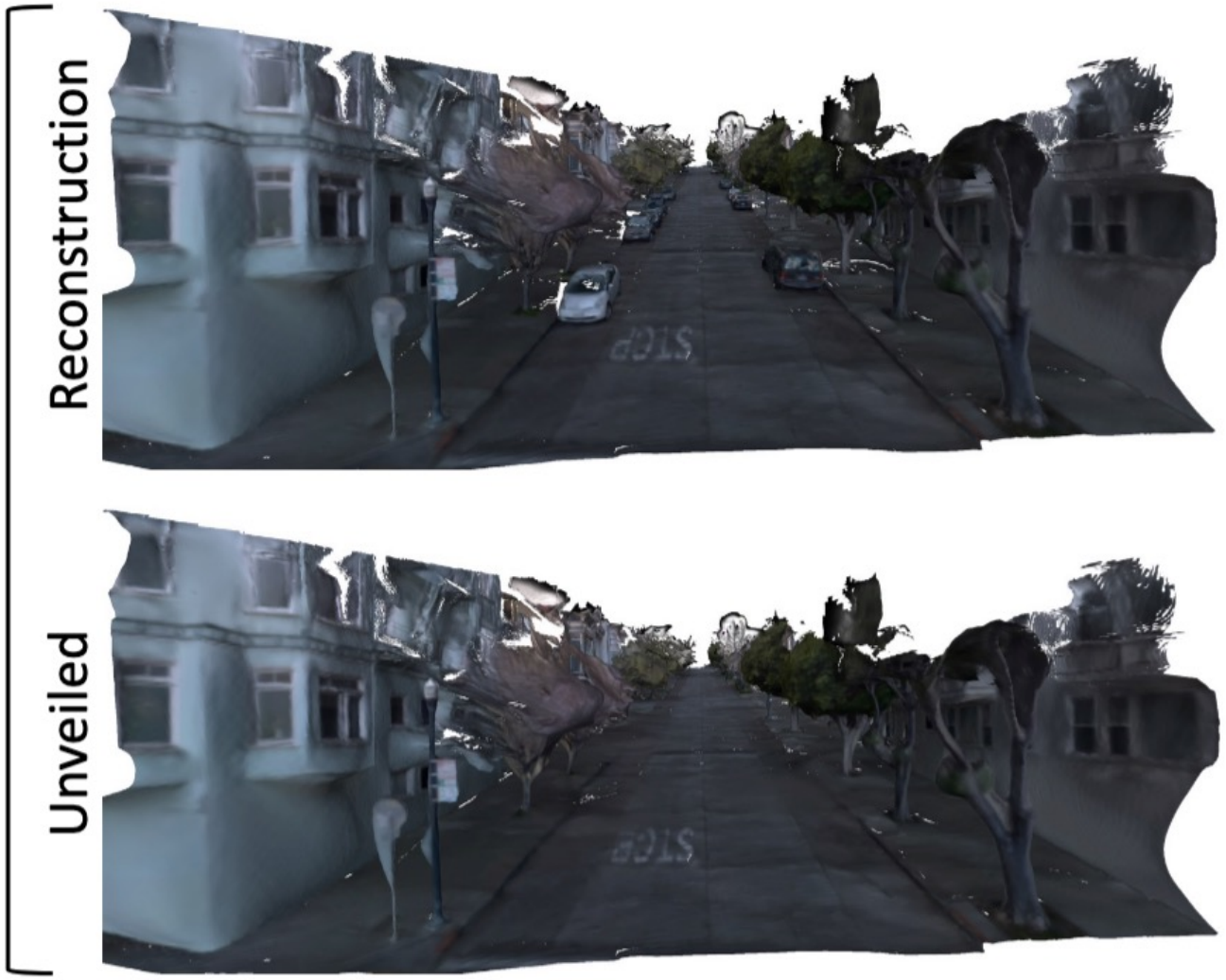
Qualitative Results (with 3D Inpaint)



InFusion [Liu et.al 2024]

Novel View Synthesis Results

Empty Street Mesh Extraction



Thanks for watching!

Project Page: https://streetunveiler.github.io/
Code: https://github.com/DavidXu-JJ/StreetUnveiler