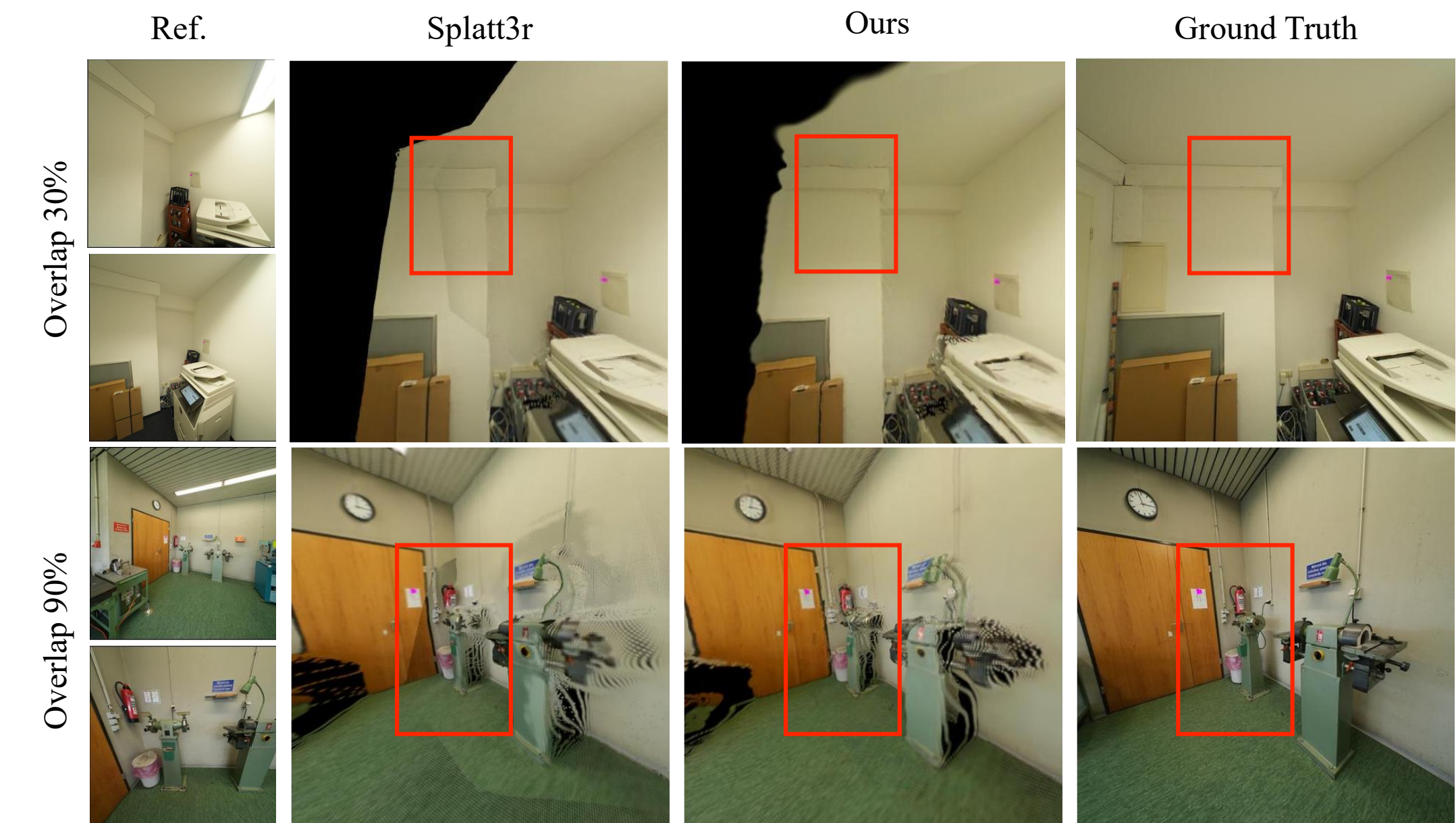


Introduction

CSG-Fusion excels on rendering quality with significantly fewer Gaussian primitives.

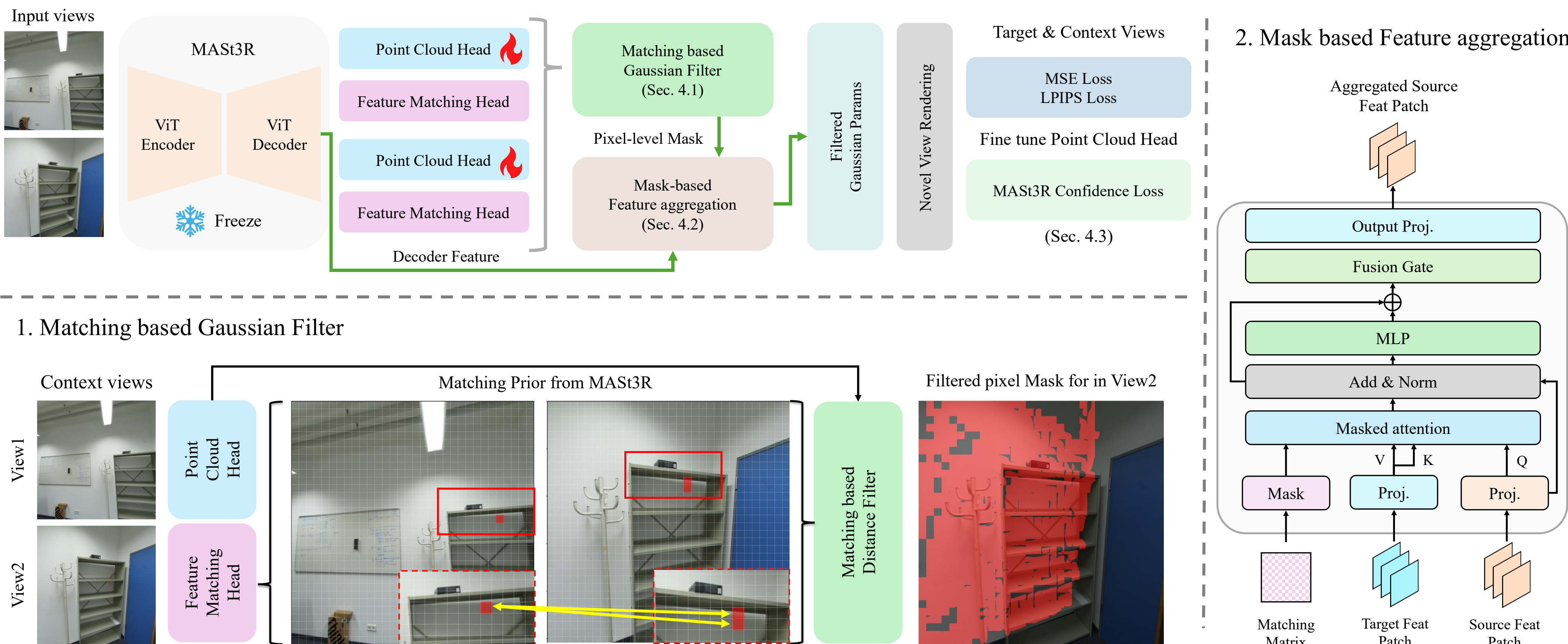


Goal: Integrate pixel-aligned pointmap to reduce redundant primitives and produce compact and consistent 3D structures.

Key design:

- Matching-based Gaussian Filter module prunes redundant Gaussians in pixel-aligned reconstruction pipelines with dense matching priors and spatial distance threshold.
- Mask-based feature aggregation module to enable consistent multi-view fusion by aggregating local features from matched Gaussian primitive groups.
- Incorporate a context-view supervision signal and fine-tune point cloud prediction to recover spatial continuity after pruning, resulting in smoother surfaces and more complete geometry.

Pipeline

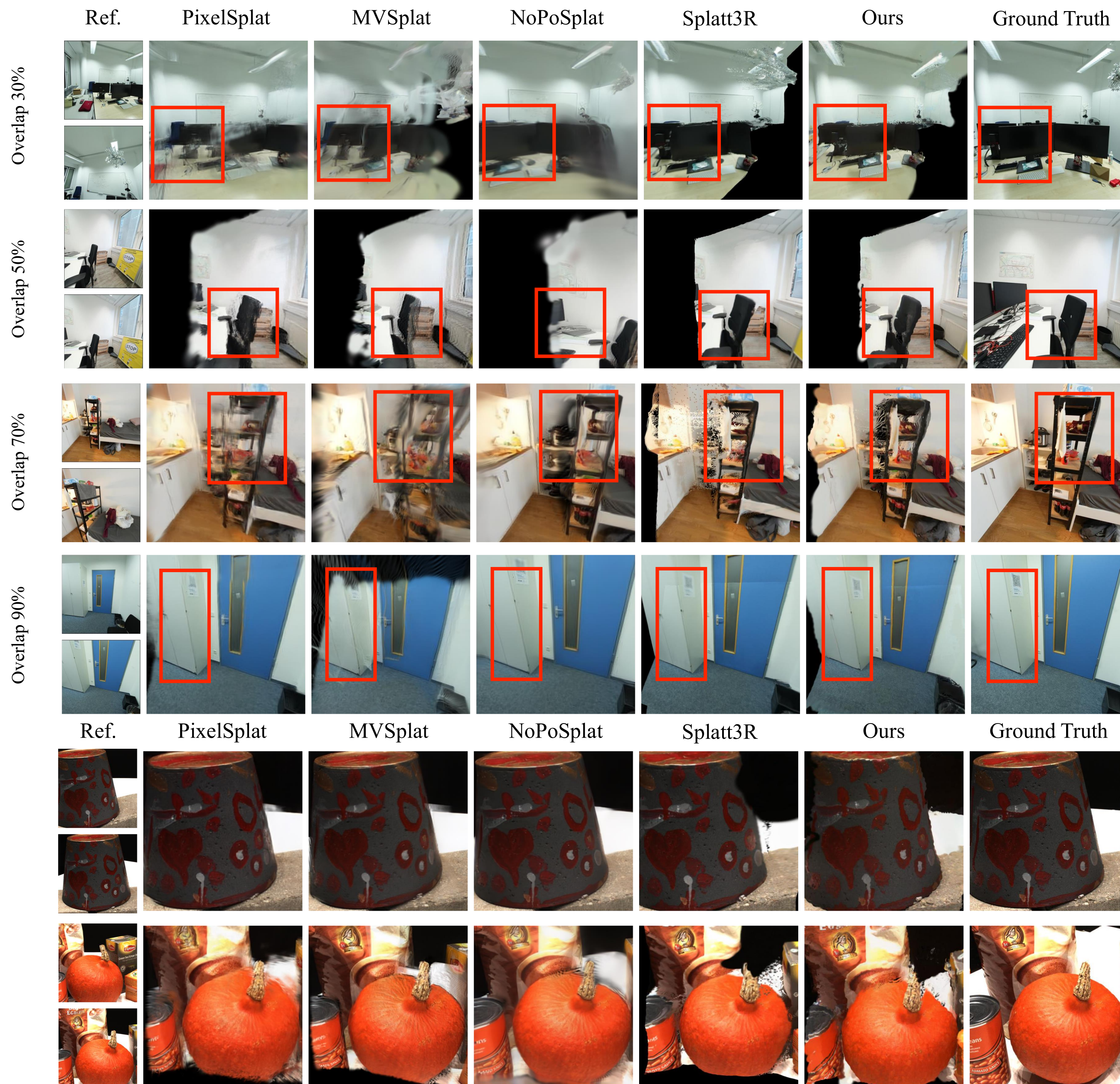


Quantitative Comparison

Method	Close 90%			Medium 70%			Wide 50%			Very Wide 30%		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
PixelSplat (GT cams) [3]	25.623	0.856	0.184	23.379	0.827	0.202	22.021	0.816	0.208	20.819	0.818	0.202
MVSplat (GT cams) [7]	16.053	0.582	0.415	14.285	0.527	0.430	13.530	0.536	0.417	13.159	0.574	0.387
Ours	19.277	0.730	0.186	19.249	0.750	0.171	19.193	0.770	0.166	18.815	0.790	0.172
Splatt3R [23]	16.479	<u>0.679</u>	<u>0.231</u>	<u>17.302</u>	<u>0.718</u>	<u>0.194</u>	<u>17.763</u>	<u>0.746</u>	<u>0.183</u>	<u>17.746</u>	<u>0.772</u>	<u>0.186</u>
NoPoSplat [31]*	<u>17.571</u>	0.653	0.343	15.401	0.610	0.373	14.616	0.621	0.368	14.669	0.667	0.331

To ensure fair comparison, all input images are centrally cropped to 256×256 . During evaluation, all model outputs are downsampled to 256×256 for metric computation.
 *: Note that NoPoSplat is trained on RealEstate10K datasets.

Qualitative Comparison



Storage & Computation Efficiency

