ErpGS

Equirectangular Image Rendering Enhanced with 3D Gaussian Regularization

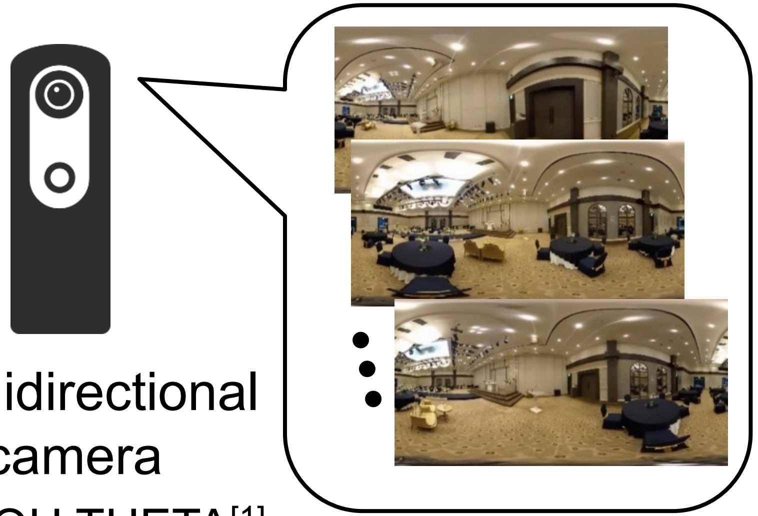
○ Shintaro Ito[1], Natsuki Takama[1] Koichi Ito[1],
Hwann-Tzong Chen[2], and Takafumi Aoki[1]

1 Graduate School of Information Sciences, Tohoku University, Japan

2 National Tsing Hua University, Taiwan

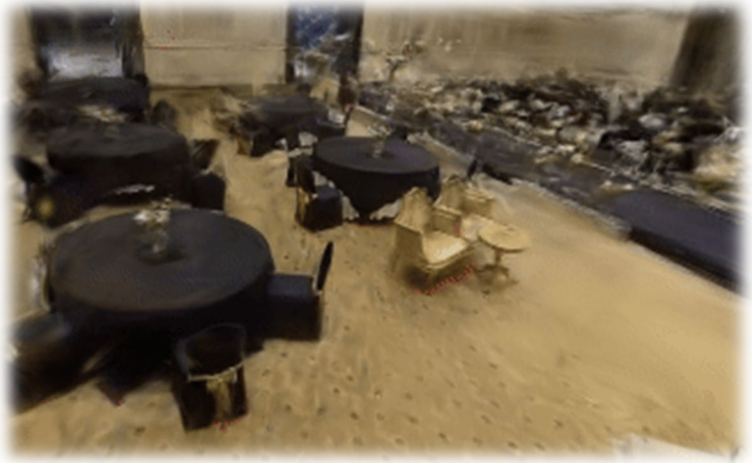
IEEE ICIP 2025

Novel View Synthesis via 3DGS[3] from ERP images



Omnidirectional camera
RICOH THETA[1]
Insta360 X5[2]

ERP images



3D scene representation by 3D Gaussian Splatting[3]

Rendered novel view image

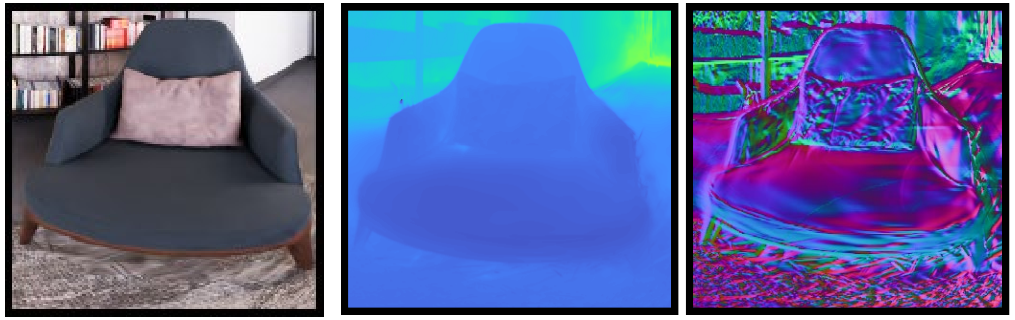
- **Novel View Synthesis (NVS)** is a promising technique with diverse applications, including VR/AR, robotics and entertainment products
- Omnidirectional camera can capture a wide field of view in a single shot with Equirectangular projection (ERP) image
- The images often contain distortion, which leads undesired artifacts in NVS

[1] https://www.ricoh360.com/ja/theta/
[2] https://www.insta360.com/jp/product/insta360-x5
[3] B. Kerbl et al., "3D Gaussian splatting for real time radiance field rendering," SIGGRAPH, July 2023.

Challenges in NVS Using ERP Images

Previous methods often show a decline in NVS accuracy due to the camera distortions and a lack of geometric consistency

3DGS's 3D scene representation



RGB Depth Normal

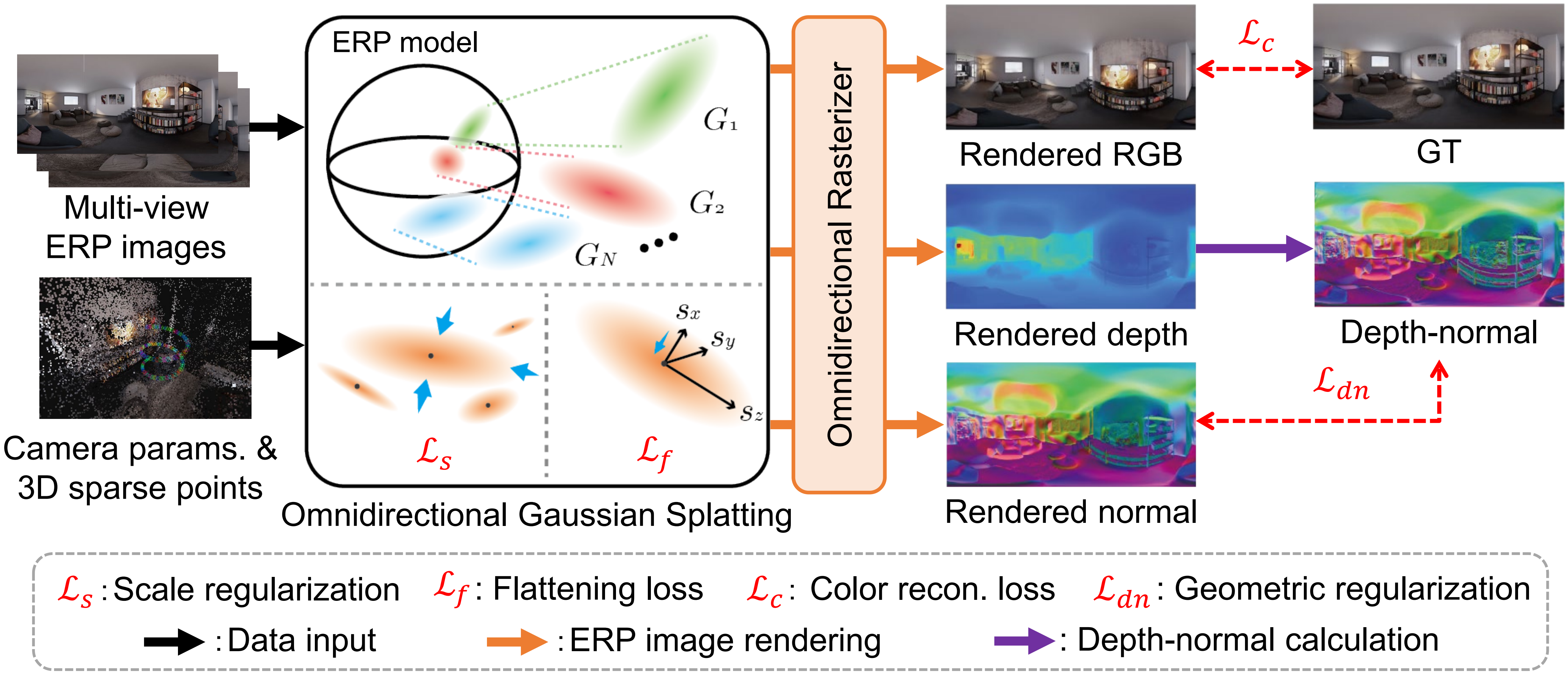
Noise caused by strong camera distortion

A lack of geometric consistency across different rendered domains

Reduce the effects with a **scale regularization**

Maintain the geometric consistency with a **geometric regularization**

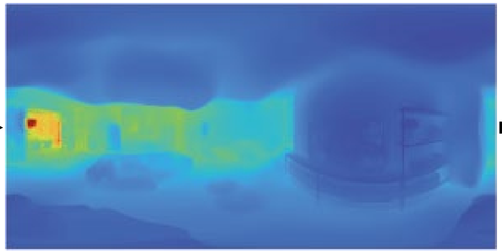
Proposed method: ErpGS



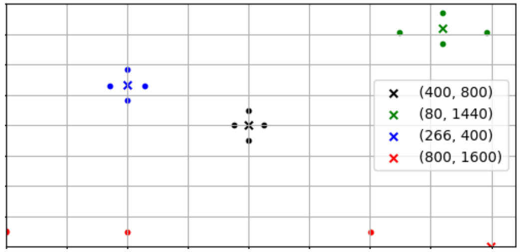
- ErpGS takes a set of omnidirectional images as input, then optimizes omnidirectional Gaussian splatting to render novel views in high accuracy

Depth-Normal Calculation

Omnidirectional Rasterizer



Rendered depth $\mathcal{D}(\boldsymbol{p})$



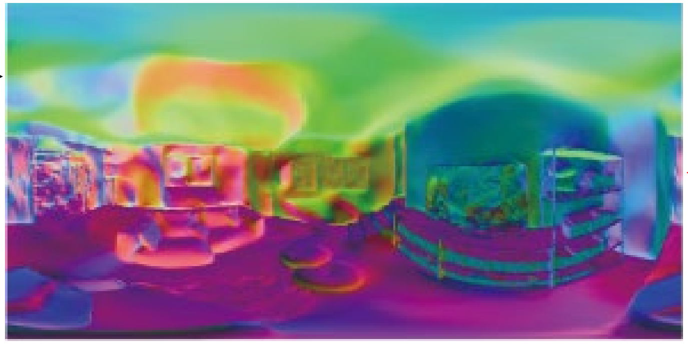
Calculate neighbor pixels based on the sphere projection

$$\mathcal{N}_d(\boldsymbol{p}) = \frac{(\boldsymbol{P}_{(1,0)} - \boldsymbol{P}_{(-1,0)}) \times (\boldsymbol{P}_{(0,1)} - \boldsymbol{P}_{(0,-1)})}{|(\boldsymbol{P}_{(1,0)} - \boldsymbol{P}_{(-1,0)}) \times (\boldsymbol{P}_{(0,1)} - \boldsymbol{P}_{(0,-1)})|}$$

$\boldsymbol{P}_{(\pm1,0)}, \boldsymbol{P}_{(0,\pm1)}$: 3D positions corresponding to the adjacent pixel $(\pm1,0), (0,\pm1)$

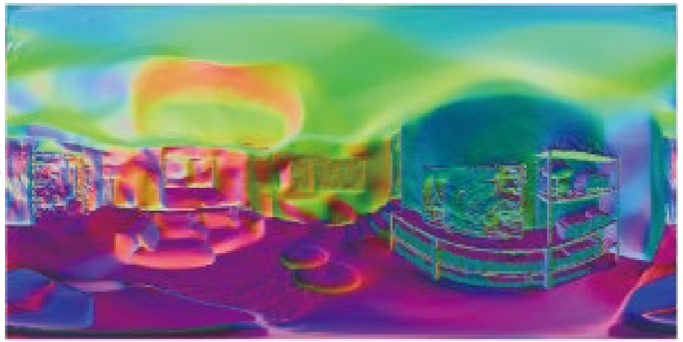
$$\mathcal{N}(\boldsymbol{p}) = \sum_{i=1}^{N_g} \boldsymbol{R}\boldsymbol{n}_i w_i$$

$\boldsymbol{R}$: Rotation matrix of a camera
$\boldsymbol{n}_i$: Normal of Gaussian G_i
w_i : Weight of G_i for the pixel $\boldsymbol{p}$



Rendered normal $\mathcal{N}(\boldsymbol{p})$

Error



Depth-normal $\mathcal{N}_d(\boldsymbol{p})$

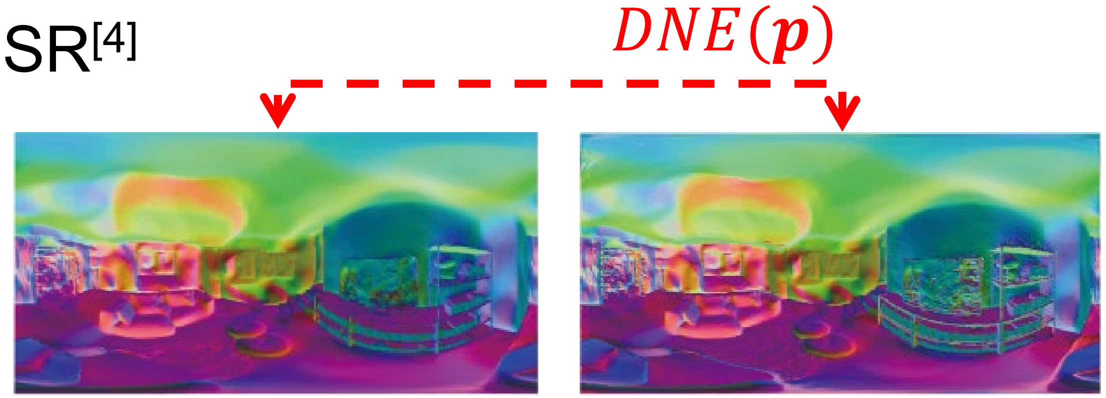
- For each pixel in an ERP image, adjacent pixels in 3D space are selected to compute a normal vector from depth
- The error of the normal map rendered by a rasterizer is minimized

Regularization Terms

- Geometric regularization
- Minimize the error between normal and depth-normal
  - Utilize the color gradient following PGSR[4]

$$DNE(\boldsymbol{p}) = \overline{|\nabla \mathcal{I}(\boldsymbol{p})|}^2 \| \mathcal{N}_d(\boldsymbol{p}) - \mathcal{N}(\boldsymbol{p}) \|$$

$\overline{|\nabla \mathcal{I}(\boldsymbol{p})|}$: The color gradient at pixel $\boldsymbol{p}$

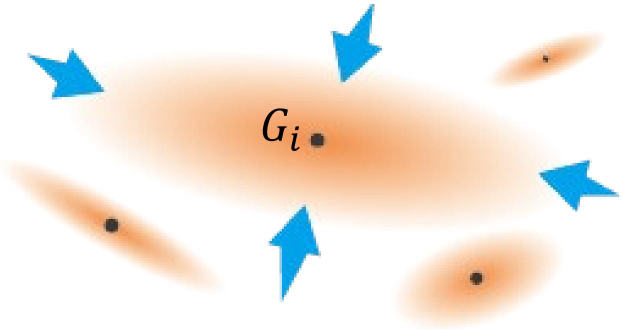


Rendered normal $\mathcal{N}(\boldsymbol{p})$ Depth-normal $\mathcal{N}_d(\boldsymbol{p})$

- Scale regularization
  - Constrain the Gaussians to be smaller so that decrease the number of excessively large Gaussians

$$\mathcal{L}_s = \frac{1}{N_g} \sum_{i=1}^{N_g} \|\boldsymbol{s}_i\|_2^2$$

N_g: The total number of Gaussians
$\boldsymbol{s}_i$: The scale vector of Gaussian G_i



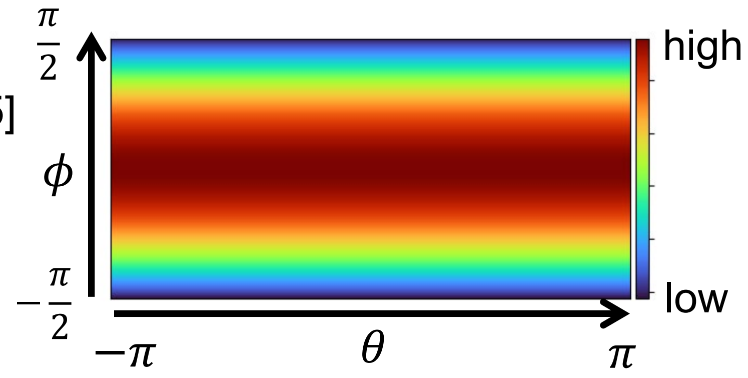
[4] D. Chen et al., “PGSR: Planar-based Gaussian splatting for efficient and high-fidelity surface reconstruction,”, TVCG, Nov. 2024.

Omnidirectional Gaussian Splatting Optimization

- Distortion-aware weight
  - Consider the weight based on the area in 3D space[5]

$$\mathcal{W} = \int_{\theta_0}^{\theta_1} \int_{\phi_0}^{\phi_1} \cos\theta \, d\theta \, d\varphi$$

θ : Longitude in ERP image
ϕ : Latitude in ERP image



Visualization of the weights

- Loss function

$$\mathcal{L} = \mathcal{L}_c + \lambda_{dn}\mathcal{L}_{dn} + \lambda_f \mathcal{L}_f + \frac{1}{2}\lambda_s \mathcal{L}_s$$

Color reconstruction loss $\mathcal{L}_c$

$$\mathcal{L}_c = \frac{\sum_{\boldsymbol{p}\in\mathcal{P}} \mathcal{W}_{\boldsymbol{p}} \odot \mathcal{M}_{\boldsymbol{p}} \odot CRE(\boldsymbol{p})}{\sum_{\boldsymbol{p}\in\mathcal{P}} \mathcal{W}_{\boldsymbol{p}} \odot \mathcal{M}_{\boldsymbol{p}}}$$

$$CRE(\boldsymbol{p}) = (1-\lambda)\left|C(\boldsymbol{p}) - C_{gt}(\boldsymbol{p})\right| + \lambda(1 - SSIM(C(\boldsymbol{p}), C_{gt}(\boldsymbol{p})))$$

Geometric regularization $\mathcal{L}_{dn}$

$$\mathcal{L}_{dn} = \frac{\sum_{\boldsymbol{p}\in\mathcal{P}} \mathcal{W}_{\boldsymbol{p}} \odot \mathcal{M}_{\boldsymbol{p}} \odot DNE(\boldsymbol{p})}{\sum_{\boldsymbol{p}\in\mathcal{P}} \mathcal{W}_{\boldsymbol{p}} \odot \mathcal{M}_{\boldsymbol{p}}}$$

Flattening loss $\mathcal{L}_f$

$$\mathcal{L}_f = \left|\min(s_x, s_y, s_z)\right|$$

$\mathcal{W}_{\boldsymbol{p}}$: Weight at $\boldsymbol{p}$ $\quad$ $\mathcal{M}_{\boldsymbol{p}}$: Mask at $\boldsymbol{p}$ $\quad$ $C(\boldsymbol{p}), C_{gt}(\boldsymbol{p})$: Rendered RGB and GT at $\boldsymbol{p}$, respectively

[5] T. Otonari et al., “Non-uniform sampling strategies for nerf on 360° images,” BMVC, Nov. 2022.

Experimental Setup

- Evaluate the quality of rendered novel views from a set of ERP images
- Dataset
  - Preprocessed data of OmniBlender[6], Ricoh360[6], and OmniScenes[7] provided by ODGS[9]
- Benchmarks and experimental settings
  - ODGS[9], OmniGS[10] , and EgoNeRF[8]
  - The number of iterations are 30,000, 30,000, and 200,000 times, respectively
- Metrics
  - PSNR [dB], SSIM
  - LPIPS (A): AlexNet is used as an encoder

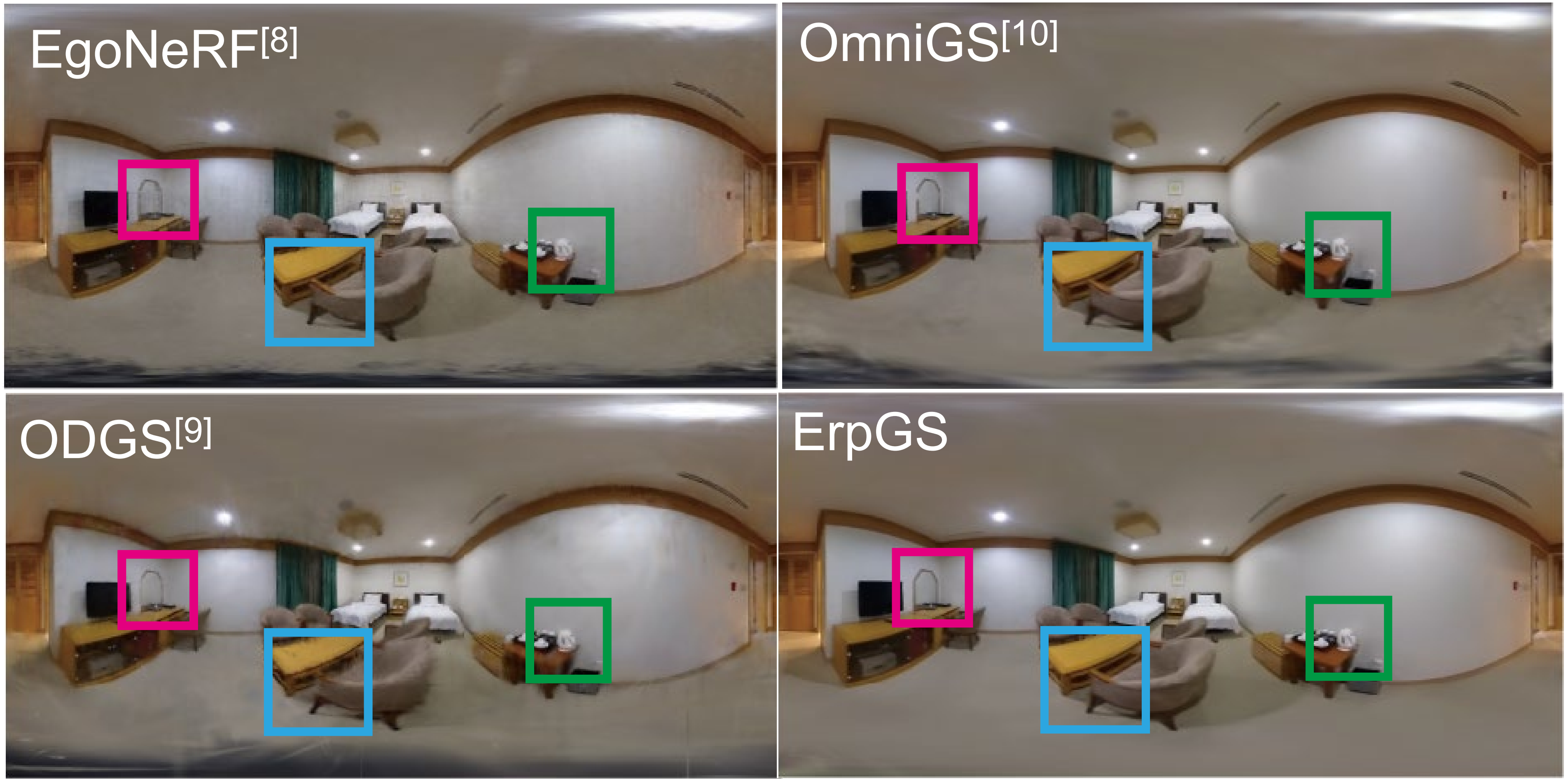
[6] https://github.com/changwoonchoi/EgoNeRF.git [7] https://github.com/82magnolia/piccolo.git
[8] C. Choi et al., “Balanced spherical grid for egocentric view,” CVPR, June 2023.
[9] S. Lee et al., “ODGS: 3D scene reconstruction from omnidirectional images with 3D Gaussian splattings,” NeurIPS, Dec. 2024.
[10] L. Li et al., “OmniGS: Omnidirectional Gaussian splatting for fast radiance field reconstruction using omnidirectional images,” WACV, Feb. 2025.

Quantitative Results in NVS

Dataset	Scene	PSNR [dB] ↑				SSIM ↑				LPIPS (A) ↓			
		EgoNeRF	ODGS	OmniGS	Ours	EgoNeRF	ODGS	OmniGS	Ours	EgoNeRF	ODGS	OmniGS	Ours
OmniBlender	barbershop	30.57	33.66	37.26	**38.71**	0.900	0.947	0.974	**0.979**	0.187	0.123	0.050	**0.040**
	lone-monk	31.10	28.58	29.00	**32.34**	0.935	0.922	0.943	**0.963**	0.073	0.098	0.067	**0.037**
	archiviz-flat	31.69	32.50	33.38	**35.95**	0.917	0.943	0.954	**0.963**	0.103	0.095	0.056	**0.040**
	classroom	26.75	26.20	33.03	**33.62**	0.770	0.798	0.906	**0.917**	0.368	0.385	0.190	**0.157**
Ricoh360	bricks	24.39	22.23	22.27	**25.03**	0.791	0.724	0.766	**0.820**	0.186	0.293	0.251	**0.173**
	center	**29.42**	24.37	26.78	28.63	0.874	0.789	0.855	**0.879**	0.144	0.396	0.188	**0.138**
	farm	**22.58**	20.31	20.04	21.66	0.695	0.630	0.654	**0.696**	0.239	0.396	0.275	**0.210**
	flower	22.09	19.61	21.74	**22.88**	0.658	0.597	0.715	**0.747**	0.291	0.474	0.258	**0.208**
OmniScenes	pyebaek	25.05	23.82	26.67	**27.08**	0.795	0.796	0.862	**0.867**	0.244	0.256	0.168	**0.156**
	room	28.69	27.25	30.29	**31.00**	0.904	0.900	0.928	**0.932**	0.202	0.225	0.155	**0.142**
	wedding-hall	26.00	24.94	26.99	**27.35**	0.826	0.831	0.868	**0.872**	0.242	0.273	0.193	**0.171**

- ErpGS (Ours) achieves the highest accuracy in novel view ERP image rendering
- In real outdoor scenes (i.e., Ricoh360), EgoNeRF demonstrates high robustness

Qualitative Results in NVS



- ErpGS renders novel view ERP image with minimal overall noise, resulting in high qualitative accuracy
  - The noise introduced by the tripod at the bottom of the image is removed

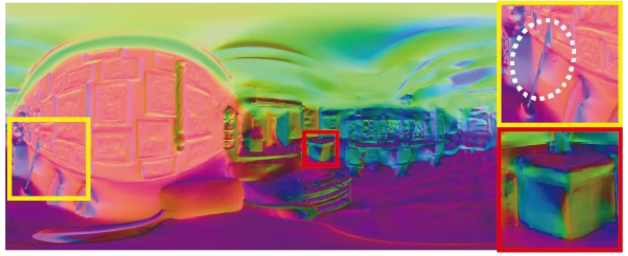
Fine Details of Qualitative Results on NVS

EgoNeRF[8] ODGS[9] OmniGS[10] ErpGS GT

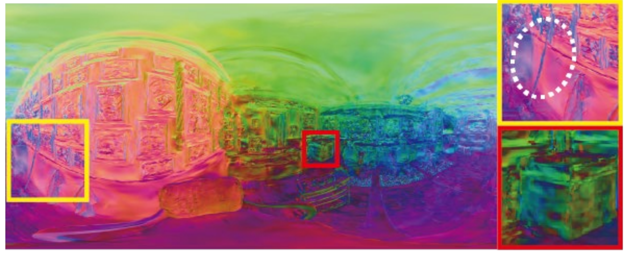
- ErpGS renders the fine details better compared to the other methods

Ablation Study: Geometric Consistency

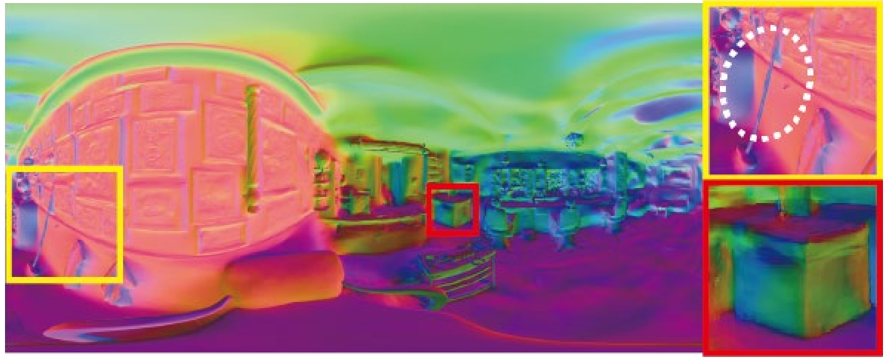
(i) w/o $\mathcal{W}$



(iii) w/o $\mathcal{L}_s$



(ii) w/o $\mathcal{L}_{dn}$



(iv) full

(v) RGB

Ablation	PSNR ↑ [dB]	SSIM ↑	LPIPS (A) ↓	LPIPS (V) ↓
w/o $\mathcal{W}$	33.86	0.949	0.0852	0.1630
w/o $\mathcal{L}_{dn}$	34.61	0.953	0.0717	0.1429
w/o $\mathcal{L}_s$	34.64	0.954	0.0723	0.1434
All	**35.16**	**0.956**	**0.0687**	**0.1374**

Quantitative results of rendered ERP images

- The ablation studies for distortion weight $\mathcal{W}$, geometric regularization $\mathcal{L}_{dn}$, and scale regularization $\mathcal{L}_s$
- Considering the distortion and preserving geometric consistency is crucial

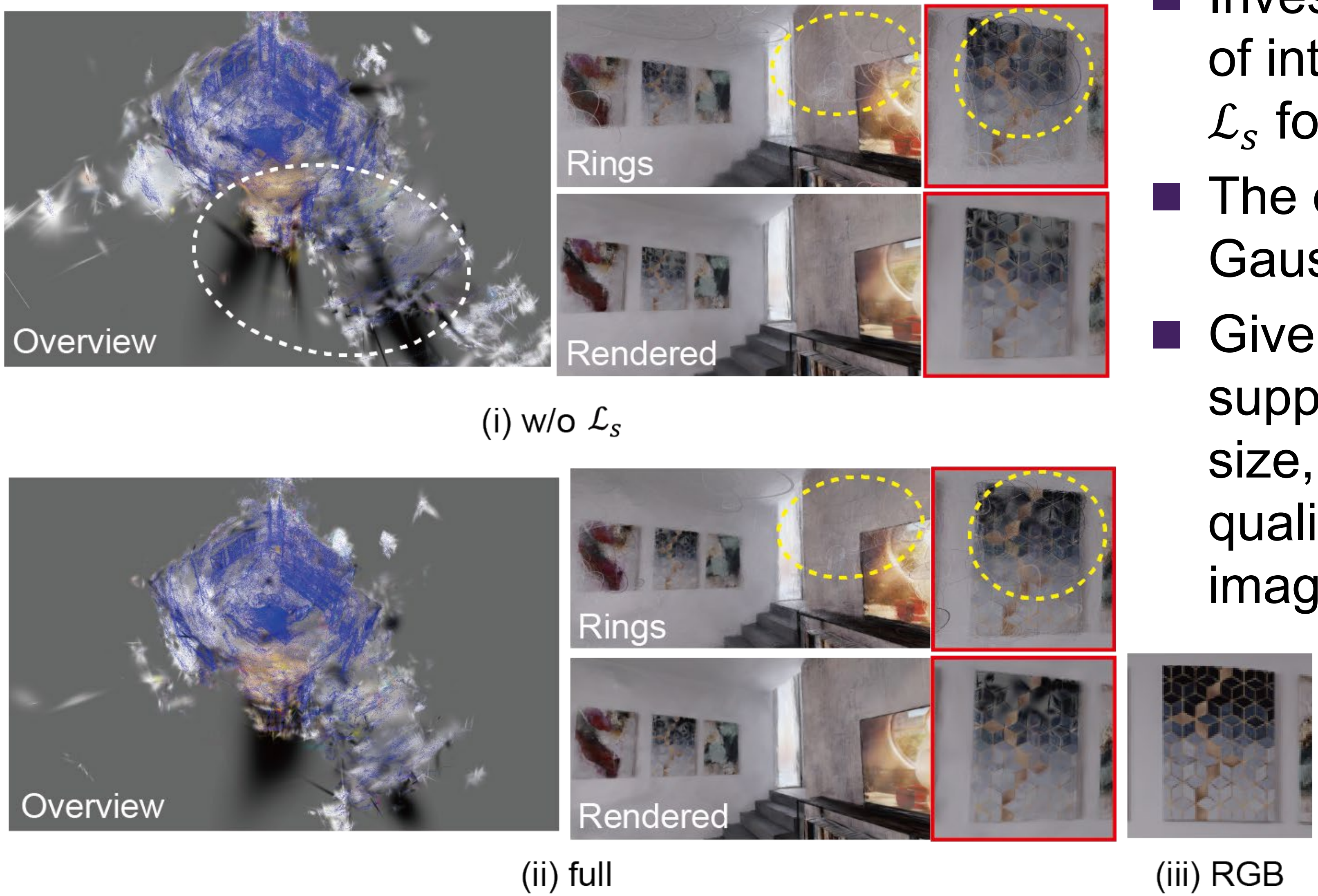
Conclusion and Future Work

- Conclusion
  - We achieved improved accuracy in novel view synthesis from a set of ERP images by introducing geometric regularization and scale regularization to optimize omnidirectional Gaussian splatting

- Future Work
  - Achieve better novel view synthesis, including camera parameter refinement
  - Realize 3D mesh reconstruction for a large-scale scene using ERP images

Thank you for your attention!!

Ablation Study: Appearance of Gaussian Ellipsoids



- Investigate the effectiveness of introducing a scaling loss $\mathcal{L}_s$ for generating Gaussians
- The occurrence of large Gaussians are suppressed
- Given that the Gaussians are suppressed to an appropriate size, it is believed that the quality of the rendered images are improved