Depth-Guided Privacy-Preserving Visual Localization Using 3D Sphere Clouds

BMVC 2024

Code

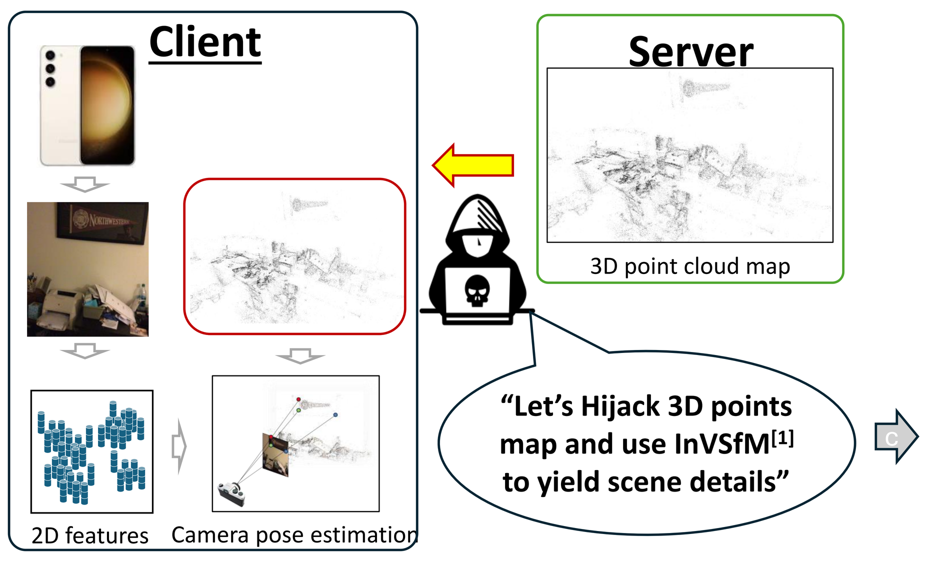
Heejoon Moon[1], Jongwoo Lee[1], Jeonggon Kim[2], Je Hyeong Hong[*,1,2]

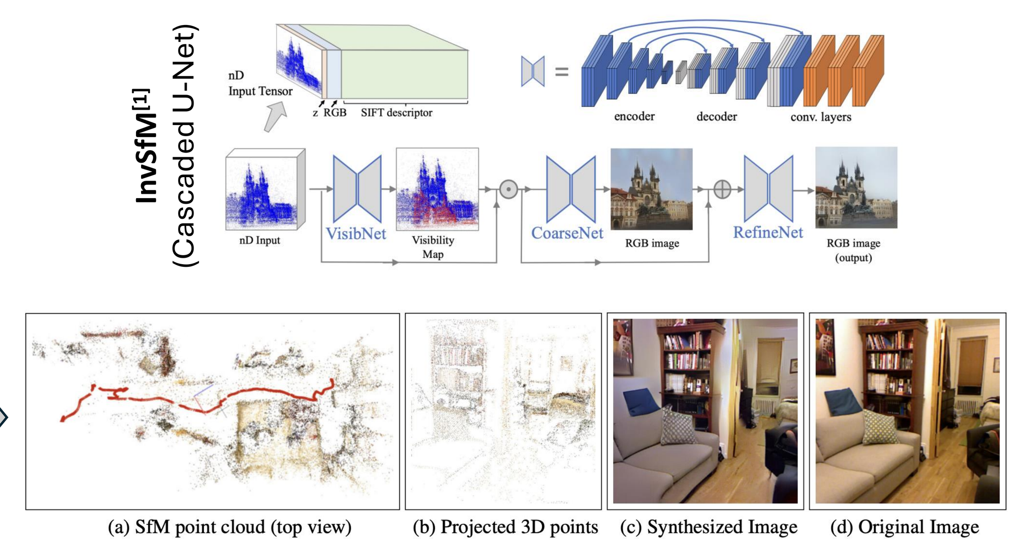
[1]Department of Artificial Intelligence, Hanyang University , [2]Department of Electronic Engineering, Hanyang University

Privacy risks in visual localization

Inversion attack - Pittaluga *et al.*[1]

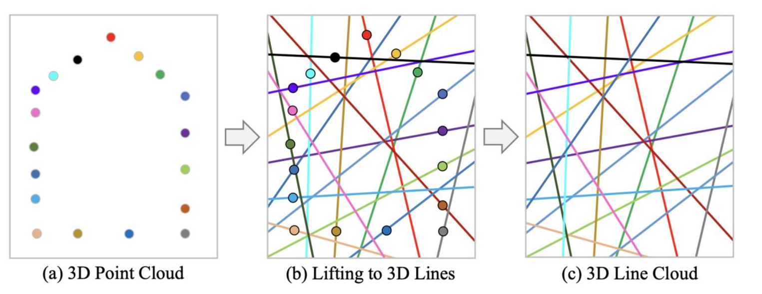
- Deep inversion model enables to generate high-fidelity scene details from sparse 3D point cloud map
- **Privacy contents in the synthesized image can be exposed! -> "*Inversion attack*"**



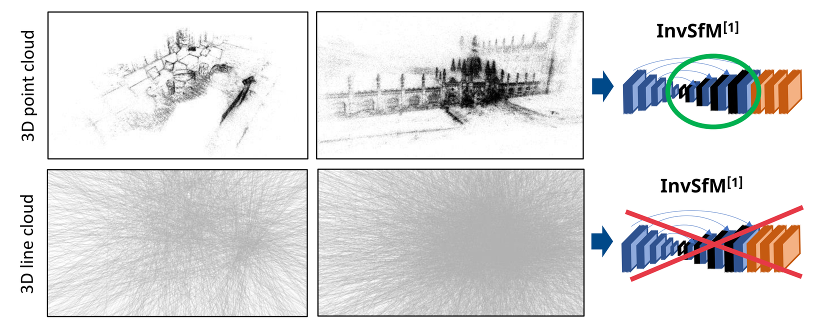


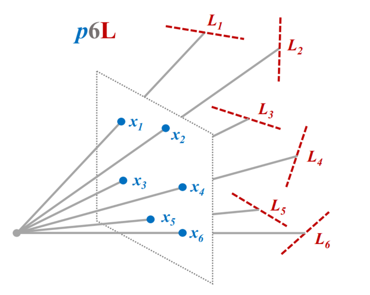
Geometric lifting into 3D lines - Speciale *et al.*[2]

- Replace 3D points into 3D randomly oriented lines passing through them -> **Prevent Inversion attack**
- Propose point-6-Line (p6L) minimal solver for visual localization



Lifting process[2]

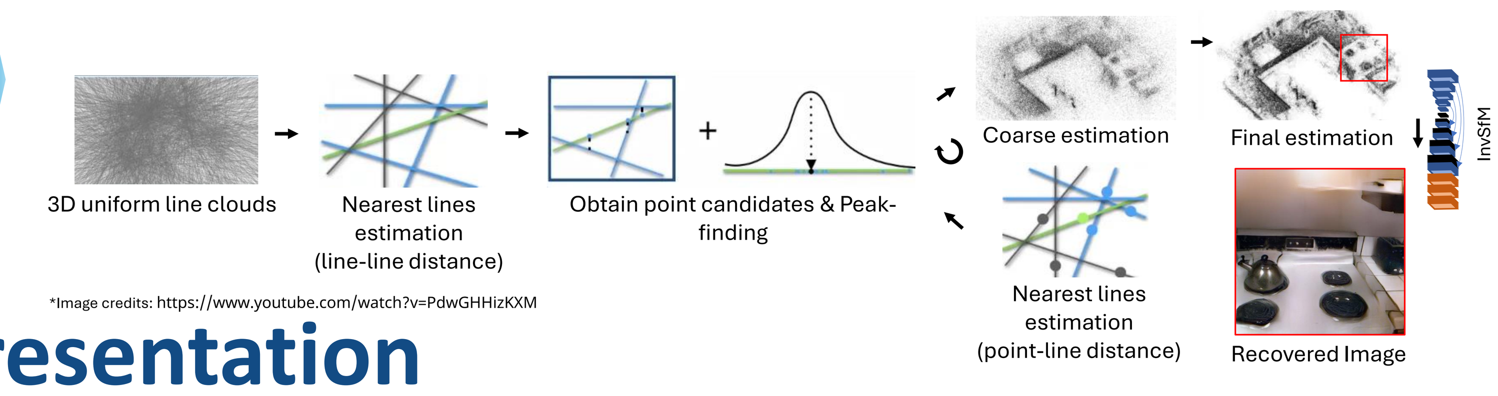




Localization (p6L solver[2])

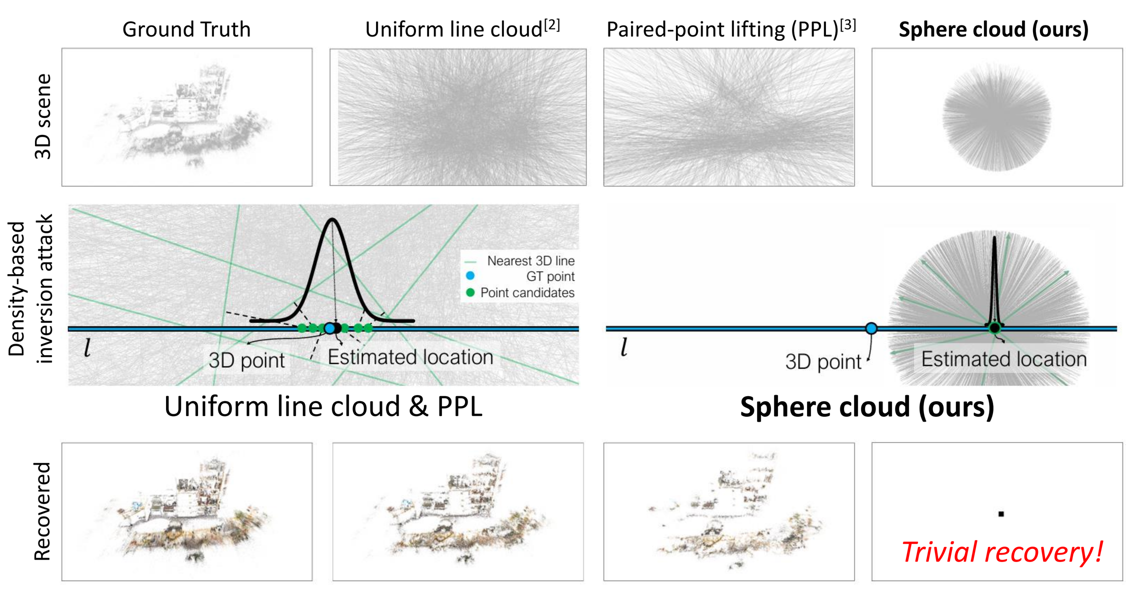
Density-based inversion attack - Chelani *et al.*[3]

- Propose to recover 3D points from (random) 3D lines
- Bayes rule: $P(\mathcal{P}|\mathcal{L}) \propto P(\mathcal{L}|\mathcal{P})P(\mathcal{P})$, $P(\mathcal{L}|\mathcal{P})$ is constant in random lines
- Maximize $P(\mathcal{P})$: Retrieving closest-point-to-other-lines & Peak-finding

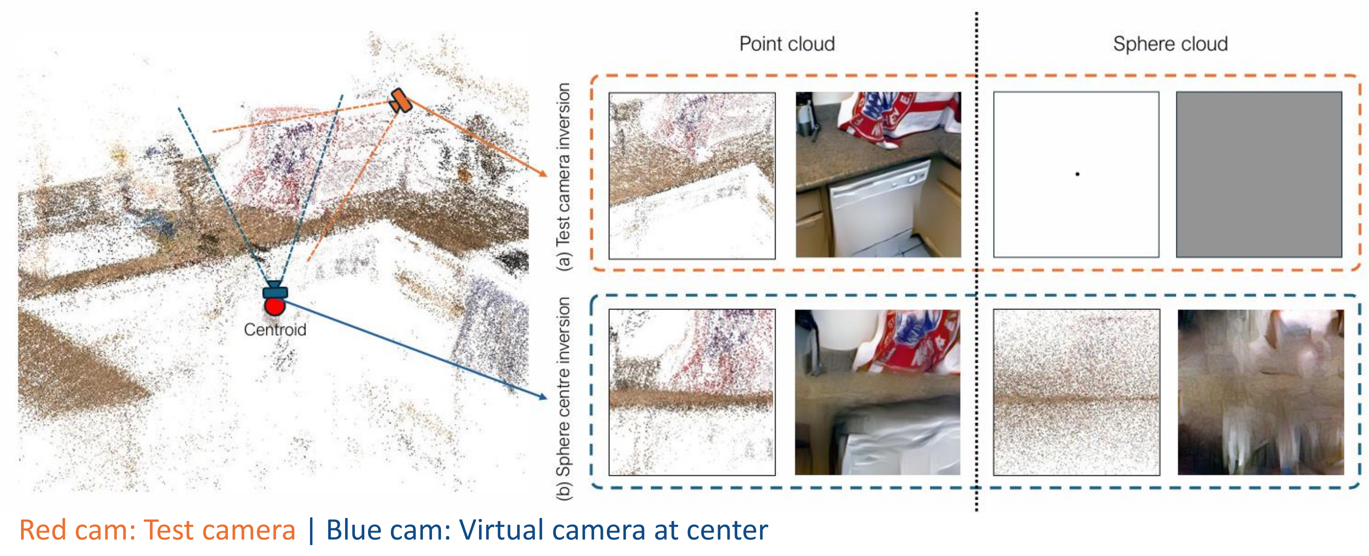


3D Sphere Clouds: A new privacy-preserving scene representation

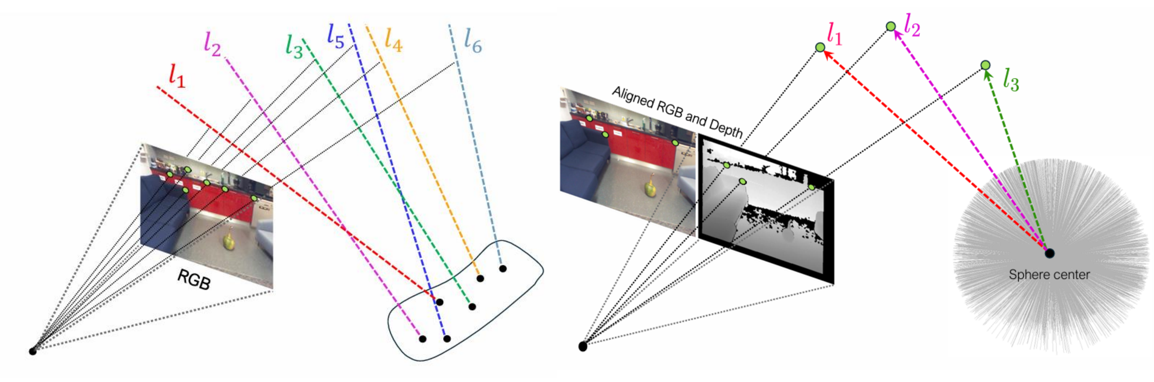
1. Completely block the density-based inversion attack[3] due to the all 3D lines intersect at a sphere center
2. Explore a new type of attack from breaching the sphere cloud and present a simple and effective strategy based on sparsification and the reusage of descriptors
3. Propose the first privacy-preserving localization framework to leverage depth observations for efficient camera pose estimation



1) Robustness to density-based inversion attack

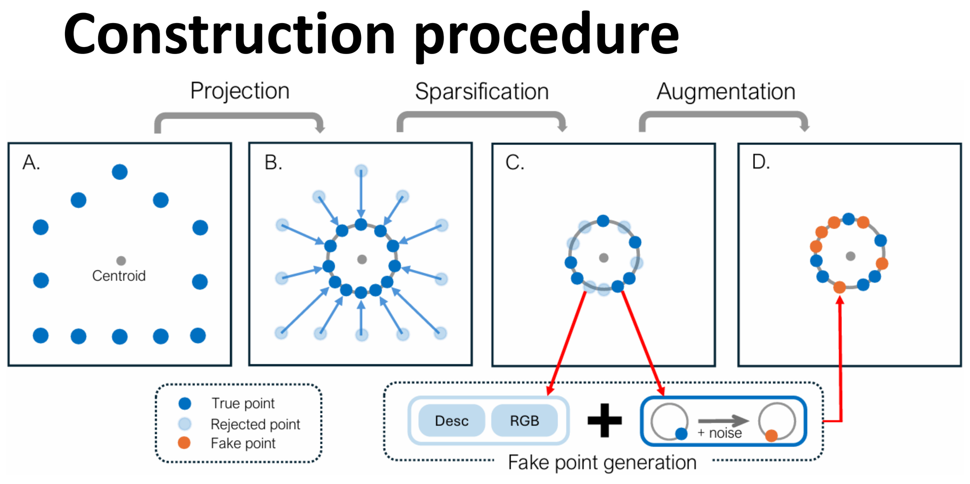


2) Exploring and addressing the potential threat in sphere cloud



3) Efficient localization via depth measurements

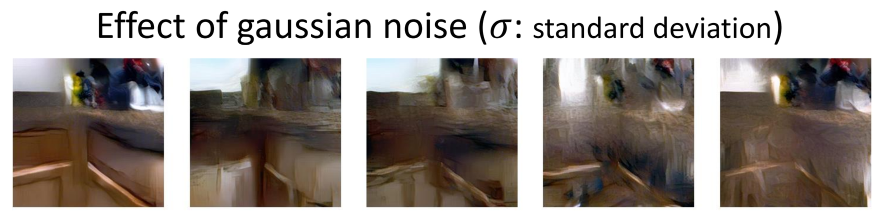
Construction procedure



1. **Project** 3D points onto the unit sphere centered at the map centroid and **discard** the portion (η) of sphere point
2. **Fake points generation**
   - **Position**: gaussian noise to the remaining point
   - **RGB, descriptors:** recycled from rejected points

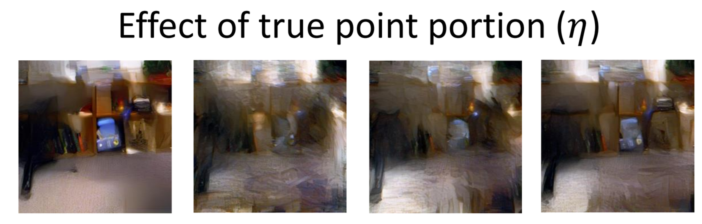
Addressing potential inversion at center by adding fake points

Effect of gaussian noise (σ: standard deviation)



(a) Pseudo-GT (b) $\sigma^2 = 0$ (c) $\sigma^2 = 0.01$ (d) $\sigma^2 = 0.1$ (e) $\sigma^2 = 1$

Effect of true point portion (η)



Pseudo-GT η = 25 % η = 33 % η = 50 %

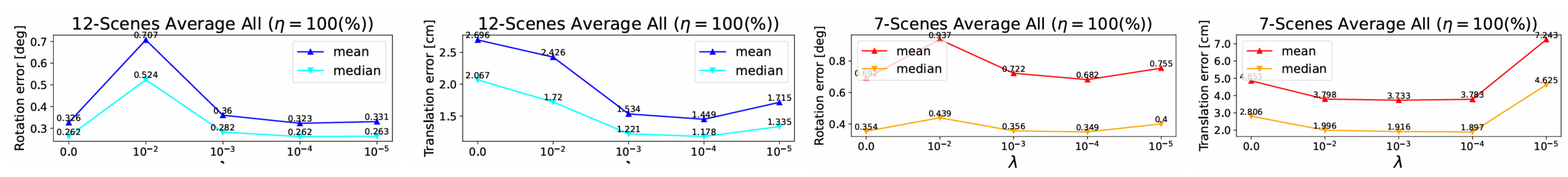
Robust pose estimation with depth regularization

- Initial pose: Aligned RGB and Depth → Get 3D keypoints -> Efficient p3p solver[5]
- Minimizing total cost function (L): LO-RANSAC pipeline with non-linear refinement

$$L = \sum_{i\in\Omega}\left(L_i^e + \lambda\, L_i^d\right) \quad \Rightarrow \quad L_i^e = \frac{([\mathbf{u}_i^\top, 1]\mathrm{K}^{-\top}\,\mathrm{E}\,\tilde{\mathbf{x}}_i)^2}{(\mathbf{e}_1^\top\tilde{\mathbf{x}}_i)^2 + (\mathbf{e}_2^\top\tilde{\mathbf{x}}_i)^2} \qquad L_i^d = (\beta_i - 1)^2, \quad \beta_i = z_i / z_i^{TOF}$$

Epipolar distance *Depth constraint*

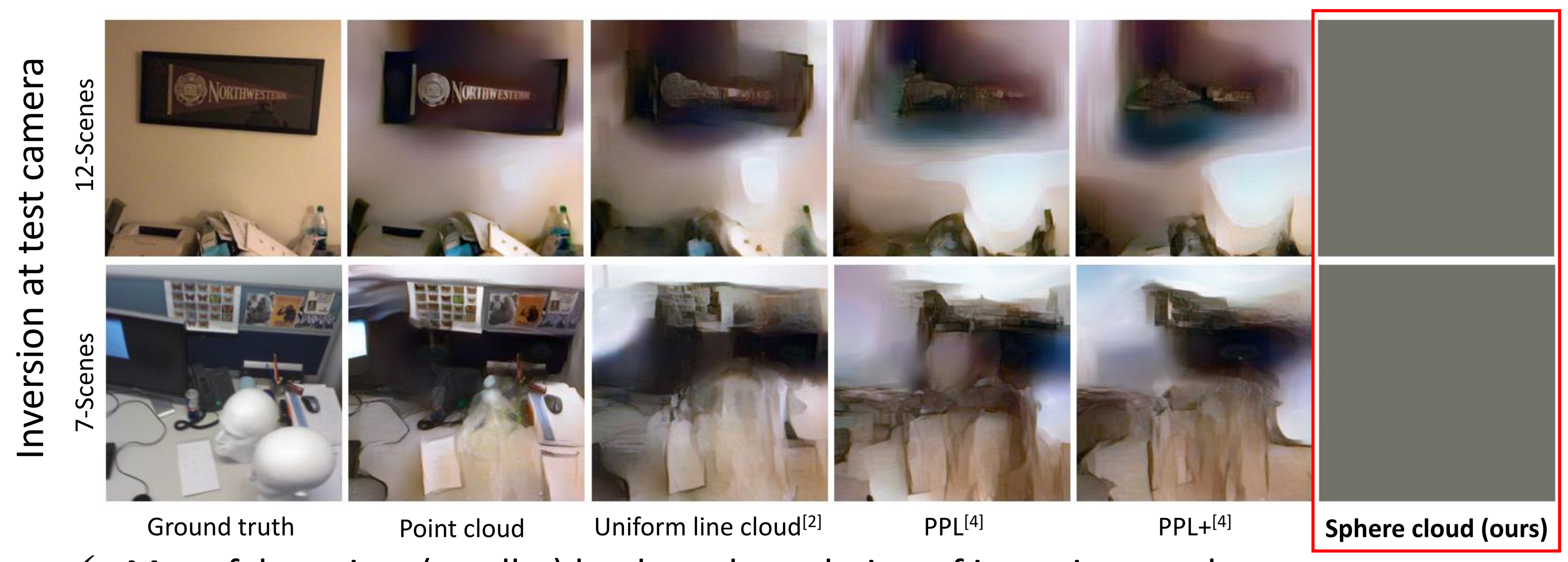
- Depth constraint ($\lambda = 10^{-4}$) leads to better localization accuracy than no constraint ($\lambda = 0$)



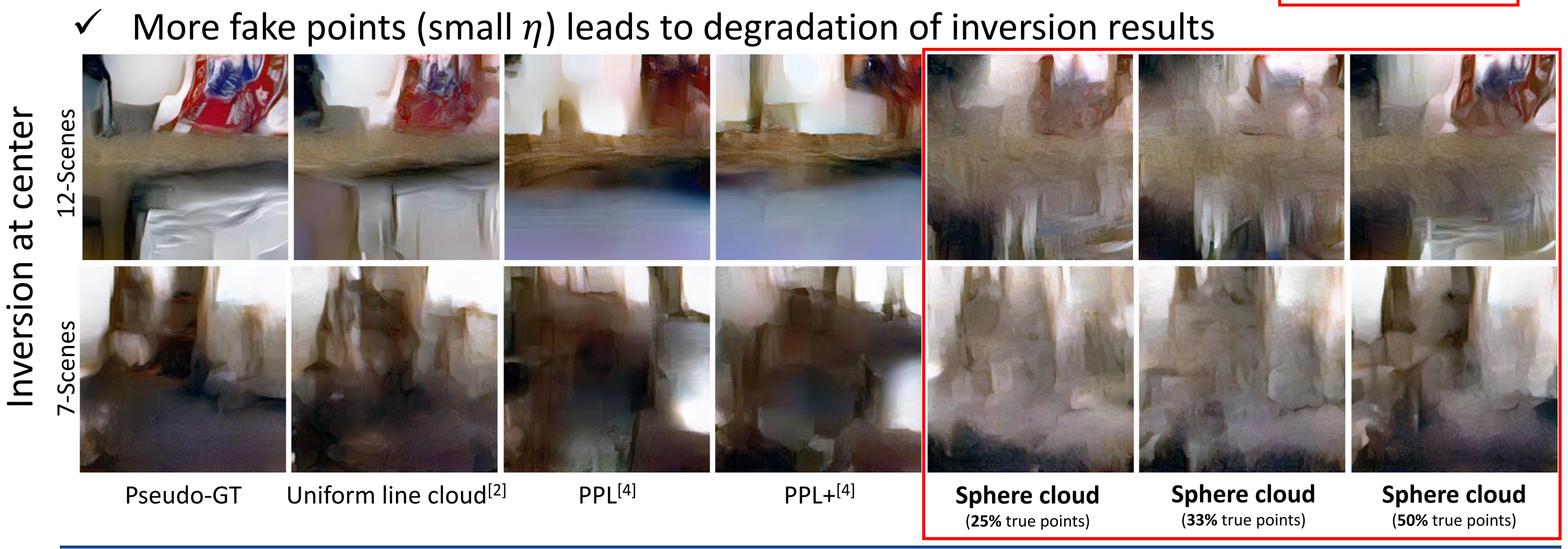
Experimental Results

Qualitative results of inversion attack[1]

✓ Sphere cloud completely blocks the scene details compared to other 3D representations



✓ More fake points (small η) leads to degradation of inversion results



[1] Pittaluga et al. Revealing Scenes by Inverting Structure from Motion Reconstructions, CVPR 2019
[2] Speciale et al. Privacy Preserving Image-Based Localization, CVPR 2019
[3] Chelani et al. How Privacy-Preserving are Line Clouds? Recovering Scene Details from 3D Lines, CVPR 2021
[4] Lee et al. Paired-point lifting for enhanced privacy-preserving visual localization, CVPR 2023
[5] Persson et al. Lambda twist: An accurate fast robust perspective three point (p3p) solver, ECCV 2018
[6] Chelani et al. Obfuscation Based Privacy Preserving Representations are Recoverable Using Neighborhood Information, arXiv 2024

Localization performance

✓ Sphere cloud shows similar localization accuracy among depth-guided localization methods and achieves real-time performance

Dataset	Metric	Image-based localization: Point cloud [25]	ULC [35]	PPL [16]	PPL+ [16]	Depth-guided localization: DVLAD* +R2D2(+D)[13]	DSAC* (+D)[2]	Sphere (η=25%)	Sphere (η=33%)	Sphere (oracle) (η=25%)	Sphere (oracle) (η=33%)
12-Scenes [38]	ΔR(°) (↓)	0.139	**0.159**	0.170	0.168	0.389	0.397	0.300	**0.288**	0.240	**0.232**
	Δt (cm) (↓)	0.627	**0.727**	0.775	0.765	0.931	**0.735**	1.310	1.282	0.601	**0.577**
	ΔR<3° (%) (↑)	100.0	**100.0**	**100.0**	**100.0**	99.73	**99.98**	99.00	99.34	99.90	**100.0**
	Δt<3cm (%) (↑)	97.94	**95.88**	95.16	95.13	97.06	**99.21**	86.97	87.86	97.22	**97.60**
	Runtime(ms) (↓)	3	96	**91**	**91**	-	84	48	**24**	22	**13**
7-Scenes [33]	ΔR(°) (↓)	0.174	**0.201**	0.206	0.207	0.966	0.655	0.438	**0.405**	0.262	**0.255**
	Δt (cm) (↓)	0.493	**0.613**	0.647	0.647	2.857	**1.573**	2.119	2.051	0.459	**0.443**
	ΔR<3° (%) (↑)	100.0	**100.0**	**100.0**	**100.0**	96.11	**99.05**	97.00	97.58	99.86	**99.93**
	Δt<3cm (%) (↑)	99.85	**99.32**	99.12	98.96	55.90	**82.81**	69.75	70.93	98.21	**98.51**
	Runtime (ms) (↓)	3	82	**78**	79	-	80	52	**25**	31	**16**

Conclusion

- Fully resilient to the density-based inversion attack and address the potential inversion at the center by injecting fake points
- Efficient and real-time performance localization with the guidance of depth measurements
- (Expected) The only privacy-preserving method against the recently proposed geometry inversion[6]

Limitation and future work

- Noises of depth measurement lead to inaccurate localization → Improvement via denoising

Acknowledgement

This work was supported by the NRF (National Research Foundation of Korea) grants funded by the Korea government (MSIT) (No. 2022R1C1C1004907)