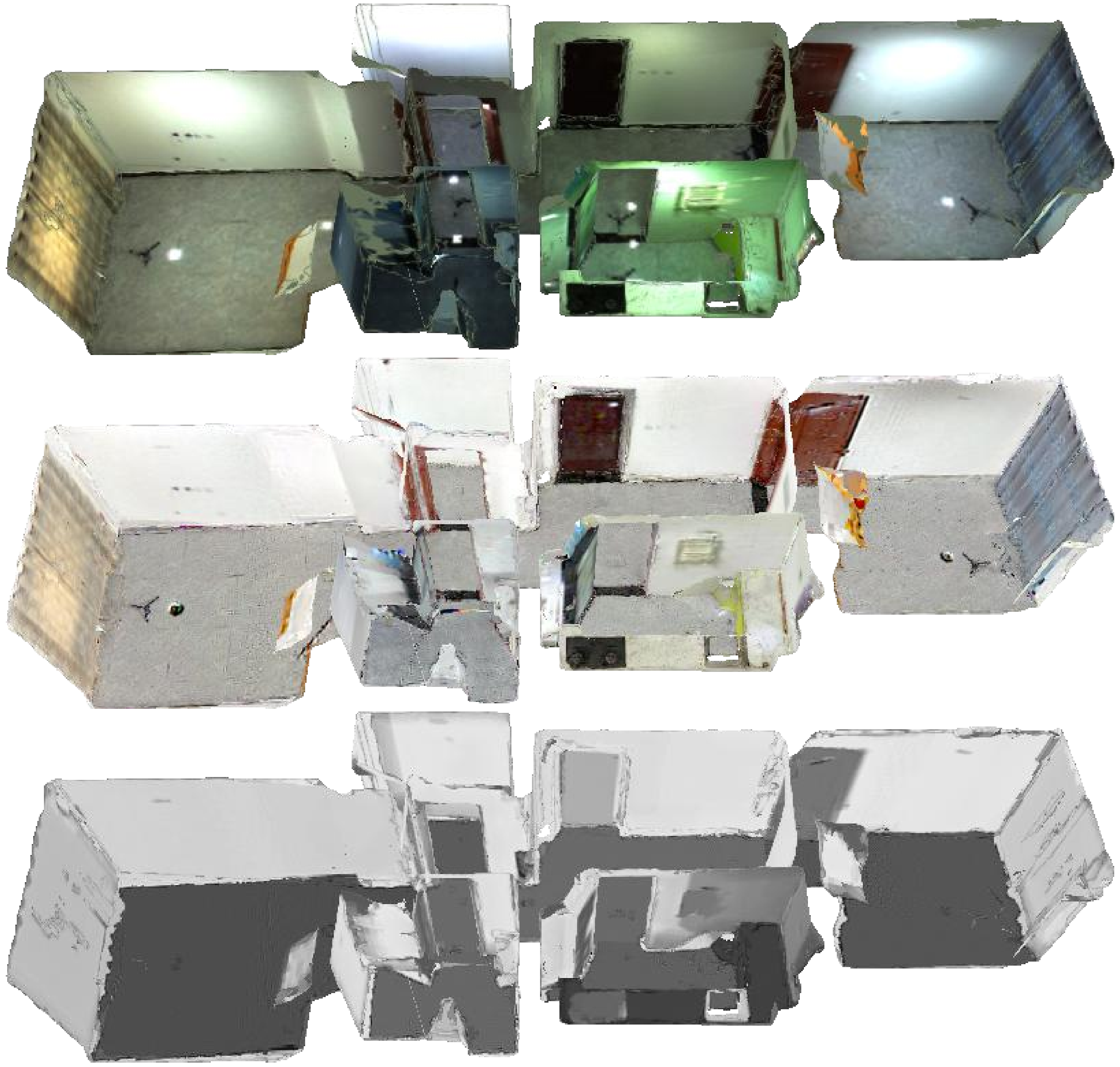
Multi-view Inverse Rendering for Large-scale Real-world Indoor Scenes

WED-PM-015

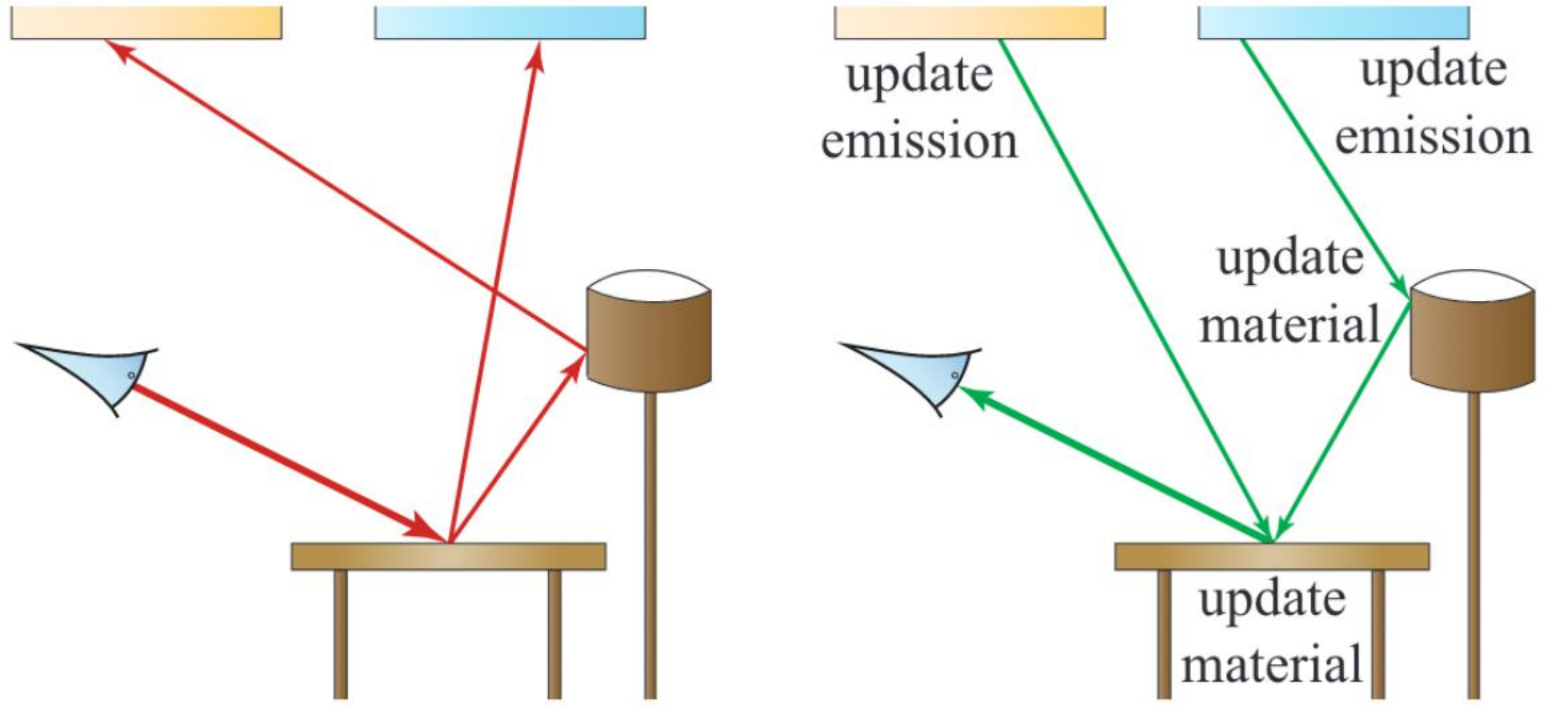
Zhen Li[1], Lingli Wang[1], Mofang Cheng[1], Cihui Pan[1], Jiaqi Yang[2]
[1]Realsee [2]Northwestern Polytechnical University

Overview

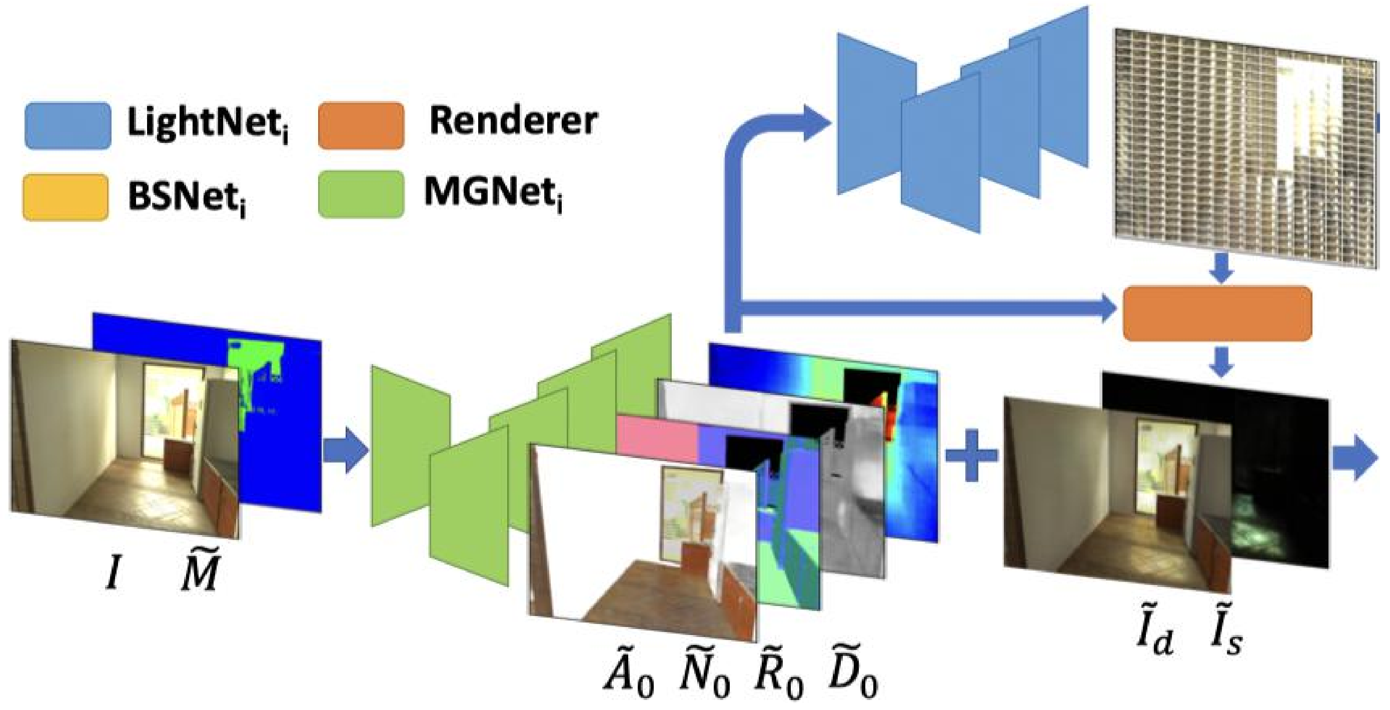


- Input: sparse posed HDR images
- Output: 3D assets, including global illumination, albedo, roughness

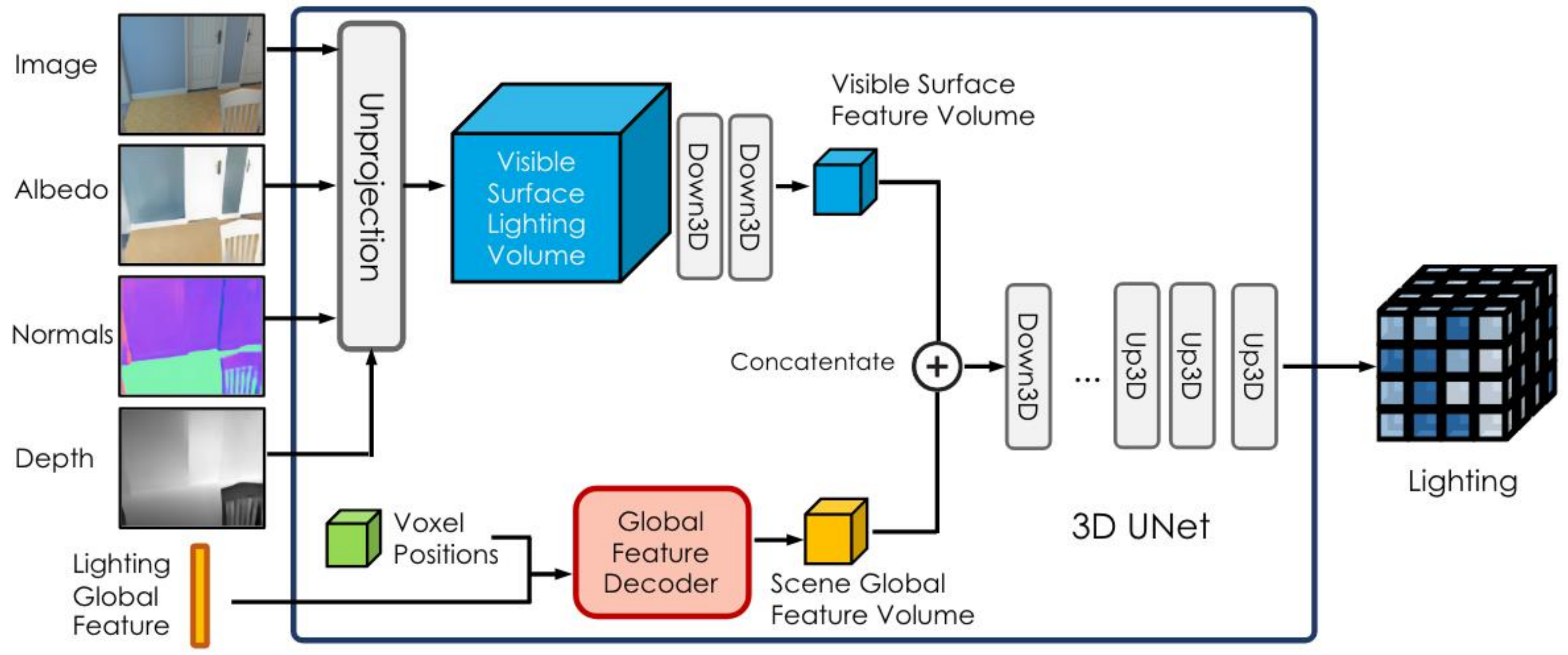
Prior Works -- **Modeling global illumination**



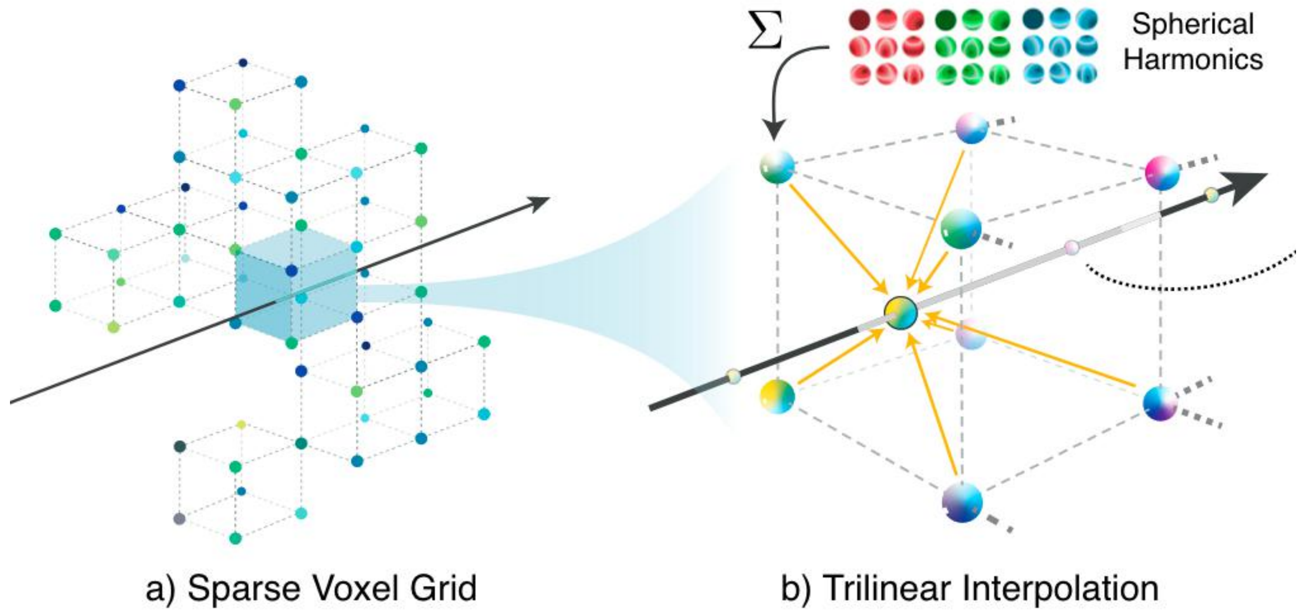
Path tracing[1]



Per-pixel IBL[2]



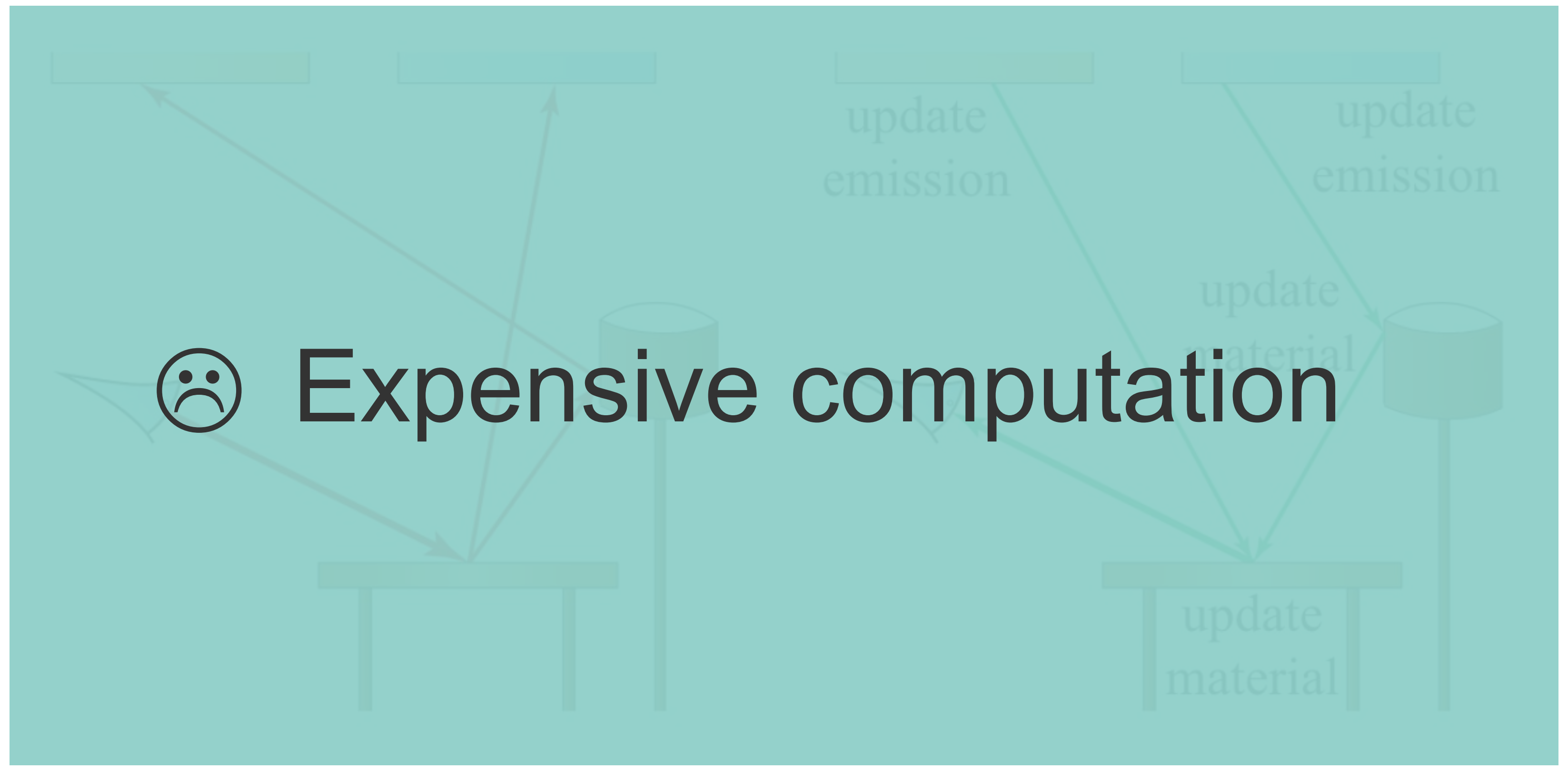
Dense voxel IBL[3]



Sparse voxel IBL[4]

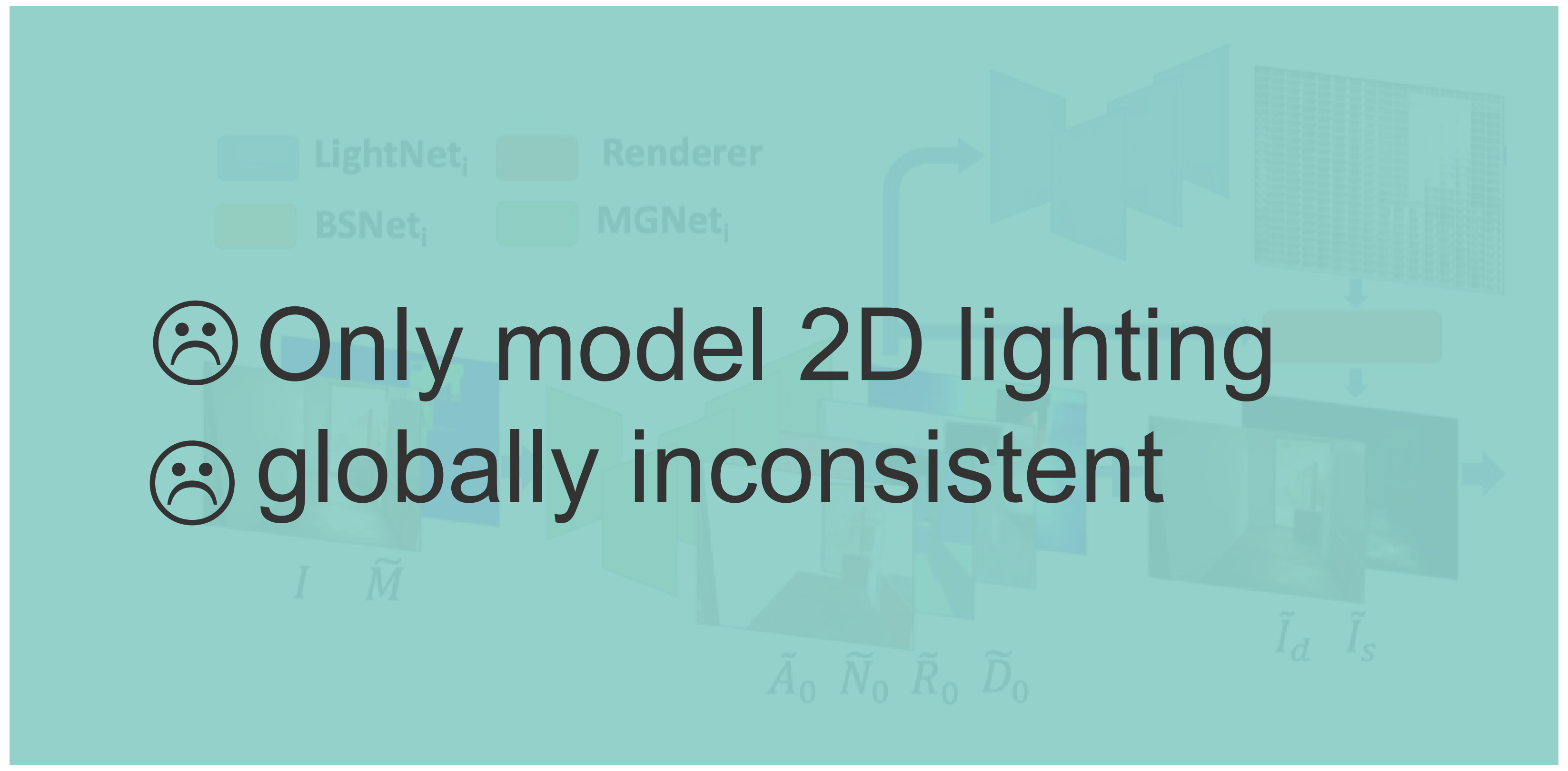
[1] Inverse Path Tracing for Joint Material and Lighting Estimation. CVPR'19
[2] Inverse Rendering for Complex Indoor Scenes: Shape, Spatially-Varying Lighting and SVBRDF From a Single Image. CVPR'20
[3] Learning Indoor Inverse Rendering with 3D Spatially-Varying Lighting. ICCV'21
[4] Plenoxels: Radiance Fields without Neural Networks. CVPR'22

Prior Works -- Modeling global illumination



☹ Expensive computation

Path tracing[1]

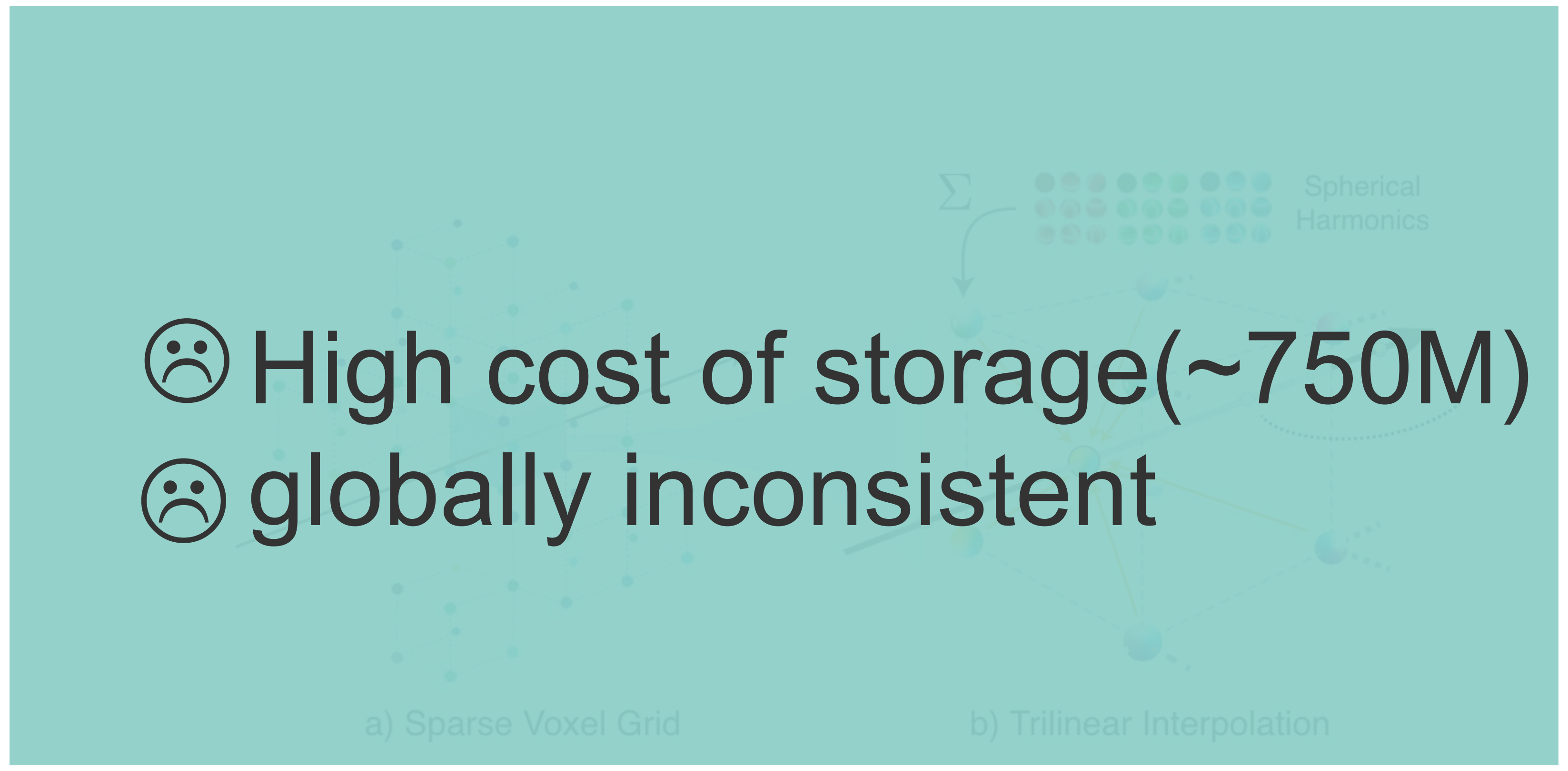


☹ Only model 2D lighting
☹ globally inconsistent

Per-pixel IBL[2]

☹ High cost of storage(~1G)
☹ globally inconsistent

Dense voxel IBL[3]



☹ High cost of storage(~750M)
☹ globally inconsistent

Sparse voxel IBL[4]

[1] Inverse Path Tracing for Joint Material and Lighting Estimation. CVPR'19
[2] Inverse Rendering for Complex Indoor Scenes: Shape, Spatially-Varying Lighting and SVBRDF From a Single Image. CVPR'20
[3] Learning Indoor Inverse Rendering with 3D Spatially-Varying Lighting. ICCV'21
[4] Plenoxels: Radiance Fields without Neural Networks. CVPR'22

Texture-based Lighting

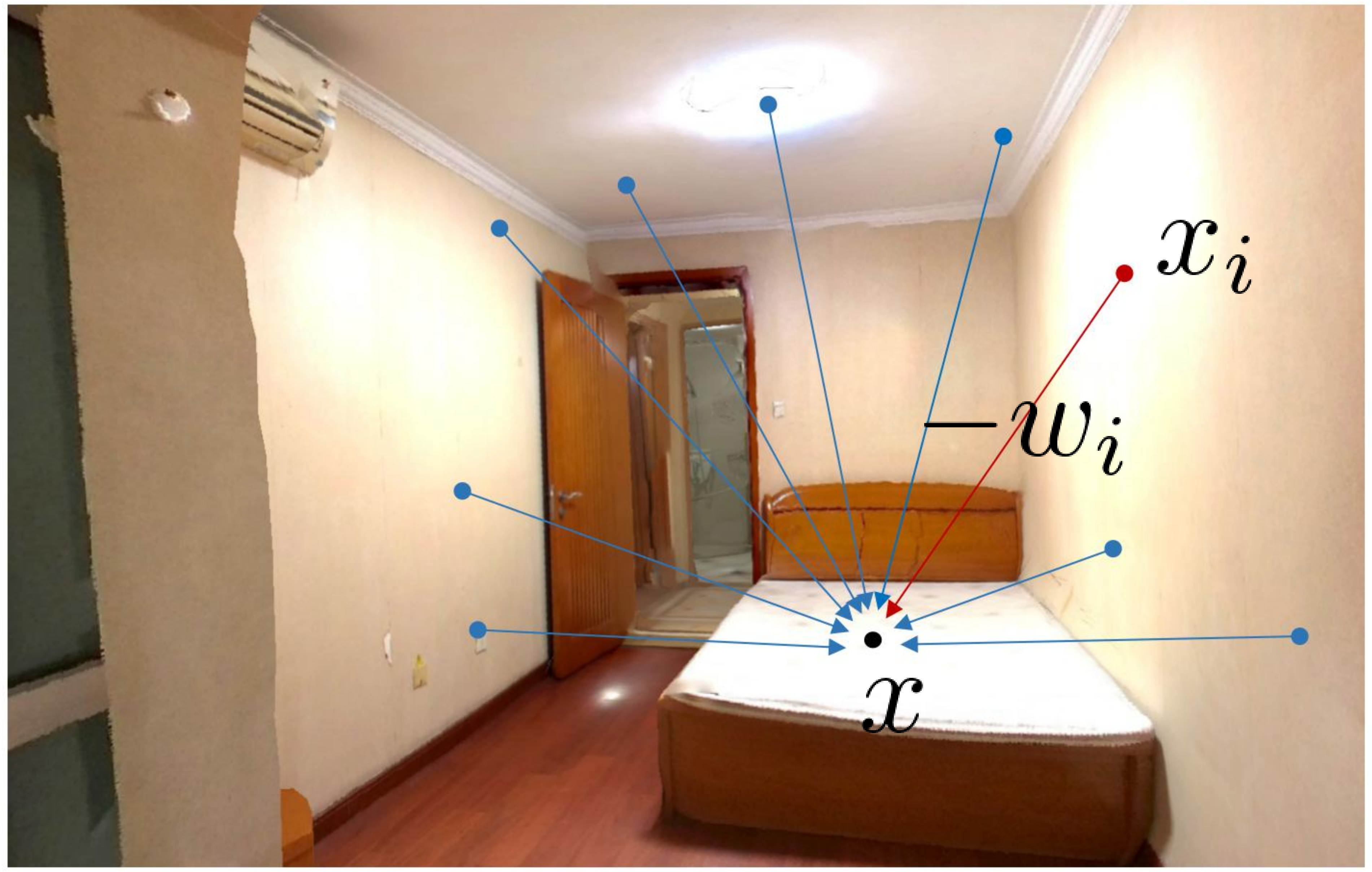
Posed HDR images

Geometry with HDR texture

Texture-based Lighting

$$L_o(x, w_o) = \int_{H^+} f_r(x, w_i, w_o) \boxed{L_i(x, w_i)} (w_i \cdot n) d_{w_i}$$

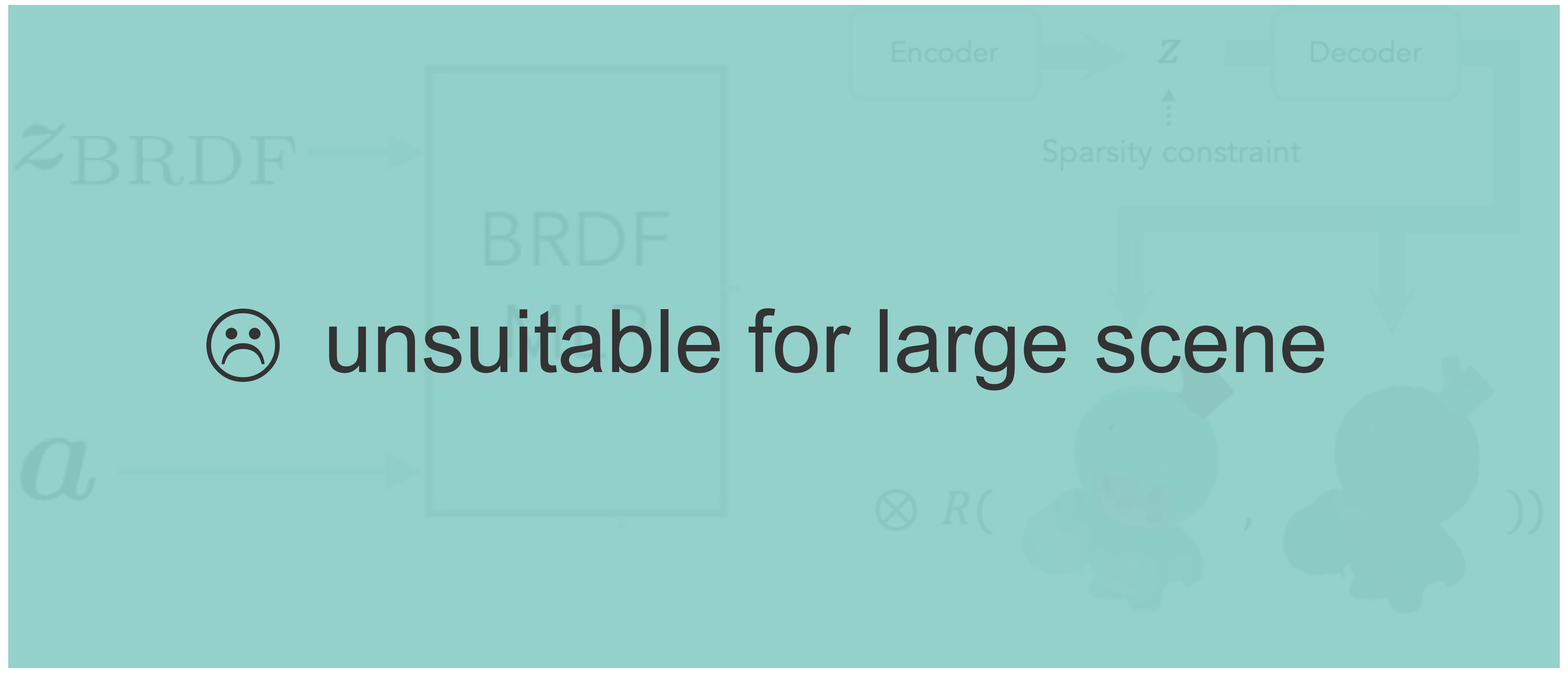
Incoming lighting



Geometry with HDR texture

precomputed irradiance

Prior Works -- Disentangling the ambiguity between materials



☹ unsuitable for large scene

Learning prior[1,2]

☹ wrong local optima

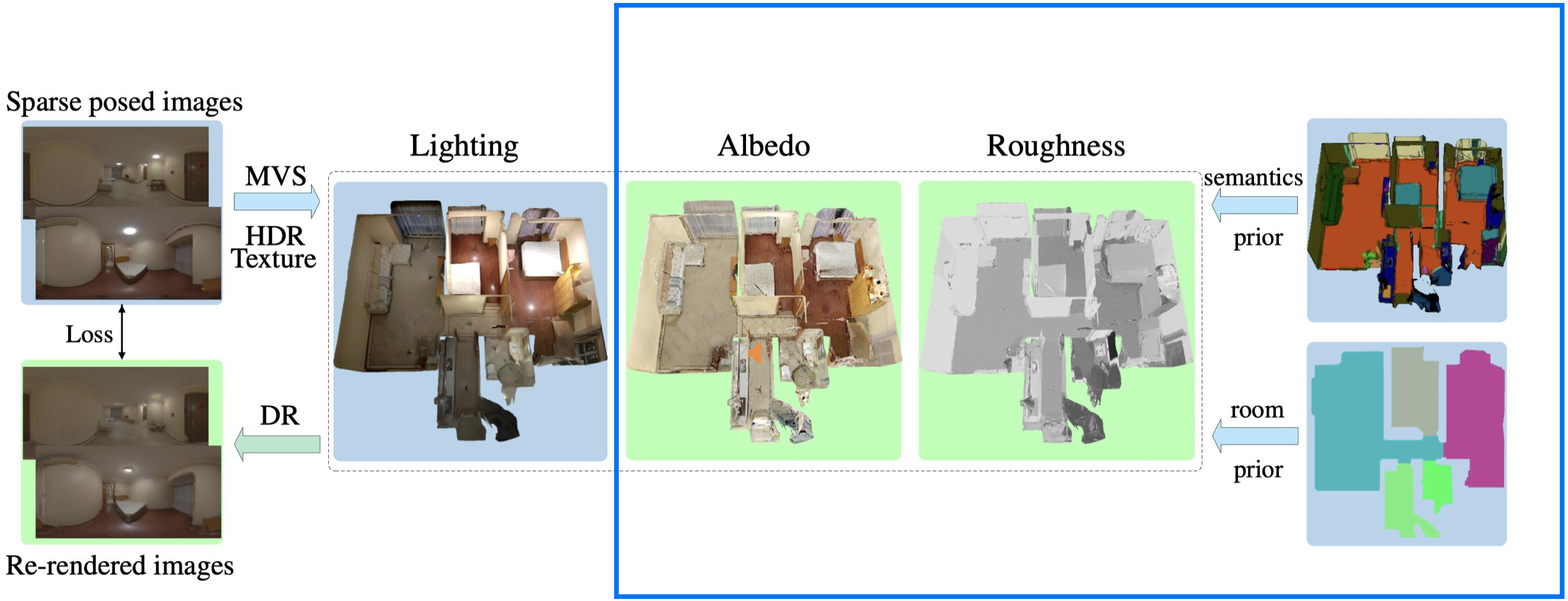
Local smooth[3]

☹ reasonable smooth but ill-posed value

Non-local bilateral smooth[4]

[1] NeRFactor: Neural Factorization of Shape and Reflectance Under an Unknown Illumination. TOG'21
[2] Modeling Indirect Illumination for Inverse Rendering. CVPR'22
[3] Extracting Triangular 3D Models, Materials, and Lighting From Images. CVPR'22
[4] NeILF: Neural Incident Light Field for Physically-based Material Estimation. ECCV'22

Architecture

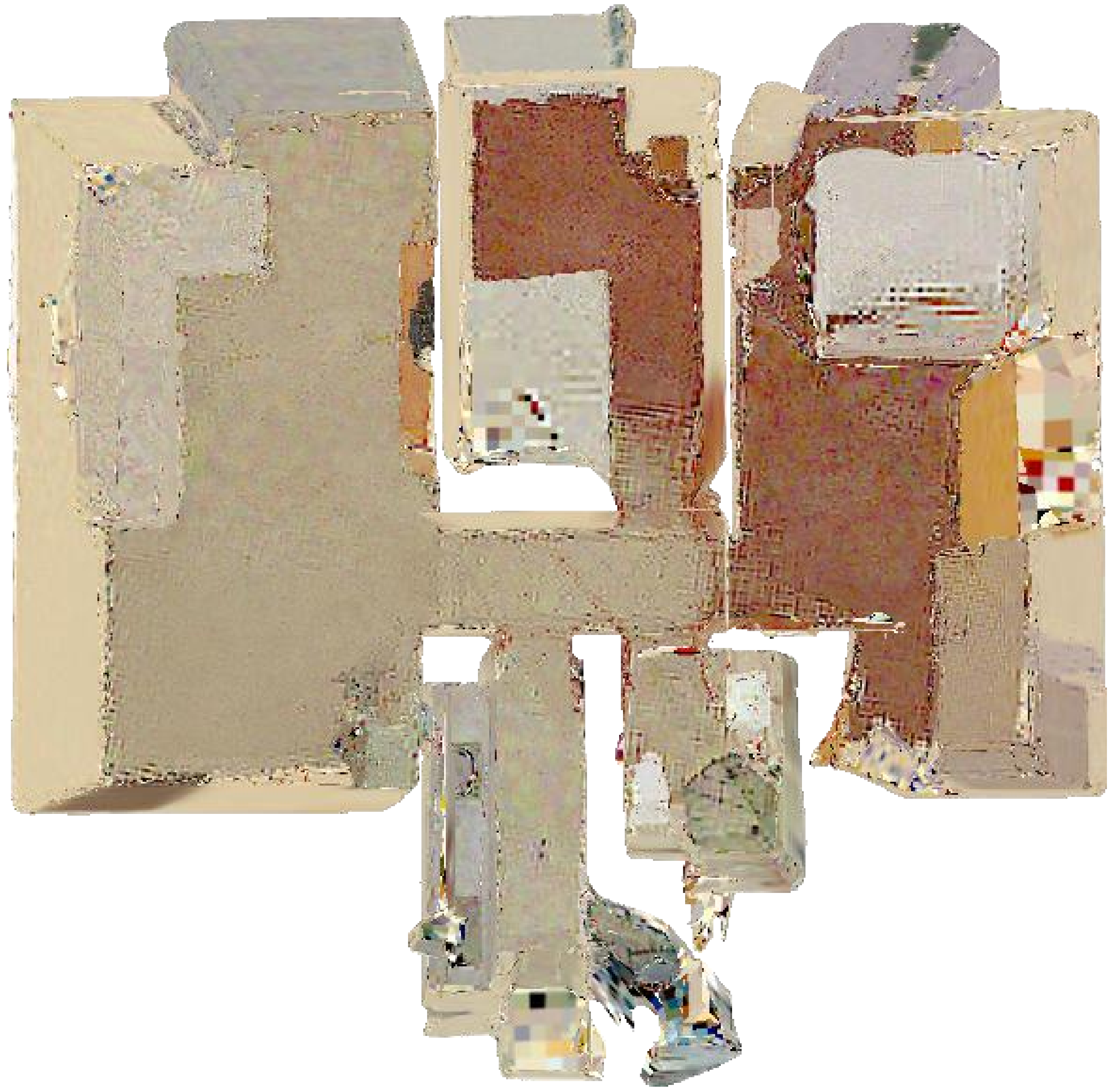


segmentation-based prior

Material Optimization -- Stage 1 Coarse Albedo

- Lambertian assumption
- semantics-based albedo smooth

$$\mathcal{L}_{ss} = \sum_{c} \left| F - \frac{\sum_p F \odot M_{seg}(c)}{\sum_p M_{seg}(c) + \epsilon} \right| \odot M_{seg}(c)$$



coarse albedo texture

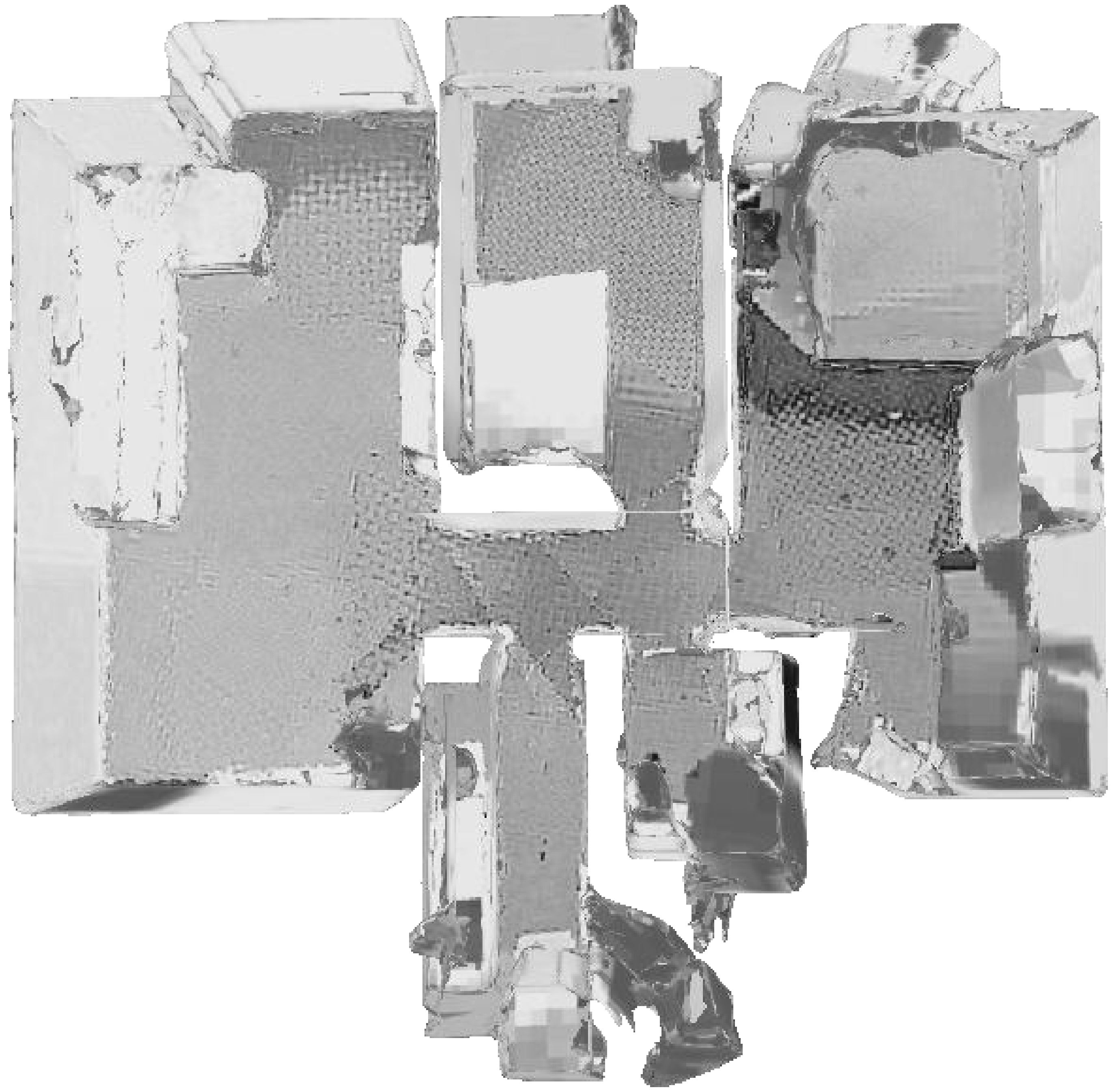
Material Optimization -- Stage 2 Coarse Roughness

- optimize roughness in virtual highlight regions
- semantics-based propagation

$$\mathcal{L}_{sp} = \sum_{c} \left| R - quantile(R \odot M_{vhl}(c), q) \right| \odot (M_{seg}(c) - M_{vhl}(c))$$

$$\mathcal{L}_{roughness} = |I - L_o| + \beta_{sp} \mathcal{L}_{sp}$$

virtual highlight regions



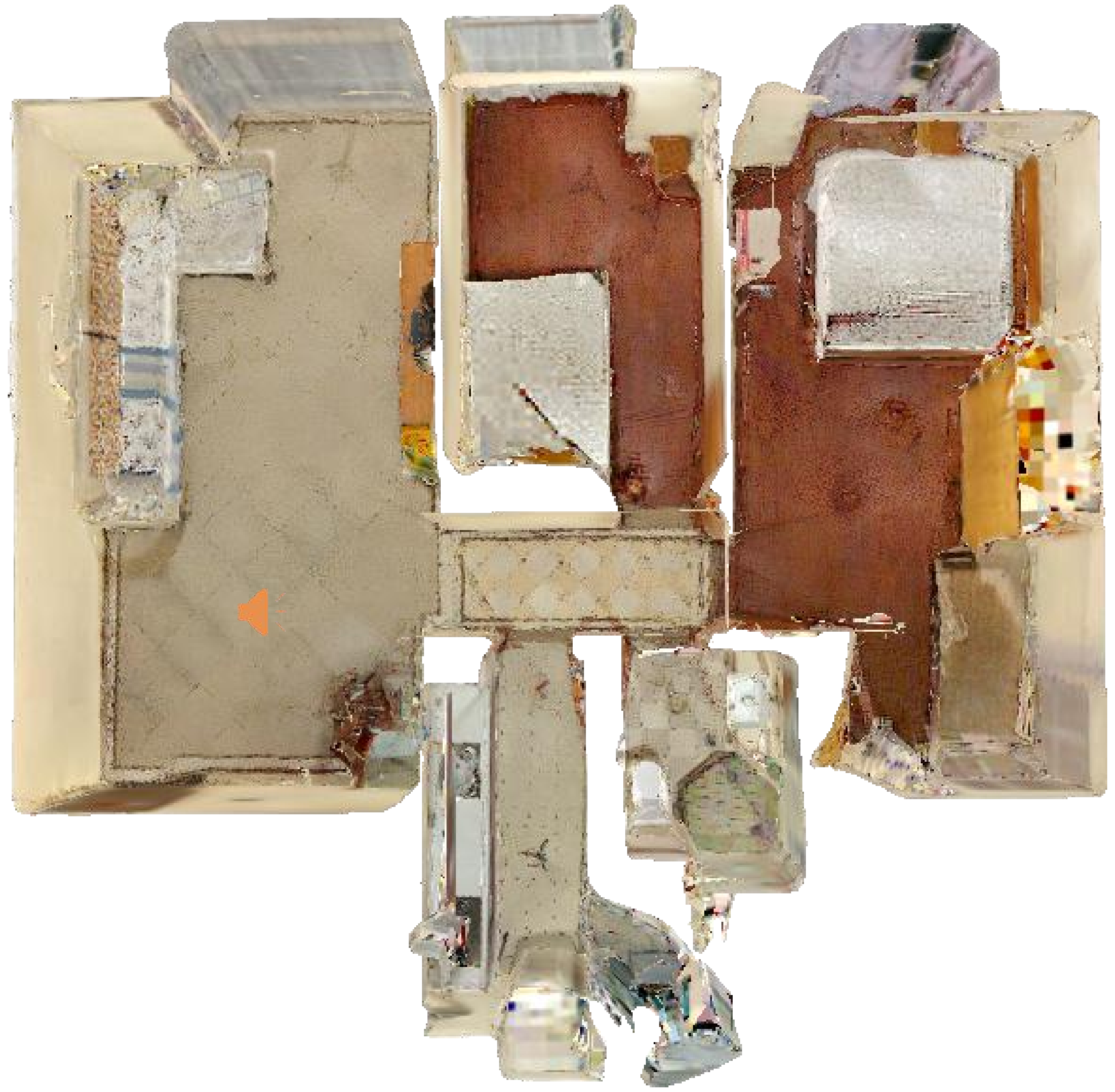
coarse roughness

Material Optimization -- Stage 3 Segmentation-based fine-tuning

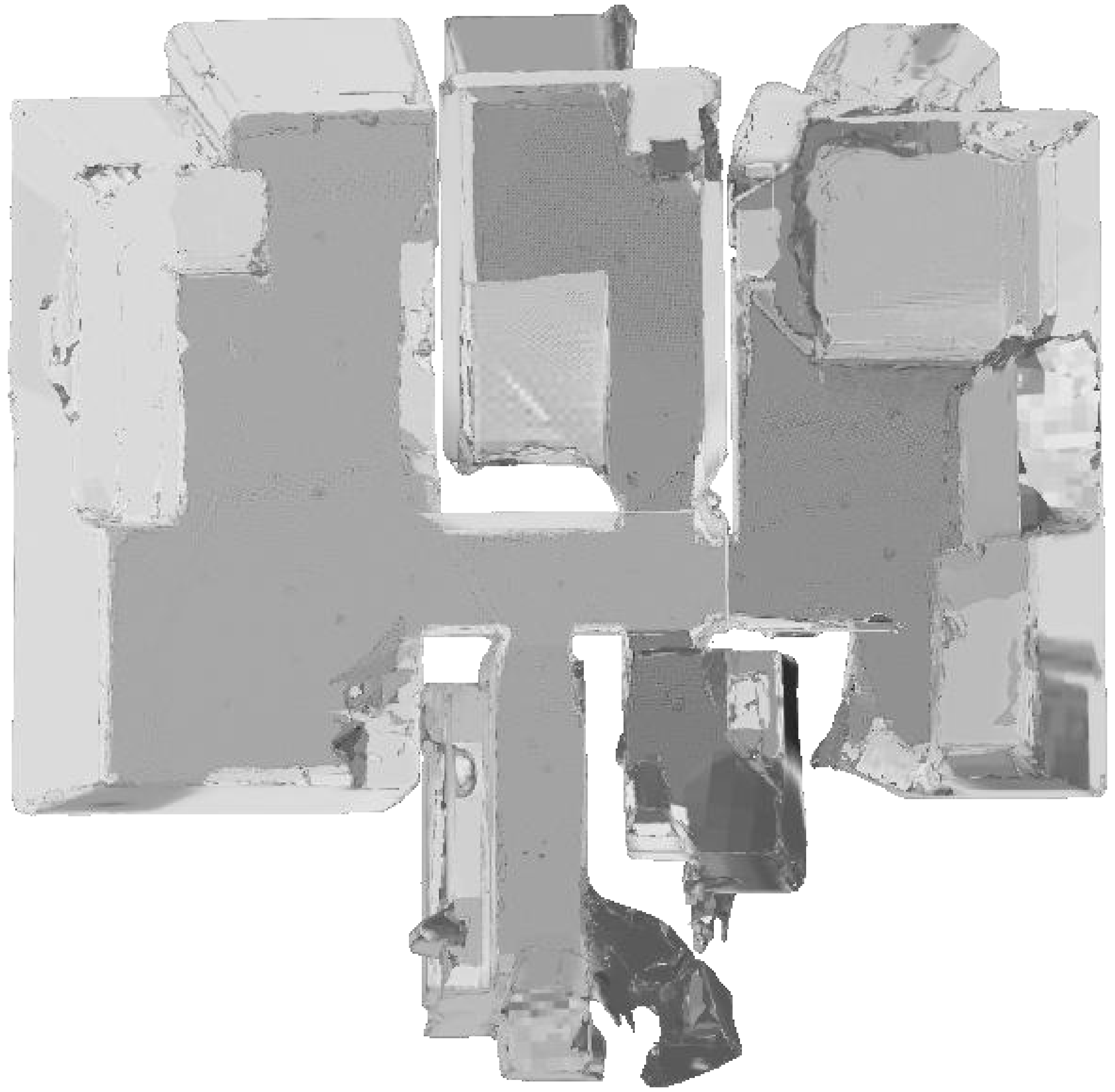
- fine-tune albedo and roughness
- Segmentation-based roughness smooth

$$\mathcal{L}_{rs} = \sum_{c} \left| R - \frac{\sum_p R \odot M_{room}(c)}{\sum_p M_{room}(c) + \epsilon} \right| \odot M_{room}(c)$$

$$\mathcal{L}_{all} = |I - L_o| + \beta_{ssr}(\mathcal{L}_{ss} + \mathcal{L}_{rs})$$

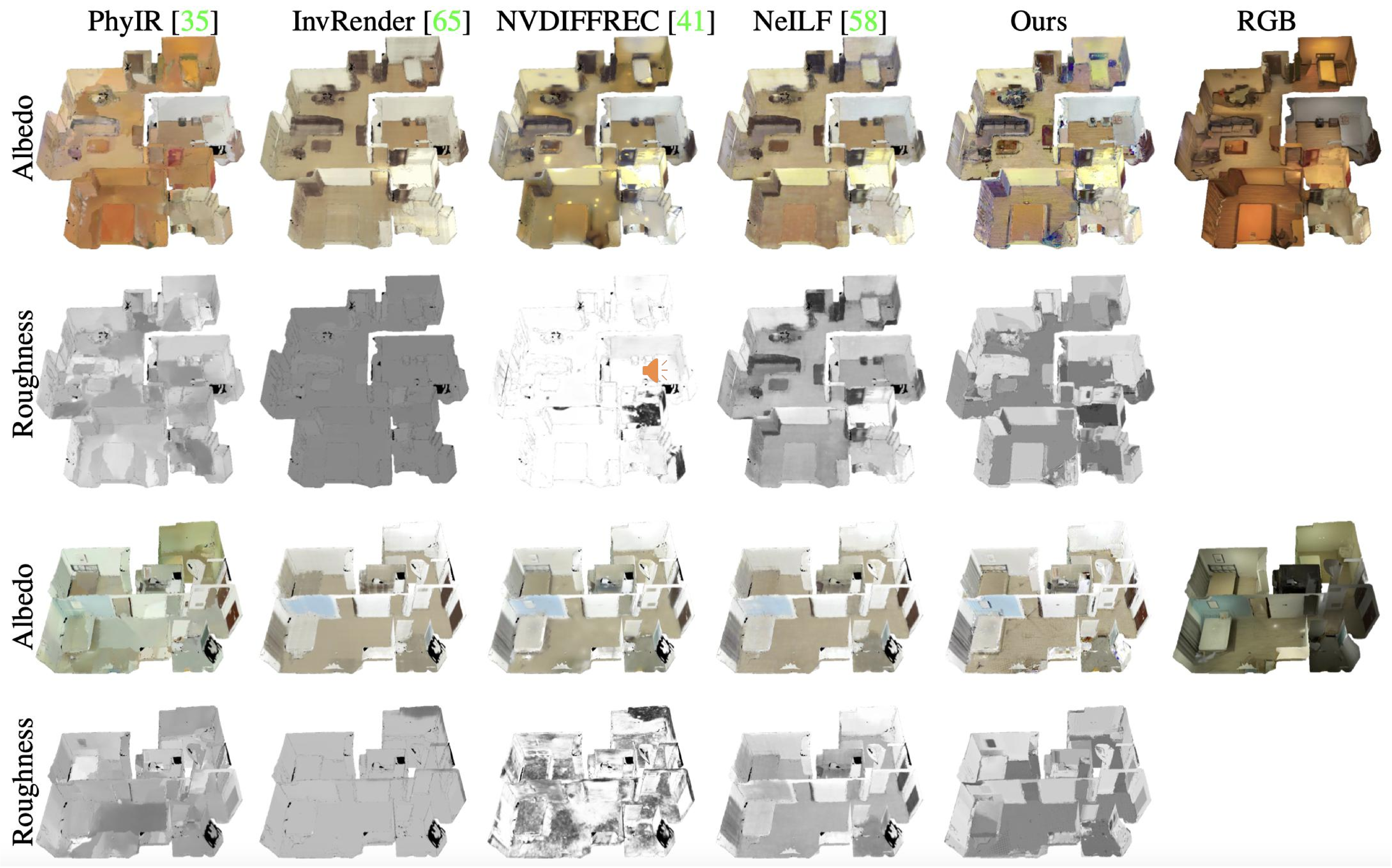


refined albedo texture

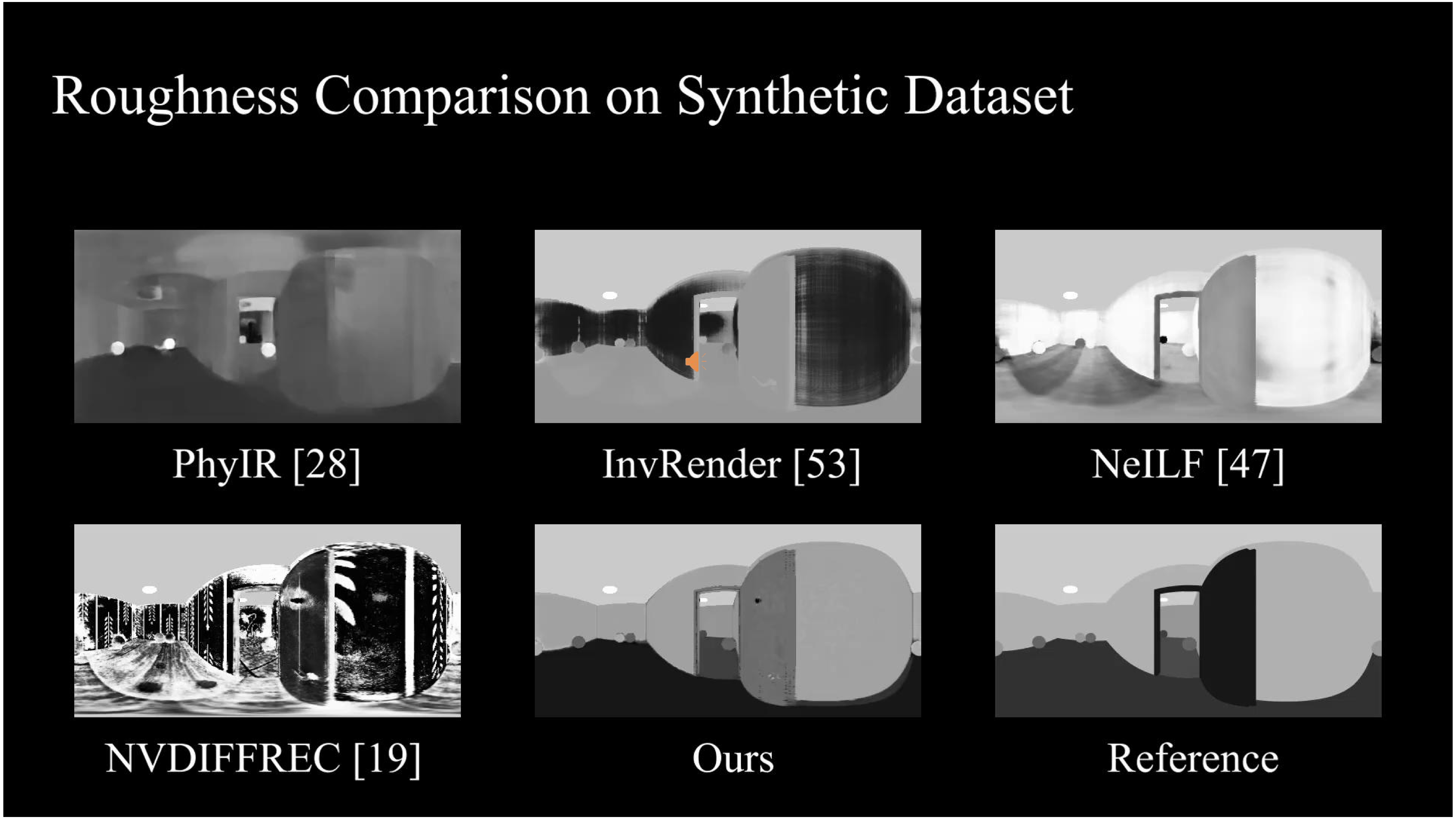


refined roughness texture

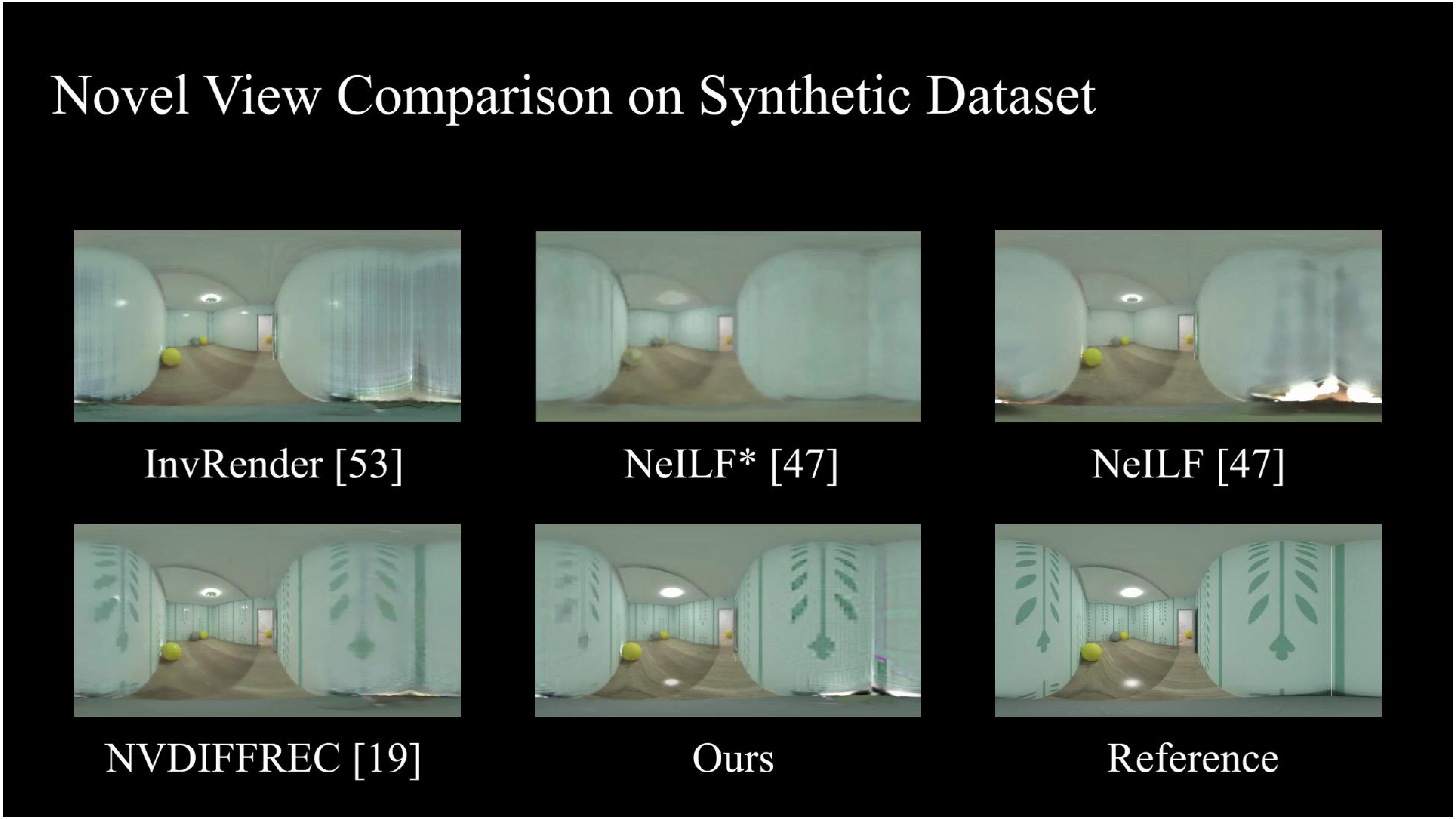
Comparisons -- Real Dataset



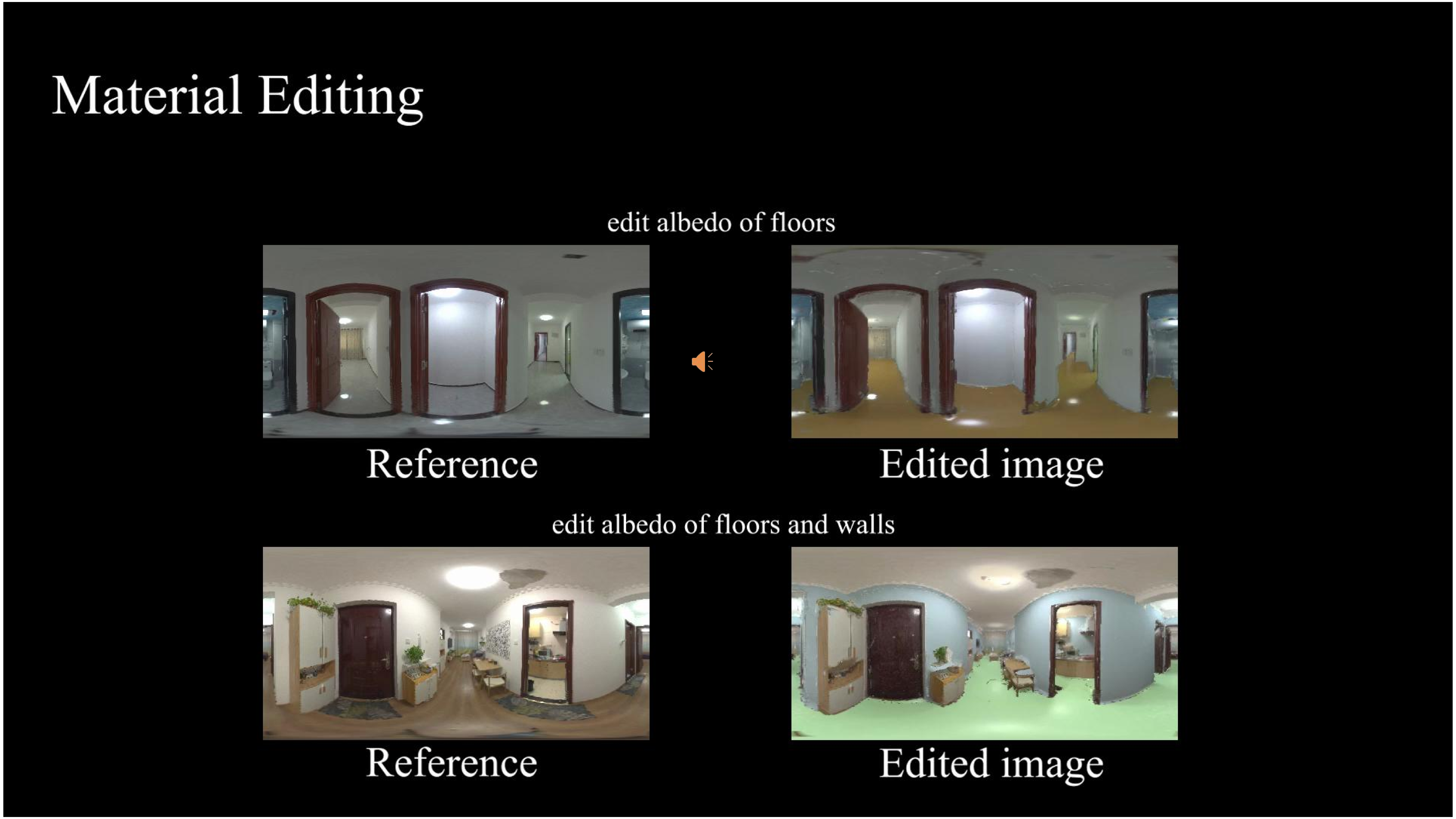
Comparisons



Comparisons



Applications



Applications

Relighting

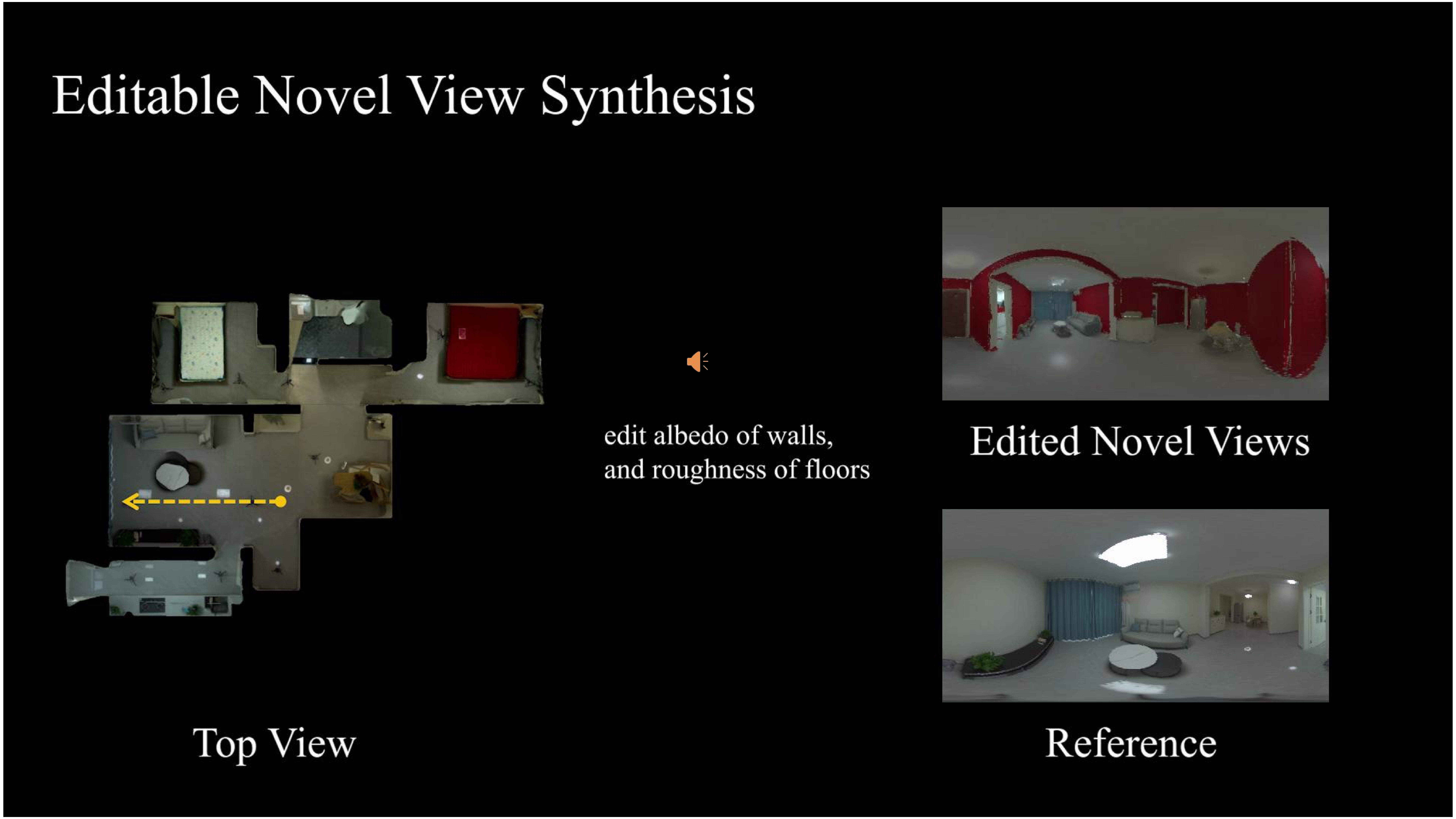
Reference

Relighted image

Reference

Relighted image

Applications



Thank you for watching

Multi-view Inverse Rendering for Large-scale Real-world Indoor Scenes

WED-PM-015

Zhen Li[1], Lingli Wang[1], Mofang Cheng[1], Cihui Pan[1], Jiaqi Yang[2]
[1]Realsee [2]Northwestern Polytechnical University