

Pano2Room

Novel View Synthesis from a Single Indoor Panorama

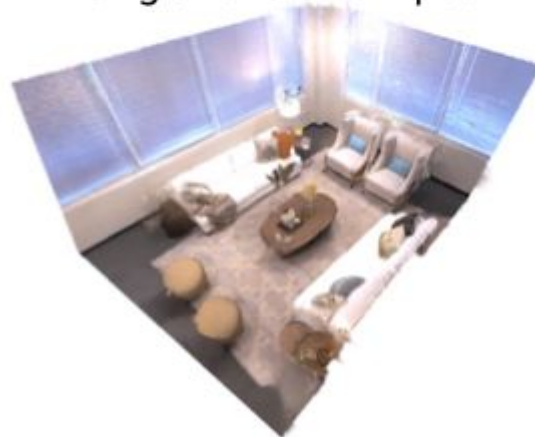
Bruno Mello

Motivação

- Métodos tradicionais de reconstrução 3D necessitam de muitas imagens
- Os métodos mais novos utilizando apenas uma imagem funcionam muito bem para objetos, mas não para cenas internas
- **Objetivo do paper:** Reconstruir uma cena interna com apenas uma imagem RGBD panorâmica



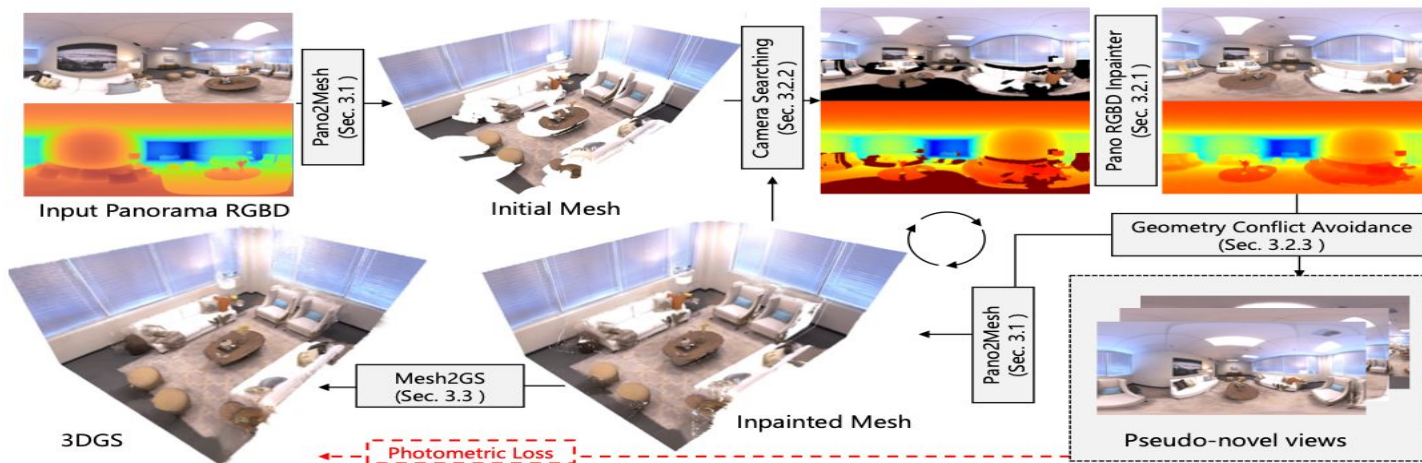
Single Panorama Input



Our Reconstructed 3DGS

Pipeline do Pano2Room

1. Converter panorama para malha inicial (Pano2Mesh)
2. **Refinar a malha usando inpainter RGBD panorâmico**
3. Converter malha final para 3D Gaussian Splatting (3DGS)
4. Treinar 3DGS com pseudo novas vistas coletadas



Pano2Mesh

- Para cada 4 vértices, cria duas faces
- Faz a projeção usando coordenadas esféricas

$$\phi = v/H \cdot \pi, \theta = u/W \cdot 2\pi - 0.5\pi,$$

$$x = \sin(\phi) \cdot \cos(\theta) \cdot D(\phi, \theta),$$

$$y = \cos(\phi) \cdot D(\phi, \theta),$$

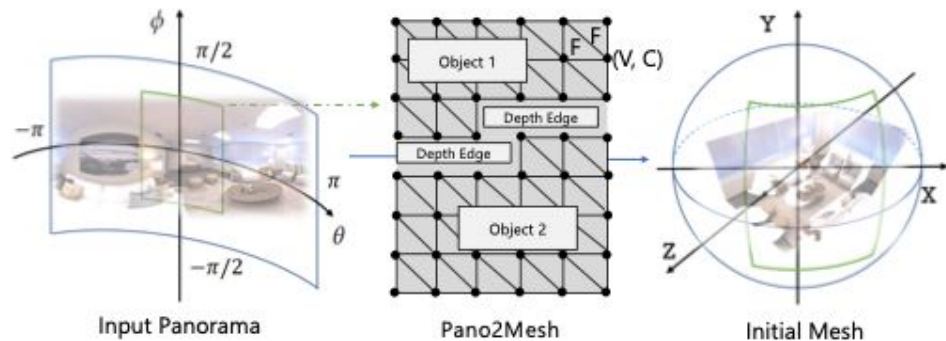
$$z = \sin(\phi) \cdot \sin(\theta) \cdot D(\phi, \theta),$$



Initial Mesh

Pano2Mesh

- Utiliza filtro baseado na distância para remover arestas de silhueta
- Não especifica exatamente como o filtro funciona mas menciona que é melhor que o do paper Text2Room, que considerava apenas o tamanho da aresta



Refinamento da malha

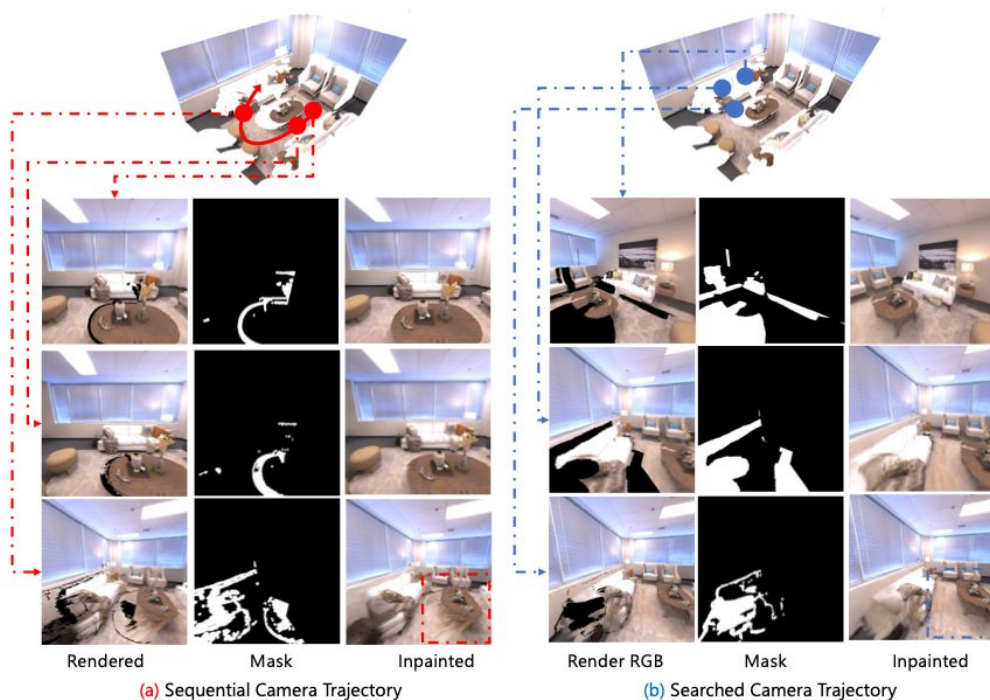
Método iterativo

1. Procura uma posição de câmera para realizar o inpainting
2. Renderiza a imagem panorâmica da malha atual nessa posição
3. Realiza o inpainting RGB e de Depth dessa panorâmica usando stable diffusion
4. Remoção de geometria conflitante
5. Adiciona nova malha

Refinamento da malha

Procura a melhor posição de câmera

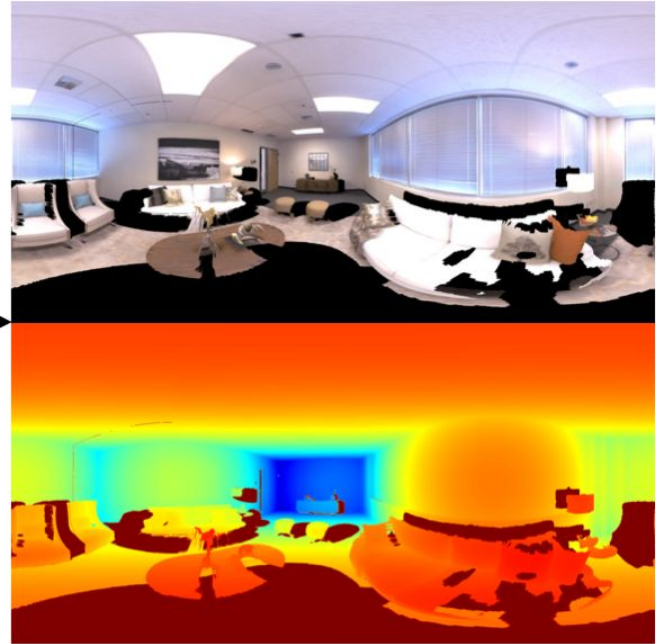
- A técnica utilizada pelo PERF não é ideal pois segue apenas uma trajetória definida
- Queremos minimizar o número de inpaintings necessários
- **Solução:** Encontrar a câmera que visualiza mais área ocluídas



Refinamento da malha

Renderização da panorâmica nessa posição

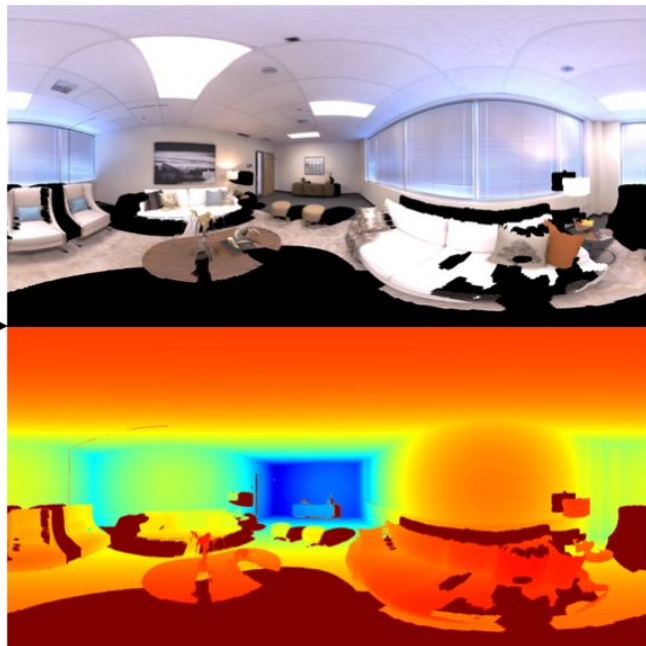
- Utilizando a malha, cria-se uma imagem panorâmica RGBD e uma máscara das áreas ocluídas



Refinamento da malha

Inpainting RGB usando stable diffusion

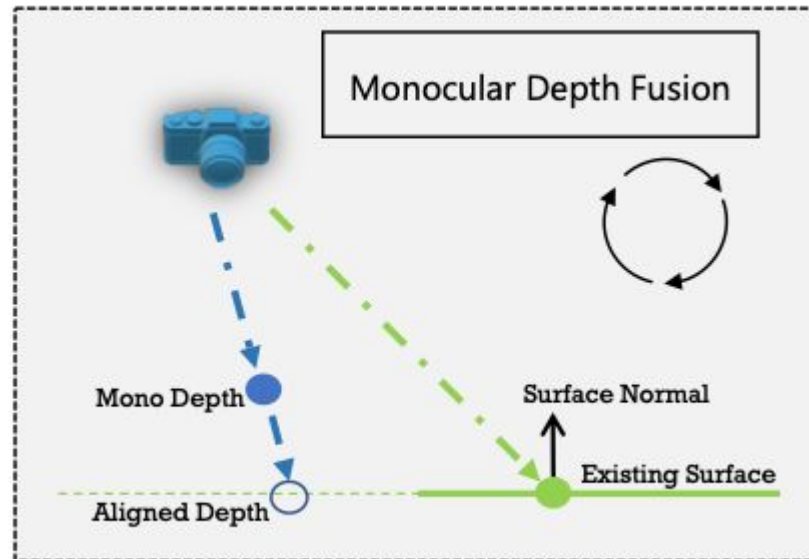
- Modelos pre treinados de stable diffusion utilizam imagens perspectivas
- **Solução:** Dividir a panorâmica em imagens perspectivas (k=20)
- Para que o modelo aprenda o estilo da cena, fazemos fine-tuning antes



Refinamento da malha

Inpainting Depth usando DPT

- Utiliza modelos pre treinados de cálculo de profundidade
- Para manter consistência geométrica, utiliza-se um modelo pre treinado de cálculo de normais para regularização



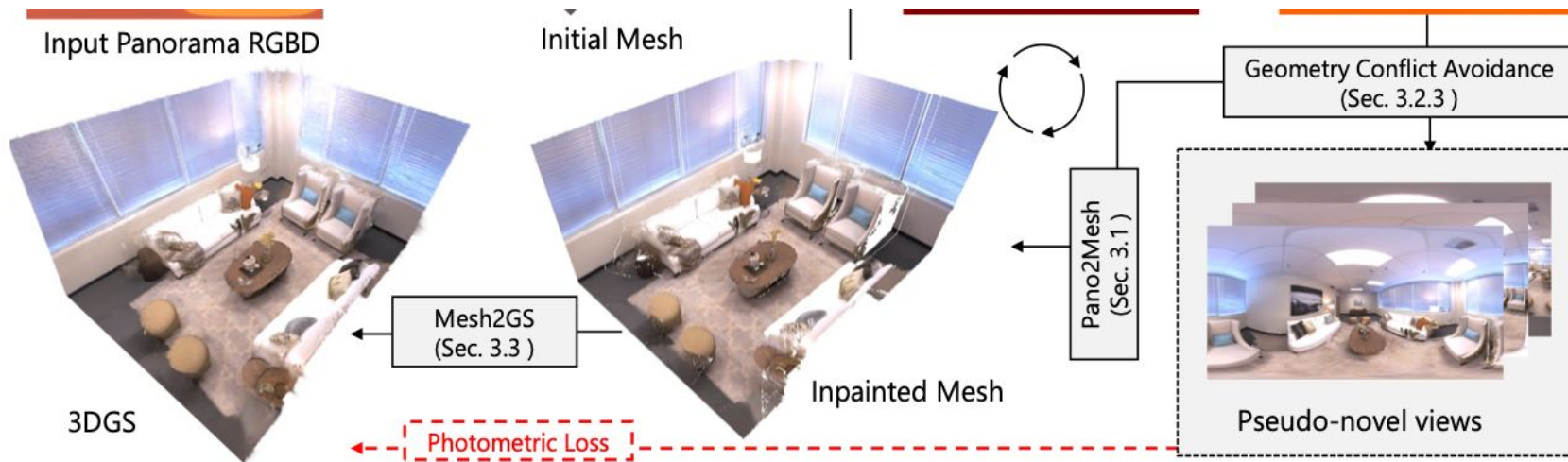
Refinamento da malha

Remoção da geometria conflitante

- Após criar a nova malha, renderiza-se ela de todas as vistas anteriores
- Se houver alguma face que não existia antes, remove-se esta face.
- Por fim, adicionamos o novo mesh ao nosso mesh atual e repetimos o processo

Mesh2GS

- A partir da malha final, inicializamos um 3DGS
- Utiliza-se para a função de perda as pseudo-novas vistas geradas no refinamento da malha



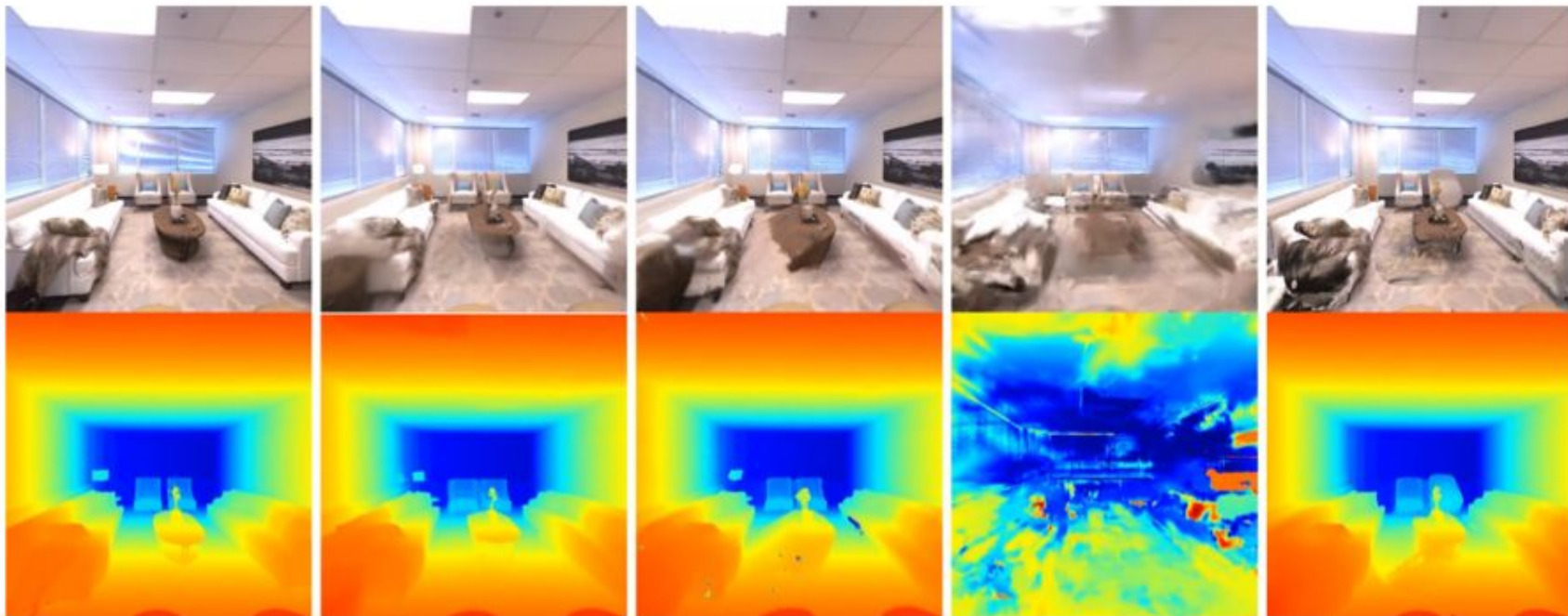
Comparações com métodos existentes

Comparou-se com os seguintes modelos:

- PERF: NeRF com panorâmica
 - Text2Room: Text-to-indoor-mesh
 - LucidDreamer: 3DGS a partir de uma imagem
-
- Como Text2Room e LucidDreamer recebem imagens perspectivas, usou-se cubemap como input
 - Utilizou-se o dataset “Replica”, que consiste de 8 cenas indoor

Comparações com métodos existentes

Comparação qualitativa



GT

Pano2Room

PERF

LucidDreamer

Text2Room

Comparações com métodos existentes

Comparação quantitativa

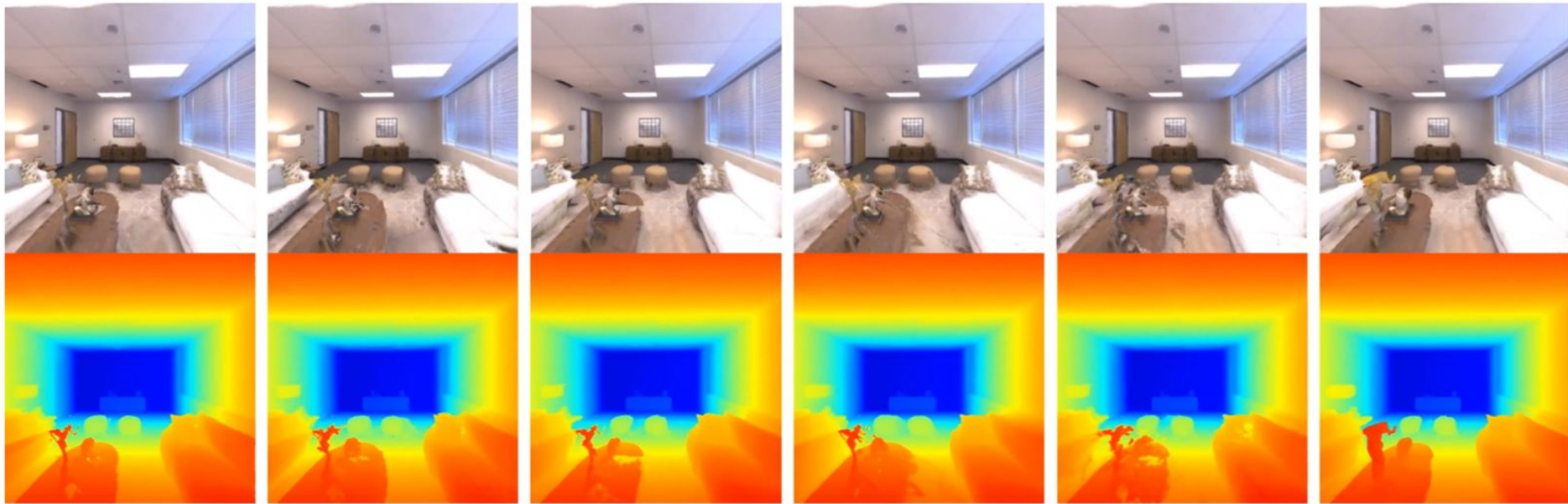
Scene Method	Pano2Room			PERF			Text2Room			LucidDreamer		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Room 0	23.29	0.788	0.168	22.58	0.767	0.167	21.55	0.779	0.151	17.50	0.611	0.470
Room 1	25.44	0.886	0.090	26.10	0.854	0.099	24.92	0.884	0.079	18.03	0.681	0.484
Room 2	25.06	0.874	0.127	23.75	0.858	0.121	22.21	0.862	0.138	17.40	0.719	0.456
Office 0	25.15	0.891	0.105	21.15	0.871	0.124	21.68	0.863	0.126	19.25	0.746	0.418
Office 1	31.77	0.953	0.043	30.98	0.938	0.070	23.92	0.935	0.078	20.91	0.788	0.332
Office 2	23.22	0.899	0.087	20.83	0.874	0.098	20.37	0.885	0.090	14.39	0.678	0.450
Office 3	20.93	0.858	0.143	19.69	0.825	0.159	15.77	0.789	0.232	13.98	0.679	0.446
Office 4	26.65	0.935	0.064	22.64	0.904	0.086	23.53	0.922	0.065	15.79	0.743	0.415

Comparações com métodos existentes

Comparação quantitativa

Scene Method	Pano2Room			PERF			Text2Room			LucidDreamer		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Room 0	23.29	0.788	0.168	22.58	0.767	0.167	21.55	0.779	0.151	17.50	0.611	0.470
Room 1	25.44	0.886	0.090	26.10	0.854	0.099	24.92	0.884	0.079	18.03	0.681	0.484
Room 2	25.06	0.874	0.127	23.75	0.858	0.121	22.21	0.862	0.138	17.40	0.719	0.456
Office 0	25.15	0.891	0.105	21.15	0.871	0.124	21.68	0.863	0.126	19.25	0.746	0.418
Office 1	31.77	0.953	0.043	30.98	0.938	0.070	23.92	0.935	0.078	20.91	0.788	0.332
Office 2	23.22	0.899	0.087	20.83	0.874	0.098	20.37	0.885	0.090	14.39	0.678	0.450
Office 3	20.93	0.858	0.143	19.69	0.825	0.159	15.77	0.789	0.232	13.98	0.679	0.446
Office 4	26.65	0.935	0.064	22.64	0.904	0.086	23.53	0.922	0.065	15.79	0.743	0.415

Ablation



Full

Sem SDFT

Sem SN

Sem CS

Sem GCA

Sem GS

Limitações

- Não funciona bem quando temos profundidade muito grande (como corredores por exemplo)
- Funciona apenas para cenas indoor
- Teste foi feito com dataset pequeno
- Depende muito de modelos pre treinados

Demo

