

Gaussian Splatting SLAM

Revisor: Daniel Morais

Arqueóloga: Fernanda

Hacker e Estudante de Doutorado: Thiago Baldivieso

Revisor Daniel Morais



Resumo

- Problema
- Principais ideias
- Contribuições e sua importância.

Problema da reconstrução SLAM

- SDF voxel grids: uso significativo de memória e resolução limitada;
 - Meshes: requerem topologia irregular para fundir novas informações;
 - Surfel clouds: são descontínuas e difíceis de fundir e otimizar;
 - Neural Fields: requerem raycasting por pixel para renderizar, que é computacionalmente caro.
-
- 3DGS não possui nenhuma dessas fraquezas. (... segundo os autores...)

Método

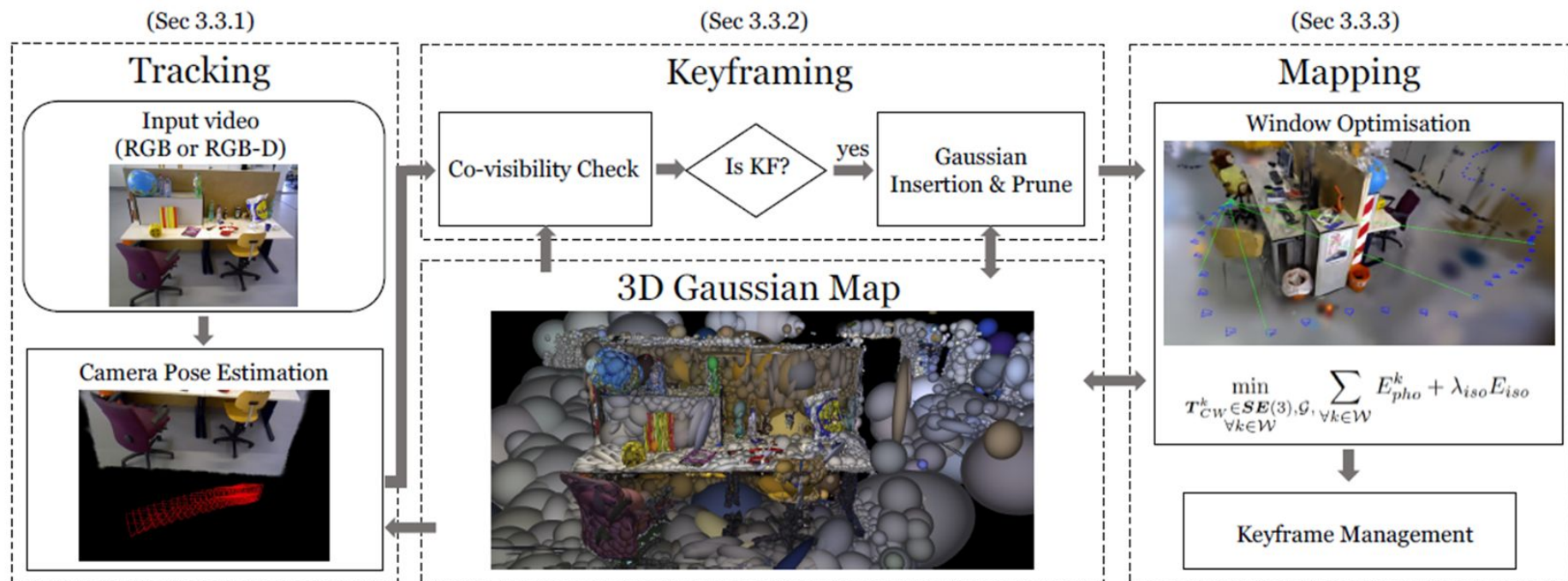


Figure 2. SLAM System Overview: Our SLAM system uses 3D Gaussians as the only representation, unifying all components of SLAM, including tracking, mapping, keyframe management, and novel view synthesis.

Método

Primeiro, é derivado o Jacobiano analítico no grupo de Lie da pose da câmera em relação a um mapa de Gaussianas 3D, para permitir que as poses da câmera sejam otimizadas junto com a geometria da cena – e isso pode ser integrado ao pipeline de rasterização diferenciável.

Colour & Opacity

$$C_p = \sum_{i \in \mathcal{N}} c_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j) .$$

3D Gaussians to 2D Gaussians

$$\mu_I = \pi(T_{CW} \cdot \mu_W) , \Sigma_I = \mathbf{J} \mathbf{W} \Sigma_W \mathbf{W}^T \mathbf{J}^T$$

Lie algebra & Lie Group

$$\frac{Df(T)}{DT} \triangleq \lim_{\tau \rightarrow 0} \frac{\text{Log}(f(\text{Exp}(\tau) \circ T) \circ f(T)^{-1})}{\tau}$$

2D Gaussian colour & Tcw camera pose (3D)

$$\frac{D\mu_C}{DT_{CW}} = [I \quad -\mu_C^\times], \frac{D\mathbf{W}}{DT_{CW}} = \begin{bmatrix} 0 & -\mathbf{W}_{:,1}^\times \\ 0 & -\mathbf{W}_{:,2}^\times \\ 0 & -\mathbf{W}_{:,3}^\times \end{bmatrix}$$

Método

Segundo, é feita uma nova regularização de forma isotrópica Gaussiana para garantir consistência geométrica, para a reconstrução incremental.

É uma forma de penalizar o parâmetro de escala si de modo a “encorajar” padrões esféricos e evitar Gaussianas alongadas.

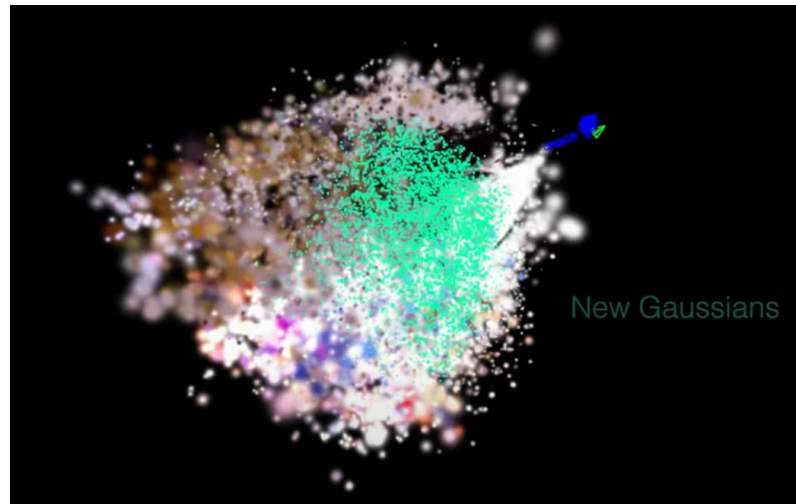
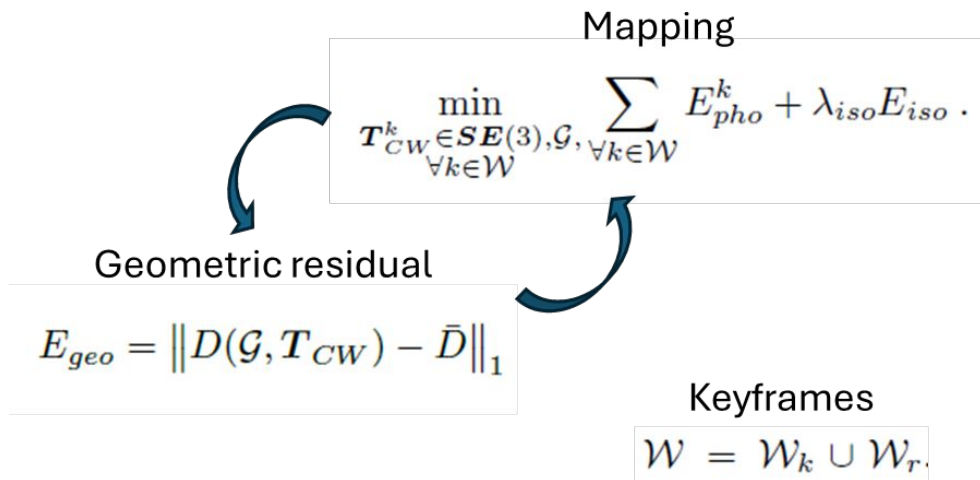
$$E_{iso} = \sum_{i=1}^{|\mathcal{G}|} \|s_i - \tilde{s}_i \cdot \mathbf{1}\|_1$$



Figure 3. **Effect of isotropic regularisation:** **Top:** Rendering close to a training view (looking at the keyboard). **Bottom:** Rendering 3D Gaussians far from the training views (view from a side of the keyboard) without (left) and with (right) the isotropic loss. When the photometric constraints are insufficient, the Gaussians tend to elongate along the viewing direction, creating artefacts in the novel views, and affecting the camera tracking.

Método

Terceiro, é utilizado um método de alocação de recursos e poda Gaussiana para manter a geometria limpa e permitir um rastreamento preciso da câmera.



Método



Pontos positivos

- Permite trabalhar com representações mínimas dos parâmetros de movimento (6 graus de liberdade para movimento 3D)
- Evita problemas de singularidades que podem ocorrer com outras representações
- Permitem representar rotações e transformações de forma computacionalmente eficiente
- Funciona bem com entrada monocular
- Unifica todo o pipeline de SLAM em uma única representação

Pontos negativos

- A qualidade final ainda não é perfeita, mas já é um grande avanço no campo
- A introdução de novas gaussianas a cada keyframe pode tornar o sistema lento
- Ainda não é capaz de lidar com cenas em larga escala.

Resultado

Artigo **Aprovado**

Arqueóloga
Fernanda



Ordem Cronológica que o paper se encaixa

Neste artigo, o GS-SLAM utiliza pela primeira vez a representação gaussiana 3D no sistema Simultaneous Localization and Mapping (SLAM).

Comparado aos métodos SLAM recentes que empregam representações neurais implícitas, o método utiliza um pipeline de renderização de splatting diferenciável em tempo real que oferece aceleração significativa para otimização de mapas e renderização RGB-D.

Ordem Cronológica que o paper se encaixa

Alcança **desempenho competitivo em comparação com métodos de última geração** em tempo real existentes nos conjuntos de dados Replica, TUMRGBD:

- Facilita um melhor equilíbrio entre eficiência e precisão.
- Estende a representação gaussiana 3D para reconstruir toda a cena, em vez de sintetizar um objeto estático nos métodos existentes.
- Otimiza a pose da câmera, resultando em redução do tempo de execução e estimativa robusta.

Ordem Cronológica que o paper se encaixa

1996: Volumetric method for building complex models from range images

2006, 2010, 2017: Simultaneous localization and mapping

2009: Parallel tracking and mapping on a camera phone.

2010: Simultaneous localization and mapping for augmented reality

2011: Real-time dense surface mapping and tracking.

2022: Visual slam algorithms and their application for ar, mapping, localization and wayfinding

2023: Efficient dense slam system based on hybrid representation of signed distance fields
3d gaussian splatting for real-time radiance field rendering.

2023: Flexible techniques for differentiable rendering with 3d gaussians.

2024: Implicit event-rgbd neural slam. (CVPR)

Artigo mais antigo citado pelo trabalho

Parallel tracking and mapping on a camera phone. ISMAR, 2009.

“Dense Visual SLAM. The existing real-time dense visual SLAM systems are typically based on discrete handcrafted features or deep-learning embeddings, and follow the mapping and tracking architecture in this paper”

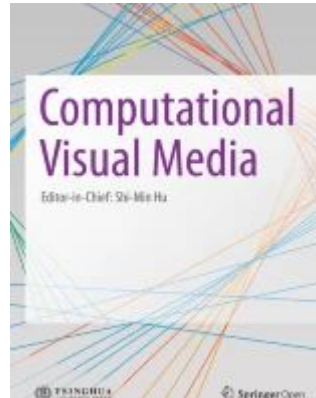
Parallel tracking and mapping on a camera phone. ISMAR, 2009.

“ This paper describes an attempt to implement a keyframe-based SLAM system on a camera phone (specifically, the Apple iPhone 3G).

We describe a series of adaptations to the Parallel Tracking and Mapping system to mitigate the impact of the device's imaging deficiencies. Early results demonstrate a system capable of generating and augmenting small maps, albeit with reduced accuracy and robustness compared to SLAM on a PC.”

Artigo mais recente que cita o atual

Recent advances in 3D Gaussian splatting



Vol. 10, No. 4, August 2024, 613–642

SPRINGER NATURE Link

Artigo mais recente que cita o atual

4.2.2 SLAM

“Further 3DGS methods target simultaneously localizing the camera and reconstructing the 3D scene. GS-SLAM [180] proposes an adaptive 3D Gaussian expanding strategy to add new 3D Gaussians into the training stage and delete unreliable ones with captured depths and rendered opacity values.”

Artigo mais recente que cita o atual

Table 4 Quantitative comparison of novel view synthesis results by different SLAM methods on the Replica [199] dataset using PSNR, SSIM, and LPIPS metrics

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
NICE-SLAM [200]	24.42	0.81	0.23
Vox-Fusion [201]	24.41	0.80	0.24
Co-SLAM [202]	30.24	0.94	0.25
GS-SLAM [180]	31.56	0.97	0.094
SplaTAM [181]	34.11	0.97	0.10
GaussianSplattingSLAM [182]	37.50	0.96	0.07
Gaussian-SLAM [183]	38.90	0.99	0.07
SGS-SLAM [185]	34.15	0.97	0.096

Papers relacionados no Google Scholar em 2024

De acordo com o Google Scholar, o artigo já foi citado 108 vezes.

Google Acadêmico

Pesquisar artigos de citação

Artigos

Aproximadamente 108 resultados (0,03 s)

A qualquer momento

Desde 2024

Desde 2023

Desde 2020

Período específico...

Ordenar por relevância

Ordenar por data

Em qualquer idioma

Pesquisar páginas em Português

✉ Criar alerta

Gs-slam: Dense visual slam with 3d gaussian splatting

☒ Pesquisar nos artigos de citação

Recent advances in 3d gaussian splatting

[T.Wu, Y.J.Yuan, L.X.Zhang, J.Yang, Y.P.Cao...](#) - Computational Visual ..., 2024 - Springer

The emergence of 3D Gaussian splatting (3DGS) has greatly accelerated rendering in novel view synthesis. Unlike neural implicit representations like neural radiance fields (NeRFs) ...

☆ Salvar 99 Citar Citado por 34 Artigos relacionados Todas as 3 versões

Scene Representations for Robotic Spatial Perception

[R.Mascaro, M.Chli](#) - Annual Review of Control, Robotics, and ..., 2024 - annualreviews.org

The ability of a robot to build a persistent, accurate, and actionable model of its surroundings through sensor data in a timely manner is crucial for autonomous operation. While ...

☆ Salvar 99 Citar Artigos relacionados

Aligning cyber space with physical world: A comprehensive survey on embodied ai

[Y.Liu, W.Chen, Y.Bai, X.Liang, G.Li, W.Gao...](#) - arXiv preprint arXiv ..., 2024 - arxiv.org

Embodied Artificial Intelligence (Embodied AI) is crucial for achieving Artificial General Intelligence (AGI) and serves as a foundation for various applications that bridge cyberspace ...

☆ Salvar 99 Citar Citado por 13 Artigos relacionados Todas as 2 versões

Sgs-slam: Semantic gaussian splatting for neural dense slam

[PDF] springer.com

[PDF] arxiv.org

[PDF] arxiv.org

Hacker Thiago Baldivieso



DEMO do paper

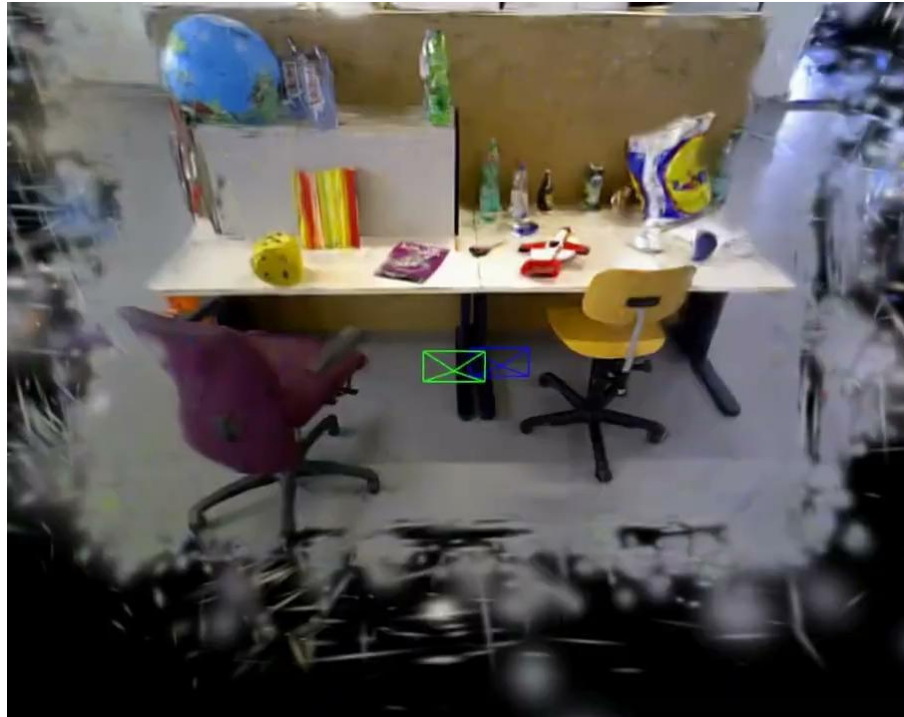
- Reprodutibilidade do método;
- Implementar uma pequena parte do artigo ou uma versão bem simplificada.
 - Por exemplo, 2D em vez de 3D;
- Rodar o código do artigo, analisar a implementação, utilizar outros datasets, e fazer experimentos adicionais.
- Focar em como as equações do paper foram implementadas.

Estudante de doutorado
Thiago Baldivieso



Proposta de projeto de pesquisa

- Aluno procurando um novo projeto na mesma área...



Gaussian Splatting Slam para uso outdoor e transmissão por Drones

Objetivo: Aplicar a reconstrução com GS-SLAM para aprimorar cenas outdoor com intuito de preservação ambiental e também em áreas com construções históricas.

- Edifício histórico Palacete da Babilônia (Casa Rosa) do CMRJ. +300 imagens de drones DJI: Mavic Pro, Mavic Air 2 e Spark. a) Nuvem de pontos; b) Nuvem densa; c) Objeto 3D



(a)



(b)



(c)

Aeronaves

Aeronaves DJI disponíveis no LIARC - IME



Fig 3: Mavic 3 Enterprise Thermal



Fig 4: Matrice 350 RTK



Fig 5: Mini 3 Pro



Fig 6: Mini 2 SE

Gaussian Splatting Slam para uso outdoor e transmissão por Drones

Dificuldades que podem ser encontradas: Latência de transmissão, treinamento em ambientes com pouca iluminação.

Fusão de dados heterogêneos

Discussão

Todos...