

Konzept

Hauptseminararbeit VCML

Johannes Decker

08.04.2025

Thema

Arbeitstitel:

RGB-based Simultaneous Localization and Mapping (SLAM): State of the Art

Beschreibung:

- Vergleich von zwei neuen Ansätzen für RGB-basiertes SLAM
- Wie kann eine 3D-Szene basierend auf Farbbildern möglichst realistisch und effizient dargestellt werden?

Literaturliste - Einstieg

How NeRFs and 3D Gaussian Splatting are Reshaping SLAM: a Survey

(Apr 2024) <https://doi.org/10.48550/arXiv.2402.13255>

- Untersucht umfassend die jüngsten Fortschritte von SLAM-Methoden (RGB-D, RGB, LiDAR)
- Schwerpunkt auf Ansätzen basierend auf Neuronalen Radianzfeldern (NeRFs) und 3D-Gaussian-Splatting (3DGS), da diese den aktuellen Stand der Technik für die Darstellung von 3D-Szenen sind

Literaturliste - Basisliteratur

NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis

(Aug 2020) <https://doi.org/10.48550/arXiv.2003.08934>

- Neuartige Methode zur Synthese fotorealistischer Ansichten komplexer 3D-Szenen durch das Training eines DNNs aus 2D-Bildern
- Eingabe: 5D-Koordinate bestehend aus 3D-Raumposition (x, y, z) und 2D-Blickrichtung $(\theta, \phi) \Rightarrow$ Lernen von $F: (x, y, z, \theta, \phi) \rightarrow (c, \sigma)$
- Ausgabe: Volumendichte und RGB-Farbe für diesen Punkt und diese Ansicht
- Durch Volumenrendering-Techniken werden diese Werte entlang der Kamerastrahlen integriert, um Bilder aus neuen Blickwinkeln zu generieren

Literaturliste - Basisliteratur

3D Gaussian Splatting for Real-Time Radiance Field Rendering (Aug 2023)

<https://doi.org/10.48550/arXiv.2308.04079>

- Neue Methode zur Ansichtssynthese ermöglicht Echtzeit-Rendering (≥ 30 FPS) in hoher visueller Qualität vgl. bisherige neuronale Rendering-Methoden
- Szenendarstellung mit 3D-Gauss-Verteilungen ermöglicht die flexible Darstellung komplexer Geometrie ohne neuronale Netze und deutlich schnelleres Training (Minuten statt Stunden)
- Ersetzt neuronale volumetrische Methoden durch einen kompakten, effizienten und hochpräzisen 3D-Gauß-basierten Ansatz

Literaturliste - RGB-SLAM-Ansätze

GIORIE-SLAM: Globally Optimized RGB-only Implicit Encoding Point Cloud SLAM (May 2024) <https://doi.org/10.48550/arXiv.2403.19549>

<https://github.com/zhangganlin/GIORIE-SLAM>

- Hochwertige 3D-Szenenrekonstruktion, -verfolgung und -rendering möglich
- GIORIE-SLAM verwendet deformierbare neuronale Punktwolke anstelle herkömmlicher dichter Gitter oder Voxelkarten
- Ermöglicht effiziente Kartenaktualisierungen und -deformationen während Schleifenschlüsse und Posenkorrekturen ohne erneutes Training auskommen

Literaturliste - RGB-SLAM-Ansätze

HI (hybrid implicit)-SLAM2: Geometry-Aware Gaussian SLAM for Fast Monocular Scene Reconstruction (Nov 2024)

<https://doi.org/10.48550/arXiv.2411.17982>

<https://github.com/Willyzw/Hi-SLAM2>

- Monokulares SLAM-System, das detaillierte 3D-Szenen ausschließlich mit RGB-Bildern rekonstruiert und dabei kein Tiefen-Signal benötigt
- Nutzt 3DGS für eine effiziente und präzise Szenendarstellung und ermöglicht so hochwertiges Rendering ohne Geschwindigkeitseinbußen
- Kombiniert online-Tracking und Loop-Closure mit offline-Verfeinerung, und gemeinsamer Optimierung von Karten- und Kamerapositionen

Datasets

- Replica dataset
- ScanNet
- eventuell TUM-RGBD
- und eventuell weitere Daten aus dem Projekt

“Smart Automobile Munich (SAM)”

=> Replica, ScanNet und TUM-RGBD enthalten größtenteils Innenraum-Szenen

Ziele bis zum nächsten Meilenstein

- Bis zum 2. Meilenstein sollen beiden SLAM-Ansätze auf mindestens 2 verschiedenen Datensätzen untersucht worden sein
- Des Weiteren soll eine Methode für den Vergleich der Performance von SLAM-Ansätzen definiert und implementiert werden

Fragen?