

ИССЛЕДОВАНИЕ ЗАДАЧИ ВИЗУАЛЬНОЙ ЛОКАЛИЗАЦИИ В НЕЯВНЫХ ПРЕДСТАВЛЕНИЯХ КАРТЫ

Егоров А. А.,

Научный руководитель – канд. техн. наук, доцент Ведяков А. А.

Университет ИТМО
vedyakov@niuitmo.ru

Введение

Неявное представление карт имеет ряд преимуществ в визуализации, плотности частиц и возможности постобработки. В частности, Neural Radiance Fields (NeRF) [1] приобрел значительную популярность. NeRF использует полносвязную нейронную сеть, обученную на коллекции монокулярных изображений. В работе Loc-NeRF [3] авторы представили метод для локализации в NeRF карте для использования неявного представления в задаче одновременной локализации и картирования. В силу достаточно большой ресурсоемкости обучения NeRF зачастую применяется подход с разделением задачи на офлайн картирование и онлайн локализацию. Так в PNeRFLoc [8] авторы представили метод локализации на основе карт признаков. Для обоих решений требуется достаточно точная инициализация с начальным приближением оценки позиции камеры

Основная часть

Задача локализации в трехмерном пространстве заключается в оценивании позиции камеры относительно окружения. Для неявных представлений карты, таких как NeRF, задача локализации может быть решена путем «инвертирования» модели. iNeRF [2] использует градиентную оптимизацию для минимизации фотометрических потерь между целевым изображением и рендером. В 2025 году был предложен гибридный метод GSplatLoc [7], комбинирующий структурное сопоставление ключевых точек на основе машинного обучения и фотометрическую оптимизацию на основе рендеринга. Для решения задачи одновременной локализации и картирования в реальном времени Loc-NeRF решает задачу локализации с помощью фильтра частиц для оценки апостериорного распределения позиции камеры методом Монте-Карло.

По результатам обзора было выявлено, что точность локализации сильно зависит от инициализации начального приближения. Современные подходы к инициализации можно разделить на методы, основанные на ключевых признаках, прямые методы и гибридные подходы. Классическим методом на основе ключевых точек является ORB [4]. Прямые методы, такие как DSO [8], минимизируют фотометрическую ошибку путем выравнивания изображений попиксельно. Для оптимизации вычислений используются пиксели с высоким градиентом интенсивности, вычисляемым оператором Собеля. Одним из современных методов выделения признаков на основе машинного обучения является XFeat [5], предложенный Г. Потье и др. в 2024 году. Далее было проведено экспериментальное сравнение вышеописанных методов для задачи инициализации метода представленном в Loc-NeRF. Методы были сравнены по метрикам скорости работы и точности оценки позиции камеры на датасете Replica [6].

Выводы

В ходе работы были исследованы методы локализации в неявных картах. Был проведен эксперимент на датасете Replica для оценки точности и скорости работы. Прямой метод [8] показал хорошие показатели точности оценки позиции при хорошем освещении и качестве рендеров. Метод на основе ключевых точек [4] показал лучшие

результаты по скорости работы. XFeat показал робастность к некачественным рендерам по точности оценки позиции и среднюю скорость работы.

Литература

1. Mildenhall B., Srinivasan P.P., Tancik M., [et al.]. Nerf: Representing scenes as neural radiance fields for view synthesis // *Communications of the ACM*. — 2021. — Vol. 65, no. 1. — P. 99–106.
2. Yen-Chen L., Florence P., Barron J.T., [et al.]. inerf: Inverting neural radiance fields for pose estimation // *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. — IEEE. 2021. — P. 1323–1330.
3. Maggio D., Abate M., Shi J., Mario C., Carlone L. Loc-nerf: Monte carlo localization using neural radiance fields // *arXiv preprint arXiv:209.09050*. — 2022.
4. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: An efficient alternative to SIFT or SURF // *2011 International conference on computer vision*. — IEEE. 2011. — P. 2564–2571.
5. Potje G., Cadar F., Araujo A., Martins R., Nascimento E.R. Xfeat: Accelerated features for lightweight image matching // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. — 2024. — P. 2682–2691.
6. Straub J., Whelan T., Ma L., [et al.]. The replica dataset: A digital replica of indoor spaces // *arXiv preprint arXiv:1906.05797*. — 2019.
7. Sidorov G., Mohrat M., Gridusov D., Rakhimov R., Kolyubin S. GSplatLoc: Grounding keypoint descriptors into 3D gaussian splatting for improved visual localization // *arXiv preprint arXiv:2409.16502*. — 2024.
8. Zhao B., Li M., Guo Z., [et al.]. Pnerfloc: Visual localization with point-based neural radiance fields // *Proceedings of the AAAI Conference on Artificial Intelligence*. — 2024. — Vol. 38, no. 7. — P. 7000–7007.

Егоров А.А. (автор)

Ведяков А.А. (научный руководитель)

Подпись

Подпись