

Оптимизация методов трехмерной реконструкции динамических сцен на основе голономных ограничений

Гридусов Д. Д.¹

Научный руководитель – доцент, доктор технических наук Колюбин С. А.¹

¹Университет ИТМО
gridden1@yandex.ru

Работа выполнена в рамках темы НИР №640105 «Физически-согласованные и семантические методы искусственного интеллекта для репрезентации динамических агентов и сред».

Введение

Задача трехмерной реконструкции является актуальной для большинства робототехнических приложений: от управления мобильными роботами [1, 2] до задач манипуляции [3]. В последние годы за счет компактности представления и фотореалистичного качества популярными стали методы на основе синтеза новых видов – NeRF [4] и 3D Gaussian Splatting [5]. Последний класс подходов за счет быстрой скорости работы чаще используется в практических приложениях, и поэтому рассматривается в данной работе. Обеспечивая высокое качество рендеринга, данные методы накладывают существенное ограничение – реконструируемая сцена должна содержать лишь неподвижные объекты. Естественным развитием стало создание алгоритмов, моделирующих движение каждой гауссианы, представляющей сцену [6, 7]. За счет добавления четвертой, временной, компоненты такие методы успешнее справляются с реконструкцией динамических сцен. Следующим шагом стала интеграция априорных знаний о механике в модель движения – это позволило добиться физически обоснованных траекторий движения объектов и улучшения качества синтеза новых видов [8, 9].

Однако большинство существующих подходов реконструкции динамических сцен моделируют движение каждой гауссианы отдельно, игнорируя голономные связи между ними. Частицы твердого тела двигаются согласованно, но в текущих подходах данное ограничение не учитывается. Подобное упущение замедляет скорость работы методов и может приводить к ухудшению качества синтеза новых видов из-за ошибок модели движения и дрейфа частиц.

Основная часть

В рамках исследования проверяется гипотеза о том, что учет ограничений на движение твердых тел способен привести к ускорению работы алгоритмов без потери качества реконструкции. В качестве базовых методов выбираются передовые подходы, такие как TRACE [9].

Предлагаемые модификации осуществляются в несколько этапов:

- добавление семантической информации на уровне каждого объекта (инстанс-сегментация);
- связь гауссиан, представляющих сцену с маской каждого объекта;
- построение прогноза движения для каждого твердого тела вместо каждой гауссианы.

Оценка результатов проводится по метрикам реконструкции (PSNR, SSIM, LPIPS) и скорости работы метода (FPS).

Выводы

Исследование поможет оценить эффективность учета ограничений на движение твердых тел в методах трехмерной реконструкции динамических сцен. Такие результаты особенно актуальны для робототехнических приложений, а также могут быть использованы для создания моделей мира.

Литература

1. Hadero B. T. et al. Beyond Implicit Representations: Exploring Gaussian Splatting for Next-Generation SLAM, Introduction and Review //IEEE Internet of Things Journal. – 2025.
2. Chen T. et al. Splat-nav: Safe real-time robot navigation in gaussian splatting maps //IEEE Transactions on Robotics. – 2025.
3. Ji M. et al. Graspplats: Efficient manipulation with 3d feature splatting //arXiv preprint arXiv:2409.02084. – 2024.
4. Mildenhall B. et al. Nerf: Representing scenes as neural radiance fields for view synthesis //Communications of the ACM. – 2021. – Т. 65. – №. 1. – С. 99-106.
5. Kerbl B. et al. 3d gaussian splatting for real-time radiance field rendering //ACM Trans. Graph. – 2023. – Т. 42. – №. 4. – С. 139:1-139:14.
6. Wu G. et al. 4d gaussian splatting for real-time dynamic scene rendering //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2024. – С. 20310-20320.
7. Xu J. et al. Grid4d: 4d decomposed hash encoding for high-fidelity dynamic gaussian splatting //Advances in Neural Information Processing Systems. – 2024. – Т. 37. – С. 123787-123811.
8. Hong H. et al. Physics-Informed Deformable Gaussian Splatting: Towards Unified Constitutive Laws for Time-Evolving Material Field //arXiv preprint arXiv:2511.06299. – 2025.
9. Li J., Song Z., Yang B. TRACE: Learning 3D Gaussian Physical Dynamics from Multi-view Videos //Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2025. – С. 8820-8829.