

РАЗРАБОТКА СИСТЕМЫ 3D ПАНОПТИЧЕСКОЙ СЕГМЕНТАЦИИ ДОРОЖНЫХ СЦЕН ДЛЯ АВТОНОМНОГО ТРАНСПОРТА НА ОСНОВЕ ДАННЫХ RGB-КАМЕР

Дженжеруха К. А.¹

Научный руководитель – докт. техн. наук, доцент Хитров Е. Г.¹

¹Университет ИТМО

kdzhenzherukha@gmail.com

Введение

Современные системы автономного транспорта сталкиваются с фундаментальной задачей – точной интерпретацией дорожной сцены в трёхмерном пространстве. Традиционно для этого используют лидарные сенсоры (LiDAR), обеспечивающие высокоточную геометрическую информацию. Однако их высокая стоимость и чувствительность к погодным условиям (дождь, снег, туман) делают подход непригодным для массового внедрения в системах автопилотов.

Camera-only подход, основанный исключительно на данных RGB-камер, позволяет существенно снизить стоимость систем восприятия. Однако он вводит принципиальные вызовы: неоднозначность глубины, сложность обработки окклюзий и восстановления полноценной 3D-структуры из 2D-изображений. Кроме того, для безопасного автономного вождения недостаточно просто восстановить геометрию сцены – система должна одновременно распознавать классы объектов (дорога, пешеход) и различать отдельные экземпляры (автомобиль 1, автомобиль 2 и пр.) Эту задачу решает интеграция паноптической сегментации, объединяющая преимущества семантической и инстанс-сегментации в едином 3D-представлении.

Основная часть

В данной работе рассматривается задача паноптической сегментации дорожных сцен [1] в 3D на основе изображений, снятых RGB-камерами автомобиля. В качестве основного набора данных используется датасет nuScenes Panoptic с полноценной 3D паноптической разметкой дорожных сцен. Целью работы является разработка эффективной системы 3D паноптической сегментации, способной работать исключительно на основе RGB-изображений без использования лидарных сенсоров при сохранении высокой точности геометрического восприятия. Для решения поставленных задач в работе рассмотрены современные архитектуры для получения 3D структуры из 2D, включая геометрические методы, к примеру, Lift-Splat-Shoot [2], трансформерные подходы (BEVFormer, DenseBEV) [3] и решения на основе разложения Гаусса, такие как GaussianCar [4]. Для практической реализации экспериментов использовался язык программирования Python и фреймворк PyTorch. Полученные результаты могут применяться при разработке экономически эффективных систем восприятия для автономного транспорта, в том числе с адаптацией под специфические условия эксплуатации, характерные для российской инфраструктуры.

Выводы

Проведен анализ современных методов паноптической сегментации дорожных сцен на основе RGB-изображений и подходов к реконструкции 3D структур из 2D изображений. Полученные результаты могут применяться для разработки экономически эффективных систем автономного транспорта. Разработанный пайплайн в дальнейшем планируется использовать для реализации исследований по оптимальному

планированию траекторий в рамках научно-исследовательской и диссертационной работы.

Литература

1. Yunge Li, Lanyu Xu. Panoptic Perception for Autonomous Driving: A Survey // arXiv:2408.15388. 2024. URL: <https://arxiv.org/abs/2408.15388> (дата обращения: 16.02.2026 г.).
2. Philion J., Fidler S. Lift, Splat, Shoot: Encoding Images from Arbitrary Camera Rigs by Implicitly Unprojecting to 3D // ECCV, 2020.
3. Li Z. et al. BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers // ECCV, 2022.
4. Santiago M.-M., Miguel A.-G., Fabio S.-G., A. Llamazares, H. Caesar, L. M. Bergasa: GaussianCaR: Gaussian Splatting for Efficient Camera-Radar Fusion // arXiv:2602.08784. 2026. URL: <https://arxiv.org/abs/2602.08784> (дата обращения: 16.02.2026 г.)