

ПРИМЕНЕНИЕ МОДЕЛЕЙ МИРА И ФУНДАМЕНТАЛЬНЫХ ПОДХОДОВ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ В РОБОТОТЕХНИКЕ

Сидоров Г.К.¹

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

¹Университет ИТМО

Введение

Современная робототехника сталкивается с фундаментальными ограничениями традиционных репрезентаций окружения при решении задач прецизионной манипуляции и навигации в неструктурированной среде. Отсутствие явной геометрической информации о трехмерной структуре сцены и динамике физических взаимодействий снижает эффективность автономных систем. В данной работе представлен обзор интеграции фундаментальных методов трехмерной реконструкции и генеративных моделей мира для создания геометрически согласованных представлений пространства.

Основная часть

Традиционные генеративные модели мира оперируют пространством пикселей, что ограничивает их применение в робототехнических задачах из-за отсутствия явной трехмерной геометрии и физической когерентности. Альтернативными являются подходы, объединяющие реконструкцию среды с предсказанием динамики в едином геометрическом пространстве. Подход на основе видеопоследовательностей цвета, глубины и нормалей поверхности с оптимизацией геометрии через интеграцию карт нормалей позволяет строить консистентные четырехмерные представления динамических сцен. Полученная реконструкция используется для обучения обратной динамической модели, непосредственно генерирующей управляющие воздействия для манипуляторов в задачах планирования и захвата [1]. Метод унификации состояния и действий в пространстве трехмерных потоков точек обеспечивает масштабируемое моделирование полной динамики сцены. Интеграция с прогнозирующим управлением позволяет выполнять перенос на реальные робототехнические платформы без дообучения для манипуляции даже с деформируемыми объектами в открытых средах [2]. Архитектура совместной оптимизации реконструкции и предсказания динамики сцены под действием управляющих воздействий и визуального планирования использует траектории камеры как геометрически информированное пространство действий. Это обеспечивает обобщение с синтетических данных на реальные условия без адаптации, одновременно решая задачи оценки глубины, позы камеры и планирования поведения [3].

Выводы

Интеграция методов трехмерной реконструкции с моделями мира обеспечивает робототехническим системам возможность пространственного планирования на основе явной геометрии сцены. Рассмотренные подходы позволяют преодолеть ограничения пиксельных представлений, обеспечивая физически корректное моделирование взаимодействий для задач автономной манипуляции.

Литература

1. Zhen H. et al. Tesseract: learning 4d embodied world models //arXiv preprint arXiv:2504.20995. – 2025.

2. Huang W. et al. PointWorld: Scaling 3D World Models for In-The-Wild Robotic Manipulation //arXiv preprint arXiv:2601.03782. – 2026.
3. Zhu H. et al. Aether: Geometric-aware unified world modeling //Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2025. – С. 8535-8546.

Сидоров Г.К. _____

Колюбин С.А. _____