

СОВРЕМЕННЫЕ МЕТОДЫ СЕМАНТИЧЕСКОГО КАРТОГРАФИРОВАНИЯ С ОТКРЫТЫМ СЛОВАРЕМ ДЛЯ МОБИЛЬНОЙ РОБОТОТЕХНИКИ

Юманов М.А. (ИТМО), Попов М.Ф. (ИТМО), Куркова Р.Е. (ИТМО)

Научный руководитель - доктор технических наук, профессор Колюбин С.А. (ИТМО)

Введение. Традиционные методы картографирования, ориентированные преимущественно на геометрическое представление пространства, оказываются недостаточными для обеспечения полноценного контекстного понимания сцены. В связи с этим, семантическое картографирование с открытым словарём, позволяющее идентифицировать и описывать широкий спектр объектов и их взаимосвязей, приобретает особую актуальность в области мобильной робототехники и автономных систем, в частности, в задачах навигации и мобильной манипуляции.

Основная часть. В рамках данной работы был проведён систематический обзор существующих методов семантического картографирования с открытым словарём, акцентируя внимание на подходах, комбинирующих трёхмерную семантическую сегментацию сцены и построение графа сцены. Ключевыми критериями оценки являлись точность семантической интерпретации и описаний объектов сцены, вычислительная эффективность алгоритмов (включая скорость обработки и потребление памяти). В результате сравнительного анализа были выявлены как преимущества, так и недостатки рассмотренных подходов, выявлены наиболее перспективные: BBQ[1], HOV-SG[2], Open3DSG[3], CLIO[4].

Выводы. Проведённый анализ методов семантического картографирования с открытым словарём позволил выделить ключевые направления развития и ограничения. Перспективными оказались подходы, использующие визуальные экстракторы перед мультимодальными моделями, иерархические графы для навигации, дистиллированные 3D-нейросети для моделирования отношений объектов, а также методы, учитывающие выполняемые задачи для определения гранулярности кластеризации. Однако у методов семантического картографирования по-прежнему есть проблемы с масштабируемостью, обработкой динамических сцен, вычислительной сложностью и интерпретируемостью. Выявленные преимущества и недостатки рассмотренных архитектур послужат основой для разработки усовершенствованных методов семантического картографирования.

Список используемых источников:

1. Linok S. et al. Beyond Bare Queries: Open-Vocabulary Object Retrieval with 3D Scene Graph //arXiv preprint arXiv:2406.07113. – 2024.
2. Werby A. et al. Hierarchical Open-Vocabulary 3D Scene Graphs for Language-Grounded Robot Navigation //First Workshop on Vision-Language Models for Navigation and Manipulation at ICRA 2024. – 2024.
3. Koch S. et al. Open3dsg: Open-vocabulary 3d scene graphs from point clouds with queryable objects and open-set relationships //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2024. – С. 14183-14193.
4. Maggio D. et al. Clío: Real-time Task-Driven Open-Set 3D Scene Graphs //arXiv preprint arXiv:2404.13696. – 2024.