

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ СЕГМЕНТАЦИИ ТРЕХМЕРНЫХ ОБЛАКОВ ТОЧЕК НА ОСНОВЕ RANSAC, HDBSCAN И ГИБРИДНОГО ПОДХОДА ДЛЯ СЕГМЕНТАЦИИ ЛИДАРНЫХ ДАННЫХ

В.В. Ковынев¹, К.И. Давыденко¹

¹*Белгородский государственный технологический университет им. В.Г. Шухова,
г. Белгород
e-mail: kovynnev_v@mail.ru*

Аннотация

В работе проводится сравнительный анализ трёх методов сегментации трёхмерных облаков точек: RANSAC для выделения плоскостей, HDBSCAN для плотностной кластеризации и разработанного гибридного алгоритма, сочетающего оба подхода. Эксперименты выполнены на реальных данных, полученных лидаром. Показано, что гибридный метод позволяет точнее разделять объекты со сложной геометрией по сравнению с чистыми методами.

Ключевые слова: сегментация облака точек, RANSAC, HDBSCAN, гибридный алгоритм, лидарные данные, сравнение методов.

Введение

Трёхмерные облака точек широко используются в системах компьютерного зрения, робототехники и промышленности. Одной из ключевых задач обработки таких данных является сегментация – разбиение неструктурированного множества точек на логические группы, соответствующие отдельным объектам или их частям [1, 2]. Существующие методы можно разделить на три большие категории: геометрические (например, RANSAC), методы кластеризации (например, HDBSCAN) и подходы на основе глубокого обучения. RANSAC эффективен для поиска параметрических примитивов, но не подходит для выделения произвольных форм. Алгоритмы плотностной кластеризации эффективно работают с кластерами произвольной формы, но чувствительны к выбору метрики и параметров.

Ранее нами был разработан гибридный алгоритм [3], последовательно применяющий RANSAC для выделения плоскостей и HDBSCAN с комбинированной метрикой для сегментации остаточных точек. В настоящей работе проводится количественное и визуальное сравнение этого гибридного подхода с чистыми методами RANSAC и HDBSCAN на реальных лидарных данных.

1. Методы сегментации

RANSAC предназначен для выделения параметрических моделей в зашумленных данных [4]. В работе используется классическая версия метода. После выделения всех плоскостей оставшиеся точки объединяются в один дополнительный сегмент.

HDBSCAN – эволюция DBSCAN, автоматически определяющая количество кластеров и устойчивая к шуму [5]. Для повышения качества разделения близко

расположенных поверхностей с разной ориентацией используется комбинированная метрика, учитывающая евклидово расстояние и различие нормалей.

Гибридный метод, предложенный в [3], работает в два этапа: сначала с помощью RANSAC выделяются доминирующие плоскости, затем HDBSCAN с комбинированной метрикой обрабатывает остаточные точки. Такой подход позволяет компенсировать недостатки каждого из вышеперечисленных методов.

2. Экспериментальная часть

2.1. Предобработка данных и параметры методов

Исходное облако точек, полученное с помощью лидара внутри помещения, содержит 24 178 точек. После предобработки осталось 9 108 точек:

- 1) воксельное понижение разрешения с размером вокселя 0,01 м;
- 2) статистическая фильтрация выбросов (число соседей – 10, порог – 1,0).

Параметры методов подбирались экспериментально для достижения наилучшего визуального качества сегментации. Для RANSAC установлены: порог расстояния до плоскости – 0,01 м, максимальное число выделяемых плоскостей – 10. Для HDBSCAN: минимальный размер кластера – 80, минимальное количество точек в окрестности – 45, вес координат – 1,5, вес нормалей – 2,0, число соседей для вычисления нормалей – 25. Для гибридного алгоритма на этапе RANSAC: порог – 0,01 м, максимальное число выделяемых плоскостей – 1; на этапе HDBSCAN: минимальный размер кластера – 135, минимальное количество точек в окрестности – 45, вес координат – 1,4, вес нормалей – 2,3, число соседей для вычисления нормалей – 40.

2.2. Количественные и визуализированные результаты

Вычислены три показателя: время выполнения (с), число выделенных сегментов и количество точек, отнесённых к шуму. Результаты представлены в таблице 1.

Таблица 1 – Сравнение производительности методов

Метод	Время (с)	Сегментов	Шум (точек)
RANSAC	0,04	11	0
HDBSCAN	1,02	10	3194
Гибридный	0,39	16	841

Гибридный метод занимает промежуточное время (0,39 с), выделяет наибольшее число сегментов (16) при количестве шума в 3,8 раза меньше, чем у HDBSCAN.

На рисунке 1 представлены результаты работы трёх алгоритмов.

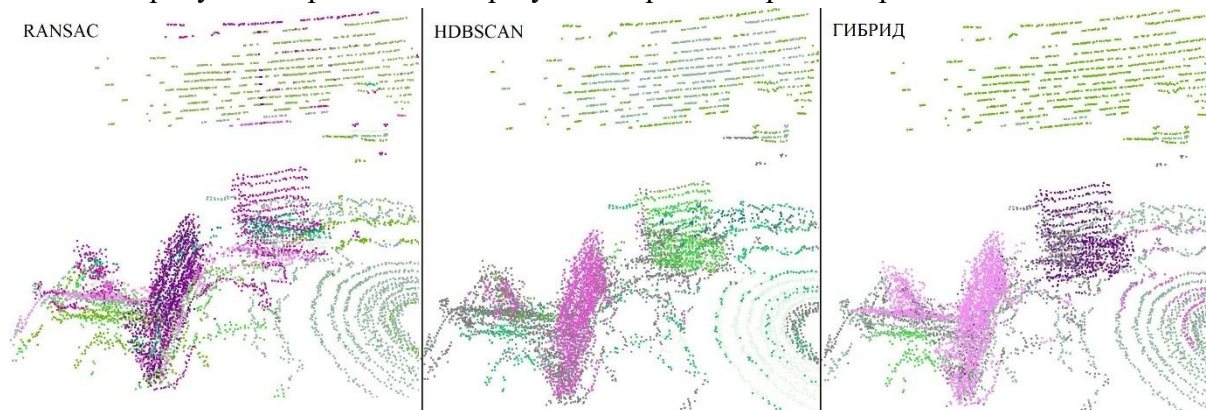


Рисунок 1 – Визуализация сегментации тремя методами

3. Анализ результатов

RANSAC демонстрирует эффективность на планарных участках сцены (например, плоскость пола), выделяя их в отдельные сегменты с четкими границами. Применительно к объектам сложной геометрии наблюдается эффект чрезмерной фрагментации. Криволинейные элементы (спинка, сиденье, ножки) аппроксимируются набором локальных плоскостей, что приводит к разбиению единого объекта на множество мелких несвязанных кластеров. Отсутствие механизма выделения шума приводит к включению выбросов в состав геометрических примитивов, искажая их форму.

HDBSCAN из-за использования плотностной кластеризации позволяет корректно объединять точки криволинейных поверхностей стула в единые семантические кластеры, сохраняя целостность объекта. Кроме того, значительная часть периферийных точек классифицируется как шум, что визуально проявляется в «разрежении» границ объектов.

RANSAC и HDBSCAN проявляют высокую чувствительность к неоднородности плотности сканирования, характерной для лидарных данных, имеющих полосчатую структуру. Это приводит к ложному разделению некоторых сплошных плоских поверхностей на несколько сегментов вдоль траектории сканирования.

Гибридный алгоритм [3] демонстрирует выделение доминирующих плоскостей методом RANSAC, что устраняет влияние крупных планарных структур на работу кластеризатора. В результате, этап HDBSCAN выполняется на остаточном облаке точек, содержащем преимущественно объекты сложной формы. Это позволяет сохранить семантическую целостность сложных объектов (части стула преимущественно выделяются как единая структура), исключить артефакты полосчатого сканирования на плоских поверхностях, существенно снизить объем ложно классифицированного шума за счет комбинированной метрики расстояния, учитывающей пространственную близость и ориентацию нормалей.

Таким образом, гибридный метод обеспечивает наиболее качественное структурное представление сцены, минимизируя недостатки, присущие каждому из базовых алгоритмов при изолированном использовании. Основным ограничением остаётся необходимость настройки параметров, однако эксперименты показали возможность их подбора в широких диапазонах.

Заключение

В работе проведено сравнение трёх методов сегментации трёхмерных облаков точек: RANSAC, HDBSCAN и разработанного гибридного алгоритма [3]. Эксперименты на реальных лидарных данных показали, что гибридный метод превосходит чистые подходы по качеству выделения объектов со сложной геометрией, обеспечивая большее число осмысленных сегментов и меньшее количество шума при приемлемом времени вычислений.

Благодарности

Работа выполнена в рамках реализации федеральной программы поддержки университетов «Приоритет 2030» с использованием оборудования на базе Центра высоких технологий БГТУ им. В.Г. Шухова.

This work was realized in the framework of the Program «Priority 2030» on the base of the Belgorod State Technological University named after V.G. Shukhov. This work was realized using equipment of High Technology Center at BSTU named after V.G. Shukhov.

Список использованных источников

1. Rabbani, T. Segmentation of point clouds using smoothness constraint / T. Rabbani, F. van den Heuvel, G. Vosselman // International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2006. – V. 36, № 5. – P. 248–253.
2. Fischler, M. A. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography / M. A. Fischler, R. C. Bolles // Communications of the ACM. – 1981. – V. 24, № 6. – P. 381–395.
3. Ковынев, В. В. Гибридный алгоритм сегментации трехмерных облаков точек для объектов со сложной структурой / В. В. Ковынев, П. С. Кабелянц // Информационные системы и технологии: научный журнал Орловского государственного университета им. И.С. Тургенева. – Принято к печати 24 марта 2026 г.
4. A density-based algorithm for discovering clusters in large spatial databases with noise / M. Ester [et al.] // Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. – 1996. – P. 226–231.
5. Campello, R. J. G. B. Density-based clustering based on hierarchical density estimates / R. J. G. B. Campello, D. Moulavi, J. Sander // Advances in Knowledge Discovery and Data Mining: 17th Pacific-Asia Conference, PAKDD 2013. – Springer, 2013. – P. 160–172.

COMPARATIVE ANALYSIS OF POINT CLOUD SEGMENTATION METHODS BASED ON RANSAC, HDBSCAN AND A HYBRID APPROACH FOR LIDAR DATA SEGMENTATION

V.V. Kovynev¹, K.I. Davydenko¹

*¹Belgorod State Technological University named after V.G. Shukhov, Belgorod
e-mail: kovynev_v@mail.ru*

Abstract

The paper presents a comparative analysis of three methods for 3D point cloud segmentation: RANSAC for plane extraction, HDBSCAN for density-based clustering, and a developed hybrid algorithm combining both approaches. Experiments are performed on real lidar data. It is shown that the hybrid method allows more accurate separation of objects with complex geometry compared to pure methods.

Keywords: point cloud segmentation, RANSAC, HDBSCAN, hybrid algorithm, lidar data, method comparison.