

DOI: 10.15514/ISPRAS-2025-37(6)-24



Сравнение классических и нейросетевых алгоритмов выделения ключевых точек на изображениях пересеченной местности для применения в SLAM алгоритмах

П.А. Ухов, ORCID: 0000-0002-3728-2262 <ukhov79@gmail.com>

М.Б. Булакина, ORCID: 0009-0008-8354-0294 <maria.b.bulakina@gmail.com>

С.С. Крылов, ORCID: 0000-0003-3267-6411 <compgra@yandex.ru>

*Московский авиационный институт (национальный исследовательский университет),
Россия, 125993, г. Москва, Волоколамское шоссе, д. 4.*

Аннотация. В работе предложена метрика для оценки качества работы алгоритмов выделения ключевых точек на изображении в условиях пересеченной местности при отсутствии однозначно определяемых ориентиров и углов. Проведено сравнение различных алгоритмов выделения ключевых точек для последующей реализации в составе SLAM алгоритма на борту беспилотного летательного аппарата. Получены значения предложенной метрики для популярных алгоритмов выделения ключевых точек и другие параметры на основе запуска решения в контролируемом окружении. Показаны преимущества алгоритмов на основе моделей машинного обучения.

Ключевые слова: алгоритмы выделения ключевых точек на изображении; алгоритмы SLAM; алгоритм SIFT; алгоритм SURF; алгоритм ORB; алгоритм SuperPoint; алгоритм R2D2; метод LightGlue; машинное обучение.

Для цитирования: Ухов П.А., Булакина М.Б., Крылов С.С. Сравнение классических и нейросетевых алгоритмов выделения ключевых точек на изображениях пересеченной местности для применения в SLAM алгоритмах. Труды ИСП РАН, том 37, вып. 6, часть 2, 2025 г., стр. 123–130. DOI: 10.15514/ISPRAS-2025-37(6)-24.

Comparison of Classical and Machine Learning Algorithms for Feature Point Extraction in Rugged Terrain Images for Application in SLAM Algorithms

P.A. Ukhov, ORCID: 0000-0002-3728-2262 <ukhov79@gmail.com>

M.B. Bulakina, ORCID: 0009-0008-8354-0294 <maria.b.bulakina@gmail.com>

S.S. Krylov, ORCID: 0000-0003-3267-6411 <compgra@yandex.ru>

Moscow Aviation Institute (National Research University)

4, Volokolamskoe highway, Moscow, 125993, Russia.

Abstract. The paper proposes a metric for evaluating the performance of feature point extraction algorithms in rough terrain conditions with no clearly defined landmarks or corners. Various feature point detection algorithms are compared for subsequent integration into a SLAM algorithm on board an unmanned aerial vehicle (UAV). The proposed metric, along with other algorithm parameters, is evaluated through experiments conducted in a controlled environment. The advantages of algorithms based on machine learning models are demonstrated.

Keywords: feature point detection algorithms; SLAM; SIFT; LightGlue; SURF; ORB; SuperPoint; R2D2; machine learning.

For citation: Ukhov P.A., Bulakina M.B., Krylov S.S. Comparison of classical and machine learning algorithms for feature point extraction in rugged terrain images for application in SLAM algorithms. *Trudy ISP RAN/Proc. ISP RAS*, vol. 37, issue 6, part 2, 2025, pp. 123-130 (in Russian). DOI: 10.15514/ISPRAS-2025-37(6)-24.

1. Введение

В последнее время с учетом недоступности глобальных навигационных систем в ряде регионов Российской Федерации и увеличившимися рисками для беспилотных летательных аппаратов (далее БПЛА) в связи с нарушением навигационных полей все большее значение приобретают задачи автономной навигации. Для БПЛА небольшого размера на текущем уровне развития технологий инерциальных навигационных систем невозможно обеспечение требуемой точности для выполнения миссий. При этом требования к БПЛА растут и необходимы системы обеспечивающие отклонения БПЛА от курса не более 5-10 метров на дистанциях более 10 км. В данных условиях одними из перспективных методов автономной навигации становятся методы на основе систем технического зрения [1].

С учетом данных требований наиболее перспективными для изучения являются методы одновременной локализации и построения карты на основе технического зрения, известные как алгоритмы SLAM (simultaneous localization and mapping). Основополагающей работой в области SLAM является исследование Смита и Чизмана по представлению и оценке пространственной неопределенности, проведенное в 1986 году [2]. Представленный метод сейчас обобщен до шести степеней свободы, и позволяет оценить взаимосвязь (положения и ориентации) между объектами, а также оценить неопределенности, связанные с этими взаимосвязями.

Для построения карт местности применяются упрощения на основе топологических карт на основе дескрипторов, фиксирующих топологию среды [3]. Данные технологии хорошо отработаны для автономных наземных транспортных средств в городской среде. Однако, для пересеченной и лесной местности, а также в применении к БПЛА многие алгоритмы работают неустойчиво как для извлечения ключевых точек из изображения, так и для последующего описания дескрипторов для SLAM алгоритма.

Традиционные алгоритмы выделения ключевых точек на изображении, такие как SIFT (Scale-Invariant Feature Transform), SURF (Speeded Up Robust Features) и ORB (Oriented FAST and

Rotated BRIEF), основываются на математических и геометрических принципах. Эти алгоритмы используют различные техники для обнаружения ключевых точек на изображении, которые являются инвариантными к изменениям масштаба, поворота и освещения. Традиционные алгоритмы выделения ключевых точек зарекомендовали себя как надежные инструменты для обработки изображений. Однако, с развитием технологий машинного обучения и нейросетевых подходов появляются новые возможности для повышения точности и устойчивости этих алгоритмов, особенно с учетом необходимости запуска данных алгоритмов на борту БПЛА в условиях ограниченных вычислительных ресурсов.

В качестве базового алгоритма для сравнения подходов к выделению ключевых точек был выбран визуальный SLAM (V-SLAM) [4], использующий изображения, полученные с камер видимого диапазона излучений.

2. Оценка алгоритмов извлечения ключевых точек на изображении

Для сопоставления алгоритмов необходимо определить ключевые характеристики, которыми должен обладать алгоритм извлечения ключевых точек на изображениях, чтобы его можно было успешно адаптировать под специфические требования и условия эксплуатации БПЛА на пересеченной, в том числе лесной, местности. По итогам проведенного анализа были выбраны алгоритмы, обладающие высокими показателями эффективности и широко применяемые в различных системах: SIFT [5], SURF [6], ORB [7], SuperPoint [8] и R2D2 [9]. Они демонстрируют хорошее качество обнаружения ключевых точек, устойчивость к шумам и изменениям в окружающей среде, однако вынуждены балансировать между точностью и способностью быстро обрабатывать данные в реальном времени.

При необходимости использовать алгоритм на борту беспилотника, важно, чтобы он работал в режиме реального времени, особенно если БПЛА выполняет задачи навигации или обнаружения объектов. Соответственно быстрота при извлечении и сопоставлении ключевых точек критически важна для применения алгоритмов на борту БПЛА. Поэтому далее введем метрику для оценки качества алгоритмов и проведем оценку скорости их работы.

Каждый алгоритм извлечения ключевых точек связан с соответствующими алгоритмами описания дескрипторов и последующего их сопоставления, поэтому особый интерес представляет комплексное сравнение данных алгоритмов по времени выполнения и качеству полученных сопоставлений. Для решения этой задачи был подготовлен тестовый набор данных.

2.1 Исходные данные для оценки

Для оценки качества работы алгоритмов был подготовлен набор данных по реальным полетам БПЛА над пересеченной местностью с учетом различных условий: времени суток (день, утро, вечер), времен года, освещения и погодных аномалий (солнечно, переменная облачность, облачно, слабый дождь и др.). Выборка составила более 300 полетов продолжительностью от 10 до 30 минут. Некоторые специфические условия съемки приведены на рис. 1. Набор данных содержит различные виды подстилающей поверхности от полей и перелесков до лесной местности. Во всех случаях наличие объектов искусственного происхождения минимально (в основном редкие сельские строения и дороги).

Для съемки использовались БПЛА мультироторного типа (преимущественно квадрокоптеры) с трехосевым подвесом и преимущественно с камерой SIYI ZR10. Полетное задание варьировалось, однако высота составляла порядка 100 метров, скорость в среднем от 5 до 10 м/с, а также был фиксирован ракурс съемки – камера направлена под углом 30 градусов к вертикали, направления камеры выбирались разными – от совпадающего с курсом полета

БПЛА до направления в сторону (углы 0, 45 и 90 градусов). Частота съемки составила 30 кадров в секунду.

Записи полетов БПЛА были разбиты по кадрам. Каждому кадру соответствовали определенные координаты БПЛА в системе WGS84, а также точные координаты подвеса камеры и параметры съемки (фокусное расстояние, время электронного затвора и др.). Указанные данные позволили формировать связанные наборы изображений с большим межкадровым смещением. В настоящее время набор данных постоянно пополняется новыми съемками полетов.



*Рис. 1. Примеры различных условий съемки:
с длинными тенями на закате (слева) и в зимний период (справа).*

*Fig. 1. Examples of shooting conditions:
with long shadows at sunset (left) and in winter (right).*

2.2 Оценка скорости работы алгоритмов

Для оценки скорости работы алгоритмов использовался микрокомпьютер Orange Pi 3 LTS 8Гб с gpu Mali-T720. Все вычисления запускаются на чистой ОС Linux Debian 3.0.8 с установленными зависимостями для выполнения соответствующих алгоритмов. Изображения считываются с встроенной карты памяти.

Записи полетов беспилотника были разбиты по кадрам в случайном порядке из расчета расстояния между крайними кадрами по времени в 1 секунду. Была составлена выборка из 100 000 последовательностей кадров в случайном порядке. Пример пары крайних кадров последовательности приведен на рис. 2.



Рис. 2. Пример двух изображений тестовой выборки (смещение беспилотника в горизонтальной плоскости порядка 10 м, шаг по времени 1 секунда, высота полета 100 м, угол наклона камеры от вертикали 30 градусов, камера направлена по курсу полета).

Fig. 2. An example of two dataset images (horizontal drone displacement - 10 m, time step - 1 second, flight altitude - 100 m, camera tilt angle from the vertical - 30 degrees, camera directed along the flight course)..

Для каждой из 100 тыс. последовательностей кадров производилось извлечение и сопоставление точек для каждого из методов по 10 раз и вычислялось среднее время,

затраченное на исполнение алгоритмов по всей выборке. Сопоставление извлеченных признаков осуществлялось методом *k*-ближайших соседей из библиотеки FLANN [10-11], а также был рассмотрен нейросетевой метод сопоставления дескрипторов LightGlue для алгоритмов SIFT и SuperPoint. Полученные средние значения времени исполнения алгоритмов на экстракцию ключевых точек и их сопоставление, представлены в табл. 1.

Табл. 1. Результаты сравнения скорости выполнения алгоритмов.

Table 1. Execution time for different algorithms.

Алгоритмы	SIFT	SURF	ORB	SuperPoint	R2D2	SIFT u LightGlue	SuperPoint u LightGlue
Время, мс.	237	188	81	101	293	184	54

По итогам сравнения можно сделать вывод, что нейросетевые методы при условии использования соответствующих аппаратных ускорителей показывают наиболее быстрые результаты. Так, наименьшее время работы показала комбинация двух нейросетевых методов – SuperPoint для извлечения ключевых точек и LightGlue для сопоставления дескрипторов. Следующим по скорости выполнения оказался алгоритм ORB, что ожидаемо, так как скорость работы и нетребовательность к вычислительным ресурсам являются главными преимуществами данного традиционного алгоритма. Медленнее всех по итогам тестирования оказались алгоритмы SIFT и нейросеть R2D2, которые требуют больше вычислительных затрат при извлечении и описании ключевых признаков.

Однако, помимо скорости выполнения важно обеспечить качество полученных дескрипторов для последующих операций построения карты и сравнения в рамках SLAM алгоритма. Для оценки качества полученных ключевых точек предлагается подход на основе метрики прослеживаемости ключевых точек между кадрами.

2.3 Оценка прослеживаемости ключевых точек между кадрами

Помимо скорости работы, при полете БПЛА для качественной навигации важно чтобы алгоритм обладал таким свойством, как способность правильно идентифицировать и отслеживать одну и ту же ключевую точку на большом количестве последовательных кадров, несмотря на смещение изображения. Для оценки прослеживаемости ключевых точек между кадрами предлагается использовать следующую метрику:

$$M_n = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{c_{0,i}}{N_i(1-\alpha)}$$

где: M_n – усредненная доля правильно сопоставленных ключевых точек на окне из n кадров, n – количество последовательно сопоставляемых кадров, $c_{0,i}$ – количество ключевых точек, успешно сопоставленных между начальным и i -тым кадром, N_i – общее количество выделенных на i -том кадре ключевых точек, α – доля площади поверхности, на которую различается i -тый и начальный кадры без учета рельефа местности.

Данная метрика направлена на оценку продолжительности сопоставления ключевых точек на последовательности кадров, что позволяет более точно оценить эффективность алгоритмов в динамичных условиях пересеченной местности для различных условий.

Для коэффициента α использованы упрощения – так как полеты выполняются в основном в равнинной местности, либо с небольшими перепадами рельефа, то поверхность можно представить в виде плоскости и вычислить разницу площадей между кадрами в упрощенной постановке.

Например, рассмотрим равномерное прямолинейное движение беспилотника над равнинной местностью. Если камера направлена вдоль оси Z , а оси X и Y определяют плоскость изображения, то с помощью библиотеки *georu* и метаданных тестовых полетов можно

рассчитать, сколько пролетает БПЛА, для того чтобы заснять полностью новый кадр. а в общем случае получить направление камеры в новом положении. Коэффициентами дисторсии в нашем случае можно пренебречь, т.к. изображения сохранены уже после соответствующих преобразований.

Предположим, что скорость БПЛА составляет 1 м/с, высота 100 метров, тогда из данных телеметрии определяется смещение между двумя соседними кадрами, например 0.96 м, изменение площадей получим через проецирование изображения камеры на поверхность Земли. Исходя из данных вычислений $\alpha = 0.0018$. То есть между двумя последовательными кадрами, совпадает примерно 98.2% площади изображенной на кадре поверхности Земли. Или по-другому, площадь, покрываемая на текущем кадре, отличается от площади на предыдущем снимке на 1.8%.

Для сравнительного анализа традиционных и нейросетевых подходов были взяты серии кадров по ограничению на коэффициент α не более 30% различия площади. На данных сериях кадров были оценены алгоритмы выделения ключевых точек по метрике M_n на собранных изображениях пересеченной местности. Были рассмотрены приведенные выше методы.

Кроме ограничений, связанных с пересечением площадей реальной поверхности Земли, некоторые алгоритмы имеют настраиваемые параметры, одним из наиболее влияющих на качество последующей работы SLAM алгоритмов является ограничение по количеству извлекаемых точек на изображении. Поэтому помимо измерения метрик на наборе данных необходимо было выбрать оптимальные параметры для каждого из алгоритмов. Для этого для рассматриваемых методов при различных ограничениях максимального количества извлекаемых точек подсчитывалась предложенная метрика M_5 . На 200 изображениях для каждого окна из 5 кадров вычислим значение метрики, а затем найдём среднее по всем вычислениям. Получившиеся результаты представлены на рис. 3.

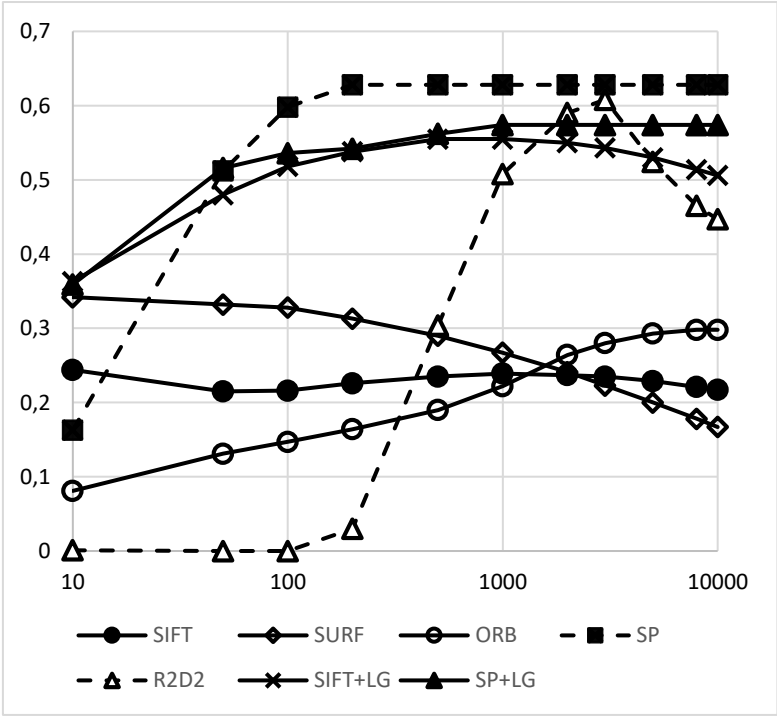


Рис. 3 Значение метрики M_5 в зависимости от ограничения числа извлекаемых ключевых точек.
Fig. 3. M_5 value depending of the extracted key points number limit.

На рис. 3 можно заметить, что самыми высокими значениями метрики обладает нейросетевая модель SuperPoint. Хотя моделью и извлекается не так много точек (табл. 2), но они являются устойчивыми и продолжительно сопоставляются на последовательности кадров. Если вместо метода k ближайших соседей из библиотеки FLANN использовать для сопоставления извлеченных признаков нейросеть LightGlue, то эксперимент тоже показывает достаточно хорошие результаты, при этом после фильтрации остается больше хорошо отслеживаемых ключевых точек, по сравнению с SuperPoint + KNN (табл. 2). Алгоритм SIFT показывает несколько худший результат в сочетании с LightGlue, а нейросеть R2D2 показывает высокое значение на ограничении порядка 3000 извлекаемых ключевых точек.

Сравнение показывает, что традиционные способы извлечения и описания ключевых точек значительно проигрывают нейросетевым подходам в случае, когда важно извлекать устойчивые признаки, которые будут продолжительно отслеживаться на последовательности кадров одной динамической сцены. По итогам сравнения различных методов извлечения ключевых точек и сопоставления полученных дескрипторов для реализации на борту БПЛА в качестве компонента SLAM алгоритма можно рекомендовать комбинацию двух нейросетевых подходов *SuperPoint* и *LightGlue*. Кроме этого, указанные решения могут быть дообучены для обеспечения большей точности извлечения ключевых точек и сопоставления дескрипторов. Также на основе проделанной работы была оценена скорость выполнения алгоритмов, что поможет в дальнейшем выбрать наиболее подходящий бортовой вычислитель для системы автономной навигации БПЛА.

Табл. 2. Реальное количество извлекаемых точек при различном значении ограничения.

Table 2. Actual number of points extracted for different limit values.

N	SIFT	SURF	ORB	SuperPoint	R2D2	SIFT + LightGlue	SuperPoint + LightGlue
10	10	10	10	10	10	7	10
50	50	50	50	49	50	34	50
100	100	100	100	88	100	70	100
200	200	200	200	128	200	142	200
500	500	500	500	129	500	362	495
1000	1000	1000	1000	129	1000	734	733
2000	2000	2000	2000	129	2000	1485	734
3000	3000	3000	3000	129	3000	2247	734
5000	5000	5000	5000	129	5000	3787	734
8000	8000	8000	8000	129	8000	6126	734
10000	10000	10000	10000	129	10000	7544	734

Список литературы / References

- [1]. Али Б., Садеков Р.Н., Цодокова В.В. Алгоритмы навигации беспилотных летательных аппаратов с использованием систем технического зрения. Гироскопия и навигация. Том 30. №4 (119), 2022, стр.87-105. DOI 10.17285/0869-7035.00105.
- [2]. Smith R.C., Cheeseman P. On the Representation and Estimation of Spatial Uncertainty. The International Journal of Robotics Research. 5 (4): 56–68. 1986. DOI: 10.1177/027836498600500404.
- [3]. Cummins M., Newman P. FAB-MAP: Probabilistic localization and mapping in the space of appearance. The International Journal of Robotics Research. 27 (6): 647–665, 2008. DOI:10.1177/0278364908090961.
- [4]. Yalan Chen, Yimin Zhou, Qin Lv, Kranthi Kumar Deveerasetty A Review of V-SLAM. Conference: 2018 IEEE International Conference on Information and Automation (ICIA). 2018. DOI: 10.1109/ICInfA.2018.8812387.

- [5]. Lowe D.G. Distinctive Image Features from Scale-Invariant Key Points. *International Journal of Computer Vision*, 2004, 60, 91-110. DOI 10.1023/B:VISI.0000029664.99615.94.
- [6]. Bay H., Tuytelaars T., Van Gool, L. SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds) *Computer Vision – ECCV 2006*. ECCV 2006. *Lecture Notes in Computer Science*, vol 3951. Springer, Berlin, Heidelberg. DOI 10.1007/11744023_32.
- [7]. Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski, Gary Bradski ORB: an efficient alternative to SIFT or SURF. *Conference: IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*. DOI: 10.1109/ICCV.2011.6126544.
- [8]. Daniel DeTone, Tomasz Malisiewicz, Andrew Rabinovich SuperPoint: Self-Supervised Interest Point Detection and Description. *Conference: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2018. DOI: 10.1109/CVPRW.2018.00060.
- [9]. Jerome Revaud, Philippe Weinzaepfel, César De Souza, Martin Humenberger. R2D2: Repeatable and Reliable Detector and Descriptor. *Conference on Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 8-14 December, 2019. DOI 10.48550/arXiv.1906.06195.
- [10]. Marius Muja and David G. Lowe, "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration", in *International Conference on Computer Vision Theory and Applications (VISAPP'09)*, Vol.1, 2009.
- [11]. Библиотека FLANN [Электронный ресурс] // GitHub. – 2009. – Режим доступа: <https://github.com/flann-lib/flann> (дата обращения: 04.05.2025).

Информация об авторах / Information about authors

Петр Александрович УХОВ – кандидат технических наук, доцент, заместитель начальника управления «ИТ-Центр» с 2020 года. Сфера научных интересов: машинное обучение, компьютерное зрение, программное обеспечение беспилотных авиационных систем, предиктивная аналитика, системы управления и навигации летательных аппаратов.

Peter Alexandrovich UKHOV – Cand. Sci. (Tech.), Associate professor, Deputy Head of «IT-Centre» MAI since 2020. Research interests: machine learning, computer vision, unmanned aircraft systems software, predictive analytics, aircraft control and navigation systems.

Мария Борисовна БУЛАКИНА – кандидат технических наук, доцент, начальник управления «ИТ-центр» Московского авиационного института. Область научных интересов – технология разработки программных систем, системный дизайн, управление ИТ-проектами.

Maria Borisovna BULAKINA – Cand. Sci. (Tech.), Associate Professor, Head of the IT Center Department at the Moscow Aviation Institute. Her research interests include software systems development technology, system design, and IT project management.

Сергей Сергеевич КРЫЛОВ – кандидат физико-математических наук, доцент, заведующий кафедрой «Вычислительная математика и программирование» Московского авиационного института. Область научных интересов – технология разработки программных систем, системы геометрического моделирования.

Sergey Sergeevich KRYLOV – Cand. Sci. (Phys.-Math.), Associate Professor, Head of the Department of Computational Mathematics and Programming at the Moscow Aviation Institute. His research interests include software systems development technology and geometric modeling systems.