

УДК 004.942

Сучков Сергей Николаевич, аспирант,
Ульяновский государственный технический университет

Научный руководитель:
Ташлинский Александр Григорьевич, д.т.н., профессор,
Ульяновский государственный технический университет

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ МЕТОДОВ ОБНАРУЖЕНИЯ И СОПОСТАВЛЕНИЯ КЛЮЧЕВЫХ ТОЧЕК НА АЭРОФОТОСНИМКАХ

Аннотация. Сопоставление изображений является одной из фундаментальных задач компьютерного зрения, служащей основой для таких задач, как визуальная локализация, определение структуры и классификация объектов. Использование сверточных нейронных сетей позволило добиться высокой точности, но их значительные вычислительные требования препятствуют практическому внедрению на устройствах с ограниченными ресурсами, таких как мобильные и встраиваемые системы. В статье приводится качественное сравнение алгоритмов обнаружения и сопоставления ключевых точек на изображениях с камеры БПЛА.

Ключевые слова: Сопоставление изображений, локальные признаки, встраиваемые системы, пространственная инвариантность.

Введение

Извлечение локальных характеристик изображения является фундаментальной методикой в компьютерном зрении, предназначенной для идентификации и извлечения репрезентативных и отличительных областей из изображений. Эти характеристики содержат важную информацию о содержании изображения и служат ключевыми компонентами в различных приложениях, включая распознавание изображений, сопоставление изображений и обнаружение объектов. В результате они играют ключевую роль во многих областях.

Традиционные методы сопоставления изображений в основном опирались на дескрипторы признаков, такие как SIFT и ORB, которые работают по ограниченным эвристическим правилам и часто демонстрируют нестабильность при изменении освещения, точки зрения или масштаба. Появление глубоких нейронных сетей значительно повысило надежность и точность извлечения и сопоставления признаков. Недавние подходы продемонстрировали превосходную производительность по различным тестам. Однако эти методы часто делают акцент на точности в ущерб вычислительной эффективности, что приводит к значительным требованиям к ресурсам, которые препятствуют развертыванию на платформах, чувствительных к задержкам или ограниченных по мощности, таких как мобильные устройства и встраиваемые системы [1-4].

В средах с ограниченными вычислительными ресурсами, таких как встраиваемые системы и мобильные устройства, где ресурсы для обработки ограничены и часто одновременно выполняется несколько задач, для практического внедрения критически важны легкие решения для сопоставления изображений.

2. Связанные работы

2.1. Извлечение признаков

В обработке изображений традиционные методы извлечения характерных точек состоят из двух ключевых компонентов: обнаружения ключевых точек и генерации дескрипторов. Основная цель обнаружения ключевых точек – идентифицировать точки изображения, которые демонстрируют уникальность, стабильность и повторяемость, что позволяет им эффективно представлять характерные особенности изображения. Генерация



дескрипторов включает кодирование локальных характеристик, таких как текстура и форма, в окрестностях этих ключевых точек, что позволяет получить отличительные векторы характеристик, которые облегчают количественную оценку сходства между изображениями.

Широко признанным методом извлечения ключевых точек является Scale-Invariant Feature Transform (SIFT) [5], который использует многомасштабные гауссовы фильтры для свертки изображения и выбирает ключевые точки на основе экстремумов в разнице гауссовых функций (DoG). Путем построения пирамиды DoG, SIFT обеспечивает эффективное обнаружение ключевых точек в нескольких масштабах. Кроме того, он генерирует инвариантные по отношению к повороту дескрипторы признаков путем вычисления гистограмм ориентаций градиента в окрестностях ключевой точки. Используя свойства масштабного преобразования гауссовых функций, SIFT обеспечивает согласованность между различными масштабами, демонстрируя отличную устойчивость при обработке поворота изображения. В качестве эталона для извлечения ручных признаков SIFT получил широкое признание за свою эффективность в академических и промышленных приложениях.

Несмотря на свою надежность, SIFT имеет заметные ограничения в вычислительной эффективности, что делает его менее подходящим для приложений реального времени и сред с ограниченными ресурсами. Для решения этой проблемы исследователи предложили альтернативные подходы, такие как FAST (Features from Accelerated Segment Test) [6] и SURF (Speeded-Up Robust Features) [7], которые направлены на достижение баланса между эффективностью извлечения признаков и снижением вычислительной сложности. FAST ускоряет обнаружение ключевых точек путем сравнения различий интенсивности пикселей с соседними пикселями, тогда как SURF основан на SIFT и использует интегральные изображения и бокс-фильтры для улучшения как обнаружения ключевых точек, так и генерации дескрипторов. Однако FAST не обладает устойчивостью к масштабным изменениям, а SURF демонстрирует ограничения в условиях экстремального вращения, что подчеркивает компромиссы, присущие традиционным методам сопоставления изображений.

Чтобы смягчить проблемы, связанные с традиционными алгоритмами сопоставления изображений, в 2011 году был представлен ORB (Oriented FAST and Rotated BRIEF) [8], интегрирующий детектор FAST с дескрипторами BRIEF для улучшения инвариантности к масштабированию и повороту при сохранении высокой скорости вывода. Однако ORB демонстрирует ограничения в областях с низкой текстурой и при значительных изменениях освещения. Кроме того, он проявляет чувствительность к шуму изображения, накладывая ограничения на стабильность сопоставления.

Дескриптор BRISK (Binary Robust Invariant Scalable Keypoints) [9] был представлен как альтернатива SIFT, имеет открытый исходный код и включён в известную библиотеку для обработки изображений OpenCV. Данный дескриптор имеет предопределенный шаблон выборки: пиксели выбираются по концентрическим кольцам. Для каждой точки выборки рассматривается небольшой фрагмент вокруг нее. Перед запуском алгоритма фрагмент сглаживается с помощью гауссова сглаживания. Для выборки используются два типа пар: короткие и длинные. Длинные пары используются для определения ориентации, а короткие – для вычисления дескриптора путем сравнения интенсивностей. BRISK обеспечивает инвариантность к вращению, проверяя ориентацию ключевой точки и поворачивая шаблон выборки на эту ориентацию.

Метод KAZE [10] повышает повторяемость результатов и различимость ключевых точек по сравнению с SIFT и SURF благодаря использованию нелинейной диффузионной фильтрации. Главный недостаток KAZE заключается в его вычислительной сложности. Поскольку аналитических решений для нелинейного уравнения диффузии не существует, необходимо использовать численные методы для аппроксимации решения. В KAZE это



делается с помощью схем аддитивного расщепления операторов (AOS). Хотя схемы AOS абсолютно стабильны для любого шага по времени и легко распараллеливаются, они требуют решения большой системы линейных уравнений для получения решения.

Дальнейшим развитием метода KAZE является AKAZE [11]. Он имеет улучшенную производительность в обнаружении ключевых точек по сравнению с ORB и BRISK, а также, увеличенную скорость вычисления по сравнению с методами SURF, SIFT, KAZE. Метод основан на построении схем Быстрой явной диффузии (FED) в пирамидальную структуру. Пирамидальная стратегия и схемы FED позволяют быстро строить нелинейное масштабное пространство, подходящее для надежного обнаружения и описания признаков.

2.1.1. Алгоритмы сопоставления изображений на основе глубокого обучения.

В последние годы методы на основе глубокого обучения стали эффективным решением для преодоления ограничений ручных средств извлечения признаков. Эти подходы объединяют обнаружение ключевых точек и генерацию дескрипторов в обучаемую и оптимизируемую структуру.

Используя мощные возможности глубоких нейронных сетей, эти методы значительно расширили глубину и объем извлечения признаков, одновременно повысив эффективность и надежность выполнения.

Среди ранних методов на основе глубокого обучения LIFT (Learned Invariant Feature Transform) [12] был одним из первых, в котором использовались сверточные нейронные сети (CNN) для полностью контролируемого сквозного обнаружения и описания ключевых точек. DISK (Deep Image Structure and Keypoints) [13] представил подход, основанный на присвоении весов, извлекая ключевые точки из тепловых карт, сгенерированных CNN, тем самым снижая зависимость от данных, помеченных вручную. Super-Point [14] реализовал самоконтролируемую сверточную модель, обученную на изображениях, сгенерированных с помощью гомографической адаптации, что улучшило надежность обнаружения признаков. Однако основным ограничением SuperPoint являются его значительные вычислительные требования, особенно в задачах сопоставления изображений, связанных с изменениями масштаба.

Другие подходы были направлены на оптимизацию извлечения признаков на основе глубокого обучения. Так, SiLK (Simple Learned Keypoints) [15] использует простую, но эффективную структуру для обучения ключевым точкам и дескрипторам, полагаясь на исходное разрешение изображения для извлечения дескрипторов. ALIKE (Accurate and Lightweight Keypoint Extraction) [16] представляет собой облегченную архитектуру, которая обеспечивает баланс между надежностью и скоростью, включая дифференцируемое обнаружение ключевых точек и потерю нейронной репроекции. Однако его зависимость от исходного разрешения изображения в окончательной карте признаков значительно увеличивает потребление памяти и вычислительные накладные расходы. Для решения этих проблем в 2023 году был представлен ALIKED [17], использующий деформируемые свертки для гибкого моделирования геометрических преобразований и обратной пропации градиентов на субпиксельном уровне для более точной генерации ключевых точек. В том же году в DeDoDe (Detect, Don't Describe – Describe, Don't Detect) [18] продемонстрирован новый подход – отделение обнаружения ключевых точек от обучения дескрипторов. Используя CNN, обученные на крупномасштабных наборах данных структуры из движения (SfM) [19], DeDoDe применяет полностью контролируемую парадигму обучения для улучшения производительности сопоставления признаков.

2.1.2. Методы сопоставления изображений на основе трансформаторов

С развитием архитектур трансформаторов многие методы сопоставления изображений на основе трансформаторов достигли высочайшей точности. LoFTR (Local Feature Transformer) [3] является первым применением трансформаторов к сопоставлению изображений, он не содержит



сложных этапов предварительной или последующей обработки. LoFTR сохраняет стабильную производительность сопоставления даже при значительных изменениях точки зрения, вариациях освещения и частичных затенениях. DISK [13] дополнительно сопряжён с CNN для обнаружения ключевых точек, одновременно включая трансформатор для обучения дескрипторов, сочетая возможности CNN по извлечению локальных признаков с преимуществами трансформатора в моделировании глобального контекста.

Хотя методы на основе трансформаторов обеспечивают превосходную точность и надежность, они требуют значительных вычислительных ресурсов, что затрудняет их внедрение в реальном времени на мобильных и встраиваемых устройствах. Компромисс между точностью и эффективностью остается критически важной проблемой при разработке практических решений для сопоставления изображений.

2.2. Сопоставление признаков

В дополнение к значительному прогрессу в области извлечения локальных признаков, были достигнуты существенные успехи в области сопоставления признаков. Традиционно сопоставление ключевых точек между изображениями выполняется с помощью метода Nearest Neighbor (NN) [20], который идентифицирует соответствующие ключевые точки путем вычисления евклидова расстояния между отдельными дескрипторами. Улучшением по сравнению с NN является подход Mutual Nearest Neighbor (MNN) [21], который требует минимальное расхождение расстояний от точек до соответствующих дескрипторов на двух изображениях, чтобы считаться действительным сопоставлением. Еще одно усовершенствование, dual-softmax matcher (DSM), повышает эксклюзивность сопоставления за счет применения операций нормализации к векторам строк и столбцов матрицы сопоставления. Это гарантирует, что каждая ключевая точка на одном изображении однозначно соответствует ключевой точке на другом, обеспечивая двунаправленную согласованность в процессе сопоставления.

С одной стороны, исследователи оптимизировали эти методы ручного сопоставления с помощью ускоренных алгоритмов поиска. Например, KD-деревья [22] и шаровые деревья эффективно сокращают пространство поиска и повышают скорость сопоставления за счет структурирования данных для эффективного поиска ближайших соседей. Кроме того, методы быстрого поиска ближайших соседей на основе хеширования, такие как локально-чувствительное хеширование (LSH) [23], повышают вычислительную эффективность за счет аппроксимации ближайших соседей при сохранении разумного уровня точности.

С другой стороны, появились инновационные стратегии сопоставления признаков на основе глубокого обучения. SuperGlue [24] использует графовые нейронные сети (GNN), на подобии работы трансформатора, для включения механизмов самовнимания и перекрестного внимания.

Это позволяет SuperGlue использовать пространственные отношения для более надежного сопоставления признаков. LightGlue [25], расширение SuperGlue, упрощает структуру GNN, снижая сложность механизма внимания. Кроме того, он вводит адаптивную стратегию сопоставления, которая динамически регулирует размер сети в зависимости от сложности задачи сопоставления, тем самым снижая вычислительные затраты при сохранении точности сопоставления.

Подход без детектора обходит традиционное обнаружение ключевых точек и напрямую генерирует плотные дескрипторы и сопоставления признаков для пар изображений. LoFTR (Local Feature Transformer) [3] использует полностью контролируемый подход на основе трансформатора, который использует самовнимание и перекрестное внимание для генерации дескрипторов признаков между двумя изображениями. Однако из-за высоких вычислительных требований была разработана усовершенствованная версия, Efficient LoFTR (E-LoFTR) [26], которая повышает эффективность за счет агрегирования механизмов внимания и включения адаптивного выбора маркеров.



OmniGlue [27] представляет предварительно обученную визуальную модель DINOv2 [28] в качестве экстрактора признаков, отображающего извлеченные признаки в общее пространство встраивания через адаптивный слой сопоставления. Аналогичным образом, Robust Dense Feature Matching (RoMa) [2] повышает точность и надежность сопоставления признаков путем встраивания DINOv2 для извлечения грубых признаков и их уточнения с помощью специализированных CNN.

Кроме того, RoMa представляет инновационный декодер сопоставления на основе трансформатора, который предсказывает вероятности опорных точек вместо использования традиционного сопоставления на основе координат, что обеспечивает более гибкие и надежные соотношения сопоставления.

3. Оценка эффективности методов выделения особых точек

Произведена оценка эффективности методов нахождения ключевых точек, наиболее релевантных для встраиваемых систем: ORB, SIFT, BRISK, KAZE, AKAZE. Для оценки использовались характерные аэрофотоснимки – фрагменты видеопоследовательности с БПЛА (рис. 1).

Тестирование производилось в онлайн-среде Spaces от Hugging Face с настройками оборудования CPU Basic (2 vCPU, 16 GB RAM, 50 GB HDD).



Рисунок 1. Фрагменты видеопоследовательности

Таблица 1

3.1 Метод ORB

Количество «особых» точек		Время работы, с
Всего обнаружено	Являются соответствующими	
216	142	0,4

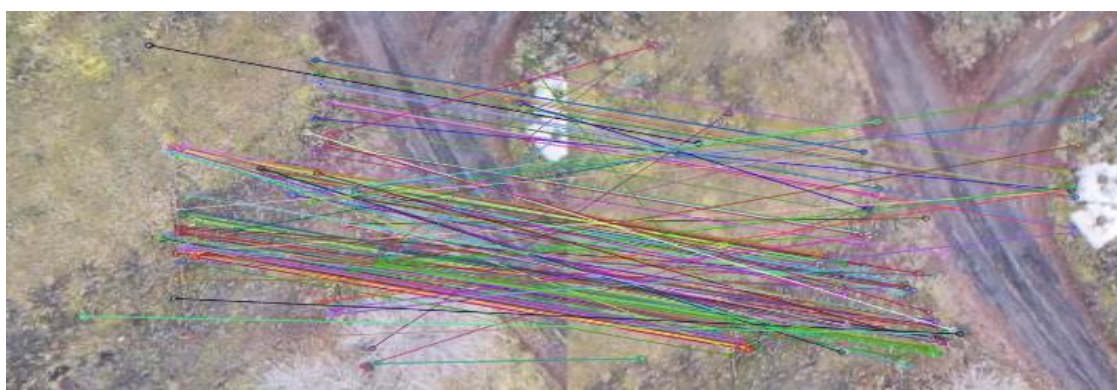


Рисунок 2. Схема расположения особых точек до этапа обнаружения соответствий (ORB)



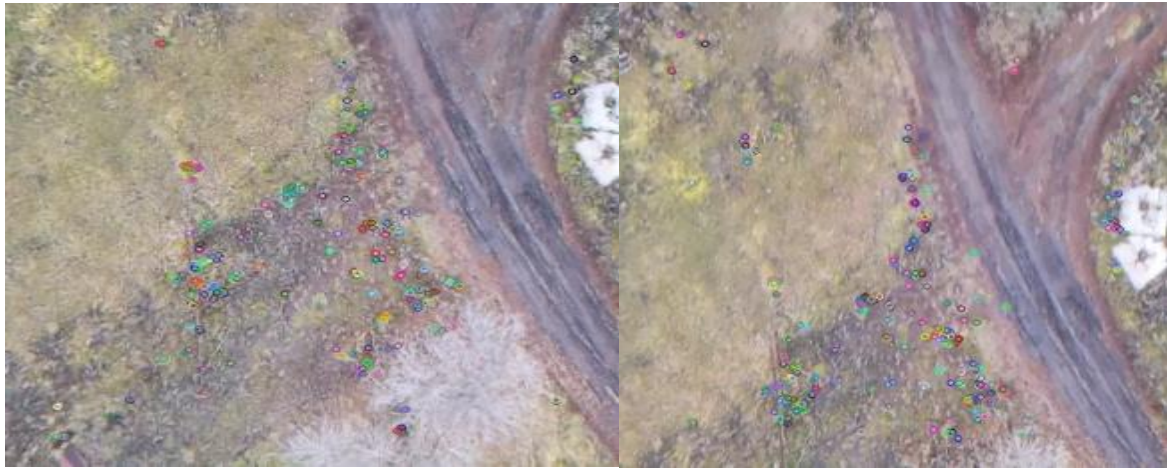


Рисунок 3. Структура обнаруженных с помощью ORB «особых точек»

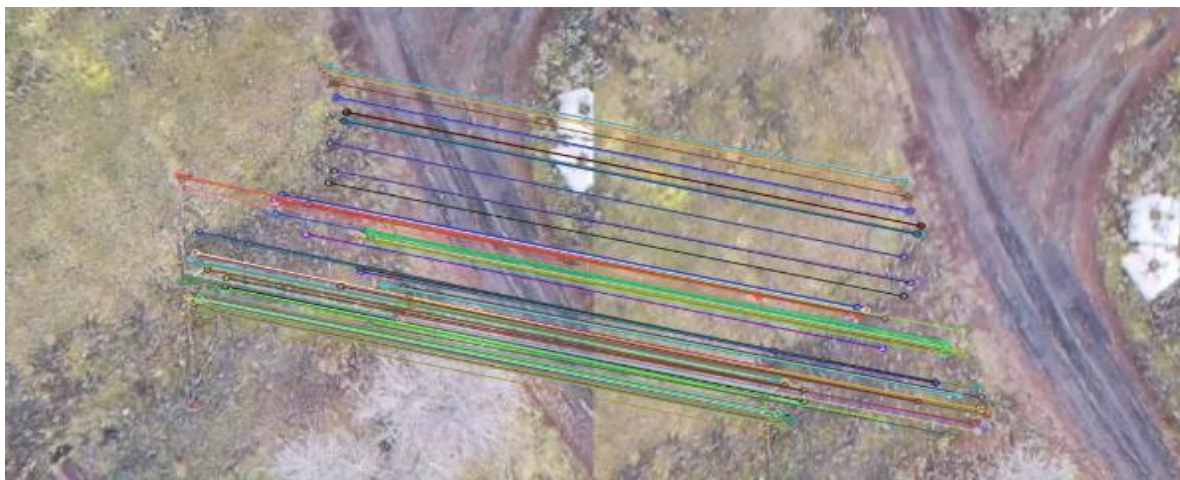


Рисунок 4. Обнаруженные соответствия с помощью метода ORB

Таблица 2

3.2 Метод SIFT

Количество «особых» точек		Время работы, с
Всего обнаружено	Являются соответствующими	
1062	399	0,9

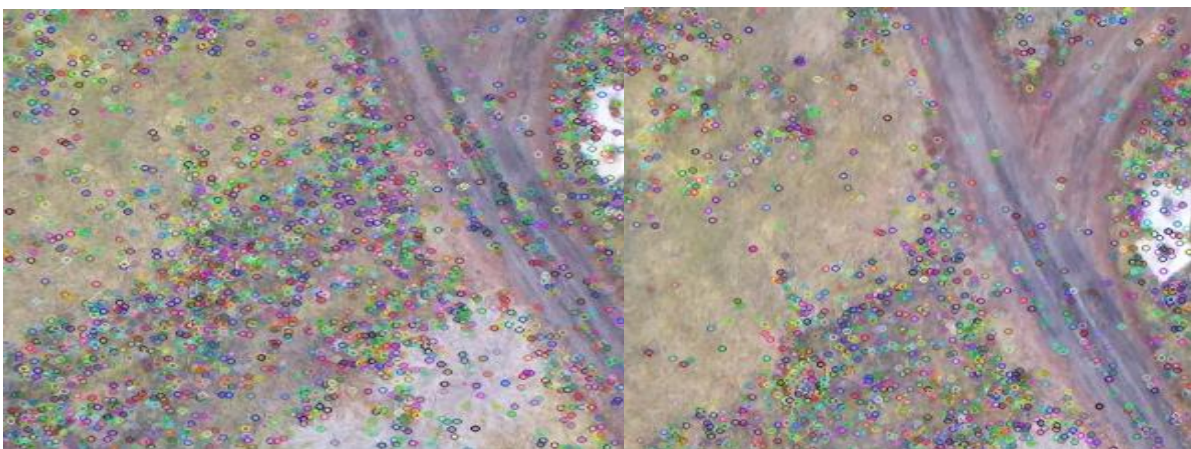


Рисунок 5. Структура обнаруженных с помощью SIFT «особых точек»



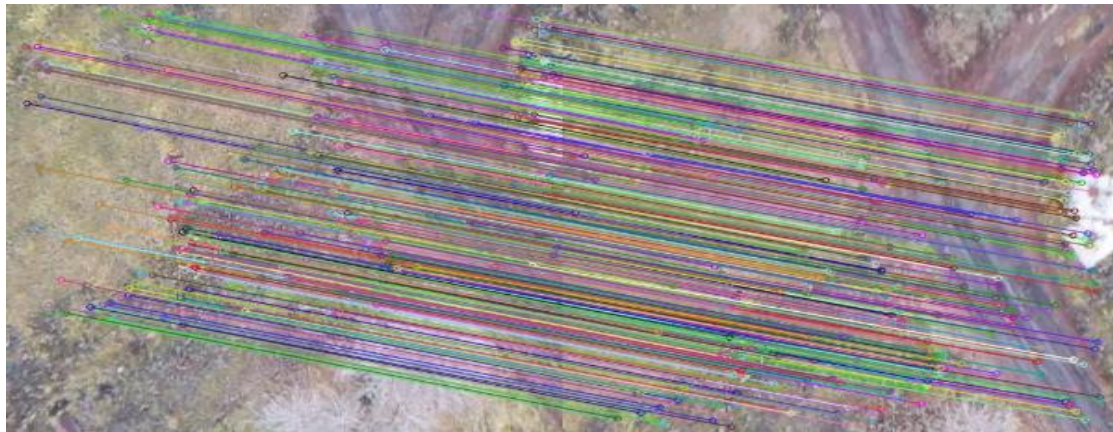


Рисунок 6. Обнаруженные соответствия с помощью метода SIFT

Таблица 3

3.3 Метод BRISK

Количество «особых» точек		Время работы, с
Всего обнаружено	Являются соответствующими	
472	195	0,6

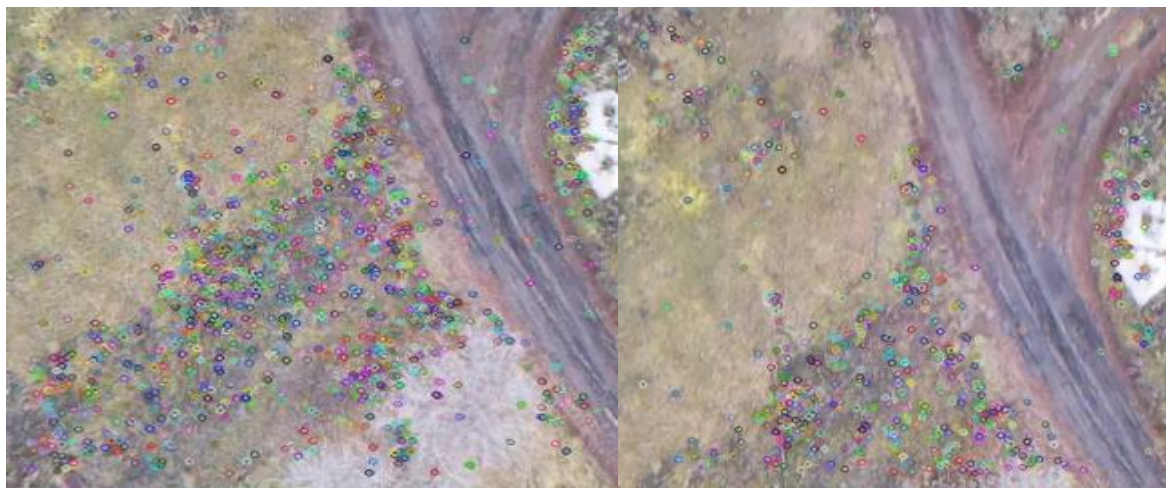


Рисунок 7. Структура обнаруженных с помощью BRISK «особых точек»

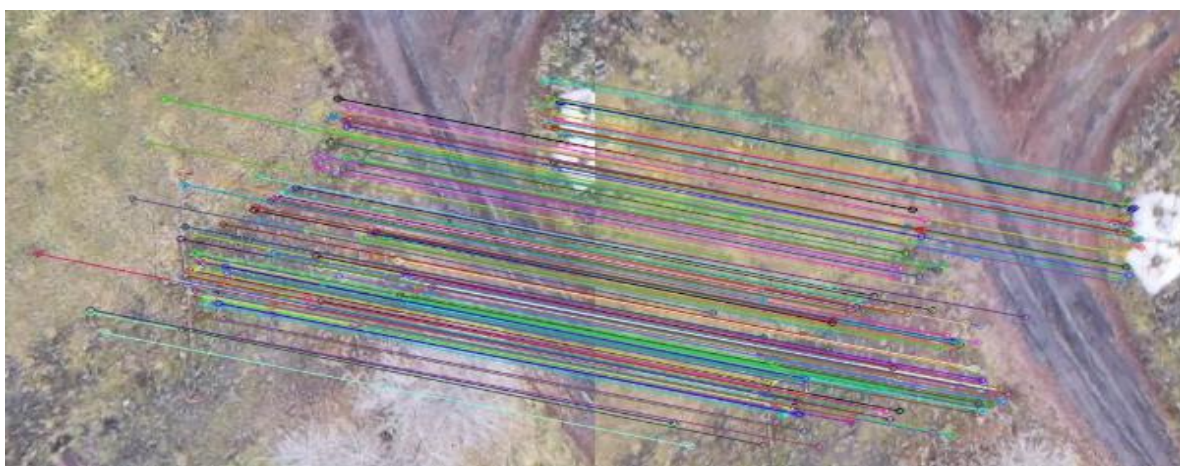


Рисунок 8. Обнаруженные соответствия с помощью метода BRISK



Таблица 4

3.4 Метод KAZE

Количество «особых» точек		Время работы, с
Всего обнаружено	Являются соответствующими	
128	99	0,8



Рисунок 9. Структура обнаруженных с помощью KAZE «особых точек»

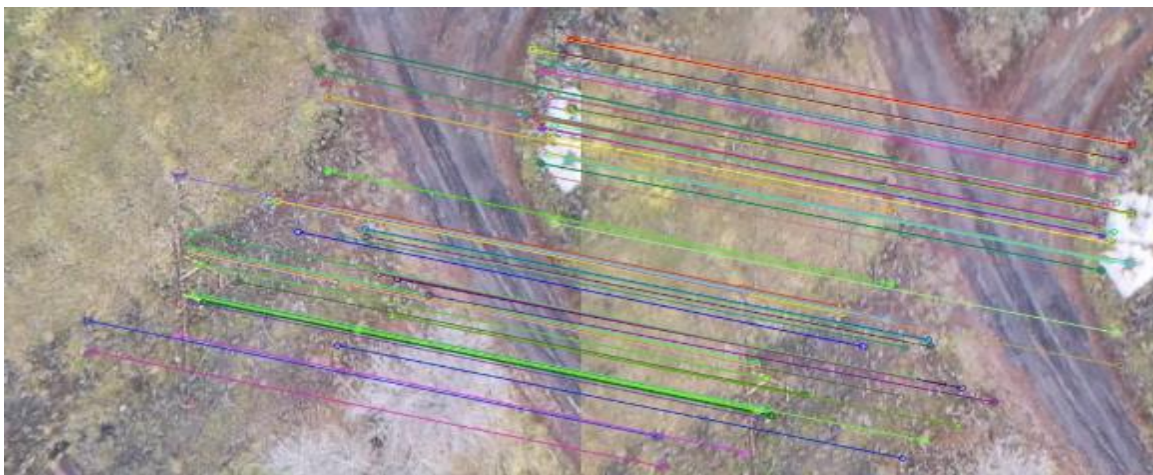


Рисунок 10. Обнаруженные соответствия с помощью метода KAZE

Таблица 5

3.5 Метод AKAZE

Количество «особых» точек		Время работы, с
Всего обнаружено	Являются соответствующими	
101	65	0,5





Рисунок 11. Структура обнаруженных с помощью AKAZE «особых точек»

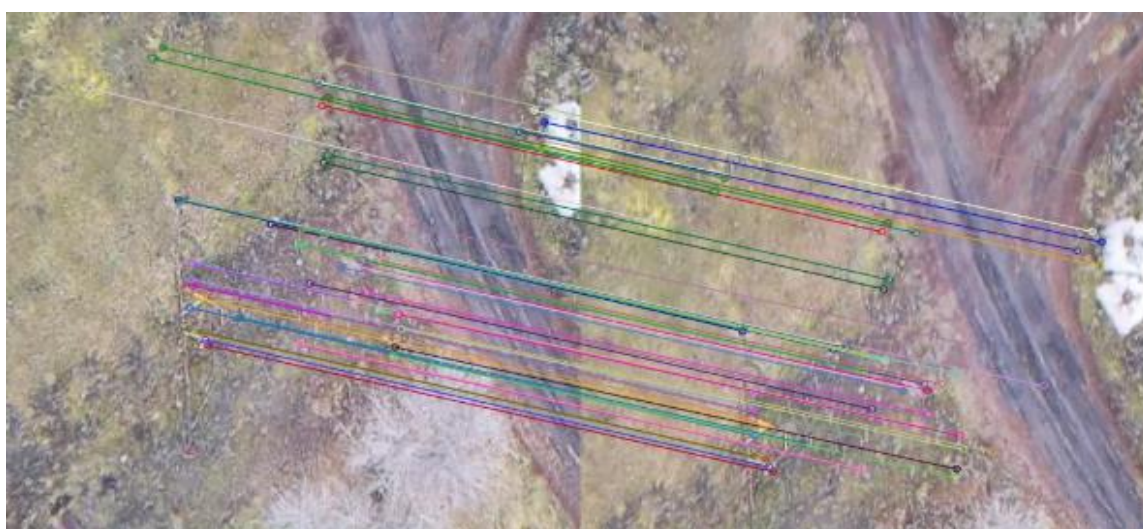


Рисунок 12. Обнаруженные соответствия с помощью метода AKAZE

Таблица 6

Метод	Обнаружено соответствий, %	Времени затрачено, с
ORB	66	0,4
SIFT	37	0,9
BRISK	41	0,6
KAZE	77	0,8
AKAZE	64	0,5

Заключение

В данной статье были рассмотрены методы сопоставления изображений на основе принципа извлечения и сопоставления признаков – дескрипторов. Данные методы можно классифицировать по принципу построения как: алгоритмы прямого извлечения признаков (SIFT, ORB и др.), свёрточные нейронные сети на основе глубокого обучения (SuperPoint, D2Net, DISK и др.), алгоритмы на основе трансформаторов (LoFTR). Каждый класс методов имеет свои преимущества и недостатки в части скорости и точности определения дескрипторов: методы на основе глубокого обучения требуют обширную тренировочную библиотеку изображений для повышения точности, а алгоритмы на основе трансформаторов



требуют наличия большой вычислительной мощности и не подходят для применения в системах реального времени. Учитывая данные ограничения, был проведён сравнительный анализ существующих классических методов определения ключевых точек и составления дескрипторов.

Наименьшее время работы и затраты на вычисления продемонстрировал алгоритм ORB, наибольшее количество соответствий установлено с помощью алгоритма SIFT. Следует отметить, что процент обнаружения соответствий у алгоритма ORB почти в два раза выше, чем у SIFT.

Наибольшее внимание заслуживает метод AKAZE – наиболее сбалансированный по рассматриваемым характеристикам. С его помощью удалось установить соответствия на большей площади изображения по сравнению с ORB, при незначительном увеличении вычислительных затрат. При этом, обнаружение соответствующих точек в процентном соотношении практически не отличается от ORB.

Список литературы:

1. DKM: Dense kernelized feature matching for geometry estimation / J. Edstedt [и др.] // In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, 17-24 июня 2023. – [Б. м.: б. и.], 2023. – С. 17765–17775.
2. RoMa: Robust dense feature matching / J. Edstedt [и др.] // In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16-22 июня 2024. – [Б. м.: б. и.], 2024. – С. 19790–19800.
3. LoFTR: Detector-free local feature matching with transformers / J. Sun [и др.] // In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20-26 июня 2021. – [Б. м.: б. и.], 2021. – С. 8922–8931.
4. PDC-Net+: Enhanced Probabilistic Dense Correspondence Network [Электронный ресурс] / P. Truong [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2023. – Т. 45, вып. 8. – С. 10247-10266. – Режим доступа: <https://doi.org/10.1109/tpami.2023.3249225>.
5. Low, D. G. Distinctive image features from scale-invariant keypoints / D. G. Low // J. Comput. Vis. – 2004. – Т. 60. – С. 91-110.
6. Viswanathan, D. G. Features from accelerated segment test (fast) / D. G. Viswanathan // In Proceedings of the 10th Workshop on Image Analysis for Multimedia Interactive Services, London, UK. – 2009. – С. 6-8.
7. Speeded-Up Robust Features (SURF) [Электронный ресурс] / H. Bay [и др.] // Computer Vision and Image Understanding. – 2008. – Т. 110, вып. 3. – С. 346-359. – Режим доступа: <https://doi.org/10.1016/j.cviu.2007.09.014>.
8. ORB: An efficient alternative to SIFT or SURF / E. Rublee [и др.] // In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6-11 нояб. 2011. – [Б. м.: б. и.], 2011. – С. 2564–2571.
9. Leutenegger, S. BRISK: Binary Robust invariant scalable keypoints [Электронный ресурс] / S. Leutenegger, M. Chli, R. Y. Siegwart // 2011 International Conference on Computer Vision. – 2011. – С. 2548-2555. – Режим доступа: <https://doi.org/10.1109/iccv.2011.6126542>.
10. Alcantarilla, P. F. KAZE Features [Электронный ресурс] / P. F. Alcantarilla, A. Bartoli, A. J. Davison // Lecture Notes in Computer Science. – 2012. – С. 214-227. – Режим доступа: https://doi.org/10.1007/978-3-642-33783-3_16.
11. Alcantarilla, P. Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces [Электронный ресурс] / P. Alcantarilla, J. Nuevo, A. Bartoli // Proceedings of the British Machine Vision Conference 2013. – 2013. – С. 1-13. – Режим доступа: <https://doi.org/10.5244/c.27.13>.



12. LIFT: Learned Invariant Feature Transform [Электронный ресурс] / К. М. Yi [и др.] // *Lecture Notes in Computer Science*. – 2016. – С. 467-483. – Режим доступа: https://doi.org/10.1007/978-3-319-46466-4_28.
13. Tyszkiewicz, M. DISK: Learning local features with policy gradient / M. Tyszkiewicz, P. Fua, E. Trulls // *Adv. Neural Inf. Process. Syst.* – 2020. – Т. 33. – С. 14254-14265.
14. DeTone, D. Superpoint: Self-supervised interest point detection and description / D. DeTone, T. Malisiewicz, A. Rabinovich // *n Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18-23 июня 2018.* – [Б. м.: б. и.], 2018. – С. 224-236.
15. Gleize, P. SiLK: Simple Learned Keypoints [Электронный ресурс] / P. Gleize, W. Wang, M. Feiszli // *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. – 2023. – С. 22499-22508. – Режим доступа: <https://doi.org/10.1109/iccv51070.2023.02056>.
16. ALIKE: Accurate and Lightweight Keypoint Detection and Descriptor Extraction / X. M. Zhao [и др.] // *IEEE Multimed.* – 2023. – Т. 25. – С. 3101-3112.
17. Aliked: A lighter keypoint and descriptor extraction network via deformable transformation / X. Zhao [и др.] // *IEEE Trans. Instrum. Meas.* – 2023. – Т. 72. – С. 1-16.
18. DeDoDe: Detect, Don't Describe – Describe, Don't Detect for Local Feature Matching [Электронный ресурс] / J. Edstedt [и др.] // *2024 International Conference on 3D Vision (3DV)*. – 2024. – С. 148-157. – Режим доступа: <https://doi.org/10.1109/3dv62453.2024.00035>.
19. Ullman, S. The interpretation of structure from motion [Электронный ресурс] / S. Ullman // *Proceedings of the Royal Society of London. Series B. Biological Sciences*. – 1979. – Т. 203, вып. 1153. – С. 405-426. – Режим доступа: <https://doi.org/10.1098/rspb.1979.0006>.
20. Eppstein, D. Raising Roofs, Crashing Cycles, and Playing Pool: Applications of a Data Structure for Finding Pairwise Interactions [Электронный ресурс] / D. Eppstein, J. Erickson // *Discrete & Computational Geometry*. – 1999. – Т. 22, вып. 4. – С. 569-592. – Режим доступа: <https://doi.org/10.1007/pl00009479>.
21. Processing Mutual Nearest Neighbor Queries for Moving Object Trajectories [Электронный ресурс] / Y. Gao [и др.] // *The Ninth International Conference on Mobile Data Management (mdm 2008)*. – 2008. – С. 116-123. – Режим доступа: <https://doi.org/10.1109/mdm.2008.17>.
22. Moore, A. An Introductory Tutorial on Kd-Trees / A. Moore // *Technical Report; Computer Laboratory, University of Cambridge: Cambridge*. – 1991. – С. 209.
23. Gionis, A. Similarity search in high dimensions via hashing / A. Gionis, P. Indyk, R. Motwani // *In Proceedings of the Vldb*. – 1999. – С. 518-529.
24. SuperGlue: Learning Feature Matching With Graph Neural Networks [Электронный ресурс] / P. Sarlin [и др.] // *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2020. – С. 4937-4946. – Режим доступа: <https://doi.org/10.1109/cvpr42600.2020.00499>.
25. Lindenberger, P. LightGlue: Local Feature Matching at Light Speed [Электронный ресурс] / P. Lindenberger, P. Sarlin, M. Pollefeys // *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. – 2023. – С. 17581-17592. – Режим доступа: <https://doi.org/10.1109/iccv51070.2023.01616>.
26. Efficient LoFTR: Semi-Dense Local Feature Matching with Sparse-Like Speed [Электронный ресурс] / Y. Wang [и др.] // *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2024. – С. 21666-21675. – Режим доступа: <https://doi.org/10.1109/cvpr52733.2024.02047>.
27. OmniGlue: Generalizable Feature Matching with Foundation Model Guidance [Электронный ресурс] / H. Jiang [и др.] // *2024 IEEE/CVF Conference on Computer Vision and*



Pattern Recognition (CVPR). – 2024. – С. 19865-19875. – Режим доступа: <https://doi.org/10.1109/cvpr52733.2024.01878>.

28. DINOv2-Based UAV Visual Self-Localization in Low-Altitude Urban Environments [Электронный ресурс] / J. Yang [и др.] // IEEE Robotics and Automation Letters. – 2025. – Т. 10, вып. 2. – С. 2080-2087. – Режим доступа: <https://doi.org/10.1109/lra.2025.3527762>.

