

АЛИ Б., Р. Н. САДЕКОВ, В. В. ЦОДОКОВА

## АЛГОРИТМЫ НАВИГАЦИИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ С ИСПОЛЬЗОВАНИЕМ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ

*В статье анализируются работы, посвященные навигации беспилотных летательных аппаратов (БПЛА) в условиях пропадания сигнала спутниковых навигационных систем с использованием систем технического зрения (камер оптического диапазона). Рассматриваемые алгоритмы основаны на сравнении изображений, сформированных БПЛА, с имеющимися геопривязанными изображениями местности, выступающими в качестве эталона. Изображения сопоставляются либо попиксельно, либо по ключевым точкам, либо с использованием нейронных сетей. Описываются этапы реализации алгоритмов, приводятся достигнутые с их помощью точности, а также используемые для апробации данные. В заключительной части статьи делаются выводы о возможностях и ограничениях рассматриваемых подходов.*

**Ключевые слова:** беспилотный летательный аппарат, автономная навигация, сопоставление ключевых точек, дескриптор, детектор, глубокое обучение, компьютерное зрение, техническое зрение, спутниковые снимки, корреляционно-экстремальные системы.

### Введение

За последние несколько лет беспилотные летательные аппараты (БПЛА) стали широко применяться при решении таких задач, как орошение сельскохозяйственных угодий [1], тушение пожаров [2, 3, 4], доставка грузов [5], поиск и спасение людей [6, 7], мониторинг районов [8], а также при ведении боевых действий [9]. БПЛА имеют ряд преимуществ перед наземными средствами: большая область покрытия территорий, возможность приземления в любую точку местности, высокая скорость полета и, как следствие, оперативность доставки грузов независимо от дорожных обстоятельств и др.

Положение БПЛА в пространстве определяется 6 степенями свободы: координатами ( $x, y, z$ ) и углами ориентации ( $\phi, \theta, \psi$ ), в качестве которых рассматриваются углы относительно горизонта  $\phi, \theta$  и азимутальный угол  $\psi$ . Определение положения объекта обычно осуществляется в глобальной системе координат WGS-84, однако в зависимости от исходных данных могут использоваться и местные системы координат.

Традиционно определение навигационных параметров БПЛА на открытом пространстве осуществляется по сигналам глобальной навигационной спутниковой систе-

---

Али Бушра. Аспирант, Национальный исследовательский университет «МИСиС» (НИТУ МИСИС) (Москва). ORCID 0000-0001-6895-1505.

Садеков Ринат Наилевич. Доктор технических наук, профессор кафедры инженерной кибернетики, НИТУ МИСиС. 0000-0001-6286-358X.

Цодокowa Вероника Владимировна. Кандидат технических наук, научный сотрудник, АО «Концерн «ЦНИИ «Электрон» (С.-Петербург).

мы (ГНСС), параметров ориентации – с использованием микромеханической инерциальной навигационной системы (ИНС). Как показывает практика [12, 13, 14], сигнал ГНСС подвержен как естественным (деревья, горы, городские строения), так и искусственным (подавление) помехам [15, 16, 17]. В условиях пропадания сигналов ГНСС оценка местоположения БПЛА становится ненадежной. При отсутствии прямой видимости спутников точность определения навигационных параметров может обеспечиваться с помощью ИНС [10, 11], однако лишь на небольшом интервале времени. Кроме того, на основе лишь информации от ГНСС невозможно решить дополнительные задачи, в частности вернуть БПЛА на базу [18] при потере сигналов управления или получить более гладкие траектории полета при выполнении задач картографирования.

В целях частичной или полной замены систем навигации, основанных на информации от ГНСС, активно разрабатываются системы навигации и управления, использующие эталонную информацию в виде карты, техническое зрение и интеллектуальные методы обработки данных [19, 20].

Системы навигации с использованием эталонной информации (карты) (map-aided navigation system) появились во второй половине прошлого столетия [21, 22, 23, 24]. В отечественной литературе они часто называются корреляционно-экстремальными системами [25, 26], что обусловлено критерием в виде корреляционной функции, который применялся при сопоставлении эталонных и измеренных фрагментов данных. Корреляционно-экстремальные системы использовались не только при наличии карты, но и при ее отсутствии [27, 28].

Системы, использующие визуальную информацию, полученную с помощью технического зрения, реализуют подход, называемый визуальной навигацией (vision-based navigation). К задачам, решаемым с их помощью, относятся [19]: визуальная локализация и построение карты (visual localization and mapping), обгибание препятствий (obstacle avoidance) и планирование пути (path planning).

Устанавливаемые на БПЛА системы визуальной локализации и построения карты, в свою очередь, делятся [19] на три типа: использующие эталонную карту (map-based), создающие карту в процессе движения (map building) и не использующие карту (mapless). К последним относятся системы, в которых реализованы алгоритмы визуальной одометрии (ВО) (mapless visual odometry) [29]. В системах, создающих карту в процессе движения, применяются алгоритмы одновременной навигации и картографирования (map-building simultaneous localization and mapping – SLAM) [30]. Указанные алгоритмы реализуются без привлечения эталонной информации: определение положения и ориентации объекта производится по последовательности изображений. Ключевыми идеями алгоритмов одновременной навигации и картографирования являются построение карты во время решения задачи навигации и оценка погрешностей определения навигационных параметров по этой карте. Так, компенсация углового ухода (дрейфа) может производиться за счет следующих техник: оптимизации траектории по ее локальному фрагменту (bundle adjustment), замыкания траектории при повторном ее пересечении (loop closing) и др.

Алгоритмы визуальной одометрии и одновременной навигации и картографирования не рассматриваются в настоящей работе, так как они в основном рассчитаны на навигацию внутри помещений и помимо систем технического зрения предполагают применение достаточно разнообразного набора других сенсоров (лидаров, камер глубины и др.).

В настоящей статье представлен обзор работ по навигации БПЛА с использованием систем технического зрения вне помещений. Анализируются алгоритмы, реализуемые в навигационных системах с применением эталонной информации. В основе рассматриваемых алгоритмов лежит сопоставление изображений с помощью различных подходов, которые предлагается разделить на:

- корреляционно-экстремальный;
- использование ключевых точек;
- использование нейронных сетей.

В соответствии с этой классификацией группируются рассматриваемые далее алгоритмы.

Первая группа алгоритмов предполагает определение взаимных ориентации и положения изображений (эталонного и формируемого на борту БПЛА) посредством их сопоставления в полном виде с помощью корреляционно-экстремального подхода. Для более надежного сравнения изображений могут использоваться контуры, области, точки и другие их характерные элементы, извлекаемые на основе как классических, так и нейросетевых алгоритмов обработки изображений.

Алгоритмы, относящиеся ко второй группе, позволяют определить взаимные ориентацию и положение изображения, формируемого БПЛА, и эталонного изображения путем сопоставления ключевых точек, предварительно извлеченных с помощью так называемых детекторов. Ключевая точка задается координатами ее положения на изображении и дескриптором – вектором, характеризующим уникальность области вокруг нее. Для построения детекторов могут использоваться как классические алгоритмы, так и основанные на нейронных сетях.

В третьей группе алгоритмов для сопоставления изображений применяются глубокие нейронные сети, обученные с помощью эталонных изображений местности. Они используются как для глобальной навигации (image-based-geolocalization), так и для локальной (image regression). В первом случае определяется местоположение БПЛА на больших пространствах, охватывающих территорию округа, штата, страны. Во втором случае, который, безусловно, ближе к тематике настоящей статьи, местоположение и ориентация БПЛА устанавливаются при полете в небольшом пространстве, например в жилом дворе или квартале из нескольких домов. Алгоритмы третьей группы являются новыми в навигации подвижных объектов.

Следует отметить, что ввиду наукоемкости и сложности существующих навигационных алгоритмов на основе эталонной карты при их реализации нередко применяются подходы, относящиеся к различным группам. Вследствие этого не всегда можно с уверенностью сказать, к какой именно группе принадлежит тот или иной алгоритм. К тому же помимо данных, полученных с помощью системы технического зрения, часто привлекается информация от инерциальных систем, которые по этой причине также будут упоминаться.

Таким образом, анализируемые в настоящей статье работы распределяются по трем блокам, которые далее рассматриваются в трех соответствующих разделах:

- алгоритмы навигации на основе сопоставления изображений с использованием корреляционно-экстремального подхода;
- алгоритмы навигации на основе сопоставления изображений с использованием ключевых точек;
- алгоритмы навигации на основе сопоставления изображений с использованием нейронных сетей.

Каждый раздел содержит описание основной идеи, анализ работ и сравнительные таблицы с информацией об используемых алгоритмах и датасетах, а также о достигнутой с их помощью точности. В общем случае каждое из представленных ниже решений предполагает использование в качестве входных данных изображения, полученного посредством системы технического зрения (камеры оптического диапазона), и карты местности. На выходе формируется вектор навигационных параметров (конкретный состав вектора состояния приводится в таблицах в столбце «Погрешность, компоненты вектора состояния»).

### **Алгоритмы навигации на основе сопоставления изображений с использованием корреляционно-экстремального подхода**

В работах, рассматриваемых в настоящем разделе, описываются алгоритмы, основанные на сравнении эталонного изображения и изображения, полученного с БПЛА. При этом сопоставляются цельные изображения попиксельно, как правило, с помощью корреляционно-экстремальных подходов, основная идея которых заключается в поиске максимума взаимной корреляционной функции изображений.

Исследования, посвященные этой группе алгоритмов, являются одними из самых ранних. В настоящее время при их разработке активно используются различные виды нейронных сетей. Например, сегментационные нейронные сети позволяют применять в качестве эталонных данных векторные карты, которые могут быть оперативно бинаризованы для выделения имеющихся на местности объектов (дороги, дома, речная сеть и др.).

На рис. 1 показаны общие этапы для алгоритмов, рассматриваемых в настоящем разделе:

- этап сегментации, на котором выделяются различные классы объектов: дороги, здания, реки и пр. (в работе [35] в качестве инструмента сегментации использовалась нейронная сеть cGAN [33, 34]);
- этап сопоставления полученных изображений;
- дополнительные процедуры для корректного сравнения сегментированных изображений, например алгоритмы ВО.

В работе [35] представлен новый подход к визуальной навигации с применением условной генеративной состязательной нейронной сети cGAN [33, 34] и метода сопоставления изображений. Полученные от БПЛА снимки сегментировались и сопоставлялись с эталонным изображением (эталонной картой) путем расчета корреляции на основе нормализованной суммы квадратов разностей интенсивностей пикселей в каждый момент времени.

В [36] описывается алгоритм навигации БПЛА, основанный на автоматическом распознавании ориентиров с использованием корреляционного подхода. На эталонной карте выбирались изображения ориентиров с известной информацией о широте и долготе. Далее с помощью ключевых точек и алгоритма AKAZE [37] производился отбор «кандидатов» – возможно, искомым ориентиров. Эталонное изображение ориентира и изображения «кандидатов» бинаризовались посредством детектора краев Канни (Canny), и по бинарным изображениям с привлечением взаимной корреляционной функции производилось окончательное распознавание ориентира. Такой подход позволил существенно снизить вероятность ложных идентификаций ориентиров.

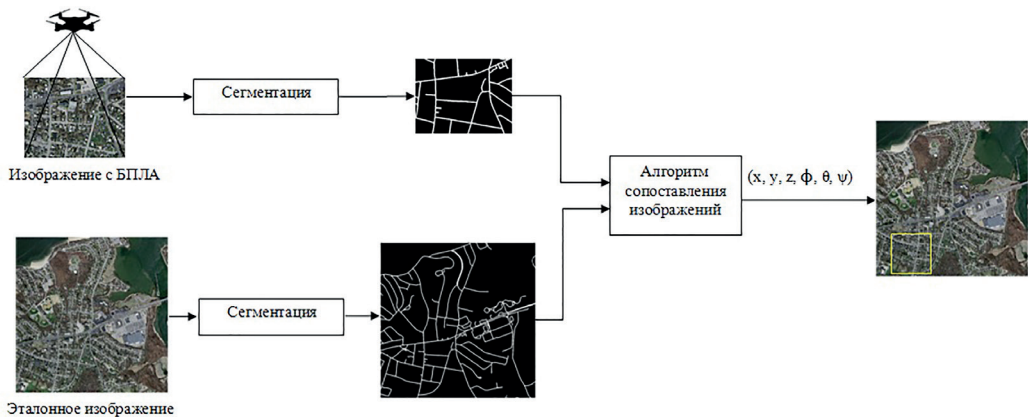


Рис. 1. Алгоритм навигации на основе сопоставления изображений с использованием корреляционно-экстремального подхода

Т а б л и ц а 1

**Сравнительные характеристики алгоритмов навигации на основе сопоставления изображений с использованием корреляционно-экстремального подхода**

Исследование	Используемые алгоритмы	Эталонные изображения / размер ш [м] × д [м]	Изображения с БПЛА / размер ш [м] × д [м]	Погрешность, компоненты вектора состояния
[35]	cGAN + корреляция изображений	Спутниковые снимки / 102 × 102	Реальный полет / 216 × 144	AED = 22,7 м
[36]	AKAZE + корреляция изображений	Спутниковые снимки (Deimos-2)	Видео с Международной космической станции (1032 кадра) / 3840 × 2160	AED = 15,24 м (x, y)
		Спутниковые снимки	Реальный полет	AED = 21,57 м (x, y)
[38]	CNN	University-1652	University-1652	–

Примечание: AED – среднее евклидово расстояние.

В работе [38] с применением нейронной сети CNN сопоставлялись изображения зданий нескольких университетов, сформированные БПЛА при различных зенитных углах линии визирования, и эталонные спутниковые снимки этих зданий. Для каждого здания в наборе данных University-1652 [39] содержится одно изображение со спутника и 54 изображения, сформированных БПЛА при различных зенитных углах. CNN была обучена таким образом, чтобы изображение БПЛА приводилось к виду, аналогичному спутниковому снимку, затем оба изображения сравнивались. В результате этих экспериментов удалось достичь уровня полноты (recall) 90,02% и средней точности распознавания (precision) 70,82%. Примечательно, что при обратном порядке сопоставления изображений – от снимка со спутника к снимку с БПЛА – различные по-

казатели точности распознавания (precision, recall) снижались на 5%. Погрешность определения навигационных параметров в данной работе не оценивалась. Вместе с тем, как заявляют авторы, при необходимости эта задача может быть решена путем добавления геопривязки спутниковых снимков.

### *Выводы*

Алгоритмы, представленные в настоящем разделе, основаны на сравнении изображений (формируемого БПЛА и эталонного) с применением корреляционных подходов. В целях повышения надежности сопоставления часто осуществляется предварительная сегментация изображений с помощью глубоких нейронных сетей.

Корреляционные подходы к сравнению изображений привлекаются все реже, их заменяют алгоритмы, основанные на сопоставлении ключевых точек (см. следующий раздел). Кроме того, несмотря на активное применение нейронных сетей для сегментации аэрофотоснимков, большое количество типов объектов по-прежнему сегментируется с низкой точностью. При этом наилучшие результаты достигаются с помощью алгоритмов, направленных на выделение на изображениях дорог и зданий и используемых прежде всего в приложениях воздушного мониторинга местности или аннотирования изображений, а также в системах дополненной реальности для операторов БПЛА. Что касается других типов объектов, высокой точностью сегментирования пока характеризуются лишь небольшие объемы размеченных данных.

### **Алгоритмы навигации на основе сопоставления изображений с использованием ключевых точек**

В настоящем разделе анализируются исследования, посвященные разработке алгоритмов, основанных на сравнении ключевых точек эталонного изображения и изображения, полученного с БПЛА. Такие алгоритмы позволяют оперировать меньшим количеством информации для сопоставления данных и обеспечивает высокую надежность сравнения. Структура алгоритмов, основанных на сопоставлении ключевых точек, схематично показана на рис. 2 и предполагает следующие этапы их реализации:

- определение начального положения БПЛА ( $x_0, y_0, z_0, \phi_0, \theta_0, \psi_0$ ) в том случае, если постановка задачи этого требует, при этом могут быть использованы другие подходы, не основанные на сопоставлении ключевых точек;
- извлечение ключевых точек и их дескрипторов из изображения, формируемого БПЛА, и из эталонного снимка с использованием различных детекторов точек;
- привлечение дополнительной информации, например измерений ИНС или алгоритмов ВО, обеспечивающих непрерывность решения задачи навигации в случаях, когда результаты сравнения эталонного снимка и снимка с БПЛА являются неудовлетворительными;
- предварительная обработка извлеченных ключевых точек с целью отбраковки части из них;
- навигация, в ходе которой определяется положение БПЛА в пространстве путем сопоставления ключевых точек на изображении, полученном с него, и эталонном снимке и определения матриц преобразования координат снимка с БПЛА в систему координат эталонного снимка. Этап может быть реализован с использованием как методов RANSAC, так и методов Монте-Карло, при этом по сути

производится поиск максимума взаимной корреляционной функции ключевых точек изображений – эталонного и формируемого на борту БПЛА. Сопоставление ключевых точек осуществляется посредством сравнения их дескрипторов с использованием квадратичной нормы [40] или расстояния Хэмминга (Hamming) (для бинарных дескрипторов) [40].

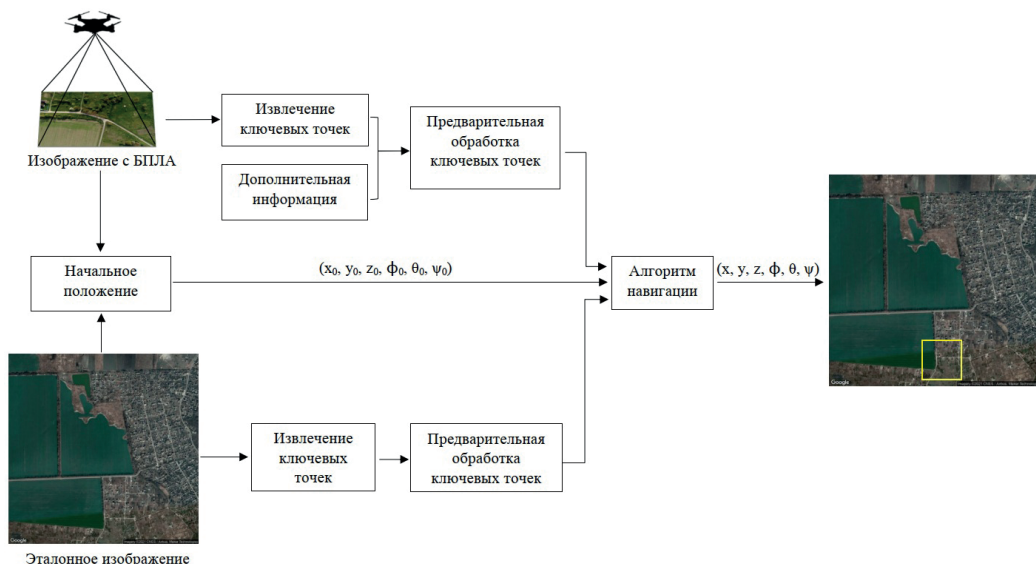


Рис. 2. Алгоритм навигации на основе сопоставления изображений с использованием ключевых точек

Обсудим работы, посвященные данному направлению, с позиций приведенной многоэтапной схемы.

В [41] предложен трехэтапный алгоритм. Сначала с использованием ВО оценивались координаты и параметры ориентации БПЛА. Далее изображения, полученные от БПЛА, и геопривязанные спутниковые снимки пропускались через сверточную нейронную сеть, скомпонованную из слоев нейронной сети VGG16 [42], с целью извлечения признаков изображений и их дальнейшего сопоставления. Для уточнения координат по ключевым точкам определялась матрица гомографии. Код алгоритма доступен по ссылке<sup>1</sup>. Алгоритм протестирован на двух наборах данных<sup>2</sup>.

В работах [46, 47] рассматриваются похожие алгоритмы для автономной навигации БПЛА. В качестве детектора ключевых точек использовался алгоритм SURF [48], алгоритм навигации основывался на методе RANSAC [43]. Другие этапы описанной выше схемы не реализовывались. Метод RANSAC позволил эффективно бороться с выбросами и находить навигационное решение даже в условиях шума. При этом количество ключевых точек, выбираемых для определения навигационных параметров, должно быть не меньше пятнадцати.

В статье [49] ключевые точки также извлекались с помощью алгоритма SURF. Из этих точек были выбраны десять с наибольшим значением определителя матрицы

<sup>1</sup> <https://www.kaggle.com/balraj98/massachusetts-buildings-dataset>

<sup>2</sup> Village Dataset. <https://www.sensefly.com/education/datasets/?dataset=1419>; Gravel Pit Dataset. <https://www.sensefly.com/education/datasets/?dataset=1421>

Гессе (Hesse), для которых вычислялся еще один дескриптор, состоящий из откликов различных вейвлетов Габора (Gabor). Алгоритм сужения области поиска реализован аналогично описанному в [50]. Сопоставление ключевых точек на эталонном изображении и изображении с БПЛА осуществлялось на основе метода RANSAC.

В работе [10] предложена модифицированная версия детектора BRIEF [52] – abBRIEF. Дескриптор, формируемый abBRIEF, в отличие от оригинала вычислялся на цветовом пространстве CIE LAB, и уменьшение шума изображения осуществлялось посредством процесса квантования вместо фильтра Гаусса (Gauss). Дескрипторы, формируемые abBRIEF, представляют собой наборы (обычно 64 или 128) пар пикселей, которые случайным образом выбираются в окрестности ключевой точки. Слежение за изменением координат точек между кадрами осуществлялось с помощью оптического потока (ОП), чтобы упростить задачу нахождения соответствия между ключевыми точками в ходе полета. Для навигации БПЛА применялся фильтр частиц (ФЧ), в котором каждая частица реализовывала гипотезу предполагаемого местоположения БПЛА. В качестве меры корреляции между ключевыми точками служило расстояние Хэмминга. По найденным ключевым точкам вычислялась матрица гомографии, посредством которой оценивалось положение БПЛА в каждой частице. Используемый в работе датасет доступен по ссылке<sup>1</sup>.

В [53] предложен алгоритм, в основе которого лежит классификация местности. Определены четыре класса местности: трава, кусты/деревья, улица/асфальт и здания. Первый и второй класс относятся к областям со сложной текстурой. Сначала изображение разделялось на небольшие прямоугольные области, затем из каждой области извлекались ORB-дескрипторы [44]. Далее с привлечением полученной информации и модели случайного леса [54] области классифицировались, что позволило обеспечить надежность сравнения ключевых точек за счет исключения «проблемных» областей, относящихся к первому и второму классам. Этап навигации реализовывался в виде ФЧ.

Для сокращения затрат на хранение данных в работе [55] предложен новый подход к навигации БПЛА, предполагающий применение автоэнкодера для преобразования изображений в их меньшее векторное представление (embedding). Автоэнкодер, представляющий собой пятислойный CNN-кодер и пятислойный CNN-декодер, в процессе обучения должен был уменьшать размер изображения и увеличивать его до исходного. В режиме полета для формирования вектора признаков изображений с БПЛА использовался только CNN-кодер. Затем извлеченный вектор сравнивался с векторами признаков эталонных изображений и на выходе формировалась вероятность соответствия изображений. Оценка местоположения вычислялась как математическое ожидание координат центров соответствующих эталонных изображений, умноженных на вероятность нахождения в них. Обучение автоэнкодера осуществлялось на изображениях Google Earth.

В [56] разработан алгоритм навигации БПЛА с использованием гистограмм ориентированных градиентов HOG [57]. Первый шаг алгоритма – инициализация ФЧ с целью получения предполагаемого местоположения БПЛА. На этом этапе корреляционный фильтр между изображением БПЛА и эталонным снимком, основанный на 2D-преобразованиях Фурье, формировал глобальную карту достоверности исходного местоположения, где положение максимума рассматривалось как начальное для взлета

<sup>1</sup> POA\_UAV dataset. <https://zenodo.org/record/1244314>; AdM\_UAV dataset. <https://zenodo.org/record/1244296>



БПЛА, а вокруг него формировался набор частиц. На следующем этапе использовались ОП [58] и измерения ИНС для прогнозирования положения БПЛА между кадрами, что существенно сокращало пространство поиска. После этого дескрипторы HOG извлекались из изображения, формируемого БПЛА, и сравнивались с дескрипторами точек, содержащимися в частицах, локализованных на эталонном изображении. Эксперименты показали, что применение ОП позволяет повысить надежность работы алгоритма.

Т а б л и ц а 2

**Сравнительные характеристики алгоритмов навигации на основе сопоставления изображений с использованием ключевых точек**

Исследование	Используемые алгоритмы	Эталонные изображения / размер ш [М] × д [М]	Изображения с БПЛА / размер ш [М] × д [М]	Погрешность, компоненты вектора состояния
[10]	abBRIEF, ФЧ	Google Earth / 1160 × 1160	Реальный полет – 3 пролета, 2 км	AED = 17,8 м (x, y, z)
[41]	VGG16, BO, SURF	USGS Earth Explorer / 7582 × 5946	Реальный полет – 0,85 км	AED = 7 м (x, y), 8 м (z) AED = 25 м (x, y), 7,7 м (z)
[46]	SURF, RANSAC	Google Earth / –	Реальный полет	MAE = 0,11 м, 0,25 м (x, y)
[47]	SURF	Google Earth / 3340 × 1880	Смоделировано с Google Earth / 496 × 278	RMSE = 1,6 м (x, y)
[49]	SURF, вейвлеты Габора, RANSAC	Спутниковый снимок / 10000 × 10000	Виртуальный полет	MAE = 4 м (x, y), 0,01° (φ, θ), 0,2° (ψ)
[51]	SIFT, ORB, RANSAC, U-Net	Google Maps и Bing	Реальный полет – 1,2 км	AED = 3,6-5,1 м (x, y)
[53]	ORB, случайный лес, ФЧ	Google Maps / 150 × 90	Реальный полет – 60 × 100	AED = 9,5 м (x, y)
[55]	Автоэнкодер	Google Earth / –	Реальный полет – 1132 м	RMSE = 3 м (x, y), 0,36° (ψ)
[56]	ОП, HOG, ФЧ	Google Maps / 300 × 150	Реальный полет – 360 м	RMSE = 6,77 м (x, y)
[59]	ResNet, RANSAC	Собственный датасет	Реальный полет – 3 пролета, 0,52 км <sup>2</sup> , 0,64 км <sup>2</sup> , 0,66 км <sup>2</sup>	AED = 3,6 м (x, y)
[64]	U-Net, SIFT, RANSAC	Лион, собственный датасет	Реальный полет – 7 км	~100 м (x, y)
Примечания: ОП – оптический поток; ФЧ – фильтр частиц; MAE – средняя абсолютная ошибка; RMSE – средняя квадратическая ошибка; AED – среднее евклидово расстояние.				

В [59] описывается формирование карт признаков для эталонного изображения и изображения с БПЛА с использованием нейронной сети ResNet [60, 61]. Далее с учетом функции косинусного сходства между извлеченными признаками обоих снимков строилась 4D-корреляционная карта, по которой определялись индексы совпадающих точек, применяемые далее в методе RANSAC (по аналогии с ключевыми точками) для вычисления матрицы преобразования. Нейронная сеть была адаптирована и обучена на наборах данных PASCAL-VOC [62] и InLoc [63]. Код алгоритма доступен по ссылке<sup>1</sup>.

В [64] предложен новый подход к коррекции навигационных показаний БПЛА, основанный на сравнении полученных от него изображений с изображениями, сформированными из векторной топографической карты. Снимок с БПЛА сегментировался с помощью сверточной нейронной сети U-Net [31], в результате получали бинарное изображение, на котором фон был черного цвета, дороги и реки – белого. Далее с использованием детектора SIFT [32] из этого изображения извлекались ключевые точки, которые сравнивались с ключевыми точками бинарного изображения векторной карты, сформированного аналогичным образом. Далее с помощью метода RANSAC [43] вычислялась матрица гомографии. Алгоритм протестирован с привлечением векторных и растровых картографических данных города Лиона<sup>2</sup>, а также данных, полученных авторами. Результаты показали возможность определения координат с погрешностью ~100 м.

В [51] для уточнения положения БПЛА помимо классического сопоставления ключевых точек производилось выделение контуров зданий на изображениях (эталонном и формируемом БПЛА) с помощью нейронной сети U-Net [31]. Полученные контуры зданий оценивались на предмет сходства путем анализа морфологических признаков и далее служили в качестве аналогов ключевых точек для расчета уточненной матрицы гомографии. Анализ точности предложенного подхода произведен по двум датасетам городов Фамагусты на Кипре и Потсдама в Германии.

### *Выводы*

В табл. 2 в обобщенном виде представлены данные исследований, о которых шла речь выше. В качестве подходов используются классические алгоритмы выделения ключевых точек, такие как SURF [49], ORB [53] и BRIEF [10]. Иногда для более точного сопоставления производится предварительная сегментация изображений [53]. В некоторых статьях описано применение нейронных сетей как для предобработки изображений [64], так и для извлечения их признаков, выступающих в роли ключевых точек [59].

В большинстве экспериментов для проверки подходов привлекались изображения, полученные во время реальных полетов, а в [47] – изображения из Google Earth. В качестве характеристик точности использовались различные метрики (MAE, RMSE и AED), что, к сожалению, не позволяет сравнивать предложенные алгоритмы с единых позиций. При этом зачастую авторы исследований приводят общую (суммарную) погрешность полученного решения, а не по отдельности для каждой из компонент вектора состояния. Кроме того, используемые в некоторых работах

<sup>1</sup> <https://github.com/m-hamza-mughal/Aerial-Template-Matching>

<sup>2</sup> Data Grand Lyon. <https://rdata-grandlyon.readthedocs.io/en/latest/index.html>

исходные данные являются закрытыми и не представлены в общем доступе. В связи с этим выводы делаются с опорой прежде всего на результаты работ с открытой исходной информацией.

Наличие большого количества исследований в рассматриваемой области обусловлено высоким спросом на создание и применение на практике алгоритмов навигации, основанных на сопоставлении ключевых точек. Точность определения навигационных параметров при помощи указанных алгоритмов может варьироваться от десятков метров до дециметров в зависимости от исходных данных: изображений, высоты и скорости полета БПЛА. При этом уровень погрешности не зависит от времени и маршрута полета. Наиболее точные результаты достигаются, если соответствуют пространственные разрешения эталонных изображений и изображений, формируемых БПЛА, а также сохраняется актуальность информации.

В условиях большого количества геоинформационных данных нейросетевые подходы показывают более высокую точность и надежность извлечения ключевых точек по сравнению с классическими подходами, что, безусловно, положительно сказывается на качестве навигационного решения. При построении алгоритмов часто вполне оправданным является применение ОП, ВО и инерциальных данных, обеспечивающих высокую надежность сопоставления точек и упрощение этой процедуры.

В качестве эталонных карт в ряде приложений можно использовать и геоинформационные ресурсы Google Earth или Bing Maps, однако необходимо учитывать возможные изменения изображений, которые могут быть обусловлены различными факторами: освещенностью, сезонными и погодными явлениями, застройкой или естественными изменениями природного ландшафта.

### **Алгоритмы навигации на основе сопоставления изображений с использованием нейронных сетей**

Все большее распространение глубокого обучения в области компьютерного зрения и робототехники в последние годы стало возможным благодаря множеству архитектур глубоких нейронных сетей, доказавших свою эффективность при решении различных задач. Например, стоит упомянуть нейронную сеть AlexNet [65], которой было присуждено первенство в крупномасштабном конкурсе визуального распознавания на наборе данных ImageNet [66]. По этой причине последние несколько лет исследователи активно применяют глубокое обучение для реализации алгоритмов навигации БПЛА.

Рассматриваемые в настоящем разделе алгоритмы построены на единственном элементе – глубокой нейронной сети, которая, получая на вход изображение, формирует на выходе координаты и параметры ориентации БПЛА. Перед полетом нейронная сеть обучается на наборе имеющихся изображений, при этом минимизируется взвешенная сумма квадратов разностей координат и параметров ориентации – предсказанных нейронной сетью и истинных (соответствующих набору имеющихся изображений).

Алгоритмы этой группы состоят из следующих этапов (рис. 3):

- обучение, в процессе которого нейронная сеть обучается на имеющихся геопривязанных изображениях или изображениях, для которых точно известно местоположение и ориентация объекта;
- навигация, во время которой определяется местоположение БПЛА с применением весов сети, полученных на первом этапе.

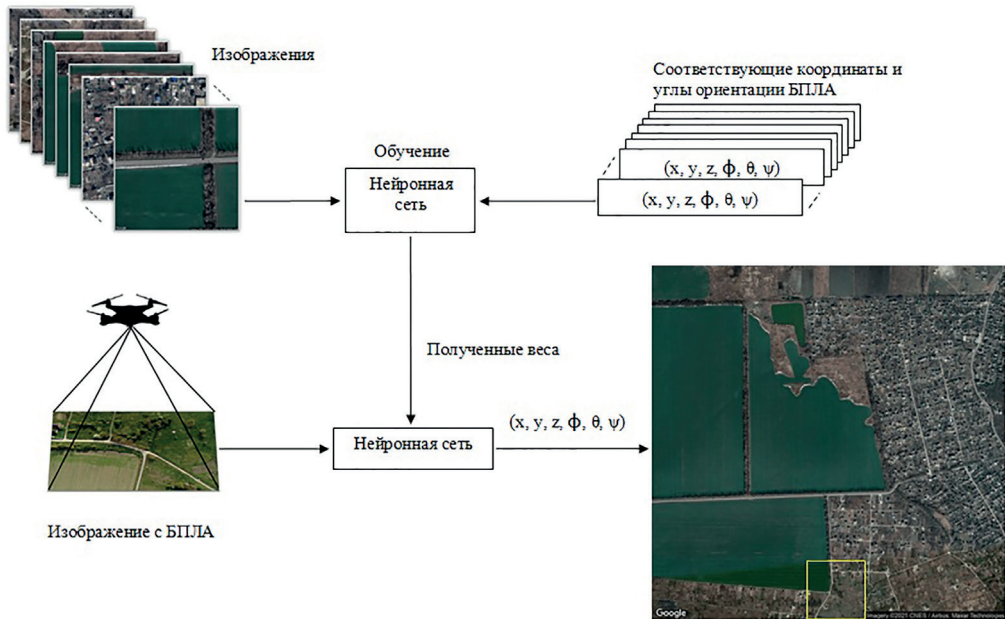


Рис. 3. Алгоритм навигации на основе сопоставления изображений с использованием нейронных сетей

В работе [67] архитектура нейронной сети Xception [68] модифицирована путем добавления полносвязных слоев исключения (dropout) и изменения функции потерь. В качестве функции потерь выступало отрицательное логарифмическое правдоподобие. Для оценки точности использовались 13 наборов данных<sup>1</sup> различных районов округа Вашингтон (на двенадцати из них проходило обучение сети). Проведенный анализ показал, что средняя квадратическая погрешность составила 166,5 м в районе площадью 32,2 км × 32,2 км. Авторы исследовали точность на крупномасштабных изображениях, при этом результаты ухудшились, что объясняется недостаточной способностью сети анализировать более мелкие объекты.

В [69] описаны два эксперимента по навигации – на уровне округа (задача классификации) и в районе (задача регрессии). В качестве нейронной сети послужила архитектура VGG16 [42], в которой обучались только два полносвязанных слоя. Для обучения использовались изображения семи различных округов Каира. Задача навигации на уровне округа заключалась в правильном его определении по изображениям с БПЛА. Сеть смогла различить семь различных округов со средней точностью распознавания 91,2%. Для навигации в районе применялся метод поиска ближайшего соседа, на вход которого подавались признаки 14-го слоя нейронной сети VGG16, при этом для определения абсолютных координат привлекалась карта признаков геопривязанного эталонного изображения выбранного района. Результаты показали, что погрешность определения координат составила 200,75 м.

В работе [70] предложен подход, основанный на использовании многоступенчатой многозадачной нейронной сети CNN, которая за один проход изображения формирует семантическое сегментированное изображение и определяет координаты

<sup>1</sup> ADOP 2017 One Foot Ortho (raster) dataset. <https://gis.arkansas.gov/>; NAIP Imagery dataset. <https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index>

местоположения. На первом этапе применялась сегментация изображений с целью выделения дорог, на втором – реализация декодера, способного определять местоположение входного изображения и рассчитывать долготу и широту точки. После построения сети каждый из ее выходов обучался независимо таким образом, чтобы функция потерь представляла собой линейную комбинацию бинарных функций кросс-энтропии и потерь с использованием коэффициента Дайса (Dice) [71]. Эксперименты проводились на доступных наборах данных, таких как Inria Aerial Image Labeling dataset<sup>1</sup> и Massachusetts Buildings<sup>2</sup>.

В представленных ниже работах с помощью глубокой нейронной сети решается задача навигации в локальных областях отдельных домов и кварталов, при этом определяются не только координаты, но и углы ориентации БПЛА. В качестве БПЛА зачастую служил квадрокоптер, осуществляющий полет в городской среде.

Так, в [72] для определения шести параметров положения квадрокоптера производилось обучение нейронной сети PoseNet. Алгоритм обеспечивал непрерывную оценку местоположения, даже когда сцена включала динамические объекты – людей или автомобили или характеризовалась меняющимися погодными условиями. В качестве основы для разработки нейронной сети PoseNet применялась нейронная сеть GoogLeNet [73], в которой слои обобщенной логистической функции (softmax) были заменены блоками для расчета аффинных преобразований, а на последний из них добавлен еще один полносвязанный слой. Для обучения и тестирования использовался датасет уличных городских объектов Cambridge Landmarks<sup>3</sup>. Эксперименты показали, что нейронная сеть PoseNet обеспечивает погрешность определения координат на уровне 2 м и ориентации на уровне 3° на площади 50000 м<sup>2</sup>.

В [74] использовалась нейронная сеть SqueezePoseNet [75] – адаптация нейронной сети SqueezeNet для более мобильных платформ [76]. В то время как SqueezeNet предназначена для решения задач классификации, SqueezePoseNet модифицирована для решения задач регрессии. Для оценки позиции привлекались два полносвязанных слоя (положение и поворот). В качестве функций активации выступали Leaky ReLU [77], а для более высокой скорости обучения применялись слои пакетной нормализации данных. Повышение обобщающей способности сети осуществлялось за счет процедуры аугментации данных. SqueezePoseNet была инициализирована с учетом параметров, обученных на основе набора данных ImageNet [66]. Последние полносвязанные слои были дообучены на наборах данных Shop Facade<sup>2</sup> и Atrium<sup>4</sup>. Эксперименты показали, что SqueezePoseNet обеспечивает погрешность определения координат на уровне 5,19 м и ориентации на уровне 29,28°.

### Выводы

Представленные в группе алгоритмы используются как для глобального позиционирования объекта на достаточно больших пространствах (район, округ, штат), так и для навигации в локальных зонах. В первом случае вырабатываются только координаты местоположения объекта, при этом оценка точности зачастую основывается не на по-

---

<sup>1</sup> Inria Aerial Image Labeling Dataset. <https://project.inria.fr/aerialimagelabeling/>

<sup>2</sup> Massachusetts Buildings Dataset. <https://www.kaggle.com/balraj98/massachusetts-buildings-dataset>

<sup>3</sup> Cambridge Landmarks dataset. [mi.eng.cam.ac.uk/projects/relocalisation/](http://mi.eng.cam.ac.uk/projects/relocalisation/)

<sup>4</sup> Atrium dataset. [https://www2.ipf.kit.edu/pcv2016/downloads/photos\\_atrium\\_recon-struction.zip](https://www2.ipf.kit.edu/pcv2016/downloads/photos_atrium_recon-struction.zip)

грешностях определения координат, а на вероятности попадания в заданную область. Во втором случае находятся не только координаты, но и параметры ориентации БПЛА.

Рассмотренные алгоритмы не могут быть реализованы в пространствах, отличных от исходных, так как требуют процедур дообучения на данных из новой локации. Кроме того, они более остальных зависят от объема данных, используемых для обучения, это заставляет исследователей создавать собственные датасеты, отличающиеся по разрешению и размеру, а также не учитывают ряд факторов, например погодные условия.

В целом применение с этими целями нейронных сетей активно развивается. Основное препятствие, которое необходимо преодолеть, заключается в расширении пространства, доступного для локальной навигации. Алгоритмы, рассмотренные в настоящем разделе, пока не обеспечивают ее высокой точности, но, возможно, новые архитектуры сетей позволят достичь и новых уровней точности.

Результаты анализа описанных выше алгоритмов представлены в табл. 3.

Таблица 3

### Сравнительные характеристики алгоритмов навигации на основе сопоставления изображений с использованием нейронных сетей

Исследование	Используемые алгоритмы	Обучающие изображения / размер ш [м] × д [м]	Изображения с БПЛА / размер ш [м] × д [м]	Погрешность, компоненты вектора состояния
[67]	Xception	12 датасетов / 1500 × 1500 6,7 м/п	1 набор данных / 1500 × 1500 6,7 м/п	AED = 166,5 м (x, y)
[69]	VGG16	Google Maps / 5251 × 9421, 3960 × 4874, 3960 × 2741	Смоделировано с Bing Maps / 595 × 595 1,19 м/п	AED = 200,75 м (x, y)
[70]	MSMT-Net	MBD, IAILD	MBD, IAILD	96,84% для 20 м допуска (x, y)
[72]	PoseNet	Cambridge Landmarks	Cambridge Landmarks	AED = 2 м (x, y, z) и AED = 3° (φ, θ, ψ)
[74]	SqueezePoseNet	Shop Facade и Atrium	Shop Facade и Atrium	AED = 5,19 м (x, y, z) и AED = 29,28° (φ, θ, ψ)

Примечание: AED - среднее евклидово расстояние.

### Заключение

В статье рассмотрены исследования в области навигации БПЛА с использованием систем технического зрения, выполненные преимущественно с 2015 года. Работы сгруппированы следующим образом:

- алгоритмы навигации на основе сопоставления изображений с использованием корреляционно-экстремального подхода;

- алгоритмы навигации на основе сопоставления изображений с использованием ключевых точек;
- алгоритмы навигации на основе сопоставления изображений с использованием нейронных сетей.

Алгоритмы, основанные на сопоставлении изображений с помощью корреляционных подходов, зачастую предполагают применение сегментационных нейронных сетей для предобработки изображений, однако исследований в этом направлении проводится достаточно мало в силу низкой производительности таких подходов. В целом эти алгоритмы могут использоваться для навигации на местности определенного типа, на которой, как правило, имеется одноэтажная жилая застройка и густая дорожная сеть, или на больших высотах с привлечением данных гидрографии. Такие алгоритмы также могут рассматриваться как корректирующие для основной системы навигации.

Алгоритмы, основанные на сопоставлении ключевых точек изображений, из всех представленных наиболее широко распространены. Детекторы позволяют извлечь из изображений самую ценную и устойчивую к искажениям информацию и обеспечивают высокую надежность и точность определения навигационных параметров. Развитие детекторов в настоящее время осуществляется в направлении использования нейронных сетей для извлечения ключевых точек или формирования их более устойчивых дескрипторов, позволяющих точно сопоставить эталонное изображение и изображение, формируемое БПЛА. В алгоритмах этой группы также применяется оптический поток при сравнении ключевых точек с целью уменьшить область их поиска на изображении. Алгоритмы, основанные на сопоставлении ключевых точек, обычно позволяют эффективно оценить координаты и ориентацию (азимутальную) БПЛА и могут надежно использоваться для навигации над протяженными пространствами, не содержащими большого количества подвижных объектов.

Алгоритмы, базирующиеся на глубоких нейронных сетях для сопоставления изображений, в настоящее время менее изучены и обеспечивают невысокую точность. Применение таких алгоритмов при глобальном позиционировании БПЛА на достаточно больших пространствах (район, округ, штат) позволяет оценивать только координаты, а при навигации в локальных зонах – все 6 параметров движения объекта. Развитие методов третьей группы в направлении минимизации размеров сетей с целью внедрения в бортовые вычислители БПЛА и улучшения обобщающих свойств для расширения области действия создаваемых навигационных систем представляется наиболее перспективным.

#### ЛИТЕРАТУРА

1. **Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., and Gattelli, M.,** Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images, *Remote Sensing*, 2015, vol. 7, no. 4, pp. 4026–4047, doi: 10.3390/rs70404026.
2. **Akhloufi, M.A., Castro, N.A., and Couturier, A.,** UAVs for wildland fires, *Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*, 2018, vol. 10643, doi: 10.1117/12.2304834.
3. **Akhloufi, M.A., Castro, N.A., and Couturier, A.,** Unmanned aerial systems for wildland and forest fires: Sensing, perception, cooperation and assistance, *Drones*, 2021, vol. 5, no. 15, doi: 10.3390/drones5010015.
4. **Мокрова М.И.** Исследование влияния сложных условий пожарной обстановки на качество наблюдения и безопасность полёта БЛА // Известия ЮФУ. Технические науки. 2021. DOI: 10.18522/2311-3103-2021-1-112-124.

5. Jordan, S., Moore, J., Hovet, S., Box, J., Perry, J., Kirsche, K., Lewis, D., and Tse, Z.T.H., State-of-the-art technologies for UAV inspections, *IET Radar, Sonar & Navigation*, 2017, vol. 12, no. 2, pp. 151–164, doi:10.1049/iet-rsn.2017.0251.
6. Scherer, J., Yahyanejad, S., Hayat, S., Yanmaz, E., Andre, T., Khan, A., Vukadinovic, V., Bettstetter, C., Hellwagner, H., and Rinner, B., An Autonomous Multi-UAV System for Search and Rescue, *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use, DroNet '15, ACM, New York, USA*, 2015, pp. 33–38, doi: 10.1145/2750675.2750683.
7. Mittal, M., Mohan, R., Burgard, W., and Valada, A., Vision-Based Autonomous UAV Navigation and Landing for Urban Search and Rescue, *Proceedings of the International Symposium on Robotics Research (ISRR)*, 2019, doi: 10.48550/arXiv.1906.01304.
8. Зоев И.В., Марков Н.Г., Рыжова С.Е. Интеллектуальная система компьютерного зрения беспилотных летательных аппаратов для мониторинга технологических объектов предприятий нефтегазовой отрасли // Известия Томского политехнического университета. Инжиниринг георесурсов. 2019. Т. 330. № 11. С. 34–49.
9. De Melo, C.F.E., Dapper e Silva, T., Boeira, F., Stocchero, J.M., Vinel, A., Asplund, M., and De Freitas, E.P., UAVouch: A Secure Identity and Location Validation Scheme for UAV-Networks, *IEEE Access*, 2021, vol. 9, pp. 82930–82946.
10. Mantelli, M., Pittol, D., Neuland, R., Ribacki, A., Maffei, R., Jorge, V., Prestes, E., and Kolberg, M., A novel measurement model based on abBRIEF for global localization of a UAV over satellite images, *Robotics and Autonomous Systems*, 2019, vol. 112, pp. 304–319, doi: 10.1016/j.robot.2018.12.006.
11. Пазычев Д.Б. Мкртчян В.И. Адаптивный субоптимальный фильтр Калмана в задаче выставки БИНС // Известия Тульского государственного университета. Технические науки 2018. № 5. С. 60–73.
12. Conte, G. and Doherty, P., An integrated UAV navigation system based on aerial image matching, *Aerospace Conference*, 2008, pp. 1–10.
13. Viswanathan, A., Pires, B.R., and Huber, D., Vision-based robot localization across seasons and in remote locations, *International Conference on Robotics and Automation*, 2016, pp. 4815–4821.
14. Шмидт Дж.Т. Эксплуатация навигационных систем на основе GPS в сложных условиях окружающей среды // Гироскопия и навигация. 2019. Том 27. № 1 (104). С. 3–21.
15. Пешехонов В.Г. Высоточная навигация без использования информации глобальных навигационных спутниковых систем // Гироскопия и навигация. 2022. Том 30. № 1 (116). С. 3–11.
16. Sabatini, R., Moore, T., Hill, C., and Ramasamy, S., Avionics-based GNSS integrity augmentation performance in a jamming environment, *AIAC16: 16th Australian International Aerospace Congress, Engineers Australia*, 2015, pp. 469–479.
17. Groves, P.D., Jiang, Z., Rudi, M., and Strode, P., A portfolio approach to NLOS and multipath mitigation in dense urban areas, *Proceedings of the 26th International Technical Meeting of The Satellite Division of the Institute of Navigation*, The Institute of Navigation, 2013.
18. Жук Р.С., Залесский Б.А., Троцкий Ф.С. Визуальная навигация автономно летящего БПЛА с целью его возвращения в точку старта // Информатика. 2020. Т. 17. № 2. С. 17–24. DOI: 10.37661/1816-0301-2020-17-2-17-24.
19. Yuncheng, L., Zhucun, X., Gui-Song, X., and Liangpei, Z., A survey on vision-based UAV navigation, *Geo-spatial Information Science*, 21:1, 2018, pp. 21–32, doi: 10.1080/10095020.2017.1420509.
20. Визильгер Ю.В., Желтов С.Ю. Использование глубоких нейронных сетей для анализа данных, управления и оптимизации в перспективных авиационных приложениях // XII мультиконференция по проблемам управления (МКПУ-2019). 2019. Т. 4. С. 17–20.
21. Белоглазов И.Н., Тарасенко В.П. Корреляционно-экстремальные системы. Москва: Сов. радио, 1974. 392 с.
22. Белоглазов И.Н., Джанджгава Г.И. Основы навигации по геофизическим полям. М.: Наука, гл. ред. физ.-мат. лит., 1985. 328 с.
23. Степанов О.А., Торопов А.Б. Методы нелинейной фильтрации в задаче навигации по геофизическим полям. Часть 2. Современные тенденции развития // Гироскопия и навигация. 2015. № 4 (91). С. 147–159.
24. Беркович С.Б., Котов Н.И., Лычагов А.В. Система технического зрения как источник дополнительной информации в задаче автомобильной навигации // Гироскопия и навигация. 2017. Т. 25. № 1 (96). С. 49–63.
25. Шолохов А.В., Беркович С.Б. Формирование траектории корреляционно-экстремальной навигационной системы по критерию минимума погрешностей координат // XXV юбилейная Санкт-Петербургская конференция по геоинформатике и навигации. 2022. Т. 25. № 1. С. 1–10.



- тербургская международная конференция по интегрированным навигационным системам. 2018. С. 175–177.
26. **Степанов О.А., Носов А.С.** Алгоритм коррекции навигационной системы по данным карты и измерителя, не требующий предварительного оценивания значений поля вдоль пройденной траектории // *Гироскопия и навигация*. 2020. Т. 28. № 2 (109). С. 70–90.
  27. **Козубовский С.Ф.** Корреляционные экстремальные системы. Справочник. Киев: Наукова думка, 1973. 224 с.
  28. **Степанов О.А.** Методы оценки потенциальной точности в корреляционно-экстремальных навигационных системах. Аналитический обзор. СПб.: ЦНИИ «Электронприбор», 1993. 85 с.
  29. **Poddar, Sh., Kottath, R., Karar, V.,** Evolution of Visual Odometry Techniques, 2018. URL: <https://arxiv.org/ftp/arxiv/papers/1804/1804.11142.pdf>.
  30. **Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I.D., Leonard, J.J.,** Simultaneous Localization and Mapping: Present, Future, and the Robust-Perception Age, *IEEE Transactions on Robotics* (cond. Accepted), 2016.
  31. **Ronneberger, O., Fischer, P., and Brox, T.,** U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015, arXiv: 1505.04597.
  32. **Lowe, D.G.,** Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 2004, vol. 60, no. 2, pp. 91–110.
  33. **Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A.A.,** Image-To-Image translation with conditional adversarial networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134, doi: 10.1109/CVPR.2017.632.
  34. **Mirza, M. and Osindero, S.,** Conditional Generative Adversarial Nets, 2014, arXiv:1411.1784 [cs, stat].
  35. **Schleiss, M.,** Translating aerial images into street-map representations for visual self-localization of UAVs, *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019, vol. 42, pp. 575–580, DOI: 10.5194/isprs-archives-XLII-2-W13-575-2019.
  36. **Silva Filho, P., Shiguemori, E.H., Saotome, O.,** UAV visual autolocalization based on automatic landmark recognition, *International Conference on Unmanned Aerial Vehicles in Geomatics*, 2017, pp. 89–94. doi:10.5194/isprs-annals-IV-2-W3-89-2017.
  37. **Alcantarilla, P.F., Nuevo, J., Bartoli, A.,** Fast explicit diffusion for accelerated features in nonlinear scale spaces, *British Machine Vision Conference (BMVC)*, 2013, doi: 10.5244/C.27.13.
  38. **Ding, L., Zhou, J., Meng, L., Long, Z.,** A Practical Cross-View Image Matching Method between UAV and Satellite for UAV-Based Geo-Localization, *Remote Sensing*, 2021, vol. 13, no. 47, doi:10.3390/rs13010047.
  39. **Zheng, Z., Wei, Y., Yang, Y.,** University-1652: A Multi-View Multi-Source Benchmark for Drone-Based Geo-Localization, *Proceedings of the 28th ACM International Conference on Multimedia*, Seattle, WA, USA, October 2020, pp. 1395–1403.
  40. **Uchida, Yusuke,** Local Feature Detectors, Descriptors, and Image Representations: A Survey, 2016, ArXiv abs/1607.08368.
  41. **Goforth, H. and Lucey, S.,** GPS-Denied UAV Localization using Pre-existing Satellite Imagery, *International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, 2019, pp. 2974–2980, doi: 10.1109/ICRA.2019.8793558.
  42. **Simonyan, K. and Zisserman, A.,** Very deep convolutional networks for large-scale image recognition, *Conference ICLR*, 2015, arXiv:1409.1556 [cs].
  43. **Fischler, M.A. and Bolles, R.C.,** Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, 1981, vol. 24, no. 6, pp. 381–395, doi: 10.1145/358669.358692.
  44. **Rublee, E., Rabaud, V., Konolige, K., and Bradski, G.,** ORB: An efficient alternative to SIFT or SURF, *International Conference on Computer Vision*, 2011, pp. 2564–2571, doi: 10.1109/ICCV.2011.6126544.
  45. **Seema, B.S., Hemanth, K., and Naidu, V.P.S.,** Geo-Registration of Aerial Images using RANSAC Algorithm, *NCTAESD-2014*, Vemana Institute of Technology, Bangalore, 2014, pp. 1–5.
  46. **Saranya, K.C., Naidu, V.P.S., Singhal, V., and Tanuja, B.M.,** Application of vision-based techniques for UAV position estimation, *International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, 2016, pp. 1–5, doi: 10.1109/RAINS.2016.7764392.
  47. **Wang, X., Kealy, A., Li, W., Jelfs, B., Gilliam, C., May, S.L., and Moran, B.,** Toward Autonomous UAV Localization via Aerial Image Registration, *Electronics*, 2021, vol. 10, no. 4, doi: 10.3390/electronics10040435.

48. Bay, H., Tuytelaars, T., and Van Gool, L., SURF: Speeded Up Robust Features, *A. Leonardis, H. Bischof, A. Pinz (Eds.), Computer Vision ECCV 2006, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2006, pp. 404–417, doi: 10.1007/11744023\_32.
49. Степанов Д.Н. Применение вейвлетов габора в задаче навигации БПЛА с использованием видеокамеры // *Фундаментальные исследования*. 2015. №12 (часть 1). С. 85–91.
50. Степанов Д.Н. Методы и алгоритмы определения положения и ориентации беспилотного летательного аппарата с применением бортовых видеокамер // *Программные продукты и системы*. 2014. Т. 1. №1.
51. Nassar, A., Amer, K., El Hakim, R., and ElHelw, M., A deep CNN-based framework for enhanced aerial imagery registration with applications to UAV geolocalization, *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1513–1523, doi: 10.1109/CVPRW.2018.00201.
52. Calonder, M., Lepetit, V., Strecha, C., and Fua, P., BRIEF: Binary Robust Independent Elementary Features, *K. Daniilidis, P. Maragos, N. Paragios (Eds.), Computer Vision ECCV 2010, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2010, pp. 778–792, doi: 10.1007/978-3-642-15561-1\_56.
53. Masselli, A., Hanten, R., and Zell, A., Localization of Unmanned Aerial Vehicles Using Terrain Classification from Aerial Images, *Intelligent Autonomous Systems, Advances in Intelligent Systems and Computing*, Springer International Publishing, 2016, vol. 13, pp. 831–842, doi: 10.1007/978-3-319-08338-4\_60.
54. Breiman, L., Random Forests, *Machine Learning*, 2001, vol. 45, no. 1, pp. 5–32, doi: 10.1023/A:1010933404324.
55. Bianchi, M. and Barfoot, T.D., UAV Localization Using Autoencoded Satellite Images, *IEEE Robotics and Automation Letters*, 2021, vol. 6, no. 2, pp. 1761–1768.
56. Shan, M., Wang, F., Lin, F., Gao, Z., Tang, Y.Z., and Chen, B.M., Google map aided visual navigation for UAVs in GPS-denied environment, *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 114–119, doi: 10.1109/ROBIO.2015.7418753.
57. Dalal, N. and Triggs, B., Histograms of Oriented Gradients for Human Detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, San Diego, CA, USA, 2005, vol. 1, pp. 886–893, doi: 10.1109/CVPR.2005.177.
58. Horn, B.K.P. and Schunck, B.G., Determining optical flow, *Artificial Intelligence*, 1981, vol. 17, no. 1, pp. 185–203, doi: 10.1016/0004-3702(81)90024-2.
59. Mughal, M.H., Khokhar, M.J., and Shahzad, M., Assisting UAV Localization Via Deep Contextual Image Matching, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, vol. 14, pp. 2445–2457.
60. He, K., Zhang, X., Ren, S., and Sun, J., Deep Residual Learning for Image Recognition, 2015, arXiv:1512.03385v1 [cs.CV].
61. Rocco, I., Cimpoi, M., Arandjelovic, R., Torii, A., Pajdla, T., and Sivic, J., Neighbourhood Consensus Networks, *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, Canada, 2018.
62. Everingham, M., Gool, L.V., Williams, C.K.I., Winn, J., and Zisserman, A., The PASCAL Visual Object Classes (VOC) Challenge, *International Journal of Computer Vision*, 2010, vol. 88, pp. 303–338.
63. Taira, H., Okutomi, M., Sattler, T., Cimpoi, M., Pollefeys, M., Sivic, J., Pajdla, T., and Torii, A., InLoc: Indoor Visual Localization with Dense Matching and View Synthesis, 2018. arXiv:1803.10368 [cs].
64. Танченко А.П., Федулин А.М., Бикмаев Р.Р., Садеков Р.Н. Алгоритм автономной коррекции навигационной системы беспилотного летательного аппарата на основе распознавания дорожной и речной сети // *Гироскопия и навигация*. 2020. Т. 28. № 3(110). С. 3–13. DOI: 10.17285/0869-7035.0038.
65. Krizhevsky, A., Sutskever, I., and Hinton, G.E., Imagenet classification with deep convolutional neural networks, *Communications of the ACM*, 2017, vol. 60, no. 6, pp. 84–90, doi: 10.1145/3065386.
66. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., et al., ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision (IJCV)*, 2015, vol. 115, no. 3, pp. 211–252, doi: 10.1007/s11263-015-0816-y.
67. Harvey, W., Rainwater, C., and Cothren, J., Direct Aerial Visual Geolocalization Using Deep Neural Networks, *Remote Sensing*, 2021, vol. 13, doi: 10.3390/rs13194017.
68. Chollet, F., Xception: Deep Learning with Depthwise Separable Convolutions, 2017. arXiv:1610.02357 [cs].
69. Amer, K., Samy, M., El Hakim, R., Shaker, M., and El Helw, M., Convolutional Neural Network-Based Deep Urban Signatures with Application to Drone Localization, *IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2017, pp. 2138–2145, doi: 10.1109/ICCVW.2017.
70. Marcu, A., Costea, D., Slusanschi, E., and Leordeanu, M., A Multi-Stage Multi-Task Neural Network for Aerial Scene Interpretation and Geolocalization, 2018, arXiv:1804.01322 [cs].

71. **Jadon, S.**, A survey of loss functions for semantic segmentation, *IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 2020, doi: 10.1109/CIB-CB48159.2020.9277638.
72. **Kendall, A., Grimes, M., and Cipolla, R.**, PoseNet: A convolutional network for real-time 6-dof camera relocalization, *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015, pp. 2938–2946.
73. **Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.**, Going deeper with convolutions, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 2015, pp. 1–9.
74. **Mueller, M.S. and Jutzi, B.**, UAS Navigation with SqueezePoseNet – Accuracy Boosting for Pose Regression by Data Augmentation, *Drones*, 2018, vol. 2, no. 7, doi: 10.3390/drones2010007.
75. **Mueller, M.S., Urban, S., and Jutzi B.**, SqueezePoseNet: Image based Pose Regression with small Convolutional Neural Networks for Real Time UAS Navigation, *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, 2017, vol. 4, pp. 49–57.
76. **Iandola, F.N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W.J., and Keutzer, K.**, SqueezeNet: AlexNet-level accuracy with 50× fewer parameters and <0.5 mb model size, 2016, arXiv:1602.07360.
77. **Maas, A.L., Hannun, A.Y., and Ng, A.Y.**, Rectifier nonlinearities improve neural network acoustic models, *Proceedings of the International Conference on Machine Learning*, Atlanta, GA, USA, 2013, vol. 30.

---

**Ali B., Sadekov, R.N.** (The National University of Science and Technology MISIS, Russia), and **Tsodokova, V.V.** (Concern CSRI Elektropribor, JSC, St. Petersburg, Russia)

A Review of Navigation Algorithms for Unmanned Aerial Vehicles Based on Computer Vision Systems, *Girokopiya i Navigatsiya*, 2022, vol. 30, no. 4 (119), pp. 87–105.

**Abstract.** The article analyzes the works devoted to outdoor navigation of unmanned aerial vehicles (UAVs) in GNSS-denied environments using computer vision systems (optical range cameras). The algorithms addressed are based on matching of UAV-generated images with the available georeferenced terrain images. Images are matched either pixel by pixel, by their key points, or using neural networks. The stages of each algorithm, as well as the accuracy achieved by the authors and the field data used for testing are considered. The paper concludes with a discussion of the capabilities and limitations of the proposed approaches.

**Key words:** unmanned aerial vehicle, autonomous navigation, key point matching, descriptor, detector, deep learning, computer vision, satellite imagery, correlation-extreme systems.

Материал поступил 09.05.2022