

УДК 004.8, 004.93

DOI 10.24412/2413-7383-2024-4-101-112

Б. В. Павленко, Я. С. Пикалёв

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

МЕТОДИКА СОЗДАНИЯ НАБОРА АЭРОФОТОСНИМКОВ ДЛЯ ЗАДАЧИ ПЕРЕКРЁСТНОЙ ГЕОЛОКАЛИЗАЦИИ*

Pavlenko B.V., Ya. S. Pikalyov

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artema str, 118-b

METHODOLOGY FOR CREATING AN AERIAL IMAGE DATASET FOR CROSS-GEOLOCATION TASKS

В данной работе описана методика создания обширного и доступного набора данных аэрофотоснимков на основе использования сторонних картографических сервисов, таких как Google Maps, Google Earth и Google Earth Studio. Предложен алгоритм автоматизированного сбора данных, включающий разбиение целевой области карты на ячейки фиксированного размера и получение двух типов изображений: со спутника и с моделируемого вида дрона. Выделены преимущества предложенной автоматизации, такие как экономия времени и ресурсов, и сложность, связанная с динамическим форматом страницы и зависимостью от функциональности используемой платформы.

Ключевые слова: аэрофотоснимки, компьютерное зрение, распознавание изображений, спутниковые изображения, 3D-визуализация, БПЛА, перекрёстная геолокализация.

Computer vision systems used in unmanned aerial vehicles (UAVs) play a crucial role in a wide range of tasks, including terrain classification, infrastructure monitoring, emergency detection, and identifying transportation and other objects. However, their application in UAV-based computer vision systems requires comprehensive aerial image datasets. The high costs and restrictions associated with aerial data acquisition under certain conditions pose significant challenges for direct data collection. This study presents a methodology for creating a comprehensive and accessible aerial image dataset using third-party cartographic services such as Google Maps, Google Earth, and Google Earth Studio. An algorithm for automated data collection is proposed, which involves dividing the target map area into fixed-size cells and obtaining two types of images: satellite-view and simulated drone-view. The advantages of the proposed automation approach include time and resource efficiency, though it also highlights challenges related to the dynamic nature of web pages and dependence on the functionality of the selected platform.

Keywords: aerial images, computer vision, image recognition, satellite images, 3D visualization, UAVs, cross-view geo-localization.

* Работа выполнена в рамках федерального проекта «Развитие человеческого капитала в интересах регионов, отраслей и сектора исследований и разработок» национального проекта «Наука и университеты» по теме научной молодежной лаборатории «Извлечение семантической информации из изображений для автономных систем навигации беспилотных летательных аппаратов (FREN-2024-0002)».

Введение

Аэрофотоснимки являются важными источниками данных для таких областей, как компьютерное зрение, дистанционное зондирование, геоинформационные системы и других. При этом системы компьютерного зрения, используемые в БПЛА, позволяют решать задачи, связанные с мониторингом и классификацией ландшафтов, обнаружением изменений в окружающей среде, развитием инфраструктуры, прецизионного земледелия [1-4], системами навигации беспилотных аппаратов. Однако, сбор реальных данных может быть затруднителен или даже невозможен по ряду причин, таких, как запреты на полеты БПЛА, затраты на проведение полетов, доступность интересующей зоны, погодные условия.

Современные картографические сервисы, такие как Google Earth [5] и Google Earth Studio [6], предоставляют доступ к обширной базе спутниковых снимков и 3D-карт местности, а также прямую совместимость с популярными системами для 3D-моделирования, такие как, Blender и Unreal Engine 5. Открытые картографические ресурсы могут быть использованы для создания экономичных, с точки зрения затрачиваемых ресурсов, наборов данных.

Постановка задачи

Программные средства сбора данных через специализированный программный интерфейс приложения (*Application Programming Interface, API*) или напрямую через скрипты позволяют применить методы автоматизации для систематического сбора изображений. **Цель данной работы** заключается в создании методики, включающей в себя пошаговый алгоритм предобработки координат, а именно – разбиения выбранного участка местности на сетку и сбор данных для каждой ячейки (или патча) с двух точек обзора: со спутника (*satellite-view*) и с видов, аналогичных рабочим высотам БПЛА (*drone- или uav-view*). При этом процесс сбора предполагает наличие пар *drone-view -> satellite-view*. Стоит отметить, что для *satellite-view* используется моделирование полёта по спирали, изменяя угол наклона и высоту с 250 м до 150 м. Такой подход позволяет обеспечить сбор разноплановых представлений местности для использования в исследованиях, в частности, для обучения моделей.

Сбор данных выполняется для создания набора данных особого формата для задачи перекрестной геолокализации (*cross-view geolocalization, CVGL*) [7-14]. Данная задача является расширением задачи поиска схожих снимков, а именно сравнения аэрофотоснимков *drone-view* с заранее подготовленными снимками *satellite-view* с соответствующими геопозициями. Дополнительно CVGL предназначена для обновления существующих снимков *satellite-view*. В работах [15], [16] был собран набор изображений зданий университетов, совмещающий спутниковые снимки местности и множество снимков зданий с БПЛА для обеспечения наличия признаков, инвариантных к точке обзора.

Таким образом, в работе были поставлены следующие задачи.

1. Проанализировать существующие наборы данных для задачи CVGL.
2. Проанализировать существующие картографические сервисы для получения снимков *drone-view* и *satellite-view*.
3. Разработать методику автоматического сбора данных:
 - алгоритм разбиения целевой области карты на фиксированные ячейки;
 - алгоритм последовательного получения визуальных и числовых данных для каждой ячейки карты;
 - алгоритм учёта зависимости данных от высоты и угла полёта.

Анализ наборов данных для задачи CVGL

Как было указано выше, снимки с БПЛА не всегда можно получить «вживую» при помощи реальных полетов, поэтому данные, как правило, собираются при помощи 3D-карт соответствующих сервисов. Краткий обзор существующих наборов данных, преимущественно видов с дрона, отраженный в [7, 13, 15-16] приведен в табл. 1. Существующие наборы могут быть полезны для типовых задач, но в некоторых случаях может потребоваться создание собственного набора аэрофотоснимков.

Таблица 1 – Характеристика наборов данных аэрофотоснимков

Наборы данных	Размер	Содержание	Высота камеры	Максимальное разрешение
UAVDT	80 тыс. изображений	транспортные средства на дороге	10 ~ 70 м	1080 × 540
VisDrone	10 тысяч изображений, 288 видео	люди и транспортные средства в повседневных сценах	-	изображение: 2000 × 1500 видео: 3840 × 2160
Stanford Drone	-	люди и транспортные средства на дороге	80 м	1400 × 1904
UAV123	110 тыс. изображений	автомобили, грузовики, корабли, люди и т.д.	Набор 1: 5~25 м Набор 2, Набор 3: недоступно	1280 × 720
Okutama-Action	77 тыс. изображений	люди, выполняющие различные действия	10 ~ 45 м	3840 × 2160
VIRAT2.0	-	люди и транспортные средства в естественных сценах	-	1920 × 1080
UCLA Aerial	27 видео	люди в сценах на открытом воздухе	25 м	-
Mini-Drone-Video	-	люди и транспортные средства на парковке	-	1920 × 1080
CARPK	1448 изображений	транспортные средства на парковках	40 м	-

Обзор картографических сервисов

Среди рассмотренных решений были такие сервисы, как Google Earth, Google Earth Studio, Google Earth Engine, Folium, Cesium. Характеристики вышеуказанных систем приведены в табл. 2. Отбор проводился по таким критериям, как качество получаемых снимков и 3D-моделей, доступность дополнительных модулей для систем, наличие API, метод и скорость сохранения извлеченной информации.

Средства построения 3D-карт с помощью фотограмметрии, ортопланов на данный момент не представляют практического интереса, т. к. подразумевают создание моделей местности на основе фотографий с БПЛА [17-19], отсутствие которых решается предлагаемой методикой в данной работе.

Таблица 2 – Характеристики систем картографии

Сервис	3D	API	Особенности
Google Earth Engine	-	+	анализ геоданных, интеграция спутниковых данных, API для научных исследований
Google Earth Studio	+	-	анимации и рендеринг спутниковых изображений, поддержка 3D-картографирования
Google Maps	ограничено	+	стандартные спутниковые карты
Cesium	+	+	поддержка 3D-картографирования, API для кастомизации, используется для анализа данных и анимации

Метод сбора данных

Предлагаемая методика основана на алгоритме разбиения целевой области карты на фиксированные ячейки и последовательное получение визуальных и числовых данных для каждой ячейки карты. Сбор выполняется с использованием библиотеки Selenium, что позволяет автоматизировать сбор данных с использованием веб-интерфейса.

Основные этапы работы алгоритма инициализации карты и сбора снимков satellite-view отображены на рис. 1. Опорной точкой служит верхний левый угол участка на карте, а ширина и высота задаются в метрах. Эта область разбивается на заданное количество ячеек. Общая область делится на сетку из n ячеек по широте и долготу. Расстояние между центрами соседних ячеек рассчитывается с учетом размеров области и количества ячеек, чтобы обеспечить равномерное покрытие.

В отличие от работ [15], [16], где ключевые метки были заданы с учётом, извлечённых метаданных из Википедии (картографическое положение зданий университетов), предложенный подход не ориентирован на отдельные объекты или здания. Достаточно получить снимки drone-view для разных перспектив (с разной высоты и наклонов камеры), т. к. все ячейки расположены рядом в одной области.

Вид со спутника собирается с Google Earth. Отображение настраивается с помощью Selenium. Каждый снимок привязывается к координатам центральной точки ячейки. Вид с дрона собирается с помощью 3D-визуализации Google Earth Studio. На этом этапе применяется методика последовательного получения скриншотов, чтобы собрать изображения объекта с нескольких ракурсов. Большое число ракурсов достигается с помощью имитации траектории полета дрона. Среда предоставляет возможность задания пути для камеры, изменение углов наклона динамически в заданном диапазоне.

Координаты карты рассчитываются в несколько шагов. Под расчетами координат понимается разбиение карты на заданное число ячеек и сохранение координат центров ячеек в список.

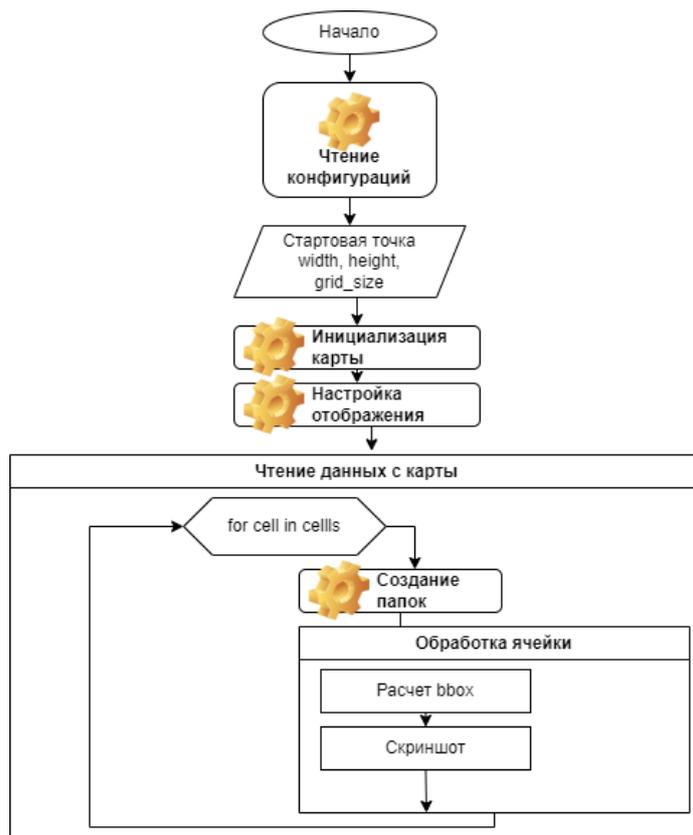


Рисунок 1 – Общая схема алгоритма

Расчеты шагов по широте и долготе требуют использования констант. Так, например, расчет размера шага долготы зависит от широты (1) [20], где L – приблизительное значение одного градуса в метрах, равное 111320 м:

$$step_{longitude} = \frac{m}{L \cdot \cos latitude}, \quad (1)$$

Используем косинус широты из-за пропорциональности изменения градуса долготы.

Координаты краев карты рассчитываются по следующим формулам:

$$p_1 = (sp_{lat}, sp_{long}) \quad (2)$$

$$p_2 = (sp_{lat}, sp_{long} + delta_{long}) \quad (3)$$

$$p_3 = (sp_{lat} - delta_{lat}, sp_{long} + delta_{long}) \quad (4)$$

$$p_4 = (sp_{lat} - delta_{lat}, sp_{long}) \quad (5)$$

$$delta_{lat} = \frac{h}{L}, \quad (6)$$

где sp_{lat} , sp_{long} – координаты стартовой точки, а $delta_{lat}$, $delta_{long}$ – шаги широты и долготы по ячейкам карты соответственно. $delta_{long}$ рассчитывается по формуле (1).

Т.к. необходимо получить виды со спутника, на рабочей зоне карты выделяется окно, соответствующее масштабу по умолчанию, установленному измерением внутри самой карты. Так как точка центрируется относительно экрана, можно рассчитать размеры и координаты ограничивающей рамки для создания скриншотов.

Для этого достаточно рассчитать отступ от центра экрана до границы самой большой панели нерабочей области, такой как меню или панель инструментов. Обозначим эту величину $half_cell_side$.

Для снимков drone-view используется расширение набора при помощи 3D-моделей, предоставляемых Google Earth Studio для моделирования съёмки камеры с БПЛА. Эта операция предусмотрена в связи с тем, что сбор подобного набора данных является дорогостоящей процедурой. 3D-модель также предоставляет возможности для манипулирования точками обзора. Google Earth Studio также предоставляет возможность для манипулирования точками обзора.

Управляющие движением БПЛА величины выражаются через переменные, определяющие положение и ориентацию БПЛА в пространстве (рис. 2). Формула для следования БПЛА траектории имеет следующий вид:

$$\begin{aligned} \ddot{x}_T &= \ddot{x}_{des} + K_{p,x}(x_{des} - \dot{x}) + K_{d,x}(\dot{x}_{des} - \dot{x}), \\ \ddot{y}_T &= \ddot{y}_{des} + K_{p,y}(y_{des} - \dot{y}) + K_{d,y}(\dot{y}_{des} - \dot{y}), \\ \phi_{des} &= \frac{1}{g}(\ddot{x}_T \cdot \sin(\psi_{des}) - \ddot{y}_T \cdot \cos(\psi_{des})), \\ \theta_{des} &= \frac{1}{g}(\ddot{x}_T \cdot \cos(\psi_{des}) + \ddot{y}_T \cdot \sin(\psi_{des})), \\ p_{des} &= 0, q_{des} = 0, r_{des} = \psi_{des}. \end{aligned}$$

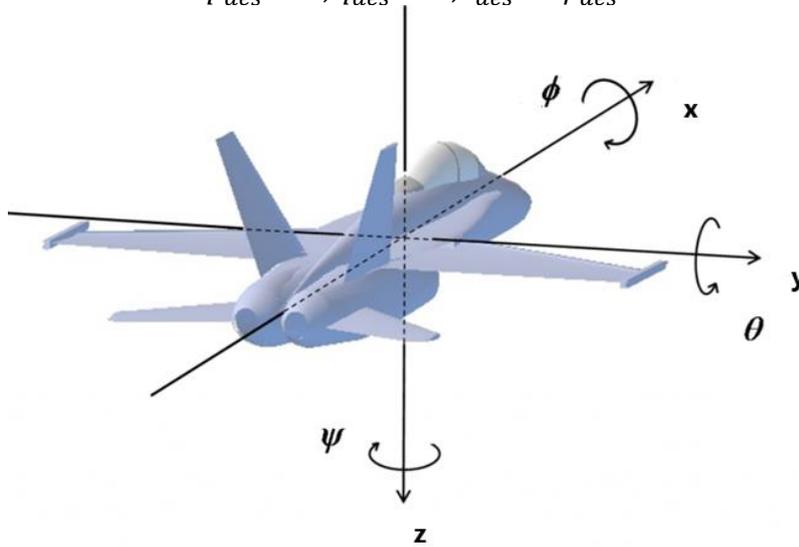


Рисунок 2 – Навигационные параметры БПЛА

С помощью данной системы уравнений реализована модель движение квадрокоптера по двум заданным траекториям: вдоль прямой линии и спирали

$$\vec{M} = \begin{bmatrix} K_{p,\phi}(\phi_{des} - \phi) + K_{d,\phi}(p_{des} - p) \\ K_{p,\theta}(\theta_{des} - \theta) + K_{d,\theta}(q_{des} - q) \\ K_{p,\psi}(\psi_{des} - \psi) + K_{d,\psi}(r_{des} - r) \end{bmatrix},$$

где p, q, r – угловые ускорения, x, y, z – подвижные оси, φ, θ, ψ – угловые положения БПЛА в виде трёх углов Крылова: угол крена, тангажа и рыскания, определяющие вращение вокруг осей.

Чтобы обеспечить возможность изменения масштаба и получения различных точек обзора, задается траектория полета в виде спирали и формируется видеозапись полета со скоростью 30 кадров в секунду. Камера совершает несколько поворотов, а высота постепенно уменьшается с максимальной h_{max} до минимальной h_{min} , что близко к высоте полета БПЛА в реальных условиях (рис.3) [[27] Troy A Rule. 2015. Airspace in an Age of Drones. BUL Rev. 95 (2015), 155.]. При этом снимки формируются с неким угловым шагом.

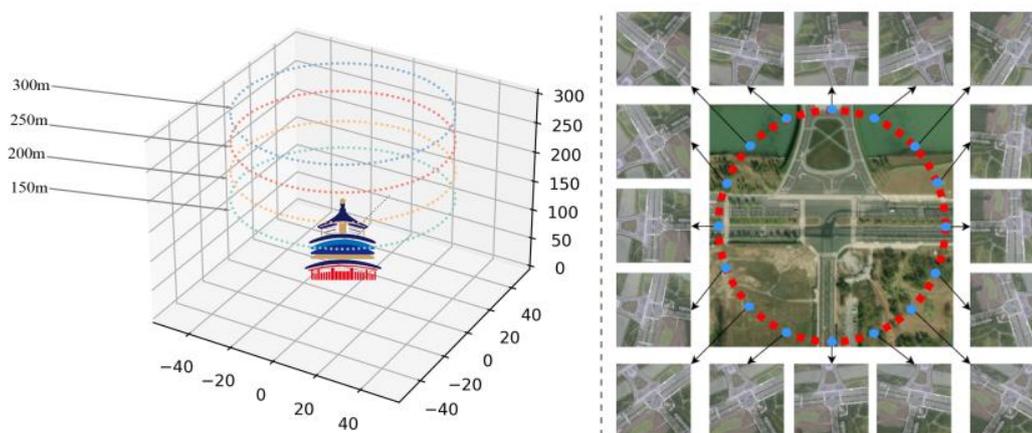


Рисунок 3 – Схема спиральной траектории БПЛА для получения drone-view снимков

Координаты центра экрана – результат деления ширины и высоты экрана на 2 без остатка. Вычтем из половины высоты нерабочую область верхней панели и получим половину ширины и высоты области захвата скриншота (7, 8). Высота h_{panel} измерена вручную.

$$hs_{half} = \frac{h_{win}}{2} - h_{panel}, \quad (7)$$

$$screen_{h/w} = 2hs_{half}, \quad (8)$$

где $screen_{h/w}$ – размеры области скриншота.

Тогда координаты верхнего и нижнего углов ограничивающей рамки будут рассчитаны как:

$$screen_{x_{topLeft}} = c_x - hs_{half}, \quad (9)$$

$$screen_{y_{topLeft}} = h_{panel}, \quad (10)$$

$$screen_{x_{bottomRight}} = screen_{x_{topLeft}} + screen_w, \quad (11)$$

$$screen_{y_{bottomRight}} = screen_{y_{topLeft}} + screen_h, \quad (12)$$

где hs_{half} – половина высоты экрана, c_x – координата центра экрана по горизонтали.

Так как ячейки сетки имеют фиксированный масштаб, зависящий от размера карты, то при необходимости захвата снимка в том же масштабе необходимо симулировать n нажатий соответствующих клавиш, где число n рассчитывается как:

$$n = \frac{|h_{cell} - h_{default}|}{\Delta h_{view}}, \quad (13)$$

где h_{cell} – расчетный размер стороны ячейки в метрах, $h_{default}$ – дистанция в метрах, попадающая в область скриншота при приближении камеры на карте по умолчанию, Δh_{view} – разница высоты камеры за один клик.

Результатом работы алгоритма является директория с собранными satellite-view снимками и снимками drone-view (рис.4).

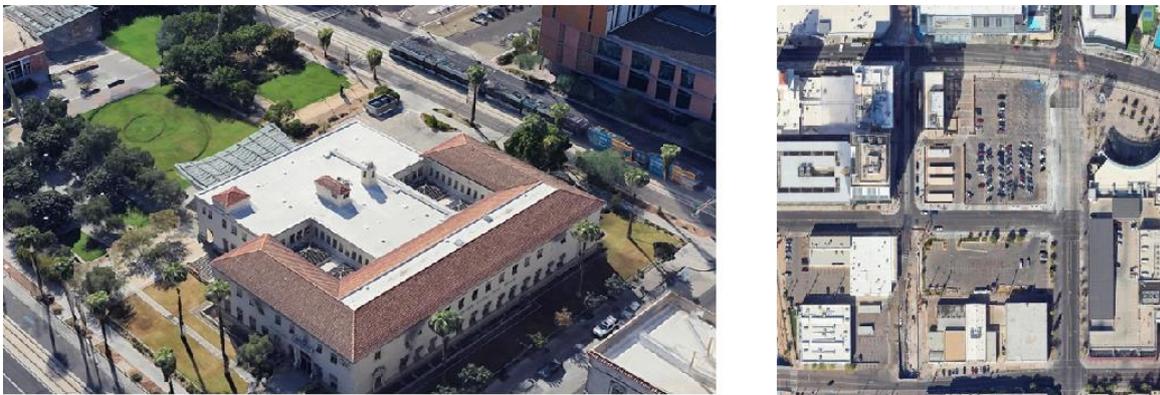


Рисунок 4 – Пример полученных снимков drone-view и satellite-view

Выводы

Несмотря на наличие API у ряда открытых источников, большинство из них (в частности, зарубежные) имеют ограниченный набор 3D-моделей для областей карт с разной степенью качества и детализации. Получение изображений при помощи API Cesium требует больше времени загрузки снимков из веб-сервиса или выгрузку в облачное хранилище Google. Разработанный метод автоматизированного сбора сохраняет снимки в организованную определенным образом структуру папок и скорость сохранения не зависит от интернет-соединения.

Снимки, формируемые при помощи платформы Google Earth, обладают достаточным качеством и возможностью приближения до низких рабочих высот БПЛА, что делает данный ресурс одним из лучших для сбора изображений. Текстуры, накладываемые на карту в 3D-режиме, обладают хорошим качеством, что является ключевым аспектом для задачи создания набора аэрофотоснимков.

К преимуществам метода также можно отнести возможность формирования конфигураций работы в виде отдельных файлов с наборами ключевых значений, таких как разрешение рабочей области, размеры карты, число ячеек, координаты стартовой точки.

К недостаткам предложенного метода можно отнести следующее: зависимость работы алгоритма от стабильности соединения, необходимость в дополнительных расчетах зоны захвата скриншота, дополнительная нагрузка на графический процес-

сор и оперативную память вычислительного устройства. Дополнительной проблемой является невозможность получения некоторых значений напрямую, в частности из-за недоступности путей к элементам внутри структуры веб-страницы. Интерактивные элементы веб-страницы требуют точного понимания алгоритма манипуляции ими для задания траектории полёта. В связи с этим, может потребоваться предварительная настройка траекторий перед автоматизированным сбором.

Список литературы

1. Ронжин, А.Л. Оптимизация технологической карты допустимых системотехнических решений задачи видеоаналитики аквакультуры [Текст] / А.Л. Ронжин, В.Н., Ле, Н.С. Шувалов // Вестник Южно-Уральского университета. Серия «Математика. Механика. Физика» . – 2024. – № 2(16). – С. 50-58 – ISSN 2075-809X. – DOI 10.14529/mmp4240205
2. Ронжин, А. Л. Интеллектуализация и роботизация научного оборудования для междисциплинарных исследований [Текст] / А.Л. Ронжин // Проблемы искусственного интеллекта. – 2024. – № 1(28). – С. 4-10 – ISSN 2413-7383. – DOI ?????
3. Верхотуров, А. Л. Использование радиолокационных данных для мониторинга состояния посевов сельскохозяйственных культур на юге Дальнего Востока России [Текст] / А. Л. Верхотуров, А. С. Степанов, Л. В. Илларионова // Информатика и Автоматизация . – 2024. – № 4(23). – С. 1221-1245 – ISSN 2713-3192. – DOI 10.15622/ia.23.4.11
4. Научно-обоснованный прогноз развития точного земледелия в России / Е.В. Рудой, М.С. Петухова, С.В. Рюмкин, Е.В. Труфляк, Н.Ю. Курченко; Новосиб. гос. аграр. ун-т, Кубан. гос. аграр. ун-т им. И.Т. Трубилина – Новосибирск: ИЦ НГАУ «Золотой колос», 2021 – 138 с
5. Google Планета Земля. [Электронный ресурс]. Режим доступа: <https://earth.google.co.zm/>. (Дата обращения: 20.10.2024)
6. Google Earth Studio. [Электронный ресурс]. Режим доступа: <https://earth.google.com/studio/docs/> (Дата обращения: 20.10.2024)
7. Али Б. Алгоритмы навигации беспилотных летательных аппаратов с использованием систем технического зрения [Текст] / Б. Али., Р.Н. Садеков, В.В. Цодокова //Гироскопия и навигация. – 2022. – Т. 30. – №. 4 (119). – С. 87). – ISSN 0869-7035. – DOI 10.17285/0869-7035.00105.
8. Зуев, В. М. Сравнение обнаружения объектов средствами искусственного интеллекта в сравнении с классическими методами [Текст] / Зуев В. М. // Проблемы искусственного интеллекта. – 2024. – № 3(34). – С. 4-10 – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-30-35.
9. Пикалёв, Я. С. О нейронных архитектурах извлечения признаков для задачи распознавания объектов на устройствах с ограниченной вычислительной мощностью [Текст] / Я.С. Пикалёв, Т.В. Ермоленко // Проблемы искусственного интеллекта.. – 2023. – № 3(30). – С. 44-54 – ISSN 2413-7383. – DOI 10.34757/2413-7383.2023.30.3.004
10. Павленко, Б. В. Интеллектуально-алгоритмический метод калибровки прицелов [Текст] / Б. В. Павленко, В. И. Бондаренко // Проблемы искусственного интеллекта. – 2024. – № 3 (34). – С. 55–63. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-55-63
11. Фролов, В. В. Анализ характеристик видеокамер оптического диапазона для работы в составе интеллектуальных систем обнаружения и оценки объектов беспилотными авиационными системами [Текст] / В. В. Фролов // Искусственный интеллект: теоретические аспекты, практическое применение : материалы Донецкого международного научного круглого стола. – Донецк : ФГБНУ «ИПИИ», 2024. – 328 с. – С. 304–310.
12. Уланов, С. И. Методы использования беспилотных летательных аппаратов в полете [Текст] / С. И. Уланов, О. А. Криводубский, А. А. Никитина // Проблемы искусственного интеллекта. – 2024. – № 2 (33). – С. 29–35. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-2-29-35.
13. Durgam A. et al. Cross-view geo-localization: a survey //arXiv preprint arXiv:2406.09722. – 2024.
14. Фаворская, М.Н. Восстановление аэрофотоснимков сверхвысокого разрешения с учетом семантических особенностей [Текст] / М.Н Фаворская, А.И. Пахирка // Информатика и Автоматизация . – 2024. – № 4(23). – С. 1047-1076.– ISSN 2713-3192. – DOI 10.15622/ia.23.4.5
15. Zhedong Zheng Yunchao Wei Yi Yang*. 2020. University-1652: A Multi-view Multi-source Benchmark for Drone-based Geo-localization. In Proceedings of the 28th ACM International Conference on Multimedia (MM '20), October 12–16, 2020, Seattle, WA, USA. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3394171.3413896>

16. Chu M. et al. Towards natural language-guided drones: GeoText-1652 benchmark with spatial relation matching //European Conference on Computer Vision. – Springer, Cham, 2025. – С. 213-231
17. Пикалёв, Я.С. Разработка системы нормализации текстовых корпусов [Текст] / Я.С. Пикалёв // Проблемы искусственного интеллекта. – 2022. – № 2(25). – С. 64-78. – ISSN 2413-7383.
18. Хакимов, Р. С. К вопросу о разработке системы аннотирования данных для задач компьютерного зрения [Текст] / Р. С. Хакимов, О. Л. Нижникова, М. В. Близно // Проблемы искусственного интеллекта. – 2024. – № 3 (34). – С. 70–79. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-70-79
19. Кришнан, Ш. Р.. Улучшение обнаружения аномалий на видео с помощью усовершенствованной технологии UNET и техники каскадного скользящего окна. [Текст] / Ш. Р. Кришнан, П. Амудха // Информатика и автоматизация. – 2024. – № 6 (23). – С. 1899-1930. – ISSN 2713-3192. –DOI 10.15622/ia.23.6.12
20. Shu X, Cheng X, Xu S, Chen Y, Ma T, Zhang W. How to construct low-altitude aerial image datasets for deep learning. Math Biosci Eng. 2021 Jan 5;18(2):986-999. doi: 10.3934/mbe.2021053. PMID: 33757171

References

1. Ronzhin, A.L. Optimizaciya tehnologicheskoy karty dopustimyh sistemotekhnicheskikh reshenij zadachi videoanalitiki akvakultury [Text] / A.L. Ronzhin, V.N., Le, N.S. Shuvalov // «Bulletin of the South Ural State University». Ser. «Mathematics. Mechanics. Physics». – 2024. – № 2(16). – P. 50-58 – ISSN 2075-809X. – DOI 10.14529/mmph240205
2. Ronzhin, A. L. Intellektualizaciya i robotizaciya nauchnogo oborudovaniya dlya mezhdisciplinarnykh issledovanij [Text] / A.L. Ronzhin // Problems of Artificial Intelligence. – 2024. – № 1(28). – P. 4-10 – ISSN 2413-7383. – DOI ?????.
3. Verhoturov A.L. Ispolzovanie radiolokacionnykh dannykh dlya monitoringa sostoyaniya posevov sel'skohozyajstvennykh kultur na yuge Dalnego Vostoka Rossii [Текст] / A.L.Verhoturov, A.S. Stepanov, L.V. Illarionova // Informatics and Automation . – 2024. – № 4(23). – P. 1221-1245 – ISSN 2713-3192. – DOI 10.15622/ia.23.4.11
4. Scientifically Based Forecast for the Development of Precision Agriculture in Russia / E.V. Rudoy, M.S. Petukhova, S.V. Ryumkin, E.V. Truflyak, N.Yu. Kurchenko; Novosibirsk State Agrarian University, Kuban State Agrarian University named after I.T. Trubilin – Novosibirsk: NGAU Publishing Center "Golden Ear", 2021 – 138 pages.
5. Google Earth. [Electronic resource]. Access mode: <https://earth.google.co.zm/> (Accessed: 20.10.2024).
6. Google Earth Studio. [Electronic resource]. Access mode: <https://earth.google.com/studio/docs/> (Accessed: 20.10.2024).
7. Ali B. Algoritmy navigacii bespilotnykh letatelnykh apparatov s ispolzovaniem sistem tehničeskogo zreniya [Text] / B. Ali., R.N. Sadekov, V.V. Tsodokova // Gyroscopy and navigation. – 2022. – Vol. 30. – №. 4 (119). – P. 87. – ISSN 0869-7035. – DOI 10.17285/0869-7035.00105
8. Zuev V.M. Sravnenie obnaruzheniya obektov sredstvami iskusstvennogo intellekta v sravnenii s klassičeskimi metodami [Text] / Zuev V.M. // Problems of Artificial Intelligence. – 2024. – № 3(34). – P. 4-10 – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-30-35.
9. Pikalyov, Ya.S. O nejronnykh arhitekturah izvlecheniya priznakov dlya zadachi raspoznavaniya obektov na ustrojstvax s ogranichennoj vychislitelnoj moshnostyu [Text] / Ya.S. Pikalyov, T.V. Yermolenko // Problems of Artificial Intelligence. – 2023. – № 3(30). – P. 44-54 – ISSN 2413-7383. – DOI 10.34757/2413-7383.2023.30.3.004
10. Pavlenko, B.V. Intellektualno-algoritmicheskij metod kalibrovki pricelov [Text] / B.V.Pavlenko, V.I. Bondarenko // Problems of Artificial Intelligence. – 2024. – № 3 (34). – P. 55–63. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-55-63
11. Frolov, V. V. Analysis of the Characteristics of Optical Range Video Cameras for Use in Intelligent Object Detection and Evaluation Systems in Unmanned Aerial Systems [Text] / V.V. Frolov // Artificial Intelligence: Theoretical Aspects, Practical Applications: Materials of the Donetsk International Scientific Round Table. – Donetsk: FSBSI "IPII", 2024. – 328 pages. – P. 304–310.
12. Ulanov, S. I. Methods of Using Unmanned Aerial Vehicles in Flight [Text] / S. I. Ulanov, O. A. Krivodubsky, A. A. Nikitina // Problems of Artificial Intelligence. – 2024. – No. 2 (33). – P. 29–35. – ISSN 2413-7383. – DOI: 10.24412/2413-7383-2024-2-29-35.
13. Durgam A. et al. Cross-view geo-localization: a survey //arXiv preprint arXiv:2406.09722. – 2024.

14. Favorskaya, M.N. Vosstanovlenie aerofotosnimkov sverhvyssokogo razresheniya s uchetom semanticheskikh osobennostej [Text] / M.N. Favorskaya, A.I. Pahirka // Informatics and Automation. – 2024. – № 4(23). – P. 1047-1076. – ISSN 2713-3192. – DOI 10.15622/ia.23.4.5
15. Zhedong Zheng Yunchao Wei Yi Yang* . 2020. University-1652: A Multi-view Multi-source Benchmark for Drone-based Geo-localization. In Proceedings of the 28th ACM International Conference on Multimedia (MM '20), October 12–16, 2020, Seattle, WA, USA. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3394171.3413896>
16. Chu M. et al. Towards natural language-guided drones: GeoText-1652 benchmark with spatial relation matching //European Conference on Computer Vision. – Springer, Cham, 2025. – P. 213-231
17. Pikalyov, Ya.S. Razrabotka sistemy normalizacii tekstovyh korpusov [Text] / Ya.S. Pikalyov // Problems of Artificial Intelligence. – 2022. – № 2(25). – P. 64-78. – ISSN 2413-7383. – DOI ?????.
18. Khakimov R.S. K voprosu o razrabotke sistemy annotirovaniya dannyh dlya zadach kompyuternogo zreniya [Text] / R.S. Khakimov, O.L. Nizhnikova, M.V. Blizno // Problems of Artificial Intelligence. – 2024. – № 3 (34). – P. 70–79. – ISSN 2075-1087. – DOI 10.24412/2413-7383-2024-3-70-79
19. Krishnan, Sh. R. Uluchshenie obnaruzheniya anomalij na video s pomoshyu usovershenstvovannoj tehnologii UNET i tehniki kaskadnogo skolzyashego okna [Text] / Sh. R. Krishann, P. Amudha //Informatics and automatization. – 2024. – № 6 (23). – P. 1899-1930. – ISSN 2713-3192. –DOI 10.15622/ia.23.6.12
20. Shu X, Cheng X, Xu S, Chen Y, Ma T, Zhang W. How to construct low-altitude aerial image datasets for deep learning. Math Biosci Eng. 2021 Jan 5;18(2):986-999. doi: 10.3934/mbe.2021053. PMID: 33757171

RESUME

B. V. Pavlenko, Ya. S. Pikalyov

Methodology for Creating an Aerial Image Dataset for Cross-View Geo-Localization

This work presents a method for automating the collection of aerial images for the cross-view geo-localization (CVGL) task using third-party cartographic services such as Google Earth and Google Earth Studio. The challenge of creating datasets for CVGL lies in the high costs associated with UAV flights and the limited access to real aerial imagery. The developed method enables the creation of an extensive dataset, including satellite-view images and simulated drone-view images, with minimal resource expenditure.

A step-by-step algorithm is proposed, involving dividing the target map area into a grid of cells and collecting images for each cell. For drone-view images, 3D modeling is employed to ensure diverse viewing angles. This approach guarantees varied representations of the terrain, essential for model training. A comparative analysis of existing datasets confirmed the importance of developing customized solutions for specific tasks.

The main advantages of the method are highlighted: resource efficiency, high-quality imagery, and flexibility in configuring data collection parameters. However, the method has certain limitations, including dependence on stable internet connectivity, high computational load, and challenges in integrating interactive web interface elements. This study demonstrates the potential of automated data collection methods and their applicability in computer vision and data analysis tasks.

РЕЗЮМЕ

Б. В. Павленко, Я. С. Пикалёв

Методика создания набора аэрофотоснимков для задачи перекрёстной геолокализации

Системы компьютерного зрения, используемые в беспилотных летательных аппаратах, играют ключевую роль в широком спектре задач: классификации местности, мониторинг инфраструктуры, обнаружение чрезвычайных ситуаций, транспорта, других объектов. Тем не менее, для их использования в системах компьютерного зрения в беспилотных летательных аппаратах необходимы наборы аэрофотоснимков.

Но высокая стоимость и ограничения на съёмку с воздуха, в определенных условиях, представляют проблему для сбора данных напрямую.

В работе разработан метод автоматизации сбора аэрофотоснимков для задачи перекрёстной геолокализации (CVGL) с использованием сторонних картографических сервисов, таких как Google Earth и Google Earth Studio. Проблема создания наборов данных для CVGL связана с высокими затратами на полёты БПЛА и ограничениями в доступе к реальным снимкам. Разработанный метод позволяет формировать обширный набор данных, включающий изображения со спутника (satellite-view) и имитацию видов с дрона (drone-view), с минимальными затратами.

Предложен пошаговый алгоритм, который включает разбиение целевой карты на сетку из ячеек и сбор изображений для каждой из них. Для снимков drone-view используется 3D-моделирование, обеспечивающее разнообразие точек обзора. Такой подход гарантирует разноплановое представление местности, необходимое для обучения моделей. Сравнительный анализ существующих наборов данных подтвердил актуальность создания собственных решений для специфичных задач.

Выделены основные преимущества метода: экономия ресурсов, высокое качество получаемых снимков и гибкость в настройке конфигураций сбора данных. К недостаткам отнесены зависимость от стабильности интернет-соединения, нагрузка на вычислительное устройство и сложность интеграции интерактивных элементов веб-интерфейсов. Работа демонстрирует перспективы развития автоматизированных методов сбора данных и их применимость в задачах компьютерного зрения и анализа данных.

Павленко Богдан Викторович – младший научный сотрудник лаборатории интеллектуальных систем и анализа данных, Федерального государственного бюджетного научного учреждения «Институт проблем искусственного интеллекта», г. Донецк. *Область научных интересов:* компьютерное зрение, машинное обучение, нейронные сети, эл. почта bogdanpav12000@mail.ru, адрес: 283048, г. Донецк, ул. Артема, д. 118 б, телефон +79494386450.

Пикалёв Ярослав Сергеевич – кандидат техн. наук, старший научный сотрудник лаборатории интеллектуальных систем и анализа данных, Федерального государственного бюджетного научного учреждения «Институт проблем искусственного интеллекта», г. Донецк. *Область научных интересов:* Цифровая обработка сигналов, анализ данных, распознавание образов, обработка естественного языка, компьютерное зрение, машинное обучение, нейронные сети, эл. почта i@pikaliov.ru, адрес: 283085, ДНР, г. Донецк, ул. Отважных, д. 19, кв.85, телефон: +7949 4287388.

Статья поступила в редакцию 03.06.2024.