

17. Luchinin A.G. Ob osnovnykh printsipakh formirovaniya izobrazheniya podvodnykh ob"ektov pri nablyudenii cherez vzvolnovannuyu poverkhnost' [On the basic principles of forming an image of underwater objects when observed through a rough surface], *AN SSSR. Fizika atmosfery i okeana* [Izvestiya AN SSSR. Physics of the atmosphere and ocean], 1996, Vol. 32, No. 2, pp. 296-302.
18. Luchinin A.G. Funktsii raspredeleniya signala v izobrazheniyakh samosvetyashchikhsya podvodnykh test-ob"ektov, nablyudaemykh cherez vzvolnovannuyu poverkhnost' [Signal distribution functions in images of self-luminous underwater test objects observed through a rough surface], *AN SSSR. Fizika atmosfery i okeana* [Izvestiya AN SSSR. Physics of the atmosphere and ocean], 2002, Vol. 38, No. 3, pp. 419-425.
19. Luchinin A.G. Veroyatnostnaya traktovka rasshchepleniya izobrazheniy podvodnykh ob"ektov [Probabilistic interpretation of splitting of images of underwater objects]. Nizhniy Novgorod: In-t prikladnoy fiziki RAN, 2002.
20. Dolin L.S., Levin I.M. Spravochnik po teorii podvodnogo videniya [Handbook of the theory of underwater vision]. Leningrad: Gidrometeoizdat, 1991, 229 p.
21. Optika okeana. T. 2 Prikladnaya optika okeana [Ocean Optics. Vol. 2. Applied Ocean Optics], pod red. A.S. Monina. Moscow: Nauka, 1983, 236 p.

Грузевич Юрий Кириллович – Открытое акционерное общество «НПО ГЕОФИЗИКА-НВ»; e-mail: yukg@mail.ru; г. Москва, Россия; тел.: +74957635492; к.т.н.; доцент; зам. генерального директора по науке.

Гордиенко Юрий Николаевич – Открытое акционерное общество «НПО ГЕОФИЗИКА-НВ»; e-mail: yuriy194421@mail.ru; г. Москва, Россия; тел.: +74957635492; зам. технического директора.

Альков Павел Сергеевич – Открытое акционерное общество «НПО ГЕОФИЗИКА-НВ»; e-mail: pavel_alkov@mail.ru; г. Москва, Россия; тел.: +79160586015; к.т.н.; доцент; зам. технического директора – начальник отдела 047.

Волков Дмитрий Владимирович – Открытое акционерное общество «НПО ГЕОФИЗИКА-НВ»; e-mail: mitya_volkov_7554@bk.ru; г. Москва, Россия; тел.: +79859521503; ведущий инженер.

Ходаковская Мария Сергеевна – Открытое акционерное общество «НПО ГЕОФИЗИКА-НВ»; e-mail: mariya.xabarova.96@mail.ru; г. Москва, Россия; тел.: +79636095192; инженер.

Gruzevich Yuri Kirillovich – Open Joint-Stock Company “NPO GEOFIZIKA-NV”; e-mail: yukg@mail.ru; Moscow, Russia; phone: +74957635492; cand. of eng. sc.; associate professor; deputy general director for science.

Gordienko Yuri Nikolaevich – Open Joint-Stock Company “NPO GEOFIZIKA-NV”; e-mail: yuriy194421@mail.ru; Moscow, Russia; phone: +74957635492; deputy technical director.

Alkov Pavel Sergeevich – Open Joint-Stock Company “NPO GEOFIZIKA-NV”; e-mail: pavel_alkov@mail.ru; Moscow, Russia; phone: +79160586015; cand. of eng. sc.; associate professor; deputy technical director – head of department 047.

Volkov Dmitry Vladimirovich – Open Joint-Stock Company “NPO GEOFIZIKA-NV”; e-mail: mitya_volkov_7554@bk.ru; Moscow, Russia; phone: +79859521503; lead engineer.

Khodakovskaya Maria Sergeevna – Open Joint-Stock Company “NPO GEOFIZIKA-NV”; e-mail: mariya.xabarova.96@mail.ru; Moscow, Russia; phone: +79636095192; engineer.

УДК 007.52

DOI 10.18522/2311-3103-2025-2-256-268

Н.В. Ким, Н.Е. Бодунков, Д.С. Гиренко, Н.А. Ляпин, Н.В. Удалова

ВИЗУАЛЬНАЯ НАВИГАЦИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ С ИСПОЛЬЗОВАНИЕМ СЕМАНТИЧЕСКИХ ОПИСАНИЙ МЕСТНОСТИ

Рассматривается задача визуальной навигации беспилотных летательных аппаратов (БЛА), которая заключается в автоматическом определении текущего положения БЛА (координат в земной (местной) системе координат), на основе сравнения и идентификации описаний принимаемых на борту текущих изображений (ТИ) подстилающей поверхности и эталонных описа-

ний, хранящихся в виде цифровой карты местности в памяти бортового вычислителя БЛА. Целью работы является повышение эффективности применения методов визуальной навигации, с точки зрения повышения вычислительной производительности, устойчивости и точности алгоритмов идентификации изображений в сложных и изменяемых условиях наблюдения, за счет использования семантических описаний наблюдаемых сцен. В данной работе под семантическими описаниями будем понимать описания, включающие классы наблюдаемых на сцене объектов, атрибуты этих объектов и отношениями между ними. Подготовка семантических описаний карты производится на этапе предполетной подготовки БЛА с помощью предварительно обученных нейронных сетей семантической сегментации. Семантические описания принимаемых ТИ формируются на борту БЛА. При этом использование нейросетевых алгоритмов позволяет реализовать этот процесс в реальном времени для широкого спектра условий наблюдений (различного времени суток и года). Использование семантических описаний карты и ТИ позволяет сократить вычисления по сравнению с традиционным попиксельным сопоставлением растровых изображений. Семантические описания сравниваются путем сопоставления классов объектов, их атрибутов и отношений. В работе приведены: общий алгоритм визуальной навигации, основные этапы методики формирования семантических описаний и алгоритм сравнения и идентификации семантических описаний ТИ и описаний карты. Предложен иерархический алгоритм сравнения и идентификации изображений, основанный на последовательном применении семантических, и растровых описаний наблюдаемых сцен. Показано, что использование процедуры сравнения семантических описаний ТИ и карты по присутствующим классам объектов позволяет существенно сократить вычисления, необходимые для идентификации изображений.

Навигационные системы; визуальная навигация; семантическая навигация; системы технического зрения.

N.V. Kim, N.E. Bodunkov, D.S. Girenko, N.A. Lyapin, N.V. Udalova

VISUAL NAVIGATION OF UNMANNED AERIAL VEHICLES USING SEMANTIC TERRAIN DESCRIPTIONS

The article addresses the problem of visual navigation for unmanned aerial vehicles (UAVs), which involves the automatic determination of the current position of the UAV (coordinates in the ground (local) coordinate system) based on the comparison and identification of descriptions of the current images (CI) received on board with reference descriptions stored in the form of a digital map in the memory of the UAV's onboard computer. The aim of this work is to improve the efficiency of visual navigation methods in terms of increasing computational performance, robustness, and accuracy of image identification algorithms in complex and changing observation conditions by using semantic descriptions of observed scenes. In this work, semantic descriptions are understood as descriptions that include classes of objects observed in the scene, their attributes, and relationships between them. The preparation of semantic descriptions of the map is carried out at the pre-flight preparation stage of the UAV using pre-trained neural networks for semantic segmentation. Semantic descriptions of the received CIs are generated on board the UAV. The use of neural network algorithms allows this process to be implemented in real-time for a wide range of observation conditions (different times of day and year). The use of semantic descriptions of the map and CI reduces computations compared to traditional pixel-by-pixel matching of raster images. Semantic descriptions are compared by matching object classes, their attributes, and relationships. The work presents a general algorithm for visual navigation, the main stages of the methodology for forming semantic descriptions, and the algorithm for comparing and identifying semantic descriptions of CIs and map descriptions. A hierarchical algorithm for comparing and identifying images based on the sequential application of semantic and raster descriptions of observed scenes is proposed. It is shown that the use of the procedure for comparing semantic descriptions of CIs and maps by the classes of objects present significantly reduces the computations necessary for image identification.

Navigation systems; visual navigation; semantic navigation; machine vision systems.

Введение. Методы визуальной навигации могут эффективно использоваться в различных полетных ситуациях, в частности, при неустойчивой работе спутниковой навигационной системы [1, 2]. Внедрение визуальной навигации расширяет возможности автономного полёта беспилотных летательных аппаратов (БЛА), что позволяет значительно повысить эффективность выполнения решаемых задач, снизить затраты на управление и обслуживание аппаратов, а также минимизировать риски, связанные с ошибками оператора.

Целью настоящей работы является повышение эффективности решения, с точки зрения повышения вычислительной производительности, устойчивости и точности, задачи визуальной навигации в сложных и изменяемых условиях наблюдения подстилающей поверхности.

В рамках данного исследования будем считать, что задача визуальной навигации заключается в определении текущего положения беспилотного летательного аппарата (БЛА), т.е. координат БЛА, с заданной точностью, в земной (местной) системе координат XOY. Высота полета определяется с помощью бортового высотомера.

Будем считать, что полет беспилотного летательного аппарата проходит в области отсутствия сигналов глобальной навигационной спутниковой системы (ГНСС). Коррекция бортовой инерциальной навигационной системы (ИНС) БЛА производится с помощью методов визуальной навигации [1], позволяющих определять координаты БЛА на основе сопоставления и идентификации принимаемых на борту текущих изображений (ТИ) подстилающей поверхности и эталонных изображений (ЭИ), хранящихся в виде цифровой карты местности (ЦКМ) в вычислительной системе БЛА [3, 4].

К основным проблемам реализации методов визуальной навигации БЛА, связанным с сопоставлением ТИ и ЭИ, относятся изменяемые условия наблюдения, в частности изменения освещенности, погодных условий, искажающие принимаемые ТИ относительно ЭИ.

Кроме того, реализацию методов визуальной навигации затрудняет высокая трудоемкость вычислений, связанных с необходимостью обработки объемных растровых изображений подстилающей поверхности, выделением, оценкой и сопоставлением необходимых признаков [5, 6].

В последние годы в различных задачах обнаружения и распознавания начинают использоваться нейросетевые алгоритмы. Формирование рабочего признакового пространства и выделение признаков в них реализуется автоматически в процессе обучения. Данный факт позволяет перейти на новые методы описания наблюдаемых сцен и повысить производительность обработки текущих изображений.

В качестве одного из вариантов решения этих проблем предлагается в отдельных процессах сопоставления ТИ и ЭИ перейти от использования растровых изображений наблюдаемых сцен к их семантическим описаниям. В данной работе под семантическими описаниями будем понимать описания наблюдаемых сцен присутствующими на них классами объектов, их атрибутами и отношениями между ними. Классификация объектов производится с помощью нейронных сетей семантической сегментации [9, 14].

Предполагается, что подобный переход существенно повысит устойчивость алгоритмов визуальной навигации к изменению условий наблюдения и вычислительную производительность вычислений.

Визуальная навигация БЛА. При использовании методов визуальной навигации оценка положения БЛА производится на основе сопоставления изображений: принимаемого на борту ТИ и подготовленного заранее ЭИ. Вероятность получения конкретного ТИ может оцениваться на основе времени полета БЛА без коррекции и расчета накопленных ошибок БИНС [7, 8].

В общем случае изображения содержат наблюдаемые сцены. В частных случаях на наблюдаемых сценах могут присутствовать неинформативные (с т.з. визуальной навигации) или информативные изображения объектов, позволяющие однозначно определять положение БЛА.

Задача привязки изображений (ТИ и ЭИ) относится к известным задачам поиска и обнаружения объектов [6], которыми являются фрагменты ЭИ, соответствующие ТИ.

В общем случае для поиска искомого участка местности или ориентира может использоваться не только исходное принятое растровое изображение, а некоторое описание, позволяющее идентифицировать этот участок. Растровые изображения являются частным случаем эталонного и текущего описаний и используются в корреляционных алгоритмах.

Другим вариантом описаний является, например, признаковое описание, в частности, на основе семантических описаний, ключевых точек и др.

Очевидно, что для сопоставления эталонного и текущего описаний подстилающей поверхности (местности) они должны быть представлены в едином формате.

Общий алгоритм оценки координат БЛА с использованием визуальной навигации может быть представлена в следующем виде (рис. 1).

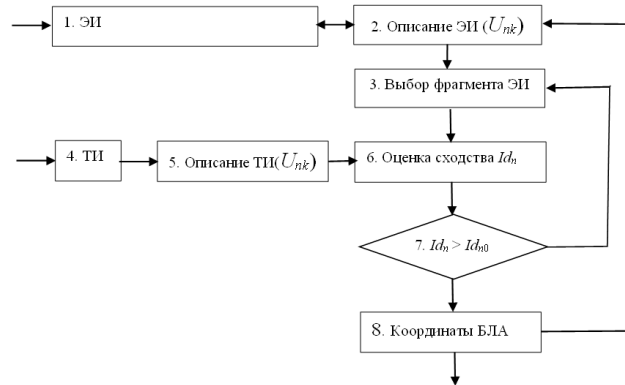


Рис. 1. Алгоритм оценки координат БЛА

На первом этапе (бл.1) определяется область местности, в которой с вероятностью отличной от нуля может находиться БЛА. Необходимо отметить, что на этапе предполетной подготовки ЦКМ может быть преобразовано в признаковое (в частности, семантическое) описание.

В бл.2 формируется описание ЭИ с учетом выбранного пространства признаков (распределения яркости, характерных черт и т.д.) U_{nk} , где n – индекс признака, k – индекс значения n -го признака.

Например, если выбрано распределение яркости изображения, то сходство ТИ и ЭИ может оцениваться корреляционной функцией [3, 4]. В этом случае в качестве сравниваемых описаний ЭИ и ТИ используются растровые изображения.

Если в качестве характерных признаков используются ключевые точки [1], то в бл.2 и бл.5 необходимо реализовать процедуры поиска этих точек.

Из принятого ТИ (бл.4) формируется описание (бл.5), базирующееся на соответствующих признаках.

В бл.3 из ЭИ выделяется фрагмент, который должен сравниваться с принятым текущим изображением. Выбор фрагмента ЭИ производится на основе выбранного метода поиска.

В бл.6 производится оценка сходства ТИ и ЭИ путем расчета некоторого выбранного показателя Id_n , где n – индекс показателя, зависящего от выбранного описания изображений. В частности, такими показателями могут быть корреляционная функция, парные функции [Ким], расстояние Махаланобиса или Хэмминга и др.

Пусть увеличение Id_n соответствует увеличению сходства сравниваемых описаний, тогда выполнение условия (бл.7)

$$Id_n \geq Id_{n0} \quad (1)$$

определяет идентичность описаний ТИ и ЭИ, где Id_{n0} – заданный порог.

При $Id_n < Id_{n0}$ (бл.7), необходимо перейти к другому фрагменту ЭИ (бл.2).

Если найден фрагмент ЭИ, идентичный принятому ТИ (1), то расчет поиска заканчивается (бл.18).

Координаты БЛА определяются, исходя из известных: положения системы наблюдения (камеры) на борту БЛА, параметров камеры, направления линии визирования, высоты полета и координат найденного фрагмента ЭИ.

Наиболее распространенными методами, реализуемыми в визуальной навигации приведенным алгоритмом (рис.1), являются корреляционно-экстремальные методы [2–5].

Преимущества данного подхода связаны с простотой реализации и высокими качествами пространственной фильтрации, свойственной обработке растровых изображений. При этом нет необходимости выделять рабочие признаки объектов, что повышает сложность алгоритма.

К недостаткам корреляционных алгоритмов можно отнести высокую чувствительности к изменению условий наблюдения и относительно низкую вычислительную производительность.

Семантические описания наблюдаемых сцен. Для повышения устойчивости и вычислительной производительности алгоритма поиска ТИ на ЭИ предлагается использовать вариант семантического описания наблюдаемых сцен.

В рамках данного описания признаками, позволяющими сравнивать схожесть и идентифицировать наблюдаемые сцены, являются классы объектов, их атрибуты, отношения между ними.

Классами объектов могут быть, например, здания, дороги, лес водная поверхность и пр.

Атрибутами объектов являются форма, размеры, текстура и пр.

В качестве *отношений между объектами* могут использоваться относительные положения объектов.

С развитием вычислительной техники и алгоритмов глубокого обучения, архитектуры сверточных искусственных нейронных сетей (НС) научились решать задачи обнаружения и распознавания искомым объектов на различных наблюдаемых сценах в режиме реального времени [9–11]. Данный факт позволяет прогнозировать существенное повышение производительности подобных алгоритмов поиска по сравнению с алгоритмами, построенными на обработке и анализе растровых изображений [18–20].

Также преимущество нейросетевых решений состоит в том, что за счет достаточно большой обучающей выборки и обучения НС на этапе предполетной подготовки можно добиться высокой устойчивости к изменению условий наблюдения, включая изменения освещенности, ракурса и масштаба наблюдаемых сцен.

Существенным недостатком предлагаемого подхода является необходимость подготовки и использования больших обучающих выборок (несколько десятков тысяч изображений на один класс объектов).

Базовой процедурой формирования семантических описаний наблюдаемых сцен является семантическое сегментирование изображения. Исходя из наблюдений последних лет, в настоящее время, несмотря на увеличивающийся прогресс в использовании сложных дескрипторов изображений [8] и более совершенных методов машинного обучения, сегментация по-прежнему остается сложной задачей.

Семантическая сегментация изображений – это процесс разделения изображения на области, где каждая область соответствует семантическому классу из заранее определенного списка [9] (алфавиту классов). Внешний вид этих классов изучается с помощью размеченных изображений. Методы обычно объединяют две подзадачи: часть распознавания, отвечающая за маркировку, и часть регуляризации, которая заставляет соседние пиксели принадлежать к одному классу. Часть распознавания основана на локальном внешнем виде, который рассматривается на уровне пикселя или на уровне участка. Для описания локального внешнего вида используются различные признаки, среди которых текстура (банки фильтров) и цветовая статистика.

Основным инструментом решения данного класса задач являются сверточные нейронные сети (СНС, CNN) [10]. При этом, в настоящее время существует огромное разнообразие таких архитектур, позволяющих решить задачу семантической сегментации изображений [11]. Большинство успешных алгоритмов семантической сегментации, разработанных в предыдущем десятилетии, опирались в основном на обычные классификаторы, такие как boosting [12], случайные леса [13] или методы опорных векторов [14]. Однако, за последние несколько лет достижения в области глубокого обучения классификации изображений были адаптированы в задачу семантической сегментации.

Базовыми архитектурами для решения задачи сегментации изображений являются архитектуры FCN, Seg-net, U-net. Для них проведена сравнительная характеристика их работы по метрике IoU (Intersection over Union) или индекс Жаккара (Jaccard index) для одного и того же датасета с изображениями спутниковых снимков, представленного ниже.

$$IoU = \frac{TP}{TP+FP+FN}, \quad (2)$$

где TP – количество правильно проклассифицированных пикселей на изображении (true positives), FP – количество пикселей, которые метод проклассифицировал как принадлежащие определенному классу, но они таковыми не являются (false positives), FN – количество пикселей, которые принадлежат определенному классу, но метод проклассифицировал их как не принадлежащие этому классу (false negatives). Значения TP , FP , FN вычисляются на основе известной матрицы ошибок [15].

В табл. 1, представлены результаты работы сравниваемых сетей.

Таблица 1

Сравнительная характеристика основных архитектур СНС для решения задачи семантической сегментации

Модель	IoU , %
FCN	65.3
Seg-net	57
U-net	76.4

Потери нейронной сети оценивались на основе двух функций:

♦ **Diceloss** – метрика, которая показывает меру несходства между предсказанной и истинной сегментацией изображения.

$$DICE = \frac{2*TP}{2*TP+FP+FN}, \quad (3)$$

♦ **CategoricalFocalLoss** – метрика, вычисляющая альфа-сбалансированную фокальную потерю кроссэнтропии.

$$FL = (1 - pr)^\gamma * \log(pr), \quad (4)$$

где pr – предсказанный класс, γ – параметр фокусировки. Когда $\gamma = 0$, фокусировка не влияет на перекрестную энтропию.

Точность модели определялась на основе следующих метрик:

♦ **Accuracy** – точность модели, рассчитывается как количество правильно предсказанных классов, разделенное на общее количество объектов.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN}, \quad (5)$$

♦ **Jaccard coefficient** – коэффициент рассчитывается как количество общих элементов в двух классах (A и B), разделённое на общее количество элементов в любом из классов.

$$Jaccard_{coef}(A, B) = \frac{A \cap B}{A \cup B}. \quad (6)$$

Исходя из полученных выше результатов, в качестве исходного датасета был размечен фрагмент карты одного из районов Санкт-Петербурга, который был размечен семантической сегментацией по пикселям в 5 классах. Общий объем набора данных составляет 20735 трехканальных изображения размером 256 x 256 пикселей.

Алфавит классов и цвет разметки:

1. Здание (З), розовый.
2. Поверхность (П) (грунтовая территория), черный.
3. Дорога (Д), серый.
4. Растительность (Р), зеленый.
5. Вода (В), синий.

На рис. 2 приведен график изменения функции потерь по эпохам после обучения нейронной сети U-net:

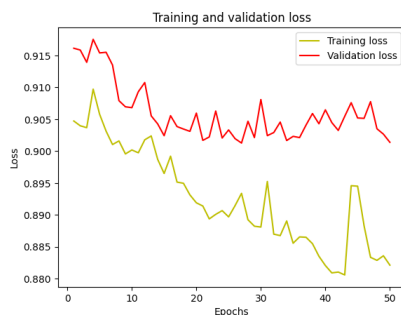


Рис. 2. График изменения функции потерь

После работы нейронной сети для каждого изображения можно определить матрицу вероятностей распознавания по каждому из классов. Для этого необходимо определить количество пикселей ($N_i, i = \overline{1..5}$), где i – номер класса, соответствующих каждому классу на истинной маске изображения. Далее необходимо определить на сегментированном изображении количество пикселей, которые соответствуют данному классу (N_{iTrue}) при по-пиксельном сравнении двух изображений (истинной и сегментированной маски) и определить вероятность $P(U_i \setminus U_i), i = \overline{1..5}$ правильного распознавания класса U_i .

$$P(U_i \setminus U_i) = \frac{N_{iTrue}}{N_i}, i = \overline{1..5}. \quad (7)$$

Затем необходимо определить условные вероятности ($P(U_i \setminus U_j), i = \overline{1..5}, j = \overline{1..5}, i \neq j$) распознавания i -го класса, при условии, что на сегментированном изображении в соответствующем пикселе был определен другой класс. Для этого, по каждому из оставшихся классов (за исключением искомого) необходимо определить количество неверно распознанных пикселей ($N_{ijFalse}$).

$$P(U_i \setminus U_j) = \frac{N_{ijFalse}}{N_i}, i = \overline{1..5}. \quad (8)$$

Пример изображения приведен на рис. 3.

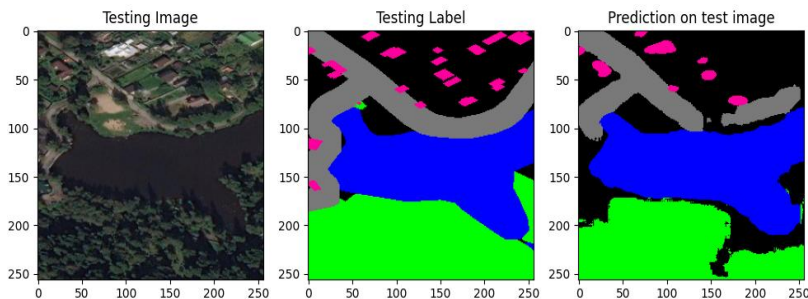


Рис. 3. Пример сегментированного изображения

Идентификация семантических описаний наблюдаемых сцен. Для примера будем считать, что ТИ имеет меньший размер, чем ЭИ.

В большинстве случаев, например, в корреляционно-экстремальных навигационных системах (КЭНС) для поиска положения ТИ (размерностью в пикселях $m*n$) на ЭИ ($M*N$) производится попиксельное обследование всего ЭИ (метод сплошного поиска). При этом количество реализуемых вычислительных операций для принятого ТИ и всего ЭИ равно

$$K\Sigma = mn(M - n + 1)(N - n + 1)K_o, \quad (9)$$

где K_o – количество элементарных вычислительных операций в одном цикле обработки изображений, в частности, для расчета выбранной критериальной корреляционной функции, но без расчета ее экстремума [4].

Для каждой сравниваемой пары изображений вычисляется показатель, значение которого определяет сходство изображений. Например, в КЭНС могут использоваться такие функции, как нормированный коэффициент корреляции или функция среднего модуля разности и пр. [4, 5]. Координаты ЭИ, при которых корреляционная функция достигает экстремума, определяют фрагмент, максимально похожий на ТИ.

Реализация на борту БЛА подобных вычислительных процессов оказывает существенную (в ряде случаев недопустимую) нагрузку на бортовые вычислители, поэтому повышение вычислительной эффективности (производительности) алгоритмов поиска и сравнения изображений является актуальной и практически значимой задачей.

Принципиальным отличием семантического описания наблюдаемых сцен от растрового является наличие в семантических описаниях выделенных классов объектов, присутствующих на сценах. Данные объекты, а также различные атрибуты объектов и отношения (в частности, геометрические) между объектами являются, в общем случае, признаками наблюдаемых сцен.

Использование семантических описаний позволяет реализовать метод выборочного сканирования ЭИ на основе использования некоторого выбранного рабочего словаря признаков для сравнения ЭИ и ТИ.

Например, в данном методе ТИ сравнивается только с фрагментами ЭИ, имеющими близкий состав по классам объектов наблюдаемой сцены. Остальные области ЭИ не рассматриваются.

Как было отмечено выше, важнейшим вопросом использования семантических описаний является выбор наиболее схожей пары и определением координат ТИ на ЭИ (идентификация ТИ).

Удобным подходом к оценке сходства изображений является использование парных функций [4, 5], ориентированных на обработку растровых изображений.

Реализация данных функций в растровых изображениях связана с вычислением количества элементов ТИ и ЭИ, совпадающих или несовпадающих по значению, т.е. значений яркости в соответствующих пикселях.

В частности, для бинарных растровых изображений предложен ряд парных критериевых функций, например, функции Рао, Джекарда, Юла и др. Предварительные исследования показали высокую эффективность функции Джекарда

$$K_D = \frac{a}{(a+b+c)}, \quad (10)$$

где a, e – количество совпадающих элементов с D значениями соответственно в ТИ и ЭИ (1,1) и (0,0); b, c – количество несовпадающих элементов со значениями соответственно (1,0) и (0,1).

Обозначим K_{PD} – количество элементарных вычислительных операций, необходимых для расчета функции Джекарда.

В случае использования семантических описаний авторами предлагается иной подход применения парных функций:

- ♦ наличие некоторого класса объектов (в общем случае признака) в описании обозначается «1», а отсутствие признака – «0»;
- ♦ сопоставление двух изображений проводится на основе расчета совпадающих или несовпадающих признаков на этих изображениях.

Так, « a » (10) может интерпретироваться, как присутствие на 2-х изображениях одинаковых признаков: классов объектов, отношений или атрибутов.

Подобный подход позволяет существенно сократить количество сравниваемых фрагментов ЭИ с ТИ. Однако выбранный словарь признаков может оказаться недостаточным для требуемой идентификации ТИ на ЭИ.

Обозначим S_0 начальную область поиска, соответствующую ЭИ, и S_i – области неопределенности положения ТИ после i -го этапа поиска.

В случае появления альтернативных вариантов решений, т.е. нахождения на ЭИ нескольких фрагментов, похожих на ТИ, производится расширение набора используемых признаков. Процесс идентификации должен повторяться до тех пор, пока вероятность принятия правильного решения не достигнет требуемого (заданного) значения.

На рис. 4 показан пример алгоритма иерархической идентификации ТИ на ЭИ.

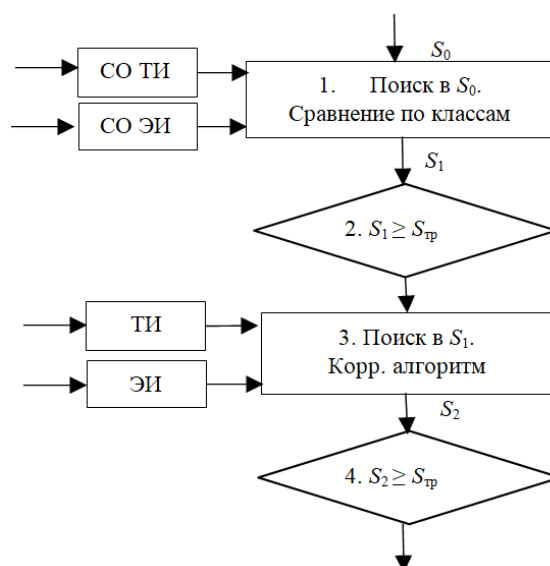


Рис. 4. Алгоритм иерархической идентификации ТИ

В бл.1 производится сравнение семантических описаний (СО ТИ и СО ЭИ), в результате которого начальная область поиска S_0 сокращается до области S_1 .

Если полученная точность идентификации, характеризующая S_1 , не хуже требуемой $S_{тр}$, то процесс заканчивается (бл.2).

В случае $S_1 > S_{тр}$ процесс продолжается и реализуется расширение словаря признаков. В частности, здесь реализован переход к растровым семантически сегментированным ТИ и ЭИ. Для их идентификации снова использована парная функция Джеккарда (10).

Данный иерархический подход является адаптивным к изменяемым условиям наблюдения наблюдаемых ТИ и ЭИ. При этом на каждом i -м цикле поиска (для определенных рабочих словарей признаков) область S_i ,

Преимуществами оценки сходства изображений на основе семантических описаний наблюдаемых сцен по сравнению с растровыми описаниями является возможность сопоставления признаков (классов объектов, отношений, атрибутов), без пространственной обработки и анализа растровой информации.

При этом, выделение признаков с помощью обученных нейросетей позволяет существенно сократить время обработки и анализа видеоинформации.

На рис. 5,а представлен пример наблюдаемой сцены, принятой в качестве ЭИ, и ее семантическое описание. ТИ и семантическое описание ТИ показаны на рис. 5,б.

Сформированное семантическое описание позволяет представлять наблюдаемые сцены в виде набора классов объектов, присутствующих в этой сцене. Так, текущее ТИ, представленное на рис. 5,б может быть описано в виде:

«Поверхность – вода – дорога» или «П – В – Д».

Для идентификации ТИ на ЭИ необходимо найти участки ЭИ, которые соответствуют этому описанию.

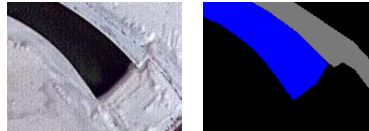
В данном описании не учитывались атрибуты объектов и отношения, требующие включения дополнительных процедур для выделения этих признаков.

На рис. 6 показаны результаты исследования семантического описания ЭИ при поиске участков, схожих по описанию ТИ.

Белым цветом на рисунке 6 отмечены области S_1 , соответствующие значению функции Джеккарда $K_D = 1$. Здесь семантические описания ТИ и ЭИ полностью совпадают. Координаты ТИ определяют координаты верхнего левого угла.



а – ЭИ и семантическое описание ЭИ



б – ТИ и семантическое описание ТИ

Рис. 5. Наблюдаемая сцена



Рис. 6. Визуализация функции Джекарда

Сплошное обследование семантического описания ЭИ с помощью функции Джекарда показало возможность существенного сокращения требуемого объема вычислений. Так, при сплошном поиске с использованием семантических описаний количество операций будет равно

$$K_{\Sigma D} = (M - n + 1)(N - n + 1)K_{fD}. \quad (11)$$

Следовательно, количество выполняемых элементарных операций уменьшилось в k -раз по сравнению с обработкой растровых изображений

$$k = \frac{mn(M - n + 1)(N - n + 1)K_o}{(M - n + 1)(N - n + 1)K_{fD}} = \frac{mnK_o}{K_{fD}}. \quad (12)$$

При выборочном сканировании ЭИ объем вычислений еще более сократится, т.к. количество обследуемых участков будет определяться плотностью расположения объектов (присутствующих на ТИ) на семантическом описании ЭИ.

В рассматриваемом примере $S_0 / S_1 \approx 7$.

Для дальнейшего повышения точности идентификации, в соответствии с алгоритмом, представленном на рис. 4, следующий этап поиска (бл.3, 4) реализован в области S_1 с использованием семантически сегментированных ТИ и ЭИ. Сопоставление изображений производилось функцией Джекарда.

Результаты расчетов показаны на рис. 7.

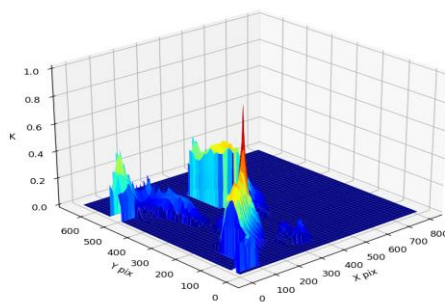


Рис. 7. Значения функции Джекарда в пространстве ЭИ

Идентификация ТИ проведена успешно и показало высокую точность оценки положения ТИ на ЭИ.

Однако необходимо отметить, что на этом этапе поиска сокращение вычислений было обеспечено только отношением количества реализуемых операций в корреляционных алгоритмах (K_0) и в функции Джекарда (K_{fD}).

$$k = \frac{mn(M - n + 1)(N - n + 1)K_0}{mn(M - n + 1)(N - n + 1)K_{fD}} = \frac{K_0}{K_{fD}}. \quad (13)$$

В целом, проведенные исследования показывают перспективность предложенного подхода.

Для оценки устойчивости семантических описаний наблюдаемых сцен предполагается проведение исследований, направленных на выявление ограничений к формированию семантических описаний, устойчивых к изменениям условий наблюдения.

Заключение. Предложен алгоритм решения задачи сравнения текущих и эталонных изображений на основе использования семантических описаний наблюдаемых сцен, позволяющий в ряде случаев заменить попиксельный анализ растровых изображений.

В работе приведены: общий алгоритм визуальной навигации, а также основные этапы методики формирования семантических описаний.

Предложены алгоритмы сравнения и идентификации семантических описаний ТИ и ЭИ по наличию классов объектов на основе функции Джекарда, а также иерархический алгоритм сравнения и идентификации изображений, основанный на последовательном применении семантических, и растровых описаний наблюдаемых сцен.

Показано, что использование процедуры сравнения семантических описаний ТИ и ЭИ по присутствующим классам объектов позволяет существенно сократить вычисления, необходимые для идентификации изображений.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования России, № темы FSFF-2024-0001.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Веремеенко К.К., Желтов С.Ю., Ким Н.В. и др. Современные информационные технологии в задачах навигации и наведения беспилотных маневренных летательных аппаратов / под ред. М.Н. Красильщикова, Г.Г. Себрякова. – М.: Физматлит, 2009. – 556 с.
2. Али Б., Садеков Р.Н., Цодокова В.В. Алгоритмы навигации беспилотных летательных аппаратов с использованием систем технического зрения // Гирскопия и навигация. – 2022. – Т. 30, № 4 (119). – С. 87-105.
3. Красовский А.А., Белоглазов И.Н., Чигин Г.П. Теория корреляционно-экстремальных навигационных систем. – М.: Наука, 1979.
4. Баклицкий В.К., Бочкарев А.М., Мусьяков М.П. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. – М.: Радио и связь, 1988.
5. Ким Н.В. Обработка и анализ изображений в системах технического зрения: учеб. пособие. – М.: Изд-во МАИ, 2001. – 164 с.
6. Горелик А.Л., Скрипкин В.А. Методы распознавания. – М.: Высшая школа, 1989. – 232 с.

7. Команда Лаборатории приборов, 2021, Оценка ошибки вычисления навигации инерциальными модулями. – <https://mp-lab.ru/navigation-error-estimation/?ysclid=lyg5vzaxqg875561868> (21 August 2024).
8. Pazychev D., Sadekov R. Simulation of INS Errors of Various Accuracy Classes. – 2020. – 1-3. – <https://doi.org/10.23919/ICINS43215.2020.9133869>.
9. Семантическая сегментация изображений с использованием видимого и ближнего инфракрасного каналов. – URL: <https://habr.com/ru/articles/723646/> (дата обращения: 17.10.2024).
10. Vijay Badrinarayanan, Alex Kendall. SegNet: A Deep Convolutional EncoderDecoder Architecture for Image Segmentation. – 2016. – С. 14.
11. Liang-Chieh Chen, George Papandreou. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. – 2017. – С. 14.
12. Shotton J., Winn J., Rother C., and Criminisi A. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context // IJCV. – 2009.
13. Shotton J., Johnson M., and Cipolla R. Semantic texton forests for image categorization and segmentation // in CVPR. – 2008.
14. Fulkerson B., Vedaldi A., and Soatto S. Class segmentation and object localization with superpixel neighborhoods // in ICCV. – 2009.
15. Старовойтов В.В., Голуб Ю.И. Об оценке результатов классификации несбалансированных данных по матрице ошибок // Информатика. – 2021. – Т. 18, № 1. – С. 61-71.
16. Лукашик Д.В. Анализ современных методов сегментации изображений // Экономика и качество систем связи. – 2022. – № 2.
17. Zhongzhi Shi. Intelligence Science. Leading the Age of Intelligence. – Tsinghua University Press, 2021. – <https://doi.org/10.1016/C2020-0-02066-9>.
18. Коновалов К.Д. Алгоритм планирования маршрута БПЛА в условиях недостатка визуальных ориентиров // Системы анализа и обработки данных. – 2024. – № 2 (94). – С. 37-54.
19. Zheng G., Jiang Z., Zhang H., Yao X. Deep semantic segmentation of unmanned aerial vehicle remote sensing images based on fully convolutional neural network // Frontiers in Earth Science. – 2023. – 11. – P. 12.
20. Pratap Singh K., Salvador J., Weihs L., Kembhavi A. Scene Graph Contrastive Learning for Embodied Navigation // IEEE/CVF International Conference on Computer Vision (ICCV). – 2023. – P. 10884-10894.

REFERENCES

1. Veremeenko K.K., Zheltov S.Yu., Kim N.V. i dr. Sovremennye informatsionnye tekhnologii v zadachakh navigatsii i navedeniya bespilotnykh manevrennykh letatel'nykh apparatov [Modern information technologies in the problems of navigation and guidance of unmanned maneuverable aerial vehicles], pod red. M.N. Krasil'shchikova, G.G. Sebryakova. Moscow: Fizmatlit, 2009, 556 p.
2. Ali B., Sadekov R.N., Tsodokova V.V. Algoritmy navigatsii bespilotnykh letatel'nykh apparatov s ispol'zovaniem sistem tekhnicheskogo zreniya [Navigation algorithms for unmanned aerial vehicles using machine vision systems], *Giroskopiya i navigatsiya* [Gyroscopy and Navigation], 2022, Vol. 30, No. 4 (119), pp. 87-105.
3. Krasovskiy A.A., Beloglazov I.N., Chigin G.P. Teoriya korrelyatsionno-ekstremal'nykh navigatsionnykh sistem [Theory of correlation-extreme navigation systems]. Moscow: Nauka, 1979.
4. Baklitskiy V.K., Bochkarev A.M., Mus'yakov M.P. Metody fil'tratsii signalov v korrelyatsionno-ekstremal'nykh sistemakh navigatsii [Methods of signal filtration in correlation-extreme navigation systems]. Moscow: Radio i svyaz', 1988.
5. Kim N.V. Obrabotka i analiz izobrazheniy v sistemakh tekhnicheskogo zreniya: ucheb. posobie [Image processing and analysis in machine vision systems: a tutorial]. Moscow: Izd-vo MAI, 2001, 164 p.
6. Gorelik A.L., Skripkin V.A. Metody raspoznavaniya [Recognition methods]. Moscow: Vysshaya shkola, 1989, 232 p.
7. Komanda Laboratorii priborov, 2021, Otsenka oshibki vychisleniya navigatsii inertial'nyimi modul'nyami [Instrument Laboratory Team, 2021, Estimation of the error in calculating navigation using inertial modules]. Available at: <https://mp-lab.ru/navigation-error-estimation/?ysclid=lyg5vzaxqg875561868> (accessed 21 August 2024).
8. Pazychev D., Sadekov R. Simulation of INS Errors of Various Accuracy Classes, 2020, 1-3. Available at: <https://doi.org/10.23919/ICINS43215.2020.9133869>.
9. Semanticheskaya segmentatsiya izobrazheniy s ispol'zovaniem vidimogo i blizhnego infrakrasnogo kanalov [Semantic segmentation of images using visible and near infrared channels]. Available at: <https://habr.com/ru/articles/723646/> (accessed 17 October 2024).

10. Vijay Badrinarayanan, Alex Kendall. SegNet: A Deep Convolutional EncoderDecoder Architecture for Image Segmentation, 2016, pp. 14.
11. Liang-Chieh Chen, George Papandreou. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, 2017, pp. 14.
12. Shotton J., Winn J., Rother C., and Criminisi A. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context, *IJCV*, 2009.
13. Shotton J., Johnson M., and Cipolla R. Semantic texton forests for image categorization and segmentation, in *CVPR*, 2008.
14. Fulkerson B., Vedaldi A., and Soatto S. Class segmentation and object localization with superpixel neighborhoods, in *ICCV*, 2009.
15. Starovoytov V.V., Golub Yu.I. Ob otsenke rezul'tatov klassifikatsii nesbalansirovannykh dannykh po matritse oshibok [On the assessment of the results of classification of unbalanced data by the error matrix], *Informatika* [Informatics], 2021, Vol. 18, No. 1, pp. 61-71.
16. Lukashik D.V. Analiz sovremennykh metodov segmentatsii izobrazheniy [Analysis of modern methods of image segmentation], *Ekonomika i kachestvo sistem svyazi* [Economics and quality of communication systems], 2022, No. 2.
17. Zhongzhi Shi. Intelligence Science. Leading the Age of Intelligence. Tsinghua University Press, 2021. Available at: <https://doi.org/10.1016/C2020-0-02066-9>.
18. Konovalov K.D. Algoritm planirovaniya marshruta BPLA v usloviyakh nedostatka vizual'nykh orientirov [Algorithm for planning a UAV route under conditions of insufficient visual reference points], *Sistemy analiza i obrabotki dannykh* [Data Analysis and Processing Systems], 2024, No. 2 (94), pp. 37-54.
19. Zheng G., Jiang Z., Zhang H., Yao X. Deep semantic segmentation of unmanned aerial vehicle remote sensing images based on fully convolutional neural network, *Frontiers in Earth Science*, 2023, 11, pp. 12.
20. Pratap Singh K., Salvador J., Weihs L., Kembhavi A. Scene Graph Contrastive Learning for Embodied Navigation, *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 10884-10894.

Ким Николай Владимирович – Московский авиационный институт (национальный исследовательский университет); e-mail: Nkim2011@list.ru; г. Москва, Россия; тел.: 89037208924; к.т.н.; профессор кафедры 704 «Информационно-управляющие комплексы ЛА».

Удалова Наталия Владимировна – Московский авиационный институт (национальный исследовательский университет); e-mail: vice1dean7@mai.ru; г. Москва, Россия; тел.: 89197724814; к.т.н.; доцент кафедры 704 «Информационно-управляющие комплексы ЛА».

Бодунков Николай Евгеньевич – Московский авиационный институт (национальный исследовательский университет); e-mail: bodunkov63@hotmail.com; г. Москва, Россия; тел.: 89168126007; к.т.н.; доцент кафедры 704 «Информационно-управляющие комплексы ЛА».

Гиренко Дмитрий Сергеевич – ООО "МИРП-ИС"; e-mail: dima@webiceberg.com; г. Дубна, Россия; инженер.

Ляпин Никита Александрович – Московский авиационный институт (национальный исследовательский университет); e-mail: lyapin2010@yandex.ru; г. Москва; тел.: 89164399972; к.т.н.; доцент кафедры 704 «Информационно-управляющие комплексы ЛА».

Kim Nikolay Vladimirovich – Moscow Aviation Institute (National Research University); e-mail: Nkim2011@list.ru; Moscow, Russia; phone: +79037208924; cand. of eng. sc.; professor of Department 704 "Aircraft Information and Control Complexes".

Udalova Natalia Vladimirovna – Moscow Aviation Institute (National Research University); e-mail: vice1dean7@mai.ru; Moscow, Russia; phone: +79197724814; cand. of eng. sc.; associate professor of Department 704 "Aircraft Information and Control Complexes".

Bodunkov Nikolay Evgen'evich – Moscow Aviation Institute (National Research University); e-mail: bodunkov63@hotmail.com; Moscow, Russia; phone: +79168126007; cand. of eng. sc.; associate professor of Department 704 "Aircraft Information and Control Complexes".

Girenko Dmitry Sergeevich – MIRP-IS LLC; e-mail: dima@webiceberg.com; Dubna, Russia; engineer.

Lyapin Nikita Alexandrovich – Moscow Aviation Institute (National Research University); e-mail: lyapin2010@yandex.ru; phone: +79164399972; cand. of eng. sc.; associate professor of Department 704 "Aircraft Information and Control Complexes".