

На правах рукописи

ГирДир

ГИРЕНКО Дмитрий Сергеевич

**ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ВИЗУАЛЬНОЙ
НАВИГАЦИИ ЗА СЧЕТ ИСПОЛЬЗОВАНИЯ СЕМАНТИЧЕСКОГО
ОПИСАНИЯ СЦЕН**

2.3.1. Системный анализ, управление и обработка
информации, статистика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук

Москва, 2025

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Московский авиационный институт (Национальный исследовательский университет)», МАИ.

Научный руководитель: **Ким Николай Владимирович**
кандидат технических наук, профессор, профессор кафедры «Информационно-управляющие комплексы летательных аппаратов» федерального государственного бюджетного образовательного учреждения высшего образования «Московский авиационный институт (национальный исследовательский университет)»

Официальные оппоненты: **Карпов Валерий Эдуардович**
доктор технических наук, доцент, начальник лаборатории робототехники Федерального государственного бюджетного учреждения «Национальный исследовательский центр Курчатовский институт»
Гласов Владислав Валерьевич
кандидат технических наук, доцент, директор проектного комплекса «Интеллектуальный борт» Федерального государственного бюджетного учреждения «Национальный исследовательский центр Институт имени Н.Е. Жуковского»

Ведущая организация: **Публичное акционерное общество «Московский институт электромеханики и автоматики», 125167, г. Москва, Авиационный переулок, д.5**

Защита состоится 25 декабря 2025 г. в 13.00 часов на заседании диссертационного совета 24.2.327.03 в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский авиационный институт (национальный исследовательский университет)» по адресу: 125993, г. Москва, Волоколамское шоссе, д. 4.

С диссертацией можно ознакомиться в библиотеке и на сайте МАИ по ссылке: https://mai.ru/events/defence/?ELEMENT_ID=186140

Автореферат разослан “ ____ ” 2025 г.

Отзывы, заверенные печатью, просим направлять по адресу: 125993, г. Москва, Волоколамское шоссе, д. 4, Отдел подготовки кадров высшей квалификации МАИ.

Ученый секретарь диссертационного совета
24.2.327.03, д.т.н., доцент



А.В. Старков

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. В настоящее время все большую роль в мире играют малые беспилотные летательные аппараты (БЛА), способные самостоятельно выполнять поставленные целевые задачи (ЦЗ).

Эффективность выполнения ЦЗ в существенной степени зависит от качества работы навигационной системы (НС), основу которой составляют приемники глобальных навигационных спутниковых систем (ГНСС), совмещенные с блоком инерциальных датчиков (ИНС - инерциальная навигационная система) пространственной ориентации.

Для коррекции работы ИНС при отсутствии сигналов ГНСС широко используются методы визуальной навигации. В основе визуальной навигации по наземным ориентирам лежит идея сравнения текущего изображения (ТИ) подстилающей поверхности, принимаемого бортовой видеокамерой БЛА, с эталонным изображением (ЭИ – цифровой картой местности, хранящейся в бортовом вычислителе). Необходимость сравнения в реальном времени растровых изображений ТИ и искомого фрагмента ЭИ предъявляет высокие требования к производительности бортового вычислителя. Кроме того, изменение условий освещенности подстилающей поверхности затрудняет оценку изображений, что может привести к недопустимым ошибкам привязки ТИ к ЭИ.

Темой диссертации является исследование и разработка алгоритмов визуальной навигации, позволяющих повысить качество навигационного решения при использовании визуальной навигации.

Целью работы является повышение эффективности алгоритмов визуальной навигации с точки зрения помехоустойчивости и вычислительной производительности за счет использования семантических описаний наблюдаемых сцен.

Объектом исследования является навигация автономного БЛА.

Предметом исследования является разработка методики построения алгоритмов визуальной навигации на основе сравнения сходства семантических описаний наблюдаемых сцен на текущих и эталонных изображениях.

Научная новизна работы заключается в следующем:

1. Разработана методика формирования семантических описаний наблюдаемых сцен.
2. Разработан алгоритм визуальной навигации на основе сравнения сходства семантических описаний наблюдаемых сцен на текущем и эталонном изображениях.
3. Разработан алгоритм оценки информативности семантических признаков для оценки задачи распознавания наблюдаемых сцен.
4. Предложен подход к повышению точности решения задачи визуальной навигации при полете над малоинформационной местностью за счет оценки информативности семантических описаний подстилающей поверхности планируемых вариантов маршрута.
5. Показано, что предлагаемые технические решения позволяют повысить вычислительную производительность и помехоустойчивость алгоритмов визуальной навигации по сравнению с корреляционно-экстремальными алгоритмами, основанными на обработке растровых изображений.

Практическая значимость работы заключается в том, что предлагаемые методики и алгоритмы позволяют повысить вычислительную производительность и помехоустойчивость алгоритмов визуальной навигации БЛА.

Результаты диссертационной работы используются в рамках государственного задания Министерства науки и высшего образования России, номер темы FSFF-2024-0001.

Методы исследования. Исследования выполнены с использованием методов математического моделирования, технологии компьютерного зрения, нейросетевых методов распознавания.

К защите представляются следующие основные положения работы:

1. Методика и алгоритм визуальной навигации на основе семантических описаний наблюдаемых сцен, обеспечивающий повышение вычислительной производительности и помехоустойчивости по сравнению с корреляционно-экстремальными алгоритмами с использованием растровых изображений.

2. Методика формирования семантических описаний наблюдаемых сцен.

3. Алгоритм идентификации семантических описаний наблюдаемых сцен на основе функции Джекарда.

4. Результаты исследований влияния изменения условий наблюдений на точность привязки семантических описаний ТИ и ЭИ.

5. Результаты моделирования, подтверждающие работоспособность и эффективность предлагаемых подходов.

Достоверность результатов, полученных в работе, подтверждается результатами математического моделирования процессов визуальной навигации БЛА на реальных изображениях земной поверхности.

Личный вклад автора и публикации

Все результаты настоящего диссертационного исследования получены автором лично. В том числе, автором разработаны: алгоритм сопоставления семантических описаний, иерархический алгоритм сопоставления сцен по семантическому описанию, алгоритм оценки информативности предполагаемых вариантов маршрута полета. По результатам диссертационной работы опубликовано 13 работ, в том числе 1 Свидетельство о государственной регистрации программ для ЭВМ, 1 Патент на изобретение, 1 отчет о НИР, 3 научно-технические статьи в изданиях, включенных в перечень ВАК, 3 доклада конференций, включенные в пучки SCOPUS и Web of Science.

Апробация работы. Результаты работы докладывались на 49-й Международной молодежной конференции «Гагаринские чтения» (Москва, 2023); 5-й Международной научно-технической конференции «Фотограмметрические методы мониторинга окружающей среды и инфраструктуры, биометрии и биомедицины» (Москва, 2023); 23-й Международной конференции по науке и технологиям РОССИЯ – КОРЕЯ – СНГ (Москва, 2023); 24-й Международной конференции по науке и технологиям РОССИЯ – КОРЕЯ – СНГ (Москва, 2024); Международной научно-технической конференции «Анализ фотограмметрических данных» (ISPRS PDA24) (Москва, 2024); 51-й Международной молодежной научной конференции «Гагаринские чтения» (Москва, 2025); 20-й Юбилейной Всероссийской научно-практической конференции «Перспективные системы и задачи управления» («Домбайская конференция»); 6-й Международной научно-технической конференции «Методы фотограмметрии и машинного зрения для мониторинга окружающей среды, биометрии и биомедицины» (PSBB25) (Москва, 2025).

Структура и объем работы

Диссертация состоит из введения, четырех глав, заключения и списка литературы. Работа изложена на 134 страницах машинописного текста, содержит 51 рисунок, 9 таблиц и 70 наименование литературных источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы, проводится анализ литературы,дается общая характеристика работы.

В **главе 1** рассматриваются особенности визуальной навигации при решении целевых задач автономным БЛА. Проведен анализ подходов к определению местоположения БЛА методами визуальной навигации.

ЦЗ, решаемые БЛА, например, мониторинг местности или доставка грузов, определяются уровнем автономности БЛА. БЛА нулевого, первого и второго уровней в основном управляются оператором. БЛА 3-го и более высоких уровней автономности должны обеспечивать самостоятельную корректировку маршрута, полет в областях с отсутствием сигналов ГНСС с использованием методов визуальной навигации, обработку и анализ принимаемой на борт видеинформации.

Методы визуальной навигации позволяют решать задачи коррекции ИНС. Однако из-за известной проблемы ухода ИНС со временем, полет в областях с неинформативными или малоинформативными участками местности требует дополнительной локальной корректировки маршрута.

Реализация алгоритмов визуальной навигации для высоких уровней автономности в существенной степени ограничивается сложностью реализуемых алгоритмов, а также возможностями современных бортовых вычислительных средств и их стоимостью. Поэтому разработка технологий, позволяющих повысить вычислительную производительность систем визуальной навигации, является актуальной и практически значимой.

В общем случае в алгоритмах визуальной навигации могут быть использованы разные варианты описаний наблюдаемых сцен. Например, растровое описание (распределение яркости и цветов в отдельных пикселях) и признаковое описание (выделяются особенности (признаки) отдельных объектов или сцен).

Результатом процесса визуальной навигации (поиска фрагмента ЭИ) является найденный (искомый) фрагмент ЭИ, идентичный (или достаточно похожий) принятому ТИ. Координаты найденного фрагмента на ЭИ позволяют произвести привязку координат БЛА к имеющейся цифровой карте местности (ЦКМ). Оценка сходства описаний ТИ и фрагментов ЭИ реализуется на основе использования разных метрик, например, расстояния Махalanобиса, корреляционных функций и т.д.

Определим основные требования, предъявляемые к решению задачи визуальной навигации. Под точностью определения местоположения БЛА в данной работе подразумевается точность определения географических координат в земной (местной) системе координат $X0Y$. Будем считать, что рассматриваемые БЛА оснащены датчиками магнитного курса и высотомерами. Примем, что условие решения навигационной задачи имеет следующий вид:

$$3\sigma_x = \Delta x \leq |\Delta x_0|, 3\sigma_y = \Delta y \leq |\Delta y_0|, \quad (1)$$

где $\Delta x_0, \Delta y_0$ – допустимые ошибки оценки координат; σ_x и σ_y – с.к.о. оценки координат.

Помимо точности оценки координат, в рамках решаемых задач целесообразно использовать показатель вероятности правильного распознавания искомого фрагмента ЭИ:

$$P(m) \geq P_0, \quad (2)$$

где $P(m)$ – вероятность правильного распознавания m -го фрагмента ЭИ из M_{Σ} возможных; P_0 – требуемая вероятность.

Требуемые значения параметров Δx_0 , Δy_0 и P_0 определяются решаемой целевой задачей. Процесс сравнения и идентификации наблюдаемых сцен на ТИ и ЭИ может рассматриваться как процесс распознавания, в котором распознаваемым объектом является наблюдаемая сцена на ТИ, а алфавитом классов – все возможные фрагменты ЭИ.

Наиболее распространенной в визуальной навигации является реализация корреляционно-экстремальных алгоритмов, использующих растровые изображения. К основным проблемам реализации корреляционных методов визуальной навигации БЛА, относятся изменяющие условия наблюдения, искажающие принимаемые ТИ относительно ЭИ, и высокая трудоемкость вычислений, связанных с необходимостью обработки объемных растровых изображений.

Возможным направлением решения проблем недостаточной помехоустойчивости методов визуальной навигации с использованием растровых изображений является применение нейросетевых технологий. Выделение классов объектов, присутствующих на наблюдаемых сценах, позволяет перейти от анализа растровых изображений к содержанию (семантическому описанию) наблюдаемых сцен.

Таким образом, для повышения производительности и помехоустойчивости алгоритмов визуальной навигации предлагается перейти от растрового описания к семантическому.

Для перехода к использованию семантических описаний наблюдаемых сцен в визуальной навигации необходимо разработать процедуры формирования семантических описаний, ориентированных на решение поставленных задач и разработать алгоритм их идентификации (распознавания).

Семантическими признаками (атрибутами) наблюдаемых сцен на ТИ и ЭИ являются: классы и количество присутствующих на сцене объектов, атрибуты объектов и пространственные отношения (взаимное положение и ориентация) между объектами.

В ряде ситуаций требуется оценить информативность ориентиров или наблюдаемых сцен с точки зрения полезности решения целевой задачи. Например, данная проблема возникает при полете над малоинформационной местностью. Наблюдаемая поверхность малоинформационной местности в существенной степени состоит из участков, на которых отсутствуют визуальные ориентиры, обеспечивающие возможность привязки БЛА. При этом присутствующие ориентиры могут иметь полезную информативность, зависящую от конкретных условий наблюдения.

Таким образом, при возникновении зон с отсутствием сигналов ГНСС и при наличии ЦКМ необходима прокладка маршрутов через наиболее информативные участки местности (с точки зрения визуальной навигации), минимизирующие накапливаемые ошибки ИНС.

Выбор оптимального маршрута с учетом полезной информативности наблюдаемых участков местности может быть рассчитан на основе энтропии, оценивающей неопределенность положения БЛА.

Примем, что положение БЛА оценивается на ЭИ (ЦКМ) в области n -го участка, используемого для визуальной навигации. БЛА (с учетом возможных ошибок ИНС) может находиться в одной из M точек ЦКМ и принимать различные изображения местности (ТИ).

Обозначим U_{nk} – значения признака (функции) наблюдаемой сцены, выделяемого на ТИ и позволяющего оценивать сходство изображений. Предполагается, что при всех возможных положениях ТИ может быть получено $K \leq M$ разных значений U_{nk} .

Информативность n -го участка местности равна:

$$I_n = H(m) - H(m|U_{nk}) = - \sum_{m=1}^M P(m) \log_2 P(m) + \sum_{l=1}^M P(l) \sum_{m=1}^M \sum_{k=1}^K P(U_{nk}|l) P(m|U_{nk}) \log_2 P(m|U_{nk}), \quad (3)$$

где m, l – точки на ЭИ из M , в которых возможно нахождение БЛА; $H(m)$ – априорная энтропия положения БЛА; $H(m|U_{nk})$ – апостериорная энтропия положения БЛА; $P(m), P(l)$ – вероятности присутствия БЛА в точках m, l соответственно; $P(U_{nk}|l)$ – условная вероятность получения значения признака U_{nk} при присутствии БЛА в точке l ; $P(m|U_{nk})$ – апостериорная вероятность присутствия объекта в точке m , при условии получения U_{nk} .

Таким образом, для достижения поставленной цели необходимо:

1. Разработать методику формирования семантических описаний наблюдаемых сцен на ЭИ и ТИ.
2. Разработать алгоритм поиска фрагмента ЭИ, идентичного принятому ТИ, на основе использования семантических описаний наблюдаемых сцен, включающий алгоритм идентификации (распознавания) семантического описания ТИ.
3. Разработать алгоритм оценки информативности семантических признаков наблюдаемых сцен.
4. Провести исследования по оценке работоспособности, помехоустойчивости и вычислительной производительности разработанного алгоритма.

Глава 2 посвящена проблеме формирования семантических описаний наблюдаемых сцен.

Под семантическим описанием наблюдаемых сцен понимается перечисление классов присутствующих на изображении объектов, описание их атрибутов и пространственных отношений между ними в словесном или кодированном виде.

Выделяемые классы объектов определяются априорно сформированным алфавитом классов. Атрибуты объектов зависят от классов объектов и, в общем случае, включают форму, размеры, текстуру и пр. Пространственные отношения между парами объектами определяются вектором – расстоянием между объектами и направлением.

Семантические описания наблюдаемых сцен могут иметь следующий **словесный вид**: «Большое здание рядом с дорогой». Здесь «здание», «дорога» являются классами объектов, «большой» – атрибут, а «рядом» – отношения между объектами.

В **кодированном виде** семантическое описание может быть представлено в следующем варианте. Описание объектов: $B(Numb)$ – класс «здание»; $R(Numb)$ – класс «дорога», где $Numb$ – номер объекта данного класса. Атрибуты ($Attr$) объектов: $B(Numb, Attr(1), Attr(2), \dots)$. Пространственные отношения (SR) между парами объектов $Ob(i), Ob(j)$ определяются расстоянием ($Dist$) и направлением (Dir): $SR(Dist((Ob(i), Ob(j)), Dir((Ob(i), Ob(j))))$.

Методика формирования семантических описаний наблюдаемых сцен включает следующие этапы:

- 1) Определение классов объектов, выделяемых на наблюдаемых сценах;
- 2) Формирование словарей признаков наблюдаемой сцены, в зависимости от цикла поиска включающих: классы объектов, атрибуты объектов, отношения;
- 3) Семантическая сегментация изображений, определяющая классы объектов, их количество и расположение на наблюдаемой сцене;
- 4) Выделение атрибутов объектов и отношений между объектами;
- 5) Формирование семантических описаний (*SD*) из семантических изображений наблюдаемых сцен в словесном или кодированном виде.

Определение классов объектов (п. 1 Методики) проводится на этапе подготовки ЭИ.

Рабочий словарь признаков (п. 2), формируется на основе признаков, выделяемых на ТИ. Так как после семантической сегментации исходных изображений известными являются классы объектов и их расположения на наблюдаемых сценах, то первично для сокращения области поиска в качестве признаков используются только они.

Для выявления атрибутов объектов и отношений требуется дополнительный анализ наблюдаемой сцены и реализация соответствующих вычислительных процедур.

На рис. 1 в выделенном прямоугольнике показан состав возможных семантических описаний наблюдаемых сцен.

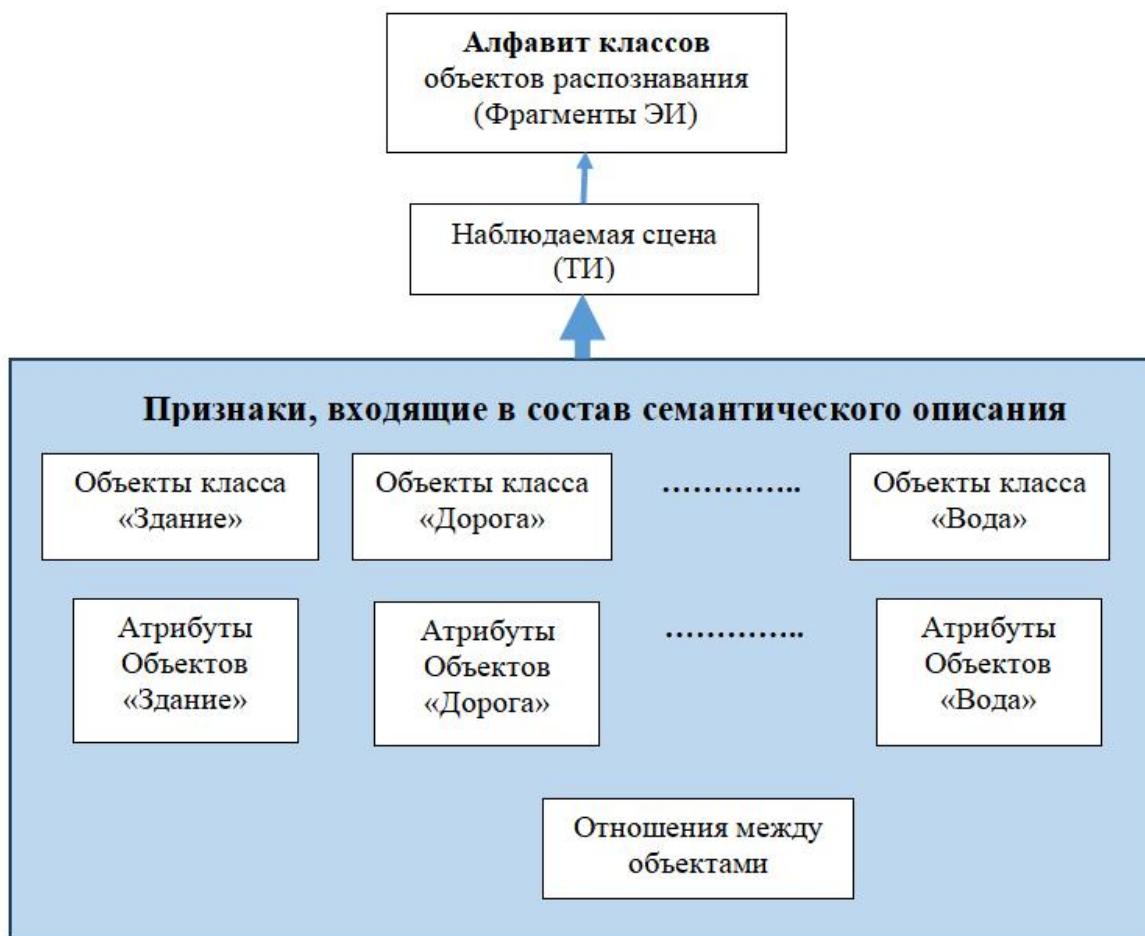


Рисунок 1 – Состав семантического описания

Семантическая сегментация изображений (п. 3). Основным инструментом решения данного класса задач являются сверточные нейронные сети. Классическими архитектурами являются *FCN*, *SegNet*, *U-net*.

Для представленных моделей проведена сравнительная характеристика их работы по метрике *IoU* (Intersection over Union) или индекс Жаккара (Jaccard index) на датасете сегментированных улиц. Результаты представлены на рис. 2.

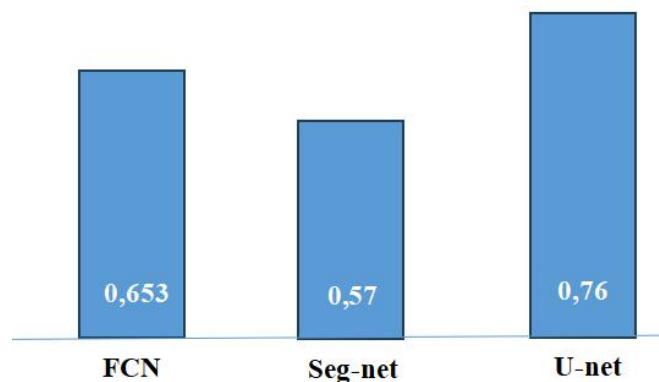


Рисунок 2 – Результаты сравнения нейронных сетей по метрике IoU

Анализируя рис. 2, отметим, что наибольший показатель качества по метрике IOU достигла нейронная сеть архитектуры U-net. Таким образом, в данной работе было принято решение использовать архитектуру U-net.

В результате работы нейронной сети каждый пиксель сегментированного изображения содержит в себе номер класса объекта. Однако на сегментированном изображении присутствуют некоторые ошибки: слияние разных объектов, разделение одного объекта на несколько, наличие нераспознанных пикселей внутри объекта, неровность границ. Данные неточности устраняются путем применения алгоритмов морфологической обработки изображений.

Результатом выполнения данного этапа (п. 3) является сегментированное семантическое изображение с выделенными и локализованными объектами, соответствующими заданному алфавиту классов.

Выделение атрибутов объектов и отношений между объектами (п. 4) необходимо, когда семантическое описание, основанное на использовании в качестве признаков только классов объектов, недостаточно для определения точного положения БЛА. В этом случае необходимо расширять словарь признаков за счет использования атрибутов объектов и/или отношений между объектами.

Выделению атрибутов (признаков) наблюдаемых сцен и отдельных объектов посвящено большое количество работ в области теории распознавания и систем технического зрения. В частности, широко внедрены алгоритмы выделения формы, контуров, ключевых точек и пр.

Например форма зданий может классифицироваться обученной нейросетью. На рис. 3 представлены типы форм зданий: прямоугольная, Г, Т, П, Н, О - форма.

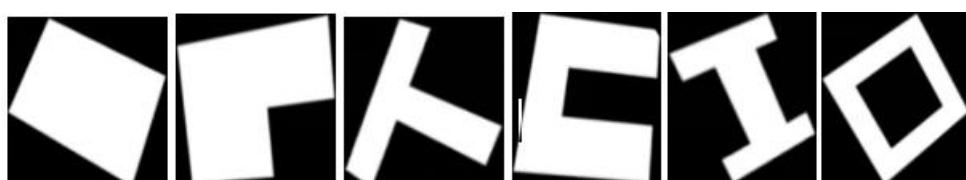


Рисунок 3 – Типы форм зданий

Одними из наиболее устойчивых к помехам атрибутов наблюдаемых сцен являются пространственные отношения между объектами, которые оцениваются расстояниями и направлениями.

В связи с возможными искажениями форм выделяемых нейросетью объектов возможно использование оценок, используемых в нечеткой логике. В терминах нечеткой логики расстояния можно оценивать, например, в следующих градациях: **вплотную (Вп)**, **близко (Бл)**, **неблизко (Нб)**, **недалеко (Нд)**, **далеко (Дл)** и пр.

Направление, оцениваемое в угловой мере, может оцениваться в радианах, угловых градусах или в градациях: **север (С)**, **юг (Ю)**, **запад (З)**, **Восток (В)** и пр.

Семантическое описание наблюдаемых сцен

Обозначим $SD_{CI}(F^r)$ – семантическое описание ТИ, $SD_{RI}(F^r, m)$ – семантическое описание фрагмента m ЭИ, где $m = 1, \dots, M$ – номер фрагмента ЭИ из M возможных. Номер фрагмента m соответствуют изображению, получаемому на борту БЛА, который находится в точке m ЦКМ. Рабочий словарь признаков F^r является переменным и может включать разные признаки наблюдаемых сцен: классы объектов, атрибуты и отношения. Семантическое описание наблюдаемой сцены зависит от рабочего словаря признаков F^r , используемого на данном r -м цикле поиска.

В **главе 3** рассматривается задача поиска фрагмента ЭИ, идентичного ТИ, на основе их семантических описаний. Оценка сходства семантических описаний производится по оценке сходства соответствующих признаков из рабочего словаря признаков F^r . Основными проблемами при разработке соответствующего алгоритма поиска являются:

- 1) выбор метрики оценки сходства наблюдаемых сцен;
- 2) формирование словаря семантических признаков распознавания;
- 3) определение процедуры принятия решения об окончании поиска и переходе к расчету координат БЛА.

Выбор метрики оценки сходства. Будем считать, что семантические описания двух сравниваемых наблюдаемых сцен сформированы полностью с достаточной степенью точности. Тогда степень сходства должна быть определена по количеству совпадающих и несовпадающих значений признаков. Данная задача для бинарных изображений решается при помощи парных критериальных функций.

Проведенный сравнительный анализ показал, что в рамках решаемых задач наиболее экономичной при достаточной точности является функция Джекарда:

$$E(F^r_l, m^r) = \frac{a_l}{a_l + b_l + c_l}. \quad (4)$$

где a_l – количество совпадающих признаков l (1, 1) – здесь первая цифра «1» показывает наличие признака в $SD_{CI}(F^r)$, а вторая – в $SD_{RI}(F^r, m^r)$; b_l, c_l – количество несовпадающих признаков l в $SD_{CI}(F^r)$ и $SD_{RI}(F^r, m^r)$ со значениями соответственно (1, 0) и (0, 1).

Сложность использования парных функций для сравнения семантических описаний состоит в том, что в парных функциях сравниваются только два возможных значения яркости: 1 и 0. Таким образом для сравнения семантических описаний ТИ и ЭИ необходимо: проводить оценку сходства отдельно для каждого признака; обозначить присутствие и отсутствие оцениваемого **признака** на наблюдаемой сцене показателями, соответственно – (1) и (0) и использовать их вместо значений яркости; проводить свертку показателей сходства по всем оцениваемым признакам.

Формирование словаря признаков распознавания. Эффективность процессов распознавания определяется информативностью используемых признаков. В сложных задачах с ограниченными вычислительными ресурсами более рациональным является реализация иерархических алгоритмов распознавания.

Определение процедуры принятия решения об окончании поиска. Сложность данной проблемы связана с тем, что окончание поиска можно реализовать по выполнению условия достижения требуемого уровня сходства некоторого $SD_{RI}(F^r, m^r)$ с $SD_{CI}(F^r)$. Для случаев когда несколько фрагментов ЭИ будут иметь одинаково высокую степень сходства с $SD_{CI}(F^r)$ в работе предлагается ввести условие окончания поиска, на основе оценки апостериорной вероятности получения искомого фрагмента ЭИ.

Таким образом, для решения указанных проблем предлагается реализовать **иерархический алгоритм поиска**. Основная идея данного алгоритма – последовательное сокращение области поиска за счет использования рабочего словаря признаков, расширяемого на каждом цикле поиска. Поиск продолжается, пока не будет достигнута требуемая (заданная) точность сходства или пока не будет использован полный словарь признаков. Блок-схема предлагаемого алгоритма представлена на рис. 4.

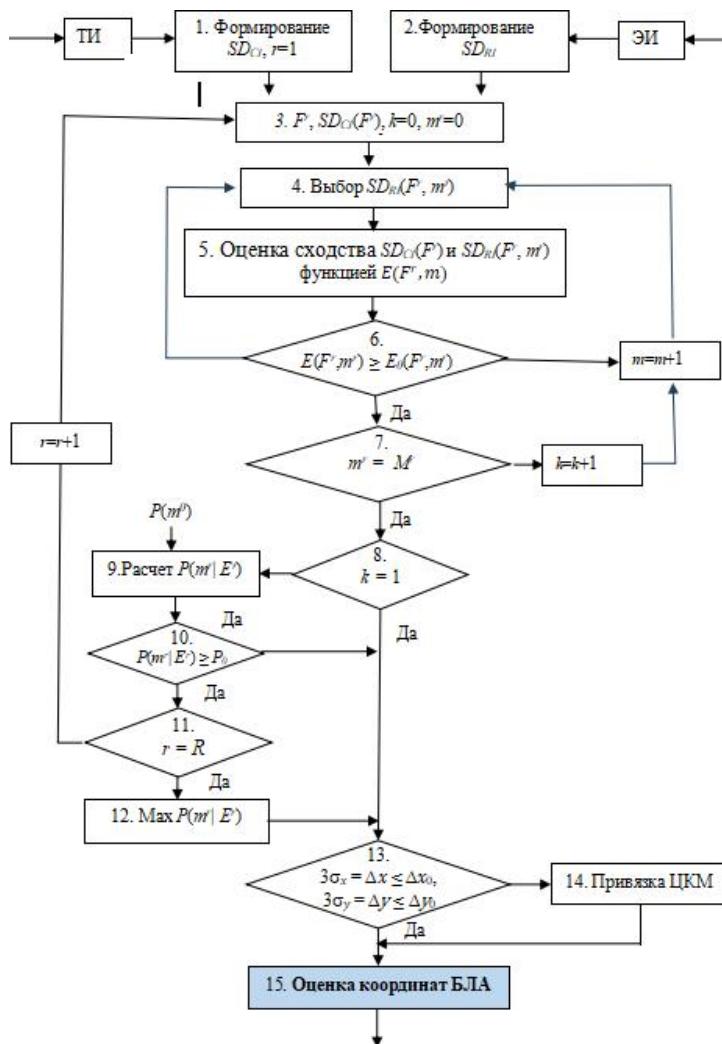


Рисунок 4 – Блок-схема алгоритма идентификации SD_{CI}

В бл. 4 производится выбор очередного фрагмента семантического описания ЭИ $SD_{RI}(F^r, m)$ для сравнения с $SD_{CI}(F^r)$.

В бл. 1, 2 (рис. 4, блоки 1, 2) на базе полученных семантических изображений формируются семантические описания сцен ТИ и ЭИ.

Для наблюдаемой сцены могут быть сформированы разные семантические описания $SD_{CI}(F^r)$ (бл. 3), базирующиеся на использовании различных словарей признаков F^r , где $r=1, \dots, R$ – индекс цикла сравнения ТИ и ЭИ.

Словарь признаков формируется на основе семантических признаков, выделяемых на ТИ. При этом признаки F^{rl} , где $l=1, \dots, L^r$ – индексы признаков, могут быть простыми и комбинированными.

Так, после рассмотренной семантической сегментации исходных ТИ и ЭИ известными являются классы объектов и их расположения на наблюдаемых сценах. Для выявления атрибутов объектов и отношений требуется реализация дополнительных вычислительных процедур. Поэтому на 1-м этапе поиска ($r=1$) в качестве признаков используются классы и их количество, найденные на этапе семантической сегментации.

В **бл. 5** проводится оценка сходства описаний $SD_{Cl}(F^r)$ и $SD_{Rl}(F^r, m^r)$ по всем L^r признакам по модифицированной функции Джекарда (МФД):

$$E(F^r, m^r) = \frac{1}{L^r} \sum_{l=1}^{L^r} E(F^r_l, m^r). \quad (5)$$

По каждому признаку в (5) l сходство оценивается на базе парной критериальной функции Джекарда (4), в которой a_l, b_l, c_l получены из описаний $SD_{Cl}(F^r)$ и $SD_{Rl}(F^r, m^r)$:

В **бл. 6** оценивается условие достижения требуемого уровня сходства, где $E_0(F, m^r)$ – заданное значение. Если условие не выполняется, то производится выбор нового фрагмента ЭИ.

Если проверены (**бл. 7**) все возможные варианты ($m^r=M^r$) и получено (**бл. 8**) единственное решение ($k=1$, где k – количество фрагментов ЭИ, в которых семантические описания $SD_{Rl}(F^r, m^r)$ идентичны $SD_{Cl}(F^r)$), то происходит переход к вычислению координат БЛА (**бл. 13**).

В противном случае производится расчет по формуле Байеса (**бл. 9**) апостериорных вероятностей приема фрагментов m^r из M^r :

$$P(m^r | E^{r-1}) = \frac{P(m^{r-1})P(E(F^{r-1}, m^{r-1}) | m^{r-1})}{\sum_{k^r}^{M^r} [P(k^{r-1})P(E(F^r, k^{r-1}) | k^{r-1})]}, \quad (6)$$

где E^{r-1} – оценка сходства $E(F^{r-1}, m^{r-1})$; $P(m^r)$ – вероятность приема фрагмента m^r на борту БЛА после реализации ($r-1$)-го цикла распознавания; m^r, k^r – номера фрагментов на r -м цикле распознавания.

На 1-м цикле поиска для расчета используется априорная вероятность $P(m^0)$, полученная на основании расчета ошибок ИНС.

Если полученная вероятность выше заданной (**бл. 10**), то происходит переход к расчету координат БЛА. В противном случае производится проверка на окончание циклов поиска (**бл. 11**). Если апостериорная вероятность ниже заданной и все циклы распознавания завершены, то определяется фрагмент с максимальной апостериорной вероятностью приема (**бл. 12**) и происходит переход к **бл. 13**.

Если циклы не завершены (**бл. 11**), то происходит переход к **бл. 3** и процесс распознавания продолжается.

В **бл. 13** проверяется условие (1). Если условие не выполняется, то в **бл. 14** производится привязка координат объектов ТИ к координатам соответствующих объектов на найденном фрагменте ЭИ (цифровой карте местности).

В **бл. 15** производится пересчет координат ТИ в координаты БЛА с учетом высоты полета и направления линии визирования.

В **главе 4** приведены результаты моделирования работы предлагаемых подходов.

Примем, что в общем случае под количественной оценкой помехоустойчивости принимается достоверность, определяемая частотой правильного распознавания обученной нейросетью классов элементов изображений объектов:

$$P_{tp}^k(q) = \frac{n_{tp}^k}{N_0^k}, \quad (7)$$

где n_{tp}^k – количество правильно распознанных элементов k -го класса;

N_0^k – количество элементов k -го класса;

q – индекс дестабилизирующего фактора.

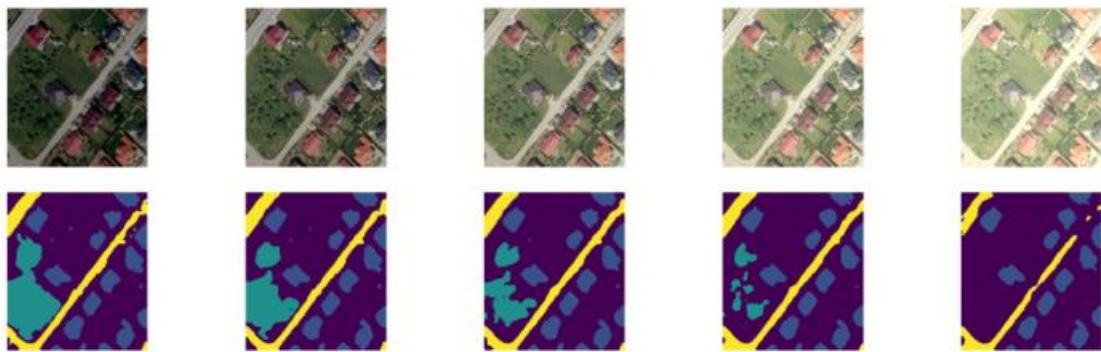
Устойчивость к белому гауссовскому шуму оценивается вероятностью $P(\beta)$, рассчитываемой по формуле (7).

Величина шума характеризуется отношением:

$$\beta = SNR = \frac{\sigma_s^2}{\sigma_n^2}, \quad (8)$$

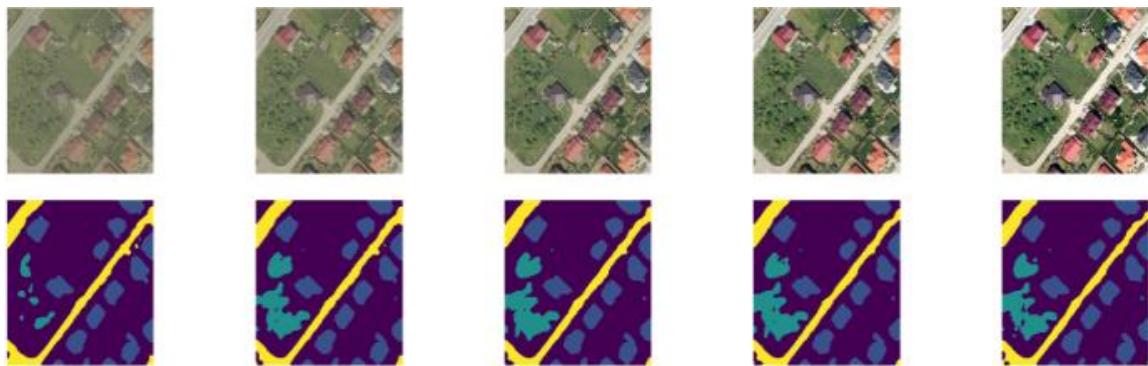
где σ_s^2 – мощность сигнала; σ_n^2 – мощность шума.

В первом эксперименте проводится оценка работы сегментационной нейронной сети. Целью данного исследования является оценка помехоустойчивости осуществления семантической сегментации изображений. Влияние освещенности (изменения математического ожидания (МО), стандартное отклонение (СКО) яркости и отношения сигнал к шуму (8)) на работу нейронной сети показано на рис. 5 - 7.



а) -50% МО б) -25% МО в) исходное МО г) +25% МО д) +50% МО

Рисунок 5 – Пример выхода нейронной сети, при изменении МО исходного изображения



а) -50% СКО б) -25% СКО в) исходное СКО г) +25% СКО д) +50% СКО
 Рисунок 6 – Пример выхода нейронной сети, при изменении СКО исходного изображения

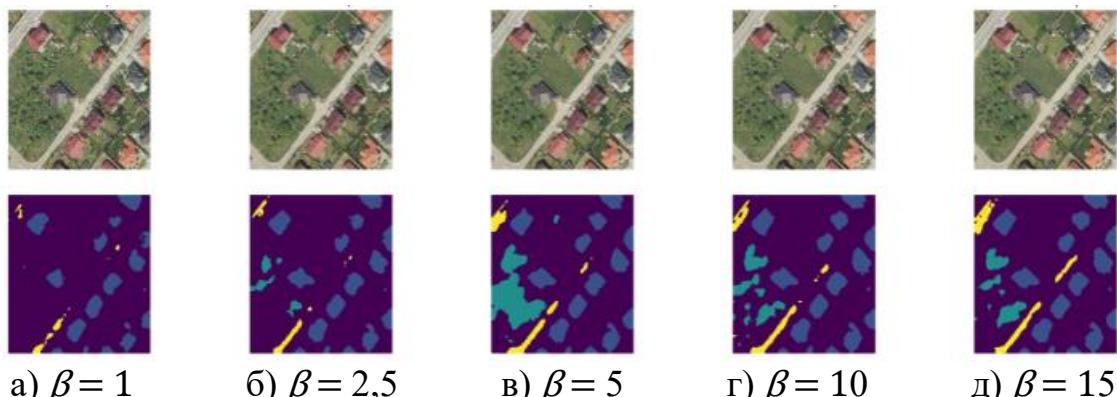


Рисунок 7 – Пример выхода нейронной сети, при изменении отношения сигнала к шуму на исходном изображении

В табл. 1 представлены результаты обработки 1000 изображений без модификации и при условии наиболее худших условий наблюдения: увеличение МО яркости пикселей на 50%, уменьшение СКО пикселей на 50% и при соотношении $\beta = 1$.

Таблица 1 – Результаты обработки 1000 изображений

Класс	$P_{tp}^k(q)$ в обычных условиях (без модификации)	$P_{tp}^k(q)$ при изменении МО яркости на +50%	$P_{tp}^k(q)$ при изменении СКО яркости на -50%	$P_{tp}^k(q)$ при $\beta = 1$
Здание	0,74	0,623	0,634	0,432
Растительность	0,91	0,71	0,85	0,002
Вода	0,93	0,95	0,496	0,648
Дорога	0,59	0,604	0,601	0,026
Фон	0,92	0,966	0,914	0,991

Результаты эксперимента показали, что использование нейросетевых алгоритмов семантической сегментации обеспечивают устойчивость распознавания пикселей в широком диапазоне вариации освещенности.

Во втором эксперименте проводится оценка помехоустойчивости алгоритма сопоставления сцен на основе семантического описания при воздействии гамма коррекции и белого шума.

В табл. 2 приведена оценка помехоустойчивости МФД в сравнении с нормированной корреляционной функции (НКК). Семантическое описание сформировано отдельно из классов, атрибутов и отношений. Также рассмотрен вариант использования попиксельной обработки семантического изображения при помощи МФД.

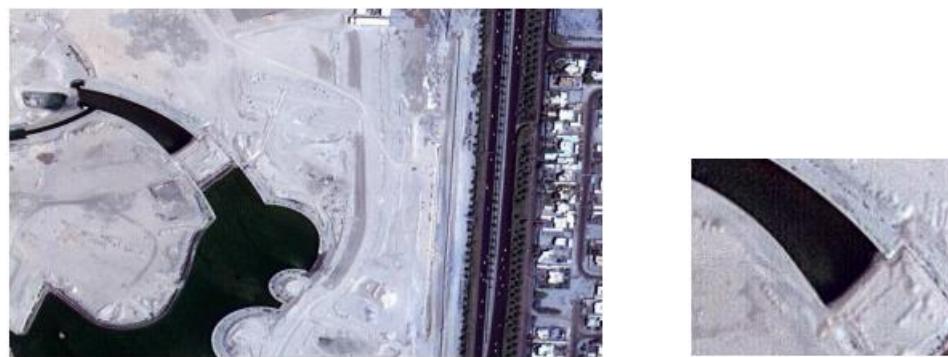
Таблица 2 – Результаты оценки помехоустойчивости

Алгоритм сопоставления	Достоверность распознавания при $\gamma = 0,5$	Достоверность распознавания при $\gamma = 1,5$	Достоверность распознавания при $\beta = 5$
МФД: Классы	0,95	0,913	0,812
МФД: Атрибуты	0,728	0,7	0,536
МФД: Отношения	0,745	0,72	0,59
Попиксельная обработка МФД	0,544	0,648	0,447
НКК	0,475	0,36	0,11

Результаты эксперимента показали, что МФД практически не подвержена увеличению ошибки в условиях вариации изображения при помощи гамма-коррекции в пределах от $\gamma = 0,5$ до $\gamma = 1,5$ и при воздействии аддитивного белого гауссовского шума с СКО, соответствующим $\beta \geq 5$.

В третьем эксперименте проводится сравнение точности корреляционных алгоритмов сопоставления растровых изображений и алгоритмов, использующих семантические описания. Целью данного исследования является сравнение точности локализации (оценки координат) ТИ существующих корреляционных алгоритмов обнаружения ориентиров с использованием ЭИ и предлагаемого алгоритма на основе семантических описаний.

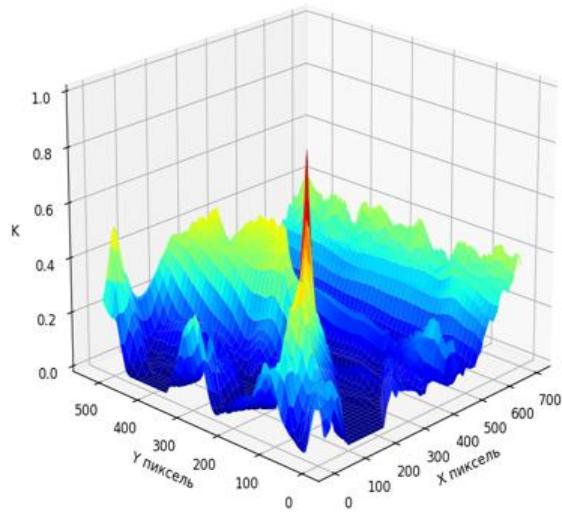
Осуществим поиск ТИ (рис. 8 б) на ЭИ (рис. 8 а) при помощи предлагаемого алгоритма и НКК.



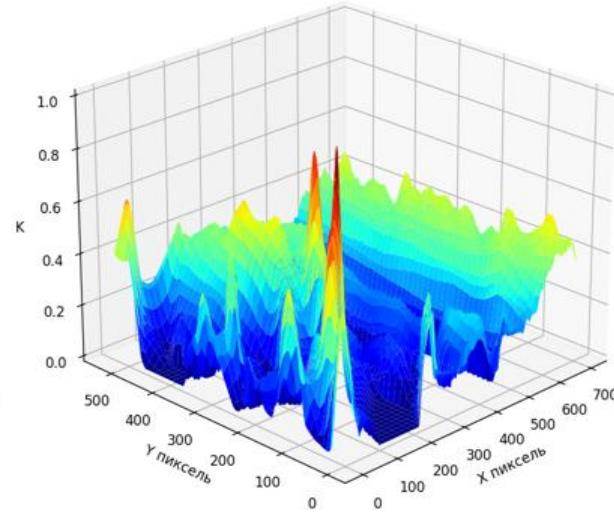
а) Эталонное изображение б) Текущее изображение

Рисунок 8 – Рассматриваемые ЭИ и ТИ

Сравнение результатов работы двух алгоритмов после обработки изображения представлено на рис. 9.



(а) МФД



(б) НКК

Рисунок 9 – Результаты работы предлагаемого алгоритма на основе семантических описаний (а) и НКК (б)

В табл. 3 представлены результаты оценки точности идентификации ТИ и ЭИ при помощи МФД и НКК на основе обработки 231 изображения.

Под точностью понимается средняя вероятность (частота) получения правильного решения задачи идентификации изображений.

Под сокращением вычисления подразумевается средний процент оставшихся возможных вариантов областей ЭИ, в которых возможно местонахождение ТИ.

Таблица 3 – Результаты обработки 231 изображения

Метод	Точность при пороге 0,8	Сокращение вычислений при пороге 0,8	Точность при пороге 0,9	Сокращение вычислений при пороге 0,9
МФД, объекты	0,672	0,0644%	0,824	0,0517%
МФД, атрибуты	0,838	0,0507%	0,874	0,0486%
МФД, отношения	0,794	0,0277%	0,846	0,0244%
МФД, иерархический алгоритм	0,848	0,0279%	0,848	0,0247%
НКК	0,5	0,0549%	0,5	0,0549%

Анализ результатов (таблица 3) показывает, что НКК дает единственное правильное решение в половине случаев, в то время как у МФД вероятность правильных решений существенно выше.

В четвертом эксперименте проводится оценка времени работы алгоритмов сопоставления сцен на основе семантических описаний и НКК.

Иерархический алгоритм при использовании в совокупности всех этапов объединяет преимущества сокращения возможных вариантов и ускорение. Результаты работы алгоритма представлены в табл. 4.

Таблица 4 – Время работы при использовании иерархического алгоритма в сравнении с временем работы корреляционной функции

Размер ЭИ и ТИ	ЭИ: 320×320 ТИ: 80×80	ЭИ: 320×320 ТИ: 160×160	ЭИ: 320×320 ТИ: 160×160	ЭИ: 480×480 ТИ: 240×240
МФД, объекты	0,753225	0,386727	0,132852	0,083849
МФД, атрибуты	0,177536	0,030938	0,010628	0,006707
МФД, отношения	0,035507	0,006817	0,00312	0,001521
МФД, иерархический алгоритм	0,966268	0,424482	0,1466	0,092077
НКК	23,711182	45,711301	22,559491	26,309172

Анализируя табличные данные, заметим, что среднее время работы при использовании иерархического алгоритма сократилось более чем в 85 раз в сравнении с НКК.

В пятом эксперименте рассматривается полет в условиях облачности с несколькими просветами. Целью исследования является подтверждение возможности оценки информативности признаков при семантическом описании сцен.

На рис. 10 показаны: а) Просветы в облачности; б) Сегментированное ЭИ.

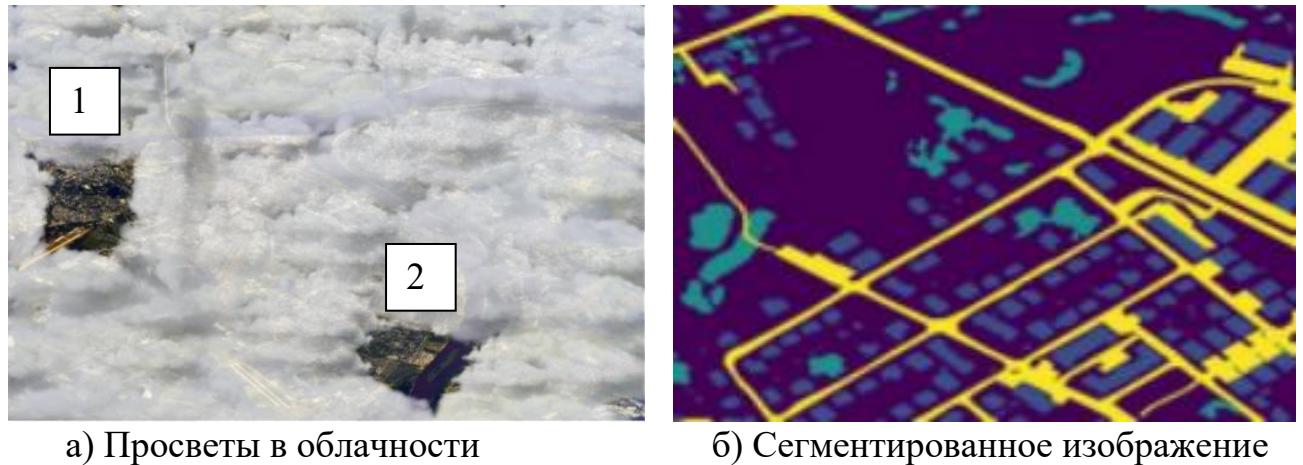


Рисунок 10 – Область полета БЛА

Маршрут 1 соответствует малоинформационному описанию вида: <здание, здание, здание, фон>, а маршрут 2 представляет описание вида <здание, здание, здание, здание, дорога, дорога, фон, фон>. Усредненные результаты моделирования, рассчитанные для разных положений просветов на ЭИ, приведены в табл. 5. Оценка энтропии и информативности представлена в битах.

Таблица 5 – Оценка информативности маршрутов

Параметр	Маршрут 1	Маршрут 2
$H(m)$	3,184	3,184
$H(m U_n)$	2,644	1,6
I_n	0,54	1,584

Таким образом, в процессе выполнения полетного задания возможно осуществление динамической оценки информативности разных вариантов маршрута с целью повышения вероятности правильной оценки координат БЛА по сравнению со случайнym выбором маршрута.

ЗАКЛЮЧЕНИЕ

Суммируя все вышесказанное, перечислим основные результаты:

1. Показано, что для повышения производительности и помехоустойчивости алгоритмов визуальной навигации необходимо перейти от растровых описаний к семантическим.

2. Разработана методика формирования семантических описаний наблюдаемых сцен на ЭИ и ТИ.

3. Разработан алгоритм поиска фрагмента ЭИ, идентичного принятому ТИ, на основе сравнения семантических описаний наблюдаемых сцен.

4. Предложен и разработан алгоритм выбора информативного маршрута полета для повышения точности определения координат БЛА.

5. Результаты исследований помехоустойчивости разработанного алгоритма показали, что алгоритм обеспечивает устойчивое формирование семантических описаний при воздействии гауссовых шумов с СКО, соответствующих $\beta \geq 5$, и изменении МО и СКО пикселей изображения в пределах $\pm 50\%$, а также при воздействии гамма-коррекции в пределах от $\gamma = 0,5$ до $\gamma = 1,5$.

6. Проведено сравнение вычислительной производительности разработанного алгоритма с корреляционным алгоритмом. Результаты показали, что предложенный алгоритм позволяет сократить объем требуемых вычислений в среднем более чем в 10 раз.

7. Представленные результаты модельных экспериментов подтвердили работоспособность предложенного алгоритма навигации по семантическому описанию сцен.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях из рекомендованного ВАК Минобрнауки России перечня

1. Бодунков Н.Е. Визуальная навигация беспилотных летательных аппаратов с использованием семантических описаний местности / Н.Е. Бодунков, Д.С. Гиренко, Н.В. Удалова, Н.В. Ким, Н.А. Ляпин // Известия ЮФУ. Технические науки. – Таганрог, 2025. – № 2 (244). – С. 256-268
2. Гиренко Д.С. Оценка эффективности идентификации изображений в визуальной навигации БЛА на основе использования их семантических описаний / Д.С. Гиренко // Навигация и управление летательными аппаратами. – М., 2025. – № 3(50). – С. 47-63.

3. Гиренко Д.С. Повышение быстродействия алгоритмов визуальной навигации за счет использования семантического описания сцен / Д.С. Гиренко // Научно-технический вестник Поволжья. – Казань: ООО «Рашин Сайнс», 2025. – № 9 – С. 41-44.

Свидетельства о государственной регистрации программ для ЭВМ

4. Бодунков Н.Е. Программа оценки информативности направления полета беспилотного летательного аппарата с точки зрения визуальной навигации / Н.Е. Бодунков, Д.С. Гиренко, Н.В. Ким // Свидетельство о государственной регистрации программы для ЭВМ № 2024691196. – 2024.

Патенты на изобретения

5. Скрибцов П.В., Тюляев Д.В., Гиренко Д.С., Шигина М.А., Шутова О.А., Сухоруков В.Ю. Способ и устройство эмуляции сигнала приемника спутниковой навигации с нейросетевой регуляризацией слияния данных // Патент на изобретение № 2846175 от 01.09.2025.

Отчеты о НИР

6. Проведение теоретических исследований и разработка алгоритмов планирования полета, визуальной навигации и адаптации параметров системы управления автономных БАС на основе элементов анализа ситуаций (этап № 1 промежуточный): отчет о НИР / МАИ; рук. Бодунков Н.Е. ; исполн.: Гиренко Д.С. [и др.]. – М., 2024. – 98 с. – Библиогр.: с. 212. № ГР FSFF-2024-0001.

Доклады конференций индексируемые в SCOPUS и WoS

7. Girenko D.S. Analysis of the situation on the observed scene containing many moving objects / D.S. Girenko, N.V. Kim // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – 2023. – Vol. XLVIII-2/W3-2023. – Pp. 65–71.

8. Girenko D.S. Improving the efficiency of visual navigation of an unmanned aerial vehicle by choosing the most informative route / D.S. Girenko, N.V. Kim // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – 2024. – Vol. XLVIII-2/W5-2024. – Pp. 67–72.

9. Girenko D.S. Improving the noise immunity of visual navigation algorithms based on the use of semantic descriptions of observed scenes / D.S. Girenko // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – 2025. – Vol. XLVIII-2/W9-2025 – Pp. 111–116.

Прочие публикации, включая тезисы докладов российских и международных конференций

10. Гиренко Д.С. Анализ ситуации на основе нейронной сети Transformer / Д.С. Гиренко // Сб. тезисов 49-й международной молодежной науч. конф. «Гагаринские чтения». – М. : Перо, 2023. – С. 508-509.

11. Гиренко Д.С. Оценка поведения людей на основе анализа видеинформации / Д.С. Гиренко, Н.В. Ким // Сб. докладов с международной конф. по науке и технологиям «Россия – Корея – СНГ». – М., 2023. – С. 73-77.

12. Гиренко Д.С. Визуальная навигация беспилотного летательного аппарата в условиях облачности / Д.С. Гиренко, Н.В. Ким // Сб. докладов с международной конф. по науке и технологиям «Россия – Корея – СНГ». – М., 2024. – С. 135-139.

13. Гиренко Д.С. Визуальная навигация по семантическому описанию сцены // Сб. тезисов 51-й международной молодежной науч. конф. «Гагаринские чтения». – М. : Перо, 2023. – С. 474-475.