

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Московский авиационный институт
(Национальный исследовательский университет)»
(МАИ)

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
(технические науки)

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата технических наук

Научный руководитель:
кандидат технических наук, профессор
Ким Николай Владимирович

Москва, 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
Глава 1. ВИЗУАЛЬНАЯ НАВИГАЦИЯ БЛА	11
1.1. Автономный полет БЛА	11
1.2. Визуальная навигация автономных БЛА	15
1.3. Информативность признаков	24
1.4. Оценка выполнения требований к решению целевой задачи	31
1.5. Постановка задачи	35
Глава 2. ФОРМИРОВАНИЕ СЕМАНТИЧЕСКОГО ОПИСАНИЯ НАБЛЮДАЕМЫХ СЦЕН	37
2.1. Семантические описания наблюдаемых сцен	37
2.2. Выбор метода сегментации изображений	47
2.3. Выделение классов объектов на основе семантической сегментации	55
2.4. Выделение атрибутов и отношений между объектами	60
Глава 3. ИДЕНТИФИКАЦИЯ СЕМАНТИЧЕСКИХ ОПИСАНИЙ	68
3.1. Оценка положения текущего изображения на эталонном изображении	68
3.2. Формирование словаря признаков семантических описаний наблюдаемых сцен	72
3.3. Оценка сходства семантических описаний наблюдаемых сцен	77
3.4. Иерархический алгоритм поиска	82
Глава 4. ИССЛЕДОВАНИЕ РАБОТОСПОСОБНОСТИ АЛГОРИТМОВ ВИЗУАЛЬНОЙ НАВИГАЦИИ БЛА С ИСПОЛЬЗОВАНИЕМ СЕМАНТИЧЕСКИХ ОПИСАНИЙ ИЗОБРАЖЕНИЙ	87
4.1. Оценка помехоустойчивости алгоритма семантической сегментации изображений в изменяемых условиях наблюдений	89
4.2. Оценка эффективности идентификации изображений на основе использования их семантических описаний	98
4.3. Оценка точности корреляционных алгоритмов сопоставления исходных растровых изображений и алгоритмов, использующих семантические описания	105
4.4. Оценка времени работы алгоритмов сопоставления сцен на основе семантических описаний и корреляционных алгоритмов	113
4.5. Оценка информативности признаков при семантическом описании наблюдаемых сцен	117
ЗАКЛЮЧЕНИЕ	123
СПИСОК СОКРАЩЕНИЙ	124
СПИСОК ЛИТЕРАТУРЫ	125

ВВЕДЕНИЕ

Актуальность темы исследования. В настоящее время все большую роль в мире играют малые беспилотные летательные аппараты (БЛА), способные самостоятельно выполнять поставленные целевые задачи (ЦЗ). БЛА активно используются для повышения эффективности или автоматизации решения множества прикладных задач, а сфера их применения постоянно расширяется. Одним из наиболее перспективных направлений в их развитии является повышение автономности [31]. Глобальный сценарий применения любого БЛА можно свести к последовательности этапов, общих для любой полезной нагрузки и целевой задачи:

- 1) предполетная подготовка – подготовка аппаратных средств БЛА и полезной нагрузки, настройка автопилота, формирование полётного задания и пр.;
- 2) взлет;
- 3) полет в зону применения полезной нагрузки (автоматический или ручной режим) – может быть реализован как полет по маршруту, сформированному на этапе 1;
- 4) применение полезной нагрузки (например, аэрофотосъемка или доставка груза);
- 5) полет в точку посадки, как и этап 3. можно представить как полет по маршруту;
- 6) посадка (при необходимости).

Особенности реализации, критерии и ограничения на выполнение каждого из этих этапов определяются особенностями целевой задачи и условиями применения и должны учитываться оператором на первом этапе – при подготовке БЛА к полету [41]. Современные средства наземных станций управления (НСУ) позволяют успешно автоматизировать процесс подготовки (например, автоматически строить маршрут) для определенных заранее условий. В свою

очередь, существующие и широко применяемые системы управления БЛА позволяют в штатном режиме решать задачи этапов 2–6 автоматически, без непосредственного вмешательства оператора. Так, например, задачи автоматического взлета и посадки, стабилизации, удержания скорости и высоты, полета по маршруту (с использованием инерциальных, спутниковых и иных навигационных систем) и управления полезной нагрузкой реализованы в функциональном программном обеспечении многих современных автопилотов (как открытых, так и проприетарных) [14, 15].

Однако априорная неопределенность условий полета (или их изменяемость) на этапе подготовки, изменение обстановки во время полета или возникновение нештатных ситуаций может привести к ухудшению качества или даже невозможности решения ЦЗ. Например, изменение погодных условий (или параметров атмосферы – температуры, ветра и пр.) непосредственно влияет на динамику БЛА и может привести к необходимости изменения параметров системы управления (для обеспечения требуемого качества управления), а отказ штатной навигационной системы (например, спутниковой) – к необходимости переключения на альтернативные системы, например, на основе систем технического зрения.

Сложность управления БЛА в этих ситуациях определяется наличием большого количества разнообразных и труднопрогнозируемых факторов, которые необходимо учитывать при принятии решений по управлению.

Повышение эффективности применения БЛА (с точки зрения автономности планирования и решения ЦЗ) возможно за счет адаптации управлений БЛА к изменению условий полета (или за счет прогнозирования и учета этих изменений при планировании).

Адаптация может быть решена за счет использования элементов *технологии анализа ситуации*, которая (в отличие от известных реализованных технологий) позволяет:

- анализировать и оценивать текущие ситуации с учетом требований к выполнению целевой задачи, состоянию среды и БЛА;

– прогнозировать и оценивать последствия возможных внешних и внутренних воздействий, а также соответствующих управлений БЛА на решение целевой задачи.

Под **анализом ситуации** (или ситуационной осведомленностью [70]) будем понимать технологию, позволяющую робототехническим системам **понимать текущую ситуацию и принимать решение в соответствии с ней** [66]. В работах [5, 21] анализ ситуации включает три основных уровня:

1. *Описание* (восприятие) текущей ситуации. Описание ситуации формируется на основе информации от внешних источников (датчиков, систем зрения и пр.), априорной информации (например, полученной на предыдущей итерации анализа ситуации), а также базы знаний. База знаний моделей описывает атрибуты, динамику (изменение собственных свойств с течением времени) и поведение (взаимодействие с другими моделями), объектов, участвующих в описании. Совокупность описаний и базы знаний формирует *модель мира*.

2. *Анализ* описанной на предыдущем этапе ситуации с целью принятия решения о дальнейших действиях. В некоторых работах [11] под анализом также понимают классификацию ситуации (из заранее сформированного конечного алфавита классов ситуаций) для формирования управления в соответствии с ее классом (для каждого класса ситуаций заранее выбирается определенная стратегия управления).

3. *Прогноз* – оценка состояния ситуации через некоторое время и/или в результате выполнения действий, выбранных на уровне анализа.

Пример системы визуального анализа ситуаций представлен в работе [47]. Очевидно, что показанное выше описание анализа ситуации слишком абстрактно и требует уточнения и адаптации для решения конкретных прикладных задач.

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 (инерционная навигационная система - ИНС) пространственной ориентации [25, 57].

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.

Степень проработанности темы исследования. Исследованию решения задачи визуальной навигации по семантическому описанию сцен посвящены работы зарубежных и отечественных авторов Junxuan Li, Yuanfang Zhang, Бодунков Н. Е., Визильтер Ю. В., Вишняков Б. В., Горбацевич В. С., Желтов С. Ю., Папулин С. Ю., Ким Н. В.

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. Формализация и постановка задачи визуальной навигации по семантическому описанию сцены с целью повышения помехоустойчивости и быстродействия.

2. Впервые предложена методика формирования и кодирования семантических описаний наблюдаемых сцен, позволяющая впоследствии осуществлять идентификацию.

3. Разработан ранее не существовавший алгоритм визуальной навигации на основе сравнения сходства семантических описаний наблюдаемых сцен на текущем и эталонном изображениях.

4. Впервые разработан иерархический алгоритм сопоставления сцен на основе анализа и сравнения семантических описаний текущих и эталонного изображений.

5. Предложен подход к повышению точности решения задачи визуальной навигации при полете над малоинформативной местностью за счет оценки информативности семантических описаний подстилающей поверхности планируемых вариантов маршрута.

6. 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 и кодирования 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.

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

Показана необходимость разработки методик и алгоритмов, обеспечивающих повышение производительности и помехоустойчивости алгоритмов визуальной навигации. В качестве решения предложено перейти к смысловому описанию сцен взамен растрового.

Глава 2 посвящена проблеме формирования семантических описаний наблюдаемых сцен.

Под семантическим описанием понимается смысловая нагрузка, заключенная в объектах, описаниях, отношениях, процессах, присутствующих на изображениях. В главе предлагается методика формирования семантических описаний наблюдаемых сцен.

Показано, что алфавит классов объектов может быть определен на этапе предполетной подготовки на основе имеющихся ЭИ, а рабочий словарь признаков необходимо адаптировать в зависимости от ТИ.

Задача выделения объектов на изображении в настоящее время успешно решается при помощи сверточных нейронных сетей. Рассмотрены несколько популярных архитектур, метрика качества и произведен выбор наилучшей для имеющегося набора данных.

Выделение атрибутов объектов и отношений между ними представляет собой обширный список возможных вариантов реализации и зависит от информативности того или иного признака при решении конкретной ЦЗ на конкретном маршруте. В главе представлены несколько алгоритмов для выделения атрибутов объектов и отношений между ними.

В главе 3 рассматривается задача идентификации ТИ и соответствующего ему фрагмента ЭИ на основе их семантических описаний. Проблема оценки сходства ТИ и ЭИ связана с разной физической природой сущностей, которые описываются при помощи семантики.

В главе предлагается в качестве оценки соответствия воспользоваться парными функциями. Для непосредственного сравнения семантические описания кодируются и сравниваются в векторном виде. При полном соответствии двух семантических описаний оценка сходства между ними равна 1, при полном отсутствии одинаковых описаний на изображениях – 0.

Также предложен иерархический алгоритм идентификации семантических описаний, позволяющий увеличить производительность за счет использования сокращенного рабочего словаря. Рассмотрен вопрос о процедуре принятия решения об окончании поиска ТИ на ЭИ.

В главе 4 приведены результаты моделирования работы предлагаемых подходов. Показана работоспособность и эффективность предлагаемых методик и алгоритмов формирования семантического описания наблюдаемой сцены. Проведены исследования по анализу помехозащищенности сверточных нейронных сетей и семантических описаний при различных изменениях условий наблюдения. Показано, что точность работы алгоритмов на основе семантического описания сопоставима алгоритмам на основе взаимной корреляционной функции. Проведена оценка производительности предлагаемого подхода в сравнении с корреляционной функцией.

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

Публикации. По результатам диссертационной работы опубликовано 13 работ, в том числе 1 Свидетельство о государственной регистрации программ для ЭВМ [28], 1 Патент на изобретение [59], 1 отчет о НИР [54], 3 научно-технические статьи в изданиях, включенных в перечень ВАК [27, 36, 37], 3 доклада конференций, включенные в перечень SCOPUS и Web of Science [7-9] и 4 прочие публикации [32-35].

Апробация результатов 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, списка сокращений и списка литературы. Работа изложена на 134 страницах машинописного текста, содержит 51 рисунок, 9 таблиц и 70 наименование литературных источников.

Глава 1. ВИЗУАЛЬНАЯ НАВИГАЦИЯ БЛА

В данной главе рассматриваются особенности визуальной навигации при решении целевых задач автономными БЛА. Проведен анализ подходов к определению местоположения БЛА методами визуальной навигации.

Показано, что использование семантических описаний наблюдаемых сцен в визуальной навигации может повысить устойчивость и вычислительную производительность алгоритмов визуальной навигации.

1.1. Автономный полет БЛА

Целевые задачи, решаемые БЛА, например, мониторинг местности или доставка грузов, определяются уровнем автономности БЛА [60].

Важным вопросом реализации соответствующих уровней автономности является вопрос построения НС БЛА [29].

Базовой основой НС БЛА является ИНС. При этом наиболее эффективными системами коррекции ИНС являются ГНСС. Однако в случаях отсутствия сигналов ГНСС для коррекции НС требуется использование методов визуальной навигации [6, 51].

Рассмотрим особенности реализации методов визуальной навигации на БЛА разных уровней автономности.

БЛА нулевого, первого и второго уровней в основном управляются оператором, который и определяет характер решаемых целевых задач.

На 3-м уровне выполнение БЛА типовых целевых задач реализуется автоматически в определенных условиях, но постоянно контролируется оператором.

Использование методов визуальной навигации на вышеперечисленных уровнях автономности БЛА является нецелесообразным.

На 4-м уровне автономности БЛА решает самостоятельно определенный набор целевых задач. Необходимость использования методов визуальной навигации на БЛА 4-го уровня автономности определяется требуемыми условиями решения типовых целевых задач. При необходимости полетов в области отсутствия сигналов ГНСС и если накапливаемые ошибки бортовых ИНС превышают допустимые, необходимо оснащать БЛА соответствующими программно-аппаратными средствами визуальной навигации.

Для оценки необходимости использования методов визуальной навигации для конкретных типов ИНС рассчитываются значения возможных ошибок определения координат.

В работе [50] показано, что линейная ошибка представляет собою квадратичную зависимость от времени:

$$\sigma(t) = at^2, \quad (1.1)$$

где a – коэффициент, определяемый для конкретных типов датчиков ИНС, t – время работы ИНС без коррекции.

В [13] авторы исследуют возможность моделирования среднеквадратической ошибки σ координат и скоростей датчиков IMU 500. Используя зависимости (1.1) и экспериментальные данные, предоставленные в работе [13], получим следующую зависимость:

$$\sigma_{IMU_{500}}(t) = 0,00013 t^2. \quad (1.2)$$

Будем считать, что вероятность положения БЛА над некоторым участком поверхности может быть рассчитана, если известны расчетные ошибки ИНС и время t полета БЛА без коррекции координат.

Например, для IMU 500: при полете без коррекции $t=1800$ с, линейные ошибки среднеквадратического отклонения (СКО) будут – $\sigma_{IMU_{500}}(t) = 421,6$ м, а при $t=3600$ с – $\sigma_{IMU_{500}}(t) = 1684,8$ м.

Следовательно, на данном уровне автономности возможно использование методов визуальной навигации, если существует вероятность полета (в рамках решаемых целевых задач) в областях с отсутствием сигналов ГНСС.

Пусть, например, линейный размер ориентира – 10 м. Линейная ширина поля зрения системы наблюдения (СН) БЛА (вертикально вниз) – 200 м.

На рисунке 1.1 показана максимальная допустимая ошибка положения БЛА, при которой гарантируется наблюдение искомого ориентира.

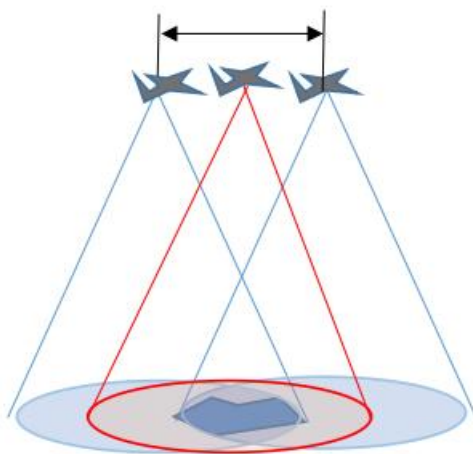


Рисунок 1.1 – Гарантированное наблюдение ориентира

В случае точного полета БЛА наземный ориентир полностью попадает в поле зрения (красная окружность) СН БЛА. При отклонении траектории полета от заданной изображение ориентира будет смещаться к границам поля зрения.

Для гарантированного наблюдения ориентира необходимо соблюдение условия $L \leq 6\sigma$.

Откуда $L \geq 6\sigma = 6(at^2)$ и время допустимого полета БЛА без коррекции координат:

$$t \leq \sqrt{\frac{L}{6a}}. \quad (1.3)$$

Для рассматриваемого примера и используемого на борту IMU 500 допустимое время без коррекции (для наблюдения ориентира): $t = \sqrt{\frac{390}{6 \cdot 0,00013}} = 707$ с.

Таким образом, визуальная навигация для 4-го уровня автономности должна обеспечивать оценку координат БЛА при полете по заданным маршрутам в известных условиях наблюдения. В качестве ЭИ могут использоваться искусственные или естественные ориентиры, изображения которых достаточно информативны и наблюдаемы в рабочих условиях полета БЛА. Расстояния между ориентирами должно обеспечивать требуемую вероятность их обнаружения при полете по заданным маршрутам. При необходимости независимой оценки скорости используются методы визуальной одометрии [63].

Для БЛА 5-го уровня автономности использование технологий визуальной навигации является необходимым, т.к. этот уровень автономности предназначен для решения различных целевых задач при любых условиях. При этом БЛА 5-го уровня автономности должен самостоятельно выбирать маршрут полета с учетом условий полета (погодных, суточных), а также наблюдаемости и информативности визуальных ориентиров.

К бортовым программно-аппаратным средствам визуальной навигации БЛА 5-го уровня автономности должны предъявляться наиболее сложные требования. В частности, в состав показателей качества планируемых маршрутов должны включаться:

- 1) наблюдаемость и устойчивость ориентиров, оцениваемые вероятностью их правильного обнаружения;
- 2) информативность ориентиров, определяющая точность привязки БЛА к цифровой карте местности (ЦКМ);
- 3) производительность программно-аппаратных средств.

1.2. Визуальная навигация автономных БЛА

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. Оценка сходства описаний ТИ и фрагментов ЭИ реализуется на основе использования различных метрик, например, корреляционных функций.

Вне зависимости от вида описаний наблюдаемых сцен (растровых, векторных, семантических) структура алгоритмов сопоставления и идентификации ТИ и ЭИ представляется в виде, показанном на рисунке 1.2.

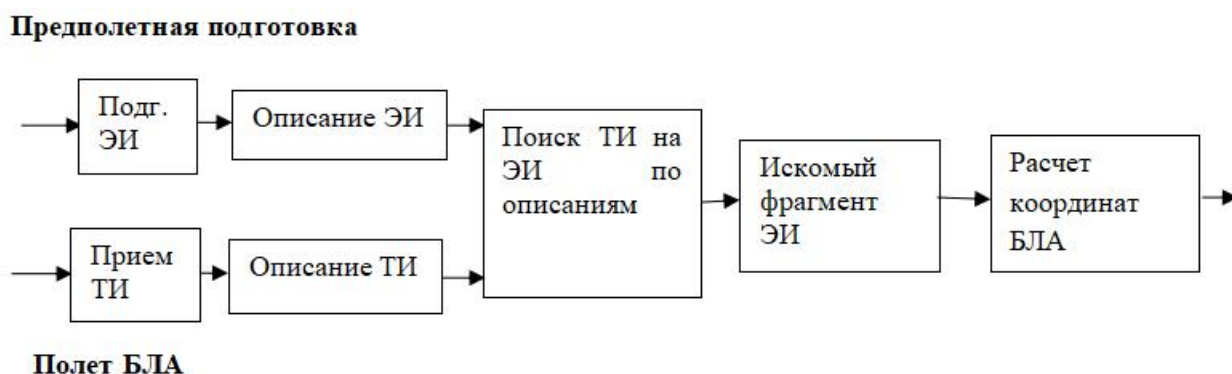


Рисунок 1.2 – Общая схема визуальной навигации БЛА

ЦКМ или ЭИ формируется на этапе предполетной подготовки и отображает всю предполагаемую область интереса, в которой БЛА должен решать заданные целевые задачи.

ТИ принимается в процессе полета БЛА бортовыми средствами наблюдения, например, видеокамерой. Размеры ТИ определяются полем зрения бортовой СН, направлением линии визирования и высотой полета БЛА.

Сравнение описаний ТИ и разных фрагментов ЭИ позволяет выбрать наиболее схожий фрагмент. Так как координаты любого фрагмента ЭИ известны, то могут быть определены и координаты БЛА.

Определим основные требования, предъявляемые к решению задачи визуальной навигации.

Будем считать, что рассматриваемые БЛА оснащены датчиками магнитного курса и высотомерами. Линия визирования СН стабилизирована и направлена вертикально вниз.

Под точностью определения местоположения БЛА в данной работе подразумевается точность определения географических координат в земной (местной) системе координат XOY . Примем, что условие решения навигационной задачи имеет следующий вид:

$$\begin{aligned} 3\sigma_x = \Delta x &\leq |\Delta x_0|, \\ 3\sigma_y = \Delta y &\leq |\Delta y_0|, \end{aligned} \quad (1.4)$$

где $\pm\Delta x_0, \pm\Delta y_0$ – допустимые ошибки оценки координат;

σ_x и σ_y – СКО оценки координат.

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 распознаваемых классов – все возможные фрагменты ЭИ. При этом сравниваемые растровые ТИ и ЭИ должны быть представлены в одном масштабе.

Так как оценка положения БЛА определяется на основе оценки сходства ТИ и ЭИ, то необходимо учитывать вероятность правильного распознавания изображений. Обозначая последовательно (слева направо и сверху вниз) от 1 до $M_{\Sigma 0}$ пиксели на ЭИ, получим:

$$P(m) \geq P_0, \quad (1.5)$$

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

Будем считать, что вероятность правильного распознавания m -го фрагмента ЭИ (1,5) эквивалентна вероятности присутствия БЛА в области m .

Требуемые значения параметров Δx_0 , Δy_0 и P_0 определяются решаемой целевой задачей.

Общий алфавит классов содержит следующее количество фрагментов ЭИ (соседние фрагменты смещены на один пиксель):

$$M_{\Sigma 0} = (M_{RI} - M_{CI} + 1)(N_{RI} - N_{CI} + 1), \quad (1.6)$$

где M_{RI} , N_{RI} – размеры ЭИ (в пикселях по вертикали и горизонтали); M_{CI} , N_{CI} – размеры ТИ.

Размеры элементарной области m равны:

$$\begin{aligned} \Delta x_m &= \frac{X_{RI}}{M_{RI}}, \\ \Delta y_m &= \frac{Y_{RI}}{N_{RI}}, \end{aligned} \quad (1.7)$$

где X_{RI} , Y_{RI} – размеры ЭИ.

Для метода **полного перебора** (сплошного поиска) требуемая вычислительная производительность процесса (количество элементарных операций в единицу времени) при частоте принимаемых кадров $f = 1/\tau$ равна:

$$A = \frac{(M_{RI} - M_{CI} + 1)(N_{RI} - N_{CI} + 1)K_0}{\Delta T}, \quad (1.8)$$

где $\Delta T = 1/f$ – время между принимаемыми кадрами видеопоследовательности;
 K_0 – количество элементарных вычислительных операций при расчете выбранного показателя сходства изображений.

Сравнение текущего и эталонного изображений.

В общем случае в алгоритмах визуальной навигации, основанных на сравнении ТИ и ЭИ, могут быть использованы разные варианты описаний наблюдаемых сцен, присутствующих на получаемых растровых изображениях.

Корреляционные алгоритмы. Наиболее распространенными в визуальной навигации являются корреляционно-экстремальные алгоритмы [48], основанные на непосредственном использовании растровых изображений, их модификации [40] и аналоги [62].

При сканировании растрового изображения (ТИ или ЭИ) большего размера окном меньшего размера, соответственно ЭИ или ТИ, вычисляется корреляционная функция, в частности нормированный коэффициент корреляции:

$$K(di, dj) = \frac{1}{M_{CI} N_{CI}} \frac{\sum_{i=1}^{M_{CI}} \sum_{j=1}^{N_{CI}} \{CI(i, j) - M[CI]\} \{RI(i + di, j + dj) - M[RI]\}}{\sigma[RI] \sigma[CI]}, \quad (1.9)$$

где $RI(i + di, j + dj)$ – эталонное изображение;

$CI(i, j)$ – текущее изображение;

i, j – координаты пикселей на ЭИ;

di, dj – смещения фрагмента ЭИ;

$M[.]$ – математическое ожидание (МО);

$\sigma[.]$ – СКО;

M_{CI}, N_{CI} – размеры текущего изображения.

В некоторых случаях используются разностные функции, например, функция среднего модуля разности и др. Далее находится экстремум, который соответствует максимальному совмещению изображений.

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

К недостаткам корреляционных алгоритмов можно отнести чувствительности ТИ к условиям наблюдения и, в ряде случаев, низкую вычислительную производительность.

Алгоритмы распознавания. Поиск искомого фрагмента ЭИ может быть реализован с помощью известных алгоритмов распознавания, в частности вероятностных [39].

В данном подходе описание наблюдаемых сцен основано на признаках, которые должны быть выделены путем дополнительной обработки исходных изображений [68].

Процесс распознавания (в частном случае – обнаружения) в рамках решаемых системами технического зрения (СТЗ) задач визуальной навигации – это **процесс принятия решения** о принадлежности распознаваемого объекта (ТИ) к конкретному классу объектов (фрагментов ЭИ) или классификации явлений по имеющейся информации и отнесение выделенной совокупности значений признаков к области, характеризующей одно из состояний источников информации.

В общем случае при принятии решений необходимо рассматривать набор или вектор признаков U , образующих пространство признаков. Это N -мерное пространство разделяется на области, каждая из которых соответствует значениям признаков конкретного класса объектов.

Признаки являются основой сообщений, несущих полезную информацию. Они могут быть простыми (непроизводными) или сложными.

К простым признакам относятся, например, такие характеристики сигналов, как амплитуда, фаза, интенсивность излучения отдельной ячейки цифрового изображения и др.

К сложным признакам, например, относятся: контур (силуэт) изображения или его элементы, значения критериальных функций, распределения яркости, используемые в корреляционно-экстремальных системах, и т.д.

В частности, к этой группе признаков относятся **признаки, являющиеся основой семантических описаний: классы объектов, их атрибуты, отношения между объектами.**

Вероятностные признаки рассматриваются как случайные величины и их конкретные значения определяются в результате проведения соответствующих оценок.

В зависимости от решаемых задач разные признаки могут содержать разные объемы полезной информации.

Выбор наиболее информативных и физически доступных признаков часто и составляет основную сложность при реализации процессов измерения, адаптивных к изменяемым условиям наблюдений.

Таким образом, при получении значений признаков наблюдаемого объекта необходимо определить, к какой области они относятся, а затем принимается решение о принадлежности наблюдаемого объекта к определенному классу.

Одной из задач распознавания является построение решающих границ – разделяющих функций $F_i(u)$, $i=1...N$, где N количество распознаваемых классов, обладающих следующими свойствами: если объект O_1 обладает значениями признаков $\bar{u}_1 = (u_1^1, u_2^1, \dots, u_m^1)$ и относится к классу Ω_1 , то величина функции $F_1(\bar{u}_1)$ должна достигать максимального значения. Она должна быть максимальной и для всех других значений признаков, относящихся к классу Ω_1 . Если через $\bar{u}_q$ обозначить вектор признаков объекта, относящегося к Ω_q -му классу, то для все значений $\bar{u}_q$:

$$F_q(\bar{u}_q) > F_g(\bar{u}_q), \quad q, g = 1, \dots, m, \quad q \neq g. \quad (1.10)$$

Тогда уравнение границы между классами q, g будет следующим:

$$F_q(\bar{u}_q) - F_g(\bar{u}_q) = 0. \quad (1.11)$$

Еще одна из задач распознавания – определение методики сравнения мер близости или сходства значений признаков объектов при сопоставлении их с эталонными представителями заданных классов.

Существует множество различных метрик, одной из которых, например, является среднеквадратичное расстояние между совокупностью из k_q объектов, представляющих собой класс Ω_q :

$$L(O, \Omega_q) = \sqrt{\frac{1}{k_q} \sum_{k=1}^{k_q} d^2(O, O_k)}, \quad (1.12)$$

причем метод вычисления расстояния между объектами $d(O, O_k)$ свободен для выбора.

Наиболее общим методом решения задач обнаружения и распознавания является метод вероятностной оценки.

В постановке задачи обнаружения для метода вероятностной оценки считаются известными априорные вероятности появления объектов $P(O_1), O_1 \in \Omega_1$ и $P(O_2), O_2 \in \Omega_2$, а также условные плотности распределения значений признака $u \in U, p(u|O_1), p(u|O_2)$.

Требуется определить, к какому классу из двух принадлежит объект, если было получено значение признака u_0 .

В зависимости от различной априорной информации существуют несколько методов определения решающей границы, разделяющей пространство признаков:

- критерий Байеса, при котором также известна матрица потерь от ошибок обнаружения и априорные вероятности появления объектов;
- критерий идеального наблюдателя Зигерта – Котельникова применяется в том случае, если отсутствует информация о возможных потерях;
- минимаксный критерий, в этом критерии отсутствует информация о вероятности появления объектов, тогда принимается решение минимизировать потери с учетом максимально неблагоприятного исхода.

- критерий Неймана – Пирсона предполагает отсутствие информации как о вероятностях появления объектов, так и о матрице потерь. В этом случае задается максимальное значение для ошибки первого рода, а далее подбирают такое значение порога, при котором минимизируется ошибка второго рода;
- критерий максимального правдоподобия Фишера, как и критерий Неймана – Пирсона, обладает минимальной априорной информацией.

Вычислительная производительность подобных алгоритмов в существенной степени определяется процессами выделения рабочих признаков на изображениях. Например, при использовании алгоритма Хафа для выделения прямых или детекторов ключевых точек и пр. [30].

Таким образом, сложность реализации вероятностных алгоритмов связана с необходимостью выделения и оценки значений признаков на наблюдаемых сценах в реальном времени, а также расчета априорных вероятностей и условных вероятностей $p(u|O_1)$, $p(u|O_2)$.

Алгоритмы, основанные на семантическом описании наблюдаемых сцен. 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 используемых нейросетей.

Данный факт позволяет перейти к алгоритмам визуальной навигации, основанным на использовании **семантических описаний наблюдаемых сцен**, и повысить производительность обработки текущих изображений.

Рассматриваемые алгоритмы подобны алгоритмам, использующим традиционные описания признаков. Однако здесь признаками являются элементы предварительно сформированных семантических описаний наблюдаемых сцен. Выделение признаков на семантических описаниях для сравнения сходства изображений существенно повышает вычислительную производительность алгоритмов.

Основная сложность реализации данного подхода связана с необходимостью формирования семантических описаний ЭИ и ТИ.

В семантическом описании наблюдаемых сцен признаками, которые могут быть использованы для сравнения сходства, являются классы объектов, их атрибуты и отношения между классами.

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, в частности направления и дальности.

С развитием вычислительной техники и алгоритмов глубокого обучения, архитектуры сверточных нейронных сетей (СНС) научились решать задачи обнаружения и распознавания различных объектов в режиме реального времени. Наиболее популярные и производительные архитектуры сверточных нейронных сетей:

- YOLO (you look only once), tiny-YOLO;
- Mask R-CNN, Fast R-CNN, Faster R-CNN;
- SSD300, SSD500.

Обученные СНС позволяют обнаруживать, распознавать и классифицировать искомые объекты на различных наблюдаемых сценах.

Преимущество нейросетевых решений состоит в том, что за счет достаточно большой обучающей выборки и обучения СНС на этапе предполетной подготовки можно добиться высокой устойчивости к изменению условий наблюдения, включая изменения ракурса и масштаба эталонных образов.

Недостатки включают в себя: необходимость больших обучающих выборок (несколько десятков тысяч изображений на один класс объектов) для стабильной работы, полученные результаты не дают оценок эффективности и достоверности, подвержены влиянию «атак».

Избыточная размерность признакового пространства приводит к непроизводительному росту трудоемкости вычислений.

Недостаточная размерность признакового пространства может привести к невозможности решения навигационной задачи.

1.3. Информативность признаков

Идентификация некоторого фрагмента ЭИ и ТИ является необходимым, но не достаточным условием успешной оценки координат БЛА.

Изображение принятого участка местности может не содержать ориентиров, обеспечивающих выполнение условий (1.4) и (1.5).

Подобные наблюдаемые сцены или отдельные ориентиры являются недостаточно информативными, т.е. не содержат объема полезной информации, достаточного для решения поставленной задачи.

Информативность признаков оценивается на основе методов теории информации [39]. Основой оценки является мера неопределенности – энтропия [19]. Ранее энтропия использовалась в работе [46] для поиска объектов.

Общее количество полезной информации при распознавании M объектов или исходов, которое может быть получено при измерении всех значений признака U_n , равно:

$$I_n = H_b(m) - H_f(m|U_n) = - \sum_{m=1}^M P(m) \log_2 P(m) + \sum_{l=1}^M P(l) \sum_{m=1}^M \sum_{k=1}^K P(U_{nk}|l) P(m|U_{nk}) \log_2 P(m|U_{nk}), \quad (1.13)$$

где $H_b(m)$ – начальная энтропия, $H_f(m|U_n)$ – конечная энтропия процесса;

$P(m)$, $P(l)$, $P(i)$ – априорные вероятности возможных исходов с индексами m , l , i соответственно;

$P(U_{nk}|l)$ – условная вероятность появления признака U_{nk} при наличии l -го исхода;

$P(m) \log_2 P(m)$ – частная энтропия исхода m ;

$P(m|U_{nk}) \log_2 P(m|U_{nk})$ – частная энтропия, апостериорной вероятности нахождения в точке m при условии получения признака U_{nk} .

Например, вариантом простого ориентира U_n может быть «дом», значения k – большой, средний и т.д.

В решаемых задачах визуальной навигации под альтернативными исходами будем понимать варианты сравниваемых пар ТИ и различных фрагментов ЭИ.

Следовательно, информативность ТИ с данным ориентиром «дом» показывает, насколько точно определяется положение ТИ на ЭИ. Чем выше информативность, тем точнее оценка положения ТИ.

В более общем случае под ориентиром может пониматься некоторая наблюдаемая сцена, содержащая различные объекты, позволяющая оценить положение БЛА.

Перед началом расчета информативности (1.13) необходимо оценить априорную (начальную) энтропию нахождения БЛА в точке (i, j) или элементарной области m (при последовательной нумерации элементов ЦКМ от 1 до M). Данная энтропия характеризует неопределенность положения БЛА до начала поиска n -го ориентира.

Например, для цифровой карты местности и нормального распределения энтропия равна

$$H(i,j)=H(m)=-\sum_{m=1}^M P(m)\log_2 P(m)=\log_2 \left(\frac{1}{2\pi e} \frac{\sigma_x \sigma_y}{\Delta_x \Delta_y} \right), \quad (1.14)$$

а для равномерного распределения:

$$H(i,j)=H(m)=-\sum_{m=1}^M P(m)\log_2 P(m)=\log_2 \left(\frac{XY}{\Delta_x \Delta_y} \right), \quad (1.15)$$

где X, Y – размеры области, в которой вероятность присутствия БЛА равна единице.

В общем случае информативность признаков в задачах распознавания, ориентиров или наблюдаемых сцен в задачах визуальной навигации и т.д. зависит

от типа признака (ориентира) и его уникальности, наблюдаемости признака, априорной информации о возможных исходах.

Тип признака. В частности, полезная информативность одиночного здания с известными координатами выше информативности прямолинейного участка дороги, т.к. для оценки координат вдоль дороги требуются дополнительные ориентиры. Будем считать, что ориентиры подразделяется (с точки зрения их полезной информативности) на три группы: информативные, малоинформативные и неинформативные ориентиры.

На рисунке 1.3 показаны подобные варианты наблюдаемых ТИ.

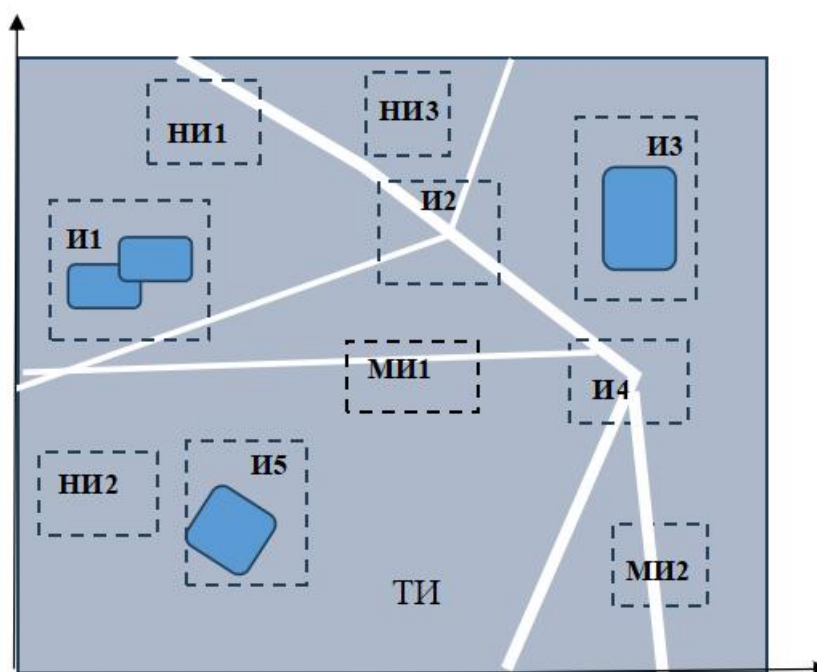


Рисунок 1.3 – Примеры ТИ, принимаемых СН БЛА

В качестве примера примем, что ошибки оценки координат Δx , Δy ограничены требованиями

$$\begin{aligned}\Delta x_{\text{тр}} &= 0,5 X_{\text{ТИ}}, \\ \Delta y_{\text{тр}} &= 0,5 Y_{\text{ТИ}},\end{aligned}\tag{1.16}$$

где $X_{\text{ТИ}}$, $Y_{\text{ТИ}}$ – линейные размеры участка местности, соответствующие ТИ.

Штриховыми прямоугольниками на рисунке 1.3 выделены примеры принимаемых ТИ, где И1, ..., И5 – информативные ТИ; МИ1, МИ2 – малоинформативные ТИ; НИ1, НИ2, НИ3 – неинформативные ТИ.

Информативные ТИ

Если ошибки оценки координат ТИ удовлетворяют заданным требованиям по точности:

$$\begin{aligned}\Delta x_0 &= \Delta x, \\ \Delta y_0 &= \Delta y,\end{aligned}\tag{1.17}$$

то будем считать подобные ориентиры **информативными** (достаточно информативными), т.к. они позволяют обеспечить заданные требования по точности оценки координат.

На рисунке 1.3 ТИ И1, ..., И5 содержат уникальные (для данного ЭИ) ориентиры (объекты), позволяющие с требуемой точностью определить координаты ТИ на ЭИ.

Малоинформативные ТИ. При выполнении условий

$$\begin{aligned}\Delta x_0 &\geq \Delta x, \\ \Delta y_0 &< \Delta y,\end{aligned}\tag{1.18}$$

Или

$$\begin{aligned}\Delta x_0 &< \Delta x, \\ \Delta y_0 &\geq \Delta y\end{aligned}\tag{1.19}$$

ориентиры – **малоинформативные** (недостаточно информативные).

Подобные ориентиры (МИ1, МИ2 на рисунке 1.3) позволяют определить положение ТИ по одной из координат, но не обеспечивают заданные требования полностью. Так, ТИ МИ1 содержит фрагмент уникального (для данного ЭИ) горизонтального участка дороги, позволяющего однозначно определить

положение ТИ по X . Положение по Y не может быть определено с требуемой точностью.

Если дорога (линия, граница) пересекает область A , в любой точке которой может находиться БЛА, то ошибки определения координат определяются точками пересечения границ области (рисунок 1.4, штриховыми прямоугольниками показаны варианты принимаемых ТИ).

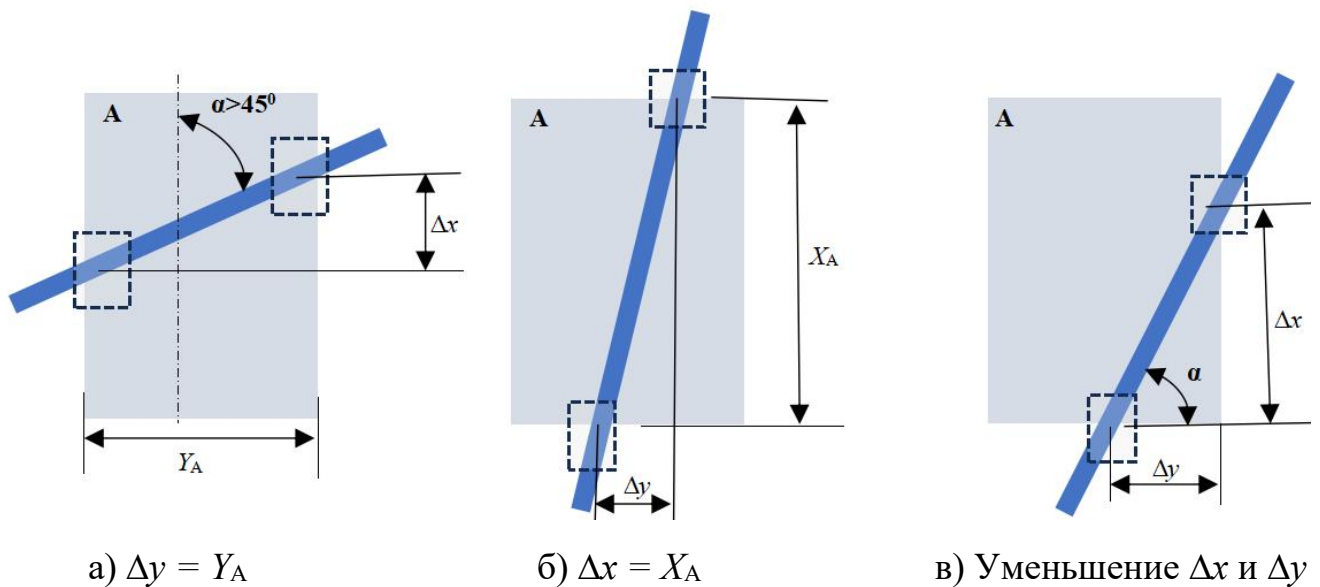


Рисунок 1.4 – Наблюдение линейных ориентиров

До начала измерений максимальные ошибки оценки координат БЛА определяются размерами X_A , Y_A области A .

На рисунке 1.4 а показан вариант, при котором ТИ содержит фрагмент прямолинейного участка дороги. Если дорога пересекает вертикальные стороны прямоугольника, а угол наклона дороги $\alpha > 45^\circ$, то ошибка Δx уменьшается, а $\Delta y = Y_A$ остается практически неизменной; на рисунке 1.4 б уменьшается ошибка Δy , на рисунке 1.4 в уменьшаются Δx и Δy .

Если линия пересекает вертикальные или горизонтальные стороны прямоугольника (Δx и Δy – известны), то выполняется следующее соотношение: $\Delta x = \Delta y \operatorname{tg}(\alpha)$.

При пересечении смежных сторон прямоугольника необходимо определить положение одной из точек пересечения.

В зависимости от требований и с учетом условий наблюдения линейные ориентиры могут быть информативными или малоинформативными.

Неинформативные ТИ

Ориентиры (НИ1, НИ2, НИ3 на рисунке 1.3) являются **неинформативными** при

$$\begin{aligned}\Delta x_0 &< \Delta x, \\ \Delta y_0 &< \Delta y,\end{aligned}\tag{1.20}$$

т.к. не позволяют оценить координаты ТИ с требуемой точностью.

Важной особенностью визуальной навигации является зависимость наблюдаемости визуальных ориентиров от условий наблюдения.

Наблюдаемость признака. Информативность признаков (1.13) в задачах распознавания или ориентиров в визуальной навигации в существенной степени зависит от их наблюдаемости – возможности их обнаружения или распознавания с учетом изменяемых условий наблюдения, в частности освещения. Так, например, признак «красная крыша дома», который днем определяет уникальность объекта, вечером или ночью теряет свою информативность.

Условная вероятность $P(U_{nk} | m)$ определяет связь между признаком U_{nk} и исходом m , оценивая вероятность получения (наблюдения, обнаружения, распознавания) признака U_{nk} при наличии m -го исхода из M возможных [39, 64].

В общем случае эта вероятность зависит от условий наблюдения (сезонных, суточных, погодных и пр.). Однако при решении многих практических задач распознавания данный факт не учитывается и считается, что $P(U_{nk} | m)$ стабильна.

Будем считать, что условия наблюдения дискретизируются и составляют (в диапазоне возможных условий функционирования БЛА) некоторый

ограниченный набор условий наблюдения $q \in Q$. Например, «Июль, 10-00 часов, солнечно, ...» или «Октябрь, 19-00 часов, ... сильный дождь» и т.д.

Тогда для учета конкретных условий условная вероятность будет иметь вид $P(U_{nk} | m, q)$.

По сути $P(U_{nk} | m, q)$ являются описаниями класса m на языке признаков U_{nk} в различных условиях наблюдения q .

Разнообразие условий применения БЛА требует подготовки и использования наборов ЭИ, адаптивных, информативных и пр. вариантов ЭИ.

Формирование описаний для каждого q реализуется разными методами: гистограммным, парзеновского окна, методом k -ближайших соседей и др. [67], в т.ч. экспериментально.

Сложность и высокая трудоемкость данных процессов ставит задачу перехода к более устойчивым описаниям наблюдаемых сцен.

В частности, семантические описания (при нейросетевом выделении объектов на наблюдаемых сценах), например: «дом около дороги» будут устойчивыми при разных внешних условиях.

Априорная информации о возможных исходах. В рамках решаемых задач визуальной навигации важной является оценка априорных вероятностей $P(m)$ присутствия БЛА в некоторой области M_0 .

Априорная вероятность в (1.13) определяет начальную энтропию (неопределенность) положения БЛА в некоторой области M_0 ЦКМ до начала визуальной навигации.

Априорная вероятность $P(m)$ нахождения БЛА в элементарной области m определяется, если известен соответствующий закон распределения вероятности.

Примем, что положение БЛА подчиняется нормальному закону, а СКО вычисляется по формуле (1.2) и для некоторого времени полета без коррекции может быть определено.

Плотность двумерного нормального распределения в системе координат XOY имеет вид

$$p(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}, \quad (1.21)$$

где μ_x, μ_y – МО по X и Y соответственно.

Если размеры и координаты элементарных областей m определены, то вероятность присутствия БЛА в области m равна

$$P(m)=\iint_m p(x,y) dx dy. \quad (1.22)$$

Расширение пределов интегрирования, в частности при объединении элементарных областей, позволяет оценивать вероятность присутствия БЛА в разных по размерам и положению областях ЖИ.

Необходимо отметить, что *информативность ориентиров, вычисляемая по (1.13), не является величиной постоянной, а зависит от ряда факторов, включая:*

- требования к точности полета на разных участках маршрута. Изменения $\Delta x_0, \Delta y_0$ могут переводить *информативные* ориентиры в *малоинформативные* и наоборот;
- бортовой аппаратуры, которая определяет накопление ошибок и точность работы СТЗ;
- накопленных ошибок перед измерениями по ориентиру;
- условиями наблюдения, в зависимости от которых могут изменяться вероятности и точности привязки СТЗ к различным ориентирам.

1.4. Оценка выполнения требований к решению целевой задачи

В рамках проводимых исследований целевой задачей БЛА является полет по маршруту при отсутствии сигналов СНС и выполнение требований (1.4), (1.5) в конечном пункте маршрута (КПМ).

Формула (1.13) позволяет оценить информативность отдельных ориентиров без учета их расположения на маршруте БЛА, в частности, при полете над местностью с переменной информативностью.

Недостатком оценки информативности ориентиров является то, что оценка по (1.13) может показывать равные значения информативности ориентиров, использование которых приводит к реализации разных по точности маршрутов БЛА.

Рассмотрим следующий пример. Наблюдаемая поверхность подобной местности может в существенной степени состоять из участков, на которых отсутствуют визуальные ориентиры, обеспечивающие возможность привязки БЛА. При этом присутствующие ориентиры могут иметь полезную информативность, зависящую от конкретных условий наблюдения.

На рисунке 1.5 показан пример полета БЛА через точки А (начальный пункт маршрута – НПМ) и Б (конечный пункт маршрута – КПМ) в области N без использования ГНСС.

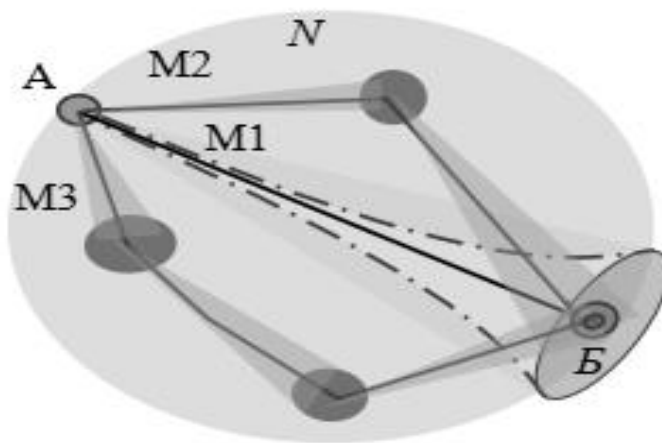


Рисунок 1.5 – Накопление ошибок ИНС в процессе полета БЛА

Рассматривается три варианта маршрутов.

Маршрут М1 – проходит из т. А в т. Б по прямой без коррекции методами визуальной навигации. Время полета t_1 без коррекции максимально. СКО оценки

координат в конце полета σ_1 , рассчитываются по (1.1) и характеризуются эллипсом или прямоугольником накопленных ошибок около КПМ (т. Б).

М2 – проходит через один участок, на котором проводится достаточно точная оценка координат по визуальным ориентирам. На следующем этапе полета без коррекции координат происходит накопление ошибок, характеризуемое треугольником от участка коррекции до КПМ. Время полета без коррекции t_2 , СКО в конце полета – σ_2 .

М3 проложен через два участка, на которых проводится визуальная достаточно точная оценка координат БЛА. Последний участок полета (t_3) без коррекции является самым коротким и σ_3 оценки координат минимальна.

Таким образом, $t_1 > t_2 > t_3$ и СКО в конце полета БЛА $\sigma_1 > \sigma_2 > \sigma_3$.

Полет по маршруту М3 обеспечивает максимальную точность оценки координат БЛА в КПМ, т.к. на этом маршруте последний участок полета без коррекции является самым коротким.

Будем считать, что ориентиры являются информативными, т.е. апостериорная (конечная) энтропия для всех ориентиров равна нулю.

Тогда информативность ориентиров будет зависеть от априорной энтропии, рассчитываемой по формулам (1.14) и (1.15).

Так как априорная энтропия последнего ориентира на М3 меньше априорной энтропии ориентира на М2, то ориентир на М2 оказывается более информативным. Однако полет по маршруту М2 в КПМ менее точен при полете по М3.

Кроме того, при полетах над малоинформативной местностью возможны ситуации, когда накапливаемые ошибки не позволяют реализовать гарантированный выход БЛА на искомый визуальный ориентир.

Будем считать, что $m \in M_0$ – область ЦКМ, в которой может находиться БЛА с вероятностью, близкой к единице.

Обозначим $n \in N_0$ подобласть из M_0 , в которой из-за накапливаемых ошибок ИНС и отклонения траектории полета от заданной в области СН БЛА не может принять изображение искомого ориентира.

Апостериорная энтропия в n , где U_{nk} не может быть принят, будет равна априорной энтропии $P(m)$ в соответствующих точках m :

$$H_{N_0}(n) = - \sum_{n=1}^{N_0} P(n) \log_2 P(n). \quad (1.23)$$

Следовательно, этот остаток энтропии $H_{N_0}(n)$ будет возрастать и добавляться к энтропии каждого очередного ориентира.

Таким образом, оценка информативности ориентиров (1.13) не позволяет учитывать влияние на точность попадания БЛА в КПМ и вероятность успешного выполнения целевой задачи.

Рассмотрим полет БЛА над малоинформативной местностью при отсутствии сигналов СНС.

Будем считать, что имеется ЦКМ и определены положение и размеры информативных областей (ориентиров). Альтернативные маршруты БЛА (Tr) разделены на этапы $St(Tr)$, связанные с последовательной привязкой к этим ориентирам.

Вероятность успешного выполнения каждого этапа $P(St^{Tr}) = 1$, если вероятность обнаружения СН БЛА искомого ориентира равна единице.

На конечном этапе полета St_f^{Tr} БЛА (к КПМ) $P(St_f^{Tr}) = 1$, если выполнены требования по обеспечению точности (1.4).

Общая вероятность выполнения целевой задачи при полете по маршруту Tr равна

$$P(Tr) = \prod_{St^{Tr}=1}^{St_f^{Tr}} P(St^{Tr}). \quad (1.24)$$

Таким образом, оценка (1.13) ориентиров (наблюдаемых сцен) позволяет выбирать наиболее информативный ориентир на каждом этапе, а (1.24) определяет выполнение требований (1.4), (1.5).

Если, в соответствии с (1.5) $P(Tr) \geq P_0$, то полет БЛА по маршруту Tr обеспечивает выполнение заданных требований.

1.5. Постановка задачи

Для повышения производительности и помехоустойчивости алгоритмов визуальной навигации предлагается перейти от растровых описаний ТИ и ЭИ к их семантическим описаниям.

Использование семантических описаний позволит реализовать «смысловое сжатие» содержания наблюдаемых сцен и перейти к более компактным алгоритмам их сравнения.

Исследуемый вариант семантического описания наблюдаемых сцен земной поверхности предназначен для поиска и распознавания (идентификации) принимаемых на борту БЛА ТИ на подготовленных заранее ЭИ.

Для оценки реализуемости предлагаемого подхода необходимо:

1. Разработать методику и алгоритмы формирования семантических описаний наблюдаемых сцен ЭИ и ТИ, включая:

- выбор метода сегментации изображений,
- выделение классов объектов на наблюдаемых сценах,
- выделение атрибутов объектов,
- определение отношений между объектами.

2. Разработать алгоритм поиска фрагмента ЭИ, идентичного принятому ТИ, на основе использования семантических описаний наблюдаемых сцен, включающий алгоритм идентификации (распознавания) семантических описаний ТИ и ЭИ. Основными задачами, которые необходимо решить при разработке соответствующего алгоритма поиска, являются:

- выбор метода обследования семантических описаний ЭИ (поиска ТИ);
- формирование словаря семантических признаков распознавания;
- выбор метрики оценки сходства наблюдаемых сцен, являющейся основой алгоритма идентификации;

- разработка алгоритма поиска семантического описания фрагмента ЭИ, включая процедуру принятия решения об окончании поиска и переходе к расчету координат БЛА.

3. Исследовать работоспособность и эффективность предлагаемых методик и алгоритмов, включая:

- формирование семантических описаний наблюдаемых сцен с учетом их устойчивости к влиянию дестабилизирующих факторов – изменений условий освещенности и влияния белого шума на исходные растровые ТИ;
- идентификацию ТИ и ЭИ на основе использования их семантических описаний сцен с учетом устойчивости к изменяемым условиям наблюдений;
- сравнительную оценку точности и быстродействия алгоритмов, использующих семантические и растровые описания изображений;
- подтверждение возможности оценки информативности признаков при семантическом описании сцен.

Глава 2. ФОРМИРОВАНИЕ СЕМАНТИЧЕСКОГО ОПИСАНИЯ НАБЛЮДАЕМЫХ СЦЕН

Семантическое описание изображения представляет собой формат представления визуальной информации наблюдаемой сцены с точки зрения его наполнения по смыслу: по наличию или отсутствию определенных объектов на изображении, по характеристикам наблюдаемых объектов и по их взаимному расположению на сцене [18].

Исследуемый вариант семантического описания наблюдаемых сцен земной поверхности предназначен для поиска и распознавания (идентификации) принимаемых на борту БЛА ТИ на ЭИ.

2.1. Семантические описания наблюдаемых сцен

В данной работе под семантическим описанием наблюдаемых сцен понимается содержание сцен, включающее перечисление классов присутствующих на изображении объектов, описание их атрибутов и пространственных отношений между ними в словесном или кодированном виде [27].

В рамках решаемой задачи, на основе использования семантических описаний, необходимо из набора фрагментов ЭИ выделить фрагмент, наиболее схожий (по заданному критерию) с принятым ТИ.

С учетом терминологии, принятой в теории распознавания образов [39], будем считать, что распознавание фрагмента ЭИ является процессом высшего уровня. Для этого уровня алфавитом классов распознаваемых объектов являются фрагменты ЭИ. Признаками, входящими в словарь признаков данного уровня распознавания, являются классы объектов, присутствующих на наблюдаемых сценах, их атрибуты и отношения между объектами.

На следующем уровне распознавания алфавитом классов являются классы присутствующих на сцене объектов, а признаками – атрибуты объектов.

Далее классами объектов распознавания могут быть атрибуты и отношения с признаками, входящими в соответствующие словари.

В общем случае классы объектов и их количество, атрибуты отдельных объектов и отношения между объектами являются признаками, определяющими индивидуальные особенности каждой наблюдаемой сцены.

На рисунке 2.1 в выделенном прямоугольнике показан состав возможных семантических описаний наблюдаемых сцен.

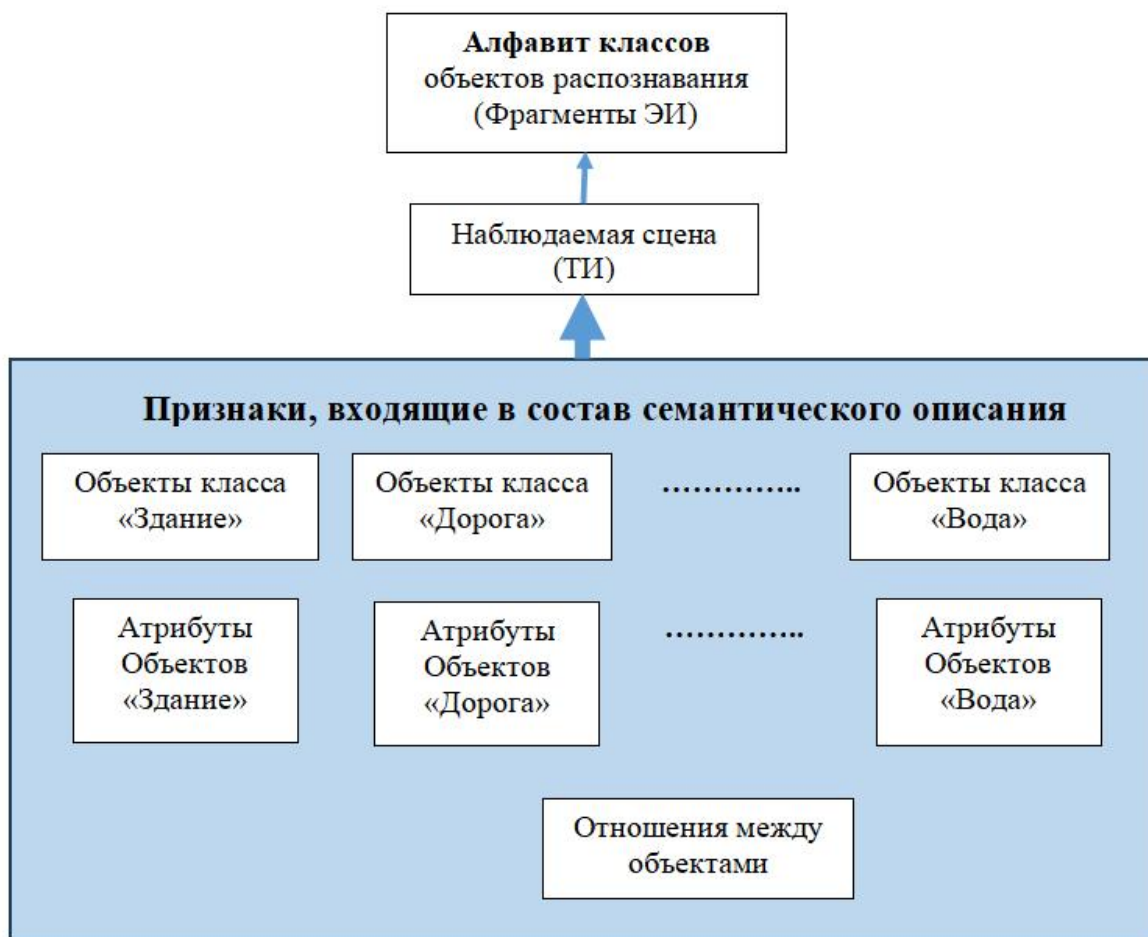


Рисунок 2.1 – Состав семантического описания

Рассматривая наблюдаемую сцену как объект распознавания (в частном случае – идентификации), можно считать классы, атрибуты и отношения признаками, входящими в соответствующий словарь признаков.

Необходимо отметить, что словарь признаков определяется из условий решаемых задач и может изменяться на отдельных этапах (циклах) распознавания.

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 различные 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.

Представленные выше семантические описания наблюдаемых сцен носят универсальный характер.

На начальных этапах разработки алгоритмов визуальной навигации более удобно использовать описания наблюдаемых сцен в словесном и/или кодированном виде .

Семантические описания наблюдаемых сцены могут иметь следующий **словесный вид**: «Большое здание рядом с дорогой» или «Маленькое здание недалеко от реки». Здесь:

- Классы объектов: «здание», «дорога», «река», ...
- Атрибуты: «большой», «маленький», ... , «белый», «черный», ..., «квадратный», «круглый», ...
- Отношения между объектами: «рядом», «недалеко», ... – расстояние между объектами, «на севере», «на юге», ... – направление.

В кодированном виде семантическое описание может быть представлено в следующем варианте:

- Классы 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): $SR(Dist((Ob(i), Ob(j))), Dir((Ob(i), Ob(j))))$.

Выделяемые классы объектов определяются априорно сформированным алфавитом классов.

Атрибуты объектов зависят от классов объектов и, в общем случае, включают признаки объектов, используемые в системах распознавания изображений: форму, размеры, ориентацию, текстуру, окраску объектов и пр.

Пространственные отношения между парами объектов определяются вектором – расстоянием между объектами и направлением.

Очевидно, что основными признаками наблюдаемых сцен являются классы объектов, т.к. атрибуты и отношения определяются только при наличии на сцене выделенных объектов.

Выделим три уровня описаний наблюдаемых сцен по выделяемым группам признаков.

Первый уровень связан с поиском объектов определенного класса на изображении, на данном этапе решается задача распознавания объектов и определение пространственного положения объекта на наблюдаемой сцене.

Второй уровень связан с выделением атрибутов – признаков найденных объектов [53].

Третий уровень – описание отношений, позволяющий определить взаимное положение объектов между собой с учетом семантики фона.

Для реализации первого уровня могут быть использованы классические методы сегментации, а также современные методы сегментации, основанные на глубоком обучении. Данные подходы позволяют решить задачу не только классификации объектов, но и задачу локализации объектов на наблюдаемой сцене [52].

Для реализации второго уровня необходимо определить состав признаков объектов для каждого класса, в частности, инвариантных признаков.

Так как условия наблюдения (время суток, ориентация системы наблюдения, положение объекта) в разные моменты фиксации изображений отличаются, то желательно, чтобы набор инвариантных признаков объектов не зависел от освещенности наблюдаемой сцены, ориентации системы наблюдения и высоты съемки. Для этих целей можно использовать топологически-семантический анализ, который представляет собой совокупность методов выделения, описания признаков объектов, полностью определяет форму объекта и содержит всю необходимую информацию для анализа объектов по их форме. Такой подход позволяет не учитывать пространственное изменение яркости пикселей объекта и тем самым позволяет значительно сократить объем перерабатываемой информации [43, 56].

Для реализации третьего уровня необходимо определить отношения между объектами. Наиболее доступными для оценки отношениями между объектами на изображении являются метрики дальности и направления.

Для работоспособности предложенного подхода необходимо, чтобы разрабатываемый алгоритм был инвариантен (в определенных задаваемых пределах) к изменениям условий наблюдения, в частности освещенности и масштаба, а также учитывал тот факт, что на наблюдаемой сцене может присутствовать только часть объектов за счет неполного перекрытия двух кадров между собой из-за перемещения системы наблюдения.

Кроме того, при решении задачи идентификации для удобства сопоставления разных признаков желательно, чтобы значения всех признаков были дискретными.

Методика формирования семантических описаний наблюдаемых сцен включает следующие этапы:

- 1) формирование алфавита классов объектов – 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. В результате выполнения данного этапа определяется перечень классов объектов, которые будут выделяться на наблюдаемых сценах.

Выбор классов объектов является сложным процессом и определяется рядом факторов:

- решаемой целевой задачей,
- характером подстилающей поверхности,
- условиями полета БЛА и пр.

Алфавит, содержащий избыточный набор классов, может привести к существенному увеличению трудоемкости подготовки описаний наблюдаемых сцен.

Недостаточное количество классов, включаемых в алфавит, не позволяет сформировать описание наблюдаемых сцен с требуемой точностью.

В качестве алфавита может быть рекомендован, например, следующий набор классов и цветов раскраски на семантическом изображении:

- 1) здание, красный;
- 2) поверхность (грунтовая территория), черный;
- 3) дорога, серый;
- 4) растительность, зеленый;
- 5) вода, синий.

Рабочий словарь признаков (п. 2), рассматриваемый в предлагаемом иерархическом алгоритме, формируется на основе признаков, выделяемых на ТИ, и является адаптивным к промежуточным результатам распознавания.

Так как после семантической сегментации исходных изображений известными являются классы объектов и их расположения на наблюдаемых

сценах, то на 1-м цикле распознавания ($r = 1$) для сокращения области поиска в качестве признаков предлагается использовать только классы объектов и их количество.

На 2-м и последующих циклах для выявления атрибутов объектов и отношений требуется дополнительный анализ наблюдаемой сцены и реализация соответствующих вычислительных процедур.

Процедуры выделения атрибутов объектов и отношений между объектами (**п. 4 Методики**) будут рассмотрены в разделе 2.3.

Необходимо отметить, что количество признаков наблюдаемых сцен, на основе использования значений атрибутов и отношений, следует ограничивать, исходя из требований к точности и вычислительной производительности решаемых задач.

На рисунке 2.2 представлена в общем виде блок-схема алгоритма формирования семантического описания наблюдаемой сцены.

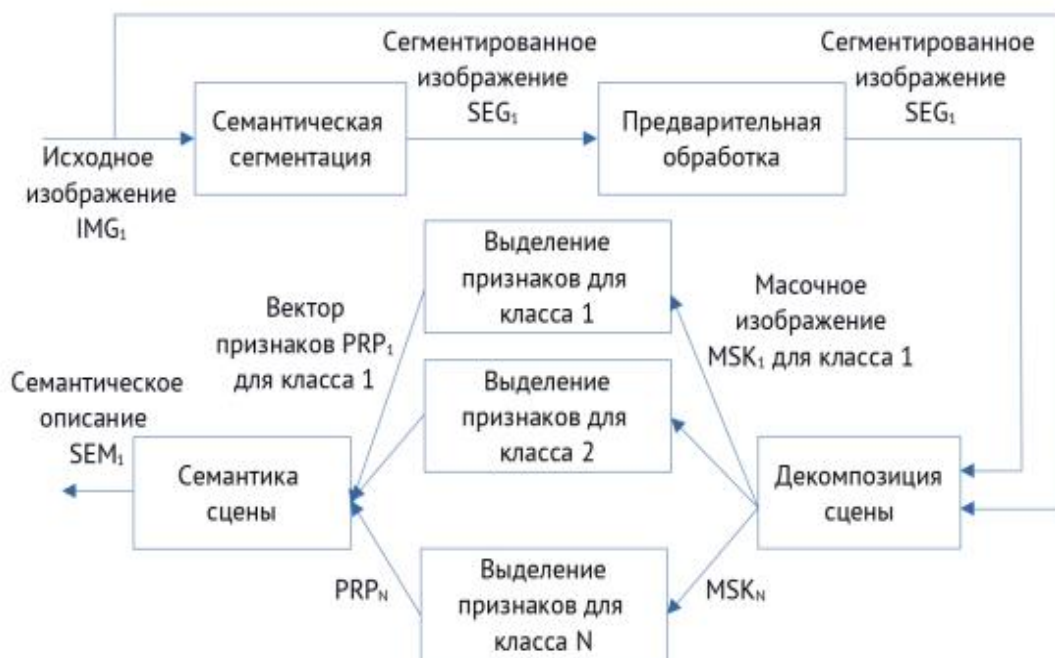


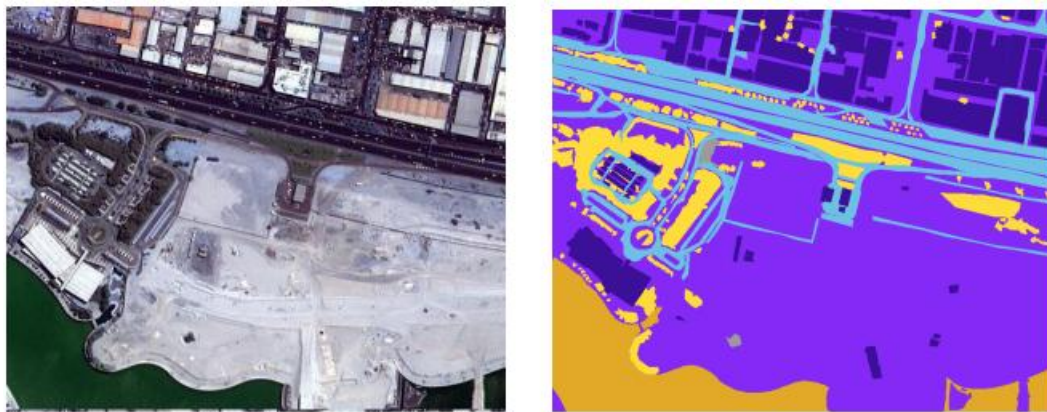
Рисунок 2.2 – Блок-схема алгоритма формирования семантического описания

На рисунке 2.2 приведены следующие обозначения:

IMG_1 – входное полноцветное исходное изображение наблюдаемой сцены, формат изображения по размеру и цветовой модели соответствует требованиям алгоритма семантической сегментации изображения.

SEG_1 – сегментированное изображение, по размеру совпадает с входным изображением и представляет собой цифровое изображение, разделенное на некоторые области, в которых цвет областей соответствует определенному классу.

Блок «Семантическая сегментация» представляет собой сверточную нейронную сеть. На вход в нейронную сеть подается растровое изображение (рисунок 2.3 а, в результате на выходе из нее получается сегментированное изображение (рисунок 2.3 б).



а) IMG_1

б) SEG_1

Рисунок 2.3 – Результат работы блока «Семантическая сегментация»

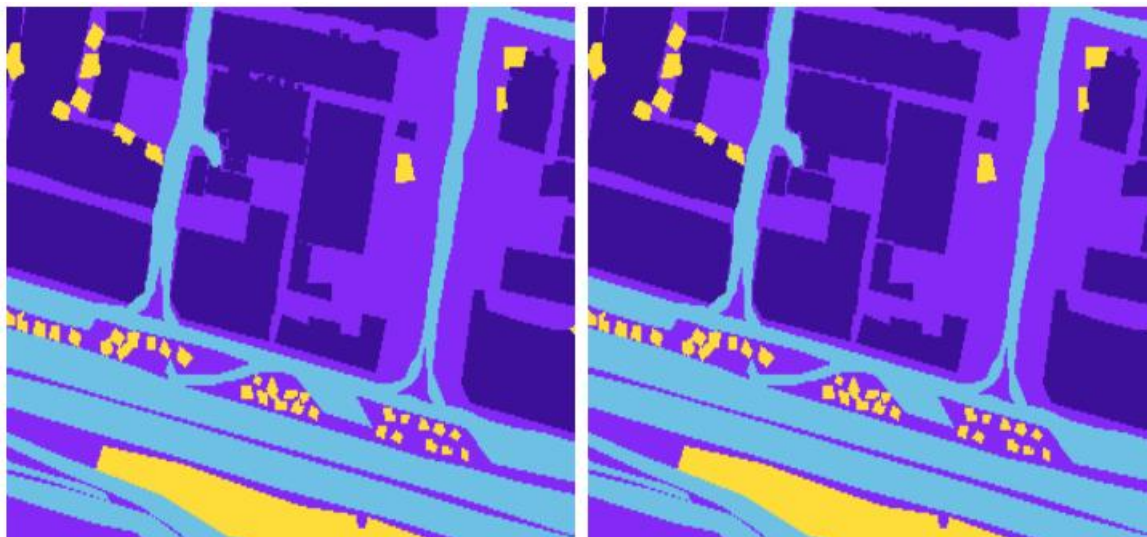
После применения процедуры «Семантическая сегментация» на сегментированном изображении могут присутствовать некоторые ошибки.

Примем, что к основным типам ошибок относятся:

- 1) слияние отдельно расположенных объектов;
- 2) разделение одного объекта на несколько объектов;
- 3) наличие «дырок» в сплошной поверхности;
- 4) деформация исходных границ объектов.

Задачей блока «Предварительная обработка» (рисунок 2.2) является устранение представленных выше ошибок на сегментированном изображении с использованием алгоритмов морфологической обработки изображений [20, 55].

На рисунке 2.4 а представлено исходное сегментированное изображение, а на рисунке 2.4 б представлено сегментированное изображение после применения блока «Предварительная обработка».



а) Исходное изображение б) После морфологической обработки

Рисунок 2.4 – Результат работы блока «Предварительная обработка»

Для устранения ошибок используются алгоритмы морфологии Серра [30]. Темно-фиолетовым цветом отмечены области, которые относятся к классу «Здания». При этом видно, что некоторые отдельстоящие здания соединены между собой перемычками того же цвета, что не позволит их в дальнейшем разделить на отдельные объекты.

В результате обработки некоторые изначально слитые здания были успешно разделены.

Далее обработанное сегментированное изображение (SEG_1) и исходное изображение (IMG_1) попадают в блок «Декомпозиция сцены» (рисунок 2.2). На выходе из блока с помощью алгоритма поиска связанных областей формируется масочное изображение MSK_i^j , где i – номер объекта, j – номер класса. Масочное изображение представляет собой полноцветное исходное изображение, в котором пикселям, не относящимся к данному объекту, присвоен черный цвет.

После декомпозиции сцены, в зависимости от выбранного класса, масочное изображение MSK_i^j передается в один из блоков «Выделение признаков для класса j » (рисунок 2.2).

Выбор информативных признаков и реализация алгоритмов выделения признаков является отдельной сложной для анализа задачей, поэтому описание этого блока будет представлено ниже. Выходом из блока «Выделение признаков для класса j » является вектор признаков PRP_i^j , размерность которого для разных классов может отличаться. В дальнейшем предполагается вектор признаков PRP_i^j использовать в алгоритме идентификации объектов.

Задачу классификации (распознавания) можно разделить на два отдельных этапа: верификация и идентификация объекта. Первый этап верификации позволяет на основе двух векторов признаков определить сходство данных объектов. А этап идентификации позволяет на основе списка объектов в базе данных определить, какие объекты находятся на текущей наблюдаемой сцене.

На рисунке 2.5 показаны два семантических изображения одной и той же наблюдаемой сцены, но наблюдаемые с разных ракурсов. Красными линиями соединены одинаковые объекты для первого и второго изображений.

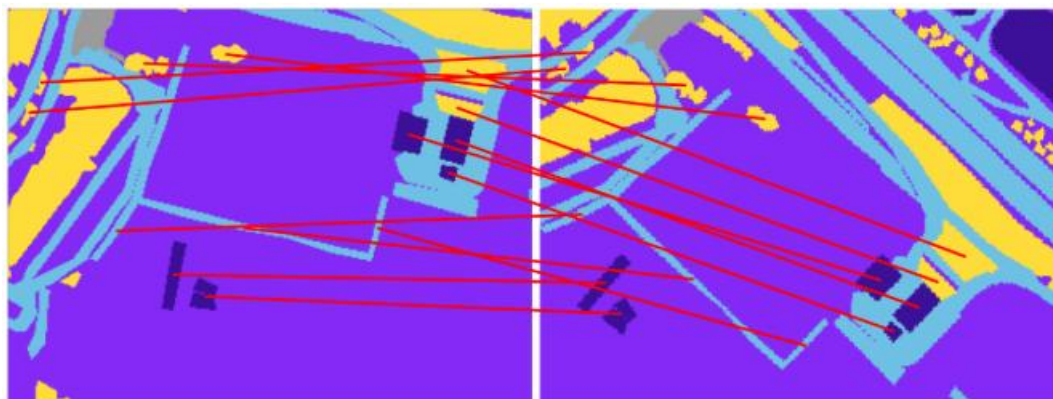


Рисунок 2.5 – Результат алгоритма идентификации объектов

Блок «Семантика сцены» (рисунок 2.2) на вход получает вектора признаков PRP_i^j всех объектов, находящихся на исходном изображении IMG_1 , и данный блок

формирует семантическое описание SEM_1 наблюдаемой сцены в определенном формате.

2.2 Выбор метода сегментации изображений

В разделе 2.1 не обсуждались вопросы выделения объектов, входящих в состав сформированного алфавита классов. Данная задача решается на основе семантической сегментации изображений.

Семантическое сегментирование (**п. 3 Методики**) изображения (сцены) является одной из задач компьютерного зрения. Исходя из наблюдений последних лет, в настоящее время, несмотря на увеличивающийся прогресс в использовании сложных дескрипторов изображений [18, 44] и более совершенных методов машинного обучения, в т.ч. нейросетевых, сегментация по-прежнему остается сложной задачей [42, 45, 58, 69].

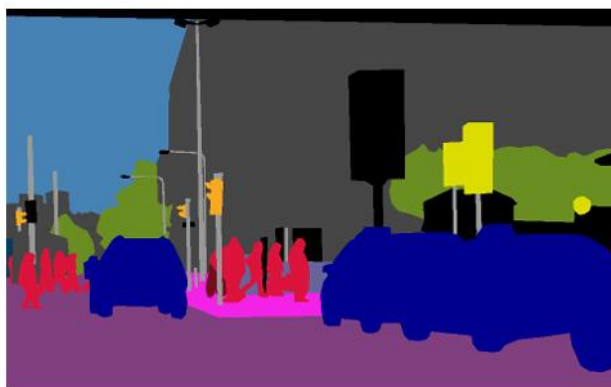
Одной из задач, связанных с нейросетевой семантической сегментацией изображений, является задача выбора рабочей архитектуры нейросети.

В данном разделе проводится сравнение работы архитектур FCN, SegNet и U-Net.

На рисунке 2.6 представлены основные типы (группы) сегментации изображений [12] на основе датасета [3]: а – исходное изображение, б – семантическая сегментация, в – сегментация (инстансовая) по экземплярам, г – паноптическая сегментация, сочетающая признаки семантической и инстансовой сегментации.



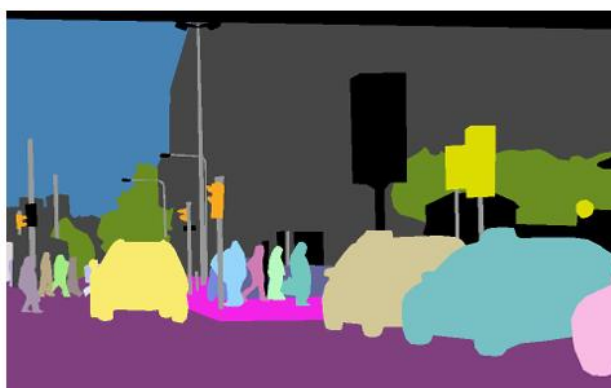
а) Исходное изображение



б) Семантическая сегментация



в) Инстансовая сегментация



г) Паноптическая сегментация

Рисунок 2.6 – Типы сегментаций

В рамках данного исследования реализуется семантическая сегментация.

Семантическая сегментация изображений – процесс деления изображения на области, где каждая область соответствует семантическому классу из заранее определенного алфавита классов [23]. Внешний вид этих классов изучается с помощью размеченных изображений.

Методы выделения объектов разных классов объединяют две подзадачи: часть распознавания, отвечающая за маркировку, и часть регуляризации, которая заставляет соседние пиксели принадлежать 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 СНС, поэтому приведем в качестве обзора эволюцию глубоких СНС к решению задачи сегментации.

Применение СНС в семантических моделях сегментации началось с огромного разнообразия. Среди различных моделей семантической сегментации на основе СНС наибольшее внимание получила полностью сверточная сеть (Fully Convolutional Network - FCN), и появилась тенденция к семантической сегментации на основе FCN. Чтобы сохранить пространственную информацию изображения, модели на основе FCN удаляли полностью связанные слои традиционного СНС. Так же использовали U-образную сеть, которая имеет сжимающее и расширяющее отображение. U-Net сжимает признаки, извлекает карты признаков и уменьшает пространственную информацию как традиционная сверточная сеть. Обратный путь принимает в качестве входных данных сжатую карту объектов и применяет операцию восходящей свертки.

Основные архитектуры U-Net и SegNet схожи, за исключением некоторых отдельных модификаций. Вторая половина этих архитектур является зеркальным отражением первой половины. Некоторые разработки использовали комбинацию архитектуры encoder-decoder и расширенной свертки для улучшения показателей сегментации. Исследователи из разных областей глубокого обучения также внедрили CNN для решения проблемы семантической сегментации. Кроме этого, существуют работы, которые использовали модели СНС для сегментации изображений на основе внимания, например, применялась нелокальная нейронная сеть DeepLab для извлечения карты признаков, а затем карты признаков подавались в рекуррентный модуль перекрестного внимания для семантической сегментации. Рассмотрим основные архитектуры СНС для решения задачи семантической сегментации.

Fully Convolutional Network (FCN). Одна из самых простых и популярных архитектур, используемых для семантической сегментации, это полностью сверточная сеть. В основе данной архитектуры (рисунок 2.7) лежит первоначальное преобразование входного изображения до меньшего размера

(одновременно с этим увеличивая количество каналов) через серию сверток. Такой набор сверточных операций обычно называется **кодировщик**. Затем выход декодируется или через билинейную интерполяцию, или через серию транспонированных сверток, который носит название **декодер**.

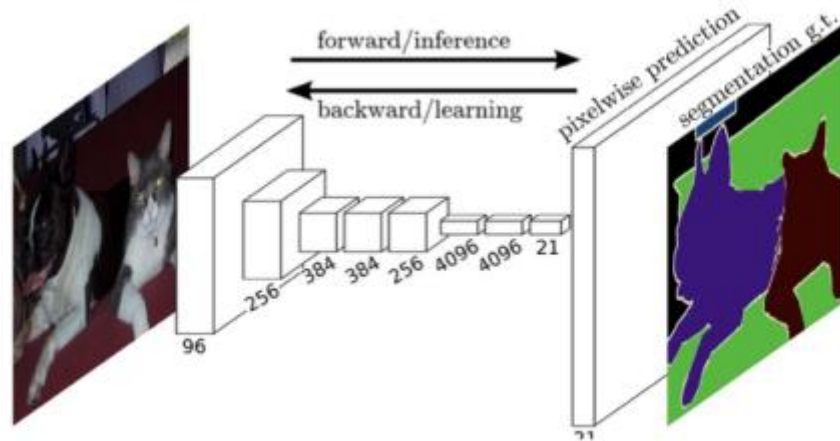


Рисунок 2.7 – Полностью сверточная нейронная сеть

Такая базовая архитектура, несмотря на ее эффективность, имеет ряд недостатков. Один из них – наличие артефактов, расположенных в шахматном порядке, связанных с неравномерным перекрытием выходов в операции транспонированной свертки. Другой недостаток связан с низкой разрешающей способностью по краям из-за потерь информации в процессе кодирования. Для увеличения качества базовой FCN-модели было предложено несколько решений, которые будут рассмотрены ниже.

Segnet имеет архитектуру кодер-декодер (рисунок 2.8), за которой следует конечный пиксельный классификационный слой. Кодировущая сеть имеет 13 сверточных слоев, как и в VGG16, а соответствующая часть декодера также имеет 13 де-сверточных слоев. В архитектуре не использовали полносвязные слои сети VGG16 для сохранения разрешения в самом глубоком слое, и это уменьшает количество параметров со 134 М до 14,7 М.

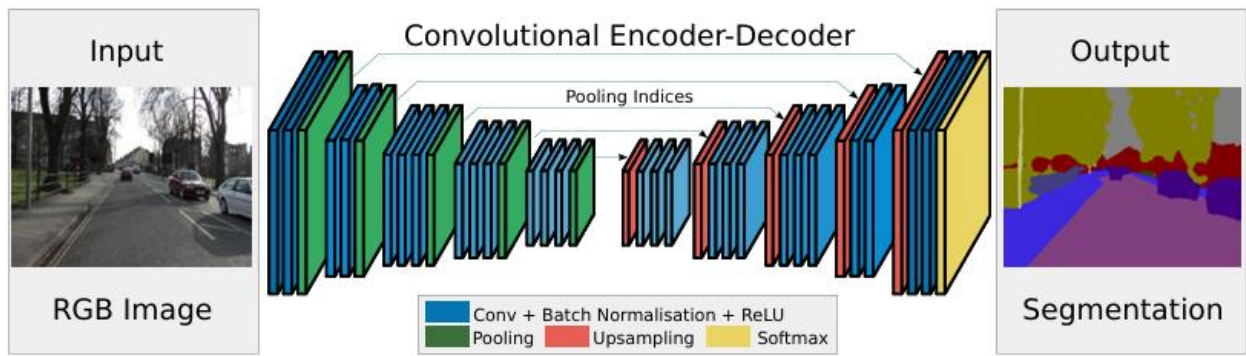


Рисунок 2.8 – Архитектура Segnet

В каждом слое сети кодировщика выполняется свертка с использованием набора фильтров для создания карт признаков. Далее для уменьшения внутреннего ковариационного сдвига используют пакетную нормализацию с последующей операцией нелинейности ReLU. Декодировочная часть в SegNet позволяет сделать более плавное разворачивание изображения после того, как отработала конволюционная (сверточная) часть архитектуры. Благодаря этому границы объектов на изображении определяются более корректно.

Чтобы сохранить граничную информацию в картах объектов, SegNet сохраняет только индексы максимального пула для каждой карты признаков кодера. Для семантической сегментации разрешение выходного изображения должно быть таким же, как и у входного. Для достижения этой цели SegNet выполняет операцию увеличения изображения в своем декодере, используя сохраненные индексы максимального объединения из соответствующей карты признаков кодера, в результате чего получается разреженная карта признаков высокого разрешения. Чтобы сделать карты объектов плотными, операция свертки выполняется с использованием обучаемого набора фильтров декодера. Затем карты признаков нормализуются. Выходная карта признаков высокого разрешения, полученная из конечного декодера, подается в обучаемый мультиклассовый классификатор *softmax* для пиксельной маркировки.

U-Net появился в 2015 г. для решения медицинских задач [16]. Эта архитектура и ее модификации являются одним из основных практических инструментов решения современных задач сегментации изображений [52].

Архитектура отличается более высокой степенью точности сегментирования, достаточной для работы с медицинскими снимками и картами, где располагаются множество мелких объектов, и других задач. *U-net*, так же как и *SegNet*, берёт входное изображение и реконструирует выход как сегментированную карту изображения. Однако в данной архитектуре используются так называемые “*skip connections*”, похожие на те, что помогли решить проблему исчезающего градиента в задаче классификации (*ResNet* архитектура). Эти “*skip connections*” позволяют сигналу не только проходить строго по траектории (сначала по нисходящей лестницы слоев, а потом по восходящей лестнице (увеличение маленькой репрезентации изображения в полноценную карту)), но и перескакивать между слоями одного уровня, что решает проблему дублирования функций слоев и улучшает эффективность обучения. Также они помогают второй половине модели, восстанавливающей результат, получать из первой половины модели пространственную информацию. На рисунке 2.9 представлена архитектура U-net.

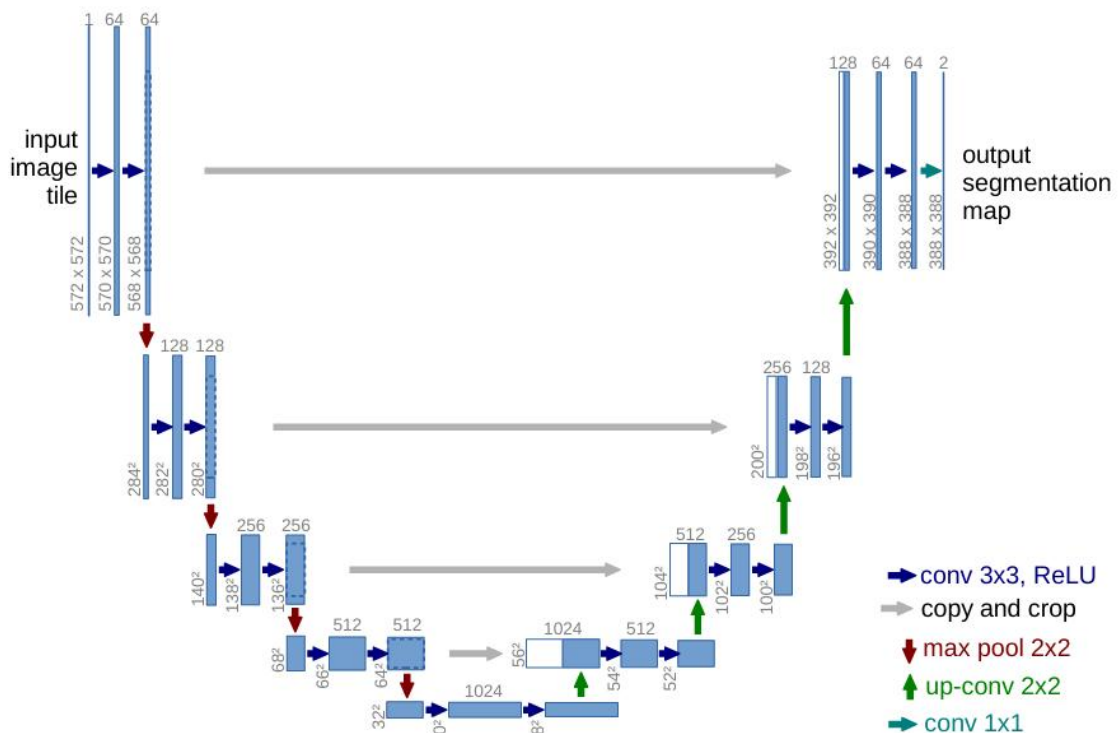


Рисунок 2.9 – Архитектура сети U-net

Путь сжатия в U-Net отвечает за идентификацию соответствующих объектов во входном изображении. Слои кодировщика выполняют сверточные операции, которые уменьшают пространственное разрешение карт признаков при одновременном увеличении их глубины, тем самым захватывая все более абстрактные представления входных данных. Этот путь сжатия аналогичен уровням прямой связи в других сверточных нейронных сетях. С другой стороны, путь развертки изображения работает над декодированием закодированных данных и определением местоположения объектов при сохранении пространственного разрешения входных данных. Слои декодера в пути развертки повышают дискретизацию карт признаков, одновременно выполняя сверточные операции. Пропуск соединений из пути сжатия помогает сохранить пространственную информацию, потерянную в пути сжатия, что помогает уровням декодера более точно определять местоположение объектов.

На рисунке 2.9 показано, как сеть U-Net преобразует входное изображение в оттенках серого размером $572 \times 572 \times 1$ пикселя в двоичную сегментированную выходную карту размером $388 \times 388 \times 2$. Можем заметить, что размер выходных данных меньше размера входных данных, потому что не используется заполнение. В процессе сжатия пути входное изображение постепенно уменьшается по высоте и ширине, но увеличивается количество каналов. Это увеличение каналов позволяет сети захватывать объекты высокого уровня по мере продвижения по этому пути. В узком месте выполняется заключительная операция свертки для создания карты объектов в форме $30 \times 30 \times 1024$. Затем путь развертки изображения извлекает карту объектов из самого крупного масштаба и преобразует ее обратно в изображение того же размера, что и исходные входные данные. Это делается с использованием слоев с повышающей дискретизацией, которые увеличивают пространственное разрешение карты объектов при одновременном сокращении количества каналов. Карты признаков из пути свертки изображения используются, чтобы помочь слоям декодера находить и уточнять объекты на изображении. Наконец, каждый пиксель на выходном изображении представляет собой метку, соответствующую определенному

объекту или классу на входном изображении. В этом случае выходная карта признаков представляет собой карту бинарной сегментации, где каждый пиксель представляет область переднего плана или фона.

Для представленных моделей 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.

$$IoU = \frac{TP}{TP + FP + FN}, \quad (2.1)$$

где TP – количество правильно проклассифицированных пикселей (true positives); FP – количество пикселей, которые метод проклассифицировал как принадлежащие классу, но они таковыми не являются (false positives); FN – количество пикселей, которые принадлежат классу, но метод проклассифицировал их как не принадлежащие классу (false negatives).

Значения TP , FP , FN вычисляются на основе матрицы ошибок, представленной на рисунке 2.10.

		Предсказание	
		True	False
Разметка	True	TP	FN
	False	FP	TN

Рисунок 2.10 – Матрица ошибок

Результаты сравнения по метрике IoU основных архитектур СНС для решения задачи семантической сегментации представлены на рисунке 2.11.

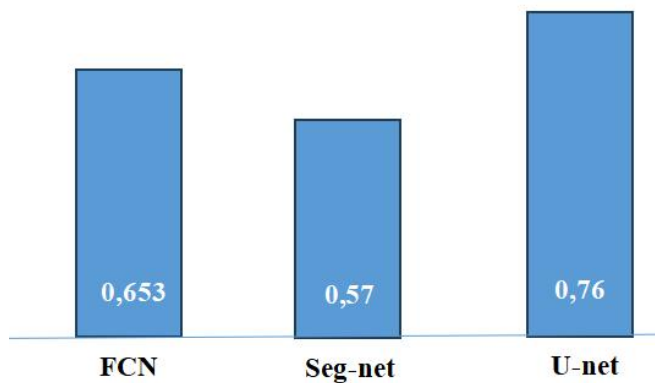


Рисунок 2.11 – Результаты сравнения СНС по метрике IoU

Анализируя рисунок 2.11, 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.

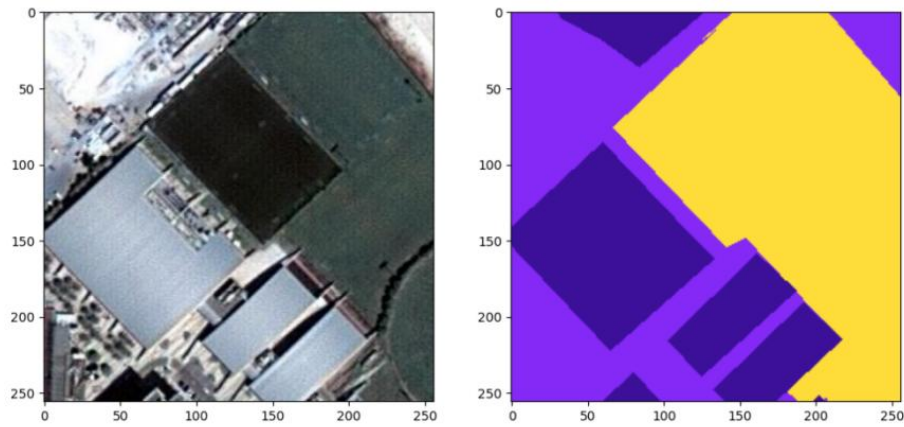
2.3. Выделение классов объектов на основе семантической сегментации

Для примера работы семантической сегментации был взят датасет, состоящий из аэрофотоснимков, полученных спутниками MBRSC и помеченных семантической сегментацией по пикселям в шести классах.

Общий объем набора данных составляет 72 трехканальных изображения размером 256×256 пикселей, сгруппированных в шесть фрагментов большего размера. Алфавит классов:

1. Здание: #3C1098 (синий)
2. Земля (грунтовая территория): #8429F6 (сине-лиловый)
3. Дорога: #6EC1E4 (аквамариновый)
4. Растительность: #FEDD3A (желтый)
5. Вода: #E2A929 (Бриллиантовый оранжево-желтый)
6. Без метки: #9B9B9B (серый)

На рисунке 2.12 а представлен пример исходного снимка и его размеченной маски (рисунок 2.12 б).



а) Исходное изображение б) Масочное изображение

Рисунок 2.12 – Оригинальное изображение и его маска

Потери нейронной сети оценивались на основе двух функций:

– Dixeloss:

$$DICE = \frac{2 * TP}{2 * TP + FP + FN}, \quad (2.2)$$

– CategoricalFocalLoss. Вычисляет альфа-сбалансированную фокальную потерю кроссэнтропии:

$$FL = (1 - pr)^{\gamma} * \log(pr), \quad (2.3)$$

где pr – предсказанный класс, γ – параметр фокусировки.

Когда $\gamma = 0$, фокусировка не влияет на перекрестную или кросс-энтропию.

Точность модели определялась на основе следующих метрик:

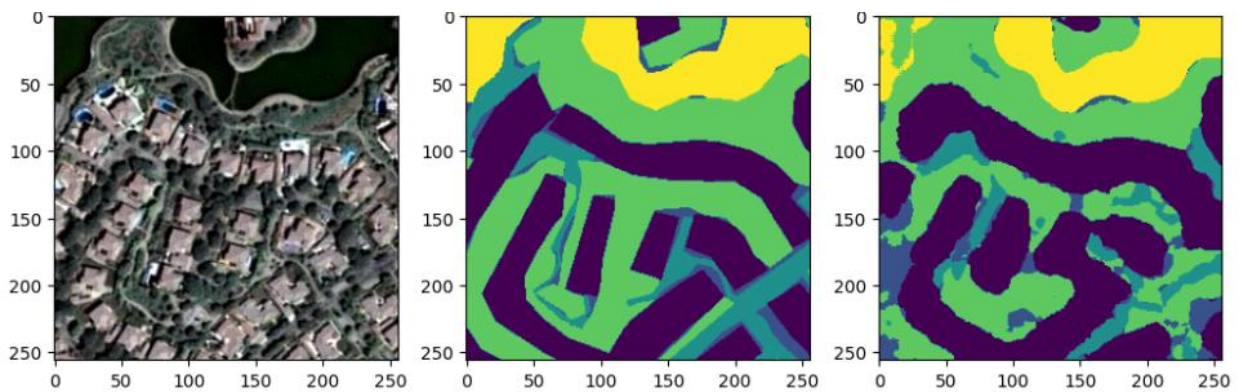
– *Accuracy* – точность модели, рассчитывается как количество правильно предсказанных классов, разделенное на общее количество объектов.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.4)$$

– *Jaccard coefficient* – коэффициент рассчитывается как количество общих элементов в двух классах (А и В), разделенное на общее количество элементов в любом из классов.

$$Jaccard_{coef}(A, B) = \frac{A \cap B}{A \cup B} \quad (2.5)$$

Пример результата сегментации приведен на рисунках 2.13 а, б, в.



а) Исходное изображение

б) Разметка

в) Выход нейронной сети

Рисунок 2.13 – Результат сегментации

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

1. Здание: **#ff009f** (розовый)
2. Земля (грунтовая территория): **#000000** (черный)
3. Дорога: **#797979** (серый)
4. Растительность: **# 00ff00** (зеленый)
5. Вода: **# 0000ff** (синий)

На рисунке 2.14 приведен график изменения функции потерь по эпохам:

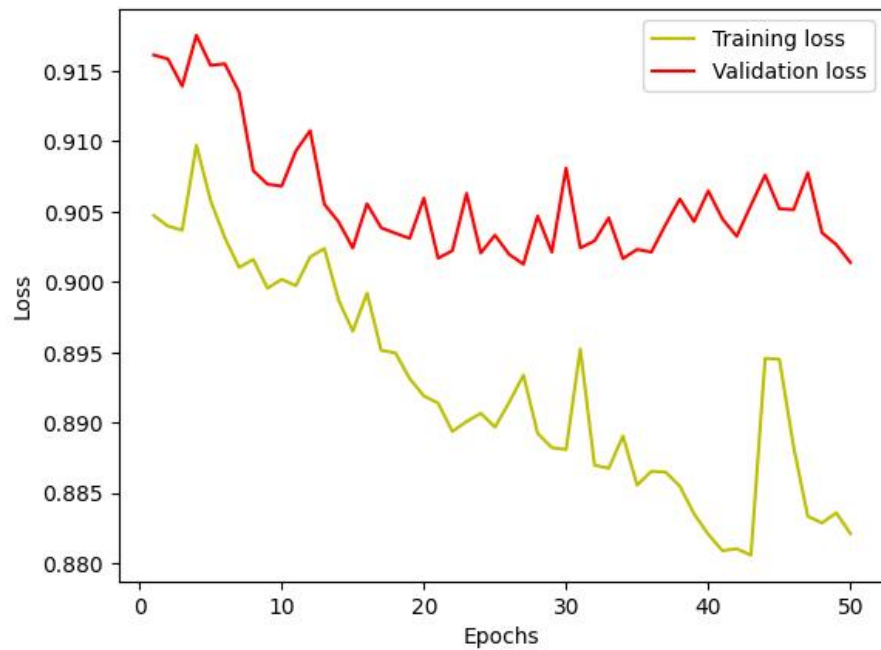
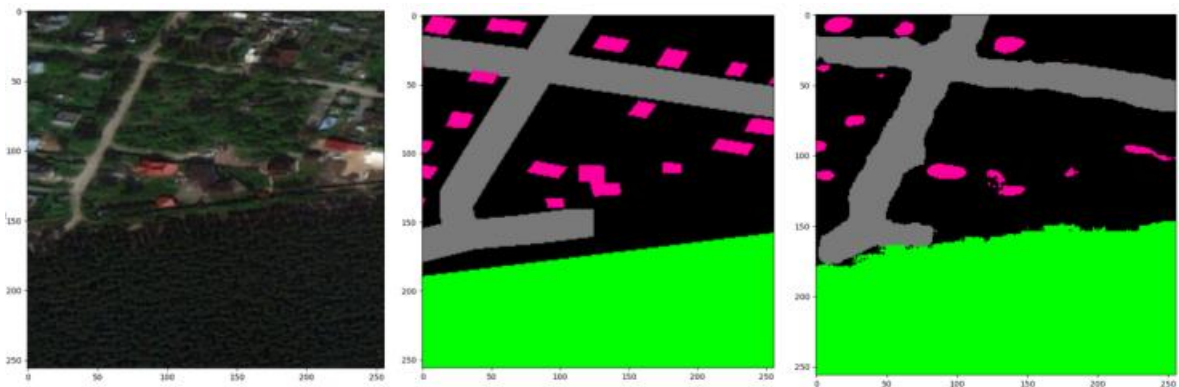


Рисунок 2.14 – График изменения функции потерь

Пример работы нейронной сети на размеченном фрагменте карты приведен на рисунках 2.15 а, б, в.



а) Растровое изображение б) Истинная маска в) Сегментация

Рисунок 2.15 – Результат сегментации нейронной сетью фрагмента карты одного из районов Санкт-Петербурга

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, соответствующий каждому классу на истинной маске изображения. Далее

необходимо определить на сегментированном изображении количество пикселей, которые соответствуют данному классу (N_{iTrue}) при попиксельном сравнении двух изображений (истинной и сегментированной маски) и определить вероятность $P(U_i \setminus U_j), i=\overline{1..5}$) правильного распознавания класса U_i при фактическом наличии класса U_i .

$$P(U_i \setminus U_j) = \frac{N_{iTrue}}{N_i}, i=\overline{1..5}. \quad (2.6)$$

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 ($N_{ijFalse}$).

$$P(U_i \setminus U_j) = \frac{N_{ijFalse}}{N_i}, i=\overline{1..5} \quad (2.7)$$

На рисунках 2.16 а, б, в изображен пример обработки изображения, содержащего все пять классов объектов. Полученная матрица ошибок приведена в таблице 2.1, где U_1 – здание, U_2 – земля, U_3 – дорога, U_4 – растительность, U_5 – вода.

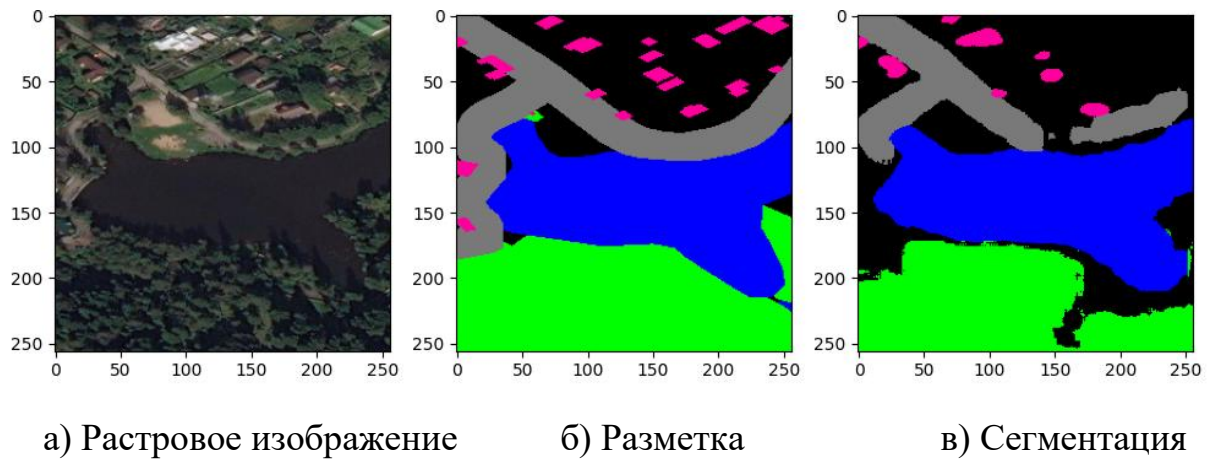


Рисунок 2.16 – Пример сегментированного изображения

Таблица 2.1 – Матрица вероятностей распознавания классов

	U_1	U_2	U_3	U_4	U_5
U_1	0,297201	0,660151	0,042648	0,00	0,00
U_2	0,033655	0,889258	0,064716	0,001397	0,010974
U_3	0,001434	0,297274	0,586448	0,00	0,114844
U_4	0,00	0,207458	0,00	0,785695	0,006847
U_5	0,00	0,048128	0,00	0,002297	0,949575

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

2.4. Выделение атрибутов и отношений между объектами

Как было показано, данные признаки наблюдаемых сцен относятся ко 2-му и 3-му уровням и выделяются (п. 4 Методики) после определения и локализации всех объектов (из заданного алфавита классов) на ТИ.

Использование этих признаков 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.

Выделение атрибутов объектов. 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. [65].

Извлечение признаков PRP_i является одним из важных этапов в задаче идентификации объектов. От выбранного набора признаков будет зависеть качество алгоритма идентификации.

Алгоритм формирования признаков можно разделить на два шага:

- 1) предварительная обработка масочного изображения – применение методов фильтрации для улучшения качества исходного изображения MSK_i^j ;
- 2) извлечение признаков – построение численного вектора на основе обработанного масочного изображения MSK_i^j и сегментированного изображения SEG_1 .

Применение методов фильтрации позволяет полностью или частично избавиться от различных искажений на изображении, обусловленных неидеальными условиями съемки, такими как неравномерность освещенности, смаз изображения, расфокусировка, а также внутренними помехами системы наблюдения, включая искажения оптической системы.

В таких случаях наиболее часто используют частотную фильтрацию, медианную фильтрацию, DoG-фильтр, вейвлет Хаара, фильтры Габора, методы восстановления изображений [38].

Формирование вектора признаков является самым сложным этапом, поскольку выбранный набор признаков влияет на качество идентификации и при этом для разных классов набор признаков может различаться.

Высокую эффективность в распознавании образов имеют типовые характеристики изображений, например: среднее значение яркости объекта: m_i $\{ m_R, m_G, m_b \}$, стандартное отклонение яркости объекта: s_i $\{ s_R, s_G, s_b \}$,

координаты центра масс объекта: $\{x_i, y_i\}$, площадь объекта: $square_i$, моменты Ху: $huM_i \{ \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7 \}$, гистограмма направленных градиентов: $gradient_i$.

Для начала перечислим признаки, которые являются общими для всех классов:

- 1) среднее значение яркости объекта: $m_i \{m_R, m_G, m_b\}$;
- 2) стандартное отклонение яркости объекта: $s_i \{s_R, s_G, s_b\}$;
- 3) координаты центра масс объекта: $\{x_i, y_i\}$;
- 4) площадь объекта: $square_i$;
- 5) моменты Ху: $huM_i \{ \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7 \}$;
- 6) гистограмма направленных градиентов: $gradient_i$.

Среднее значение яркости объекта для цифрового изображения вычисляется по формуле (2.8):

$$m = \frac{1}{n} \sum_x \sum_y I(x, y), \quad (2.8)$$

где $I(x, y)$ – значение яркости для изображения MSK_i^j в точке с координатами (x, y) ; n – размер объекта.

Размер объекта n рассчитывается на основе сегментированного изображения SEG_1 , в котором значение яркости 1 принадлежит объекту, а значение яркости 0 – фону. По сути, размер объекта n соответствует площади объекта $square_i$.

Цветное изображение в формате RGB обычно имеет три канала, поэтому среднее значение яркости можно рассчитать отдельно для каждого канала: m_R – среднее значение яркости для красного канала (*red*), m_G – среднее значение яркости для зеленого канала (*green*), m_b – среднее значение яркости для синего канала (*blue*). В программной реализации алгоритма формирования вектора признаков будет предусмотрен доступ к средним значениям по трем каналам m_R ,

m_G , m_b , но в алгоритме идентификации цветовая составляющая для некоторых классов объектов не будет учитываться.

Стандартное отклонение яркости объекта – оценка СКО случайной величины относительно ее МО на основе несмещенной оценки дисперсии. В данном случае в качестве случайной величины рассматриваются значения яркости пикселей изображения. Формула вычисления стандартного отклонения (2.9) для цифрового изображения выглядит следующим образом:

$$s = \sqrt{\frac{1}{n-1} \sum_x \sum_y (I(x, y) - m)^2}, \quad (2.9)$$

где $I(x, y)$ – значение яркости на изображении MSK_j^i в точке с координатами (x, y) ;
 n – размер объекта;
 m – среднее значение яркости объекта.

Для цветного изображения, как правило, стандартное отклонение для трех RG -каналов имеет примерно один порядок, поэтому принято, что при формировании вектора признаков будет использоваться стандартное отклонение только для яркости объекта. Для получения матрицы яркостей необходимо перевести цветное изображение в градации серого.

Моменты X_y

Момент изображения – некоторые частные средневзвешенные яркости пикселей изображения. Моменты X_y являются инвариантными к ряду преобразований, таких как изменение размера, поворот и др. Строятся на основе центральных моментов [17]. Применительно к цифровым изображениям простой момент M_{pq} вычисляется по формуле (2.10):

$$M_{pq} = \sum_x \sum_y x^p \cdot y^q \cdot I(x, y), \quad (2.10)$$

где p, q – порядок момента;

$I(x, y)$ – значение яркости на изображении MSK_j^i в точке с координатами (x, y) .

Центральные моменты для цифрового изображения рассчитываются по формуле (2.11):

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y), \quad (2.11)$$

$$\text{где } \bar{x} = \frac{M_{10}}{M_{00}},$$

$$\bar{y} = \frac{M_{01}}{M_{00}}.$$

Нормированные центральные моменты η_{pq} инвариантны к переносу и к изменению масштаба за счет деления соответствующего центрального момента на правильно масштабированный момент порядка (00) и вычисляются по формуле (2.12):

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\left(1 + \frac{p+q}{2}\right)}}. \quad (2.12)$$

Ниже представлены семь моментов Ху, которые инвариантны не только к переносу и масштабированию, но одновременно к вращению:

$$\phi_1 = \eta_{20} + \eta_{02};$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + (2 \cdot \eta_{11})^2;$$

$$\phi_3 = (\eta_{30} - 3 \cdot \eta_{12})^2 + (3 \cdot \eta_{21} - \eta_{03})^2;$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2;$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3 \cdot \eta_{12}) \cdot (\eta_{30} + \eta_{12}) \cdot [(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2] + \\ & + (3 \cdot \eta_{21} - \eta_{03}) \cdot (\eta_{21} + \eta_{03}) \cdot [3 \cdot (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]; \end{aligned}$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \cdot [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4 \cdot \eta_{11} \cdot (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03});$$

$$\begin{aligned} \phi_7 = & (3 \cdot \eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{21} + \eta_{03})^2] + \\ & + (\eta_{30} - 3 \cdot \eta_{12})(\eta_{21} + \eta_{03})[3 \cdot (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned}$$

Площадь объекта вычисляется по следующей формуле (2.13):

$$square = \sum_x \sum_y I(x, y), \quad (2.13)$$

где $I(x, y)$ – значение яркости на сегментированном изображении SEG_1 в точке с координатами (x, y) для заданного объекта.

Координаты центра масс объекта. Координаты центра масс объекта рассчитываются с использованием моментов M_{pq} и имеют следующий вид:

$$\begin{aligned} x &= \frac{M_{10}}{M_{00}}; \\ y &= \frac{M_{01}}{M_{00}}. \end{aligned} \quad (2.14)$$

В формуле для расчета M_{pq} (2.10) в качестве исходного изображения $I(x, y)$ используется сегментированное изображение SEG_1 для заданного объекта.

Гистограмма направленных градиентов. Метод гистограмм направленных градиентов основан на подсчете количества направлений градиента в локальных областях изображения [4]. Алгоритм основывается на допущении, что внешний вид и форма объекта на участке изображения могут быть описаны распределением градиентов яркости или направлением краев. В этом случае изображение объекта делится на пересекающиеся области и для каждой области вычисляется градиент, который представляет собой n -канальную гистограмму. Далее полученные гистограммы преобразуются в единый вектор, который может быть использован в качестве одного из элементов вектора признаков объекта.

Форма объектов класса «Здание». Важным атрибутом зданий является их форма, которая может классифицироваться обученной нейросетью.

На рисунке 2.17 представлены типы распознаваемых форм зданий: прямоугольный, Г-форма, Т-форма, П-форма, Н-форма, О-форма.



Рисунок 2.17 – Типы форм зданий

Выделение отношений объектов на наблюдаемой сцене

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

Отношения между объектами наблюдаемых сцен хранятся в соответствующих матрицах отношений.

При оценке отношений с такими объектами, как здание или дорога, удобно использовать соответственно центры или центральные (осевые линии объектов). Расстояние и направление до границ площадных объектов (береговая линия, граница леса и пр.) удобнее оценивать по расстоянию до ближайшей точки границы (линии).

Расстояние. Наиболее известной и распространенной мерой является евклидово расстояние между точками А и Б:

$$D(A,B)=\sqrt{(x_A-x_B)^2+(y_A-y_B)^2}, \quad (2.15)$$

где x_A, x_B, y_A, y_B – координаты точек А, Б.

В зависимости от условий наблюдения, а также искажений форм объектов, выделяемых нейросетью, возможно отклонение расстояний между точками на ТИ $D_{CI}(AB)$ и на ЭИ $D_{RI}(AB)$. Поэтому для оценки сходства данных признаков (отношений) на ТИ и ЭИ, требуется введение некоторого порога $\Delta D(AB)$. Выбор значения $\Delta D(AB)$ зависит от условий решения целевой задачи.

Тогда условие сходства расстояний на ЭИ и ТИ будет иметь вид:

$$|D_{RI}(AB) - D_{CI}(AB)| \leq \Delta D(AB). \quad (2.16)$$

В ряде случаев возможно использование других вариантов оценок расстояний. Например, в терминах нечеткой логики расстояния можно оценивать в следующих градациях: *вплотную, близко, не близко, недалеко, далеко и пр.*

Для большей определенности оценок будем считать, что расстояние между объектами в пределах одной градации равно

$$\Delta X = \frac{X_{CI}}{N_D}, \quad \Delta Y = \frac{Y_{CI}}{N_D}, \quad (2.17)$$

где X_{CI} , Y_{CI} – размеры ТИ;

N_D – количество градаций.

Направление, оцениваемое в угловой мере, может оцениваться в радианах, угловых градусах или в градациях.

Например, может использоваться приближенная мера восьми направлений (градаций): $0 \div 7$ или *Север, Юг, Запад, Восток, Северо-запад, Северо-восток, Юго-запад, Юго-восток*, где за нулевое направление или Север на ЦКМ выбирается направление вертикально вверх.

Глава 3. ИДЕНТИФИКАЦИЯ СЕМАНТИЧЕСКИХ ОПИСАНИЙ

В данной главе рассматривается задача поиска фрагмента ЭИ, идентичного ТИ, на основе их семантических описаний.

Идентификация семантических описаний ТИ и ЭИ является задачей распознавания, алфавит классов в которой включает возможные фрагменты ЭИ. В результате оценки сходства семантических описаний по признакам из рабочего словаря признаков F^r , производится выбор наиболее похожего семантического описания m -го фрагмента ЭИ $SD_{RI}(F^r, m^r)$ на описание $SD_{CI}(F^r)$ ТИ, где r – индекс цикла поиска.

Основными задачами, которые необходимо решить при разработке соответствующего алгоритма поиска, являются:

- выбор метода обследования семантических описаний ЭИ;
- формирование словаря семантических признаков распознавания;
- выбор метрики оценки сходства наблюдаемых сцен;
- разработка алгоритма поиска семантического описания фрагмента ЭИ, включая процедуру принятия решения об окончании поиска и переходе к расчету координат БЛА.

3.1. Оценка положения текущего изображения на эталонном изображении

Оценка положения координат БЛА в визуальной навигации состоит в поиске положения принимаемого на борту ТИ на большем по размеру ЭИ подстилающей поверхности – ЦКМ.

Другим вариантом поиска является поиск заранее определенного ориентира (ЭИ) на большом, принимаемом на борту ТИ.

Будем считать, что ТИ имеет меньший размер, чем ЭИ. При этом могут использоваться различные описания ТИ и ЭИ, в частности семантические.

Процесс поиска наиболее схожего с ТИ фрагмента ЭИ состоит из перебора фрагментов ЭИ, сравнения их сходства с ТИ и выделение наиболее похожего фрагмента (рисунок 3.1).

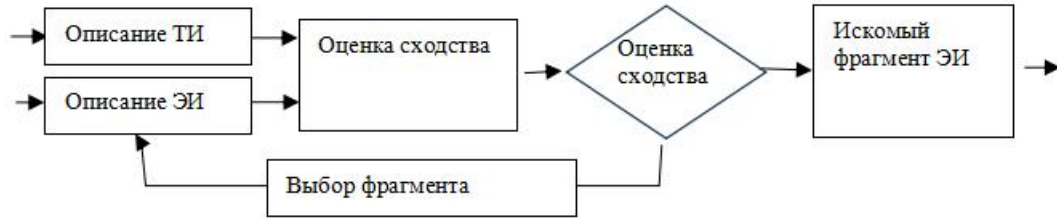


Рисунок 3.1 – Схема привязки ТИ к ЭИ

Данный процесс подобен известному процессу поиска наземных стационарных объектов с помощью авиационных носителей СН [22].

Нахождение фрагмента ЭИ, наиболее схожего с ТИ, является процессом распознавания (идентификации), где

- объектом распознавания является наблюдаемая сцена на принятом ТИ;
- распознаваемыми классами являются сцены, выделяемые на ЭИ;
- признаками являются особенности, характеристики, свойства наблюдаемых сцен.

Описание наблюдаемой сцены на выделенном фрагменте ЭИ зависит от размеров участка местности, попадающего в поле зрения бортовой СН.

Будем считать, что для получения изображений подстилающей поверхности используется видеокамера, установленная так, что ее линия визирования направлена вертикально вниз и стабилизирована.

Размеры участка местности S , попадающего в ТИ, определяются высотой полета БЛА H , угловым полем зрения углами по вертикали – α_x и горизонтали – α_y :

$$S = 4H^2 \operatorname{tg} \alpha_x \operatorname{tg} \alpha_y. \quad (3.1)$$

Размеры участка поверхности, попадающие в пиксель: по вертикали – $2H^2 \operatorname{tg} \alpha_x / m$, по горизонтали – $2H^2 \operatorname{tg} \alpha_y / n$, где m, n – размеры приемной матрицы.

Поиск положения ТИ на ЭИ. Рассмотрим три метода поиска ТИ на ЭИ: сплошной поиск, выборочное сканирование, иерархический поиск.

Метод сплошного поиска (полного перебора). Сплошным называется такое наблюдение (поиск), при котором обследованию подвергаются все без исключения элементы изучаемой совокупности, в частности все фрагменты ЭИ соответствующего размера (размер ТИ).

Данный метод широко используется, например, в корреляционно-экстремальных навигационных системах (КЭНС) [24, 26, 49] для поиска положения ТИ (размерностью в пикселях $m \times n$) на ЭИ ($M \times N$). В реализуемых алгоритмах производится попиксельное обследование всего ЭИ.

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.

Преимуществами метода является простота использования и его универсальность, что позволяет реализовать алгоритмы сплошного поиска при любых описаниях наблюдаемых сцен.

Однако данный метод является наиболее трудоемким и в ряде случаев предъявляет высокие требования к бортовым вычислителям БЛА.

Метод выборочного сканирования. Выборочное наблюдение (поиск) является одним из наиболее широко применяемых видов несплошного наблюдения. Выбор элементов (фрагментов ЭИ) производится на основе оценки дополнительных факторов, например, априорной вероятности их информативности и др.

В данном методе:

- на предполетном этапе формируется признаковое описание ЭИ с разделением признаков сцены (или объектов интереса) по соответствующим классам;
- в процессе полета формируется описание наблюдаемой сцены на принятом ТИ;

– производится сопоставление описаний ТИ и различных фрагментов ЭИ с учетом присутствующих на ТИ объектов (рисунок 3.1). Определяется искомый фрагмент ЭИ.

В процессе поиска на ЭИ выбираются и обследуются только те фрагменты (участки), которые содержат признаки, аналогичные признакам на ТИ. Словарь признаков должен быть постоянным и достаточно информативным, чтобы обеспечить решение поставленной задачи.

Отметим, что метод случайного поиска, который также относится к выборочному сканированию, используемый в ряде случаев поиска мобильных объектов, в рассматриваемых задачах не эффективен.

К преимуществам метода следует отнести существенное сокращение вычислений по сравнению с методом сплошного поиска.

Сложность реализации метода связана с необходимостью предварительного формирования *общего словаря признаков*, выделения этих признаков на ЭИ на предполетном этапе и выделения *рабочего словаря признаков* на ТИ в процессе полета.

Для оценки сходства ТИ и фрагментов ЭИ используется рабочий словарь признаков.

Иерархический поиск основан на предположении иерархической связности кластеров исследуемых данных. Обследуемая область интереса последовательно сжимается за счет сокращения элементов (фрагментов ЭИ) в кластерах. При этом кластеры фрагментов ЭИ формируются на основе принадлежащих им признаков.

Таким образом, на каждом цикле поиска определяется новая область поиска, сжатая по сравнению с предыдущей областью поиска.

Степень сжатия определяется обновляемым набором признаков – *адаптивным рабочим словарем признаков*, характеризующих текущий кластер выбираемых фрагментов ЭИ.

В рамках данного подхода, в случае появления альтернативных вариантов решений, т.е. нахождения на ЭИ нескольких фрагментов, похожих на ТИ,

производится расширение набора используемых признаков: геометрических отношений между объектами, атрибутами объектов.

Далее формируются (или выбираются из заранее подготовленной базы данных) описания ТИ и выбранных фрагментов ЭИ.

Процесс идентификации должен повторяться до тех пор, пока вероятность принятия правильного решения не достигнет требуемого (заданного) значения.

Данный подход является адаптивным к изменяемым условиям наблюдения, содержанию наблюдаемых ТИ и ЭИ. Поиск продолжается, пока не будет найден искомый фрагмент ЭИ.

Преимуществами данного метода являются:

- высокая вычислительная производительность,
- возможность комбинированного использования с другими методами на разных этапах поиска,
- возможность использования разных вариантов словарей признаков.

К недостаткам метода следует отнести высокую чувствительность вычислительной производительности иерархических алгоритмов поиска к выбору адаптивного рабочего словаря признаков.

Как показывает анализ рассмотренных методов, ключевым вопросом, определяющим производительность поиска фрагмента ЭИ, наиболее схожего с ТИ, является выбор словаря признаков сравниваемых сцен.

Недостаточный, с точки зрения полезной информативности, словарь признаков приводит к неоднозначности принимаемых решений, а избыточный – к непроизводительному расходованию вычислительных ресурсов. При этом следует учитывать необходимость выделения соответствующих признаков на ТИ в реальном времени, что повышает требуемые объемы вычислений.

3.2. Формирование словаря признаков семантических описаний наблюдаемых сцен

Описание наблюдаемой сцены должно обеспечивать уникальность описания искомой сцены и возможность принятия решения о распознавании с требуемой

вероятностью. При этом уникальность конкретной сцены на ТИ (в рамках используемого описания) определяется содержанием сцены и может изменяться в широких пределах.

Расширение пространства признаков (в случае их полезной информативности) может привести к улучшению их разделяющей способности в задачах обнаружения и распознавания.

В общем случае эффективность процессов распознавания объектов определяется информативностью используемых признаков [39].

В сложных задачах с ограниченными вычислительными ресурсами наиболее рациональным является реализация алгоритмов на основе иерархического метода или метода выборочного сканирования.

В частности, в иерархических алгоритмах рабочий словарь признаков обновляется на каждом цикле распознавания, пока поставленная задача не будет решена.

Однако неправильный выбор словаря признаков может привести к невозможности решить задачу распознавания с требуемой точностью или к избыточному расходу вычислительных ресурсов. При этом выделение и оценка значений разных признаков связаны с реализацией соответствующих алгоритмов разной трудоемкости.

Так, при распознавании наблюдаемых сцен в визуальной навигации признаки могут быть разнофизичными (тип объектов наблюдения, количество, цвет, размеры объектов и пр.) и принимать значения из разных множеств, например [39]:

- $\{0,1\}$, 0 – признак не выражен, 1 – признак выражен;
- $\{0,1,x\}$, x – информация о признаке отсутствует;
- $\{0,1, ..., d\}$ – степень выраженности признака имеет разные градации;
- $[a,b]$ – признак принимает значения из числового значения;
- $p_i(y_1, ..., y_n)$ – условная плотность распределения значений признаков.

Выделение подобных признаков и оценка их значений требует применения различных трудоемких алгоритмов обработки растровых изображений.

Развитие нейросетевых технологий позволило разработать более продуктивные, чем традиционные, методы выделения объектов заданных классов на изображениях.

Принципиальным отличием семантического описания наблюдаемых сцен от растрового является наличие в семантических описаниях выделенных объектов, присутствующих на наблюдаемых сценах.

Данные объекты, входящие в заранее определенные классы, а также различные атрибуты объектов и отношения (в частности, геометрические) между объектами являются, в общем случае, признаками наблюдаемых сцен и входят в общий и рабочие словари признаков.

В данной работе классы объектов, входящие в общий словарь признаков, выделяются нейросетевыми методами.

В качестве примера примем, что классами являются:

- здание (З), прямоугольники желтого цвета;
- дорога (Д), линии голубого цвета;
- поверхность (П), области зеленого цвета.

Как было показано в разделе 2.4, содержание семантических описаний может основываться на разных вариантах используемых словарей признаков.

Рассмотрим некоторые варианты семантического описания наблюдаемых сцен, показанных на рисунке 3.2, где наблюдаемые сцены выделены штриховыми прямоугольниками.

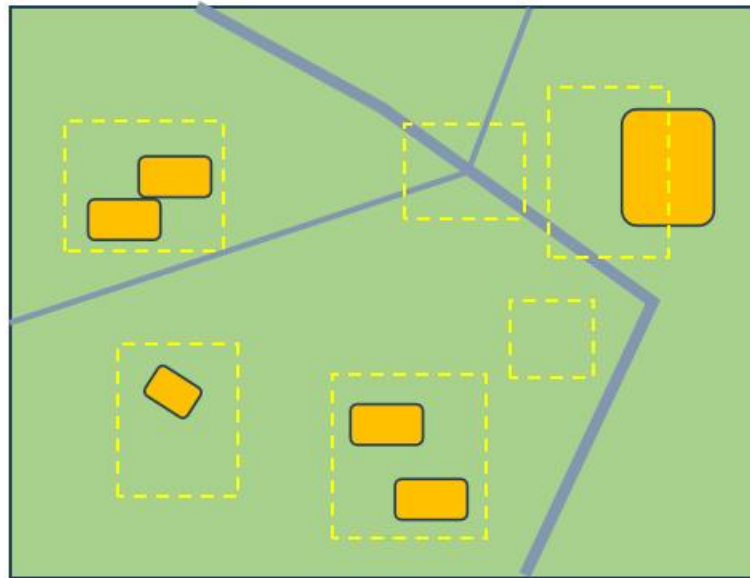


Рисунок 3.2 – Примеры наблюдаемых сцен

Вариант 1. Классы объектов

- | | |
|---|--------------------|
| – Сцена 1. На сцене присутствуют 3 объекта. | Описание: 3-3-П. |
| – Сцена 2. На сцене присутствуют 4 объекта. | Описание: Д-Д-Д-П. |
| – Сцена 3. На сцене присутствуют 2 объекта. | Описание: 3-П. |
| – Сцена 4. На сцене присутствует 1 объект. | Описание: П. |
| – Сцена 5. На сцене присутствуют 2 объекта. | Описание: 3-П. |
| – Сцена 6. На сцене присутствуют 3 объекта. | Описание: 3-3-П. |

Подобные семантические описания позволяют сопоставлять разные изображения и оценивать меру их сходства.

Описание сцены 4 не позволяет локализовать положение данного участка изображения. Здесь проявляется факт малой информативности навигационного поля [61].

Описание сцены 2 позволяет однозначно идентифицировать и решить навигационную задачу.

Описания пар сцен 1, 6 и 3, 5 не позволяют определять их уникальность, т.к. в данном варианте описания не учитываются атрибуты объектов и отношения между ними.

Вариант 2. Классы объектов и межобъектные отношения

Межобъектные отношения показывают относительное расположение объектов на наблюдаемой сцене. Оценка значений признаков может производиться на основе использования различных мер.

Одним из простейших вариантов являются качественные оценки дальности, например: вплотную (впл), рядом (ряд), близко (бл), недалеко (нд), далеко (д). При этом оценка «рядом» может определять дальность, например, как «не дальше максимальной длины объекта», а «близко» – «не дальше 5-ти длин объекта» и т.д.

Направление может оцениваться как: север, юг и т.д. или справа, слева, выше, ниже и пр.

Более точными являются векторные оценки взаимного положения объектов. Например, может оцениваться минимальное расстояние между объектами и угол, определяющий направление вектора.

При использовании данного варианта сцены пары 1, 6 (рисунок 3.2) легко идентифицируются:

- Сцена 1. Описание: 3-впл-3-П.
- Сцена 6. Описание: 3-ряд-3-П.

Необходимо отметить, что межобъектные отношения являются наиболее устойчивыми (к изменяемым условиям наблюдения) признаками наблюдаемых сцен.

Вариант 3. Классы объектов и атрибуты

Атрибутами (признаками) объектов являются: форма, размеры, текстура, распределение яркости, цвет объектов.

В частности, простейшим описанием атрибута «размеры» являются описания: маленький (м), средний (с), большой (б) или меньше, больше, равны.

Использование данного атрибута позволяет разделить описания сцен 3, 5 на рисунке 3.2.

- Сцена 3. Описание: 3-б-П.
- Сцена 5. Описание: 3-м-П.

Более полными являются описания, в которых атрибуты оцениваются количественно, например, в пикселях, метрах и пр.

При формировании рабочего словаря признаков необходимо учитывать, что атрибуты имеют разную устойчивость к изменяемым условиям наблюдения.

Вариант 4. Классы объектов, отношения и атрибуты

Наиболее полным семантическим описанием наблюдаемых сцен является описание, в котором присутствуют классы объектов, их атрибуты и отношения между объектами.

Сложность использования подобного описания связана с проблемами его формирования (выделение и оценка значений атрибутов и отношений) и применения в процессе полета в реальном времени.

3.3. Оценка сходства семантических описаний наблюдаемых сцен

Преимуществами оценки сходства изображений на основе семантических описаний наблюдаемых сцен по сравнению с растровыми описаниями является возможность сопоставления признаков (объектов, отношений, атрибутов) с более высокой вычислительной производительностью.

Выделение признаков, в частности объектов, с помощью обученных нейросетей позволяет существенно сократить время обработки и анализа видеoinформации.

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

Для оценки сходства семантических описаний возможно использование различных методов, ориентированных на особенности описаний.

Будем считать, что семантические описания двух сравниваемых наблюдаемых сцен сформированы полностью с достаточной степенью точности: классы используемых для идентификации ТИ и ЭИ признаков (объектов,

отношений, атрибутов) определены и могут быть представлены в виде (1,0) – признак присутствует или отсутствует.

В этом случае степень сходства должна быть определена по количеству совпадающих и несовпадающих значений признаков. При этом описание каждого отдельного признака является бинарным.

Существует ряд функций, позволяющих решить указанную задачу.

Например, распознавание может быть основано на вычислении степени похожести (оценки) распознаваемого объекта, на объекты, принадлежность которых к конкретным классам известна. Процедура распознавания включает в себя два этапа: подсчет оценок для каждого объекта и расчет суммарных оценок. При этом процедура вычисления оценок может иметь различные варианты.

Для бинарных растровых изображений предложен ряд парных критериальных функций. Например, широкое распространение получили функции:

Рао	$Kn_1 = \frac{a}{(a+b+c+e)}$;	
Джекарда	$Kn_2 = \frac{a}{(a+b+c)}$;	
Роджерса – Танимото	$Kn_3 = \frac{a}{(d_1+d_2-a)}$;	(3.2)
Соукала – Мишнера	$Kn_4 = \frac{(a+e)}{(a+b+c+e)}$;	
Юла	$Kn_5 = \frac{(ae-bc)}{(ae+bc)}$;	

где a, e – количество совпадающих элементов со значениями соответственно (1,1) и (0,0);

b, c – количество несовпадающих элементов со значениями соответственно (1,0) и (0,1);

d_1 и d_2 – число «1» в ТИ и ЭИ.

Данные функции предназначены для сравнения сходства растровых бинарных изображений по размерам, форме, расположению объектов интереса и могут быть неэффективны при сравнении семантических описаний.

Так, в знаменателе функций Рао, Соукала – Мишнера и Юла значение e не используется, потому что соответствующие признаки отсутствуют в семантическом описании.

Значения d_1 и d_2 в функции Роджерса – Танимото предназначены для сравнения размеров объектов на изображениях, не в семантических описаниях размеры (атрибуты) оцениваются иначе.

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

$$Kn_2 = \frac{a}{(a+b+c)}. \quad (3.3)$$

Сложность использования парных функций для сравнения семантических описаний состоит в том, что в парных функциях сравниваются только два возможных значения яркости: 1 и 0. При сравнении семантических описаний ТИ и ЭИ необходимо сравнивать наличие разных несопоставимых признаков, например, цвета и размеров объектов или объектов разных классов.

Для реализации функции Джекарда в семантических описаниях предлагается:

- дискретизировать непрерывные значения признаков наблюдаемых сцен, например, расстояния между объектами: «близко», «далеко»;
- обозначить присутствие и отсутствие оцениваемых **признаков** на наблюдаемой сцене показателями, соответственно – (1) и (0);
- заменить в функции Джекарда значения яркости на показатели присутствия или отсутствия оцениваемого признака: (1,1) – признак присутствует в ТИ и ЭИ, (1,0) – признак присутствует в ТИ и отсутствует в ЭИ и т.д.;
- проводить оценку сходства отдельно для каждого признака;

– проводить свертку показателей сходства по всем оцениваемым признакам.

Оценка сходства семантических описаний функцией Джекарда

Обозначим наличие некоторого признака в описании «1», а отсутствие признака – «0».

Например, признак «здание» на ТИ в семантическом описании ТИ обозначается «1» вне зависимости от размеров здания.

Реализация данных функций связана с вычислением количества признаков ТИ и ЭИ, совпадающих или не совпадающих по значению.

Так, «а» может интерпретироваться как присутствие на двух изображениях одинаковых признаков: классов объектов, отношений или атрибутов.

В таблице 3.1 представлены оценки сходства сцены 1 (рисунок 3.2) вариантов 1–4, рассмотренных в разделе 3.2, семантических описаний сцен на основе оценки функции Джекарда (3.3).

1. Вариант 1. Классы объектов.
2. Вариант 2. Классы объектов и межобъектные отношения.
3. Вариант 3. Классы объектов и атрибуты.
4. Вариант 4. Классы объектов, отношения и атрибуты.

Например, для **варианта 1 (сопоставление классов)** при сравнении сцен 1 и 1 получим $Kn_2 = 1$, так как $a = 3$ – совпадают объекты 3, 3, П; $b, c = 0$ – несовпадающих объектов нет.

При сравнении сцен 1 и 2 (рисунок 3.2) получим $Kn_2 = 0,14$, так как $a = 1$ – совпадают объекты П; $b = 2$ – не совпадают объекты 3, 3; $c = 4$ – не совпадают объекты Д, Д, Д, Д.

Для **варианта 2 (сопоставление классов и отношений)** при сравнении сцен 1 (описание: 3-впл-3-П) и 1 получим $Kn_2 = 1$, так как $a = 4$ – совпадают объекты 3, 3, П и отношение между объектами 3 и 3 – вплотную; $b, c = 0$ – несовпадающих объектов нет.

При сравнении сцен 1 и 2 (описание: Д-Д-Д-Д-П, отношения не учитываются) получим $Kn_2 = 0,14$.

Результаты аналогичных расчетов для вариантов 3, 4 представлены в таблице 3.1.

Таблица 3.1 – Сравнение сцены 1 со сценами 1–6 функцией Джекарда

Сцена	1	2	3	4	5	6
Вариант 1	1	0,14	0,67	0,33	0,67	1
Вариант 2	1	0.12	0,5	0,25	0,5	0,6
Вариант 3	1	0,11	0,33	0,2	0,33	1
Вариант 4	1	0.1	0,28	0,17	0,28	0,71

Как видно из таблицы 3.1, варианты 1, 3 описаний не позволяют однозначно идентифицировать сцену 1. Сцены 1 и 6 имеют одинаковый состав объектов, причем атрибуты – размеры присутствующих зданий – также одинаковы.

Различие сцен состоит в разном расстоянии между зданиями, что показывает признак отношения в варианте 4.

Сцены 3, 5 и 1, 6 в вариантах 1, 2, 3 описаний имеют соответственно равные значения функции Джекарда, что показывает недостаточность описаний по этим вариантам.

Наилучшую разделяющую способность имеет семантическое описание сцены по варианту 4, позволяющее однозначно идентифицировать сцену 1.

Таблица 3.1. также показывает, что расширение словаря признаков приводит к повышению точности оценки сходства наблюдаемых сцен.

Критериальные функции Kn_i (3.2), рассчитываемые для каждого типа признаков, имеют разную чувствительность к частотам появления признаков, объему словаря признаков и алфавиту классов.

Так, рассмотренный пример основан на использовании функции Джекарда, которая учитывает оценку «с» отсутствующих на ТИ признаков, но присутствующих в общем словаре признаков. Очевидно, что при увеличении общего количества признаков чувствительность значения «а» к изменению совпадении признаков будет падать.

3.4. Иерархический алгоритм поиска

Алгоритм предназначен для поиска фрагмента ЭИ, идентичного ТИ, на основе сравнения семантических описаний изображений.

Основная идея данного алгоритма – последовательное сокращение области поиска (фрагментов ЭИ) за счет использования рабочего словаря признаков, расширяемого на каждом цикле поиска. При этом на каждом r -м цикле производится оценка сходства описаний фрагментов ЭИ $SD_{RI}(F^r, m^r)$ с ТИ $SD_{CI}(F^r)$, где F^r – словарь признаков на r -м цикле, m – индекс фрагмента ЭИ.

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

При адаптивном расширении словаря признаков более удобной по сравнению с (3.2) будет модифицированная функция Джекарда (МФД), применяемая для оценки сходства наблюдаемых сцен:

$$E(F, m^r) = \frac{1}{L^r} \sum_{l=1}^{L^r} E(F_l, m^r), \quad (3.4)$$

где $l \in L$ – индекс класса признаков;

$E(F_l, m^r) = \frac{a_l}{a_l + b_l + c_l}$ – парная критериальная функция на основе функции Джекарда (3.3) для оценки сходства семантических описаний ТИ и m -го фрагмента ЭИ по l -му признаку на r -м цикле поиска.

В (3.4) вычисления локальной оценки сходства проводятся для каждого l -го признака, далее оценки суммируются.

В общем случае поиск продолжается, пока не будет достигнута требуемая (заданная) точность сходства

$$E(F, m^r) \geq E_0, \quad (3.5)$$

где E_0 – заданное значение оценки сходства.

При необходимости требуемые оценки сходства могут задаваться для каждого цикла поиска $E(F', m') \geq E_0(F', m')$.

Введение данного условия может упростить процедуру принятия решения о переходе алгоритма на новый цикл поиска.

Необходимо отметить, что переход на новый цикл требует обновления словаря признаков, что может быть связано с существенным увеличением требуемых вычислений. Если объекты, входящие в основной словарь признаков, выделяются нейросетью, то выделение на ТИ отдельных атрибутов или отношений могут потребовать подключения дополнительных трудоемких алгоритмов. Поэтому целесообразно переходить на новый цикл вычислений, максимально сократив область поиска, не изменяя рабочий словарь признаков.

Также одной из проблем при формировании алгоритма является определение процедуры принятия решения об окончании поиска.

Сложность данной проблемы связана с тем, что окончание поиска не всегда можно реализовать по выполнению условия (3.5), т.к. нельзя исключать возможность того, что несколько фрагментов ЭИ будут иметь одинаково высокую степень сходства с $SD_{CI}(F^r)$.

Кроме того, полученные значения оценки сходства могут не удовлетворять заданные требования, например, из-за ухудшения условий наблюдения.

В подобных ситуациях необходимо принимать решения по выбору искомого фрагмента ЭИ на основе выполнения других требований.

Поэтому в работе предлагается ввести дополнительное условие окончания поиска, связанное с оценкой апостериорной вероятности получения искомого фрагмента ЭИ.

Блок-схема предлагаемого алгоритма представлена на рисунке 3.3.

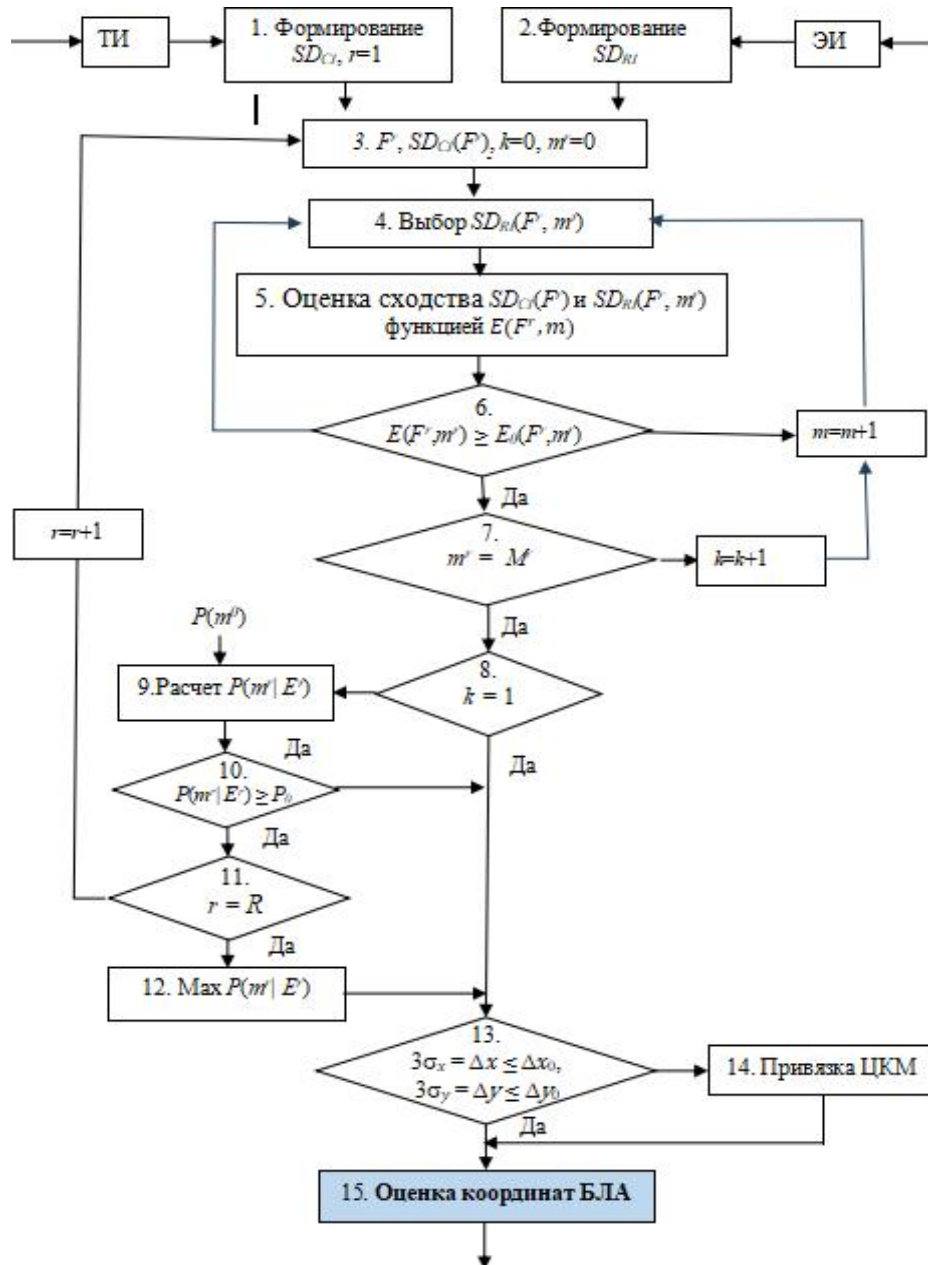


Рисунок 3.3 – Блок-схема алгоритма идентификации SD_{CI}

В **блоках 1, 2** (рисунок 3.3, блоки 1, 2) на базе полученных семантических изображений формируются семантические описания сцен ТИ и ЭИ.

Для наблюдаемой сцены могут быть сформированы разные семантические описания $SD_{CI}(F^r)$ (**блок 3**), базирующиеся на использовании разных 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 $F^r_{l'}$, 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 блоке 4 производится выбор очередного фрагмента семантического описания ЭИ $SD_{RI}(F^r, m)$ для сравнения с $SD_{CI}(F^r)$.

В блоке 5 проводится оценка сходства описаний $SD_{CI}(F^r)$ и $SD_{RI}(F^r, m^r)$ по всем L^r признакам по функции (3.4).

По каждому признаку в (3.4) l сходство оценивается на базе парной критериальной МФД, в которой a_l, b_l, c_l получены из описаний $SD_{CI}(F^r)$ и $SD_{RI}(F^r, m^r)$:

$$E(F^r, m^r) = \frac{a_l}{a_l + b_l + c_l}, \quad (3.6)$$

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).

В блоке 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 (блок 7) все возможные варианты ($m^r = M^r$) и получено (блок 8) единственное решение ($k=1$, где k – количество фрагментов ЭИ, в которых семантические описания $SD_{RI}(F^r, m^r)$ идентичны $SD_{CI}(F^r)$), то происходит переход к вычислению координат БЛА (блок 13).

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

$$P(m^r | E^{r-1}) = \frac{P(m^{r-1})P(E(F^{r-1}, m^{r-1}) | m^{r-1})}{\sum_{k^r}^{M^r} [P(k^{r-1})P(E(F^r, k^{r-1}) | k^{r-1})]}, \quad (3.7)$$

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 (глава 1).

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

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

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

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

Глава 4. ИССЛЕДОВАНИЕ РАБОТОСПОСОБНОСТИ АЛГОРИТМОВ ВИЗУАЛЬНОЙ НАВИГАЦИИ БЛА С ИСПОЛЬЗОВАНИЕМ СЕМАНТИЧЕСКИХ ОПИСАНИЙ ИЗОБРАЖЕНИЙ

Целью данной главы является исследование работоспособности и эффективности предлагаемых методик и алгоритмов:

- формирования семантических описаний наблюдаемых сцен;
- идентификации ТИ и ЭИ с использованием семантических описаний сцен;
- оценки информативности признаков при семантическом описании наблюдаемых сцен.

Исследования проводятся на основе математического моделирования процессов обработки исходных растровых изображений, семантической сегментации и сопоставления наблюдаемых сцен на ТИ и ЭИ при их растровом и семантическом описаниях.

Работоспособность. Работоспособный алгоритм должен обеспечивать выполнение поставленной задачи с требуемой точностью в изменяемых условиях наблюдений, в частности, в условиях воздействия основных дестабилизирующих факторов: изменений МО и СКО яркости ТИ, а также при воздействиях аддитивного белого шума.

Оценка помехоустойчивости показывает, в каком диапазоне изменений условий наблюдения возможно выполнение поставленных задач (распознавания классов объектов) с требуемой точностью.

Так, изменениям математического ожидания яркости соответствуют изменения освещенности наблюдаемых сцен: максимальная яркость (для средней полосы) – лето, полдень, солнечный день (более 50000 лк (люкс)); средняя яркость – лето, 14–15 ч (около 10000 лк); минимальная яркость – лето, 19:00 ч (около 2000 лк).

Изменения СКО яркости соответствует в реальных условиях изменениям контрастности изображений, например, в ясный солнечный день, а также в туман и пасмурную погоду.

Примем, что *рабочий диапазон* изменений МО должен соответствовать освещенности земной поверхности с 6:00 до 20:00 ч летом (высота положения солнца $\geq 15^\circ$ над горизонтом).

Следовательно, проверка алгоритмов должна проводиться в соответствующих условиях.

Для обеспечения требований по точности определим расстояние между фрагментами ЭИ Δx , Δy таким образом, чтобы выполнялось условие (1.4):

$$\Delta x \leq |\Delta x_0|, \Delta y \leq |\Delta y_0|, \quad (4.1)$$

где $\pm \Delta x_0, \pm \Delta y_0$ – допустимые ошибки оценки координат.

Следовательно, если принимается единственное решение и некоторый выбранный фрагмент ЭИ соответствует ТИ, то условие (1.4) выполняется.

В этом случае условие работоспособности алгоритмов распознавания, в соответствии с (1.5), будут иметь вид:

$$P(m) \geq P_0, \quad (4.2)$$

где $P(m)$ – вероятность правильного распознавания m -го исхода из M_Σ возможных; P_0 – требуемая вероятность.

Эффективность алгоритмов идентификации ТИ и ЭИ будем оценивать на основе сравнения точности, помехоустойчивости и быстродействия алгоритмов, использующих семантические и растровые описания изображений.

Для анализа работоспособности и эффективности исследуемых методик и алгоритмов проведена оценка:

1. Устойчивости семантических описаний сцен к изменяемым условиям наблюдения (раздел 4.1):

- изменениям МО яркости ТИ,
- изменениям СКО яркости ТИ,
- воздействию аддитивного белого шума.

2. Влияния условий наблюдения на точность идентификации изображений (раздел 4.2).

3. Точности корреляционных алгоритмов сопоставления растровых изображений и алгоритмов на основе семантических описаний (раздел 4.3).

4. Вычислительных затрат на реализацию предложенного подхода идентификации $SD_{CI}(F^r)$ и $SD_{RI}(F^r, m^r)$ в сравнении с аналогичным подходом сопоставления ТИ и ЭИ на основе корреляционной функции (раздел 4.4).

5. Реализуемости расчета информативности признаков (1.13) при семантическом описании наблюдаемых сцен (раздел 4.5).

Общая методика моделирования состоит из следующих этапов:

1. Определение области присутствия БЛА на ЭИ.
2. Выбор ТИ в области присутствия БЛА.
3. Модификация ТИ, отображающая изменение условий наблюдения.
4. Сравнение точности и помехоустойчивости алгоритмов визуальной навигации.
5. Оценка работоспособности разработанного алгоритма.

4.1. Оценка помехоустойчивости алгоритма семантической сегментации изображений в изменяемых условиях наблюдений

Целью данного исследования является оценка помехоустойчивости осуществления семантической сегментации изображений.

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

$$P_{tp}^k(q) = \frac{n_{tp}^k}{N_0^k}, \quad (4.3)$$

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

N_0^k – количество элементов k -го класса;

q – индекс дестабилизирующего фактора.

В качестве дестабилизирующих факторов, влияющих на точность распознавания, рассматриваются:

- изменения МО яркости изображения (освещенности наблюдаемых сцен);
- изменения СКО яркости изображений;
- влияние белого гауссовского шума на изображении.

Устойчивость к белому гауссовскому шуму оценивается вероятностью $P(\beta)$, рассчитываемой по формуле (4.3).

Величина шума характеризуется отношением:

$$\beta = SNR = \frac{\sigma_s^2}{\sigma_n^2}, \quad (4.4)$$

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

Указанные дестабилизирующие факторы ухудшают точность выявления, распознавания и оценивания признаков распознаваемых объектов.

Перед проведением исследований необходимо произвести обучение нейронной сети на выборке, состоящей из изображений земной поверхности и разметки для этих изображений.

Методика эксперимента включает следующие этапы:

1. Обучение нейронной сети для семантической сегментации изображения на нескольких выборках. Оценка качества обучения по IoU.

2. Варьирование МО яркости входных изображений, оценка влияния на точность сегментирования (4.3).

3. Варьирование СКО яркости входных изображений, оценка влияния на точность сегментирования (4.3).

4. Добавление гауссовского белого шума к входному изображению, оценка влияния на точность сегментирования (4.3).

5. Анализ полученных результатов.

Обучение нейронной сети (этап 1).

Как было показано в главе 2, наилучший показатель точности, по метрике IOU, среди рассматриваемых сетей показала сеть U-net.

С целью подтверждения возможностей U-net при работе с различными данными были проведены исследования с использованием как спутниковых изображений, так и аэрофотоснимков.

В качестве исходных данных для обучения были взяты датасеты открытого доступа: [1], [2], [10] и размеченные участки Санкт-Петербурга.

Каждый из представленных датасетов содержит разметку для классов: здания, растительность, водная поверхность, дорога. Остальные классы датасетов были помещены в класс с кодовым значением “0”, который обозначает фон.

В частности, датасет [2] содержит изображения аэрофотосъемок различных участков поверхности Польши. Изначальные данные содержат 33 ортофотоснимка (снимок, снятый перпендикулярно поверхности Земли) с разрешением 25 сантиметров на пиксель и 8 ортофотоснимков с разрешением 50 сантиметров на пиксель.

Размер полной выборки после разбиения исходных данных и их аугментации: 68320 изображений. Все эти изображения были поделены на обучающую, валидационную и тестовую выборки в соотношении 0,7; 0,2 и 0,1 соответственно.

Условием остановки обучения было выбрано достижение каждым классом значения метрики IOU более 0,5. В результате обучения была получена нейронная сеть архитектуры U-net, показывающая достаточную эффективность

осуществления семантической сегментации. На рисунке 4.1 представлены результаты оценки метрики IOU по классам в момент окончания обучения.

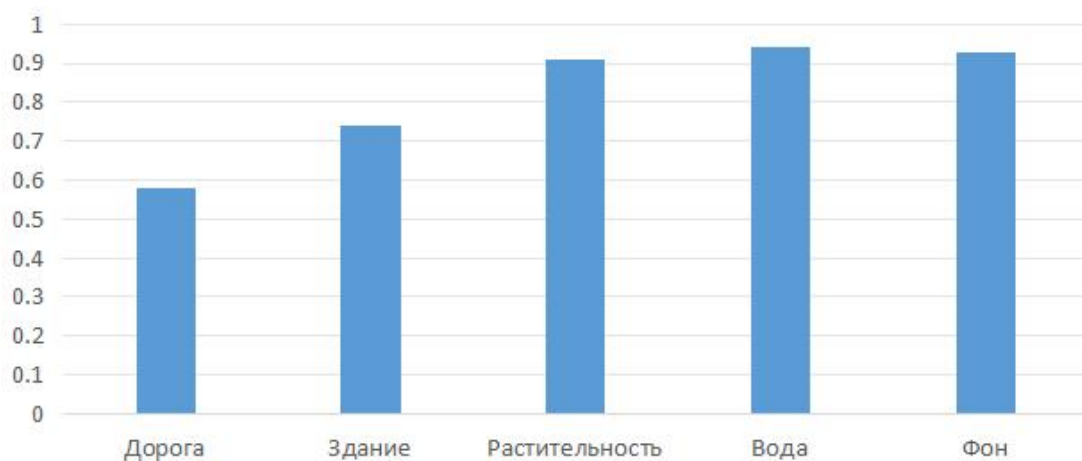
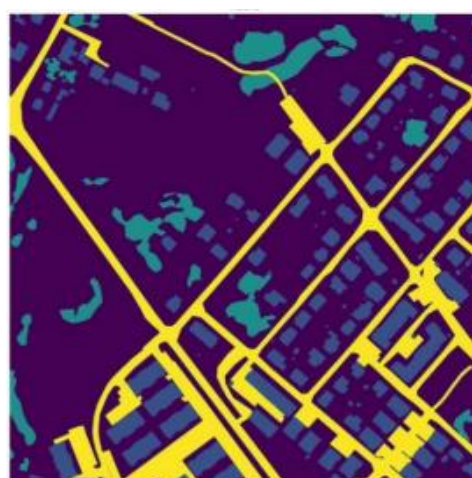


Рисунок 4.1 – Результаты обучения нейронной сети

На рисунке 4.2 а представлен пример аэрофотоснимка из обучающей выборки, а на рисунке 4.2 б – разметка к нему. На масочном изображении желтым отображены объекты класса «дорога», синим – «здание», зеленым – «растительность», фиолетовым – «фон».



а) Изображение

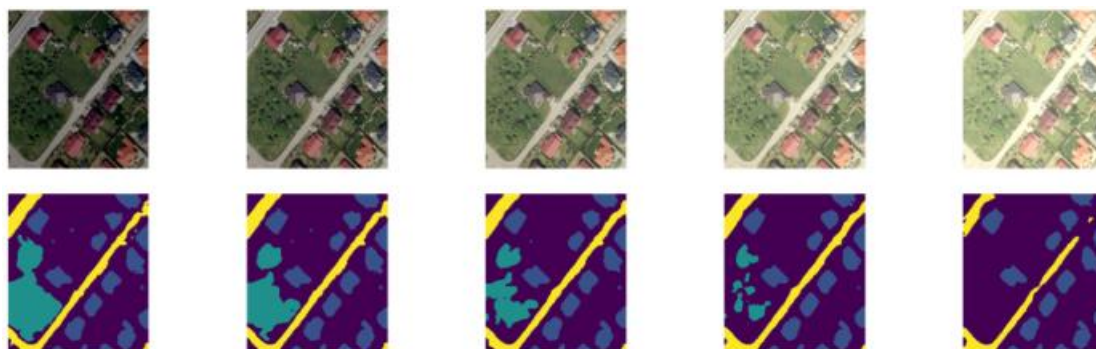


б) Разметка

Рисунок 4.2 – Аэрофотоснимок из обучающего набора и маска к нему

Оценка влияния изменений МО яркости (этап 2).

Влияние на качество работы нейронной сети путем варьирования МО яркости пикселей показано на рисунках 4.3 а, б, в, г, д.



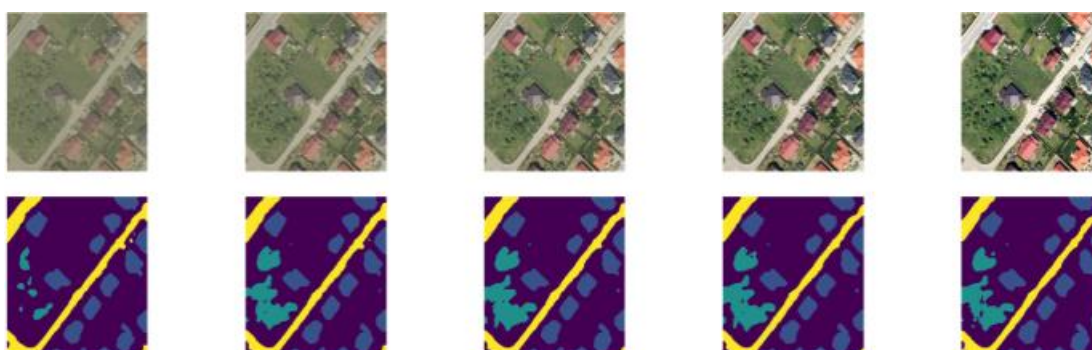
а) -50% МО б) -25% МО в) исходное МО г) +25% МО д) +50% МО

Рисунок 4.3 – Пример выхода нейронной сети, при изменении МО
исходного изображения

Изменения МО яркости приводит к деформации формы краев (границ) объектов. При этом может происходить слияние/разделение объектов, а также появление разрывов на объектах малой ширины, в частности, на дорогах.

Оценка влияния изменений СКО яркости (этап 3).

Влияние изменения СКО яркости на качество работы нейронной сети показано на рисунках 4.4 а, б, в, г, д.



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

Рисунок 4.4 – Пример выхода нейронной сети, при изменении СКО
исходного изображения

Изменения СКО яркости изображений в широких пределах в целом обеспечивает достаточно высокую устойчивость форм сегментированных объектов. При наименьших значениях СКО (низкой контрастности), соответствующих условиям высокой пасмурности, увеличиваются ошибки сегментации, связанные с потерей информативности краев слабоконтрастных объектов.

Оценка влияния аддитивного белого гауссовского шума (этап 4).

Влияние изменения отношения сигнала к шуму на качество работы нейронной сети показано на рисунках 4.5 а, б, в, г, д.

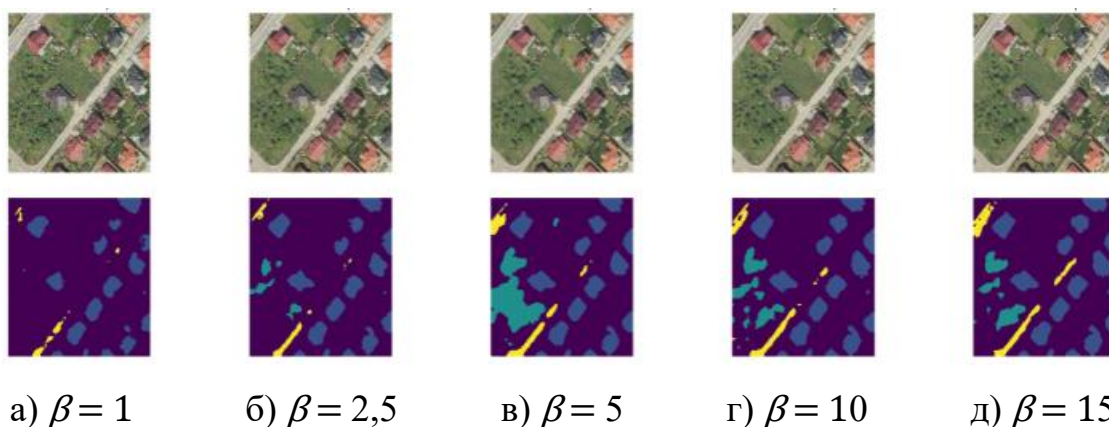


Рисунок 4.5 – Пример выхода нейронной сети, при изменении отношения сигнала к шуму на исходном изображении

Основное влияние высокочастотный аддитивный белый шум оказывает на края объектов. Возможно слияние близкорасположенных объектов и разделение объектов с тонкими соединениями. При этом объекты малого размера, в частности узкие дороги, могут терять целостность.

Анализ полученных результатов (этап 5).

Ниже представлены примеры графиков, иллюстрирующих влияние дестабилизирующих факторов на частоту правильных распознаваний пикселей при семантической сегментации изображений.

На рисунке 4.6 отображена зависимость частоты правильного распознавания (рассчитанная по формуле 4.3) класса «здание» от изменения МО яркости изображения.

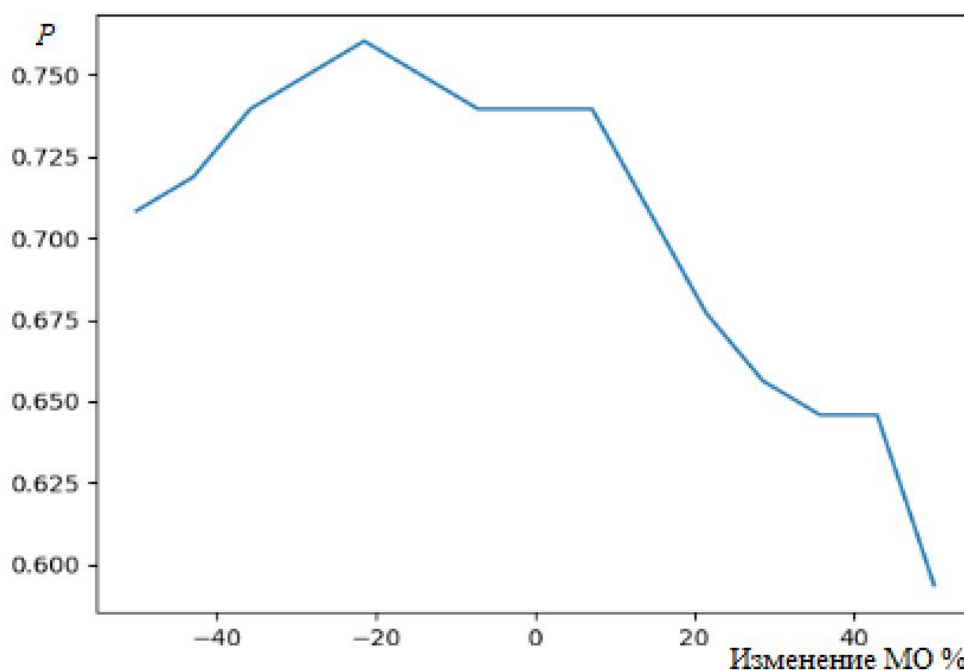


Рисунок 4.6 – Зависимость частоты верного определения здания от изменения МО

На рисунке 4.7 отображена зависимость частоты правильного распознавания класса «здание» от изменения СКО яркости изображения.

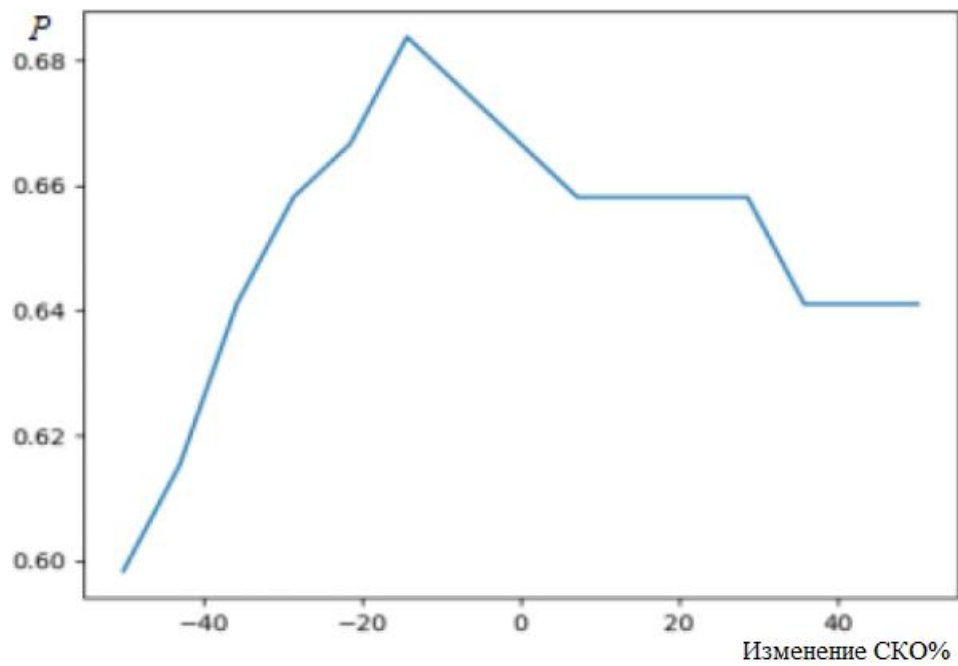


Рисунок 4.7 – Зависимость частоты верного определения здания от изменения СКО

На рисунке 4.8 отображена зависимость частоты правильного распознавания класса «здание» от изменения соотношения β .

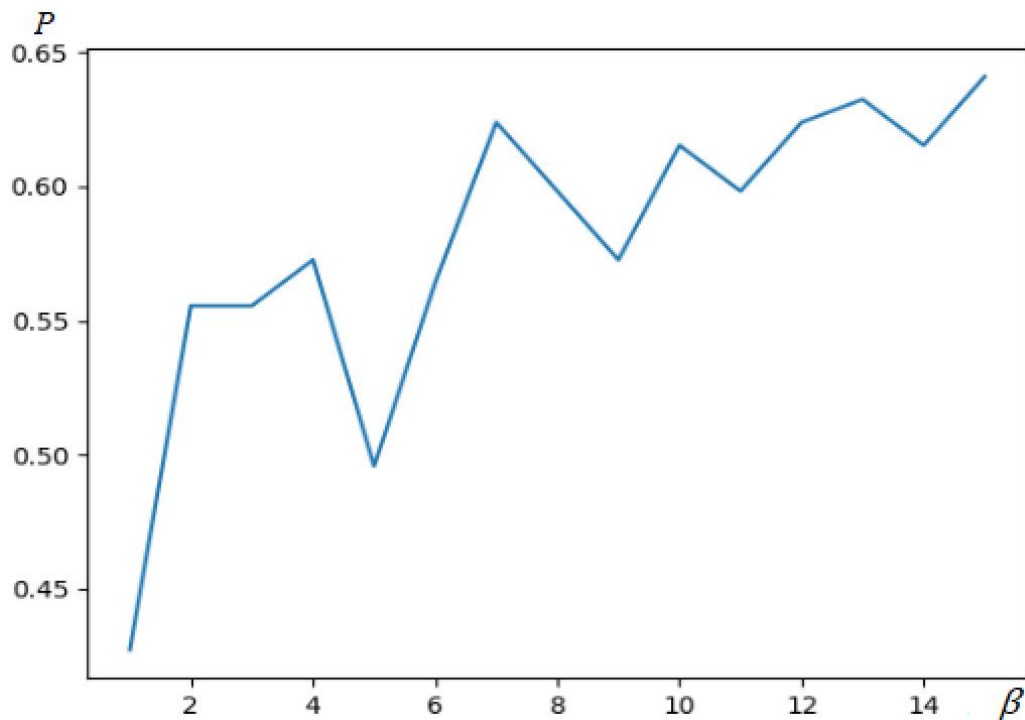


Рисунок 4.8 – Зависимость частоты верного определения здания от изменения соотношения β

В таблице 4.1 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 4.1 – Результаты обработки 1000 изображений

Класс	Частота правильного распознавания в обычных условиях (без модификации)	Частота правильного распознавания при изменении МО яркости на +50%	Частота правильного распознавания при изменении СКО яркости на -50%	Частота правильного распознавания при $\beta = 1$
Здание	0,74	0,623	0,634	0,432
Раститель- ность	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 проведенных исследований показывают, что:

1. Наилучшие результаты распознавания при сегментации пикселей (в обычных условиях и изменении МО) были получены для площадных объектов классов: вода, лес, фон. Классы «здания» и «дороги» обладают меньшей частотой правильного распознавания, однако показывают высокую устойчивость к изменениям условий наблюдения.

2. Влияние изменений СКО (контрастности) изображений на ухудшение распознавания объектов класса «вода» связано с уменьшением информативности высокочастотных признаков.

3. Относительно низкие значения частоты правильного распознавания объектов класса «дорога» связаны с малыми размерами объектов (шириной), соизмеримыми с размерами масок.

4. Влияние аддитивного белого гауссовского шума (при значениях $\beta \geq 3$) на частоту правильного распознавания несущественно.

Таким образом, результаты эксперимента подтверждают, что использование нейросетевых алгоритмов семантической сегментации обеспечивают устойчивость распознавания пикселей в широком диапазоне вариации освещенности.

4.2. Оценка эффективности идентификации изображений на основе использования их семантических описаний

Целью данного исследования является оценка эффективности предлагаемого подхода идентификации изображений (сцен), реализованного алгоритмом идентификации семантического описания ТИ (раздел 3.4).

В качестве метрики, оценивающей сходство семантических описаний сцен, используется МФД (3.4).

В качестве показателя эффективности рассматривается устойчивость алгоритма к изменению условий наблюдения и влиянию шумовых воздействий, оцениваемая достоверностью правильного распознавания (сопоставления) изображений (4.3).

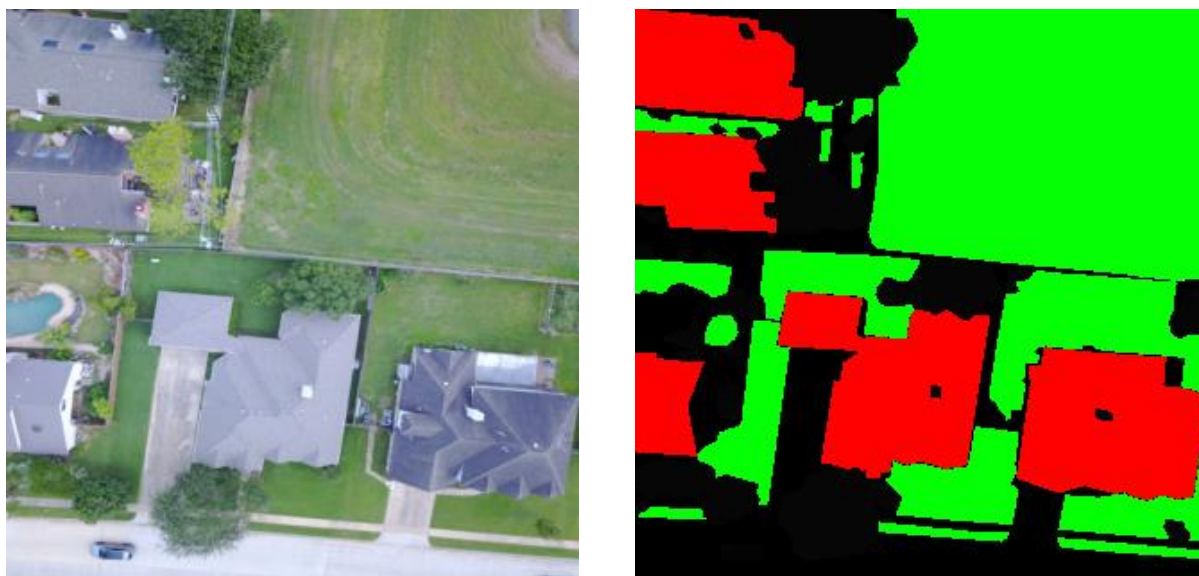
Проведено сравнение результатов идентификации ТИ и ЭИ на основе использования семантических и растровых описаний.

Методика эксперимента включает следующие этапы:

1. Моделирование ТИ с разными условиями освещенности.
2. Моделирование воздействия белого шума на ТИ.
3. Оценка достоверности правильного распознавания ТИ на разных этапах иерархического алгоритма с помощью МФД.
4. Оценка достоверности правильного распознавания ТИ при помощи нормированного коэффициента корреляции (НКК).
5. Анализ результатов.

В качестве ЭИ используется немодифицированное ТИ. На рисунке 4.9 а представлен вариант такого ЭИ, а на рисунке 4.9 б – разметка к нему. Красным

цветом обозначены объекта класса «здание», зеленым – «растительность», черным – «фон».



а) Изображение

б) Разметка

Рисунок 4.9 – Принятое ЭИ

Моделирование разной освещенности

Для моделирования изменений освещенности и контрастности изображений был использован механизм гамма-коррекции.

Гамма-коррекция цифрового изображения с помощью степенной функции позволяет изменить контрастность и яркость изображения. Данное преобразование может быть представлено следующим образом:

$$V_{\text{out}} = A * V_{\text{in}}^{\gamma}, \quad (4.5)$$

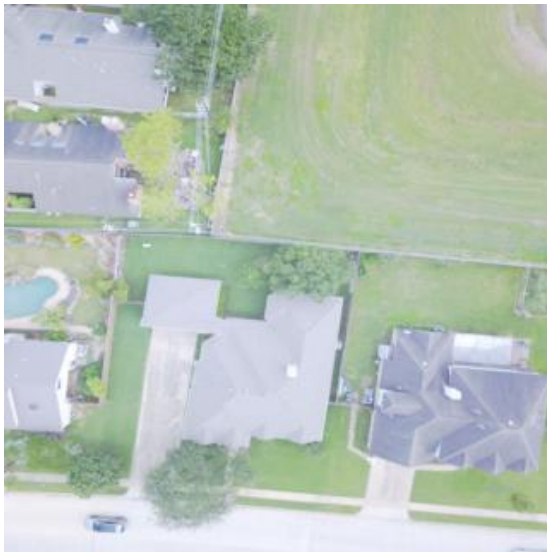
где V_{out} – выходная яркость пикселя;

A – коэффициент пропорциональности;

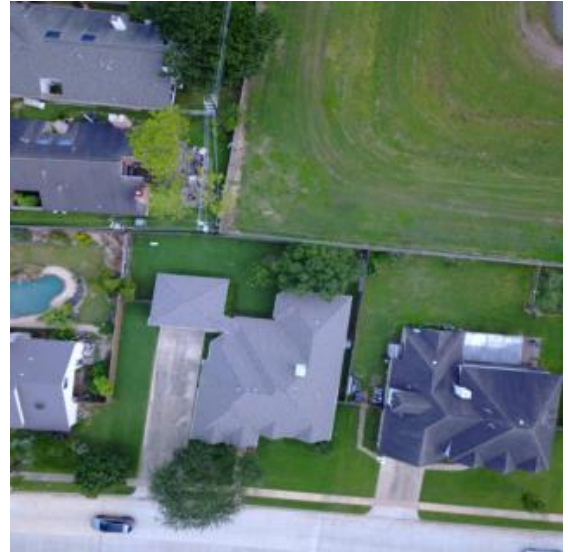
V_{in} – входная яркость пикселя в диапазоне от 0 до 1;

γ – параметр алгоритма.

На рисунках 4.10 а, б представлены результаты обработки изображения при помощи алгоритма гамма-коррекции с $\gamma = 0,5$ и $\gamma = 1,5$ соответственно.



а) $\gamma = 0,5$



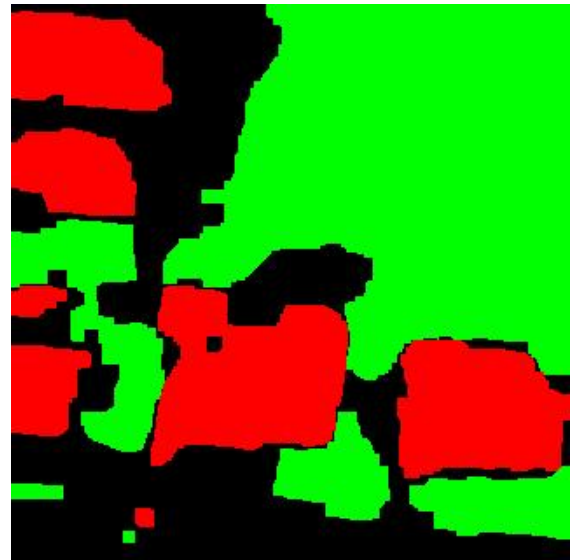
б) $\gamma = 1,5$

Рисунок 4.10 – Результат применения алгоритма гамма-коррекции

На рисунках 4.11 а, б представлены результаты семантической сегментации для соответствующих изображений с рисунка 4.10.



а) $\gamma = 0,5$



б) $\gamma = 1,5$

Рисунок 4.11 – Сегментированные изображения после гамма-коррекции

Графики зависимости значений критериальных функций сходства ЭИ и ТИ от коэффициента γ (изменение освещенности) ТИ представлен на рисунке 4.12.

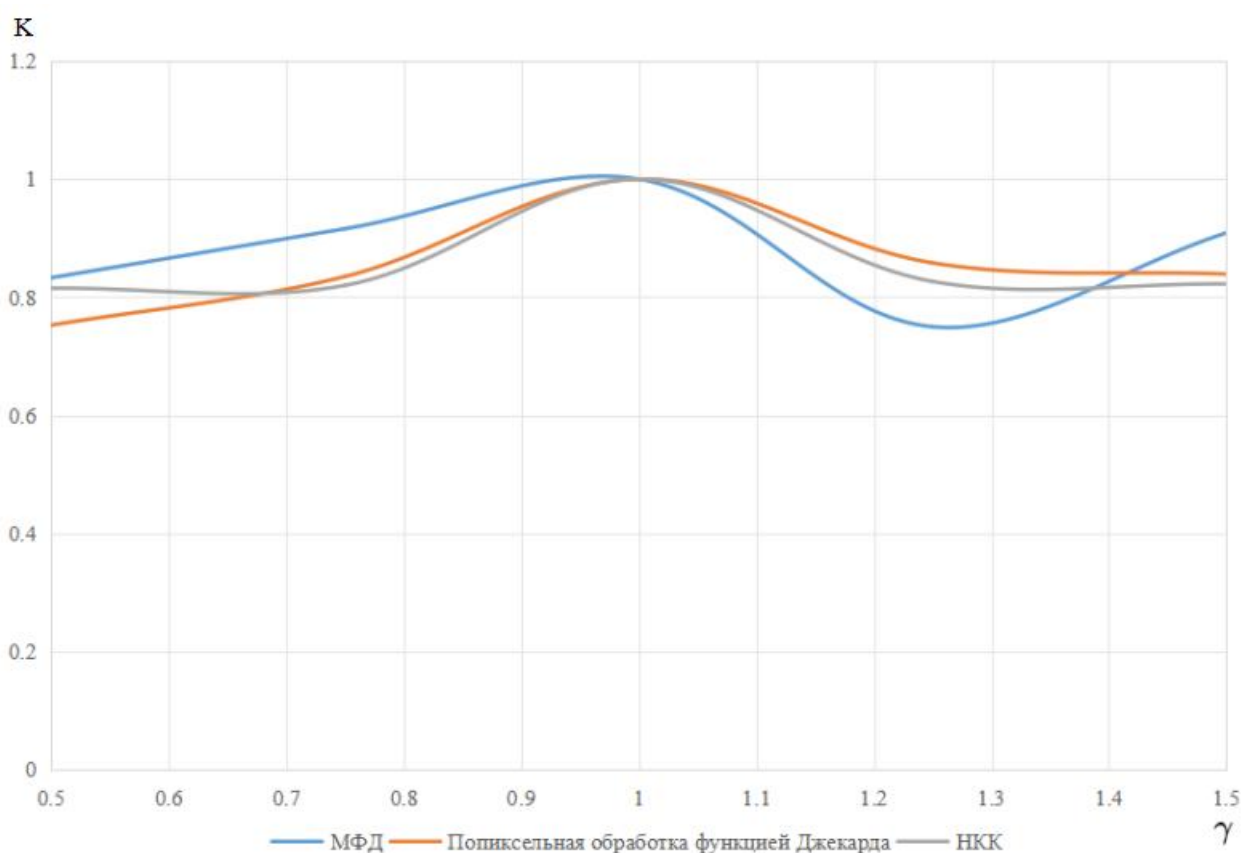


Рисунок 4.12 – Зависимость значений критериальных функций сходства ЭИ и ТИ от изменения освещенности

При использовании алгоритма МФД производится сопоставление по классам объектов, атрибутам и отношениям. Отметим, что максимальные значения рассматриваемых методов, соответствующие абсолютному сходству изображений, равны единице и достигаются при $\gamma = 1$.

Проанализировав график, отметим, что использование МФД для рассматриваемого ЭИ дает результаты, сопоставимые с результатами НКК. Спад достоверности при $\gamma = 1,25$ обусловлен недостаточным качеством обучения нейронной сети.

Моделирование воздействия белого шума на ТИ

При моделировании воздействия шума к каждому пикселю изображения прибавляется случайная величина, соответствующая воздействию аддитивного белого гауссовского шума, имеющая нормальный закон распределения:

$$\text{noise} \sim N(0, \sigma), \quad (4.6)$$

где noise – случайная величина, отображающая шумовое воздействие;

$N(0, \sigma)$ – нормальный закон распределения с нулевым средним и изменяемым значением СКО шума (σ).

На рисунке 4.13 изображено исходное текущее изображение и результат применения аддитивного гауссовского белого шума с СКО, соответствующим $\beta = 5$.



Рисунок 4.13 – Пример воздействия аддитивного гауссовского шума с СКО, соответствующим $\beta = 5$.

На рисунке 4.14 представлен результат семантической сегментации для зашумленного ТИ, представленного на рисунке 4.13.

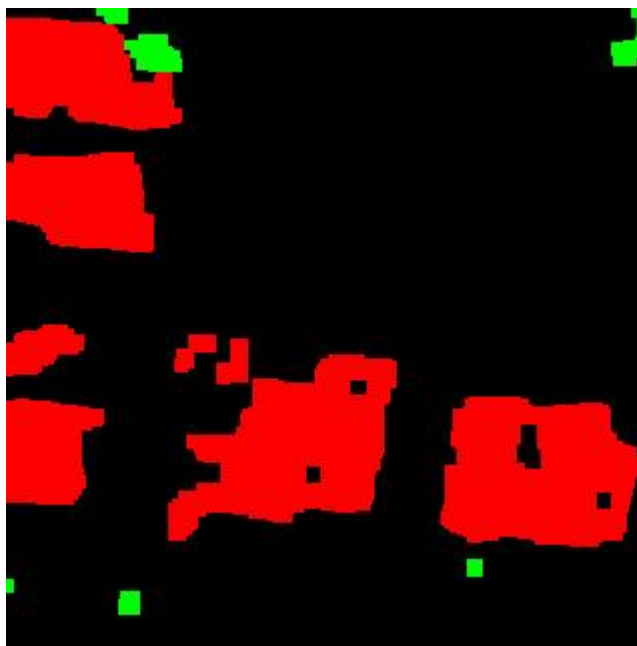


Рисунок 4.14 – Результат семантической сегментации зашумленного ТИ

Проанализировав рисунок 4.14, отметим практически полное отсутствие объектов класса «растительность». Однако объекты класса «здание» сохранили свои очертания и могут быть использованы для осуществления идентификации при помощи МФД.

Графики зависимости значений критериальных функций сходства ЭИ и ТИ от соотношения сигнал/шум представлены на рисунке 4.15.

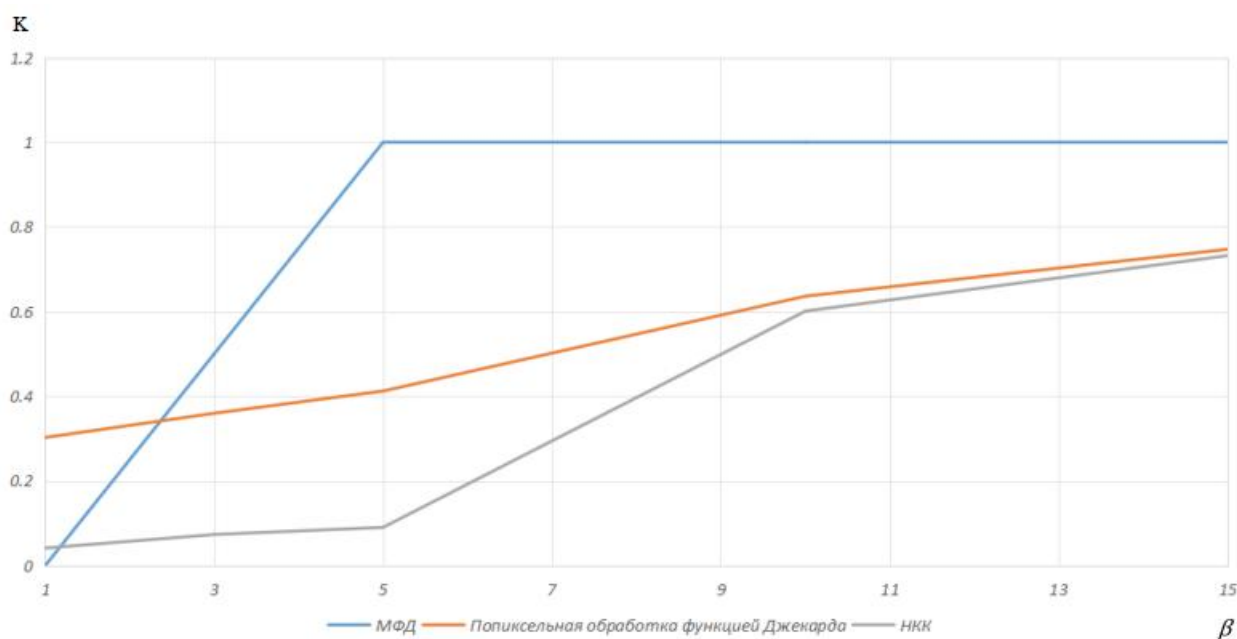


Рисунок 4.15 – Зависимость значений критериальных функций сходимости ЭИ и ТИ от β

Стоит отметить, что на графике значение $K = 1$ для МФД алгоритма при $\beta=5$ достигается за счет использования семантического описания классов «здание».

Анализ результатов

Полученные данные показывают, что использование семантических описаний позволяет существенно повысить устойчивость алгоритмов оценки сходства ТИ и ЭИ по сравнению с использованием растровых изображений.

Оценка сходства изображений по содержанию, реализуемая семантическими описаниями, уменьшает чувствительность МФД к искажениям (деформациям) изображений (рисунки 4.11, 4.14), связанным с воздействием дестабилизирующих факторов.

В то же время подобные искажения изображений влияют на значения функций, оценивающих сходство растровых изображений.

В таблице 4.2 представлены результаты оценки частоты правильного распознавания признаков (классов, атрибутов, отношений) при обработке 231 изображения (при гамма-коррекции с $\gamma = 0,5$ и $\gamma = 1,5$; а также при воздействии гауссовского шума с СКО, соответствующим $\beta = 5$).

Таблица 4.2 – Результаты оценки помехоустойчивости

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

Результаты эксперимента показали, что иерархический алгоритм, основанный на использовании МФД, практически не подвержен увеличению ошибки при 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 $\beta \geq 5$.

4.3. Оценка точности корреляционных алгоритмов сопоставления исходных растровых изображений и алгоритмов, использующих семантические описания

Целью данного исследования является сравнение точности локализации (оценки координат) ТИ существующих корреляционных алгоритмов обнаружения ориентиров с использованием эталонного изображения и предлагаемого алгоритма на основе семантических описаний.

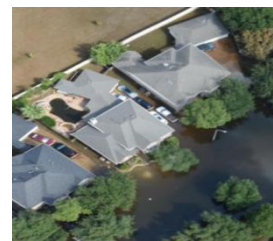
Методика эксперимента включает следующие этапы:

1. Составление набора изображений, использующихся как ЭИ.
2. Выделение на каждом ЭИ фрагментов, использующихся как ТИ.
3. Оценка точности сопоставления изображений при помощи МФД.
4. Оценка точности сопоставления исходных (несегментированных) изображений при помощи корреляционных алгоритмов.
5. Сравнение полученных результатов.

Пример 1. На рисунке 4.16 а представлен вариант ЭИ, а на рисунке 4.16 б – ТИ.



а) Эталонное изображение



б) Текущее изображение

Рисунок 4.16 – Первый пример

Положение ТИ соответствует верхнему левому участку ЭИ. На рисунке 4.17 а изображена разметка для ЭИ, 4.17 б – результат сегментации при помощи нейронной сети, 4.17 в – результат морфологической постобработки.

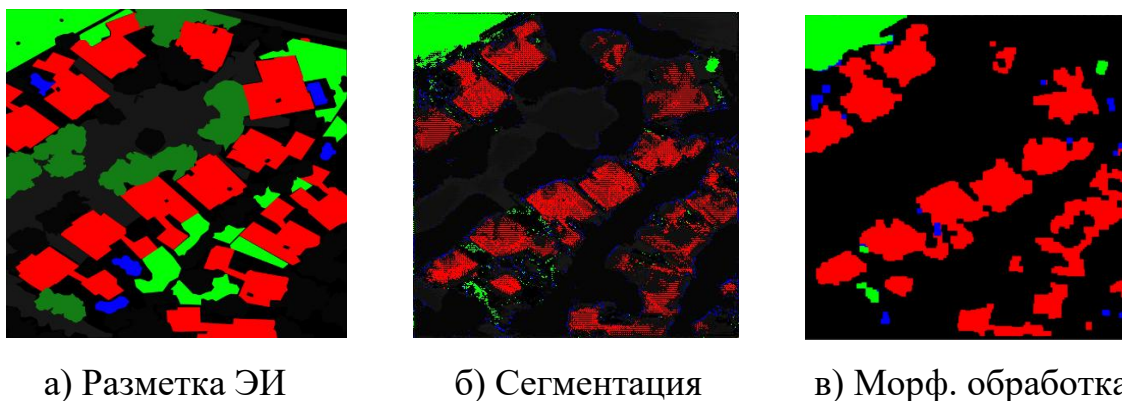


Рисунок 4.17 – Семантическая сегментация

Красным цветом обозначены объекты класса «здание», зеленым — «растительность», синим — «вода», белым — «дорога». Остальные цвета объединены в единый класс — «фон».

Проанализировав рисунок 4.17 б, отметим, что результат сегментации нейронной сети содержит множество неточностей по сравнению с исходной разметкой, так как в процессе сегментирования производится распознавание каждого пикселя в отдельности.

Результат морфологической постобработки обеспечивает получение более корректного описания сцены.

Так как обученная нейронная сеть лучше всего распознает пиксели растительности и зданий, далее используются только объекты данных классов.

Оценка работы МФД по семантическому описанию сцен

На рисунке 4.18 показан результат работы МФД (3.3), описание которой представлено в разделе 3.4:

$$E(F, m') = \frac{1}{L'} \sum_{l=1}^{L'} E(F_l, m'_l).$$

В качестве семантического описания использовались признаки вида класс, атрибут.

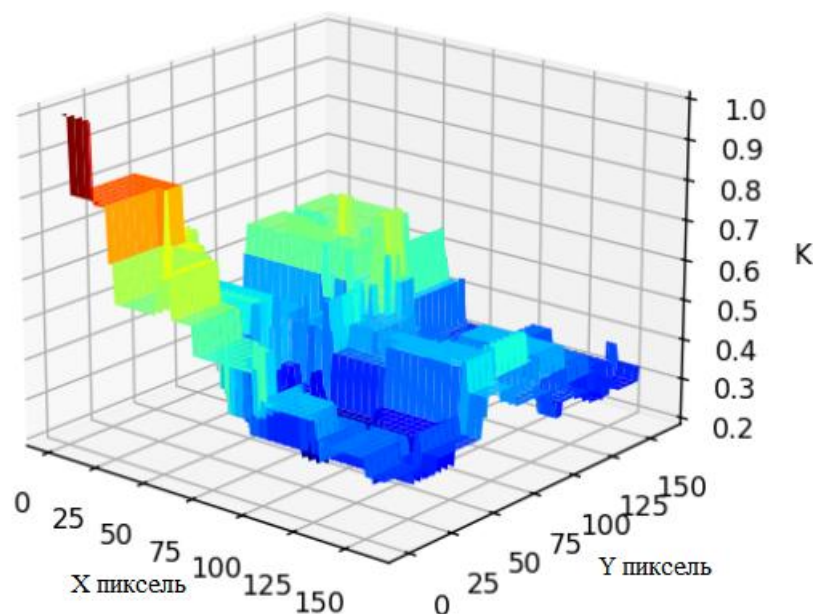


Рисунок 4.18 – Результат сопоставления ЭИ и ТИ МФД

Максимальное значение функция для оценки сходства семантических описаний на рисунке 4.18 соответствует истинному положению ТИ на ЭИ (рисунок 4.16). Следовательно, предлагаемый алгоритм (раздел 3.4) позволяет корректно определять координаты ТИ.

Сравнительная оценка работы алгоритмов идентификации на основе использования растровых изображений

Алгоритм с использованием семантических описаний сцен

В отличие от предыдущего эксперимента, здесь сегментированные изображения сравнивались как растровые.

В качестве основной метрики использовался расчет МФД по всем пикселям сегментированных изображений в отдельности.

Результаты расчетов представлены на рисунке 4.19.

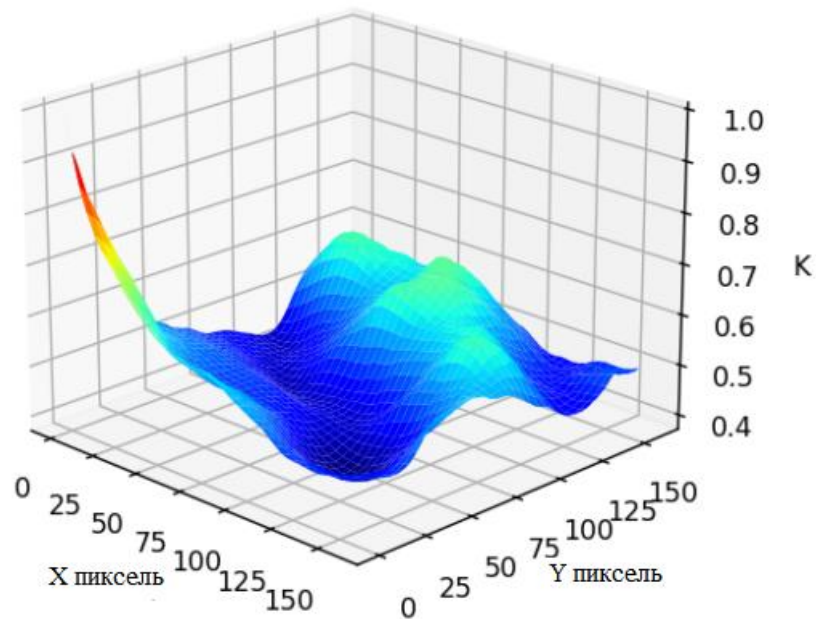


Рисунок 4.19 – Результат попиксельного сопоставления ЭИ и ТИ МФД

Оценка координат ТИ на ЭИ проведена успешно.

Алгоритм с использованием исходных растровых изображений

В качестве функции сопоставления использовалась НКК. Результаты расчетов показаны на рисунке 4.20.

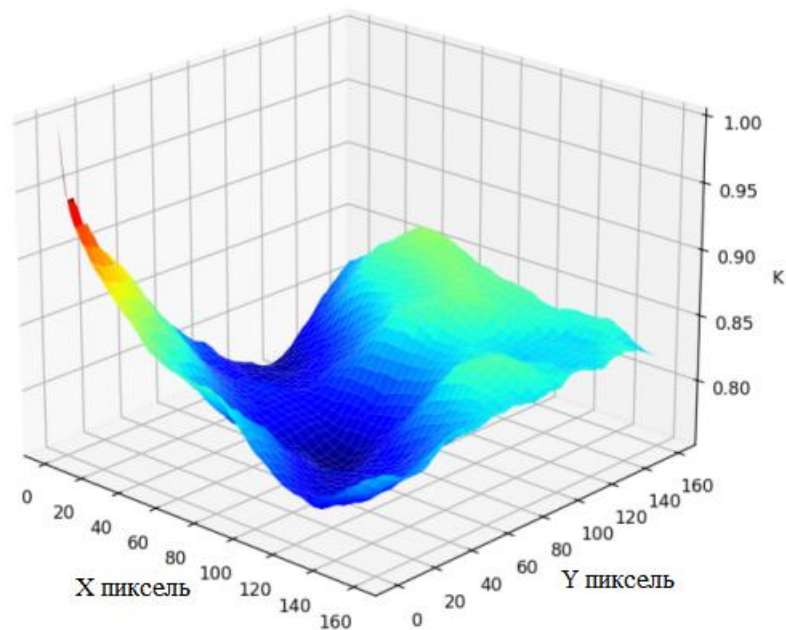
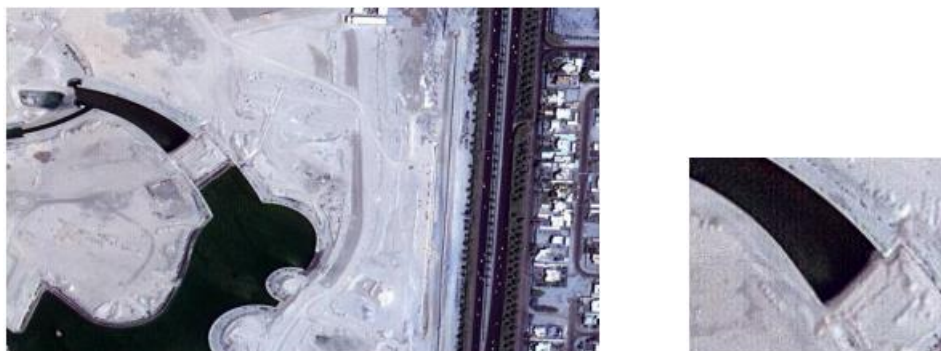


Рисунок 4.20 – Результат работы корреляционной функции

Полученные координаты ТИ совпадают с координатами, найденными в предыдущих экспериментах. Однако стоит отметить, что корреляция участков ЭИ, не соответствующих ТИ, не менее 0,75, что может приводить к ложному распознаванию в условиях наличия возмущающих воздействий.

Пример 2. Проведем эксперимент для ЭИ (рисунок 4.21 а) и для ТИ (рисунок 4.21 б).



а) Эталонное изображение

б) Текущее изображение

Рисунок 4.21 – Второй пример

Результаты сопоставления растровых изображений при помощи оценки семантических описаний МФД алгоритмом представлен на рисунке 4.22 а, при помощи НКК – на рисунке 4.22 б.

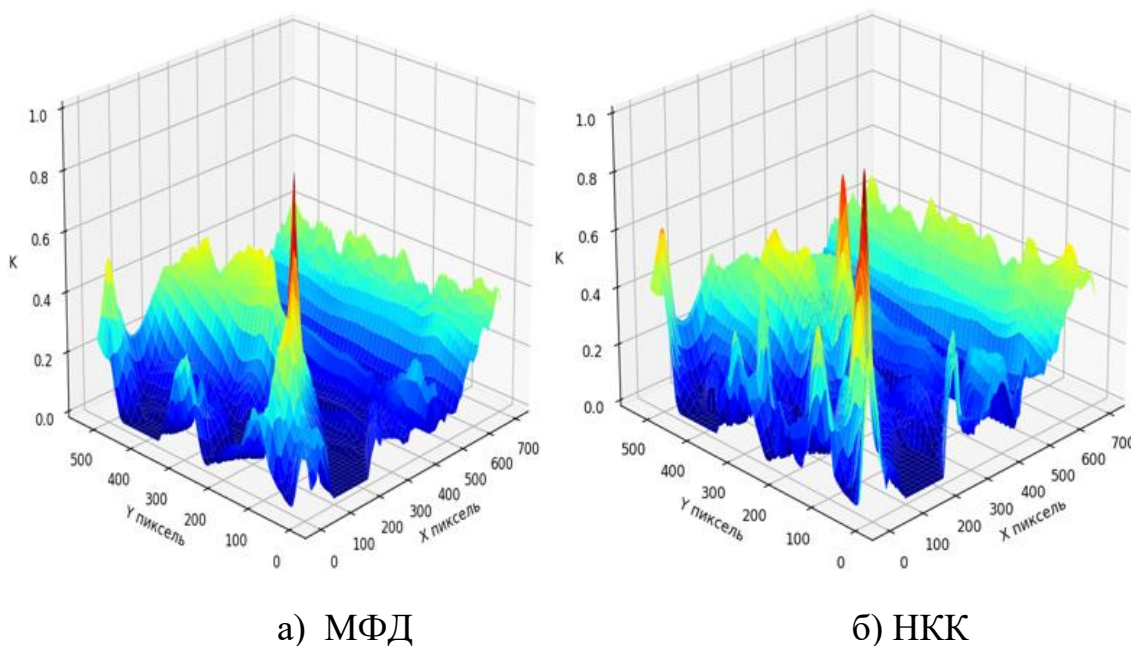


Рисунок 4.22 – Результаты работы алгоритмов сопоставления

Результаты исследований показали, что оба алгоритма решают задачу оценки положения ТИ на ЭИ при обработке растровых изображений с достаточно высокой и близкой степенью точности. Максимальные значения (экстремум) критериальных функций более 0,9.

Алгоритм, основанный на использовании сегментированных изображений МФД, оказался более чувствителен к наличию несоответствий между сопоставляемыми изображениями. Вторые локальные максимумы составляют:

- для МФД – 0,475;
- для НКК – 0,8.

Кроме того, более низкая чувствительность НКК в ряде случаев может привести к появлению близких локальных максимумов (рисунок 4.22), что в случае наличия оптических помех может привести к принятию ошибочных решений.

В таблице 4.3 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. Задаваемый порог

служит для определения, считать ли локальный экстремум, получаемый в результате обработки изображений, одним из возможных решений.

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

Таблица 4.3 – 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
Семантическое описание, объекты	0,672	0,0644%	0,824	0,0517%
Семантическое описание, атрибуты	0,838	0,0507%	0,874	0,0486%
Семантическое описание, отношения	0,794	0,0277%	0,846	0,0244%
Семантическое описание, иерархический алгоритм, МФД	0,848	0,0279%	0,848	0,0247%
НKK	0,5	0,0549%	0,5	0,0549%

Анализ результатов (таблица 4.3) показывает, что НКК дает единственное правильное решение в половине случаев, в то время как у МФД вероятность правильных решений существенно выше.

Наиболее информативным семантическим описанием оказалось семантическое описание атрибутов. При этом использование полного семантического описания позволяет еще больше увеличить вероятность решения поставленной задачи.

В целом, использование семантических описаний повышает точность алгоритма сопоставления изображений.

4.4. Оценка времени работы алгоритмов сопоставления сцен на основе семантических описаний и корреляционных алгоритмов

Целью данного исследования является сравнение вычислительных затрат предложенного подхода на основе семантических описаний и корреляционной функции на основе обработки растрового изображения.

Как было сказано в главе 3, каждый из этапов иерархического алгоритма может быть использован в отдельности для сокращения вычислений.

В данном случае рассматриваются три варианта ТИ, осуществляется поиск на ЭИ и оценивается количество найденных возможных вариантов решения навигационной задачи.

Рассмотрены варианты сопоставления сцен:

- по классам объектов, присутствующих на $SD_{Cl}(F^r)$ и $SD_{RI}(F^r, m^r)$;
- по атрибутам объектов;
- по отношениям между объектами;
- на основе сравнения по всем признакам.

В качестве корреляционной функции рассмотрен нормированный коэффициент корреляции (1.9).

Методика эксперимента включает следующие этапы:

1. Оценка сокращения возможных вариантов решения навигационной задачи на основе семантического описания классов.

2. Сравнение времени работы отдельных этапов с постоянными словарями признаков и корреляционной функции.

3. Сравнение времени работы иерархического алгоритма и корреляционной функции.

4. Анализ результатов.

Сокращение вычислительной сложности задачи сопоставления. В некоторых случаях нет возможности решить задачу сопоставления на основе семантических описаний вследствие недостаточной информативности наблюдаемой сцены. Тогда важным свойством алгоритма может выступать возможность его использования в качестве предварительного фильтра. На рисунке 4.23 представлен рассматриваемый вариант ЭИ.



Рисунок 4.23 – Вариант ЭИ

На рисунке 4.24 изображены исследуемые варианты ТИ.



Рисунок 4.24 – Варианты ТИ

В таблице 4.4 отображены результаты последовательного поиска рассматриваемых ТИ на ЭИ при помощи МФД. Изначально производится сравнение по наличию на фрагменте ЭИ требуемого списка объектов. Затем идет фильтрация по количественному совпадению конкретных классов объектов.

Таблица 4.4 – Сокращение возможных вариантов решения задачи сопоставления на основе семантического описания.

Показатель сходства	Количество фрагментов (штук / % от общего количество фрагментов на ЭИ)		
	Эксперимент 1	Эксперимент 2	Эксперимент 3
Сравнение состава объектов	741 / 16,05%	99 / 2,14%	325 / 7,04%
Сравнение объектов класса «Здание»	728 / 15,76%	98 / 2,14%	47 / 1,02%
Сравнение объектов класса «Дорога»	4 / 0,09%	53 / 1,15%	3 / 0,06%
Сравнение объектов класса «Растительность»	4 / 0,09%	50 / 1,08%	3 / 0,06%
Сравнение объектов класса «Фон»	4 / 0,09%	21 / 0,45%	3 / 0,06%
Сравнение объектов класса «Вода»	2 / 0,04%	6 / 0,13%	3 / 0,06%

Проанализировав таблицу 4.4, отметим, что при наилучших условиях использование МФД позволяет сократить последующие вычисления в 2500 раз, оставляя лишь два возможных варианта местоположения БЛА (эксперимент 1). При этом сравнение состава объектов позволяет в среднем сократить количество вычислений до 8%.

Рассмотрим время работы каждого этапа иерархического алгоритма в отдельности. Результаты представлены в таблице 4.5.

Таблица 4.5 – Время работы алгоритмов выделения и сопоставления семантического описания и алгоритма 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
Семантическое описание, объекты	0,753225	0,386727	0,132852	0,083849
Семантическое описание, атрибуты	2,219206	1,756542	0,853270	0,752762
Семантическое описание, отношения	5,521837	7,797435	8,890382	9,158547
Семантическое описание, объекты + атрибуты + отношения	10,014251	11,606667	10,794207	11,111801
НКК	23,711182	45,711301	22,559491	26,309172

Анализируя табличные данные, получим, что среднее время работы при использовании семантического описания сократилось примерно в 10 раз.

Иерархический алгоритм при использовании в совокупности всех этапов объединяет преимущества сокращения возможных вариантов и ускорение. Результаты работы представлены в таблице 4.6.

Таблица 4.6 – 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
Семантическое описание, объекты	0,753225	0,386727	0,132852	0,083849
Семантическое описание, атрибуты	0,177536	0,030938	0,010628	0,006707
Семантическое описание, отношения	0,035507	0,006817	0,00312	0,001521
Семантическое описание, иерархический алгоритм, МФД	0,966268	0,424482	0,1466	0,092077
НKK	23,711182	45,711301	22,559491	26,309172

Анализируя табличные данные, получим, что среднее время работы при использовании иерархического алгоритма сократилось более чем в 85 раз.

Сравнение результатов исследований показывает, что предлагаемый подход, основанный на использовании семантических описаний наблюдаемых сцен, позволяет сократить объем вычислений в среднем в 10 раз (в наилучшем случае – более чем в 85 раз).

4.5. Оценка информативности признаков при семантическом описании наблюдаемых сцен

Целью исследования является подтверждение возможности оценки информативности признаков при семантическом описании сцен.

Как было показано в разд. 1, общее количество полезной информации при распознавании M объектов или исходов, которое может быть получено при измерении всех значений признака U_n , равно (1.13):

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

где $P(m)$, $P(l)$, $P(i)$ – априорные вероятности возможных исходов с индексами m , l , i соответственно;

$P(U_{nk} | l)$ – условная вероятность появления признака U_{nk} при наличии l -го исхода;

$P(m) \log_2 P(m)$ – частная энтропия исхода m ;

$P(m | U_{nk}) \log_2 P(m | U_{nk})$ – частная энтропия, апостериорной вероятности нахождения в точке m при условии получения признака U_{nk} .

В рассматриваемом подходе признаком U_n , позволяющим определить положение БЛА (исход m), является наблюдаемая сцена.

При семантических описаниях сцен в зависимости от словаря могут использоваться различные локальные признаки – классы объектов, атрибуты и отношения, определяющие особенности каждой сцены.

Таким образом, задача перехода к использованию семантических описаний при расчете информативности I_n наблюдаемой сцены (или признака U_n) состоит в оценке условной вероятности $P(U_{nk} | l)$ для используемых семантических признаков.

Суть проводимых экспериментов состоит в повышении вероятности решения задачи визуальной навигации, за счет планирования маршрута. При этом рассматриваются участки маршрута, на которых может отсутствовать навигация с помощью ГНСС.

Если ситуация с отсутствием сигналов ГНСС возникает в процессе полета БЛА, то реализуется оперативное планирование выбора локального участка маршрута.

Конечной точкой планируемого маршрута может быть точка, в которой происходит минимизация текущей энтропии процесса или возобновление решения навигационной задачи при помощи ГНСС.

Методика эксперимента включает следующие этапы:

1. Выбор участка местности на ЦКМ, соответствующего области предполагаемого полета.
2. Оценка СКО ошибки оценки координат инерциальных датчиков БЛА по формулам (1.1), (1.2).
3. Составление списка информативных областей.
4. Определение точки старта и точки конца маршрута.
5. Построение вариантов маршрута, проходящих через информативные области, например, таких как просветы в облачности, пересечения дорог, мосты, одиночные здания и пр.
6. Оценка информативности по формуле (1.10) каждого из вариантов маршрута.
7. Выбор наиболее информативного варианта маршрута.

Необходимо отметить, что в ряде случаев по объективным причинам не удастся обеспечить выполнение требований (1.4), (1.5). В подобных ситуациях предлагается принимать решения по управлению, минимизирующим конечную энтропию процесса.

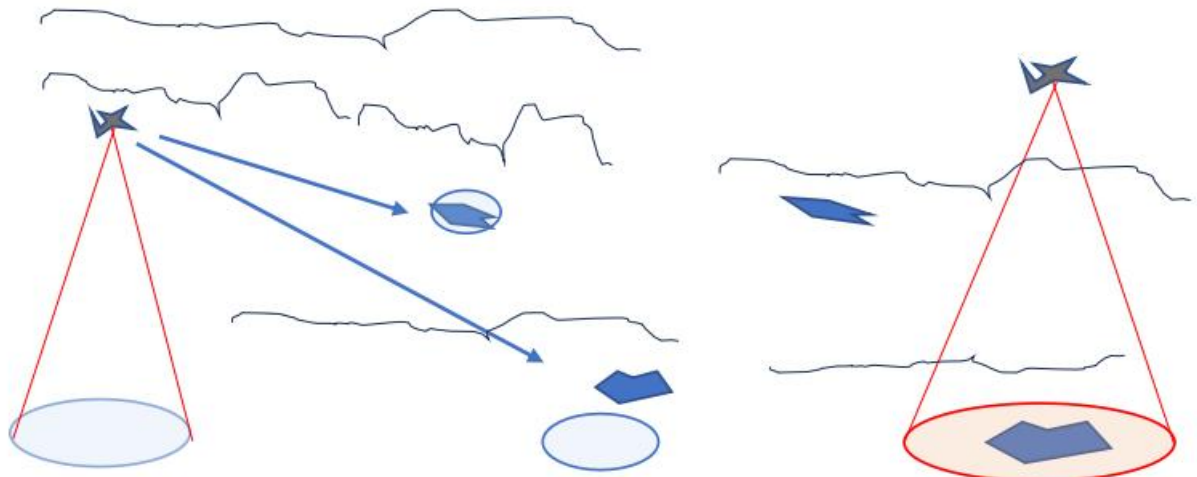
Оценка работоспособности алгоритма в условиях облачности

Рассмотрим ситуацию, когда БЛА с двумя видеокамерами летит над облаками с просветами. Первая камера направлена вперед-вниз и используется для поиска просветов. Вторая камера направлена вертикально вниз для осуществления привязки к бортовой карте местности.

Требуется выбрать локальный маршрут БЛА из N возможных, проходящий над просветом в облачности. Предлагается выбирать наилучший маршрут по максимальному значению информативности при помощи расчета энтропии.

На рисунке 4.25 а представлен БЛА, оснащенный двумя камерами, направленными: 1-я камера – вертикально вниз и 2-я – вперед-вниз. При этом 2-я камера обнаруживает просветы в облачности.

При пролете над выбранным просветом 1-я камера принимает ТИ, по которому уточняется положение БЛА (рисунок. 4.25 б).



а) Сканирование облачности

б) Получение ТИ

Рисунок 4.25 – Поиск просветов в облачности

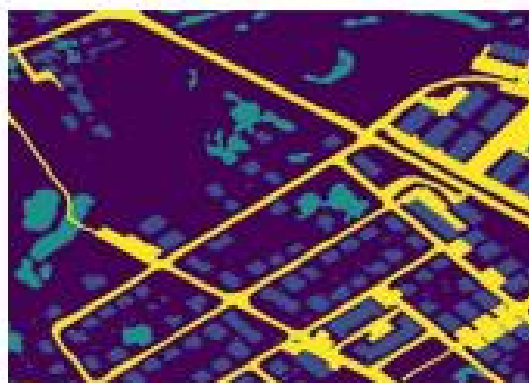
На рисунке 4.26 показаны: а) Просветы в облачности; б) Карта местности области полета БЛА (ЭИ); в) Сегментированное изображение ЭИ. На сегментированном изображении белой штриховой и точечной линиями показаны границы просветов облачности.



а) Просветы в облачности



б) ЭИ



в) Сегментированное изображение

Рисунок 4.26 – Область полета БЛА

Просвет 1 (штриховая линия) может находиться в любой точке левой половины ЭИ. Просвет 2 (точечная линия) располагается в правой половине ЭИ.

На карте местности возможно присутствие четырех классов объектов: дорога, здание, растительность, фон.

При переходе к семантическим описаниям условной вероятности $P(U_{nk} | I)$ в (1.10) заменяется условной частотой появления признака U_{nk} , рассчитываемой для каждого класса объектов.

В соответствии с количеством просветов рассматриваются два варианта маршрутов, проходящих над каждым просветом.

Требуется выбрать наиболее информативный маршрут БЛА.

Информативность маршрута I_n показывает, насколько уменьшается начальная неопределенность (энтропия) положения БЛА.

Априорные вероятности $P(m)$ появления в просвете m -го участка определяются из известной плотности распределения ошибок ИНС $p(x, y, t)$ (1.2). Примем, что распределение ошибок представляет собой равномерный закон с вероятностью реализации $P(m) = 0,11$.

Апостериорные вероятности $P(m|U_{nk})$ определяются по формуле Байеса при предположении возможного получения информации U_{nk} о присутствии k -го ориентира. Вероятность $P(m|U_{nk})$ определяет уникальность наблюдаемых сцен на каждом маршруте.

Усредненные результаты моделирования, рассчитанные для разных положений просветов на ЭИ, приведены в таблице 4.7. Оценка энтропии и информативности представлена в битах.

Таблица 4.7 – 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

Приведенные в таблице 4.7 расчеты показывают, что предлагаемый подход позволяет оценивать информативность семантических описаний разных наблюдаемых сцен. Маршрут 1 соответствует малоинформативному описанию вида: < здание, здание, здание, фон >, а маршрут 2 представляет описание вида < здание, здание, здание, здание, дорога, дорога, фон, фон >.

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

ЗАКЛЮЧЕНИЕ

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 гауссовских 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 СОКРАЩЕНИЙ

FCN	- Fully Convolutional Network
FN	- false negative
FP	- false positive
IoU	- Intersection over Union
TP	- true positive
YOLO	- You Only Look Once
БЛА	- Беспилотный летательный аппарат
ГНСС	- глобальная навигационная спутниковая система
ИНС	- Инерциальная навигационная система
КПМ	- конечный пункт маршрута
КЭНС	- корреляционно-экстремальная навигационная система
МО	- математическое ожидание
МФД	- модифицированная функция Джекарда
НKK	- нормированный коэффициент корреляции
НПМ	- начальный пункт маршрута
НС	- навигационная система
НСУ	- наземная станция управления
СКО	- среднеквадратическое отклонение
СН	- система наблюдения
СНС	- сверточная нейронная сеть
СТЗ	- система технического зрения
ТИ	- текущее изображение
ЦЗ	- целевая задача
ЦКМ	- цифровая карта местности
ЭИ	- эталонное изображение

СПИСОК ЛИТЕРАТУРЫ

1. Alchimowicz J. Semantic_segmentation_satellite_imagery [Электронный ресурс] : collection / J. Alchimowicz // Figshare. – 2022. – DOI: 10.6084/m9.figshare.c.6026765.v1.
2. Boguszewski, Adrian, Batorski Dominik, Ziemba-Jankowska Natalia, Zambrzycka Anna, Dziedzic Tomasz. LandCover.ai : Dataset for Automatic Mapping of Buildings, Woodlands and Water from Aerial Imagery. – 2020. – DOI: 10.48550/arXiv.2005.02264.
3. Cordts M., Omran M., Ramos S., Rehfeld T., Enzweiler M., Benenson R., Franke U., Roth S., Schiele B. The Cityscapes Dataset for Semantic Urban Scene Understanding [Электронный ресурс] // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – DOI: 10.48550/arXiv.1604.01685.
4. Dalal, N., Triggs B. Histograms of oriented gradients for human detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). – 2005. – Vol. 1. – Pp. 886-893 – DOI: 10.1109/CVPR.2005.177.
5. Endsley, M., Daniel J. Garland. Theoretical Underpinning of Situation Awareness: Critical Review // Situation Awareness Analysis and Measurement. – 2000. – P. 3-32.
6. Geng, K., Chulin A. UAV Navigation Algorithm Based on Improved Algorithm of Simultaneous Localization and Mapping with Adaptive Local Range of Observations [Электронный ресурс] // Herald of the Bauman Moscow State Technical University. Series Instrument Engineering. – 2017. – URL: <https://doi.org/10.18698/0236-3933-2017-3-76-94> (дата обращения: 27.05.2025).
7. 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.

8. 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.
9. 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.
10. Humans in the Loop. Semantic segmentation of aerial imagery (CC0: Public Domain) [Электронный ресурс] / Humans in the Loop // Kaggle. – 2024. – URL: <https://www.kaggle.com/datasets/humansintheloop/semantic-segmentation-of-aerial-imagery> (дата обращения: 02.09.2024).
11. Kim, N., Bodunkov N. Automated Decision Making in Road Traffic Monitoring by On-board Unmanned Aerial Vehicle System // Computer Vision in Control Systems-3. – 2018. – P. 149-175.
12. Kirillov, A., et al. Panoptic segmentation // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – P. 9404-9413.
13. Pazychev, D., Sadekov R. Simulation of INS Errors of Various Accuracy Classes. 1-3 [Электронный ресурс]. – 2020. – URL: <https://doi.org/10.23919/ICINS43215.2020.9133869>.
14. Plane architecture overview [Электронный ресурс] / ArduPilot Dev Team // ArduPilot Documentation. – URL: <https://ardupilot.org/dev/docs/plane-architecture.html> (дата обращения: 02.09.2024).
15. PX4 Architecture Overview [Электронный ресурс] / PX4 Development Team // PX4 Autopilot User Guide. – URL: <https://docs.px4.io/main/en/concept/architecture.html> (дата обращения: 02.09.2024).

16. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention. – Cham : Springer international publishing, 2015. – P. 234-241.
17. Sabhara, R. K., Lee C. P., Lim K. M. Comparative study of hu moments and zernike moments in object recognition // SmartCR. – 2013. – Vol. 3. – no. 3. – P. 166-173.
18. Salamati, N., Larlus D., Csurka G., Süssstrunk S. Semantic Image Segmentation Using Visible and Near-Infrared Channels. In: Fusiello, A., Murino, V., Cucchiara, R. (eds) Computer Vision – ECCV 2012. Workshops and Demonstrations. ECCV 2012 [Электронный ресурс] // Lecture Notes in Computer Science. – 2012. – Vol. 7584. – Springer, Berlin, Heidelberg. – URL: https://doi.org/10.1007/978-3-642-33868-7_46.
19. Shannon, C., Weaver W. The Mathematical Theory of Communication // Urbana. – Univ. of Illinois Press, 1949.
20. Soille, P. Morphological Image Analysis: Principles and Applications / P. Soille. – 2nd ed. – Berlin, Heidelberg : Springer-Verlag, 2003. – 391 p. : ill. – (Texts in Computer Science). – ISBN 978-3-642-07696-1. – DOI: 10.1007/978-3-662-05088-0.
21. Абросимов В., Хейден М. Имитационная модель формирования ситуационной осведомленности в группе автономных роботов в условиях потенциальных угроз / В. Абросимов // Известия ЮФУ. Инженерные науки. – Таганрог, 2010. – № 3 (104). – С. 14-20.
22. Абчук, В. А. Поиск объектов / В. А. Абчук, В. Г. Суздаль. – М. : Советское радио, 1977. – 336 с.
23. Али, Б., Садеков Р.Н., Цодокова В.В.. Алгоритмы навигации беспилотных летательных аппаратов с использованием систем технического зрения // Гироскопия и навигация. – 2022. – Т. 30. – № 4 (119). – С. 87-105.

24. Баклицкий, В.К. и др. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации / В.К. Баклицкий, А.М. Бочкарев, М.П. Мусьяков; под ред. В.К. Баклицкого – М. : Радио и связь, 1986 – 216 с., ил.
25. Балабаев О. С., Прохорцов А. В. Методика начальной выставки и коррекции БИНС ЛА в полёте // Известия ТулГУ. Технические науки. – 2025. – № 1. – URL: <https://cyberleninka.ru/article/n/metodika-nachalnoy-vystavki-i-korreksii-bins-la-v-polyote> (дата обращения: 18.08.2025).
26. Белоглазов, И.Н., Джангджава Г.И., Чигин Г.П. Основы навигации по геофизическим полям. – М. : Наука, Главная редакция физико-математической литературы, 1985. – 327 с.
27. 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
28. 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.
29. Веремеенко, К.К. Современные информационные технологии в задачах навигации и наведения беспилотных маневренных летательных аппаратов / К. К. Веремеенко, С. Ю. Желтов, Н. В. Ким и др. // Под ред. М. Н. Красильщикова, Г. Г. Себрякова. – М. : ФИЗМАТЛИТ, 2009. – 556 с.
30. Визильтер, Ю. В. Обработка и анализ изображений в задачах машинного зрения / Ю. В. Визильтер, С. Ю. Желтов, А. В. Бондаренко, М. В. Ососков, А. В. Моржин. – М. : Физматкнига, 2010. – 672 с. – ISBN 978-5-89155-201-3.
31. Гайдук, А. Р., Капустян С. Г., Дьяченко А. А., Плаксиенко Е. А. Автономное осуществление миссий БЛА [Электронный ресурс] // Известия ЮФУ. Технические науки. – 2017. – № 1 (186). – URL: <https://cyberleninka.ru/article/n/avtonomnoe-osuschestvlenie-missiy-bla> (дата обращения: 27.08.2025).

32. 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.
33. 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.
34. 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.
35. Гиренко Д. С. Оценка поведения людей на основе анализа видеoinформации / Д. С. Гиренко, Н. В. Ким // Сб. докладов с международной конф. по науке и технологиям «Россия – Корея – СНГ». – М., 2023. – С. 73-77.
36. 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.
37. 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 – 148 с.
38. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс ; пер. с англ. под ред. П. А. Чочиа. – 3-е изд. – М. : Техносфера, 2012. – 1104 с. : ил. – (Мир цифровой обработки). – ISBN 978-5-94836-300-3.
39. Горелик, А. Л. Методы распознавания : учебное пособие для вузов / А. Л. Горелик, В. А. Скрипкин. – М. : Высшая школа, 1984. – 232 с. : ил. – Библиогр.: с. 227-228. – Предм. указ.: с. 229-230.

40. Гороховатский В. А., Передрий Е. О. Корреляционные методы распознавания изображений путем голосования систем фрагментов // Радиоэлектроника, информатика, управління. – 2009. – №1 (20). URL: <https://cyberleninka.ru/article/n/korrelyatsionnye-metody-raspoznavaniya-izobrazheniy-putem-golosovaniya-sistem-fragmentov> (дата обращения: 16.09.2025).
41. ГОСТ Р. Беспилотные авиационные системы. Технические средства обучения внешнего экипажа. Общие требования. – М., 2024. – 24 с.
42. Друки А. А., Спицын В. Г., Аркалыков Е. У. Алгоритмы семантической сегментации снимков земной поверхности на основе нейронных сетей // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2020. – № 51. – URL: <https://cyberleninka.ru/article/n/algoritmy-semanticheskoy-segmentatsii-snimkov-zemnoy-poverhnosti-na-osnove-neyronnyh-setey> (дата обращения: 18.08.2025).
43. Еремеев, С. В., Абакумов А. В. Классификация объектов на изображениях с учетом искажений на основе двухэтапного топологического анализа [Электронный ресурс] // Научно-технический вестник информационных технологий, механики и оптики. – 2022. – № 1. – URL: <https://cyberleninka.ru/article/n/klassifikatsiya-obektov-na-izobrazheniyah-s-uchetom-iskazheniy-na-osnove-dvuhetapnogo-topologicheskogo-analiza> (дата обращения: 18.09.2024).
44. Захаров А. А., Жизняков А. Л., Титов В. С. Метод нахождения соответствий на изображениях с использованием структур дескрипторов // Компьютерная оптика. – 2019. – №5. – С. 810-817 URL: <https://cyberleninka.ru/article/n/metod-nahozhdeniya-sootvetstviy-na-izobrazheniyah-s-ispolzovaniem-struktur-deskriptorov> (дата обращения: 16.09.2025).

45. Камаева А. А., Тепаева Ю. Е., Ротанов А. С., Васькин В. А., Ямашкин С. А. Анализ методов и алгоритмов семантической сегментации графических данных // E-Scio. – 2022. – №11 (74). – URL: <https://cyberleninka.ru/article/n/analiz-metodov-i-algoritmov-semanticheskoy-segmentatsii-graficheskikh-dannyh> (дата обращения: 16.09.2025).
46. Ким, Н. В., Бодунков Н. Е., Михайлов Н. А. Энтропийный подход к решению задачи поиска наземных объектов с учетом фактора наблюдаемости [Электронный ресурс] / Н. Ким // Вестник компьютерных и информационных технологий. – 2019. – С. 22-29. – DOI: 10.14489/vkit.2019.01.pp.022-029 (дата обращения: 27.06.2025).
47. Ким Н. В., Червоненкис М. А. Ситуационный контроль беспилотных летательных аппаратов для мониторинга дорожного движения / Н. Ким // Современная прикладная наука. – 2015. – № 9 (5). – С. 1-13 – DOI: 10.5539/mas.v9n5p1
48. Ким, Н. В. Обработка и анализ изображений в системах технического зрения : учебное пособие. – М. : Изд-во МАИ. – 2001. – 164 с.
49. Кирпичников А. П., Мифтахутдинов Д. И., Ризаев И. С. Решение задачи корреляционной привязки изображения и цифровой карты местности // Вестник Казанского технологического университета. – 2015. – №17. URL: <https://cyberleninka.ru/article/n/reshenie-zadachi-korrelyatsionnoy-privyazki-izobrazheniya-i-tsifrovoy-karty-mestnosti> (дата обращения: 16.09.2025).
50. Команда Лаборатории приборов. Оценка ошибки вычисления навигации инерциальными модулями [Электронный ресурс] // Сайт Лаборатории приборов. – 2021. – URL: <https://mp-lab.ru/navigation-error-estimation/> (дата обращения: 21.08.2024).
51. Красовский, А. А. Теория корреляционно-экстремальных навигационных систем / А. А. Красовский, И. Н. Белоглазов, Г. П. Чигин. – М. : Наука, 1979. – 424 с.

52. Лукашик, Д. В. Анализ современных методов сегментации изображений [Электронный ресурс] // Экономика и качество систем связи. – 2022. – № 2 (24). – URL: <https://cyberleninka.ru/article/n/analiz-sovremennyh-metodov-segmentatsii-izobrazheniy> (дата обращения: 18.09.2024).
53. Нистратов, А. А., Семина Т. В., Тарасов Е. Ф. Семантические характеристики восприятия изображений и их описаний [Электронный ресурс] // Вопросы психолингвистики. – 2016. – № 3 (29). – URL: <https://cyberleninka.ru/article/n/semanticheskie-harakteristiki-vozpriyatiya-izobrazheniy-i-ih-opisaniy> (дата обращения: 18.09.2024).
54. 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.
55. Рябых М. С., Сойникова Е. С., Батищев Д. С., Синюк В. Г., Михелев В. М. Высокопроизводительный метод анализа и морфологической обработки изображений // Научный результат. Информационные технологии. – 2016. – № 3. – URL: <https://cyberleninka.ru/article/n/vysokoproizvoditelnyy-metod-analiza-i-morfologicheskoy-obrabotki-izobrazheniy> (дата обращения: 18.08.2025).
56. Самарина О. В. Топологически-семантический анализ изображения [Электронный ресурс] // Вестник ЮГУ. – 2017. – № 3 (46). – URL: <https://cyberleninka.ru/article/n/topologicheskii-semanticheskii-analiz-izobrazheniya> (дата обращения: 18.09.2024).
57. Семенова Л. Л. Современные методы навигации беспилотных летательных аппаратов // Наука и образование сегодня. – 2018. – № 4 (27) – С. 6-8.

58. Каздорф С. Я., Першина Ж. С. Алгоритм семантической сегментации трехмерных сцен // Cloud of science. – 2019. – №3. URL: <https://cyberleninka.ru/article/n/algoritm-semanticheskoi-segmentatsii-trehmernyh-stsen> (дата обращения: 15.09.2025).
59. 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.
60. Соколов, С. М. Сравнительный анализ степени автономности робототехнических комплексов [Электронный ресурс] // Известия ЮФУ. Технические науки. – 2023. – № 1 (231). – URL: <https://cyberleninka.ru/article/n/sravnitelnyy-analiz-stepeni-avtonomnosti-robototekhnicheskikh-kompleksov> (дата обращения: 28.08.2025).
61. Степанов, О. А., Носов А. С., Торопов А. Б. Навигационная информативность геофизических полей и выбор траекторий в задаче уточнения координат с использованием карты [Электронный ресурс] // Известия ТулГУ. Технические науки. 2018. – № 5. – URL: <https://cyberleninka.ru/article/n/navigatsionnaya-informativnost-geofizicheskikh-poley-i-vybor-traektoriy-v-zadache-utochneniya-koordinat-s-ispolzovaniem-karty> (дата обращения: 27.08.2025).
62. Супрун Д. Е. Алгоритм сопоставления изображений по ключевым точкам при масштабировании и вращении объектов // Вестник МГТУ им. Н.Э. Баумана. Серия «Приборостроение». – 2016. – №5 (110). URL: <https://cyberleninka.ru/article/n/algoritm-sopostavleniya-izobrazheniy-po-klyuchevym-tochkam-pri-masshtabiruемости-i-vraschenii-obektov> (дата обращения: 16.09.2025).
63. Терехов М. А. Обзор современных методов визуальной одометрии // Компьютерные инструменты в образовании. – 2019. – №3. – С. 5-14. URL: <https://cyberleninka.ru/article/n/obzor-sovremennyh-metodov-vizualnoy-odometrii> (дата обращения: 16.09.2025).

64. Фомин, Я. А. Статистическая теория распознавания образов / Я. А. Фомин, Г. Р. Тарловский. – М. : Радио и связь, 1986. – 264 с.
65. Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс; пер. с англ. А. В. Назаенко, И. Ю. Дорошенко. – М. : Вильямс, 2004. – 926 с. – ISBN 5-8459-0542-7.
66. Фридман А. Я., Кулик Б. А. Количественная оценка ситуационной осведомленности в системе концептуального моделирования / А. Фридман // Системный анализ в инженерии и управлении. – 2019. – № 3. – С. 449-460.
67. Фукунага К. Введение в статистическую теорию распознавания образов / К. Фукунага; пер. с англ. Ю. П. Ивановой, В. М. Четыркина; под ред. В. М. Четыркина. – М. : Наука, 1979. – 368 с.
68. Хотилин М. И., Парингер Р. А., Рыцарев И. А., Кравцова Н. С. Разработка и анализ методов выделения объектов изображения // Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2018). – Самара: Новая техника. – 2018. – С. 3581-3589.
69. Шабалина Д. Е., Ланчуковская К. С., Лях Т. В., Чайка К. В. Семантическая сегментация изображений в проекте «duckietown» // Вестник НГУ. Серия: Информационные технологии. – 2021. – №3. – С. 26-39. URL: <https://cyberleninka.ru/article/n/semanticheskaya-segmentatsiya-izobrazheniy-v-proekte-duckietown> (дата обращения: 16.09.2025).
70. Эндсли М., Джонс Д. Проектирование для информирования о ситуации - авиационная перспектива / М. Эндсли // CRC Press. – 2016. – С. 1-13