

Министерство науки и высшего образования Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и кибербезопасности

Работа допущена к защите
Директор ВШТИИ
_____ В.А. Мулюха
« _____ » _____ 2024 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
РАБОТА БАКАЛАВРА
РАЗРАБОТКА АЛГОРИТМА ЛОКАЛИЗАЦИИ ГРУЗОВЫХ
АВТОМОБИЛЕЙ НА ПЕРЕСЕЧЕННОЙ МЕСТНОСТИ НА ОСНОВЕ
ОПТИЧЕСКИХ И ИНФРАКРАСНЫХ СНИМКОВ С ИСПОЛЬЗОВАНИЕМ
МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

по направлению подготовки 02.03.01 Математика и компьютерные науки
Направленность (профиль) 02.03.01_01 Системы искусственного интеллекта и
суперкомпьютерные технологии

Выполнил
студент гр. 5130201/00101

П.Е. Глезова

Руководитель
доцент ВШТИИ ИКНК,

М.А. Курочкин

Консультант
по нормоконтролю

Ю.П. Хотякова

Санкт-Петербург
2024

РЕФЕРАТ

На 44 с., 4 рисунка, 4 таблицы, 0 приложений

КЛЮЧЕВЫЕ СЛОВА: ЛОКАЛИЗАЦИЯ, ИНФРАКРАСНОЕ ИЗЛУЧЕНИЕ, СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ, ОПТИЧЕСКИЙ ПОТОК.

Тема выпускной квалификационной работы: «Разработка алгоритма локализации грузовых автомобилей на пересеченной местности на основе оптических и инфракрасных снимков с использованием методов машинного обучения»

В данной работе предложен интеллектуальный метод обеспечения безопасности охраняемых территорий с использованием беспилотных летательных аппаратов для интеллектуального обнаружения грузовых автомобилей. Проведен обзор существующих технических методов обеспечения безопасности территории, в частности, охраны периметра и обнаружения несанкционированного доступа. Изучены методы машинного обучения для локализации объектов на оптических и инфракрасных снимках, проведен их сравнительный анализ для использования в условиях ограниченных вычислительных мощностей, которые можно расположить на беспилотных летательных аппаратах малых размеров. Разработана конкретная программная реализация алгоритма локализации грузовых автомобилей на основе оптических и инфракрасных снимков.

ABSTRACT

44 pages, 4 figures, 4 tables, 0 appendices

KEYWORDS: LOCALIZATION, INFRARED RADIATION, CONVOLUTIONAL NEURAL NETWORKS, OPTICAL FLOW.

The subject of the graduate qualification work is «Development of an Algorithm for Localizing Trucks in Rugged Terrain Based on Optical and Infrared Images Using Machine Learning Methods».

This study presents an intelligent method for ensuring the security of protected areas through the use of unmanned aerial vehicles (UAVs) for the intelligent detection of trucks. A review of existing technical methods for territorial security, particularly perimeter protection and unauthorized access detection, is conducted. Machine learning methods for object localization on optical and infrared images are examined, with a comparative analysis performed under the constraints of limited computational resources

that can be mounted on small UAVs. A specific software implementation of the algorithm for truck localization based on optical and infrared images has been developed.

СОДЕРЖАНИЕ

| | |
|--|----|
| Введение | 5 |
| Глава 1. Обзор предметной области..... | 6 |
| 1.1. Задача обеспечения безопасности | 6 |
| 1.2. Существующие методы обеспечения безопасности | 8 |
| 1.2.1. Физическая охрана..... | 8 |
| 1.2.2. Инженерно-технические средства охраны | 8 |
| 1.2.3. Технические средства охраны..... | 10 |
| 1.2.4. Технические средства обнаружения несанкционированного доступа | 11 |
| 1.3. Выводы | 13 |
| Глава 2. Разработка метода решения задачи..... | 14 |
| 2.1. Измерение инфракрасного излучения | 14 |
| 2.1.1. Погрешности измерения инфракрасного излучения | 19 |
| 2.1.2. Коэффициент доверия..... | 19 |
| 2.2. Использование методов машинного обучения..... | 22 |
| 2.3. Анализ методов машинного обучения | 24 |
| 2.3.1. Сверточные нейронные сети..... | 24 |
| 2.3.2. <i>Aggregate Channel Features</i> | 25 |
| 2.3.3. <i>Deformable Part Models</i> | 26 |
| 2.3.4. Случайный лес | 28 |
| 2.3.5. Сравнительный анализ | 29 |
| 2.4. Анализ классических алгоритмов | 31 |
| 2.4.1. Вычитание фона..... | 31 |
| 2.4.2. Оптический поток | 33 |
| 2.4.3. Сравнительный анализ..... | 34 |
| 2.5. Разработка гибридного алгоритма | 36 |
| 2.5.1. Математическая постановка задачи..... | 36 |
| 2.5.2. Обзор гибридного алгоритма..... | 37 |
| 2.5.3. Оценка параметров движения | 37 |
| 2.5.4. Гибридный классификатор | 40 |
| 2.6. Оценка точности обнаружения..... | 42 |
| Список использованных источников..... | 43 |

ВВЕДЕНИЕ

Эффективная локализация транспортных средств, в частности грузовых автомобилей на пересеченной местности, представляет собой сложную задачу, требующую использования современных технологий.

Пересеченная местность включает в себя разнообразные природные условия, например, смешанный лес, и характеризуется наличием 20% и более различного рода препятствий, затрудняющих передвижение. Эти препятствия могут быть как естественного происхождения (деревья, склоны, реки), так и искусственно созданными объектами (заборы, рельсовые пути), что значительно усложняет процесс обнаружения и идентификации объектов. Одним из перспективных подходов к решению данной задачи является анализ оптических и инфракрасных снимков.

Инфракрасные снимки, полученные с тепловизоров, позволяют обнаруживать объекты на основе их тепловых характеристик. Оптические снимки, полученные с помощью видеокамер, дают четкое визуальное представление об объектах на местности. Эти данные хорошо дополняют тепловизионные, обеспечивая многосторонний анализ.

Для размещения видеокамер и тепловизоров используются беспилотные летательные аппараты. Они обеспечивают высокую мобильность, позволяя эффективно обследовать труднодоступные зоны и получать данные в режиме реального времени.

Методы машинного обучения являются важным инструментом для обработки и анализа данных, полученных с видеокамер и тепловизоров. Современные алгоритмы машинного обучения позволяют эффективно совмещать и интерпретировать данные, выявляя ключевые признаки, указывающие на наличие грузовых автомобилей на исследуемой территории. Это позволяет создавать интеллектуальные системы, способные автоматически обнаруживать и локализовать транспортные средства с высокой точностью.

Разработка алгоритма локализации грузовых автомобилей на пересеченной местности имеет критическое значение для обеспечения безопасности на охраняемых территориях. Предлагаемая интеллектуальная система обнаружения транспортных средств позволит своевременно выявлять грузовые автомобили, проникающие на охраняемые территории. Данный подход будет способствовать предотвращению несанкционированного проникновения, повысит уровень безопасности объектов и территорий, находящихся под охраной.

ГЛАВА 1. ОБЗОР ПРЕДМЕТНОЙ ОБЛАСТИ

В данной главе рассмотрены основные методы обеспечения безопасности объектов в условиях современных угроз. Описаны различные формы охраны, такие как государственная, ведомственная и вневедомственная, а также способы организации охраны объектов по периметру и внутри территории. Особое внимание уделено существующим инженерно-техническим средствам охраны, системам видеонаблюдения, контроля доступа и сигнализации. Также подробно проанализированы технические средства обнаружения несанкционированного доступа. В заключение представлены выводы о необходимости разработки автономных интеллектуальных методов охраны территорий для повышения уровня безопасности[0].

1.1. Задача обеспечения безопасности

В современных условиях резкого осложнения криминогенной обстановки, роста числа террористических и диверсионных актов проблема обеспечения безопасности объектов входит в разряд приоритетных задач, как для государственных организаций, так и для организаций любой другой формы собственности и для собственников любых видов недвижимости.

По субъектам организации охранной деятельности различаются:

- государственная охрана, представляющая собой специализированные автономные организационные структуры, предназначенные для охраны объектов особой государственной важности, перечень которых устанавливается специальными нормативными актами правительства;
- ведомственная охрана, представляет собой специализированные, вооруженные (как правило) подразделения, осуществляющие охрану различных объектов, входящих в структуру определенного ведомства;
- вневедомственная охрана это специализированные подразделения, осуществляющие охрану объектов, принадлежащих различным ведомствам и частным лицам, на контрактной, возмездной основе.

Формы организации, номенклатура охранных услуг, методы и средства реализации охранной деятельности в основном определяются тем, какому субъекту охранной деятельности подведомственен данный объект, кем он охраняется - государством, ведомством, вневедомственной государственной или частной охраной.

На подступах к объектам охраны создаются активные и пассивные защитные препятствия, например: система физических препятствий (инженерные заграждения), специальное оборудование мест хранения секретных документов, контрразведывательное обеспечение.

Надежность охраны достигается детальным построением системы охраны, правильной организацией и бдительным несением службы нарядами.

В зависимости от местности, характера и категории объекта и других особенностей охрана объектов может быть организована следующими способами:

1. По периметру. Технические средства охраны выставляются на границе охраняемой территории и преграждают доступ к объекту вне пропускных пунктов (именно таким способом, как правило, охраняются некоторые режимные объекты).
2. По отдельным объектам. Личный состав выставляется непосредственно на охраняемом объекте (примером такого способа охраны может быть порядок организации охраны складов МВД по хранению боеприпасов).
3. Смешанным способом. По периметру и отдельным объектам одновременно.
4. Способом оперативного дежурства. Охранные функции осуществляются комплексом инженерно-технических средств охраны при дежурном состоянии сил охраны (примером такого способа охраны является организация охраны любой атомной электростанции).

На основе изученных статистических данных, можно сделать вывод, что охраняемые объекты наиболее часто подвергаются следующим видам угроз:

- несанкционированное проникновение на территорию;
- несанкционированное получение информации об объекте или иной закрытой информации путем установки на объекте скрытых средств негласного получения информации;
- нападение на охраняемый объект с целью хищения материальных ценностей;
- угрозы жизни и здоровью персонала и посетителей объекта, в том числе взятие заложников с целью достижения иных целей;
- нарушение инфраструктуры и линий жизнеобеспечения объекта охраны;
- нарушение режима работы объекта, с целью прекращения его функционирования;
- саботаж технических средств охраны.

1.2. Существующие методы обеспечения безопасности

Методы обеспечения безопасности объектов включают в себя широкий спектр мер, направленных на предотвращение и минимизацию рисков, связанных с несанкционированным доступом, диверсиями, террористическими актами и другими угрозами. Рассмотрим основные существующие методы обеспечения безопасности.

1.2.1. Физическая охрана

Физическая охрана объектов заключается в использовании специально обученного персонала для защиты территории и имущества от различных угроз. Может быть реализована различными способами.

1. Патрулирование территории.

Сотрудники охраны проводят регулярные обходы по территории объекта с целью выявления и предотвращения возможных угроз. Патрулирование может осуществляться как пешком, так и на транспортных средствах.

2. Посты охраны.

Размещение стационарных постов охраны на ключевых точках объекта, таких как въезды и выезды, зоны повышенного риска или наиболее уязвимые участки периметра.

3. Контрольно-пропускные пункты (КПП).

Организация пропускного режима с использованием систем идентификации (карты доступа, биометрические данные), журналов регистрации посетителей и персонала, а также физического контроля со стороны сотрудников охраны.

1.2.2. Инженерно-технические средства охраны

В современных условиях для обеспечения охраны территорий используются разнообразные инженерно-технические средства. Эти средства могут включать как естественные, так и искусственные барьеры, которые препятствуют незаконному проникновению на охраняемую территорию. Основное внимание уделяется искусственным заградительным сооружениям, которые обеспечивают физическую защиту периметра объекта, элементов зданий и помещений от несанкционированного доступа.

А. Заграждения и противотаранные устройства.

1. Колючая проволока и армированная колючая лента. Используются для создания заграждений, которые затрудняют или делают невозможным преодоление препятствия. АКЛ обладает высокой прочностью, упругостью и стойкостью к коррозии благодаря оцинкованному покрытию. Этот тип заграждения является одним из самых распространенных и недорогих средств защиты.
2. Сварные сетчатые панели. Применяются для ограждения промышленных объектов, объектов городской инфраструктуры и частной собственности. Конструкции таких заграждений могут включать различные дополнительные технические средства обнаружения, имеют минимальные сроки монтажа и хорошо вписываются в городскую инфраструктуру.
3. Просечная вытяжная сетка. Обеспечивает устойчивость к ветровым нагрузкам и может быть выполнена из различных материалов, таких как низкоуглеродистая сталь, алюминий или оцинкованная сталь.
4. Сварные трубы и радиопрозрачные заграждения. Сварные трубы часто используются для создания прочных ограждений, покрытых полимерными материалами. Радиопрозрачные заграждения, выполненные из пластика и стеклопластика, предназначены для защиты радиотехнических комплексов, так как они не препятствуют приему и передаче электромагнитных волн.
5. Электрошоковые заграждения. Используются для создания высокоэффективных барьеров с применением безопасного электрошокового воздействия. Эти заграждения питаются от напряжения 220В и вызывают болезненные ощущения, вынуждая злоумышленника отказаться от противоправных действий.
6. Железобетонные противотаранные заграждения. Обеспечивают надежную защиту от таранных атак. Внутри железобетонных плит могут прокладываться кабели для систем сигнализации и видеонаблюдения.

В. Средства регулирования доступа.

1. Шлагбаумы. Используются для контроля въезда и выезда автотранспорта на охраняемую территорию. Управление шлагбаумами может осуществляться с пульта охраны, пульта-брелока водителя или с помощью бесконтактных карт и жетонов.
2. Ворота. Существуют различные типы ворот, такие как распашные, откатные и консольные. Они могут быть оснащены датчиками контроля положения, электроприводами и дополнительными заградительными элементами.
3. Противотаранные устройства и блокираторы. Включают мобильные блоки, выдвижные столбы и стационарные дорожные блокираторы, предназначенные для предотвращения несанкционированного проезда автотранспорта. Некоторые устройства могут оснащаться датчиками для обнаружения ударов или вибраций.
4. Дорожные шипы и козырьковые заграждения. Дорожные шипы предназначены для принудительной остановки автотранспорта, пробивая шины. Козырьковые заграждения устанавливаются на верхней части ограждений для предотвращения перелазов.

1.2.3. Технические средства охраны

Современные технологии позволяют существенно повысить уровень безопасности объектов. Технические средства охраны включают в себя следующее.

А. Системы видеонаблюдения.

Установка камер видеонаблюдения по периметру и внутри объекта позволяет осуществлять круглосуточный мониторинг и запись событий. Современные системы видеонаблюдения имеют разные уровни сложности: от одной камеры с монитором до многокамерных компьютерных систем с цифровой обработкой изображения в реальном времени. Основные типы видеокамер представлены ниже.

1. Цветные, черно-белые и камеры с режимом «день-ночь». Черно-белые камеры обладают большим разрешением и чувствительностью, что делает их подходящими для наблюдения за далеко удаленными объектами. Камеры с режимом «день-ночь» автоматически переключаются между цветным и черно-белым

режимами в зависимости от уровня освещенности, что позволяет вести наблюдение даже в полной темноте.

2. Видеокамеры различной конструкции: стационарные, поворотные, цилиндрические, купольные и уличные. Поворотные камеры позволяют наблюдать за обширными зонами, а купольные высокоскоростные камеры оснащены трансфокаторами, позволяющими изменять масштаб изображения.
3. Беспроводные системы видеонаблюдения, которые используются в случаях, когда прокладка кабелей невозможна или нежелательна, например, в зданиях с архитектурной ценностью или в кабинах лифтов. Беспроводное видеонаблюдение может быть организовано с использованием Wi-Fi модулей, но требует учета возможных проблем с устойчивостью сигнала.

Важным элементом системы видеонаблюдения являются видеорегистраторы, которые записывают, обрабатывают и хранят видеоданные. Существуют различные типы видеорегистраторов: цифровые *DVR*, сетевые *NVR* и гибридные *HDVR*, а также программные комплексы для ПК.

В. Системы контроля доступа.

Использование электронных замков, турникетов, шлюзов и других устройств для регулирования доступа на территорию объекта. Данные системы интегрируются с базами данных сотрудников и посетителей, обеспечивая индивидуальные уровни доступа и фиксируя все попытки входа и выхода.

С. Системы сигнализации.

Установка охранных, пожарных и тревожных сигнализаций, которые автоматически оповещают службу безопасности и экстренные службы о возникновении угрозы. Сигнализация может быть оснащена датчиками движения, разбития стекла, дыма, газа и других параметров.

1.2.4. Технические средства обнаружения несанкционированного доступа

Современные системы безопасности играют ключевую роль в охране периметров объектов от несанкционированного доступа. Эти системы включают

разнообразные технические средства, направленные на своевременное выявление попыток вторжения и предотвращение угроз.

Одним из основных элементов таких систем являются объемные ультразвуковые извещатели (рис.1.1a). Эти устройства предназначены для обнаружения проникновения в охраняемую конструкцию и перемещения предметов внутри нее. Ультразвуковой извещатель состоит из блока обработки сигнала, акустического излучателя и приемника. Излучающий элемент преобразует электрическое напряжение в акустические колебания, которые заполняют охраняемый объем. Приемный элемент фиксирует эти колебания и передает их в блок обработки сигнала, который, в зависимости от заложенного алгоритма, формирует извещение о нарушении.

Для охраны периметров также широко используются цифровые оптико-электронные пассивные инфракрасные извещатели (рис.1.1b). Эти устройства устанавливаются на открытых территориях и предназначены для обнаружения движения в пределах охраняемой зоны. Принцип их работы основан на фиксации инфракрасного излучения, излучаемого объектами, движущимися в зоне контроля. Прибор анализирует изменение теплового поля и, при выявлении подозрительной активности, передает сигнал тревоги на центральный пульт охраны.

Важным компонентом систем обнаружения несанкционированного доступа являются вибрационные извещатели, устанавливаемые на ограждениях и других конструкциях периметра. Эти извещатели реагируют на механические колебания и вибрации, возникающие при попытках преодоления или повреждения барьеров. Высокая чувствительность позволяет фиксировать даже незначительные попытки воздействия, обеспечивая своевременное реагирование на угрозы.

Системы радиолучевого обнаружения применяются для создания невидимых барьеров на границе охраняемой территории. Они состоят из передатчика и приемника, между которыми создается радиочастотное поле. Нарушение этого поля вследствие движения объекта приводит к генерации сигнала тревоги. Радиолучевые системы эффективны на открытых пространствах и могут покрывать большие периметры.

Современные средства обнаружения также включают оптоволоконные датчики, которые интегрируются в структуры ограждений. Эти датчики фиксируют изменения в световом потоке, передаваемом по оптоволоконным кабелям, вызванные деформацией или вибрацией ограждений. Высокая точность и устойчивость к

внешним условиям делают оптоволоконные системы надежным инструментом в обеспечении периметровой безопасности.

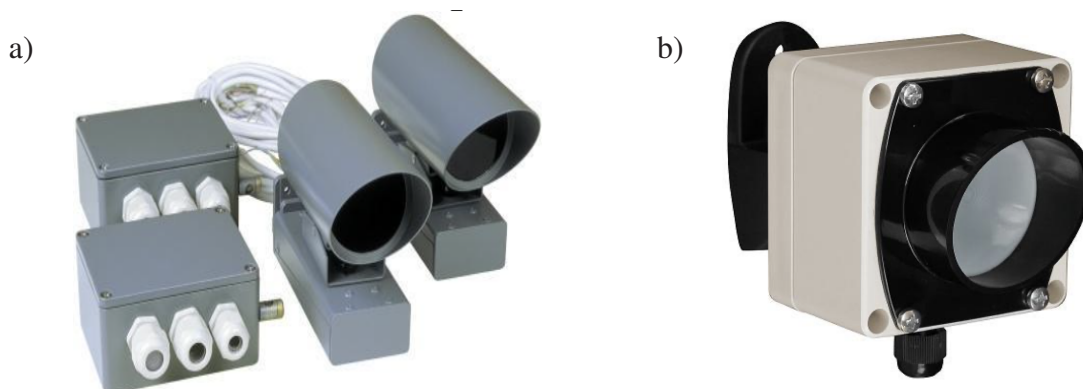


Рис.1.1. Примеры систем охраны периметра: *a* — активный инфракрасный извещатель; *b* — уличный цифровой оптико-электронный пассивный инфракрасный излучатель

Каждое из этих технических средств обладает своими преимуществами и особенностями, которые определяют их использование в зависимости от конкретных требований объекта охраны. Комплексное применение различных типов извещателей позволяет создавать многоуровневую систему безопасности, способную эффективно противостоять различным видам угроз.

1.3. Выводы

Представленные методы обеспечения охраны территорий, включая инженерно-технические средства, системы видеонаблюдения, контроля доступа и сигнализации, представляют собой многообразные и технологически продвинутые решения для предотвращения несанкционированного доступа и защиты объектов. Однако, несмотря на их эффективность и многоуровневый подход, эти методы имеют ряд существенных недостатков и ограничений.

Физическая охрана и технические средства требуют значительных финансовых затрат на установку, обслуживание и модернизацию. Они также зависят от человеческого фактора, что может привести к ошибкам и нарушениям безопасности. Инженерно-технические ограждения, которые обеспечивают физические барьеры, не всегда могут реагировать на динамические угрозы и требуют постоянного мониторинга.

Системы контроля доступа и сигнализации обеспечивают высокий уровень защиты, но часто сложны в интеграции и эксплуатации. Они требуют постоянного обновления и адаптации к новым угрозам и технологиям. Более того, в случае

кибератак или технических сбоев эти системы могут быть выведены из строя, что существенно снижает уровень безопасности.

Таким образом, представленные методы, хотя и обеспечивают базовый уровень защиты, не всегда могут эффективно и оперативно реагировать на угрозы. Для повышения уровня безопасности необходимо разрабатывать и внедрять автономные интеллектуальные методы охраны территорий. Такие системы, основанные на искусственном интеллекте и машинном обучении, способны самостоятельно анализировать большие объемы данных, предотвращать потенциальные угрозы, адаптироваться к изменяющимся условиям и минимизировать зависимость от человеческого фактора. Интеллектуальные системы охраны могут значительно повысить уровень безопасности, обеспечивая более надежную и устойчивую защиту объектов в современных условиях.

ГЛАВА 2. РАЗРАБОТКА МЕТОДА РЕШЕНИЯ ЗАДАЧИ

2.1. Измерение инфракрасного излучения

Все объекты, температура которых превышает температуру абсолютного нуля излучают электромагнитное тепловое излучение. Согласно распределению длин волн, излучаемая энергия имеет зависимость от температуры поверхности объекта, которую можно описать законом излучения Планка:

$$M(\lambda, T) = \frac{c_1 \lambda^{-5}}{e^{c_2/\lambda T} - 1} \quad (2.1)$$

где $M(\lambda, T)$ величина излучения абсолютно черного тела. c_1 и c_2 первая и вторая константы излучения соответственно. λ длина волны излучения и T абсолютная температура черного тела. Когда $\exp(c_2/\lambda T) \gg 1$, формулу Планка можно заменить следующей формулой смещения Вина:

$$M_b(\lambda, T) = c_1 \lambda^{-5} e^{-c_2/\lambda T} \quad (2.2)$$

Закон смещения Вина указывает на то, что чем выше температура объекта, тем короче длина волны его спектра излучения, и центральный пик смещается в сторону коротких волн. Однако излучаемая энергия, поступающая на чувствительную поверхность датчика, в реальных измерениях включает не только излучаемую энергию целевого объекта, но также энергию окружающих объектов и атмосферы.

Поэтому спектральную излучательную способность поверхности целевого объекта можно выразить следующим образом:

$$L_{\lambda} = \varepsilon_{\lambda} M_b(\lambda, T_{\text{obj}}) + (1 - \alpha_{\lambda}) M_b(\lambda, T_{\text{sur}}) \quad (2.3)$$

где ε_{λ} и T_{obj} излучательная способность и температура целевого объекта соответственно; $\varepsilon_{\lambda} M_b(\lambda, T_{\text{obj}})$ и $M_b(\lambda, T_{\text{sur}})$ это спектральная яркость целевого объекта и окружающей среды. α_{λ} поглощающая способность поверхности целевого объекта. T_{sur} температура окружающей среды.

Первый множитель в правой части уравнения (3.3) представляет собой спектральную яркость поверхности целевого объекта, а второй множитель — спектральную яркость окружающей среды, отраженную от целевого объекта. Излучение, действующее на систему измерения инфракрасного излучения, представлено на рисунке 2.1, основными источниками которого являются окружающая среда, объект и атмосфера. Это может быть выражено как:

$$E_{\lambda} = A_{\text{obj}} d^{-2} [\tau_{\alpha\lambda} \varepsilon_{\lambda} M_b(\lambda, T_{\text{obj}}) + \tau_{\alpha\lambda} (1 - \alpha_{\lambda}) \cdot M_b(\lambda, T_{\text{sur}}) + \varepsilon_{\alpha\lambda} M_b(\lambda, T_{\text{atm}})] \quad (2.4)$$

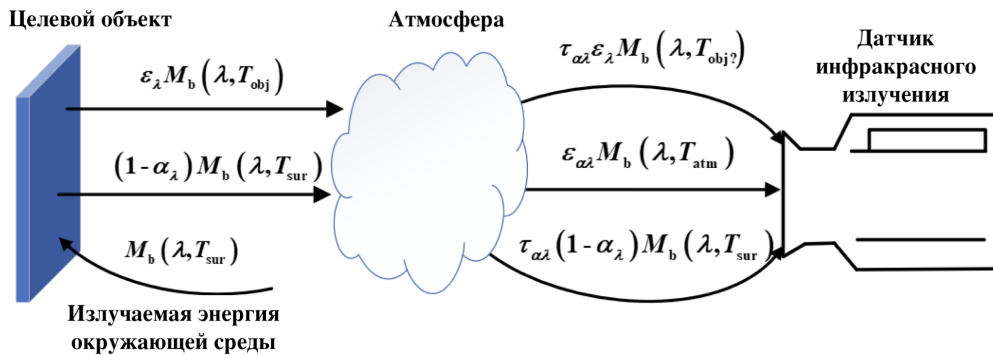


Рис.2.1. Принципиальная схема получаемой энергии системы измерения инфракрасного излучения

где A_{obj} — видимая площадь целевого объекта, соответствующая минимальному углу тензора датчика инфракрасного излучения, а d — расстояние от целевого объекта до датчика инфракрасного излучения. В общем случае $A_{\text{obj}} d^{-2}$ является константой. $\tau_{\alpha\lambda}$, $\varepsilon_{\alpha\lambda} M_b(\lambda, T_{\text{atm}})$, $\varepsilon_{\alpha\lambda}$ и T_{atm} — это спектральная пропускная способность, излучение, коэффициент излучения и температура атмосферы соответственно.

Полученная инфракрасная излучаемая энергия преобразуется датчиком в сигнал тока, то есть падающая инфракрасная излучаемая энергия интегрируется по

полосе пропускания $\Delta\lambda$. Следовательно, зависимость между излучаемой энергией и током можно выразить следующим образом:

$$I_0 = \int_{\Delta\lambda} A_R R_\lambda E_\lambda \tau_f d\lambda \quad (2.5)$$

где I_0 — выходной сигнал тока датчика. E_λ — освещенность радиации, полученная инфракрасной измерительной системой. A_R — площадь инфракрасной фокусирующей линзы. R_λ — спектральная чувствительность инфракрасного датчика. τ_f — пропускная способность оптической системы. Выходной ток можно преобразовать в сигнал напряжения через цепь преобразования I/V, что можно выразить следующим образом:

$$V_{\text{out}} = \int_{\Delta\lambda} R A_{\text{obj}} d^{-2} A_R \tau_f R_\lambda [\tau_{\alpha\lambda} \varepsilon_\lambda M_b(\lambda, T_{\text{obj}}) + \tau_{\alpha\lambda} (1 - \alpha_\lambda) M_b(\lambda, T_{\text{sur}}) + \varepsilon_{\alpha\lambda} M_b(\lambda, T_{\text{atm}})] d\lambda \quad (2.6)$$

где R — нагрузка.

Уравнение (3.6) указывает, что существует множество факторов, влияющих на измерение инфракрасной температуры. На самом деле, эффективная площадь линзы, пропускная способность оптической системы и нагрузка определяются после спецификации аппаратного обеспечения системы измерения температуры. Когда $K = R A_R T \tau_f$, уравнение (3.6) можно упростить следующим образом:

$$V_{\text{out}} = K A_{\text{obj}} d^{-2} \int_{\Delta\lambda} R_\lambda [\tau_{\alpha\lambda} \varepsilon_\lambda \cdot M_b(\lambda, T_{\text{obj}}) + \tau_{\alpha\lambda} (1 - \alpha_\lambda) M_b(\lambda, T_{\text{sur}}) + \varepsilon_{\alpha\lambda} M_b(\lambda, T_{\text{atm}})] d\lambda \quad (2.7)$$

В таблице 2.1 представлено подробное описание переменных уравнения (3.7).

Из уравнения (3.7) видно, что точность измерений зависит от температуры окружающей среды, расстояния и угла измерения при определенной излучательной способности черного тела.

Таблица 2.1

Описание переменных уравнения (3.7)

| Переменная | Описание | Источник значения | Единица измерения |
|-----------------|---|--|----------------------------|
| V_{out} | Выходное напряжение на инфракрасном датчике | Результат измерений | Вольт (V) |
| K | Константа, определяющая параметры оптической системы инфракрасного датчика: площадь линзы, пропускная способность линзы | Задается характеристиками инфракрасного детектора | Безразмерная величина |
| A_{obj} | Видимая площадь наблюдаемого объекта | Определяется геометрией объекта и углом зрения инфракрасного датчика | Квадратные метры (m^2) |
| d | Расстояние от наблюдаемого объекта до инфракрасного датчика | Измеряется в экспериментальных условиях | Метры (м) |
| $\Delta\lambda$ | Спектральный диапазон измерений | Задается характеристиками инфракрасного датчика | Микрометры (μm) |
| R_λ | Спектральная чувствительность инфракрасного датчика | Задается характеристиками инфракрасного датчика | Безразмерная величина |

Таблица 2.1

Описание переменных уравнения (3.7)

| Переменная | Описание | Источник значения | Единица измерения |
|--|---|--|--|
| $\tau_{\alpha\lambda}$ | Спектральная пропускная способность атмосферы | Табличные или экспериментальные данные | Безразмерная величина |
| ε_{λ} | Излучательная способность объекта | Табличные данные | Безразмерная величина |
| $M_b(\lambda, T_{obj})$ | Спектральная плотность излучения объекта при температуре T_{obj} | Вычисляется по закону Планка | Ватты на квадратный метр на микрометр (Вт/м ² /μ м) |
| $M_b(\lambda, T_{sur})$ | Спектральная плотность излучения окружающей среды при температуре T_{sur} | Вычисляется по закону Планка | Ватты на квадратный метр на микрометр (Вт/м ² /μ м) |
| $\tau_{\alpha\lambda}(1 - \alpha_{\lambda})$ | Фактор отражения излучения окружающей среды | Табличные или экспериментальные данные | Безразмерная величина |
| $M_b(\lambda, T_{atm})$ | Спектральная плотность излучения атмосферы при температуре T_{atm} | Вычисляется по закону Планка | Ватты на квадратный метр на микрометр (Вт/м ² /μ м) |
| $\varepsilon_{\alpha\lambda}$ | Спектральная излучательная способность атмосферы | Табличные или экспериментальные данные | Безразмерная величина |

2.1.1. Погрешности измерения инфракрасного излучения

В практическом применении при проведении измерения инфракрасного излучения существенное значение имеет выявление и устранение систематических и случайных погрешностей, оказывающих влияние на результаты измерения.

Систематические погрешности заключены в конструкции измерительного прибора, а также зависят от его выбора в соответствии с требованиями к совершенству измерения (разрешающей способности, поля зрения и т.п.).

Случайными погрешностями, возникающими при проведении измерения инфракрасного излучения, могут являться:

- коэффициент излучения материала;
- солнечная радиация;
- расстояние до объекта;
- тепловое отражение и т.п.

2.1.2. Коэффициент доверия

Для вычисления уровня доверия к результатам измерения на основе различных параметров, можно выбрать ключевые изменяемые параметры из формулы (3.7).

Основными параметрами, влияющими на измерения, являются:

- расстояние до объекта (d);
- угол наблюдения (θ);
- температура атмосферы (T_{atm});
- спектральная пропускная способность атмосферы ($\tau_{\alpha\lambda}$);
- спектральная излучательная способность объекта (ε_{λ})

Формула для вычисления уровня доверия

Предположим, что каждый параметр влияет на уровень доверия линейно и независимо. Можно назначить каждому параметру весовой коэффициент, отражающий его относительное влияние на общий уровень доверия.

$$D = w_d \cdot f_d(d) + w_{\theta} \cdot f_{\theta}(\theta) + w_{T_{\text{atm}}} \cdot f_{T_{\text{atm}}}(T_{\text{atm}}) + w_{\tau_{\alpha\lambda}} \cdot f_{\tau_{\alpha\lambda}}(\tau_{\alpha\lambda}) + w_{\varepsilon_{\lambda}} \cdot f_{\varepsilon_{\lambda}}(\varepsilon_{\lambda}) \quad (2.8)$$

Где:

- d — расстояние до объекта,
- θ — угол наблюдения,
- T_{atm} — температура атмосферы,

- $\tau_{\alpha\lambda}$ — спектральная пропускная способность атмосферы,
- ε_λ — излучательная способность объекта

Функции и веса

Для каждого параметра определим нормированные функции f_i , принимающие значения от 0 до 1, и весовые коэффициенты w_i , сумма которых равна 1.

1. Расстояние до объекта (d):

$$f_d(d) = \frac{1}{1 + k_d \cdot d}$$

Где k_d — коэффициент, определяющий, как быстро снижается уровень доверия с увеличением расстояния.

2. Угол наблюдения (θ):

$$f_\theta(\theta) = \cos(\theta)$$

Так как по закону Ламберта интенсивность радиации пропорциональна косинусу угла наблюдения.

3. Температура атмосферы (T_{atm}):

$$f_{T_{\text{atm}}}(T_{\text{atm}}) = 1 - \left| \frac{T_{\text{atm}} - T_{\text{opt}}}{T_{\text{max}} - T_{\text{min}}} \right|$$

Где T_{opt} — оптимальная температура для измерений, T_{max} и T_{min} — максимальная и минимальная температуры, соответственно.

4. Спектральная пропускная способность атмосферы ($\tau_{\alpha\lambda}$):

$$f_{\tau_{\alpha\lambda}}(\tau_{\alpha\lambda}) = \tau_{\alpha\lambda}$$

Прямо пропорциональна пропускной способности атмосферы.

5. Излучательная способность объекта (ε_λ):

$$f_{\varepsilon_\lambda}(\varepsilon_\lambda) = \varepsilon_\lambda$$

Прямо пропорциональна излучательной способности объекта.

Пример весовых коэффициентов

Пусть весовые коэффициенты распределены следующим образом:

$$w_d = 0.3$$

$$w_\theta = 0.25$$

$$w_{T_{\text{atm}}} = 0.2$$

$$w_{\tau_{\alpha\lambda}} = 0.15$$

$$w_{\varepsilon_{\lambda}} = 0.1$$

С учетом вышеуказанных весов и функций, получаем формулу:

$$D = 0.3 \cdot \frac{1}{1 + k_d \cdot d} + 0.25 \cdot \cos(\theta) + 0.2 \cdot \left(1 - \left| \frac{T_{\text{atm}} - T_{\text{opt}}}{T_{\text{max}} - T_{\text{min}}} \right| \right) + 0.15 \cdot \tau_{\alpha\lambda} + 0.1 \cdot \varepsilon_{\lambda}$$

Пример набора параметров и расчет уровня доверия

Пусть:

$$d = 20 \text{ м}$$

$$\theta = 30^\circ$$

$$T_{\text{atm}} = 25^\circ \text{C}$$

$$T_{\text{opt}} = 20^\circ \text{C}$$

$$T_{\text{max}} = 35^\circ \text{C}$$

$$T_{\text{min}} = 5^\circ \text{C}$$

$$\tau_{\alpha\lambda} = 0.85$$

$$\varepsilon_{\lambda} = 0.9$$

Тогда:

$$f_d(20) = \frac{1}{1 + k_d \cdot 20}, \text{ предположим } k_d = 0.05, \text{ тогда } f_d(20) = \frac{1}{1 + 1} = 0.5$$

$$f_{\theta}(30^\circ) = \cos(30^\circ) = \sqrt{3}/2 \approx 0.866$$

$$f_{T_{\text{atm}}}(25) = 1 - \left| \frac{25 - 20}{35 - 5} \right| = 1 - \frac{5}{30} = 0.833$$

$$f_{\tau_{\alpha\lambda}}(0.85) = 0.85$$

$$f_{\varepsilon_{\lambda}}(0.9) = 0.9$$

Подставляя значения в формулу:

$$D = 0.3 \cdot 0.5 + 0.25 \cdot 0.866 + 0.2 \cdot 0.833 + 0.15 \cdot 0.85 + 0.1 \cdot 0.9$$

$$D = 0.15 + 0.2165 + 0.1666 + 0.1275 + 0.09$$

$$D = 0.7506$$

2.2. Использование методов машинного обучения

Во многих современных оптических системах для беспилотных летательных аппаратов применяются алгоритмы для распознавания и выбора объектов. Если целевой объект находится, например, на большой высоте, его можно обнаружить с помощью оптического контрастирования из-за заметного отличия от неба или в инфракрасном спектре благодаря значительной разнице температур между объектом и окружающей средой. Однако эти методы не столь эффективны для оптических систем беспилотников, сканирующих земную поверхность, так как сцены становятся многообъектными и автоматическое распознавание целевого объекта усложняется помехами. В таких случаях целесообразно использовать методы распознавания образов, основанные на алгоритмах машинного обучения и технологии искусственного интеллекта.

При построении систем управления движущимися объектами различного назначения особое внимание уделяется выбору типа и набора бортовых сенсорных устройств и их характеристик, обеспечивающих получение достоверной информации о состоянии окружающей среды и, как следствие, эффективное решение поставленной задачи управления в различных условиях применения. Большое распространение получили оптико-электронные системы для вывода на экран оператора видеоизображения, полученного при помощи тепловизионной и телевизионной камер (рис. 2.2)[bezzubov] .

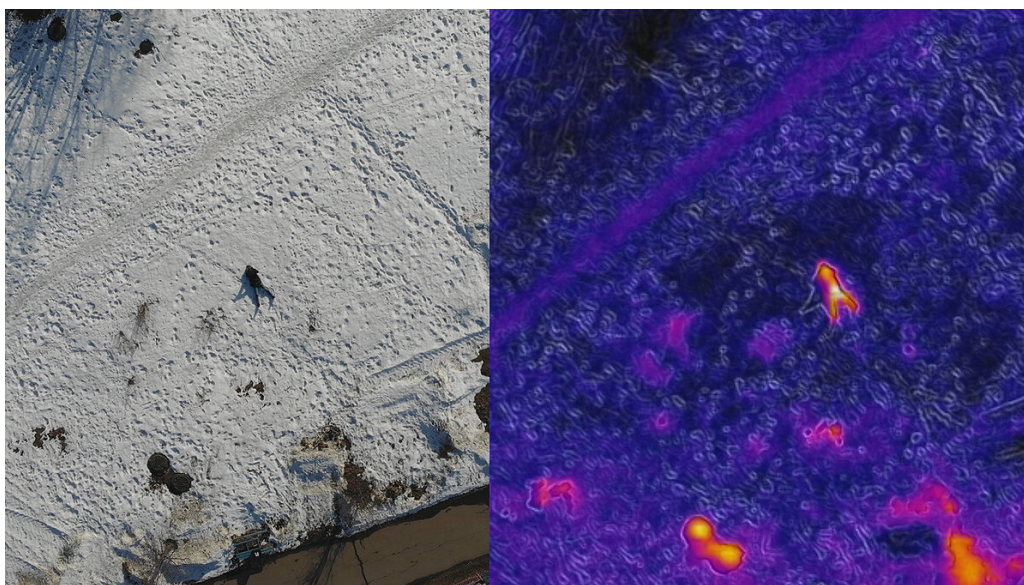


Рис.2.2. Пример съемки на Mavic 2 Enterprise Dual с высоты 40 метров

Телевизионные камеры имеют существенный недостаток: погодные условия, такие как дождь, снег и туман, затрудняют видимость или делают её невозможной. Проблему видимости в экстремальных условиях можно решить с помощью тепловизоров. Однако тепловизоры также имеют недостатки. Из-за низкой пространственной разрешающей способности они не показывают мелкие детали наблюдаемой сцены. Кроме того, тепловизоры эффективны только для объектов с температурой, отличающейся от температуры окружающей среды, то есть при наличии температурного контраста.

Полезными признаками для телевизионных изображений являются [0]:

- форма,
- размеры,
- текстура,
- внутренняя структура объектов,
- окружение.

Полезными признаками для тепловизионных изображений являются:

- форма,
- максимальное/минимальное значение эмиссии,
- количество и расположение горячих пятен,
- окружение (среда).

Распознавание трехмерных объектов по их двумерным изображениям стало одной из ключевых задач анализа сцен и компьютерного зрения. Объектом считается не только цифровое представление фрагмента двумерной сцены, но и его приближенное описание в виде набора характерных свойств (признаков). Признаки – это измерения, используемые для классификации объектов [0].

При реализации задачи локализации решение может формулироваться следующим образом: на основе априорной информации о рассматриваемой сцене (участке земной поверхности) и апостериорной информации, полученной в процессе полета, сопоставляются текущее и эталонное изображения. Это позволяет локализовать заданные объекты сцены на текущем изображении и определить их текущие координаты для формирования управляющих сигналов для беспилотного летательного аппарата [0].

2.3. Анализ методов машинного обучения

2.3.1. Сверточные нейронные сети

Сверточные нейронные сети представляют собой класс искусственных нейронных сетей, специально разработанных для обработки и анализа изображений. Их архитектура и механизмы позволяют эффективно распознавать сложные паттерны в визуальных данных, что делает их незаменимыми в различных областях компьютерного зрения [0].

Основными компонентами являются сверточные слои, слои объединения (пулинга) и полностью связанные слои.

1. Сверточные слои выполняют операцию свертки, применяя фильтры или ядра к исходному изображению. Фильтры помогают выделять различные признаки, такие как края, углы и текстуры. Каждый фильтр генерирует карту признаков, которая подчеркивает определенные аспекты входного изображения.
2. Слои объединения пулинга уменьшают размерность карт признаков, что снижает вычислительную нагрузку и вероятность переобучения. Наиболее распространенными методами объединения являются максимальный пулинг (максимальное значение в каждом подокне) и средний пулинг (среднее значение в каждом подокне).
3. Полностью связанные слои находятся в конце сети и служат для окончательной классификации. Они соединяют все нейроны предыдущего слоя с каждым нейроном текущего слоя, позволяя учитывать все выделенные ранее признаки.

Сверточные нейронные сети обладают рядом преимуществ, делающих их эффективными для обработки изображений. В первую очередь свойства сверточных слоев позволяют учитывать локальные пространственные отношения между пикселями, что важно для распознавания объектов. Использование методов, таких как нормализация локальных откликов и Dropout, помогает уменьшить вероятность переобучения. Также, сверточные нейронные сети могут эффективно обучаться на больших объемах данных благодаря использованию графических процессоров (*GPU*), что значительно ускоряет процесс обучения.

Сверточные нейронные сети широко применяются для решения различных задач в области классификации изображений. Например, используются для иден-

тификации и верификации лиц на изображениях и в видео; помогают в анализе медицинских изображений, таких как рентгеновские снимки и МРТ, для обнаружения патологий; используются для распознавания объектов и ситуаций на дорогах, что необходимо для безопасного управления транспортом.

Современные исследования в области сверточных нейронных сетей направлены на улучшение архитектуры и алгоритмов обучения. Разрабатываются новые методы нормализации и регуляризации, такие как *Batch Normalization* и новые типы нелинейностей (например, *ReLU*). Ведутся работы по созданию более эффективных моделей, которые требуют меньше вычислительных ресурсов и могут работать в реальном времени.

2.3.2. *Aggregate Channel Features*

Метод *ACF* (*Aggregate Channel Features*) является инновационным подходом в области распознавания лиц, который направлен на решение проблем, связанных с большими изменениями внешнего вида лиц в естественных условиях. Данный метод опирается на концепцию каналов изображения, расширяя их до различных типов, таких как градиенты и гистограммы ориентированных градиентов, что позволяет кодировать разнообразную признаками информацию в простой форме [0].

Исторически, одним из самых влиятельных подходов к распознаванию лиц был метод *Viola-Jones*, который использует прямоугольные признаки, похожие на признаки Хаара, и обучает классификатор с помощью алгоритма *Adaboost*. Этот метод достиг значительных успехов в распознавании лиц в реальном времени, но его производительность все еще оставалась ограниченной из-за большой изменчивости внешнего вида лиц в неконтролируемых условиях. Для преодоления этого ограничения метод *ACF* использует расширенные каналы изображения, такие как градиенты и гистограммы ориентированных градиентов, что позволяет более эффективно кодировать информацию о лицах.

Метод *ACF* включает в себя несколько ключевых аспектов. Во-первых, он использует расширенные каналы изображения для кодирования разнообразной информации. Например, цветовые каналы в пространстве *LUV*, каналы градиентной величины и гистограммы градиентов в *RGB* пространстве показывают наилучшие результаты при распознавании лиц. Эти каналы обеспечивают богатую репрезен-

тационную способность, что особенно важно для обработки лиц с различными выражениями и позами.

Во-вторых, метод *ACF* предусматривает мультискейлинг представления признаков, что позволяет еще больше обогатить репрезентацию. В оригинальной версии метода все признаки имели одинаковый масштаб, но эксперименты показали, что мультискейлинг улучшает производительность. Это достигается за счет изменения масштаба восприятия, локального масштаба и масштаба интеграции признаков.

Третьим важным аспектом является подход к мультивидовому распознаванию лиц, который использует повторную ранжировку оценок и корректировку обнаружений. Это помогает эффективно справляться с различными позами лиц и улучшает точность локализации лиц в изображениях. В результате, метод *ACF* демонстрирует конкурентоспособные результаты на сложных наборах данных, таких как *AFW* и *Fddb*, показывая высокую точность и скорость распознавания (до 42 кадров в секунду на изображениях *VGA*).

Метод *ACF* значительно улучшает производительность распознавания лиц благодаря тщательному исследованию и оптимизации различных параметров признаков. Он предлагает более быструю и точную альтернативу традиционным методам распознавания лиц, обеспечивая богатую репрезентацию и высокую эффективность вычислений.

2.3.3. *Deformable Part Models*

Метод деформируемых частей (*Deformable Part Models, DPM*) представляет собой один из эффективных подходов к обнаружению объектов на изображениях. Этот метод был разработан для улучшения точности и эффективности поиска объектов, особенно в условиях сложного фона и объектов в различных позах [0].

Основная идея метода *DPM* заключается в моделировании объектов как совокупности частей, которые могут деформироваться относительно друг друга. Каждая часть описывается своим собственной подмоделью, которая оценивает вероятность нахождения этой части объекта в определённом месте изображения. Главным преимуществом такого подхода является возможность учёта различных положений и деформаций частей объекта, что значительно улучшает результаты обнаружения по сравнению с более простыми моделями машинного обучения.

Основой метода *DPM* является использование каскадных классификаторов. В данном контексте каскадный классификатор представляет собой последовательность всё более сложных моделей, которые последовательно отсекают маловероятные гипотезы обнаружения объектов. Сначала применяется простая модель, которая быстро отбрасывает большую часть областей изображения. Затем, на оставшихся областях, применяются более сложные модели, которые более точно определяют положение и конфигурацию частей объекта.

Для реализации *DPM* с каскадными классификаторами используется несколько ключевых компонент.

1. Использование моделей звездной структуры, в которых все части объекта привязаны к одной "корневой" части, упрощает вычисления и позволяет эффективно использовать динамическое программирование для поиска оптимальных конфигураций частей.
2. На ранних этапах каскада используются упрощённые модели частей, которые быстрее вычисляются. Это позволяет быстро отсеивать маловероятные гипотезы без значительных затрат вычислительных ресурсов.
3. Методы динамического программирования и обобщённого дистанционного преобразования позволяют эффективно учитывать зависимости между частями объекта. Вместо того чтобы рассматривать каждую часть отдельно, используются общие характеристики и взаимосвязи, что значительно ускоряет процесс обнаружения.
4. Для отсеивания гипотез на каждом этапе каскада используются вероятно допустимые пороги (*PAA thresholds*). Они выбираются на основе статистики частичных гипотез по положительным примерам и гарантируют, что с высокой вероятностью не будут отсеяны гипотезы, ведущие к правильным обнаружениям.

Данный метод показывает высокую эффективность на различных наборах данных, таких как *PASCAL VOC*, демонстрируя способность точно обнаруживать объекты в условиях разнообразных поз, масштабов и фонов.

Метод *DPM* обладает рядом преимуществ, включая гибкость и способность учитывать сложные деформации объектов, что делает его эффективным инструментом для задач обнаружения объектов в условиях реального мира. Однако, несмотря на свою эффективность, метод требует значительных вычислительных ресурсов, что может ограничивать его использование в некоторых задачах.

2.3.4. Случайный лес

Случайный лес (*Random Forest*) — метод ансамблевого обучения, используемый для задач классификации и регрессии. Состоит из множества независимых деревьев решений, которые работают совместно для повышения точности прогноза и уменьшения вероятности переобучения [0].

Алгоритм Случайного леса был впервые предложен в 1995 году Тин Кам Хо и позднее расширен Лео Брейманом и Адель Катлер. Основной принцип метода заключается в объединении большого числа деревьев решений, обученных на различных подмножествах данных и случайных подмножествах признаков, чтобы обеспечить более стабильные и надежные предсказания. В случае задач классификации итоговый результат определяется большинством голосов деревьев, а в задачах регрессии — средним значением их предсказаний.

Одним из ключевых преимуществ Случайного леса является его устойчивость к переобучению. Это достигается за счет использования метода *bagging* (*bootstrap aggregating*), при котором для каждого дерева выбирается случайная подвыборка данных с возвращением, что создает множество различных обучающих наборов. Каждый из этих наборов используется для обучения отдельного дерева, а затем результаты их предсказаний объединяются. Добавление случайности на этапе выбора признаков при построении деревьев также снижает корреляцию между деревьями и способствует уменьшению переобучения.

Среди других преимуществ Случайного леса можно отметить:

- благодаря объединению множества деревьев, метод обеспечивает высокую точность предсказаний;
- может применяться как для задач классификации, так и для регрессии;
- позволяет легко определить, какие признаки оказывают наибольшее влияние на предсказание, используя такие метрики, например, как *Gini importance* и *mean decrease in impurity (MDI)*;
- эффективно справляется с пропущенными данными, сохраняя высокую точность предсказаний.

Однако, у метода есть и свои недостатки. Он может быть ресурсоемким и требовать значительных вычислительных мощностей и времени на обучение, особенно при работе с большими наборами данных. Кроме того, интерпретируемость модели Случайного леса ниже по сравнению с отдельными деревьями решений, что может затруднить объяснение полученных результатов.

2.3.5. Сравнительный анализ

В таблице 2.2 представлено сравнение методов машинного обучения для решения задачи обнаружения объектов на изображении. Рассмотрены преимущества и недостатки таких методов, как Сверточные нейронные сети, Aggregate Channel Features, Deformable Parts Models, Случайный лес.

Таблица 2.2

Сравнение методов машинного обучения

| Метод | Описание | Преимущества | Недостатки |
|---|--|--|--|
| <i>Convolutional Neural Network (CNN)</i> <i>Сверточные нейронные сети</i> | Глубокая нейронная сеть, использующая сверточные слои для автоматического извлечения признаков из изображений. | Высокая точность распознавания, автоматическое извлечение признаков, устойчивость к деформации распознаваемых изображений. | Необходимо большое количество размеченных данных для обучения, требуются большие вычислительные ресурсы (GPU). |

Таблица 2.2

Сравнение методов машинного обучения

| Метод | Описание | Преимущества | Недостатки |
|---|---|--|--|
| <i>Aggregate Channel Features (ACF)</i> | Метод, который извлекает признаки из нескольких каналов изображения (например, оттенки серого, RGB) и агрегирует их для обнаружения объектов. | Подходит для работы в реальном времени, хорошо работает с небольшим количеством данных, высокая точность локализации объектов на изображениях при использовании совместно с бустингом. | Более низкая точность по сравнению с методами глубокого обучения, ограничен простыми признаками, не подходит для сложных сцен. |
| <i>Deformable Parts Models (DPM)</i> | Использует графическую модель частей объекта (неориентированный граф) и их пространственные отношения для обнаружения и локализации объектов. | Может моделировать деформации объектов, хорошо работает с представлением объектов на основе частей. | Вычислительно затратен, может быть медленнее, чем CNN, производительность сильно зависит от качества моделей частей. |

Таблица 2.2

Сравнение методов машинного обучения

| Метод | Описание | Преимущества | Недостатки |
|--|---|---|---|
| <i>Random Forest</i> <i>Случайный лес</i> | Метод ансамблевого обучения, в ходе которого строится большое количество деревьев решений, а результатом классификации является наиболее часто встречающийся класс. | Устойчив к переобучению, хорошая производительность при ограниченных вычислительных ресурсах, может обрабатывать различные типы данных. | Более низкая точность по сравнению с методами глубокого обучения, требует тщательного выбора признаков, может быть чувствителен к шуму. |

Сверточные нейронные сети обеспечивают высокую точность и могут обрабатывать сложные сцены, но требуют больших вычислительных ресурсов и больших наборов данных для обучения. Метод *ACF* эффективен и быстр, подходит для работы в реальном времени, но имеет меньшую точность в сложных сценах. Метод *DPM* хорошо справляется с моделированием деформаций объектов, однако он вычислительно затратен. Случайный лес устойчив к переобучению и работает с ограниченными ресурсами, но может быть менее точным по сравнению с методами глубокого обучения.

2.4. Анализ классических алгоритмов

2.4.1. Вычитание фона

Метод вычитания фона (*background subtraction*) является популярным подходом для выделения движущихся объектов на видео, снятом стационарной камерой. Он служит первым шагом в многоступенчатых системах компьютерного зрения, которые решают задачи мониторинга, распознавания и отслеживания объектов. Основная задача этого метода заключается в том, чтобы на каждом кадре видео отличать пиксели, принадлежащие движущимся объектам, от пикселей фона. Это

достигается путем сравнения текущего кадра с моделью фона, которая представляет собой статическое изображение сцены без движущихся объектов.

Основное уравнение вычитания фона можно представить следующим образом:

$$M_t(s) = \begin{cases} 1, & \text{если } d(C_t(s), B_t(s)) > \theta \\ 0, & \text{иначе} \end{cases}$$

где $M_t(s)$ — метка движения в момент времени t для пикселя s , $B_t(s)$ — модель фона для пикселя s в момент времени t , $C_t(s)$ — цвет пикселя s в момент времени t , d — метрика расстояния между цветом пикселя и моделью фона, θ — пороговое значение.

Методы вычитания фона можно классифицировать на две основные группы: методы, основанные на временных различиях между последовательными изображениями, и методы, использующие сравнение текущего изображения с заранее подготовленными моделями фона. Рассмотрим три наиболее распространенных метода: метод интеркадровых различий, метод приблизительного медианного фильтра и модель гауссовых смесей (*GMM*).

Метод интеркадровых различий использует разницу между последовательными кадрами для выделения движущихся объектов. Его основное преимущество заключается в скорости выполнения, однако он чувствителен к изменениям освещения и подвержен эффекту "камуфляжа", когда объект сливается с фоном по цвету. Обновление модели фона в этом методе можно описать формулой:

$$B_{t+1}(s) = (1 - \alpha)B_t(s) + \alpha C_t(s)$$

где α — константа, определяющая скорость обновления модели фона.

Метод приблизительного медианного фильтра предлагает более устойчивую к шуму модель фона, обновляя её путем медианной фильтрации нескольких последовательных кадров. Этот метод лучше справляется с динамическими изменениями фона, однако требует большего объема вычислений и памяти.

Модель гауссовых смесей (*GMM*) представляет фон как набор гауссовых распределений, что позволяет эффективно справляться с изменениями освещения и динамическими фонами. Несмотря на высокую точность, этот метод требует значительных вычислительных ресурсов и памяти, что делает его менее пригодным для применения в системах с ограниченными ресурсами, таких как встраиваемые системы и *IP*-камеры.

Эффективность рассмотренных методов может оцениваться по таким параметрам, как время вычислений, потребление памяти и точность (*precision/recall*). Результаты исследований [0] показывают, что методы интеркадровых различий и приближительного медианного фильтра значительно быстрее и требуют меньше памяти по сравнению с *GMM*. Однако *GMM* обеспечивает более высокую точность за счет учета сложных изменений фона и освещения.

2.4.2. Оптический поток

Метод оптического потока является важным инструментом в компьютерном зрении, предназначенным для оценки движения объектов в последовательности изображений. Он основывается на предположении о сохранении яркости: интенсивность пикселей переносится из одного кадра в следующий без изменений [0]. Основное уравнение, используемое в данном методе, выглядит следующим образом:

$$I(\mathbf{x}, t) = I(\mathbf{x} + \mathbf{u}, t + 1)$$

где $I(\mathbf{x}, t)$ — интенсивность изображения в точке \mathbf{x} в момент времени t , а $\mathbf{u} = (u_1, u_2)^T$ — двумерная скорость.

Основным подходом к оценке оптического потока является градиентный метод. Он предполагает, что изменения интенсивности во времени малы, что позволяет использовать линейную аппроксимацию:

$$I_t + \nabla I \cdot \mathbf{u} = 0$$

где I_t — временная производная интенсивности, а ∇I — пространственные производные. Для решения этого уравнения используется метод наименьших квадратов, минимизирующий функцию ошибки:

$$E(\mathbf{u}) = \sum_k w_k (\nabla I(\mathbf{x}_k) \cdot \mathbf{u} + I_t(\mathbf{x}_k))^2$$

где w_k — весовые коэффициенты, определяющие вклад каждого пикселя в общую ошибку.

Однако, при использовании этого метода возникают проблемы с неоднозначностью решений, особенно в областях с низкой текстурой. Для преодоления этой проблемы применяются итеративные методы, такие как метод Гаусса-Ньютона, которые уточняют оценку оптического потока за несколько шагов, уменьшая остаточную ошибку.

Одной из важных задач при оценке оптического потока является обработка границ объектов и областей с разными движениями. Для этого используются модели смешивания, которые предполагают наличие нескольких движущихся объектов и учитывают вероятность принадлежности каждого пикселя к той или иной модели движения. Формула функции ошибки для смеси моделей выглядит следующим образом:

$$E(\mathbf{u}_n) = \sum_k q_n(\mathbf{x}_k) [\nabla I(\mathbf{x}_k, t) \cdot \mathbf{u}_n + I_t(\mathbf{x}_k, t)]^2$$

где $q_n(\mathbf{x}_k)$ — вероятность того, что пиксель \mathbf{x}_k принадлежит модели движения \mathbf{u}_n .

Для улучшения точности оценки оптического потока используются много-масштабные представления изображений, которые позволяют избежать локальных минимумов, и устойчивые функции стоимости, способные обрабатывать выбросы. Дополнительно могут использоваться фазовые методы, которые основываются на разложении изображения на полосовые каналы и предполагают сохранение фазы в каждом канале:

$$\nabla \varphi(\mathbf{x}, t) \cdot \mathbf{u} + \varphi_t(\mathbf{x}, t) = 0$$

где $\varphi(\mathbf{x}, t)$ — фаза комплексного полосового сигнала $r(\mathbf{x}, t)$.

Метод оптического потока является мощным инструментом для анализа движения, однако требует дальнейших исследований для улучшения учета изменений яркости в длительных последовательностях изображений.

2.4.3. Сравнительный анализ

Проведем сравнение классических алгоритмов обнаружения объектов на изображениях или в видеопотоке (таблица 2.3).

Таблица 2.3

Сравнение классических алгоритмов

| Метод | Описание | Преимущества | Недостатки |
|------------------|---|--|--|
| Вычитание фона | Обнаруживает движущиеся объекты путем вычитания текущего кадра из эталонной модели фона. | Простой, быстрый и вычислительно эффективный, хорошо подходит для статических настроек камеры. | Проблемы с динамическим фоном, изменениями освещения и сложными сценами, неустойчив к движению камеры. |
| Оптический поток | Вычисляет движение объектов между последовательными кадрами на основе изменений интенсивности пикселей. | Может захватывать информацию о движении, полезен для обнаружения движущихся объектов независимо от фона. | Вычислительно затратен, чувствителен к шуму и быстрым движениям, может иметь проблемы со сложными принципами движения. |

Наилучшие результаты могут быть достигнуты при использовании гибридных подходов. Например, при комбинации сверточных нейронных сетей с алгоритмом оптического потока можно объединить преимущества глубокого обучения для точной классификации и оптического потока для детекции движения. Такой подход позволяет улучшить точность и надежность системы, справляясь как с детекцией сложных объектов, так и с анализом движения в динамичных сценах. Впервые данный подход был представлен в статье [0] как алгоритм под названием «С-ФО».

Внедрение гибридных методов позволяет учитывать сильные стороны каждого из алгоритмов, минимизируя их недостатки. Это особенно важно в условиях

реального времени, где требуется высокая производительность и точность при ограниченных вычислительных ресурсах.

2.5. Разработка гибридного алгоритма

2.5.1. Математическая постановка задачи

Для постановки задачи обнаружения введем стохастический процесс $Y(t)$:

$$Y(t) = \{Y_t \mid M \times A \times I, t\}, \quad (2.9)$$

где:

- Y_t — набор параметров, описывающий позицию обнаруживаемого объекта, в данном случае ограничивающий прямоугольник в измерении изображения с вложенным в него участком этого изображения, в комбинированном пространстве,
- M — пространство признаков движения,
- A — пространство признаков внешнего вида,
- I — пространство признаков инфракрасного излучения,
- t - дискретный временной шаг.

Несмотря на то, что точное математическое описание движения и изображения объекта возможно, в большинстве случаев оно является вычислительно трудноразрешимым, потому процесс является стохастическим.

Функция τ осуществляет отображение признаков объекта в пространство его положений Y

$$\tau : \mathbf{N} \rightarrow M \times A \times I, t \rightarrow \tau(t). \quad (2.10)$$

В рассматриваемой нами задаче τ формируется как сложная функция:

$$\tau(t) = \delta(X(t)), \quad (2.11)$$

где $X(t)$ – последовательность изображений из пространства L_c :

$$X(t) = \{X_t \in L_c, t \in \mathbf{N}\}, \quad (2.12)$$

обрабатываемая функцией:

$$\delta : L_c \rightarrow M \times A \times I, X(t) \rightarrow \delta(X(t)). \quad (2.13)$$

С учетом измерений, учитываемых функцией τ , уравнение (2.11) представим в следующем виде:

$$\tau(t) = \delta(X(t), Y(t - 1)). \quad (2.14)$$

Использование уравнения (2.14) предполагает значительное использование вычислительных ресурсов, поэтому на практике вводится ограничивающий параметр n :

$$\tau(t) = \delta(X(t) \setminus X(t - n), Y(t - 1) \setminus Y(t - n)). \quad (2.15)$$

Суть данной работы заключается в поиске оператора δ , который является алгоритмом обнаружения объектов.

2.5.2. Обзор гибридного алгоритма

Основная задача, которую необходимо решить при обнаружении движущихся объектов с борта беспилотного летательного аппарата (БПЛА), заключается в необходимости различать изменения в кадре, вызванные движением БПЛА, от изменений, вызванных движущимися объектами. Для этого проводится первый этап предобработки изображений, в ходе которого извлекаются особые точки. На следующем этапе происходит классификация объектов с использованием гибридного классификатора.

2.5.3. Оценка параметров движения

Блок-схема алгоритма оценки параметров движения представлена на рис. 2.3. Задача данного алгоритма заключается в предобработке получаемых кадрах, а именно, в формировании перечня потенциальных объектов, которые имеют перемещение во времени. Для этого производится двухэтапная оценка на соответствие фона в текущем и предыдущем кадрах видеопотока: расчет глобальной модели движения на каждом изображении, а затем выравнивание за счет модели движения. Таким образом, разница моделей движения двух последовательных изображений дает представление о подвижных объектах, имеющих непостоянное и отличающееся от фонового изменение параметров движения, что дает возможность детально изучать полученные области на предмет поиска искомых объектов путем формирования особых точек.

Алгоритм оценки движения реализована в пять этапов: выбор точек привязки, локальная оценка движения, глобальная оценка движения, вычитание фона, извлечение особых точек.



Рис.2.3. Блок-схема алгоритма оценки параметров движения

Выбор точек привязки

Для выбора точек привязки используется метод *Good Features to Track*, который был предложен Карло Томази и Джианбо Ши в 1994 году [goodTruck]. Данный метод направлен на выявление таких характеристик в изображениях, которые могут быть надежно отслежены в последовательных кадрах.

Метод основывается на анализе малых матриц автокорреляции изображения. Рассматривается 2×2 матрица Z , элементы которой зависят от градиентов интенсивности изображения в окне размером $W \times W$. Формула для элементов матрицы Z :

$$Z = \sum_{x,y \in W} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2.16)$$

где I_x и I_y — градиенты интенсивности изображения по осям x и y соответственно.

Хорошие характеристики должны удовлетворять двум условиям:

- значения собственных чисел матрицы Z должны быть достаточно большими, чтобы преодолеть уровень шума изображения;

- собственные числа не должны сильно отличаться по величине, что обеспечивает хорошую обусловленность матрицы.

Если собственные числа λ_1 и λ_2 матрицы Z велики и примерно равны, то окно содержит текстуру, которая хорошо отслеживается. В противном случае, окно либо содержит однородную область (оба собственных числа малы), либо одно из собственных чисел значительно меньше другого (одномерный паттерн).

Локальная оценка движения

Извлеченные из предыдущего кадра точки привязки алгоритму оптического потока Лукаса-Канаде с пирамидальной структурой [0]. Цель алгоритма состоит в отслеживании точки u на кадре I_k (где k - номер кадра) и её перемещении в другое местоположение v на кадре I_{k+1} . Разница между точками u и v представляет собой вектор оптического потока для этой точки, который описывает как угол движения, так и величину, с которой точка u перемещается с кадра I_k на v кадр I_{k+1} . Этот процесс применяется ко всем ключевым точкам, извлечённым из предыдущего кадра изображения, для определения их соответствующих положений на текущем кадре изображения.

Глобальная оценка движения

После вычисления векторов смещения оптического потока они нормализуются и объединяются с помощью алгоритма кластеризации *ELM* [0]. Координаты x и y векторов оптического потока, которые охватывают как угол движения, так и смещение оптического потока, используются в качестве признаков для кластеризации. Метод *Evolving Local Means* работает за один проход, рекурсивно вычисляя локальное среднее и дисперсию кластеров.

Кластер с наибольшим количеством векторов оптического потока, связанных с ним, представляет движение камеры относительно фона. Признаки, содержащиеся в самом большом кластере, который отражает движение камеры, затем используются для вычисления матрицы гомографии с использованием алгоритма случайной выборки для улучшения устойчивости и удаления оставшихся выбросов. Выбор *ELM* вместо альтернатив, таких как алгоритм сдвига среднего значения, был обусловлен его вычислительной эффективностью, причем оба подхода обеспечивают сопоставимую точность.

Введение небольших ошибок *ELM* минимизируется при помощи алгоритма случайной выборки консенсуса (*RANSAC*) при вычислении гомографии. Минимальный размер объекта ограничивается процессом кластеризации, удаляются кластеры размером менее 10 пикселей для подавления шума. Если движущийся

объект останавливается на продолжительный период времени, подход перестает его обнаруживать, так как *RDE* является детектором на основе движения. Чистое вращение представляет сложность для предложенного подхода из-за использования оптического потока при вычислении матрицы гомографии.

Вычитание фона

На этом этапе происходит вычитание движущегося фона, чтобы выделить области, в которых не произошли изменения. Это необходимо из-за того, что параметры движения движущихся объектов отличаются от модели движения фона.

Для того чтобы вычислить изображение фона, оценивается предыдущий кадр:

$$\overbrace{X_{n-1}}^n(s) = X_n(T(H_n, s)), \quad (2.17)$$

где $\overbrace{n-1}^n(s)$ - оценка предыдущего кадра, s координаты точки. Затем вычитаемое изображение фона E_{n-1} для X_{n-1} вычисляется как:

$$E_{n-1} = |X_{n-1} - \overbrace{n-1}^n|. \quad (2.18)$$

Извлечение особых точек

На этом этапе определяются точки фона вычитаемого изображения и извлекаются области на исходном изображении и фоне вычитаемого изображения, расположенного вокруг особых точек. Вычитая расчетный фон, движущиеся объекты остаются заметными на вычитаемом изображении. Таким образом определяется местоположение областей, определяющих выделенную точку на вычитаемом фоне изображения.

2.5.4. Гибридный классификатор

Классификатор внешнего вида

Архитектура нейронной сети классификатора внешнего вида состоит из 16 фильтров с размером сверточных ядер 3x3, после которых применяется функция активации ReLU, затем из 32 фильтров с размером ядер 3x3, активацией ReLU и операцией объединения (pooling). Аналогично применяется еще 64 фильтра. В конце сети идет полносвязный слой с функцией активации softmax, где для каждой особой точки $q_{n,i}$ получаем вероятность принадлежности к классу $p_{n,i}^a$.

Также после каждой функции активации были использованы исключаяющие слои (dropout) и пакетная нормализация для регуляризации процесса обучения.

Классификатор движения

Разница в движении $d_{n,i}$ для движущегося объекта при $q_{n-1,i}$ между перспективным и локальным движением определяется следующим образом:

$$d_{n,i} = T(H_n, q_{n-1,i}) - \tilde{q}_{n,i}.$$

где $\tilde{q}_{n,i}$ - соответствующая точка в текущем кадре, полученная с помощью оценки параметров локального движения особых точек с помощью согласования оптического потока Лукаса - Канаде, $v_{n,i}$ - обратное локальное движение, которое можно вычислить с помощью уравнения

$$v_{n,i} = \arg \min_u \sum_{s \in N(\tilde{p}_{n,i})} |X_n(s) - X_{n-1}(s - u)|^2, \quad (2.19)$$

Кроме того, были введены следующие характеристики: $u_{n,i} = \tilde{q}_{n,i} - q_{n-1,i}$ и $h_{n,i} = T(H_n, q_{n-1,i}) - q_{n-1,i}$. Эта разница в движении представляет собой фактическую скорость движущегося объекта по отношению к параметрам движения фона.

Таким образом, можно ввести параметры движения, которые описывают характер движения объекта в оптическом потоке. Данные параметры приведены в таблице 2.4.

| № | Параметр | Описание |
|----|--|---|
| 1. | $l_{n,i} = \ d_{n,i}\ _2$ | Величина разности движений |
| 2. | $\alpha_{n,i} = \arctan(d_{n,i})$ | Угол разности движений |
| 3. | $\varepsilon_{n,i} = \ u_{n,i} - v_{n,i}\ _2$ | Двунаправленное контрольное расстояние |
| 4. | $\theta_{n,i} = \arctan(h_{n,i}) - \arctan(u_{n,i})$ | Перспективно-локальная разница углов движения |
| 5. | $\delta_{n,i} = h_{n,i}\ _2 - \ u_{n,i}\ _2 $ | Перспективно-локальная разница амплитуды движения |

Таблица 2.4

Значения параметров движения объекта

После составления из данных значений вектор-признаков был обучен классификатор селекции ложных объектов, основанный на полносвязной нейронной сети.

Структурная схема архитектуры данной нейронной сети представлена на рисунке 4 и состоит из 8, 16 и 32 последовательных полносвязных слоев, после каждого из которых применяется функция активации. В конце сети идет полносвязный слой с функцией активации softmax, где для каждой особой точки $q_{n,i}$ получаем вероятность принадлежности к классу $p_{n,i}^0$.

Аналогично классификатору внешнего вида после каждой функции активации были использованы исключаяющие слои (dropout) и пакетная нормализация для регуляризации процесса обучения.

Ансамблирование

На этом этапе используется принцип бустинга моделей, в частности алгоритм AdaBoost [0] для модели классификатора внешнего вида и модели селекции ложных объектов.

В качестве модели бустинга выбрана модель решающего дерева с максимальной глубиной, равной 2. Таким образом, получается вычислительно простое решение бустинга:

$$\beta_{n,i} = \sum_{m=1}^{M_0} w_m * g_m \left(p_{n,i}^a, p_{n,i}^0 \right), \quad (2.20)$$

где g_m является классификатором дерева решений, $p_{n,i}^a$ и $p_{n,i}^0$ - вероятности, получаемые из соответствующих моделей классификации внешнего вида и селекции ложных объектов, $w_* = \{w_i\}_{i=1}^{M_0}$ означает веса для ансамбля моделей.

Таким образом, финальная метка класса определяется, как:

$$y_{n,i} = \begin{cases} 1, & \text{если } \beta_{n,i} > 0,5 \\ 0, & \text{иначе} \end{cases}, \quad (2.21)$$

где 1 означает, что движущийся объект при $q_{n,i}$ является истинным, а 0 - ложным.

2.6. Оценка точности обнаружения

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

0. Aggregate channel features for multi-view face detection / B. Yang [и др.] // IEEE International Joint Conference on Biometrics. — 2014. — С. 1—8. — DOI 10.1109/BTAS.2014.6996284.

0. Bouguet J.-Y. Pyramidal implementation of the lucas kanade feature tracker //. — 1999. — URL: <https://api.semanticscholar.org/CorpusID:9350588>.

0. Brutzer S., Höferlin B., Heidemann G. Evaluation of background subtraction techniques for video surveillance // CVPR 2011. — 2011. — С. 1937—1944. — URL: <https://api.semanticscholar.org/CorpusID:206591471>.

0. Dutta Baruah R., Angelov P. Evolving local means method for clustering of streaming data //. — 2012. — С. 1—8. — DOI 10.1109/FUZZ-IEEE.2012.6251366.

0. Felzenszwalb P. F., Girshick R. B., McAllester D. Cascade object detection with deformable part models // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — 2010. — С. 2241—2248. — DOI 10.1109/CVPR.2010.5539906.

0. Fleet D., Weiss Y. Handbook of Mathematical Models in Computer Vision //. — 2006.

0. Freund Y., Schapire R. E. A Short Introduction to Boosting //. — 1999. — URL: <https://api.semanticscholar.org/CorpusID:9621074>.

0. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet classification with deep convolutional neural networks // Communications of the ACM. — 2012. — Т. 60. — С. 84—90. — URL: <https://api.semanticscholar.org/CorpusID:195908774>.

0. On-Board Object Detection: Multicue, Multimodal, and Multiview Random Forest of Local Experts / A. González [и др.] // IEEE Transactions on Cybernetics. — 2017. — Т. 47, № 11. — С. 3980—3990. — DOI 10.1109/TCYB.2016.2593940.

0. Гришин Н.А. Б. А. АЛГОРИТМ РАСПОЗНАВАНИЯ ЛЕТАЮЩИХ ОБЪЕКТОВ В ОПТИЧЕСКОМ ВИДЕОПОТОКЕ В УСЛОВИЯХ ДИНАМИЧЕСКИ ИЗМЕНЯЮЩЕГОСЯ ФОНА // Вестник Концерна ВКО Алмаз-Антей. — 2022. — Т. 3 (42). — URL: <https://cyberleninka.ru/article/n/algoritm-raspoznavaniya-letayuschih-obektov-v-opticheskom-videopotoke-v-usloviyah-dinamicheskii-izmenyayuschegosya-fona>.

0. Инсаров В.В. Т. С. Использование градиентного подхода в задаче выделения контуров крупных техногенных объектов на изображениях многообъектных

наземных сцен // Мехатроника, автоматизация, управление. — 2015. — Т. 6. — С. 415—420. — URL: <https://doi.org/10.17587/mau.16.415-421>.

0. *М. М. Л.* Математические методы распознавания образов. — Курс лекций, МГУ, ВМиК, кафедра «Математические методы прогнозирования», 2004.

0. *Полякова Л.А. Прожерин В.Г. С. Я.* Средства инженерно-технического обеспечения охраны объектов. — СПб: Университет ИТМО, 2014. — 56 с.

0. *Ю. В. Визильтер С. Ю. Ж.* Проблемы технического зрения в современных авиационных системах // Техническое зрение в системах управления мобильными объектами. — 2010. — Т. 4. — С. 11—44.