

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**
(**Н И У « Б е л Г У »**)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК
КАФЕДРА ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ
СИСТЕМ И ТЕХНОЛОГИЙ

**ИССЛЕДОВАНИЕ МЕТОДОВ И РАЗРАБОТКА АЛГОРИТМА
РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ВИДЕО-ДАННЫХ**

Выпускная квалификационная работа
обучающегося по направлению подготовки
11.04.02 Инфокоммуникационные технологии и системы связи,
магистерская программа «Системы и устройства радиотехники и связи»
очной формы обучения, группы 07001636
Наумчик Анастасии Александровны

Научный руководитель
кандидат технических наук, доцент,
доцент кафедры ИТСиТ Заливин А. Н.

Рецензент
Соловьев Виктор Иванович к.т.н.,
начальник отдела программного
обеспечения информационных средств
ООО "НПП "ЭИТ" БелГУ"

БЕЛГОРОД 2018

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	4
ГЛАВА 1. АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ.....	6
1.1 Распознавание объектов с помощью нейронных сетей искусственного происхождения.....	6
1.2 Интенциональные методы	10
1.3 Экстенциональные методы.....	11
1.4 Метод оценки эффективности распознавания образов.....	14
1.5 Области интереса гистограммы направленных градиентов (HOG – дескрипторы)	17
1.6 Метод SIFT	23
1.7 Метод SURF	26
1.8 RANSAC метод.....	30
ГЛАВА 2. СИСТЕМА ОБРАБОТКИ ИЗОБРАЖЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ	37
2.1 Предварительная обработка перед распознаванием.....	40
2.2 Определение ключевых особенностей изображения.....	42
2.3 Детектирование объектов.....	43
2.4 Воздействие шума на точность детектирования.....	46
2.5 Распознавание объектов с применением обнаружения области и выделением признаков	48
ГЛАВА 3. РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ	52
3.1 Алгоритм RANSAC.....	52
3.2 Алгоритм распознавания объектов с применением обнаружения области и выделением их признаков	55
3.3 Распознавание методом SIFT	56
ГЛАВА 4. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ РАСПОЗНАВАНИЯ.....	58
4.1 Распознавание методом SURF	58
4.2 Вычисление методом RANSAC	60

4.3 Распознавание методом SIFT	61
ЗАКЛЮЧЕНИЕ	69
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	70
ПРИЛОЖЕНИЕ А	78
ПРИЛОЖЕНИЕ Б.....	79

ВВЕДЕНИЕ

Распознавание визуальных образов представляет собой один из важнейших компонентов систем управления и обработки информации, автоматизированных систем и систем принятия решений. Задачи, связанные с классификацией и идентификацией предметов, явлений и сигналов, характеризующихся конечным набором некоторых свойств и признаков, возникают в таких отраслях как робототехника, информационный поиск, мониторинг и анализ визуальных данных. Алгоритмическая обработка и классификация изображений применяются в системах безопасности, контроля и управления доступом, в системах видеонаблюдения, виртуальной реальности и информационных поисковых системах.

В настоящее время существует множество задач, в которых требуется принять некоторое решение в зависимости от присутствия на изображении объекта или классифицировать его. Это связано с массовым внедрением информационных технологий в различные области человеческой жизни. Обычно они представляют собой набор взаимосвязанных функциональных систем, таких как системы сбора информации с детекторов транспорта и телекамер. Вследствие многообразия внешних условий, в которых приходится работать указанным системам, методики, используемые при их создании, могут сильно отличаться друг от друга. Однако большинство существующих на сегодняшний день систем имеет в своем составе два основных блока: получения изображения (видеокамера) и системы его последующего анализа.

Так, актуальной проблемой, остается распознавание объектов под действием преобразования плоскости изображения, способных значительным образом изменить его форму, не влияя при этом на принадлежность объекта к категории распознавания. Существуют и успешные продвижения при выстраивании видео фиксации и комплексе выявления любого рода

объектов, но есть затруднения при перемещении к наибольшему разрешению вводной в видеоряде по причине использования закрепленных характеристик объема и позиционирования субъекта в снимке. Актуальность данной проблемы особенно высока в отраслях, где распознавание образов применяется в естественной среде (видеонаблюдение, анализ данных камер мониторинга, робототехнические зрительные системы), где зрительный сенсор может иметь произвольный ограниченный угол обзора по отношению к искомому объекту.

Для решения поставленной цели определения предметов, важным фактором является поиск приемлемого алгоритма, который не только определяет данные об особенностях какой-либо определенной области, а также позволяет распознавать любой объект в присутствии влияния шума на изображении.

Основной целью работы является исследование и разработка алгоритма для прецедентного распознавания объектов на видео-данных.

Объектом исследования являются видео-данные, применяемые для распознавания объектов.

Предметом исследования являются методы и алгоритмы распознавания объектов на видео-данных с задачей улучшения систем их видеофиксации.

Задачи дипломной работы:

- произвести анализ существующих методов распознавания на видео-данных;
- выбор методов для исследования распознавания объектов;
- сравнение выбранных методов распознавания объектов;
- разработать алгоритмы распознавания объектов на видео-данных;
- выполнить экспериментальные исследования разработанных алгоритмов распознавания объектов;
- выполнить общий анализ результатов исследуемых алгоритмов распознавания.

ГЛАВА 1. АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

1.1 Распознавание объектов с помощью нейронных сетей искусственного происхождения

Нейронные сети искусственного происхождения имеют представление математической модели, которая работает за счет функционирования биологических нейронных сетей – это сети нервных клеток живого организма. Так же, как и в биологической нейронной сети, главной составляющей искусственного происхождения нейронной сети считается нейрон. Объединенные нейроны, формируют, так сказать, пласты, их численность имеет возможность различаться в зависимости от трудности нейронной сети, а также и задач, которые она решает. Абстрактные почвы программирования подобных нейронных сетей имеют описание во множественных работах.

Нейросетевые методы – это методы, которые основываются на том, что используются всевозможными разновидностями нейронных сетей. Нейронные сети [1, 2, 3] уже издавна и весьма эффективно применяются для решения задачи распознавания. Главный замысел, существующий в базисе нейронных сетей – это очередность трансформации сигнала, которая взаимодействует с параллельно с работоспособными компонентами, и, невзирая на значительные расхождения, отдельные виды нейронных сетей они располагают некоторыми общими чертами. В основании каждой нейронной сети располагаются, в некоторой степени простые, и в большей части однотипные элементы, так сказать ячейки, которые копируют действие нейронов мозга.

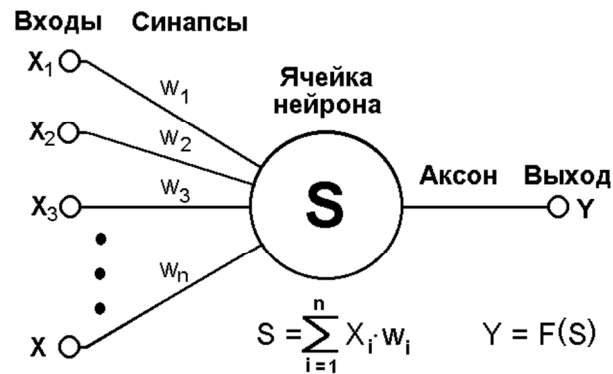


Рисунок 1.1 - Строение формального нейрона

Нейроны характеризуются своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов. У каждого синапса есть такая характеристика как величина синоптической связи или ее вес w_i . Совокупный вид нейрона приведен на рисунке 1.1.

Состояние нейрона производится по формуле взвешенной суммы его входов:

$$s = \sum_{i=1}^n x_i \cdot w_i . \quad (1.1)$$

Выход нейрона есть функция его состояния:

$$y = f(s) . \quad (1.2)$$

Нелинейная функция f называется задеистойванной и принимает разноплановый вид. Из всех широко распространенных известной является нелинейная функция – функция с насыщением, ее еще можно назвать логической функцией, функцией сигмоид – функция S-образного вида.

$$f(x) = \frac{1}{1 + e^{-\alpha x}}. \quad (1.3)$$

Другой особенностью свойственной нейронным сетям является принцип параллельной обработки сигналов, такого рода принцип достигается путем ассоциации крупного числа нейронов в слое и состыковки конкретным способом нейронов различных уровней, а при неких комплектациях и нейронов в общем слое между ними, причем, в таком случае, процессинг совместной работы всех нейронов осуществляется послойно. На рисунке 1.2 указана структура однослойной нейронной сети.

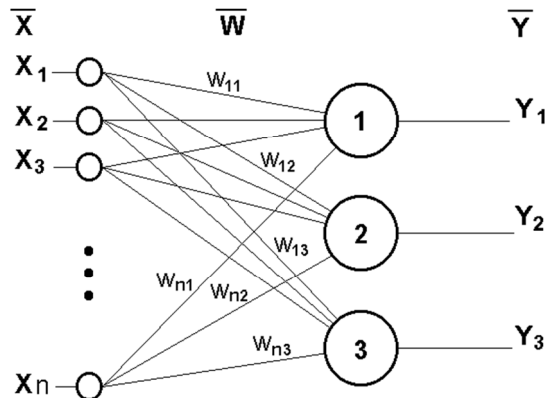


Рисунок 1.2 - Структура связей однослойной нейронной сети

В применении нейронных сетей есть преимущество и недостаток. В первом случае для распознавания человеческого лица – способность принятия классификатора, который отлично создает затруднительную функцию расклада изображений человеческих лиц $p(x | face)$. Второй случай – недостаток – потребность при тщательной конфигурации нейросети, в

последствии устраиваемого итога классификации. Основные преимущества, которыми обладают нейронные сети следующие.

Несмотря на достоинства нейронных сетей, есть факт, что при их использовании в плане изображений нужны особые старания. Что в первую очередь говорит о сложном характере изображений, в частности трехмерных объектов реального мира, а это лица людей.

Существует ориентировочно десяток ключевых видов нейронных сетей, также их основы производительности структур, сделанных на базе автоассоциативной памяти (АП). Это сформировывается на ответной реакции к некоему входящему статистическому показателю данных, их можно назвать «ключом», а также обеспечить к заключению хранящуюся в сети близко расположенную к вводной по величине статистическое выражение такого же размера [4, 5]. Когда происходит такой момент как опознание по снимку, то зацепкой является изображение человеческого облика [4, 6, 7, 8, 9, 10].

Процесс работоспособности нейронных сетей основанный на базисе автоассоциативной памяти заключается в некой последовательности:

- 1) Изображение оцифровывается и кодируется в виде вектора;
- 2) Все координаты вектора размещаются в собственной ячейке, которые соединяются с другими ячейками;
- 3) Векторы перерабатываются нейронной сетью, и на ее выходе образовывается близко расположенное к вводу сохраненному в АП фотоснимку.

Линейная автоассоциативная память является одним слоем нейронной сети, а всякий нейрон такого слоя сравнивается с некоторым элементом, который происходит от декомпозиции изображения лица вектора, по аналогии сравним с методом «собственных лиц». Из этого следует что при параметрах изображения $m \times n$ пикселей, слои такой сети будут иметь $m \times n$ нейронов. Помимо этого, нейроны имеют некую связь с остальными нейронами, а линейная АП конструируется на расчете $m \times n$ весов

соединений нейронной сети. Это обуславливается на уровне обучения, при нем изображения, которые представляются АП, остаются в ее памяти.

1.2 Интенциональные методы

Отличительной особенностью интенциональных методов является то, что в качестве элементов операций при построении и применении алгоритмов распознавания образов они применяют различные характеристики признаков и их связей. Такими элементами бывают отдельные значения или интервалы значений признаков, средние величины и дисперсии, матрицы связи признаков, над которыми производятся действия, выражаемые в аналитической или конструктивной форме. Причем объекты в данных методах не рассматриваются как целостные информационные единицы, а играют роль индикаторов для оценки взаимодействия и поведения своих атрибутов. Группа данных методов распознавания объектов велика, и ее деление на подклассы носит в определенной мере условный характер [11].

1) Методы, основанные на оценках плотностей распределения значений признаков

Такие методы распознавания заимствованы из классической теории статистических решений, где объекты исследования рассматриваются как реализации многомерной случайной величины, распределенной в пространстве признаков по какому-либо закону. Они базируются на байесовской схеме принятия решений, апеллирующей к априорным вероятностям принадлежности объектов к тому или иному распознаваемому классу и условным плотностям распределения значений вектора признаков. Данные методы сводятся к определению отношения правдоподобия в различных областях многомерного пространства признаков.

Главными проблемами использования этих методов считаются необходимость запоминания всей обучающей выборки для вычисления оценок локальных плотностей распределения вероятностей и высокая чувствительность к не респектабельности обучающей выборки [12].

2) Методы, основанные на предположениях о классе решающих функций

В данной группе методов считается известным общий вид решающей функции и задан функционал ее качества. На основании чего, по обучающей последовательности находят наилучшее приближение решающей функции. Самыми распространенными являются представления решающих функций в виде линейных и обобщенных нелинейных полиномов. Функционал качества решающего правила обычно связывают с ошибкой классификации [11].

Основным достоинством методов, основанных на предположениях о классе решающих функций, является ясность математической постановки задачи распознавания, как задачи поиска экстремума. Разнообразие методов этой группы описывается широким спектром используемых функционалов качества решающего правила и алгоритмов поиска экстремума. Выводом рассматриваемых алгоритмов, к которым относятся, в частности, алгоритм Ньютона, алгоритмы перцептронного типа и др., является метод стохастической аппроксимации. Потенциалы градиентных алгоритмов поиска экстремума, особенно в группе линейных решающих правил, достаточно хорошо изучены. Схожесть этих алгоритмов доказана только для случая, когда распознаваемые классы объектов отображаются в пространстве признаков компактными геометрическими структурами [13].

1.3 Экстенциональные методы

В методах данной группы, в отличие от интенционального направления, каждому изучаемому объекту в большей или меньшей мере придается самостоятельное диагностическое значение. По своей сути эти

методы близки к клиническому подходу, который рассматривает людей не как проранжированную по тому или иному показателю цепочку объектов, а как целостные системы, каждая из которых индивидуальна и имеет особенную диагностическую ценность [14]. Такое бережное отношение к объектам исследования не позволяет исключать или утрачивать информацию о каждом отдельном объекте, что происходит при применении методов интенционального направления, использующих объекты только для обнаружения и фиксации закономерностей поведения их атрибутов.

Основными операциями в распознавании образов с помощью данных методов являются операции определения сходства и различия объектов. Объекты в указанной группе методов играют роль диагностических прецедентов. При этом в зависимости от условий конкретной задачи роль отдельного прецедента может меняться в самых широких пределах: от главной и определяющей и до весьма косвенного участия в процессе распознавания. В свою очередь условия задачи могут требовать для успешного решения участия различного количества диагностических прецедентов: от одного в каждом распознаваемом классе до полного объема выборки, а также разных способов вычисления мер сходства и различия объектов. Этими требованиями объясняется дальнейшее разделение экстенциональных методов на подклассы [11].

1) *Метод сравнения с прототипом*

Это наиболее простой экстенциональный метод распознавания. Он применяется, например, в том случае, когда распознаваемые классы отображаются в пространстве признаков компактными геометрическими группировками. В таком случае обычно в качестве точки — прототипа выбирается центр геометрической группировки класса (или ближайший к центру объект).

Для классификации неизвестного объекта находится ближайший к нему прототип, и объект относится к тому же классу, что и этот прототип.

Очевидно, никаких обобщенных образов классов в данном методе не формируется.

В качестве меры близости могут применяться различные типы расстояний. Часто для дихотомических признаков используется расстояние Хэмминга, которое в данном случае равно квадрату евклидова расстояния. При этом решающее правило классификации объектов эквивалентно линейной решающей функции.

Указанный факт следует особо отметить. Он наглядно демонстрирует связь прототипной и признаковой репрезентации информации о структуре данных. Пользуясь приведенным представлением, можно, например, любую традиционную измерительную шкалу, являющуюся линейной функцией от значений дихотомических признаков, рассматривать как гипотетический диагностический прототип. В свою очередь, если анализ пространственной структуры распознаваемых классов позволяет сделать вывод об их геометрической компактности, то каждый из этих классов достаточно заменить одним прототипом, который фактически эквивалентен линейной диагностической модели.

На практике, безусловно, ситуация часто бывает отличной от описанного идеализированного примера. Перед исследователем, намеревающимся применить метод распознавания, основанный на сравнении с прототипами диагностических классов, встают непростые проблемы.

Во-первых, это выбор меры близости (метрики), от которого может существенно измениться пространственная конфигурация распределения объектов. Во-вторых, самостоятельной проблемой является анализ многомерных структур экспериментальных данных. Обе эти проблемы особенно остро встают перед исследователем в условиях высокой размерности пространства признаков, характерной для реальных задач [11].

2) *Метод k ближайших соседей*

Метод k ближайших соседей для решения задач дискриминантного анализа заключается в следующем.

При классификации неизвестного объекта находится заданное число (k) геометрически ближайших к нему в пространстве признаков других объектов (ближайших соседей) с уже известной принадлежностью к распознаваемым классам. Решение об отнесении неизвестного объекта к тому или иному диагностическому классу принимается путем анализа информации об этой известной принадлежности его ближайших соседей, например, с помощью простого подсчета голосов.

Первоначально метод k ближайших соседей рассматривался как непараметрический метод оценивания отношения правдоподобия. Для этого метода получены теоретические оценки его эффективности в сравнении с оптимальным байесовским классификатором. Доказано, что асимптотические вероятности ошибки для метода k ближайших соседей превышают ошибки правила Байеса не более чем в два раза.

При использовании метода k ближайших соседей для распознавания образов исследователю приходится решать сложную проблему выбора метрики для определения близости диагностируемых объектов. Эта проблема в условиях высокой размерности пространства признаков чрезвычайно обостряется вследствие достаточной трудоемкости данного метода, которая становится значимой даже для высокопроизводительных компьютеров. Поэтому здесь так же, как и в методе сравнения с прототипом, необходимо решать творческую задачу анализа многомерной структуры экспериментальных данных для минимизации числа объектов, представляющих диагностические классы.

Необходимость уменьшения числа объектов в обучающей выборке (диагностических прецедентов) является недостатком данного метода, так как уменьшает представительность обучающей выборки [11].

1.4 Метод оценки эффективности распознавания образов

генерируются как реализации многомерной случайной величины с заданным законом распределения по одному из известных алгоритмов [15].

Перед тем как поступить на вход алгоритма распознавания, численные значения параметров представляющие собой распознаваемый объект, подвергаются случайному искажению, что создает результат воздействия различных помех в ходе определения признаков при использовании соответствующих технических средств с конкретными точностными характеристиками. Искаженные значения представляющие наблюдаемый объект в том виде, в каком его воспринимает система, поступают на вход блока распознавания, в котором определяется принадлежность объекта одному из классов сопоставляет номер класса, к которому отнесен объект блоком распознавания, с точным номером, т. е. с тем, который задавался на первом этапе формирования объекта, определяет верность распознавания объекта и систематизирует соответствующую информацию для подсчета оценок вероятностей верных и ошибочных решений. При распознавании объектов из класса, оценкой вероятности получения правильного решения служит отношение количества правильных ответов к общему числу испытаний над объектами класса

$$\omega_i \approx \frac{N_{\text{пр}}^i}{N^i} . \quad (1.4)$$

Число испытаний определяется доверительной вероятностью, которая задается при формулировке задачи исследования.

В зависимости от задачи исследования под влияние искажения попадают также априорные данные о классах объектов, т. е. функции распределения информация о признаках может урезаться, что соответствует отсутствию некоторых средств определения признаков, и т. п.

Если априорные вероятности появления объектов из разных классов известны, то

$$W = \sum_{i=1}^m \omega_i P(\Omega_i), \quad (1.5)$$

представляющая безусловную вероятность правильного решения задачи распознавания системой, может быть выбрана в качестве показателя эффективности системы распознавания.

Рассмотренная статистическая модель дает возможность найти зависимость показателя эффективности системы от вида и от количества привлекаемых для распознавания признаков и точности технических средств, которыми оснащается распознающая система:

$$W = W(x_1, \dots, x_N; \sigma_1, \dots, \sigma_S). \quad (1.6)$$

Данные сведения исходные для задач об определении состава технических средств наблюдений системы распознавания, необходимых точностей их работы; об оптимальном с точки зрения экономических соображений распределении точностей по средствам и т. д. Таким образом, экспериментирование со статистической моделью системы распознавания позволяет без обращения к лабораторным и натурным экспериментам решать задачи, связанные с построением оптимальных систем распознавания [15].

1.5 Области интереса гистограммы направленных градиентов (НОГ – дескрипторы)

Крайне важным, хоть и промежуточным, этапом решения задач компьютерного зрения является описание изображения. В время описания изображения происходит изменение визуальных данных, которые находятся на изображении, в доступную для алгоритма классификации форму.

Ввиду того, что в визуальных данных как правило содержится масса избыточной информации, задача описания в том числе состоит в устранении наибольшего количества малоинформативных признаков при сохранении данных обо всех значительных для задачи распознавания параметрах исследуемого фрагмента изображения.

Гистограмма направленных градиентов (HOG) – дескрипторы особых точек, применяемые в технической классификации объектов и обработке изображений для распознавания. Этот метод базируется на подсчитывании количества направлений градиента в локальных областях изображения, также он однороден с гистограммами направления края и дескрипторами SIFT. Отличается он тем, что в нем вычисление происходит на полной сетке, на которой равномерно распределены ячейки, также данный метод применяет нормализацию перекрывающегося локального контраста, чтобы увеличить точность.

Основа предоставляемого алгоритма – внешний вид и форма распознаваемого объекта на какой-либо зоне изображения, отображаются распределением градиентов интенсивности и направлением на нем краев. Осуществление таких дескрипторов может быть произведено следующим образом: происходит деление изображения на небольшие связующие между собой области, так называемые, ячейками; далее рассчитывается для каждой такой ячейки гистограмма направлений градиентов, а для направлений краев – пиксели, которые располагаются внутри ячейки. Сочетание гистограмм и есть дескриптор $h = f(h_1, \dots, h_k)$, где данное представление нормализуется по яркости (L_1 и L_2 норма). Для того чтобы увеличить точность, локальные гистограммы переходят в операцию нормализации по контрасту. Для этого происходит вычисление меры интенсивности на фрагменте изображения, именуемый блоком, а выведенное значение применяется для выполнения нормализации. Дескрипторы, которые прошли нормализацию, обладают лучшей инвариантностью по отношению к освещению.

НОГ дескриптор обладает несколькими преимуществами в отличие от других. Так как он работает локально, то происходит поддержание инвариантности геометрических и фотометрических преобразований, исключением является ориентация объекта. Аналогичные изменения появляются лишь в больших частях изображения. По данным исследований, разбиение пространства, точный расчет направлений градиента и локальная световая нормализация дают возможность избегать движения объектов, в том случае, если они сохраняют вертикальное положение. Таким образом, НОГ дескриптор является удобным средством нахождения требуемых объектов на изображениях.

Ниже представлена реализация алгоритма:

1) Вычисление градиента. Первым этапом вычисления в детекторах является нормализация цвета и гамма-коррекция. В своем опыте Далал и Триггс выяснили, что для НОГ дескриптора данное действие выполнять не обязательно, потому как следующая нормализация покажет тот же самый результат. В итоге, просто рассчитываются значения градиентов. Здесь самым наиболее применяемым методом считается применение одномерной дифференцирующей маски в горизонтальном и/или вертикальном направлении. Данная процедура нуждается в цветовой или яркостной составляющей фильтрации, используя вспомогательные ядра, такие как: $[-1, 0, 1]$ и $[-1, 0, 1]^T$.

Для эксперимента применялись более сложные маски, такие как Собель 3x3 (оператор Собеля) или, их еще называют, диагональные маски. Но, как показал результат, данные маски показали более низкий коэффициент полезного действия для данной задачи. Также проводилось экспериментирование с размытием по Гауссу перед потреблением дифференцирующей маски, но также выяснилось, что пропуск данного этапа увеличивает работоспособность без очевидной потери качества.

Макси Собеля, применяемые для анализа градиента функции интенсивности, выглядят следующим образом:

$$M_x \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad M_y \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ -1 & -2 & -1 \end{pmatrix}$$

2) Группировка направлений. Здесь происходит вычисление гистограммы ячеек, где каждый пиксель принимает участие во взвешенном голосовании для каналов гистограммы направлений, которое базируется на значении градиентов. Ячейки могут применять как круглую, так и прямоугольную форму. Каналы гистограммы равномерно распределяются от 0 до 180° , либо от 0 до 360° , в зависимости от того, как вычисляется «знаковый» или «без знаковый» градиент. В результате обнаружилось, что «без знаковый» градиент вместе с девятью каналами гистограммы показывает при распознавании положительные результаты. При распределении значений в голосовании, значение пикселя может быть задано либо абсолютным значением градиента, либо некоторой функцией от него (на практике в тестах абсолютное значение градиента дает положительные результаты).

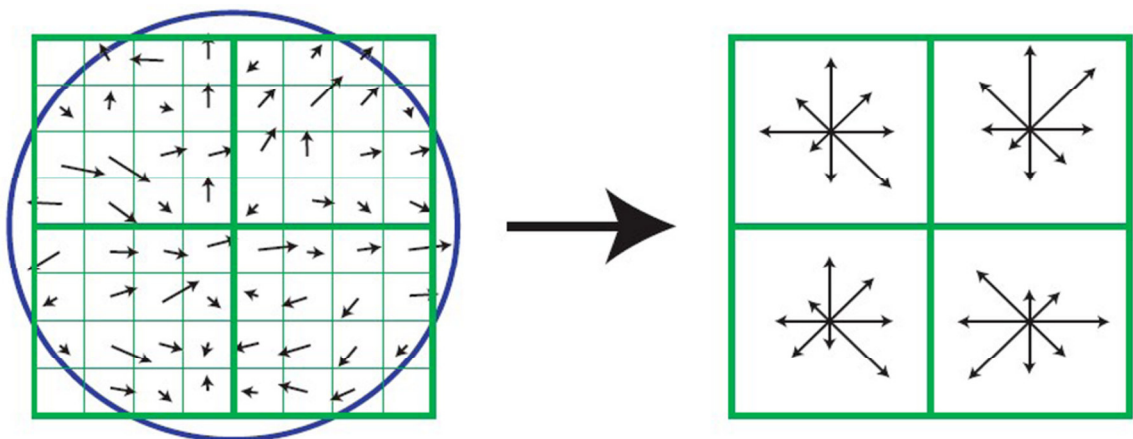


Рисунок 1.4 – Вычисление гистограмм направлений градиента в ячейке изображения

3) Блоки дескрипторов. Чтобы учесть яркость и контрастность, градиенты требуется локально нормировать, для этого группируются ячейки

в более объемные связанные между собой блоки. HOG дескриптор, в данном случае, будет выполнять задачу вектора компонента нормированных гистограмм ячеек их все частей блока. Чаще всего блоки перекрываются, то есть каждая ячейка будет попадать более чем в один конечный дескриптор, применяя при этом две геометрии блока – прямоугольные R-HOG и круглые C-HOG. Блоки R-HOG, как правило, имеют вид квадратной сетки, имеющей следующие характеристики: количеством пикселей на ячейку, количеством каналов на гистограмму ячейки, количеством ячеек на блок. Наиболее приемлемыми параметрами являются блоки 16x16, ячейки 8x8 и 9 каналов на гистограмму. Отсюда вывод, что можно немного поднять скорость расчетов, используя гауссов фильтр внутри каждого блока, это будет снижать показатели значений пикселей на границах блоков. Блоки R-HOG в некоторой степени идентичны с SIFT – дескрипторами, но в отличие от них, блоки R-HOG рассчитываются на плотных сетках закрепленного масштаба без закрепленного направления, тогда как SIFT-дескрипторы рассчитываются на плотных сетках определенного масштаба без определенного направления, тогда как SIFT-дескрипторы рассчитываются в разреженных, не чувствительных к масштабу ключевых точках и применяют оборот для выравнивания. Помимо этого, чтобы закодировать информацию о модели объектов, совокупно используются блоки R-HOG, тогда как SIFT-дескрипторы применяются по одному.

Блоки C-HOG обладают двумя видами: с единой основной ячейкой; с ячейкой, делящейся на сектора. Данные блоки характеризуются четырьмя параметрами: численность областей и колец, радиус центрального кольца, коэффициент увеличения для радиусов прочих колец. Эти две разновидности дают один и тот же результат, а, в свою очередь, деление на два кольца и четыре сектора с радиусом в четыре пикселя, коэффициентом увеличения на два, выдает положительный результат. К тому же, гауссово взвешивание не предоставляет никаких повышений качества, используя C-HOG блоки, также эти блоки схожи на отрывки формы, но также они содержат в себе одно

отличие – блоки C-HOG имеют ячейки с несколькими каналами направлений, тогда как контексты формы применяют лишь только присутствие одного края.

4) Нормализация блоков. Всего в эксперименте четыре метода нормализации. Пусть v — ненормированный вектор, имеющий все гистограммы данного блока, $\|v\|_k$ — его k -норма при $k = 1, 2$ и e — некая малая величина. Тогда нормировочный множитель можно получить одним из следующих способов:

$$\text{L2-норма: } f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$$

L2-hys: L2-норма ограничивается сверху (значения v , бóльшие 0,2, полагаются равными 0,2) и перенормируется

$$\text{L1-норма: } f = \frac{v}{(\|v\|_1 + e)}$$

$$\text{корень из L1-нормы: } f = \sqrt{\frac{v}{(\|v\|_1 + e)}}$$

Экспериментально выявлено, что употребление нормы L1 приводит к выводу наименее положительных результатов, чем применение некоторых приведенных норм. Нормы L2 и корень из L1-нормы предоставляют идентичные выводы, доставляя результаты почти одинаковой точности. Причем применение какого угодно из данных методов гораздо повышают положительный результат в отличие ситуации отсутствия нормировки.

Приобретенные в конечном итоге расчеты HOG дескрипторов имеют несколько значительных достоинств перед остальными часто встречающимися признаками. Так как гистограммы направлений градиента рассчитываются на ограниченных участках изображения, то они инвариантны к неким геометрическим изменениям формы, к примеру, поворот и сдвиг. Процесс нормировки гистограмм в пределах ячеек предоставляет инвариантность к световым изменениям, таким как перемена яркости и контраста изображений. Недостатки вышеописанного подхода -

недостаток инвариантности к преобразованию масштаба и ориентации объекта.

Заключительный этап расчетов HOG–дескрипторов – получение вектора признаков. Он получается с помощью группировки всех имеющихся элементов нормированных блоков гистограмм.

На рисунке 1.5 схематично показан процесс вычисления HOG–дескрипторов на примере распознавания номера.

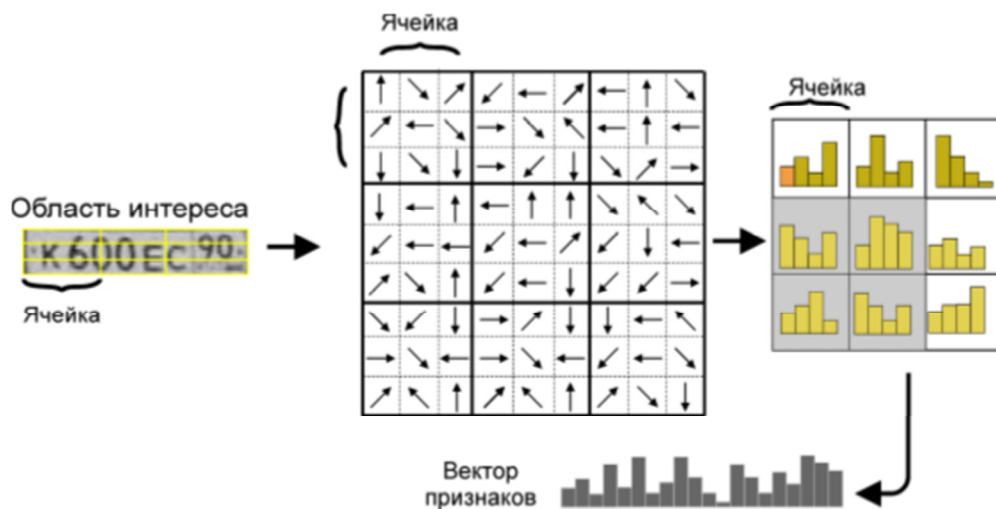


Рисунок 1.5 – Схема вычисления HOG – дескрипторов

В результате расчетов, получившийся вектор ставится в соответствии первичному изображению области интереса. Этот вектор во всем описывает данное изображение, в том числе и его наиболее информативные признаки, являющиеся нужными в рамках поставленной задачи.

1.6 Метод SIFT

Метод Scale Invariant Feature Transform – распознает локальные характеристики изображения и описывает их. С помощью этого метода,

можно получить инвариантные признаки в отношении размера и поворота, которые не зависимы от шума и ряда афинных преобразований.

В алгоритме используется пирамида Гаусса, строящаяся на изображении. После этого, изображения формируются в один размер и происходит вычисление разницы между ними (DoG, difference-of-Gaussian images). На рисунке 1.6 показан данный процесс.

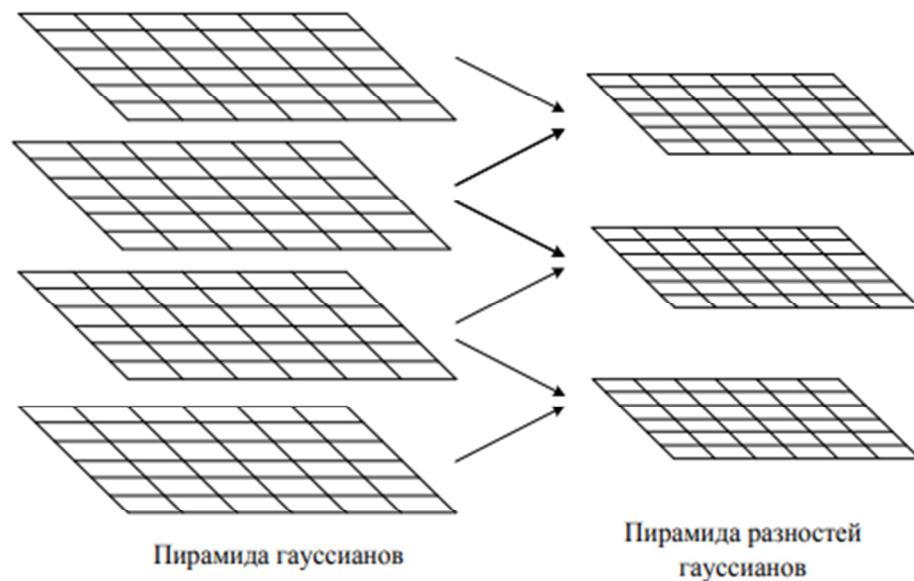


Рисунок 1.6 – Процесс поиска точек интереса

В роли кандидатов точек интереса больше всего подходят сильно отличающиеся пиксели, от других. Это выполняется способом сопоставления каждого имеющегося пикселя на изображении с некоторым количеством других соседних пикселей. Данная операция изображена на рисунке 1.7.

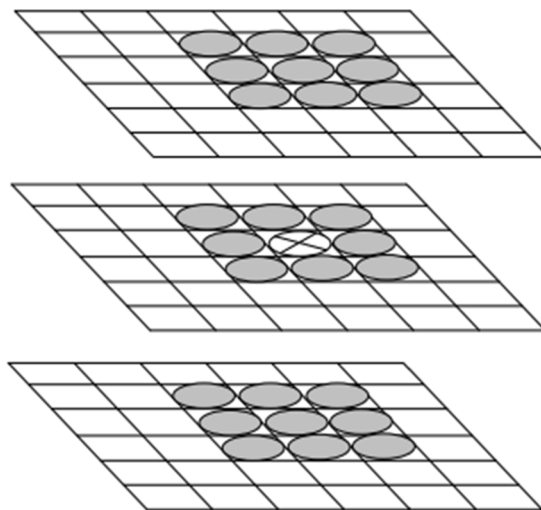


Рисунок 1.7 – Процесс поиска локальных экстремумов

Изображение размером 500x500 пикселей, обычно дает увеличение в 2000 постоянных функций [28]. Однако, такое количество точек обычно не требуется для обнаружения объекта и его оценки расположения. К примеру, требуется только 5 соответствий точек к восстановлению объекта, учитывая параметры внутренней камеры. Согласно тестированию SIFT, среднее число времени отсеивания характерный точек 640x480 пикселей изображения, составляет 1.75 секунды, в то время как для изображения 320x240 пикселей – 0.63.

Преимущество SIFT – полная инвариантность, по отношению к четырем параметрам аффинных преобразований, в них входят: размер изображения, его вращение и передвижение.

Недостатки SIFT – при использовании данного метода, не все приобретенные точки и дескрипторы этих точек соответствуют нужным потребностям, что сказывается последующим выполнением поставленной задачи сравнения изображений.

SIFT не будет выполнять поставленную задачу при следующих обстоятельствах:

- 1) На изображении присутствует освещение в зависимости от времени суток (день или ночь);

- 2) Распознаваемый объект содержит отбрасывающую в обратном направлении поверхность;
- 3) Угол обзора распознаваемого объекта отличается.

1.7 Метод SURF

Метод SURF (Speeded Up Robust Features) включает в себя выполнение двух задач: исследование особых точек изображения; создание их дескрипторов, инвариантных к размеру и циркуляции. Следовательно, представление ключевой точки будет равным, несмотря на то что образец может поменять свою форму и станет развернутым. Помимо этого, исследование ключевых точек должно владеть инвариантностью, для того чтобы перевернутый объект на изображении содержал в себе идентичный набор ключевых точек.

Операция объединения особых точек на изображении производится на базе матрицы Гессе (FAST – Hessian detector). Ее применение обеспечивает инвариантность в отношении к изменению размера. В связи с чем, SURF использует фильтры разного размера для расчетов Гассиана. К примеру, если исходное изображение задается матрицей M , то выбранный пиксель, взятый под рассмотрение, будет обозначаться $X = (x, y)$, а σ – размер фильтра. Исходя из этого, матрица будет выглядеть следующим образом:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix}, \quad (1.7)$$

$$L_{xx}(X, \sigma) = I(x, y) * \frac{\partial^2}{\partial x^2} g(\sigma), \quad (1.8)$$

где $L_{xx}(X, \sigma), L_{xy}(X, \sigma), L_{yy}(X, \sigma)$ – свертки аппроксимации второй производной Гауссова ядра с изображением при $X = I^*(x, y)$ во входном изображении, имеющем масштаб σ . Кроме того, $L_{xy}(X, \sigma)$ и $L_{yy}(X, \sigma)$ представлена свертка второй производной Гауссова фильтра и входного изображения для x у направление (диагональное) и направление y (вертикальное).

Метод использует адаптированное приближение фильтра окна, свертка второй производной Гауссовых, чтобы решить проблему увеличения времени обработки. [29]

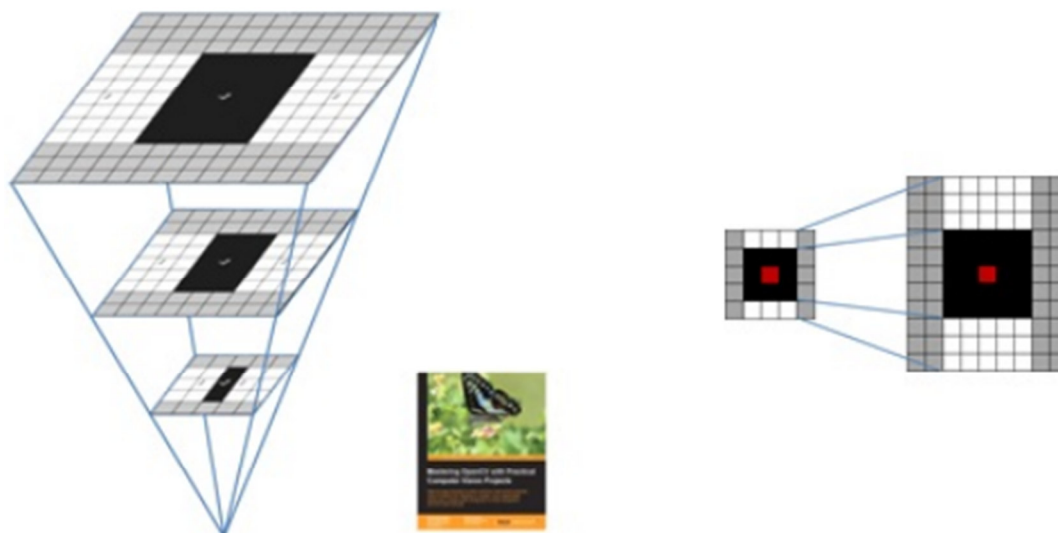


Рисунок 1.8 - Пирамида изображения и отфильтрованное окно

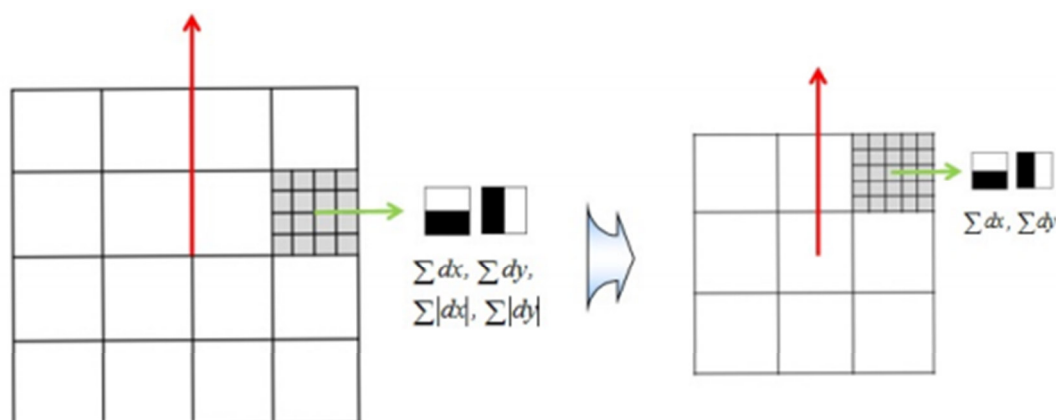


Рисунок 1.9 – сокращение размера в дескрипторе функции

Рисунок описывает предложенный метод сложности уменьшения, используя сокращение размера дескриптора функции. Обычные алгоритмы, использующие дескрипторы с размером 64 не подходят для условия реального времени, так как их вычисление сложно для высокого извлечения характерных точек. Поэтому сокращение размера в дескрипторе функции необходимо для эффективного уменьшения сложности вычисления, чтобы выполнить распознавание объектов в режиме реального времени [30].

Уменьшение размера в дескрипторе функции используется для расчета вектора направления через масштаб s , чтобы определить доминирующую ориентацию и расширение ее окна $\pi/2$ для оценки направления точной ориентации из множества информации.

Прямоугольное окно разделено 3×3 подобласти, а затем они повторно делятся на области 5×5 подобласти. В уравнении 1.9, $18 \cdot (3 \times 3 \times 2)$ – дескриптор функции размерности в сегментированной области, составляет два характеристических вектора.

$$V_{sub} = [\Sigma dx, \Sigma dy], \quad (1.9)$$

где Σ – сумма небольшого вейвлет-преобразования Хаара вычисления направления по горизонтали (dx) и по вертикали (dy). Так как в результате Хаара – надежное условие, а этот метод предлагает уменьшение сложности вычисления.

Определитель матрицы Гессе достигает экстремума в точках, где происходит максимальное изменение яркости. В итоге, SURF проходит фильтром с Гауссовым ядром по всему изображению, где обнаруживает эти самые точки, в которых добивается максимального значения определителя матриц Гессе. Следует заметить, что данный метод определяет, как темные пятна на белом фоне, так и наоборот.

Следующим этапом в отдельности для найденных точек рассчитывается преобладающее направление перепада яркости на изображении. Данное понятие схоже с направлением градиента, но для вычисления направления особой точки применяется фильтр Хаара.

После того, как были найдены ключевые точки, SURF формирует их дискриминанты:

1) Вокруг области точки выстраивается квадратное пространство, имеющее размер $20s$, где s – это размер, на которое получено максимальное значение определителя матрицы Гессе;

2) Принятая квадратная область делится на соответствующие блоки, где она будет делиться на 4×4 региона;

3) Для каждого такого блока будут рассчитываться несложные признаки. И в результате, на выходе получается вектор, который имеет 4 компоненты: 2 – это суммарный градиент по квадрату; 2 – сумма модулей точечных градиентов;

4) Формирование дескриптора происходит в результате объединения взвешенных описаний градиента для 16-ти квадратов вокруг особой точки. Части этого дескриптора взвешиваются на коэффициенты Гауссова ядра, а взвешивание необходимо для лучшей стабильности по отношению к шумам на изображении в удаленных точках;

5) В дополнение к дескриптору прибавляется след матрицы Гессе, это требуется для отличия темных и светлых пятен на изображении, т.е. у светлых точек след на темном фоне отрицательный, а у темных наоборот – положительный.

Недостаток метода заключается в том, что SURF применяется для нахождения объектов на изображении, где он самостоятельно работать с объектами не может. Также он никаким образом не выделяет объект общего фона, а считает изображение целым и производит поиск его характеристик. При всем этом, характеристики могут быть внутри объекта, и вне его, также в точках границы фона и объекта. Это значит, что данный метод не очень

хорошо будет производить работу для объектов простой формы и без четкой текстуры. Исходя из этого, объект не будет распознан в другом изображении на другом фоне, а значит, он слабо чувствителен по отношению к шуму.

Неизменность и положительный вывод при преобразовании освещенности изображения, или углом, под которым был выполнен снимок для обработки. Но тем не менее, SURF в несколько раз быстрее и лучше работает по сравнению с SIFT. Также описанный метод отлично выполняет задачу изменения яркости на изображении.

1.8 RANSAC метод

Основы RANSAC

RANSAC (RANdom SAmples Consensus) – неизменный метод оценки параметров модели на базе случайных выборок. Когда возникает задача обработки данных, где нужно вычислить параметры модели, удовлетворяющей первоначальной информации. Такая информация делится на две части, первая из которой – верные точки, удовлетворяющие модели, также именуемые как «инлаеры» (inlier) вторая часть – ложные точки, куда также входит шум, они называются «выбросами/аутлаерами» (outlier).

В дополнение к этому, данные могут подвергнуться шуму. Выбросы могут происходить, например, от экстремумов шума или от ошибочного измерения, неправильных гипотез об интерпретации данных. RANSAC также предполагает, что, данный (обычно маленький) набор инлаеров, где существует процедура, которая может оценить параметры модели, оптимально объясняет или соответствует этим данным.

Поскольку есть неправильные соответствия из-за неоднозначных признаков, у отдельных соответствующих признаков есть более низкая вероятность правильности, чем у группы признаков. Это считается, по

крайней мере, тем что, тремя групповыми признаками возможно достигнуть надежного распознавания.

RANSAC может быть использован при проверке соответствия группы точек к геометрическому моделированию. От соответствующих точек, полученных SIFT методом, выбираются три пары точек в произвольном порядке, чтобы создать матрицу преобразования, которая соответствует к 2D плоскости. Тогда порог настраивается расстоянием правильной позиции точки от предыдущей позиции точки, вычисленной матрицей преобразования. RANSAC достигает своей цели, многократно выбирая случайное подмножество исходных данных, которые означают гипотетические инлаеры. Эта гипотеза выглядит следующим образом:

- 1) Модель адаптирована к гипотетическим инлаерам, т.е. для всех свободных параметров модель восстановлена от набора данных;

- 2) Тогда все другие данные протестированы против подходящей модели. Если точка соответствует к хорошо предполагаемой модели, это также рассматривается как гипотетический инлаер;

- 3) Предполагаемая модель довольно хороша, если достаточно большое количество точек было классифицировано как гипотетический инлаер;

- 4) Модель повторно оценивается от всего гипотетического инлаера, поскольку она оценивается только от начального набора гипотетических инлаеров;

- 5) Наконец, модель оценивается, оценив ошибку инлаера относительно модели.

Эта процедура повторяется от постоянного числа времени, каждый раз производя любую модель, которая отклонена, потому что слишком мало точек классифицировано инлаерами, либо модель усовершенствована вместе с соответствующим ошибочным показателем. В последнем случае сохраняется усовершенствованная модель, если ее ошибка ниже, чем

последняя сохраненная модель. На рисунке 1.10 приведена блок-схема работы процедуры.

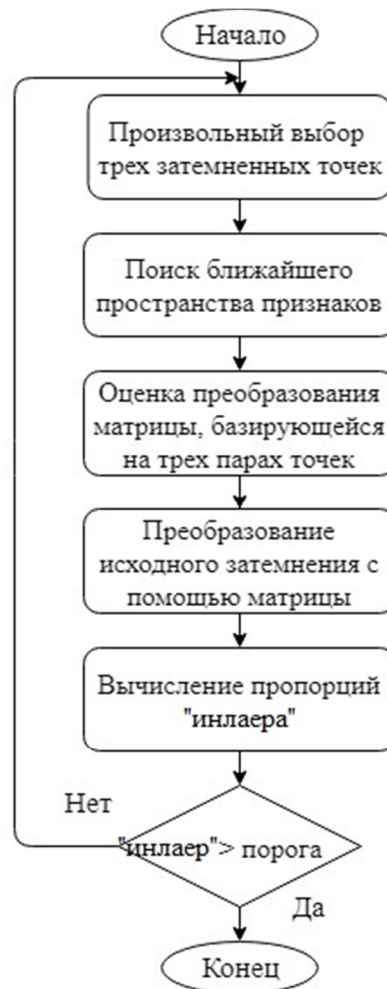


Рисунок 1.10 – Блок-схема основного алгоритма RANSAC для затемненной регистрационной точки

К примеру, работы алгоритма состоит в занесении прямой в 2D точки. Если предположить, что есть выбросы среди имеющихся данных, то оценка размеров обычным способом, достигнет того результата, когда определится ложная модель, в связи с тем, что она базируется на основе всех точек. Описанный метод принимает за основу две ему требуемые точки для того, чтобы построить прямую, а также с помощью этих точек и саму модель. После этого, он производит проверку соответствия количества точек к модели, при этом применяя функцию оценки с предложенным ему порогом.



Рисунок 1.8 – Группа данных, куда нужно занести прямую. Выбросы пребывают в значительном количестве

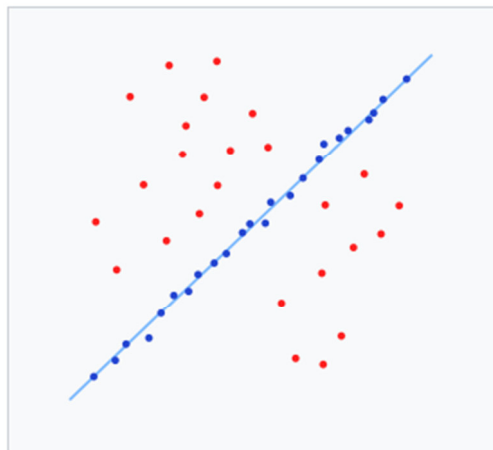


Рисунок 1.9 – Предложенная алгоритмом RANSAC прямая. Выбросы не воздействуют на ожидаемый результат

Аффинное преобразование сохраняет коллинеарность (т.е., все точки, лежащие на строке первоначально все, еще находятся на строке после преобразования) и отношения расстояний (например, средняя точка линейного сегмента остается средней точкой после преобразования). Это значит, что он аффинно указывает на специализированный класс проективных преобразований, которые не перемещают никакие объекты от аффинного пространства до плоскости в бесконечности или, с другой стороны. Геометрическое сокращение, расширение, отражение,

вращение, сдвиг, преобразования подобия, спиральные общие черты, и перевод – это все аффинные преобразования, как их комбинации.

В целом аффинное преобразование – это преобразование вращений, расширения и сдвига. В то время как аффинное преобразование сохраняет пропорции на строках, при этом не обязательно сохраняет ли он углы или длину. Любой треугольник может быть преобразован в любой другой аффинным преобразованием, таким образом, все треугольники аффинные и, в этом смысле, аффинно обобщение подходящих и подобных. Основные аффинные преобразования показаны на рисунке 1.10, а их соотношения преобразования, показаны в таблице 1.1.

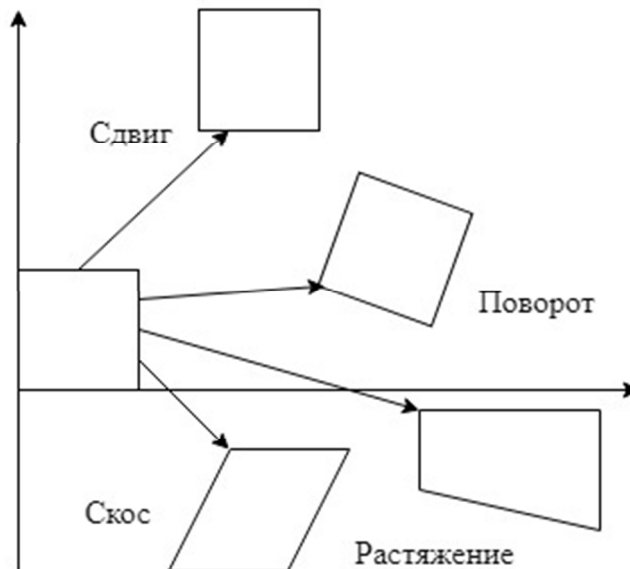


Рисунок 1.10 – Основные аффинные преобразования

Таблица 1.1 – Аффинные преобразования

	Рисунок	Матрица
Сдвиг		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
Увеличение		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
Поворот		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
Движение		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & dx \\ \sin \theta & \cos \theta & dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
Подобие		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos \theta & -s \sin \theta & dx \\ s \sin \theta & s \cos \theta & dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
Аффинное преобразование		$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos \theta & -\sin \theta & dx \\ s \sin \theta & \cos \theta & dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

Преимущество RANSAC – возможность сделать устойчивую оценку параметрам модели, т.е. это может оценить параметры с высокой степенью точности даже когда существенное количество выбросов присутствует в наборе данных.

Недостаток RANSAC – то, что нет никакой верхней границы на времени, которое требуется, чтобы вычислить их параметры. Когда верхняя граница используется с указанием срока (максимальное количество

итераций), то полученное решение оптимальным не будет, также есть вероятность того, что ни одна модель не сойдется с исходными данными. Фактическая модель может быть произведена RANSAC только с определенной вероятностью, которая будет больше используемых итераций. Другой недостаток RANSAC – то, что он требует установки конкретных пороговых значений. RANSAC может только оценить одну модель для определенного набора данных. Также можно сказать, что для любого подхода для одной модели, есть некая проблема, заключающаяся в ситуации, когда присутствуют две модели (или больше), тогда RANSAC не может не найти никакую из них.

ГЛАВА 2. СИСТЕМА ОБРАБОТКИ ИЗОБРАЖЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

К распознаванию объектов относятся задачи построения и использования работ над числовыми или же буквенными демонстрациями объектов окружающего мира, их следствие отображает взаимосвязь эквивалентности между субъектами. Взаимосвязи равенства изъясняют некое приспособление рассматриваемых объектов к классам разного вида, они рецензируются как самостоятельные семантические единицы. Можно сказать, о «распознавании с преподавателем», т.е. когда компьютерная структура решает поставленную задачу систематизации не завися от внешней информации, предлагающей знания, тогда это уже называется автоматической систематизацией («распознавание без преподавателя»). Большому количеству алгоритмов требуется добавление существенных вычислительных объемов, которые могут содержать в себе высокую результативность ЭВМ. Главная существующая топология методов для распознавания образов:

- методы, базирующиеся на концепции деления;
- статистические методы;
- методы, которые выстраиваются на базисе «потенциальных функций»;
- методы, которые рассчитывают оценки, т.е. голосование.

Внутри базиса выше описанной систематизации есть разница, которая находится в формальных методах распознавания объектов, так что обзор эвристического подхода к распознаванию, имеющего полноценную и разумную разработку в специализированных системах, понижено. Такой подход базируется на сложно оформляемых знаниях, научный работник самостоятельно выявляет те данные и способы, которые структура должна применить для того чтобы добиться нужного эффекта распознавания.

Аналогичная топология методов распознавания с уровнем подробности может встречаться во многих исследовательских работах, но не входит во внимание характеристика, которая отображает конкретику пути воззрения познания о предметной площади с применением алгоритма распознавания объектов.

В особенности следует акцентировать внимание на то, что бытие методов распознавания углубленно правомерно, потому что они взаимодействуют и с внешними атрибутами, и с предметами. Если эти методы использовать отдельно друг от друга, то результат не будет образовывать достойное осмысление тематики. Это означает, что системы распознавания в будущем, должны гарантировать осуществление этих методов, а не одного из них.

В результате всего этого, в базис систематизации методов распознавания, заложены прочные основы правомерности, располагающиеся в манере человеческого познания. Это выдвигает ее в льготную позицию по отношению к другим систематизациям, они на контексте всего этого кажутся более облегченными и ненатуральными.

Автоматизированное распознавание объектов является актуальной задачей, имеющей отношение с широким классом практических приложений. В наши дни существует большое количество устройств для распознавания объектов, эти аппараты отдалены от безупречности и каждый день они видоизменяются. Но все же, суммарное построение и заключение предоставленной задачи сейчас сформулированы.

Видеокамера – это предмет, который совершает съемку. Для размещения видеокамеры существуют некие условия. У нее оптическое разрешение ориентируется так, что объект помещается в кадре по горизонтали. Распознавание камерой, базируется от обзора угла наклона камеры, этот угол обязан быть на 40^0 . Размер объекта по вертикали сокращается пропорционально величине косинуса угла наклона камеры, в следствие, требуется увеличение разрешения камеры, а это означается то, что

произойдет уменьшение ширины прицела. Наиболее удобным размещением камеры заключается горизонтальное – это используется в системах, когда доступ ограничен. [22]

Следующий, не менее важный параметр – минимально допустимая контрастность изображения государственного регистрационного номера автомобиля. Устанавливают дополнительный модуль, который состоит из инфракрасного освещения и соответствующего фильтра [23]. Это значит, что повышается контрастность номерной пластины по отношению к изображению. Основывается это на том факторе, что номерной знак имеет специальную светоотражающую поверхность, при котором отражающийся свет будет распространяться в противоположном направлении распространению падающего света. В итоге, камера будет принимать в общем случае только инфракрасный свет, а также отраженный свет от гос. номера автомобиля. А это значит, что снимок получится монохромным, без деталей, за исключением государственного регистрационного номера автомобиля.

Предварительная обработка в результате получаемого изображения, имеет несколько этапов:

- 1) Корректирование изображения – эквализация, т.е. выравнивание условий передачи всех частот в пределах полосы пропускания. Ограничение экстремальных значений яркости, модификация гистограммы распределения яркостей.

- 2) Устранение эффекта размытия изображения.

- 3) Ограничение избыточных данных – задействование инфракрасной (ИК) подсветки, бинаризация (двухуровневое разбиение изображения на отдельные цветовые регионы).

- 4) Употребление программного обнаружения движения для определения месторасположения объекта на изображении.

Основной недостаток двухуровневого разбиения фотоснимка на отдельные цветовые регионы — это то что выбираемый порог бинаризации

не дает требуемого качества для всяческого типа изображения. Такие факторы, как подсветка и цвет воздействуют на качество фотоснимка. Применение методов приспособляющегося двухуровневого разбиения изображения на отдельные цветовые регионы позволяет устранить проблему эффективно.

Из методов для локализации объектов есть метод выявления границ фотоснимка с применением алгоритмов Кенни, Собеля, Робинсона и т.д. Алгоритм выявления границ должен выявлять горизонтальные и вертикальные края, в итоге, изображение после этапа выделения краев должно будет иметь определенное количество линий в площади распознаваемого объекта, потому что содержит определенные границы.

Для локализации объекта выстраивается окно, которое примерно будет равным к его размеру на изображении. Затем, это окно применяется для того, чтобы сделать оценку количеству окантовок во всех областях фотоснимка, у которых есть наибольшая контрастность. Окно накладывается на результирующий фотоснимок в большей степени контрастных зонах охвата. Если же, может получиться так, что количество граней будет располагаться в заданном диапазоне, то это будет означать что данная область будет маркироваться как область, а количество граней находится экспериментально. Итогом работы такого окна будет являться список возможных областей, т.е. кандидатов, в которых присутствует объект. Главный недостаток этого – медленная обработка выбора кандидата, так как значения пикселей в отведенном окне будут многократно складываться. Также есть еще один недостаток такого подхода – это набор огромного количества кандидатов в объекте.

2.1 Предварительная обработка перед распознаванием

Известно, чтобы получить некое изображение, используется камера, которая должна работать в постоянном режиме, или же включаться, когда возникает в настроенной области распознаваемого объекта [24, 25]. Сейчас часто используются камеры, которые срабатывают по особому сигналу, исходящему от триггерного устройства, оно контролируется датчиком, имеющим индуктивную петлю. И здесь имеются свои недостатки [24, 26]: не всякий раз получается конкретно постановить локацию объекта, увеличивается цена этой системы, а также, на принимаемых изображениях объект может быть виден не полностью. Структура с одной камерой, непрерывно работающей, выгоднее и плотнее, но по своей схеме она затруднительна, потому что из всего течения поступающих изображений, она должна выделять только индивидуальные объекты [26].

В пределах того, какой способ применяется для того чтобы получить изображение, главным атрибутом является разновидность употребляемой видеокамеры [24]. Ограниченность динамического спектра обширно используемых камер считается весомым вопросом в плане разработки робастных систем, т.е. совокупности способов теории управления, целью которых считается обеспечение хорошего качества управления [27], механического распознавания, например, просматриваемые кадры как правило располагают значительно более важными пределами конфигурации яркостей, чем стабилизирует камера [25, 26]. Имеется пара методов чтобы решить такую проблему, а это: применение видеокамеры с усовершенствованными параметрами, либо развитые квалификационными алгоритмами тестирования изображений: существенный прирост стоимости системы, а это нежелательно и прием изображений с динамическим диапазоном, который увеличен в отличии от камеры [26]. Для того чтобы не было размытия изображения, которое происходит из-за того, что объект быстро движется, требуется использование специальных методов обработки.

Когда изображение принято, идет его пересылка узлу обработки и анализа фотоснимка, он содержит пару этапов, а это детектирование [24, 27].

2.2 Определение ключевых особенностей изображения

Для решения подзадачи выделения границ объекта на цифровом изображении в рамках данной работы применяется метод, основанный на нахождении ключевых особенностей, которыми он обладает, а именно её углов [34]. Для этого применяется уголкового детектор Харриса [35]. Является текущая точка изображения угловой или нет, алгоритм определяет на основе статистики значений градиента интенсивности изображения в рамках окрестности данного пикселя [7, 35]. Результатом работы алгоритма является карта угловых особенностей изображения.

Далее карта уголков Харриса подвергается процедуре пороговой обработки (бинаризации), после чего происходит извлечение связных областей, обнаруженных на предыдущей стадии [4, 7, 36]. Одна из данных областей и будет соответствовать номерной пластине.

Для найденной области необходимо решить задачу бинарной классификации данных областей [7, 8]. Эта задача состоит из описания обнаруженных областей с помощью определённого набора признаков и непосредственной классификации с использованием алгоритма машинного обучения [24, 37, 38]. Данный алгоритм на основе некоторого решающего правила определяет, является найденная на изображении область номерной пластиной или нет [23]. На выходе алгоритм выдает координаты местоположения номерной пластины в кадре.

Детектор углов, или в более общей терминологии детектор точечных особенностей, является подходом, который применяется в системах компьютерного зрения для нахождения определённых особенностей на цифровом изображении [26, 27]. В литературе описано достаточно много различных детекторов точечных особенностей: детекторы Моравеца, Ши

Томаси, детектор Харриса, и др. Одним из главных недостатков многих алгоритмов является их высокая вычислительная сложность при не самой высокой точности [34, 35]. Среди них наиболее точные результаты показывает алгоритм Харриса, который и применяется в данной работе.

2.3 Детектирование объектов

В первую очередь, что само по себе означает слово детектирование – это обнаружение, преобразование электрических колебаний. Но в данной работе рассматривается распознавание объектов, поэтому цель детектирования – обозначить присутствие объекта на изображении, отыскать расположение объекта в системе отсчета пикселей начального снимка. Решение такой задачи дает возможность разобрать добротную структуру сюжета, которая выявлена на снимке и принять данные об обоюдном месте размещения объектов.

Существуют методы решения задачи детектирования объектов. Они делятся на три группы:

1) Методы, которые нужны для того чтобы охарактеризовать объект, в нем применяются признаки, в большей степени подходящие для объектов. В виде этих признаков подбираются аккуратные, или даже можно сказать скрупулёзные, черты объекта, также и признаки, которые сооружены для снимка, имеющего единственный объект.

2) Методы, для обнаружения объектов, сопоставимые шаблону, т.е. характеристике объектов.

3) Методы детектирования движения объектов, т.е. на базе некоего количества изображений или видеок кадров одинаковой сцены происходит выделение объектов.

Итак, рассмотрим методы для обнаружения объектов поиска по шаблону. Детектирование объектов на базисе какого-то шаблона

свидетельствует о том, что есть снимок объекта с ассигнованными внешними атрибутами – шаблон, испытываемый снимок, сравниваемый с таким шаблоном, см. рисунок 2.3.

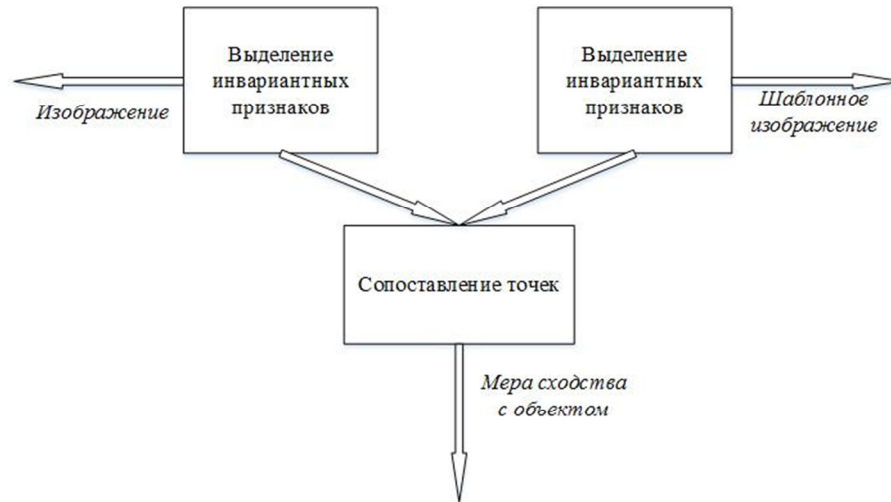


Рисунок 2.3 - Схема решения задачи детектирования объектов с использованием методов поиска объектов

Следствие этого сравнения – мера сходства, учитывается то, что, если такая мера, увеличенная от какого-то рода порога, тогда испытываемый снимок будет снимком объекта. При элементарном событии в виде шаблона служит снимок объекта – массив выраженности цветности, в большей степени типичных к объекту. Гораздо более трудные методы рецензированной группы в виде шаблона применяют комплекты векторов внешних атрибутов, т.е. дескрипторов, абстрактное воззрение объекта либо вероятностные образцы объектов, вмещающие в себя сведения об раздате насыщенности пикселей [24].

Во время поиска реализуется прохождение, так сказать, «бегущее окно», располагающим величиной шаблона, по снимку и сопоставлением характеристики детали первоначального снимка, укрываемого окном, и шаблона. Сравнение с шаблоном означает сопоставление характеристики испытанного и шаблонного снимка на некой отобранной метрике [25], из регламентации следует, что отбирается Евклидово расстояние, стандарт L_1 ,

сбалансированная свертка квадратических просчетов или же корреляция, т.е. соотношение.

К примеру, есть шаблонная характеристика объекта $I_0(X)$ в дискретной (прерывистой) среде пикселей $\{X_i = (x_i, y_i)\}$. Затем цель розыска объекта относится к цели сворачивания совокупного просчета. А если в виде меры схожести применено Евклидово расстояние, тогда цель будет внесена нижеуказанным способом:

$$E(u) = \sum (I(X_i + u) - I_0(X_i))^2 - \sum e_1^2 \rightarrow \min \quad (2.9)$$

где u – сдвиг шаблонной характеристики в комплексе осей первоначального снимка. В результате, без учета отобранной метрики происходит переход до функции упорядочения.

К алгоритмам, применяющих сравнение дескрипторов основных пунктов, в большей степени значимым служит подбор порога, который применяется в виде эталона соблюдения, другими словами, когда дистанция меж дескрипторами уменьшенная от порога, тогда пункты будут учитываться как релевантными. Приумножение порога влечет за собой к приумножению количества обнаруженных сравнений, в другом случае это влечет за собой приумножение количества неверных откликов. Минимизация порога наравне к произрастанию количества верных распознанных несоответствий проведет к увеличению количества выброшенных верных согласований.

Следует обратить внимание, что вышеописанные методы детектирования по указанному шаблону результативно срабатывают во время розыска одинарного объекта. В случае появления настила в «бегущем окне» пропадают некие атрибуты в характеристике. Исходя из вышеизложенного во время сравнивания окна шаблону вносится порог, по нему отрезаются бесперспективные окна, умышленно не имеющие объектов.

Рассмотрим распознавание на примере номерного знака. В первую очередь, чтобы система автоматического распознавания объектов осуществляла работу, производится детектирование знака в кадре, пример показан на рисунке 2.3. Это очень важно, поскольку, насколько аккуратно выделен субъект распознавания на изображении, то, соответственно, все результаты зависят от всех последующих стадий [14, 16]. Для выделения объектов есть целый ряд методов, которые могут быть сгруппированы в зависимости от используемых характерных признаков изображения. Основными признаками изображения являются текстуры, цвет и симметрия, краевые точки [14].

Существуют факторы, которые оказывают негативное влияние на корректное детектирование распознаваемого объекта, а это: низкое качество изображения, меняющиеся внешние условия освещения, перспективные искажения, отражение света, близкое сходство объектов между собой.

При выделении объекта для распознавания на изображении, используется информация о цвете. В разных странах номерные знаки обладают определенными цветами фона, установленные государством, и расположенных на нем букв. В качестве характерных признаков номера используется последовательность цветовых кодов, которая получается при поперечном сканировании изображения слева направо. После выделения номерных знаков применяются методы классификации, а это: искусственные нейронные сети и генетические алгоритмы, которые позволяют локализовать область номерного знака. Следует отметить, что использование цветовой информации дает ряд преимуществ, и в то же время сохраняются проблемы, связанные с использованием освещенности сцены, близостью цветов номера, а также автомобиля.

2.4 Воздействие шума на точность детектирования

Любая реальная радиотехническая система работает в условиях воздействия шумов. Их влияние на работу алгоритмов обработки изображений также сложно переоценить. Поэтому рассмотрим зависимость точности детектирования от величины, которая выступает в качестве оценки шумового воздействия на входное изображение.

Суть эксперимента сводится к следующему. Все изображения тестовой базы данных подвергаются зашумлению аддитивным белым гауссовским шумом (АБГШ). Данная модель шума характеризуется постоянной во всём частотном диапазоне спектральной плотностью мощности шума. После процедуры наложения шума на изображения происходит тестирование рассматриваемого алгоритма на тестовой базе из изображений, содержащих шум.

В качестве оценки вклада шума в изображение используется величина пикового отношения сигнал/шум (ПОСШ). Для исходного тестового изображения X и зашумлённого изображения Y вычисляются величины по следующим формулам:

$$\text{СКО} = \frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2, \quad (2.10)$$

$$\text{ПОСШ} = 20 \log_{10} \frac{255}{\sqrt{\text{СКО}}}. \quad (2.11)$$

По итогам тестирования алгоритма построен график зависимости точности детектирования от величины ПОСШ (рисунок 2.4). Величина ПОСШ измеряется в децибелах (дБ). Данный график показывает, что уменьшение шума на цифровом изображении приводит к росту точности детектирования автомобильных номеров.

При снижении величины ПОСШ возрастает влияние шума на работу алгоритма: появляются ложные области на карте откликов алгоритма Харриса, шум проникает в НОГ-дескрипторы. Это приводит к тому, что появляются ложные срабатывания на выходе алгоритма детектирования аномалий. При низких значениях ПОСШ на изображении теряются границы номерного знака, что приводит к тому, что данная область теряется из виду.

Следовательно, теряются точечные особенности (углы), которые должен обнаруживать алгоритм Харриса.

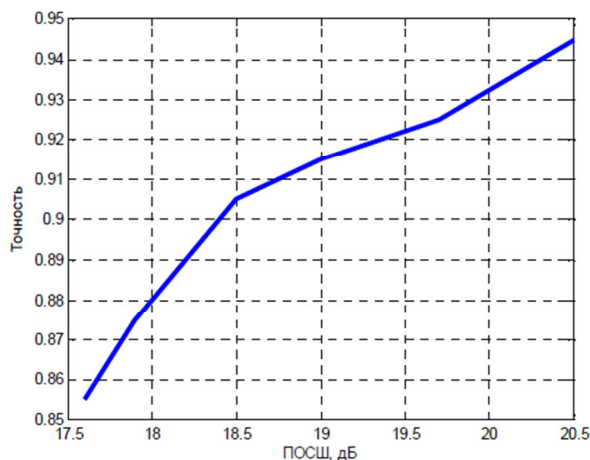


Рисунок 2.4 – Зависимость точности детектирования от величины ПОСШ при действии АБГШ

2.5 Распознавание объектов с применением обнаружения области и выделением признаков

Алгоритмы дескриптора детектора SURF и (U) SURF [31] (Speeded-Up Robust Features) разработаны с целью ускорения локализации ключевых точек. Вместо пространства масштаба проанализирован масштабирующийся размер фильтра, многократно уменьшая размер изображения, как это происходит в SIFT методе.

Вместо того чтобы вычислить инвариантность функции целого изображения, цель представленного метода – уменьшить размер изображения, которое будет обработано для извлечения части изображения, что с высокой вероятностью будет содержать большинство фундаментальных свойств. Далее представлено два шага этого подхода.

1. Извлечение области интереса

Для извлечения области интереса используется несколько методов. Априорное знание объектов, которые будут идентифицированы, может

использоваться, например, форма или информация о цвете [70]. Вместо этого, подход может быть обобщенным, сканируя изображение для непрерывного соединения области или блоба. Блоб (Binary Large Object) – массив двоичных данных, является областью затрагивающих пикселей с той же логической структурой. Удаление блоба, также известный метод как обнаружение области или маркировки сегментации изображения, пиксели которого отображаются как принадлежащие одной из многих дискретных областей. Процесс состоит из сканирования и нумерации любых новых областей, с этим также встречается и слияние старых областей, когда они оказываются объединенными на более низкой строке. Поэтому отсканированное изображение и каждый пиксель индивидуально маркированы идентификатором, который показывает область, к которой он принадлежит.

Распознавание блоба обычно выполняется на получающемся двухуровневом изображении от шага пороговой обработки. Вместо этого применяется SUSAN (Smallest Univalued Segment Assimilating Nucleus) – при использовании данного алгоритма для каждого пикселя изображения рассматривается круговой регион фиксированного радиуса, центр региона называется ядром, его значение интенсивности запоминается и сравнивается со значениями интенсивности других точек региона [69], он считается более быстрым и стабильным оператором.

Процесс экстракции блоба может дать много различных областей для единственного изображения. Для блоба, который будет подтвержден как кандидат, результат процесса его обнаружения должен фильтроваться и быть отброшены ложные положительные стороны. Могут использоваться различные методы фильтрации, например, [70] информация о цвете в области кандидата используется для идентификации дескриптора. Альтернатива этому – [71] связанная с расположением информация о пикселе, используемая для дискриминации блоба, где цель состоит в том, чтобы считать объекты в изображениях взятыми системой наблюдения.

Аналогичный подход используется в предложении где блобы, которые не являются непротиворечивыми с определенными ограничениями размера. Ограничения размера зависят от близости изображения.

В зависимости от задачи, которая будет решена, изображения не должны быть ограничены единственным объектом, который будет идентифицирован и, следовательно, мог содержать больше чем одну область, обнаруживаемую и извлекаемую. Как только самые интересующие блобы расположены, используются длина блоба и значения ширины, чтобы впоследствии найти их центр, а квадратное изображение извлечено для каждого. Размер квадрата определен максимальным значением между длиной и шириной блоба кандидата, тогда изображение масштабируется к фиксированному размеру, чтобы получить его часть — это значит, что он с высокой вероятностью должен содержать предмет интереса, т.е. ROI. Затем ключевые точки извлекаются и выполняется процедура соответствия к каждому извлеченному ROI.

2. Выделение признаков

SIFT функции извлекаются согласно следующей процедуре:

1) Обнаруживается экстремальное значение масштабного пространства, ищущее по всем масштабам и расположения изображения. Потенциальный интерес указывает на инвариантность, чтобы масштаб и ориентация эффективно вычислялись, используя DoG (дифференциал Гауссовых) функцию.

2) Локализируются ключевые точки, обнаружив локальное экстремальное значение и удаляются точки с низкой контрастностью.

3) Присваивается ориентация кандидатам ключевых точек, базирующихся на локальных направлениях градиента изображения.

После того, как ключевые точки локализованы для каждого из них вычисляется дескриптор, рассчитывая гистограмму градиентов ориентированной локальной переменной вокруг точки интереса и хранения

контейнеров в размере векторов 128. Тогда эти дескрипторы могут быть по сравнению с сохраненными в целях распознавания объектов.

Для онлайн-приложений, каждого из трех шагов (обнаружение локального экстремального значения, вычисление ключевых точек, их описания и соответствие), должен быть быстро вычислен.

SURF [72] дескрипторы вычисляются в два шага. Восстанавливаемая ориентация находится на основе информации, а круговая область вокруг точки интереса. Это выполняется использованием вейвлет-преобразования Хаара в X , и был обнаружен в направлениях Y масштабируемых точек интереса. Тогда оценивается доминирующий признак ориентации, вычислив сумму во всех отвечающих скользящих окнах. Затем, область делится на меньшие квадратные подобласти и вычисляются некоторые простые функции. Это приводит к дескриптору длины 64, половины размера оригинала SIFT дескриптора и, следовательно, предложения в менее дорогом вычислительном отношении соответствия процесса.

Вертикальная версия SURF, названного USURF, пропускает первый шаг процесса вычисления дескриптора, приводящего к более быстрому варианту. USURF предлагается для тех случаев, в которых инвариантность вращения не обязательна. Аналогичный подход – примененный SURF алгоритм Ledwich и Williams [73], где предполагается, что точка обзора относительно стабильна к вращению вокруг представления оси. Во всех этих трех методах процесс соответствия ссылается на базу данных, где ключевые точки интересующего объекта хранятся вместе с соответствующими критериями.

ГЛАВА 3. РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

3.1 Алгоритм RANSAC

Алгоритм RANSAC часто используется в компьютерном зрении, например, если нужно одновременно решить проблему соответствия изображения и оценить основную матрицу, связанную с параметрами расположения камеры.

В целом алгоритм RANSAC работает следующим образом:

Вход алгоритма:

- 1) Набор исходных данных
- 2) Модель, которая может быть адаптирована к данным
- 3) n - минимальное количество данных, требуемых для соответствия модели
- 4) k - максимальное количество итераций для алгоритма
- 5) t - пороговое значение для определения, когда данная величина соответствует модели
- 6) d - количество близких значений данных, требуемое для убеждения, что модель хорошо соответствует к параметрам

Алгоритм содержит только один цикл, где каждую итерацию возможно логически разделить в два шага.

Первый шаг – выбор точек и модели.

- 1) Из количества исходных точек X произвольно выбирается n разных точек.
- 2) На базе этих точек производится расчет параметров модели, с применением функции M , данная модель будет именоваться как гипотеза.

Второй шаг – проверка полученной гипотезы.

- 1) Для имеющихся точек производится проверка ее соотношения к гипотезе, при этом применяя функцию оценки и порога;
- 2) Точка имеет отметку инлаера или выброса;
- 3) Когда проведена проверка всех имеющихся точек, тогда выполняется проверка для гипотезы, считается ли она хорошей на данный момент времени, если да, то она заменяет предыдущую гипотезу.

В конце цикла остается последняя хорошая гипотеза.

Возможные варианты алгоритма RANSAC:

1. Перерыв основного цикла, если достаточно хорошая гипотеза была найдена, т.е. одна с достаточно небольшой ошибкой. Это может сохранить некоторое время вычисления за счет дополнительных параметров.

2. Вычисление ошибки непосредственно из модели, повторно не оценивая, а устанавливая согласно модели, что может сэкономить некоторое время за счет сравнения – имели ли ошибки отношение к моделям, которые оценены от маленького числа точек и поэтому более чувствительный к шуму.

Значения параметров t и должны быть определены в зависимости от требований, связанных с набором данных, возможно на основе экспериментальной оценки. Параметр k (количество итераций), может быть определен еще до выполнения алгоритма способом теоретической оценки. Пусть p – это вероятность того, что алгоритм RANSAC в некоторых итерациях выбирает только инлайнеры из входного набора данных, когда выбираются точки n , из которых оцениваются параметры модели. Когда это происходит, получающаяся модель вероятна, чтобы быть точной, таким образом, p дает вероятность, что алгоритм производит точный результат. Пусть w – вероятность выбора инлайнеров каждый раз, когда единственная точка выбрана, т.е.

$$w = I/T, \quad (3.1)$$

где I – количество инлаеров, T – общее количество точек.

В большинстве случаев, часть инлаеров w будет неизвестна еще до того, как алгоритм начнет работу.

Предположим, что точки n необходимы для оценки модели, и выбраны случайно, w^n – вероятность того, что все точки n - инлайнеры, и $1 - w^n$ – вероятность того, что по крайней мере одна из точек n - изолированная часть. Случай, когда подразумевается, что это не лучшая модель, будет оценен от этого набора точек, то вероятность приводит к k вероятности, это значит, что алгоритм никогда не выбирает ряд n точки, которые все инлайнеры и это должно совпасть с $1 - p$. Следовательно,

$$1 - p = (1 - w^n)k \quad (3.2)$$

к которому, после принятия логарифма, приводит

$$k = \frac{\log(1-p)}{\log(1-w^n)}. \quad (3.3)$$

Нужно отметить, что этот результат предполагает, что n точки данных выбраны независимо, т.е. точка, которая была выбрана один раз, заменена и может быть выбрана снова в той же итерации. Часто это не лучший подход, и полученное значение для k должно быть взято в качестве верхнего предела в случае, когда точки выбраны без замены. Например, в случае нахождения строки, которая соответствует набору данных проиллюстрирован в вышеупомянутой формуле 3.3, алгоритм RANSAC обычно выбирает 2 точки в каждой итерации, и вычисляет модель как строку между точками, что приводит к выводу инлайнеров.

3.2 Алгоритм распознавания объектов с применением обнаружения области и выделением их признаков

Был разработан двухступенчатый алгоритм, который стремится убрать лишние ключевые точки, полученные методами SIFT и (U)SURF.

На рисунке 3.1 показан вывод описываемого алгоритма.

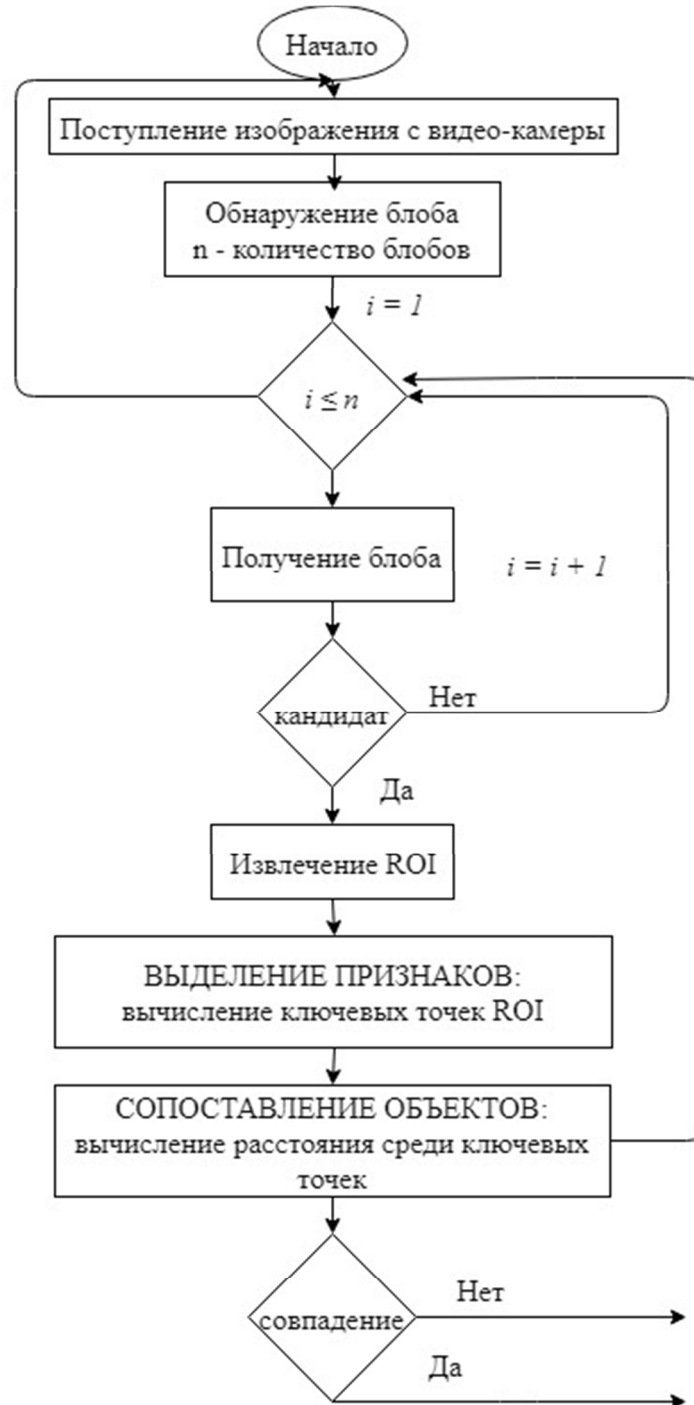


Рисунок 3.1 – Блок-схема алгоритма распознавания обнаружения области и выделением признаков

3.3 Распознавание методом SIFT

Алгоритм SIFT достаточно результативен, он имеет особенность большой сложности обработки изображения. Он делится на две части: нахождение точек интереса; формирование дескрипторов окрестностей данных точек. Для определения точек интереса существует несколько способов, и алгоритм SIFT один из известных и хорошо используемых.

1) Для обнаружения объекта в ситуации с захватом, обычно целевой объект занимает лишь небольшую часть изображения, следовательно, он нужен для того, чтобы найти целое изображение.

2) Привлечь внимание только на маленькое окно, которое содержит объект и свойства этого окна в площади отслеживания, применяя предыдущее описание, т.е. после распознавания объекта можно вычислить его расположение и размер, из этого будут извлечены только характерные точки.

3) Если объект успешно обнаружен, то расположение и размер этого окна будут обновлены и можно будет дальше продолжать процедуру распознавания.

4) Если же размер окна стал слишком маленьким, это будет означать, что есть некая потеря, в таком случае, при распознавании система может автоматически повторно отслеживать определение параметров процесса, обнаруживая, при этом, целое изображение, где располагается окно, содержащее нужный объект.

Блок-схема процесса показана на рисунке 3.2.



Рисунок 3.2 – Блок-схема процедуры отслеживания методом SIFT

ГЛАВА 4. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ РАСПОЗНАВАНИЯ

4.1 Распознавание методом SURF

Тестировалось видео с множеством последовательностей, задача вычисления акцентировалась на целевом моделировании и соответствии, обнаружении характерной точки соответствия, где находятся пропорции к отслеживанию размера окна. У SURF есть низкая вычислительная сложность для распознавания объектов в режиме реального времени. Поэтому, предложенный метод извлекает функции и находит соответствующие точки в видеопоследовательности. Также уменьшалась вычислительная сложность, чтобы эффективно снизить размер в дескрипторе функции для выполнения распознавания объектов.

Рисунки 3.2, 3.3, 3.4 описывают распознавание объектов и видимый результат, используя подход этого исследования, который улучшает SURF и алгоритм CamShift. Рисунок 3.5 показывает, что цель интереса точно обнаружена в сложной среде. Таким образом, низкая сложность и устойчивое распознавание объектов является усовершенствованным соответствием функции. Также это может эффективно выполнить отслеживание. Поскольку изменяется объект интереса, то его область может адаптироваться к изменению правильных размеров окна.

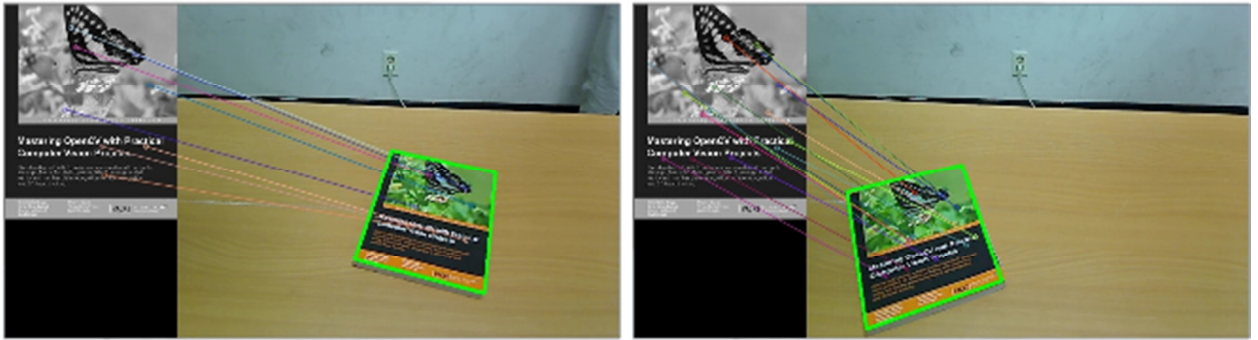


Рисунок 4.1 - Результаты распознавания и отслеживания объектов для книги

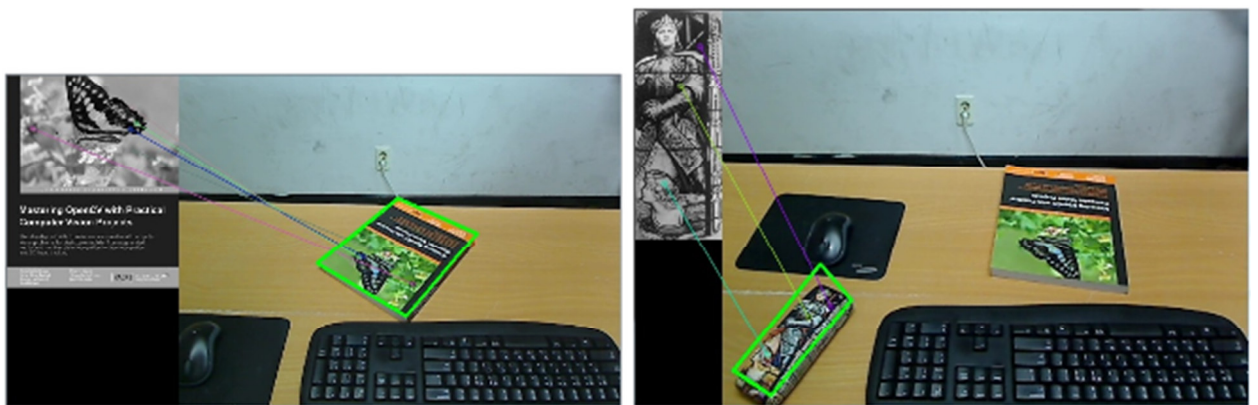


Рисунок 4.2 - Распознавание объектов и отслеживание результатов в сложной среде

В таблице 4.1 показана улучшенная производительность алгоритма по сравнению с существующими. Предложенный метод эффективно уменьшает время обработки, чтобы найти соответствующую точку через получение конкретной информации об ориентации через расширенное окно и сокращение размера дескриптора функции. Поэтому данный метод может улучшить проблему, такую как потеря интереса, когда похожий цветовой фон существует, а высокая вычислительная сложность распознает объект, используя характерные точки.

Таблица 4.1 – Результаты эксперимента

Метод	Коэффициент распознавания	Время распознавания
Предлагаемый метод	95%	0.49
Метод SURF	94%	0.65
Метод SIFT	96%	4.82

4.2 Вычисление методом RANSAC

Во время выполнения экспериментов были рассмотрены различные комбинации параметров: размер облаков-гипотез S , порог T , период выборки кадров N , количество итераций алгоритма RANSAC k . Каждый набор параметров запускался 10 раз, затем считалось среднее значение для каждого момента времени. Во всех запусках длина серии была принята в 16 кадров.

В ходе экспериментов в качестве значений брались 0.01, 0.03, 0.05, 0.10. Значения 0.01, 0.03 давали достаточно большой разброс результатов, а количество «хороших» особенностей могло уменьшаться в несколько раз относительно общего количества. При значениях 0.05 и 0.10 результаты получались лучше, однако вследствие того, что при значении 0.10 количество особенностей в «хороших» облаках практически совпадало с общим числом особенностей (попадали особенности с большой ошибкой измерения), результаты при $T = 0.05$ были наилучшими. Параметры k и S явным образом связаны между собой. В качестве параметра S использовались различные значения: от минимального $S = 3$, до $S = 50$, а параметр k непосредственно отвечает за соответствие времени работы алгоритма в реальном времени. При $S = 3$ работа алгоритма оказывалась неустойчивой даже при максимальных значениях параметра k , однако, тенденция к уменьшению разброса результатов при увеличении параметра S позволила уменьшать параметр k . Наилучшие результаты были достигнуты при значениях $k = 100$, $S = 50$. Для параметра N также пробовались несколько значений 1, 4. Значения $N = 1$ и $N = 4$ дали интересные результаты: при значении $N = 4$ показатели улучшились по сравнению с $N = 1$. Основными факторами, которые повлияли на поведение алгоритма в данной ситуации, являются характер сцены и движения камеры. При минимальном перемещении особенность может попадать в тот же пиксел изображения камеры, что и на предыдущем шаге, и таким образом давать дополнительную погрешность.

При $N = 4$ особенности становятся более «различимы» на изображениях, но при этом расстояние между одной и той же особенностью на кадрах с соседних шагов находится в рамках корректной работы отслеживания, результаты выведены в таблице 4.2.

Таблица 4.2 – Вычисление методом RANSAC

	T = 1		T = 4	
	S = 50	Без RANSAC	S = 50	Без RANSAC
Ошибка положения, м	0.0901	0.18	0.0750	0.1298
Ошибка поворота, м	0.0991	0.2642	0.0963	0.1381
Время работы (реальное = 128 сек.)	114.7312	94.3912	58.9953	39.1953

4.3 Распознавание методом SIFT

Чтобы протестировать возможности предложенного подхода, было применено использование базы данных к проблеме идентификации дорожных знаков. Эта база данных содержит 360×270 измеряли 48 изображений, и три сигнала состоят в том, чтобы распознать: пешеход, велосипед и перекресток. Рисунок 2 показывает пример из обнаруженных блобов и изображений, которые впоследствии были извлечены.



а)

б)

Рисунок 4.3 - Обнаруженные blobs и извлеченные сигналы

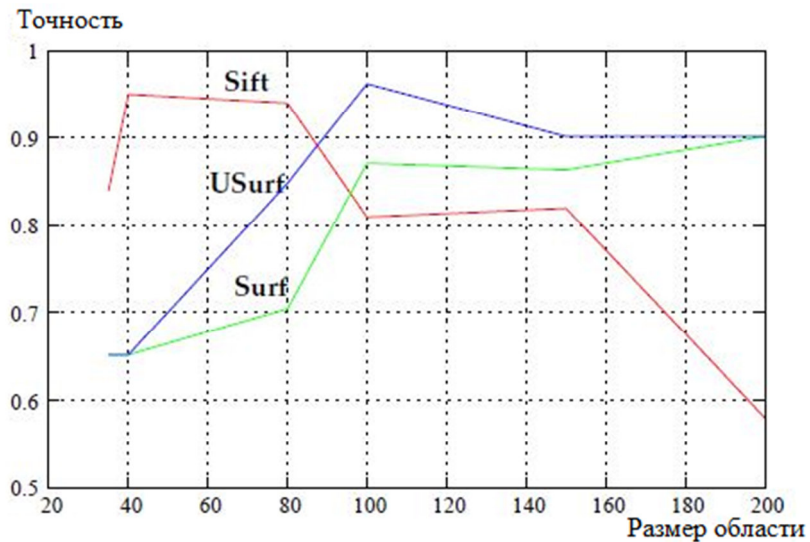
Для расчета точности классификации предполагается, что изображение может содержать не более одного сигнала каждого типа, т. е. они не могут быть больше трех сигналов на изображение. Таким образом, вычисляется точность суммирования истинных положительных и истинных отрицательных случаев. Но точность считается довольно грубой оценкой, которая не дает много информации о производительности классификатора. Вместо этого используется мера $F1$ в качестве основного оценочного показателя при объединении точности и возврата в единую метрику, благоприятствует сбалансированной производительности двух показателей [3]

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4.1)$$

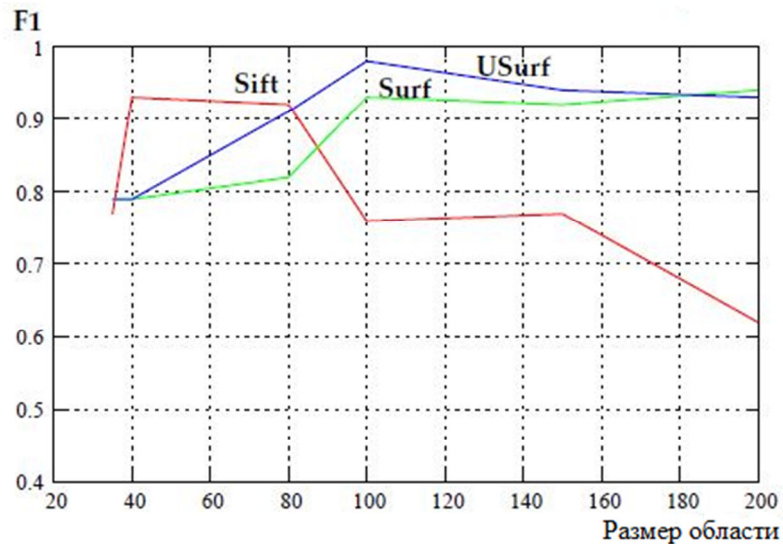
где $Precision$ – точность классификации, $Recall$ – возврат в единую метрику.

На рисунке 4.4 показаны полученные результаты. Точность и Размер F1 показаны для того, чтобы оценить сильные и слабые стороны подходов. Для размера большой области интересов число 150 ложных срабатываний значительно увеличивается при применении SIFT. Лучшие результаты

получены для малых размеров области, потому что размер сигнала на изображениях приближается. Несмотря на сложность задачи, точность увеличена до 95% для размера области 40×40 . Для большей рентабельности масштабирование размеров изображения резко влияет на стабильность ключевых моментов, производя ухудшение на эффективность классификации.



а)



б)

Рисунок 4.4 - Результаты для изменения размеров области

Однако, для SURF и USURF нужны большие размеры области интереса по сравнению с SIFT, чтобы достичь своих лучших результатов классификации. Небольшие размеры области не подходят, потому что количество извлеченных точек очень мало, как видно из таблицы 4.3.

Таблица 4.3 – Среднее количество базы данных ключевых точек на сигналах изображений

Размер \ Метод	30	40	80	100	150	200
SIFT	0.66	18.66	49.66	72	119.33	159
(U)SURF	0.33	1.66	14	24	48.33	59.66

Распознавание на примере дорожного знака

Вместо обработки всего изображения, целью данного метода является то, чтобы сначала извлечь область интереса (ROI) и дальнейшей его обработки. Некоторые ограничения размера были введены к обнаруженному объекту, чтобы гарантировать, что этот выбранный объект соответствует с большой вероятностью. На рисунке 4.5 показаны примеры результатов извлечения области интересов для объекта.



а)



б)

Рисунок 4.5 - Извлечение массива двоичных данных и масштабирование области интереса

Характеристики баз данных были следующими:

Так как объект для распознавания был выбран прямоугольной формы, то размер дескрипторов был приблизительно 5000 изображений и база данных, содержала 19 изображений масштабируемых к соответствующему размеру области интереса. Использовались соответствие критериев ключевых точек для эксперимента.

Таблица 4.4 - Результаты эксперимента без извлечения области интереса

Метод \ Метрика	Прямоугольная форма объекта	
	Точность	F1
SIFT	55.1	36.22
SURF	42.34	32.41
USURF	47.0	34.44

Таблица 4.5 - Результаты эксперимента обозначения области + извлечение признаков

Метод \ Метрика	Прямоугольная форма объекта		
	Точность	F1	Размер
SIFT	91.52	66.92	80
SURF	87.77	40.6	40
USURF	87.77	40.6	40

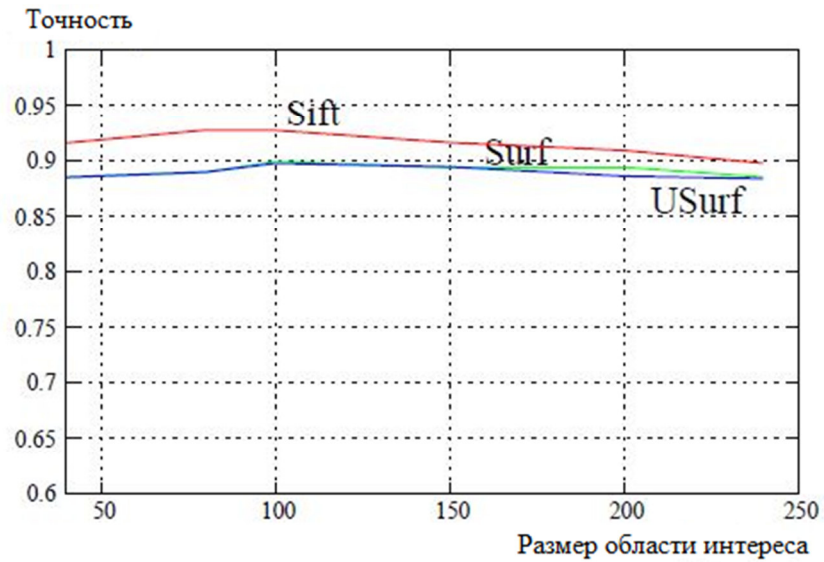
Таблица 4.6 - Результаты эксперимента обозначения области + извлечение признаков + морфологическая реструктуризация

Метод \ Метрика	Прямоугольная форма объекта		
	Точность	F1	Размер
SIFT	91.67	69.96	80
SURF	89.87	66.84	100
USURF	89.71	66.62	100

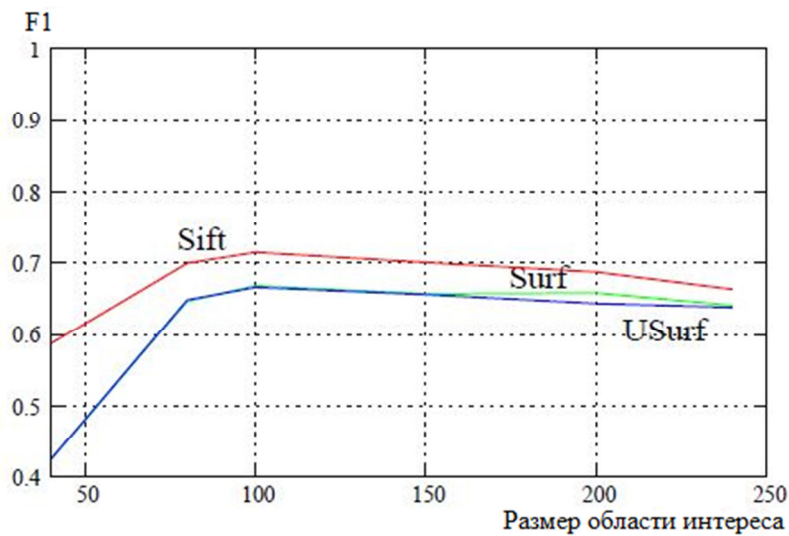
Таблица 4.4 показывает самую лучшую точность и соответствия F, измеренных значений, полученных для различных методов, вместе с соответствующим размером области интереса.

Таблица 4.5 показывает результаты, полученные с применением ключевого пункта методов извлечения исходных изображений (без области интереса) Эти представления были нужны для сравнения вычисления. USURF дает результат лучше SIFT в обоих случаях.

Таблица 4.6 показывает улучшение полученного метода, предложенного в работе. Опять же, USURF, кажется самый нормальным методом извлечения характеристики для объекта, но SIFT более лучший метод, который нужно приложить к полученной области интереса. Точность повышается почти на 30% для SIFT, и около 20-22% для SURF и USURF, соответственно для объекта. Прямоугольная форма объекта лучше с увеличением на 35% для SIFT и около 40% для SURF. Следует обратить внимание, что это улучшение также отражается в измерении F1, доказывая адекватность методов.



а)



б)

Рисунок 4.6 – Изменение показателей прямоугольного объекта

На рисунке 4.6 показано изменение показателей. Обе классификации точность и мера F1 отображаются при увеличении размера области интереса. Эти значения соответствуют предложенному двухэтапному методу с ранее упомянутым морфологическим ограничением. SIFT улучшает ускорять-вверх варианты когда прямоугольные ручки потребность быть обнаруженным, даже для более больших размеров области интересов.

В таблице 4.7 показано среднее число ключевых моментов увеличения вместе с размером области интереса. Эти значения были рассчитаны с использованием эталонов баз данных, используемых в каждом эксперименте

испытания и, следовательно, являются лишь ориентировочными. Хотя количество ключевых точек увеличивается вместе с размером области интересов, стабильные характерные точки найдены на небольших размерах изображения. Возрастающий масштаб области интереса дает большее количество точек, но повторяемость этих новых точек не положительна.

В таблице 4.8 показано среднее время, необходимое для обработки одного изображения (расположения объектов, поиска точки и сопоставления против эталонной базы данных). Стоит обратить внимание, что необходимое время выше для больших размеров области интереса из-за большого количества ключевых точек, которые происходят в изображениях.

Таблица 4.7 - Среднее количество ключевых точек

Размер	40	80	100	150	200	320x240
Объект прямоугольной формы	5.95	23.47	38.79	72.10	97.78	139.89

Таблица 4.8 – вычисления изображения с использованием SIFT

Размер	40	80	100	150	200	Без области интереса
Объект прямоугольной формы	0.06	0.13	0.17	0.42	0.80	0.69

ЗАКЛЮЧЕНИЕ

На основании проведенных исследований в магистерской работе получены следующие основные результаты:

1. Произведен анализ существующих методов распознавания на видео-данных.

2. Выполнен выбор методов для исследования распознавания объектов. Из существующих на сегодняшний день методов, более приемлемыми и результативными являются методы RANSAC, SIFT и SURF.

3. Исследуемые методы распознавания имеют свои достоинства и недостатки, более того, методы SIFT и SURF дополняют друг друга.

4. Разработаны алгоритмы распознавания объектов:

- На основе метода SIFT, который распознает локальные характеристики изображения и описывает их. С помощью данного метода, можно получить инвариантные признаки в отношении размера и поворота, которые не зависят от шума и ряда аффинных преобразований. Достоинством метода является устойчивость к небольшим окклюзиям и инвариантность определения движущегося объекта.

- На основе метода SURF, который включает в себя выполнение двух задач: поиск особых точек изображения; создание их дескрипторов, инвариантных к размеру и циркуляции.

Результат проведенных экспериментов показывают, что работа алгоритма быстра и надежна, чем традиционные методы, отслеживающие объекты в различных средах.

5. Выполнена программная реализация разработанных алгоритмов распознавания объектов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Edwards G. Interpreting face images using active appearance models / G. J. Edwards, C. J. Taylor, T. F. Cootes // Proc. Intl. Conf. on Automatic Face and Gesture Recognition – 1998. – P. 300–305.
2. Matthews I. 2D vs. 3D Deformable Face Models: Representational Power, Construction and Real-Time Fitting / I. Matthew, J. Xiao, S. Baker // Intl. Journal of Comp. Vision – Springer, 2007. – Vol. 75, No. 1. – P. 93–113.
3. Matthews I. Active Appearance Models Revisited / I. Matthews, S. Baker // Intl. Journal of Computer Vision – Hingham, MA: Kluwer Academic Publishers. – 2004. – Vol. 60, No. 2. – P. 135–164.
4. Golomb L. A., Lawrence D. T. and Sejnowski T. J. / SexNet: A neural network identifies sex from human faces// Advances in Neural Information Processing Systems, Morgan Kaufmann Publishers. San Mateo. USA. 1991. P. 77-83.
5. Takas B. and Wechsler H. /Locating features using SOFM// Proceedings of IAPR. 1994. Vol. 2. P. 55-60.
6. Feraud R., Bernier O. J., Viallet J-E., Collobert M A fast and accurate face detector based on neural networks// IEEE Trans. on Pattern Analysis and Machine Intel. 2001. Vol. 23. № .1. P.42-52.
7. Krüger N. An algorithm for the learning of weights in discrimination functions using a priori constrains// IEEE Trans. on Pattern Analysis and Machine Intel. 1997. Vol. 19. № .7. P.764-768.
8. Lades M., Vorbruggen J., Buhmann J., Lange J., Masburg C. and Wurtz R. / Distortion invariant object recognition in the dynamic link architecture// IEEE Trans. on Computers. — 1993. Vol. 42. P. 300-311.
9. Valentin D. and Abdi H. / Can linear autoassociator recognize faces from new orientations// Journal Opt. Soc. Am. A.. 1996. Vol. 13. P. 522-530.

10. Wiskott L. /Phantom faces for analysis // Pattern Recognition. 1997. Vol. 30. № .6. P. 837-846.
11. Попова Л. П., Датьев И. О.. Обзор существующих методов распознавания образов [Текст] / — М.: Сборник научных трудов, 2007. — 11 с.
12. Лифшиц Ю. Курс «Современные задачи теоретической информатики» — лекции по статистическим методам распознавания образов, распознаванию лиц, классификации текстов [Электронный ресурс] / – 2005. – 12 с. <http://yury.name/modern/07modernnote.pdf> (дата обращения 10.04.2018)
13. Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии [Текст] / – М.: Наука, 1988. – 280 с. (дата обращения 19.02.18)
14. Дюк В.А. Компьютерная психодиагностика [Электронный ресурс] / – СПб: Братство, 1994. – 365 с.
15. Бусленко Н. П., Голенко Д. И. Метод статистических испытаний (метод Монте-Карло) [Электронный ресурс] / . Физматгиз, 1961. – 228 с.
16. Путятин В.М. Распознавание изображений на основе вычисления их признаков корреляции между ними [Электронный ресурс] // Известия Южного федерального университета. Технические науки. Выпуск №2/том 79/2008.
17. Давыдов А.В. Распознавание объектов изображений: Конспект лекций. 2008. [Электронный ресурс] — <http://refdb.ru/look/1526317.html>. (дата обращения 04.04.2018)
18. Журавель И.М. Краткий курс теории обработки изображений. Распознавание объектов на основе вычисления их признаков. [Электронный ресурс] / - <http://matlab.exponenta.ru/imageprocess/book2/58.php> (дата обращения 04.04.2018 г.)
19. Прэтт У. Цифровая обработка изображений: В 2 т. [Текст] / М.: Мир, 1982.

20. Писаревский А. Н., и др. Системы технического зрения: (Принципиальные основы, аппаратное и математическое обеспечение). [Электронный ресурс] // Л.: Машиностроение, 1988.
21. Halarick R. M., Shapiro L. G. Computer and Robot Vision // V. I. Addison-Wesley, 1992.
22. Department of Electrical and Computer Web Pages. License Plate Recognition Final Report. http://www.ece.cmu.edu/~ee551/Final_Reports/Gr18.551.S00.pdf (дата обращения 30.04.2018 г.)
23. Kim K.I., Kim K.K., Park S.H., Jung K., Park M.H., Kim H.J. VE-GA VISION: A Vision System For Recognizing License Plates // IEEE Intern. Symp. on Consumer Electronics (ISCE 99). – 1999. – V. 2. – P. 176–181.
24. Shapiro L. G., Stockman G. / Computer vision. Prentice-Hall / - 2001.
25. Szeliski R. / Computer vision: algorithms and applications. Springer / , 2010.
26. Martinsky O. / Algorithmic and mathematical principles of automatic number plate recognition systems / , B. Sc. Thesis, Brno University of Technology, 2007.
27. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход [Текст] / Форсайт Д., Понс Ж. – М.: Издательский дом «Вильямс», 2004. – 465с.
28. Lowe, D. (2004) Distinctive image features from scale-invariant keypoints, / International Journal of Computer Vision, 60: 91–110.
29. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. / In Proc. of the 2001 // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01), Kauai, Hawaii, USA, volume 1, pages 511–518, December 2001.
30. D. G. Lowe. Distinctive image features from scale-invariant keypoints. / International Journal of Computer Vision / , 60(2):91–110, 2004.

31. 1 implementation. In CSCS14, volume 1, pages 243–248, 2003. [2] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded up robust features. In Proc.
32. Пименов В.Ю. Метод поиска нечетких дубликатов изображений на основе выявления точечных особенностей // Труды семинара РОМИП 2007 [Электронный ресурс]- 2008. СПб.: НУ ЦСИ, 2008. С. 145-159.
33. Гладышева Ю.С. Поиск изображений по признаку самоподобных локальных дескрипторов [Текст], 2011.
34. Ke Y., Suthanakar R. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors // Comput. Vis. And Pattern Recogn. — 2004. Vol. 2.- P. 506-513.
35. Никитин А.Н. Идентификация и локализация клеток на микроскопических изображениях с использованием SIFT-алгоритма [Электронный ресурс]// Методы микроскопического анализа. Сборник статей. Выпуск 1. - Москва: Издание ЗАО "Медицинские Компьютерные Системы (МЕКОС)", 2009 . - С. 114 – 125.
36. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва: Техносфера, 2005. – 1072с 6. [Электронный ресурс]// Гармаш М.И. Подходы к решению задачи визуального определения местоположения в помещении. ВЕСТНИК ОГУ №9 (115) /сентябрь 2010.
37. Гороховатский В.А. Метрики на множествах ключевых точек изображений [Электронный ресурс]// БИОНИКА ИНТЕЛЛЕКТА. 2008. № 2 (69).
38. Tuytelaars T., Mikolajczyk K. Local Invariant Feature Detectors: A Survey // Foundation and Trends in Computer Vision, Vol.3, No. 3, 2007. – pp. 177-280.
39. Viola P., Jones M.J. Robust Real-Time Face Detection // international Journal of Computer Vision 57(2). – 2004. – pp. 137-154.
40. Vedaldi Andrea. An Implementation of Multi-Dimensional Maximally Stable Extremal Regions. February 7, 2007.

41. Mikolajczyk K., Schmid C. A Performance Evaluation of Local Descriptors // IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.27, No.10, 2005. – pp. 1615-1630.
42. Tola E., Lepetit V., Fua P. A Fast Local Descriptor for Dense Matching // IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08), 2008. – pp. 1-8.
43. Calonder M., Lepetit V., Strecha C., Fua P. BRIEF: Binary Robust Independent Elementary Features // 11th European Conference on Computer Vision (ECCV), 2010.
44. Mood, A.; Graybill, F.; Boes, D. (1974). Introduction to the Theory of Statistics (3rd ed.). McGraw-Hill. p. 229.
45. Ojala T., Pietikäinen M., and Harwood D. A Comparative Study of Texture Measures with Classification Based on Feature Distributions// Pattern Recognition, vol. 29, pp. 51-59, 1996.
46. M.A.Fischler, R.C.Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography. CACM 24.381-395. 1981.
47. D. Nister. Preemptive RANSAC for Live Structure and Motion Estimation, // IEEE International Conference on Computer Vision, pp. 199-206, 2003.
48. D. Fontanelli, L. Ricciato, and S. Soatto. A fast ransacbased registration algorithm for accurate localization in unknown environments using lidar measurements, in Proc.// IEEE Int. Conf. on Automation Science and Engineering, Scottsdale, AZ, USA, 22-25 September 2007, pp. 597-602.
49. B. Kitt, A. Geiger and H. Lategahn. Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme, // IEEE Intelligent Vehicles Symposium (IV), pp. 486-492. 2010.
50. D. Scaramuzza , F. Fraundorfer and R. Siegwart. Realtime monocular visual odometry for on-road vehicles with 1-point RANSAC, Proc. // IEEE Int. Conf. Robot. Autom., pp.4293 -4299 2009.

51. K. Berthold, P. Horn. Closed-form solution of absolute orientation using unit quaternions. 1987.
52. Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. / Carnegie Mellon University Technical Report CMU-CS-91-132, / April 1991.
53. Jianbo Shi and Carlo Tomasi. Good Features to Track. // IEEE Conference on Computer Vision and Pattern Recognition, pages 593–600, 1994.
54. O. Chum, J. Matas, J. Kittler. Locally Optimized RANSAC. / DAGM 2003.
55. H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, / 110(3):346–359, 2008.
56. M. M. Bhajibhakare and P. K. Deshmukh. Detection and tracking of moving object for surveillance system. / *International Journal of Application of Innovation in Engineering and Management*, / 2(12):298–301, 2013.
57. M. Brown and D. Lowe. Invariant features from interest point groups. // In Proc. of the 2002 British Machine Vision Conference (BMVC'02), Cardiff, UK, pages 656–665. // British Machine Vision Association, September 2002.
58. D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. // *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):564–577, 2003.
59. S. A. Dave, M. Nagmode, and A. Jahagirdar. Statistical survey on object detection and tracking methodologies. / *International Journal of Scientific and Engineering Research*, / 4(3):1–8, 2013.
60. M. Du, J. Wang, J. Li, H. Cao, G. Cui, J. Lv, and X. Chen. Robot robust object recognition based on fast surf feature matching. In Proc. of the 2013 Chinese Automation Congress (CAC'13), Changsha, China, pages 581–586. // IEEE, November 2013.
61. D. Exner, E. Bruns, D. Kurz, and A. Grundhofer. Fast and Robust CAMShift Tracking. In Proc. of the 2010 // IEEE Computer Society Conference on

Computer Vision and Pattern Recognition Workshops (CVPRW'10), San Francisco, California, USA, pages 9–16. // IEEE, June 2010.

62. A. R. Francois. Camshift tracker design experiments with intel opencv and sai. Technical Report IMSC-04- 423, / Institute for Robotics and Intelligent Systems, University of Southern California, / August 2004.

63. S.-W. Ha and Y.-H. Moon. Multiple object tracking using sift features and location matching. / International Journal of Smart Home, / 5(4):17–26, 2011.

64. K. Huang, L. Wang, T. Tan, and S. Maybank. A real-time object detecting and tracking system for outdoor night surveillance. / Pattern Recognition, 41(1):432–444, 2008.

65. L. Juan and O. Gwun. A comparison of sift, pca-sift and surf. / International Journal of Image Processing, / 3(4):143–152, 2009.

66. Y. Ke and R. Sukthankar. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. In Proc. of the 2004 // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), Washington, DC, USA, volume 2, pages 506–513. // IEEE, June-July 2004.

67. Y.-H. Lee, J.-H. Park, and Y. Kim. Comparative analysis of the performance of sift and surf. Journal of the Semiconductor & Display Technology, / 12(3):59–64, 2013.

68. I. Leichter, M. Lindenbaum, and E. Rivlin. Mean shift tracking with multiple reference color histograms. / Computer Vision and Image Understanding, / 114(3):400–408, 2010

69. Живрин Я. Э., Алкзир Н. Б. Методы определения объектов на изображении // Молодой ученый. — 2018. — №7. — С. 8-19. — URL <https://moluch.ru/archive/193/48447/> (дата обращения 19.05.2018).

70. E. Jauregi, J. M. Martinez-Otzeta, B. Sierra, and E. Lazkano. Door handle identification: a three-stage approach. In IAV-07: International Conference on Intelligent Autonomous Vehicles, volume I, 2007.

71. W. Ye and Z. Zhong. Robust people counting in crowded environment. In Proceedings of the 2007 // IEEE International Conference on Robotics and Biomimetics, / pages 1133–1137, 2007.
72. H. Bay, T. Tuytelaars, and L. Van Gool. / SURF: Speeded up robust features. In Proc/.
73. L. Ledwich and S. Williams. Reduced SIFT features for image retrieval and indoor localisation. / In Australian Conference on Robotics and Automation, / 2004.

ПРИЛОЖЕНИЕ А

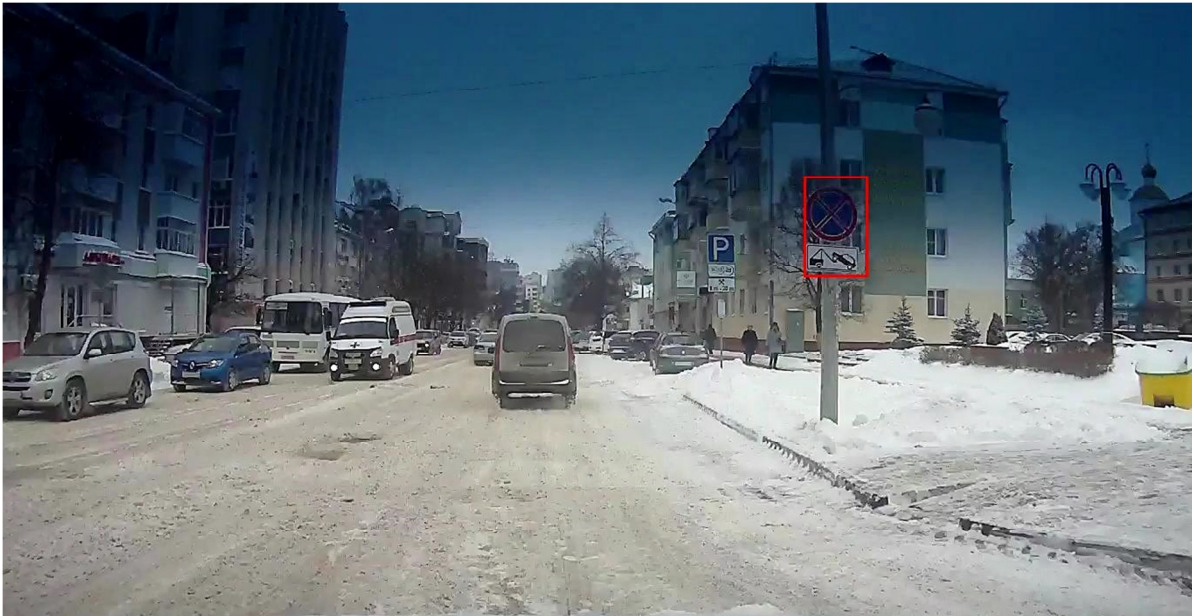


Рисунок 1 – Результат распознавания дорожного знака

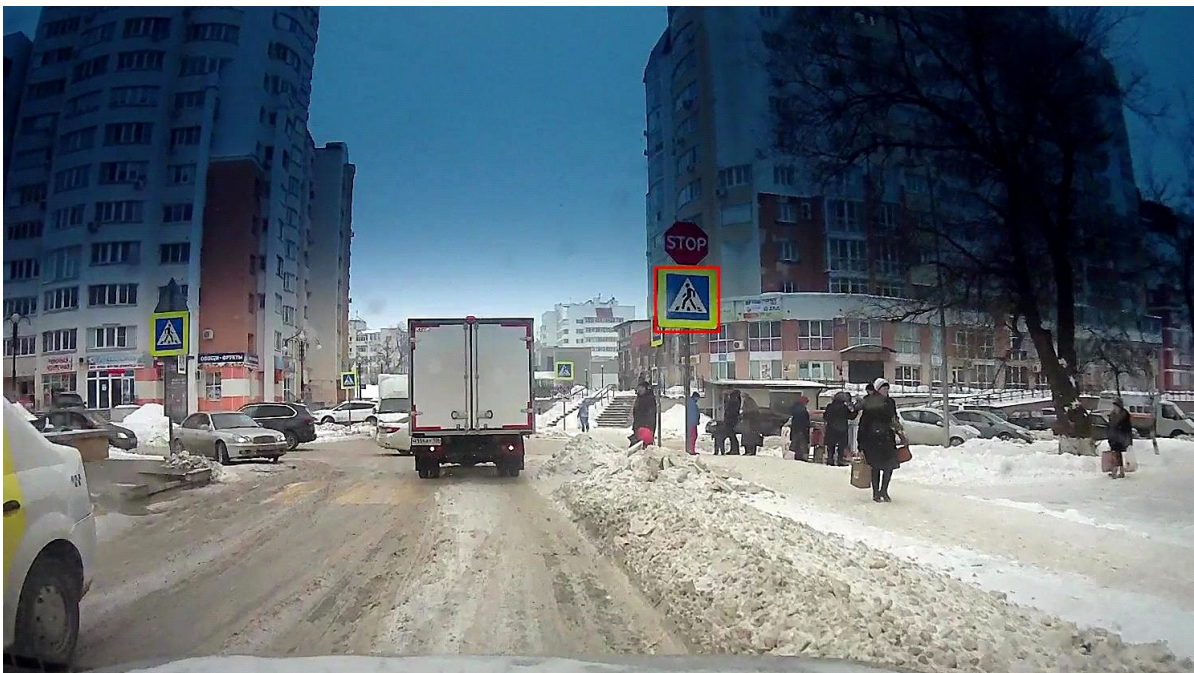


Рисунок 2 – Результат распознавания дорожного знака

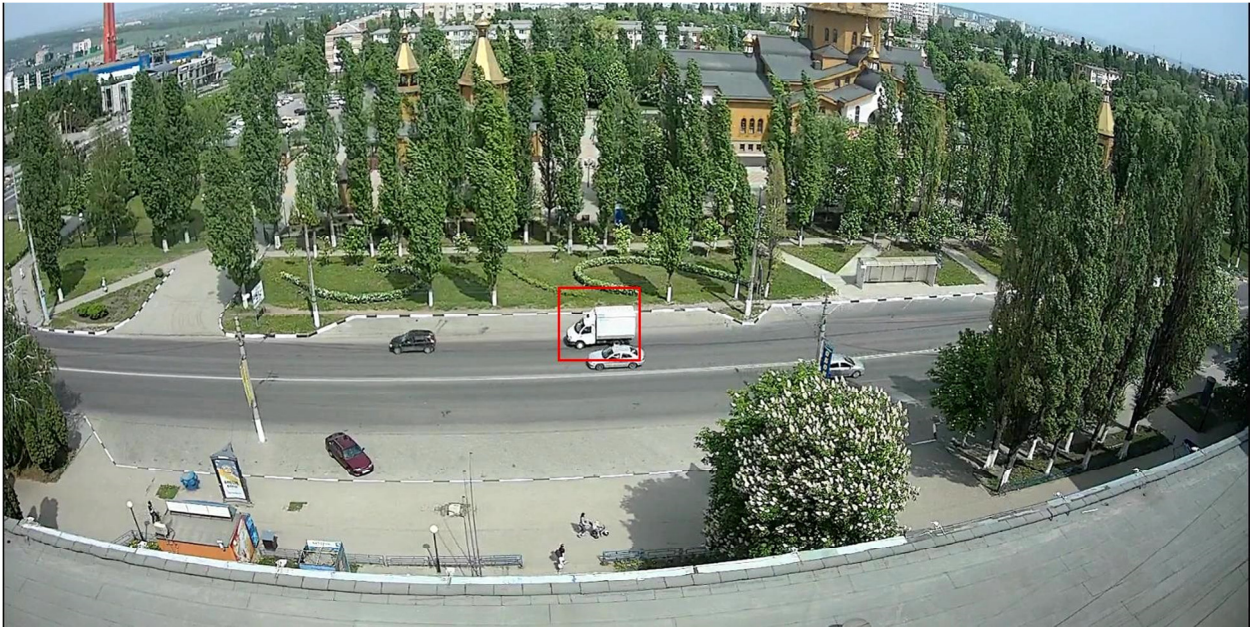
ПРИЛОЖЕНИЕ Б

Рисунок 1 – Результат распознавания транспортного средства



Рисунок 2 – Результат распознавания квадрокоптера