

## Оглавление

<b>Введение .....</b>	<b>4</b>
<b>Глава 1. Обзор существующих методов .....</b>	<b>7</b>
1.1. Структура алгоритмов обнаружения объектов на изображении.....	7
1.2. Обзор существующих методов построения признакового описания объекта на изображении .....	8
1.2.1. Алгоритм GIST .....	9
1.2.2. Алгоритм HOG.....	10
1.2.3. Алгоритм COHOG .....	12
1.2.4. Моментные инварианты Ху.....	13
1.3. Методы локализации объектов на изображении .....	16
1.3.1. Методы, основанные на поиске формы объекта.....	16
1.3.2. Методы, основанные на цвете объекта.....	17
1.3.3. Методы, основанные на машинном обучении .....	18
1.4. Выводы по главе 1 .....	20
<b>Глава 2. Информационная модель обнаружения объектов на изображении .....</b>	<b>22</b>
2.1 Общая структура системы распознавания объекта на изображении .....	22
2.1.1. Предварительная обработка изображения .....	24
2.1.2. Детектирование объекта на изображении .....	26
2.1.3. Формирование глобального признакового описания изображения с помощью теории активного восприятия .....	27
2.1.4. Принятие решения.....	34
2.2. Выводы по Главе 2 .....	35
<b>Глава 3. Вычислительный эксперимент.....</b>	<b>36</b>
3.1. Описание эксперимента .....	36
3.2. Описание программного продукта.....	37
3.2.1. Калибровка цветовой маски .....	37
3.2.2. Предварительная обработка входного изображения.....	41
3.2.3. Постобработка результатов локализации .....	42
3.2.4. Классификация обнаруженных объектов .....	44
3.3. Постановка вычислительного эксперимента .....	46
3.4. Анализ результатов вычислительного эксперимента.....	47
3.5. Выводы по главе 3 .....	52
<b>Заключение.....</b>	<b>53</b>
<b>Список литературы.....</b>	<b>54</b>

					<b>ВКР-НГТУ-09.04.01-(М17-ИВТ-3)-014-2019(ПЗ)</b>			
<b>Изм.</b>	<b>Лист</b>	<b>№ докум</b>	<b>Подп.</b>	<b>Дата</b>				
Разраб.		Смирнов А.В.			Модель и алгоритмы обнаружения объектов на изображении с использованием глобального признакового описания	Литера	Лист	Листов
Пров.		Гай В.Е.					3	55
Т. контр.						НГТУ им. Р.Е. Алексеева		
Н. контр.								
Утв.								

## **Введение**

### **Актуальность темы**

Распознавание объектов на изображениях стало в последнее время одной из важнейших задач анализа сцен и машинного зрения. Под объектом понимаем не только цифровое представление локального фрагмента двумерной сцены, а некоторое его приближенное описание, в виде набора характерных свойств (признаков). Основное назначение описаний (образов объектов) - это их использование в процессе установления соответствия объектов, осуществляемого путем сравнения (сопоставления).

Исследования по распознаванию образов пространственных объектов отличаются большим разнообразием в постановке задач и выборе средств их решения (методов обработки соответствующих фрагментов изображений), что является следствием разнообразия областей практического применения. Традиционными задачами, решавшимися еще в первых опытных разработках систем машинного зрения, служат задачи обнаружения и распознавания объектов, имеющих заданную форму на основе искаженных (зашумленных, повернутых, деформированных) изображений.

Исходя из этого, можно сделать вывод о том, что разработка новых моделей и методов для решения задачи обнаружения объектов на изображении является вполне актуальной.

### **Цель работы и задачи исследования**

Целью данной работы является разработка и исследование новых моделей и алгоритмов решения задачи обнаружения объектов на изображении с использованием глобального признакового описания. Для достижения этой цели необходимо решить следующие задачи:

- Обзор и анализ существующих известных методов решения задачи обнаружения объектов с использованием глобального признакового описания.
- Создание информационной модели описания объекта с использованием глобального признакового описания.
- Создание нового алгоритма формирования глобального признакового описания изображения.
- Проведение исследования с целью выявления наилучшей комбинации параметров разрабатываемой системы
- Проведение вычислительного эксперимента для установления корректности работы созданных моделей и алгоритмов.

### **Объект исследования**

Объектом исследования являются изображения, представленные в одном из доступных цифровых форматов.

### **Предмет исследования**

Предметом исследования являются модели и методы решения задачи обнаружения объектов на изображении.

### **Методы исследования**

В ходе выполнения работы были использованы методы формирования глобального признакового описания изображений на основе теории активного восприятия, методы принятия решения о локализации объекта на изображении на основе алгоритмов цветовой фильтрации. Также был использован метод вычислительного эксперимента, для проведения которого был разработан программный продукт на языке программирования Python.

### **Положения, выносимые на защиту**

На защиту выносятся следующие положения:

- Информационная модель глобального признакового описания изображения искомого объекта, построенная на основе положений теории активного восприятия.
- Алгоритм локализации объекта на изображении на основе цветовой фильтрации в пространстве HSV.
- Результаты вычислительного эксперимента.

### **Научная новизна**

Научная новизна предлагаемого метода решения задачи обнаружения объекта на изображении заключается в использовании нового способа формирования глобального признакового описания на этапе формирования признакового описания. В данной работе, таким подходом является использование инструментария теории активного восприятия.

### **Теоретическая и практическая ценность**

Теоретическая ценность работы заключается в разработанной информационной модели глобального признакового описания объекта, а также в предложенном алгоритме его локализации на изображении.

Практическая ценность работы заключается в созданном программном продукте, реализующем указанный выше алгоритм, а также в результатах проведенного вычислительного эксперимента.

### **Апробация работы**

Апробация данной работы была проведена в ходе защиты отчета по распределенной научно-исследовательской работе, который включал в себя основные положения проведенного исследования и анализ полученных результатов, а так же в ходе защиты публикации по данной работе на XXV международной научно-технической конференции «Информационные системы и технологии - 2019».

### **Обоснованность и достоверность**

Обоснованность и достоверность проведенного исследования доказывается проведением вычислительного эксперимента, анализом его результатов, и сравнением их с результатами существующих методов решения данной задачи.

### **Публикации**

1. А.В. Смирнов, В.Е. Гай - Модель и алгоритмы обнаружения объектов на изображении с использованием глобального признакового описания // Материалы XXV международной научно-технической конференции «Информационные системы и технологии - 2019», ИСТ-2019, Россия, Н. Новгород, 2019г.

### **Объём и структура**

Выпускная квалификационная работа состоит из введения, трёх глав основной части, заключения и списка литературы. Общий объём работы – 55 страниц. Диссертация содержит 25 рисунков, 8 таблиц, 30 формул. Список литературы включает в себя 27 наименований.

## Глава 1. Обзор существующих методов

### 1.1. Структура алгоритмов обнаружения объектов на изображении

Большая часть актуальных методов, решающих задачу обнаружения объектов на изображении, включают в себя три основных этапа:

- Предварительная обработка изображения.
- Построение признакового описания искомого объекта.
- Принятие решения на основе полученного признакового описания.

Предварительная обработка изображения является первым этапом в решении данной задачи. Цель этого этапа – представление изображения, содержащего искомый объект в виде, подходящем для последующего формирования его признакового описания. Выявление сегментов изображения, предположительно содержащих объекты, для которых необходимо получение глобального признакового описания, происходит при помощи алгоритмов локализации объектов на изображении.

Следующий этап решения данной задачи - построение признакового описания искомого объекта. На данном этапе для изображения объекта формируется набор признаков, позволяющий однозначно описать данный объект. Цель этого этапа – представление изображения искомого объекта в подходящем виде для последующего принятия решения о классификации данного объекта.

Третий этап заключается в поиске эталонного изображения объекта, соответствующего объекту, обнаруженному на изображении. Для этого этапа необходимо иметь базу изображений-эталонов, для каждого из которых необходимо получить признаковое описание. На третьем этапе необходимо принять решение о том, какому из признаковых описаний эталонных изображений соответствует данное признаковое описание.

Схема описанных выше этапов решения задачи представлена на рисунке 1.1.



Рисунок 1.1. – Структура системы обнаружения объектов на изображении

## 1.2. Обзор существующих методов построения признакового описания объекта на изображении

Процесс детектирования объектов на изображении представляет собой сложную и интересную задачу. Общие принципы работы существующих систем по обнаружению объектов достаточно схожи. Как правило, в процессе поиска производится сравнение признаков данного объекта или группы данных объектов в используемой базе изображений. Степени схожести объектов задаются параметрами системы и задачами, которые она выполняет. Главное преимущество глобальных дескрипторов заключается в том, что такого рода признаки демонстрируют достаточно высокую производительность, при этом обладая существенно меньшей сложностью вычислений. К этой категории можно отнести большую часть дескрипторов для выявления форм и текстур объектов. Каждый объект, представленный как дескриптор, соответствует точке в многомерном признаковом пространстве – этим объясняется особенность данных дескрипторов.

К локальным дескрипторам относятся методы признакового описания, которые выделяют определенные области изображения, а затем строят его полное описание, руководствуясь имеющейся информацией об обнаруженных объектах.

Напротив, для получения глобального признакового описания, используются низкоуровневые операции, которые производятся при структурном анализе изображения. Благодаря этому, такого рода дескрипторы могут нести информацию текстуре объекта, о его цвете, или возможном движении объекта. Глобальное признаковое описание изображений, как альтернатива локальному описанию, обычно применяется для

проектирования систем компьютерного зрения. К примеру, одной из таких задач может быть задача о классификации статических сцен на изображении.

Таким образом, глобальный дескриптор изображения показывает общую, низкоуровневую структуру изображения объекта, при этом, локальные структуры игнорируются и не учитываются таким дескриптором. Для классификации таких дескрипторов, удобно пользоваться евклидовым расстоянием, так как наименьшее евклидово расстояние между дескрипторами объектов означает, что изображения объектов, от которых были получены сравниваемые дескрипторы, имеют отношение к одной общей группе или же к единому общему для них классу.

Сегодня имеется несколько алгоритмов, которые применяются для вычисления глобального признакового описания объектов. Рассмотрим самые основные и наиболее популярные из них:

- GIST [1]
- HOG [2]
- COHOG [3]
- Моментные инварианты  $X_u$  [4]

### 1.2.1. Алгоритм GIST

GIST – Изначально метод вычисления GIST дескрипторов был описан в статье [1]. После её публикации, данный дескриптор нашел применение в различных областях, связанных с компьютерным зрением. В работе [5] рассматривается SLAM-система, основанная на многочастичных фильтрах для определения ранее посещенных локаций, описанных с помощью GIST-дескриптора. Система с использованием GIST-дескриптора для панорамных изображений представлена в работе [6]. GIST-дескриптор также использовался для поиска изображений в базе данных с миллионами записей [7].

Алгоритм базируется на допущении, что при помощи распределения градиентов интенсивности или направления краев можно описать внешний вид и форму объектов на изображении. Реализация такого рода дескриптора может осуществляться разделением изображения на сегментированные области, именуемые ячейками, и расчетом для каждой ячейки гистограммы направлений градиентов или направлений краев для пикселей, находящихся внутри ячейки. Комбинация такого рода гистограмм именуется дескриптором. С целью повышения точности, гистограммы нормализуются по контрасту. Для этого вычисляется мера интенсивности на так называемом блоке (большой фрагмент изображения), и результирующее значение используется для этапа нормализации. По

отношению к освещению, нормализованные дескрипторы обладают лучшей инвариантностью.

GIST суммирует градиентную информацию (масштабы и ориентацию) для разных частей изображения. Вычисления происходят следующим образом:

1. Свертка изображения, используя 32 фильтра Габора в 4 масштабах, 8 ориентациях.
2. Разделение каждой карты на 16 областей и усреднение значений в каждой области.
3. Конкатенация всех усредненных значений [8].

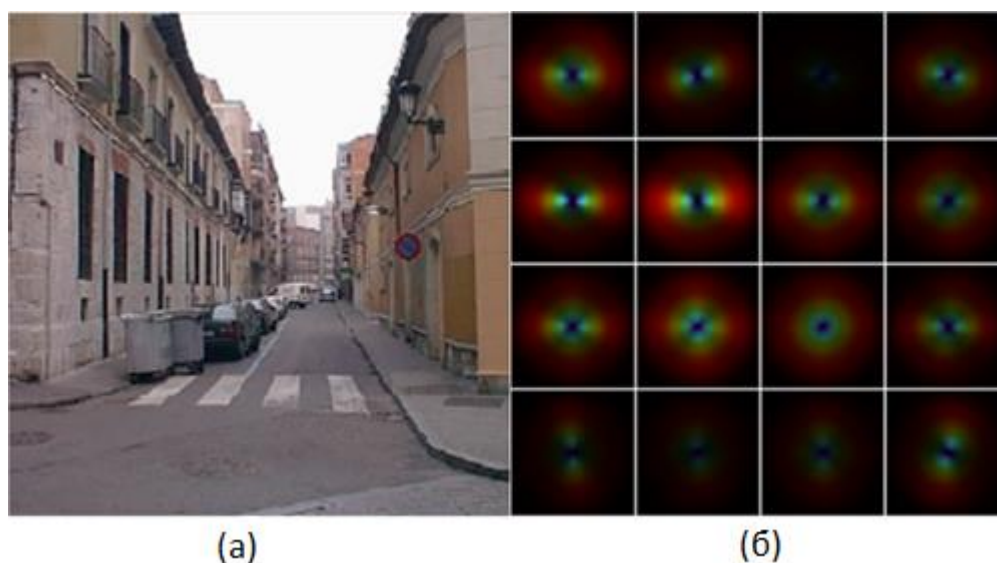


Рисунок 1.2. – Входное изображение (а) и его дескриптор GIST (б)

### 1.2.2. Алгоритм HOG

HOG – Гистограмма направленных градиентов – дескрипторы особых точек, которые используются в компьютерном зрении и обработке изображений с целью распознавания объектов. Этот метод похож на гистограммы направления края, дескрипторы SIFT и контексты формы, но отличается тем, что вычисляется на плотной сетке равномерно распределенных ячеек по всему изображению и использует нормализацию перекрывающегося локального контраста для увеличения точности.

Дескриптор HOG обладает некоторыми преимуществами по сравнению с прочими дескрипторами. Ввиду того, что HOG работает локально, данный метод обеспечивает инвариантность как геометрических, так и фотометрических преобразований, за исключением лишь ориентации самого объекта в пространстве. Такого рода искажения могут появиться только в больших участках изображения. Кроме этого, достаточно грубая сегментация, точность вычисления направлений и высокая степень локальной



фотометрической нормализации дают возможность игнорировать движения объектов, если они поддерживают вертикальное положение. Таким образом, данный дескриптор хорошо проявляется свои качества в задачах по обнаружению людей на изображениях.

В основе метода HOG лежит предположение, что вид распределения градиентов интенсивности изображения позволяет достаточно точно определить наличие и форму присутствующих на нем объектов [9]. При описании изображения оно разбивается на несколько небольших участков, далее называемых ячейками. В ячейках вычисляются гистограммы  $h_i$  направленных градиентов внутренних точек. Обычно они объединяются в одну гистограмму  $h = f(h_1, \dots, h_k)$ , после чего она нормализуется по яркости ( $L_2$  или  $L_1$  норма), где  $\varepsilon$  — некоторая малая константа.

$$h_{L_2} = \frac{h}{\sqrt{|h|_2^2 + \varepsilon}}, h_{L_1} = \frac{h}{\sqrt{|h|_1 + \varepsilon}}, h_{\sqrt{L_1}} = \sqrt{h_{L_1}}$$

Таким образом, данный дескриптор содержит пространственную информацию о фрагменте и инвариантен к освещению. При вычислении градиентов производится свертка изображения с ядрами  $[-1, 0, 1]$  и  $[-1, 0, 1]^T$ , в результате чего образуются две матрицы  $Dx$  и  $Dy$  производных вдоль осей  $x$  и  $y$  соответственно. Эти матрицы используются для вычисления углов и величин (модулей) градиентов в каждой точке изображения. Пусть множество углов  $(-\pi, \pi]$  разбивается на  $n$  равных интервалов вида  $(k-1n\pi, kn\pi]$ , где  $k = \{1, \dots, n\}$ . Каждому интервалу ставится в соответствие бин гистограммы. Тогда гистограмма ячейки заполняется так, что величина градиента в каждой внутренней его точке добавляется к величине бина, соответствующего интервалу, содержащему угол данного градиента.

На рисунке 1.3 представлена визуализация процесса получения дескриптора HOG:

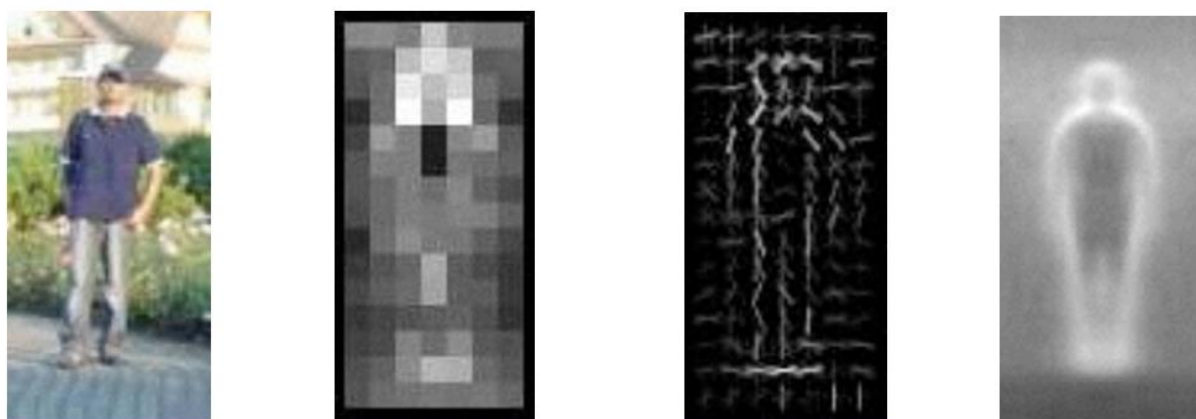


Рисунок 1.3. – Визуализация процесса получения дескриптора HOG

### 1.2.3. Алгоритм СОНОГ

СоНОГ – совместная гистограмма ориентированных градиентов – предложенная в статье [3], представляет собой дескриптор, основанный на наборе из нескольких пар допустимых градиентных ориентаций.

Таким образом, СоНОГ - это пары градиентных ориентаций. В НОГ для одного градиента ориентация имеет только восемь разновидностей, но пара из них имеет гораздо больше возможных разновидностей. Таким образом, СоНОГ может выражать формы более подробно, чем НОГ, который использует ориентацию с одним градиентом, как показано на рисунке 1.3 - набор допустимых ориентаций градиентов для дескрипторов НОГ (а) и СОНОГ (б). Как можно заметить, одиночный набор ориентаций градиентов (а) представляет собой набор из восьми доступных положений, в то время как набор пар, составленный из их комбинаций (б) имеет намного больше доступных положений.

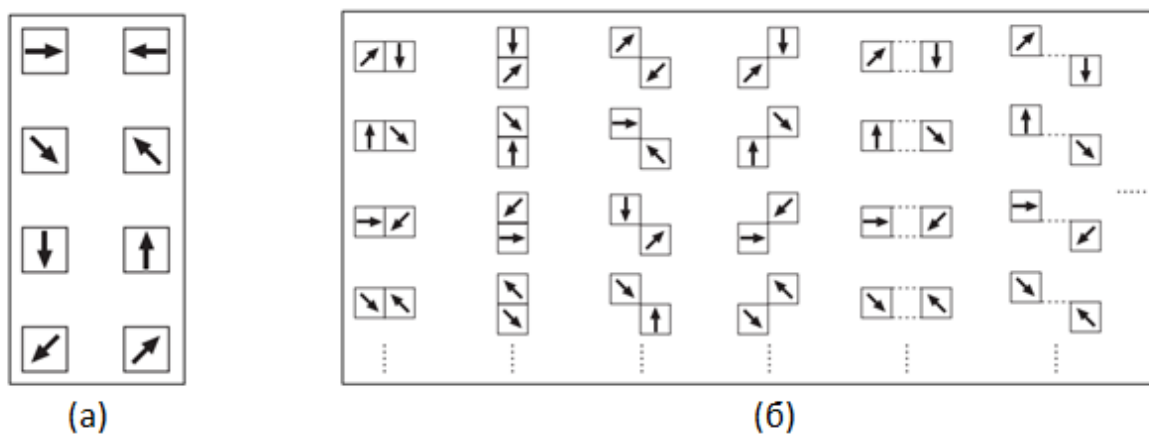


Рисунок 1.3. – Одиночный (а) и множественный (б) набор ориентаций градиентов

#### 1.2.4. Моментные инварианты Ху

Одним из хорошо известных подходов в распознавании образов на цифровом изображении является использование моментов различных порядков, а также моментных инвариантов. За последние сорок лет моментные инварианты, предложенные Ху, стали классическим инструментом распознавания образов. Приведенный им в [4] набор характеристик включает семь инвариантов.

Множество работ было посвящено обобщению и улучшению инвариантных признаков, предложенных Ху, а также попыткам применить их в приложениях из разных областей знаний. Например, в [10] и [11] было предложено использовать моментные инварианты для распознавания силуэтов самолетов на изображении, в [12] моментные инварианты применяются при сравнении с эталоном и совмещении снимков, полученных со спутника. Ряд авторов используют инварианты для распознавания символов. В [13] и [14] предложено использовать инварианты для изменений контраста, в [15] для классификации текстур, а в [16] описаны инварианты к линейным преобразованиям.

Алгоритм [17] получения инвариантов содержит следующие этапы:

- 1) определение центральных моментов порядка не выше третьего
- 2) получение моментов, инвариантных к операциям поворота, переноса и зеркального отображения [18]
- 3) получение моментов, инвариантных к полной группе аффинных преобразований

Рассмотрим подробнее основные его этапы и используемые формулы:

Для двумерной непрерывной функции  $f(x, y)$  момент (иногда называемый «необработанным моментом») порядка  $(p + q)$  определяется как

$$M_{pq} = \iint_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

где  $p, q = 0, 1, 2, \dots$ . Адаптируя это к скалярному (в оттенках серого) изображению с интенсивностью пикселей  $I(x, y)$ , необработанные моменты изображения  $M_{ij}$  вычисляются как

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y)$$

В некоторых случаях это можно рассчитать, рассматривая изображение как функцию плотности вероятности, то есть путем деления вышеуказанного на  $\sum_x \sum_y I(x, y)$ .

Теорема единственности [4] гласит, что если  $f(x,y)$  кусочно-непрерывна и имеет ненулевые значения только в конечной части плоскости  $x,y$ , то существуют моменты всех порядков, а последовательность моментов  $(M_{pq})$  однозначно определяется по  $f(x,y)$ . Наоборот,  $(M_{pq})$  однозначно определяет  $f(x,y)$ . На практике изображение суммируется с функциями нескольких моментов более низкого порядка. Центральные моменты определяются как:

$$\mu_{pq} = \iint_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

где  $\bar{x} = \frac{M_{10}}{M_{00}}$  и  $\bar{y} = \frac{M_{01}}{M_{00}}$  являются компонентами центра масс. Если  $f(x, y)$  является цифровым изображением, то предыдущее уравнение имеет вид:

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

Центральными моментами порядка до 3 являются:

$$\mu_{00} = M_{00}$$

$$\mu_{01} = 0$$

$$\mu_{02} = 0$$

$$\mu_{11} = M_{11} - \bar{x}M_{01} = M_{11} - \bar{y}M_{10}$$

$$\mu_{20} = M_{20} - \bar{x}M_{10}$$

$$\mu_{02} = M_{02} - \bar{y}M_{01}$$

$$\mu_{21} = M_{21} - 2\bar{x}M_{11} - \bar{y}M_{20} + 2\bar{x}^2M_{01}$$

$$\mu_{12} = M_{12} - 2\bar{y}M_{11} - \bar{x}M_{02} + 2\bar{y}^2M_{10}$$

$$\mu_{30} = M_{30} - 3\bar{x}M_{20} + 2\bar{x}^2M_{10}$$

$$\mu_{03} = M_{03} - 3\bar{y}M_{02} + 2\bar{y}^2M_{01}$$

Таким образом, можно показать, что:

$$\mu_{pq} = \sum_m^p \sum_n^q \binom{p}{m} \binom{q}{n} (-\bar{x})^{(p-m)} (-\bar{y})^{(q-n)} M_{mn}.$$

Центральные моменты являются трансляционно-симметричными. Моменты хорошо известны своим применением в анализе изображений, поскольку они могут использоваться для получения инвариантов относительно определенных классов преобразования. Следует заметить, что описанные ниже инварианты являются таковыми только в непрерывной области. В дискретной области ни масштабирование, ни вращение не определены четко: дискретное изображение, преобразованное таким образом, обычно является приближенным, и преобразование не является обратимым. Следовательно, эти инварианты являются лишь приблизительно инвариантными при описании формы в дискретном изображении. Инварианты  $\eta_{ij}$  относительно сдвига и масштаба могут быть построены из центральных моментов путем деления на правильно масштабированный нулевой центральный момент:

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(1+\frac{i+j}{2})}}$$

где  $i + j \geq 2$ . Следует заметить, что трансляционная инвариантность напрямую следует только за счет использования центральных моментов.

Как показано в работе Ху, могут быть построены инварианты относительно перемещения, масштаба и вращения:

$$I_1 = \eta_{20} + \eta_{02}$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].$$

Они хорошо известны как моментные инварианты Ху. Первый,  $I_1$ , аналогичен моменту инерции вокруг центра масс изображения, где интенсивности пикселей аналогичны физической плотности. Последний,  $I_7$ , является косоинвариантным, что позволяет ему отличать зеркальные изображения от других идентичных изображений.

### 1.3. Методы локализации объектов на изображении

Данный раздел посвящен методам, использующимся на этапе принятия решения о локализации объектов на изображении. Были подробно рассмотрены следующие алгоритмы:

1. Методы, основанные на поиске формы объекта;
2. Методы, основанные на цвете объекта;
3. Методы, основанные на машинном обучении;
  - 3.1 Нейронные сети;
  - 3.2 Метод опорных векторов;
  - 3.3 Деревья принятия решений;
  - 3.4 Каскадный классификатор.

#### 1.3.1. Методы, основанные на поиске формы объекта

Сегментацией называют процесс разделения целого изображения на несколько сегментов или множество пикселей, которые еще называют суперпикселями. Целью сегментации является упрощение представления изображения для последующего анализа. Обычно сегментацию используют для выделения объектов и границ (кривые, линии и т.п.) на изображениях. В ходе выполнения сегментации изображений происходит процесс присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками обладают общими визуальными характеристиками.

Пороговая обработка является наиболее простым методом сегментации, что привлекает к нему большое внимание ученых. Данный метод ориентирован на обработку изображений, отдельные однородные участки которых различаются средней яркостью. Самым простым, но при этом наиболее часто применяемым видом сегментации называют бинарную сегментацию, когда имеется только два типа однородных участков.

Еще одним методом из данной группы является преобразование Хафа [19]. Оно используется для поиска объектов, которые принадлежат строго определённого класса фигур. Важным моментом в данном методе является использование процедуры голосования, которая применяется к пространству параметров, из которого и получаются объекты определённого класса фигур по локальному максимуму. При вычислении трансформации Хафа получается так называемое, накопительное пространство.

Преобразование Хафа базируется на представлении объекта в виде параметрического уравнения. Параметры этого уравнения представляют фазовое или аккумуляторное пространство. Затем, берётся двоичное изображение, например, результат работы детектора границ Кенни. Все точки границ перебираются и делается

предположение о том, что точка принадлежит линии искомого объекта. Таким образом, мы рассчитываем нужное уравнение для каждой точки изображения и получаем необходимые параметры, которые сохраняются в пространстве Хафа. Заключительным шагом является обход пространства Хафа и выбор его максимальных значений, за которые «проголосовало» больше всего пикселей картинки, что и даёт нам параметры для уравнений искомого объекта.

### 1.3.2. Методы, основанные на цвете объекта

Наиболее распространенный подход в обнаружении объектов основан на нахождении области изображения, который содержит интересующий нас цвет. На полученных областях определяются либо объекты, либо области, где они могут гипотетически располагаться.

Главный недостаток такого подхода заключается в том, что в зависимости от времени суток, погодных условий, теней и т.д. цвет знака может быть искажен. Стандартное RGB цветовое пространство очень чувствительно к освещению, поэтому исследователи стараются использовать другие цветовые пространства, такие как, например, HSV. Но, тем не менее, на основе данного подхода было создано довольно много разнообразных методов. Альберто Бругги предложил способ, основанный на зависимости от источника света. Он состоит в том, чтобы найти белый объект в сцене и вычислить различие между изображением, белым и теоретически белым цветом. Однако, в реальных условиях нельзя рассчитывать на наличие белого ориентира. [20]

Серхио Эскалера предложил подход для обнаружения красного в цветовом пространстве HSV. Входное изображение сначала преобразовывается из цветового пространства RGB в HSV. Для каждого пикселя значение оттенка и насыщенности вычисляется так, чтобы диапазон красного был ярко выражен. Авторы предполагают, что значение оттенка и насыщенность измерена в диапазоне от 0 до 255. Полученные значения перемножаются; авторы заявляют, что подобный подход позволяет обезопаситься в случае ошибки нахождения одной из величин. [21]

Ксиоли Хао в своих исследованиях использовал цветовую модель CIECAM97. Изображения сначала преобразуются из RGB к CIE, и затем к LCH. Авторы рассматривают четыре случая: среднестатистические условия, а также условия во время солнечного, облачного и дождливого дня. Используя приемлемые диапазоны, изображение рекурсивно разделяется на сектора, благодаря чему достигается предопределенный размер пикселя. [22] Результаты подобных исследований говорили о высоком потенциале данного подхода, но различные факторы, такие как изменение

освещения, тени, неблагоприятные погодные условия значительно влияют на процент детектирования. Несмотря на это, использование методов по распознаванию дорожных знаков имеет место быть, как дополнительный фактор в алгоритме.

### **1.3.3. Методы, основанные на машинном обучении**

В настоящее время становятся все более и более популярными разнообразные алгоритмы машинного обучения. Такие алгоритмы и методы находят широкое применение в системах различного назначения: поисковых системах, алгоритмах распознавания, анализа и синтеза речи, медицинской диагностике, биоинформатике, финансовом прогнозировании и т.д.

Исключением не является и компьютерное зрение. Например, подавляющее большинство современных систем детектирования объектов на изображениях и видео основано на применении алгоритмов машинного обучения. Такой подход позволяет компьютерной системе самой «научиться» отличать изображения, содержащие искомый объект, от остальных, используя для этого лишь примеры таких изображений. Также кластеризация и обучение с учителем успешно применяется в алгоритмах классификации изображений, опять же позволяя автоматически установить неявные различия между изображениями различных типов. Рассмотрим подробнее несколько самых распространённых алгоритмов машинного обучения.

#### **1.3.3.1. Нейронные сети**

Нейросетевые методы включают в себя целый класс алгоритмов. Основная идея, которая лежит в их основе, заключается в последовательном преобразовании сигнала параллельно работающими функциональными элементами, или, как их еще называют, нейронами. Процесс обучения таких нейронных сетей сводится к уменьшению среднеквадратичной ошибки. Системы обнаружения объектов на изображениях, основанные на нейронных сетях, используют иерархическую структуру. Сначала вектор признаков обрабатывается грубой сетью с высоким уровнем ошибок второго рода, а затем, если вектор не был классифицирован как не объект, решение корректируется более точной, но при этом и более медленной сетью. Достоинством нейронных сетей является высокая точность обнаружения при правильной настройке параметров сети. Из недостатков можно выделить сильную чувствительность к шуму, необходимость в тщательной настройке параметров нейронной сети для получения хороших результатов. Самым же главным недостатком является высокая вычислительная сложность, и, как следствие, скорость работы, которая недостаточна для обработки в реальном времени.



### 1.3.3.2. Метод опорных векторов

Это набор схожих алгоритмов вида «обучение с учителем», который используется при решении задач классификации и регрессионного анализа. Данный метод принадлежит к семейству линейных классификаторов. Суть метода заключается в нахождении гиперплоскости, которая разделяет пространство объектов на два класса. Из всех возможных гиперплоскостей, разделяющих два класса, нужно выбрать такую гиперплоскость, расстояние до которой от каждого класса максимально. Эту гиперплоскость называют оптимальной разделяющей гиперплоскостью, а соответствующий ей линейный классификатор называется оптимально разделяющим классификатором. К достоинствам метода опорных векторов относят высокую устойчивость к переобучению, возможность уменьшения чувствительности к шуму за счет снижения точности. При этом, поиск оптимального соотношения данных параметров требует точной настройки. А также точность работы метода уступает другим методам машинного обучения.

### 1.3.3.3. Деревья принятия решений

Это один из алгоритмов принятия решений, который используется для различных моделей. Структура дерева состоит из ребер и узлов: на ребрах («ветках») дерева решения записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах — атрибуты, по которым различаются объекты. Чтобы классифицировать новый объект, необходимо спуститься по дереву до листа и выдать соответствующее значение. Подобные деревья решений широко используются в интеллектуальном анализе данных. Достоинствами алгоритма являются: высокая скорость обучения, высокое качество получаемых моделей (сравнимое с нейронными сетями и ансамблями нейронных сетей). Вместе с тем, можно отметить и недостатки алгоритма: необходимость в большом количестве памяти. Также обученная модель работает несколько медленнее других алгоритмов (если в модель входит 100 деревьев, мы должны пройти по всем, чтобы получить результат), что делает данный алгоритм непригодным для поиска объектов в видеоряде в реальном времени.

### 1.3.3.4. Каскадный классификатор

Каскадный классификатор или метод Виола-Джонса был предложен Полом Виолой и Майклом Джонсом [23] и стал первым методом, который демонстрировал высокую точность распознавания объектов на изображениях в реальном времени. В данном алгоритме используется набор признаков, близких к признакам Хаара совместно с одной из вариации алгоритма AdaBoost.

AdaBoost (Adaptive Boosting) – алгоритм усиления классификаторов путем объединения их в отдельные группы. Впервые данный алгоритм был предложен Йоавом Фройндом и Робертом Шапире. Обычно он используется в сочетании с другими алгоритмами классификации для улучшения их эффективности. Данный алгоритм называется адаптивным потому, что каждая следующая группа классификаторов строится по объектам, неверно классифицированным предыдущими группами. Достоинствами подобного метода являются высокая скорость обнаружения за счет использования каскадного классификатора, высокая точность обнаружения с низким процентом ложных срабатываний, в целом сравнимая с точностью обнаружения у гораздо более медленных алгоритмов. Кроме того, алгоритм подстраивается под проблемные элементы из обучающей выборки и теоретически метод достигает нулевой ошибки обучения за конечное число итераций. Но при этом данному алгоритму требуется большая обучающая выборка из позитивных и негативных изображений и большое время обучения.

#### 1.4. Выводы по главе 1

В данной главе был приведен обзор нескольких распространенных методов, используемых при решении задачи обнаружения объектов на изображении. При написании этой главы были описаны основные подзадачи, которые входят в общую задачу обнаружения объектов на изображении. Кроме того, были приведены существующие способы решения такого рода подзадач.

В процессе написания обзора были проанализированы разнообразные существующие методы, решающие поставленную задачу детектирования объекта на изображении. Это говорит о том, что данная задача имеет актуальность в современной теоретической информатике. В том числе было выявлено, что большая часть рассмотренных методов имеют схожую фундаментальную основу и отличаются только деталями реализации на различных этапах решения присущих каждому из них подзадач. Одной из целей составления приведенного обзора было выявление проблем, которые появляются при решении задачи обнаружения объектов на изображении.

Выделим основные проблемы, обнаруженные при составлении данного обзора:

- Проблема формирования глобального признакового описания объекта. Эта проблема имеет связь с выбранной системой признаков, используемой для единообразного и содержательного описания точек детектируемого объекта. Основная сложность при поиске оптимальной системы признаков – поиск компромисса между размером признакового описания и его информативностью.

Некоторые из популярных на сегодняшний день алгоритмов обычно формируют признаковые описания точек объекта, представленного как вектор чисел с плавающей точкой. Размерность такого вектора может быть 64, 128 и т.д. Такого рода дескрипторы используют существенный объем памяти для хранения, соответственно, при вычислении и дальнейшей обработке для них требуются высокие вычислительные затраты.

- Проблема принятия решения о локализации объекта. Эта проблема проявляется с детальным рассмотрением анализируемого изображения. Необходимые вычислительные ресурсы на этом этапе растут по мере увеличения углубления анализа изображения. При этом, выбранная ранее система признаков для формирования глобального признакового описания непосредственно сказывается на объеме времени, необходимом для вычислений. Это означает высокую степень важности правильного выбора системы признаков, которая позволила бы сохранять дескриптор искомого объекта в компактном, удобном для хранения виде.

На основании выявленных этапов и проанализированных имеющихся подходов к решению задачи обнаружения объектов на изображении, в данной работе предлагается новый метод, разработанный для решения данной задачи. Предлагаемый метод описывает собственную комбинацию подходов к решению указанных выше проблем. Теоретические основы предлагаемого в настоящей работе метода решения задачи детектирования объекта на изображении будут подробно описаны во второй главе.

## Глава 2. Информационная модель обнаружения объектов на изображении

Предлагаемый метод решения задачи обнаружения объекта на изображении комбинирует метод локализации объекта по цвету в цветовом пространстве HSV на этапе детектирования объекта с применением теории активного восприятия (ТАВ) [24] на этапе формирования глобального признакового описания выделенного объекта. На этапе принятия решения о распознанном на изображении объекте применяется алгоритм поиска наименьшего евклидова расстояния между данным глобальным признаковым описанием объекта и базой эталонных признаковых описаний известных объектов в пространстве признаков.

### 2.1 Общая структура системы распознавания объекта на изображении

Общий алгоритм работы системы следует разделить на два этапа:

1. Первичная инициализация системы для обучения и построения модели
2. Распознавание изображений с использованием построенной на первом этапе модели

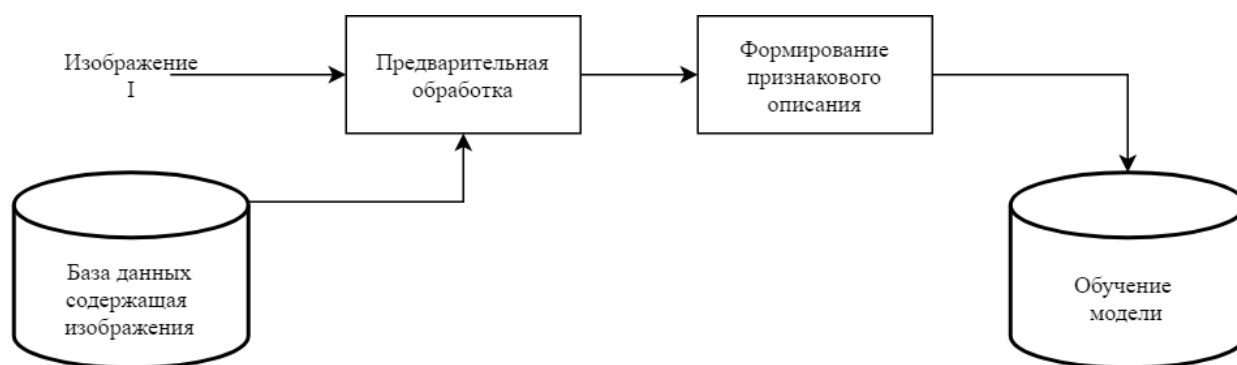


Рисунок 2.1. – Структура этапа обучения системы



Рисунок 2.2. – Структура этапа применения созданной системы

Дадим некоторые общие пояснения к каждому из обозначенных выше этапов. На первом этапе, под набором данных, содержащей изображения, подразумевается база данных, содержащая эталонные изображения объектов, которые в дальнейшем система будет распознавать. Целью данного этапа является получение при помощи теории активного восприятия глобального признакового описания для каждого из эталонных изображений, которое вычисляется после предварительной обработки каждого из входных изображений. Полученное глобальное признаковое описание сохраняется на диске для дальнейшего использования на втором этапе.

Второй этап подразумевает, что на вход системе подается изображение, содержащее объект, который необходимо распознать. Прежде всего, входное изображение, как и на первом этапе, проходит предварительную обработку. Затем происходит детектирование всех присутствующих на изображении объектов, после чего с каждого из обнаруженных детектором объектов при помощи теории активного восприятия снимается глобальное признаковое описание. Далее, полученное глобальное признаковое описание каждого из объектов поступает на блок принятия решения. Данный блок в своей работе использует полученные ранее на первом этапе дескрипторы изображений, содержащих эталонные объекты. Осуществляется принятие решения об объекте, присутствующем на данном изображении. Результатом принятия решения является одно из эталонных изображений, показывающее, какой объект был распознан на данном изображении.

В качестве распознаваемых объектов в данной работе будут использоваться дорожные знаки. Таким образом, база эталонов будет включать в себя непосредственно

изображения самих дорожных знаков, а база данных анализируемых изображений будет сформирована из изображений, покадрово полученных с видеорегистратора, находящегося в салоне автомобиля. Системе будет необходимо проанализировать входное изображение, выделить на нем область, в которой присутствует дорожный знак, а затем – распознать, какой именно знак представлен на изображении.

Перейдем к детальному описанию каждого из блоков, приведенных на информационной модели.

### 2.1.1. Предварительная обработка изображения

Пусть изображение искомого объекта закодировано согласно модели RGB. В этом случае каждый пиксель изображения представляется в виде числового вектора из трёх элементов, которые описывают яркости красного, зеленого и синего компонентов модели RGB соответственно. На основе такого описания пикселей задается функция яркости изображения:

$$I(x, y) = \frac{(R_{xy} + G_{xy} + B_{xy})}{3}, x \in X, y \in Y,$$

где  $R_{xy}, G_{xy}, B_{xy}$  – яркости соответствующих компонент модели RGB для пикселя с координатами  $(x, y)$ ;  $X, Y$  – множества соответствующих координат изображения.

Таким образом, значение функции яркости изображения для определенного пикселя равно среднему значению яркостей компонент модели RGB для данного пикселя. Тогда матрица значений яркости изображения  $I$  формируется как набор значений функции  $I(x, y)$  на всем множестве  $X$  и  $Y$ .

С целью устранения шумов на изображениях будем использовать фильтр Гаусса. Фильтр Гаусса – это фильтр размытия изображения, который использует нормальное распределение (также называемое Гауссовым распределением) для вычисления преобразования, применяемого к каждому пикселю изображения. Нормальное распределение для двух измерений описывается формулой [25]:

$$G(u, v) = \frac{1}{2\pi r^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}},$$

где  $r$  – радиус размытия,  $\sigma$  – стандартное отклонение распределения Гаусса.

Данная формула задает поверхность, имеющую вид концентрических окружностей с нормальным распределением от центральной точки. Пиксели, где распределение отлично от нуля используются для построения матрицы свертки, которая применяется к исходному изображению. Значение каждого пикселя становится средне взвешенным для окрестности. Исходное значение пикселя принимает наибольший вес (имеет наивысшее

Гауссово значение), и соседние пиксели принимают меньшие веса, в зависимости от расстояния до них. Пример работы размытия по Гауссу для одномерного массива можно увидеть на рисунке 2.3.

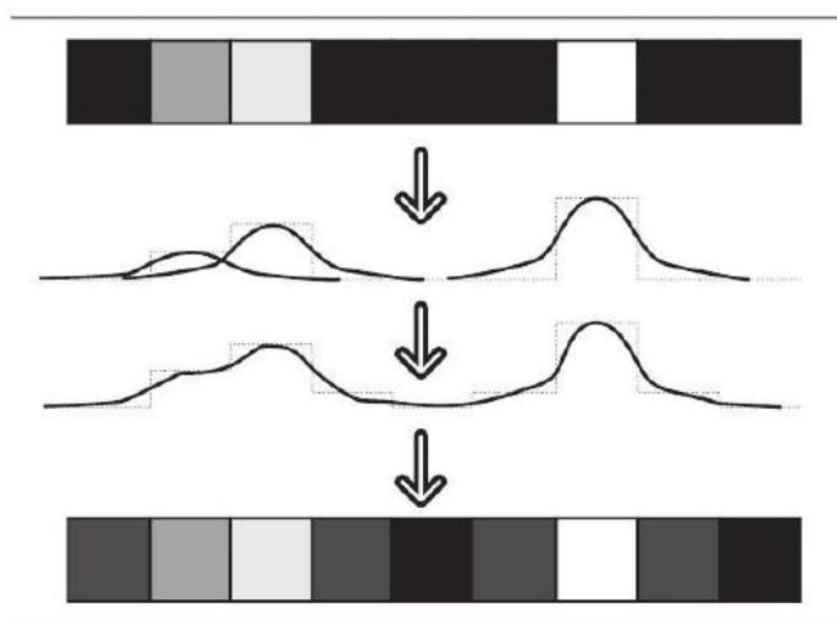


Рисунок 2.3. – Размытие по Гауссу на одномерном массиве

Таким образом, размытие по Гауссу позволит избавиться от нежелательных шумов на изображениях, что сведет к минимуму их влияние при дальнейшей классификации знаков. Результат применения фильтра Гаусса к целому изображению продемонстрирован на рисунке 2.4.

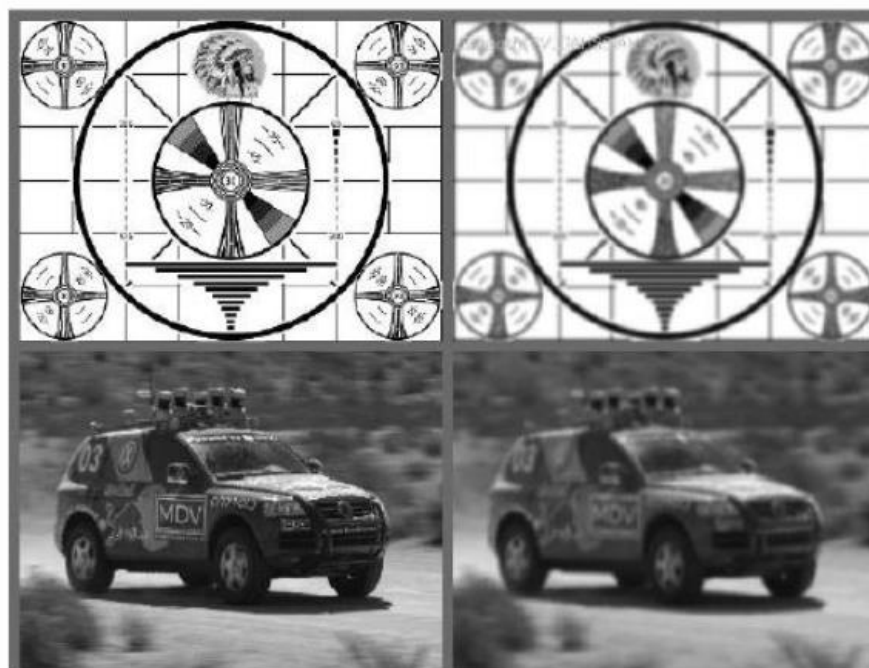


Рисунок 2.4. – Результат применения фильтра Гаусса

### 2.1.2. Детектирование объекта на изображении

Специфические условия освещённости значительно влияют на возможность корректного восприятия цвета объектов снимаемой сцены. При съёмке реальной дорожной ситуации на знаках возникает целый ряд различных условий освещения, сильно усложняющих детектирование объектов: прямые солнечные лучи, отражённый свет, затенения, свет автомобильных фар в тёмное время суток. Более того, различные искажающие эффекты могут присутствовать на одном дорожном знаке одновременно. Таким образом, на изображениях, получаемых с камеры, установленной в автомобиле, не всегда возможно выделить интересующую область по цвету с помощью простого применения порогового фильтра напрямую в цветовом пространстве RGB (Red, Green and Blue). Для извлечения красного, синего и желтого цвета из входного изображения необходимо использовать информацию о цвете каждого пикселя независимо от его неконтролируемого показателя освещённости. Для этих целей было выбрано цветовое пространство HSV (Hue, Saturation and Value). Цветовое пространство HSV принадлежит к группе цветовых пространств, изначально основанных на человеческом восприятии цветов. HSV в основном используется в области компьютерного зрения и компьютерной графики. Рисунок 2.5 иллюстрирует стандартное представление цветового пространства HSV в виде конуса. [26]

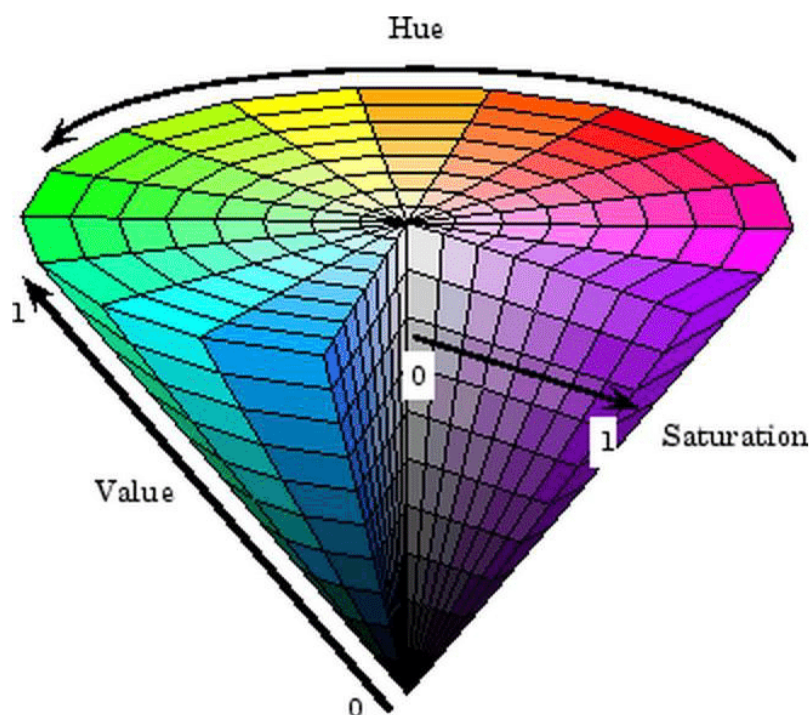


Рисунок 2.5. - Представление цветового пространства HSV



Оттенок (H) изменяется по окружности конуса, насыщенность (S) цвета возрастает с удалением от оси конуса, а яркость (V) – с приближением к его основанию. Различные оттенки и цвета определены при движении против часовой стрелки от 0 до 360 градусов. Насыщенность варьируется в пределах от 0 до 1 и представляет собой «чистоту цвета» – чем ближе этот параметр к 0, тем ближе цвет к нейтральному серому. V содержит информацию о яркости пикселя, принимает значения от 0 до 1.

### **2.1.3. Формирование глобального признакового описания изображения с помощью теории активного восприятия**

#### **2.1.3.1. Введение в теорию активного восприятия**

Теория активного восприятия (ТАВ) [24] была разработана Владимиром Александровичем Утробиним, профессором НГТУ им. Р.Е. Алексеева. Данная теория используется для формирования признакового описания и последующей обработки и анализа изображений.

Теория активного восприятия базируется на операции  $U$ -преобразования, применяемой к исследуемому изображению.  $U$ -преобразование заключается в последовательном применении к изображению операций интегрирования и дифференцирования.

В результате этого формируется набор спектральных коэффициентов, который и является признаковым описанием изображения. Также теория активного восприятия включает в себя так называемую «Алгебру групп». Данный раздел разработан для выявления зависимостей между спектральными коэффициентами, полученными в результате вычисления  $U$ -преобразования. Данные зависимости также могут быть использованы при формировании признакового описания изображения.

#### **2.1.3.2. $U$ -преобразование и $Q$ -преобразование**

Исследуемое изображение представлено в виде функции яркости  $I(x, y)$ ,  $x \in X, y \in Y$ ,  $X$  и  $Y$  - конечные множества. В результате применения к такому изображению  $U$ -преобразования вычисляется спектральное представление изображения:

$$D = \{d_i\}, i = \overline{1, M}$$

где  $M$  – число сегментов, на которые было разбито исследуемое изображение;  $d_i$  – спектральное представление сегмента  $i$ , включающее  $L$  спектральных коэффициентов ( $L$  – число фильтров, используемых на этапе дифференцирования);  $d_i\{k\}$  –  $k$ -й спектральный коэффициент представления сегмента  $i$ ,  $k = \overline{1, L}$ .

Формирование исходного описания с позиции теории активного восприятия изображения понимается как операция интегрирования, и включает в себя два следующих этапа:

- 1) Нормирование изображения
- 2) Вычисление матрицы «визуальных масс»

Нормирование изображения происходит путем вычитания из каждого элемента матрицы минимального элемента матрицы, а затем делением каждого элемента матрицы на максимальный элемент матрицы.

Для получения матрицы «визуальных масс» необходимо выполнить так называемое в теории активного восприятия Q-преобразование.

Для выполнения Q-преобразования необходимо разбить матрицу на блоки размером  $\frac{N}{4} * \frac{M}{4}$

Q-преобразование выполняется по формуле:

$$B_{kl} = \sum_{i=1}^2 \sum_{j=1}^2 A_{kl}[i, j]; k = \overline{1,4}; l = \overline{1,4}$$

0.1	0.1	0	0.3	0.1	0.1	0	0		0.6	0.3	0.6	0.7
0.2	0.2	0	0	0.2	0.2	0	0.7					
0	0	0.8	0.7	0	0.2	0.8	0.7		0.1	3.4	0.2	3.4
0.1	0	0.9	1	0	0	0.9	1					
0.1	0.1	0	0	0.1	0.1	0	0		0.6	0	0.6	0
0.2	0.2	0	0	0.2	0.2	0	0					
0	0	0.8	0.7	0	0	0.8	0.7		0.9	3.4	0	3.4
0	0.9	0.9	1	0	0	0.9	1					

Рисунок 2.6. – Q-преобразование

Формирование системы признаков с позиции теории активного восприятия изображения понимается как операция дифференцирования. Результатом операции дифференцирования является вектор «спектральных коэффициентов»  $\mu = (\mu_0, \mu_1, \mu_2, \dots, \mu_{15})$ . Операция дифференцирования реализуется по средствам 16 фильтров (рисунок 2.7), применяемых к матрице визуальных масс, где темный элемент означает умножение исходного элемента матрицы визуальных масс на -1, светлый элемент – умножение исходного элемента матрицы визуальных масс на +1. Все полученные в результате применения фильтра к матрице значения суммируются между собой, а результирующая сумма называется спектральным коэффициентом.

Рассмотрим пример вычисления  $U$ -преобразования для одного сегмента изображения с использованием 16 фильтров на этапе дифференцирования. Первым этапом вычисления  $U$ -преобразования является этап интегрирования ( $Q$ -преобразование). На данном этапе изображение (или его рассматриваемый сегмент) разбивается на 16 равных частей по числу фильтров. Яркости пикселей внутри каждой из полученных частей изображения складываются, таким образом, получается матрица визуальных масс. Визуализация данного этапа представлена на рисунке 2.6.

Далее выполняется этап дифференцирования. В ходе выполнения данного этапа матрица визуальных масс поэлементно умножается на фильтры, представленные на рисунке 2.7. Светлый элемент фильтра соответствует значению «1», темный элемент – значению «-1».

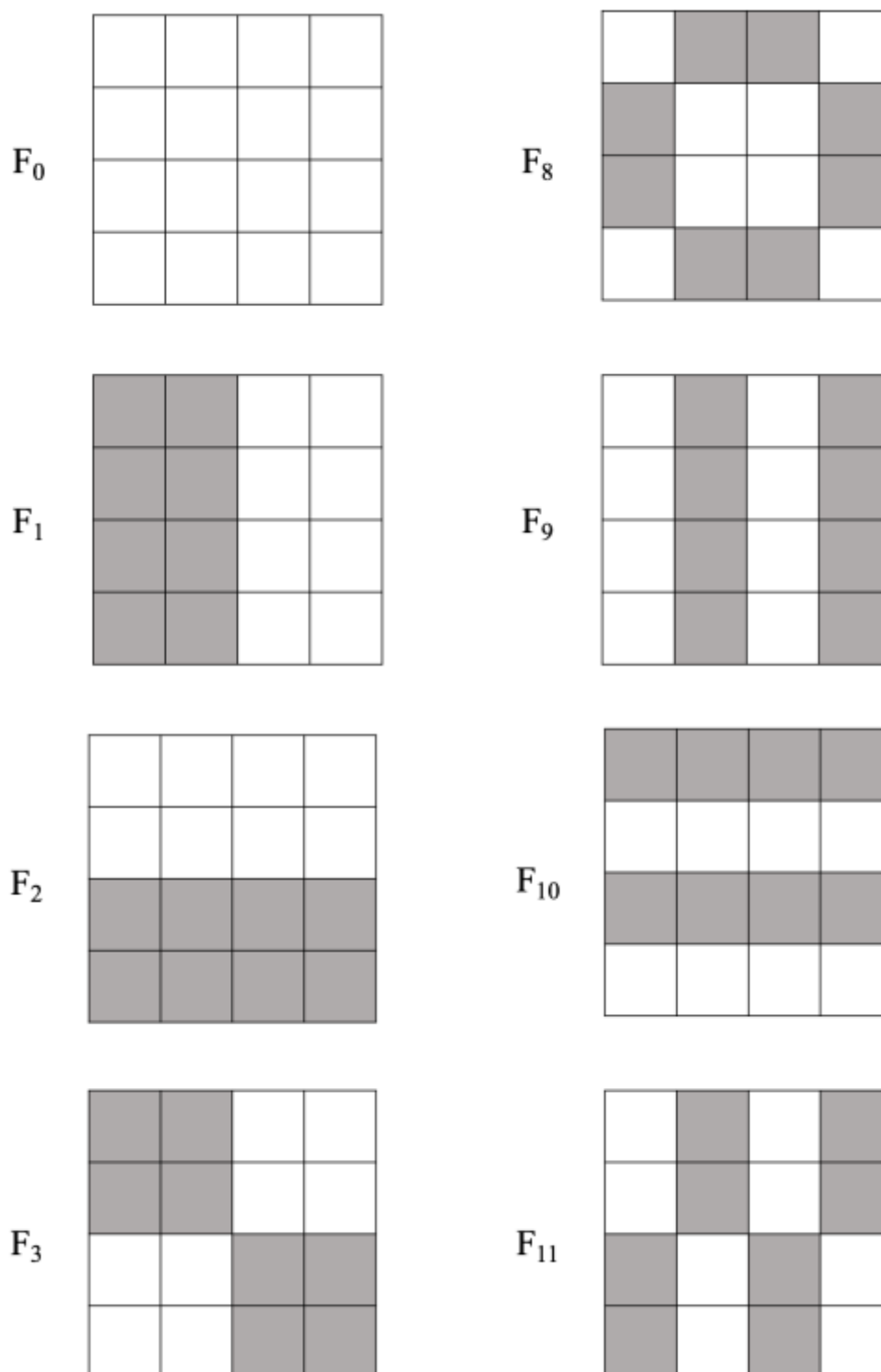


Рисунок 2.7. - Фильтры, использующиеся для вычисления  $U$   
преобразования

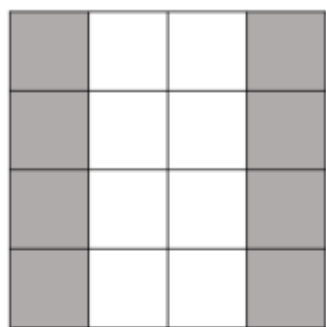
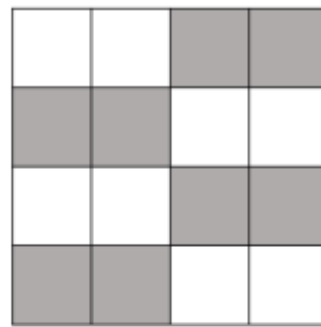
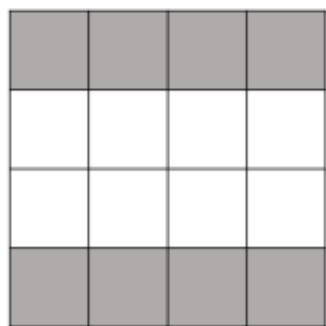
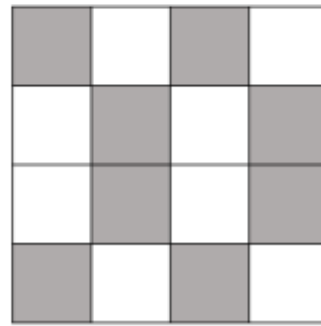
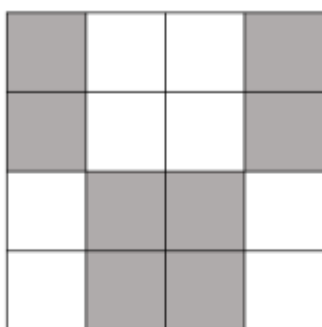
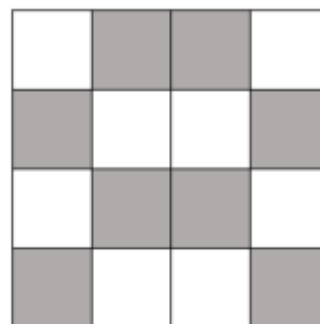
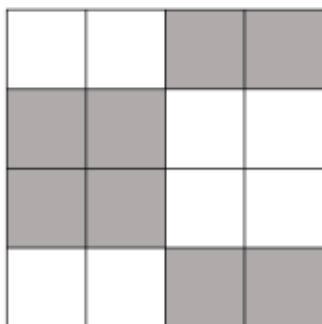
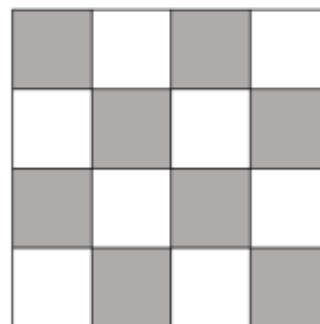
$F_4$  $F_{12}$  $F_5$  $F_{13}$  $F_6$  $F_{14}$  $F_7$  $F_{15}$ 

Рисунок 2.7. - Фильтры, использующиеся для вычисления  $U$   
преобразования

На рисунке 2.8 приведен пример получения спектрального коэффициента  $\mu_1$  после применения к матрице визуальных масс фильтра  $F_1$ .

$$\Sigma \left[ \begin{array}{c} \text{Матрица} \\ \text{«визуальных масс»} \end{array} \times \begin{array}{c} \text{Фильтр } F_1 \end{array} \right] = -0.4$$

0.6	0.3	0.6	0.7
0.1	3.4	0.2	3.4
0.6	0	0.6	0
0.9	3.4	0	3.4

-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1

Рисунок 2.8. – Получение спектрального коэффициента  $\mu_1$

Аналогично примеру с фильтром  $F_1$ , рассчитываются все 16 спектральных коэффициентов. После выполнения этой операции будет получен вектор из 16 спектральных коэффициентов – данный вектор является результатом U-преобразования. Пример такого вектора приведен на рисунке 2.9. Данные спектральные коэффициенты могут быть использованы как глобальное признаковое описание проанализированного сегмента изображения.

$$\mu =$$

$F_0$	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$	$F_6$	$F_7$	$F_8$	$F_9$	$F_{10}$	$F_{11}$	$F_{12}$	$F_{13}$	$F_{14}$	$F_{15}$
18.2	-0.4	0.4	1.4	-1.2	-1.6	0.6	0.6	1.4	-11	11.4	-1.6	-1.2	0.4	-0.4	-13.8

Рисунок 2.9. – Пример вектора спектральных коэффициентов

### 2.1.3.3. Алгебра групп

Алгебра групп – раздел теории активного восприятия, предназначенный для выявления зависимостей между спектральными коэффициентами, полученными на этапе вычисления U-преобразования. Для этого в алгебре групп вводятся понятия операторов, полных и замкнутых групп.

Множество бинарных операторов  $\{V_i\}$  вычисляется на основе множества фильтров  $\{F_i\}$ , представленных на рисунке 2.7. Для получения оператора на основе фильтра необходимо поставить следующее соответствие между значениями элементов фильтра и оператора:  $(+1 \rightarrow 1), (-1 \rightarrow 0)$ . Пример вычисления бинарного оператора  $V_3$  на основе фильтра  $F_3$  приведен на рисунке 2.10.

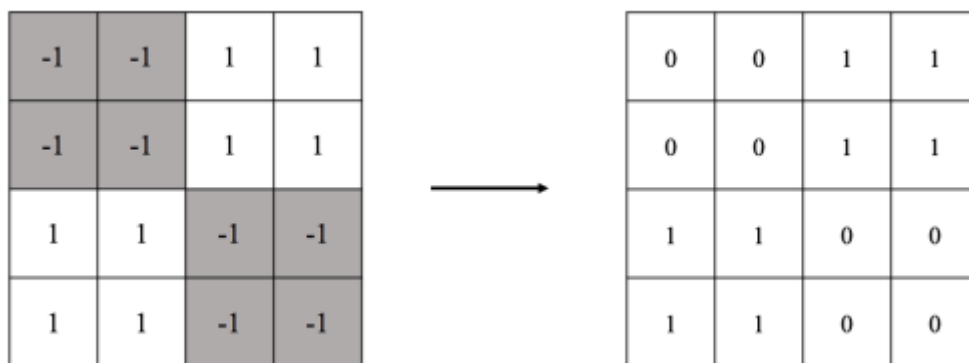


Рисунок 2.10. – Вычисление бинарного оператора  $V_3$  на основе фильтра  $F_3$

На основании полученных операторов формируются множества полных и замкнутых групп:

- Множество полных групп  $P_n = \{P_{ni}\}$ ,  $P_{ni} = \{V_i, V_j, V_k\}$  мощностью 35;
- Множество замкнутых групп  $P_s = \{P_{si}\}$ ,  $P_{si} = \{V_i, V_j, V_k, V_r\}$  мощностью 105.

Количество существующих полных и замкнутых групп может быть расширено за счет использования в их составе инверсных операторов. Таким образом, общее количество полных групп – 140, замкнутых групп – 840. Полные и замкнутые группы позволяют произвести спектрально-корреляционный анализ изображения. Полные группы выявляют корреляционные связи между операторами, замкнутые группы – связи между полными группами. Иерархия операторов, полных и замкнутых групп представлена на рисунке 2.11.

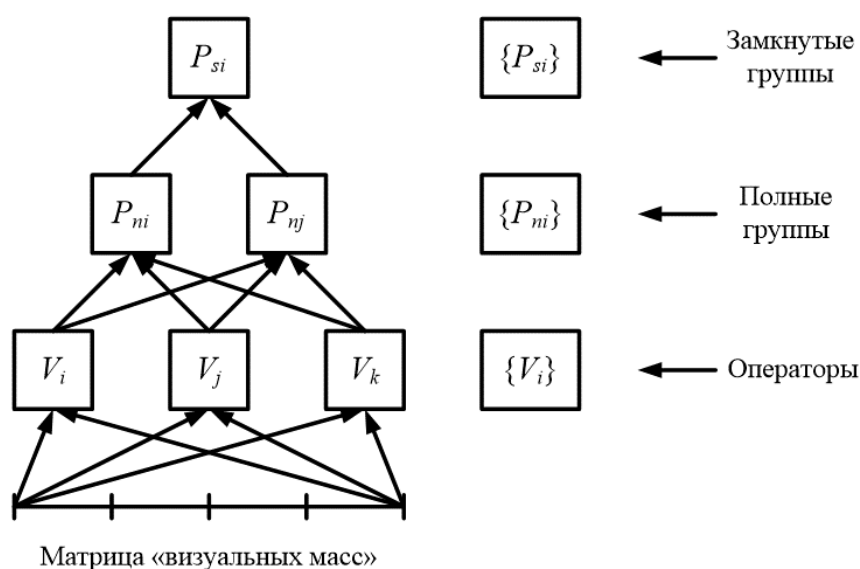


Рисунок 2.11. – Иерархическая структура операторов, полных и замкнутых групп

Каждому оператору или группе ставится в соответствие понятие образа оператора или группы. Образ оператора  $V_3$  представлен на рисунке 2.10. На основании подобных образов происходит вычисление масс операторов, полных и замкнутых групп. Пример вычисления массы замкнутой группы по сегменту изображения представлен на рисунок 2.12.

$$\Sigma \left( \begin{array}{c} \text{Матрица визуальных масс} \\ \begin{array}{|c|c|c|c|} \hline 1209 & 1446 & 1027 & 741 \\ \hline 1786 & 2686 & 1955 & 1273 \\ \hline 426 & 688 & 500 & 317 \\ \hline 288 & 365 & 281 & 243 \\ \hline \end{array} \end{array} \right) \times \left( \begin{array}{c} \text{Образ группы } V_1 V_3 V_7 \overline{V_{12}} \\ \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 \\ \hline \end{array} \end{array} \right) = 6466$$

Рисунок 2.12. – Вычисление массы замкнутой группы по сегменту изображения

Таким образом, вычисляются массы операторов, полных и замкнутых групп, полученных на основе спектрального представления исследуемого изображения. Информация о наличии тех или иных полных и замкнутых групп в спектре изображения или его сегмента, а также массы данных групп являются базисом, на котором формируется еще один вид признакового описания изображения, предлагаемый теорией активного восприятия.

Данный вид признакового описания является более сложным и массивным по сравнению с спектральным представлением изображения, полученным в результате  $U$ -преобразования. Однако признаковое описание на основе алгебры групп несет в себе больше информации, в том числе о корреляционных связях между спектральными коэффициентами, что позволяет иметь более точное описание исходного изображения.

#### 2.1.4. Принятие решения

Этап принятия решения заключается в получении соответствия между данным объектом и одним из эталонных объектов, известных системе. С этой целью, в пространство признаков, полученное по глобальному признаковому описанию объектов-эталонов, внесем глобальное признаковое описание объекта, полученное после его обнаружения на изображении. Далее будем искать ближайшее признаковое описание



эталонного объекта к данному объекту. В качестве меры близости будем использовать евклидово расстояние, вычисляемое по формуле:

$$(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}$$

Таким образом, необходимо будет провести вычисления для каждого эталонного объекта. Это означает, что количество вычислений по формуле равно количеству эталонных объектов в системе. Другими словами, вычислительная сложность этапа принятия решений равняется  $O(n)$ , однако, среди его преимуществ – простота реализации и хорошая точность.

В результате вычисления евклидова расстояния от глобального признакового описания данного объекта до глобального признакового описания каждого из эталонных объектов получим числовой вектор, в котором нам нужно найти минимальный элемент, представляющий собой минимальное евклидово расстояние. Этот элемент указывает на эталонный объект, который оказался в признаковом пространстве ближе всего к данному объекту. Технически это реализовано через вектор экземпляров класса, объекты которого имеют два атрибута: глобальное признаковое описание объекта и само изображение объекта. Таким образом, обнаружив в векторе расстояний минимальный элемент, не составляет труда найти как само изображение эталонного объекта, так и его глобальное признаковое описание, после чего можно вывести его на экран или проверить на предмет ошибки.

## 2.2. Выводы по Главе 2

В данной главе были рассмотрены теоретические основы метода обнаружения объекта на изображении на основе теории активного восприятия. Была предложена информационная модель создаваемой системы и были предложены варианты реализации каждого из этапов решения задачи обнаружения объекта на изображении:

- Предварительная обработка изображения – нормализация, вычисление функции яркости, фильтр Гаусса.
- Формирование глобального признакового описания объекта – теория активного восприятия ( $U$ -преобразование).
- Принятие решения о локализации объекта – минимальное евклидово расстояние от данного объекта до эталонных объектов.

Использование данного набора вариантов реализации каждого из этапов решения задачи обнаружения объекта на изображении свидетельствует о научной новизне предлагаемого метода. Экспериментальные данные о результатах работы данного метода на различных наборах исходных данных будут подробно описаны в главе 3.

### Глава 3. Вычислительный эксперимент

Вычислительный эксперимент – это обязательная часть исследования по разработке метода решения задачи обнаружения объектов на изображении. Анализ результатов вычислительного эксперимента дает возможность построить выводы о том, насколько успешно предлагаемый в настоящей работе метод справляется с решением задачи по сравнению с уже известными аналогами, необходимы ли предлагаемому методу доработки, имеется ли смысл применять данный метод на практике и т.д. Данная глава посвящена описанию проведенного вычислительного эксперимента и анализу его результатов.

#### 3.1. Описание эксперимента

Для проведения эксперимента была сформирована база изображений объектов. Как было отмечено выше, в роли объектов выступают дорожные знаки. В базе содержатся изображения 156 различных дорожных знаков, в качестве которых взяты наиболее основные и часто встречающиеся в реальных дорожных условиях. Каждый объект в базе имеет четыре шаблона, которые используются непосредственно при обнаружении данного объекта на изображении. Каждый из такого рода шаблонов отличается от остальных погодными условиями и условиями освещения при которых проведена съемка объекта. Рассмотрим их подробнее:

1. Изображение объекта при благоприятной погоде и хорошем освещении
2. Изображение объекта при неблагоприятной погоде и хорошем освещении
3. Изображение объекта при благоприятной погоде и плохом освещении
4. Изображение объекта при неблагоприятной погоде и плохом освещении

В данном списке, под термином «благоприятная погода» подразумеваются такие климатические условия съемки, при которых на изображении отсутствует дождь, снег, град, туман, песочная буря и прочие факторы, искажающие изображение, которые по сути можно классифицировать как естественный шум.

Напротив, под термином «неблагоприятная погода» подразумевается наличие любого из перечисленных выше климатических явлений, или любой их комбинации на изображении. Хорошее освещение, как правило, присуще утреннему и дневному времени суток, а плохое – вечернему и ночному времени суток.

Рисунок 3.1 иллюстрирует все 4 шаблона знака «Пешеходный переход» в соответствии с приведенными описаниями каждой из категории шаблонов:



Рисунок 3.1. – Пример четырех шаблонов знака «Пешеходный переход»

Для проведения вычислительного эксперимента была использована Российская база изображений автодорожных знаков [27], известная так же как Russian Traffic Sign Dataset (RTSD). Примеры кадров из этой базы представлены на рисунке 3.2. Данная база содержит:

- 179138 размеченных кадров с авторегистратора, содержащих дорожные знаки
- 156 типов дорожных знаков
- 104358 изображений знаков (полученных путем кадрирования)

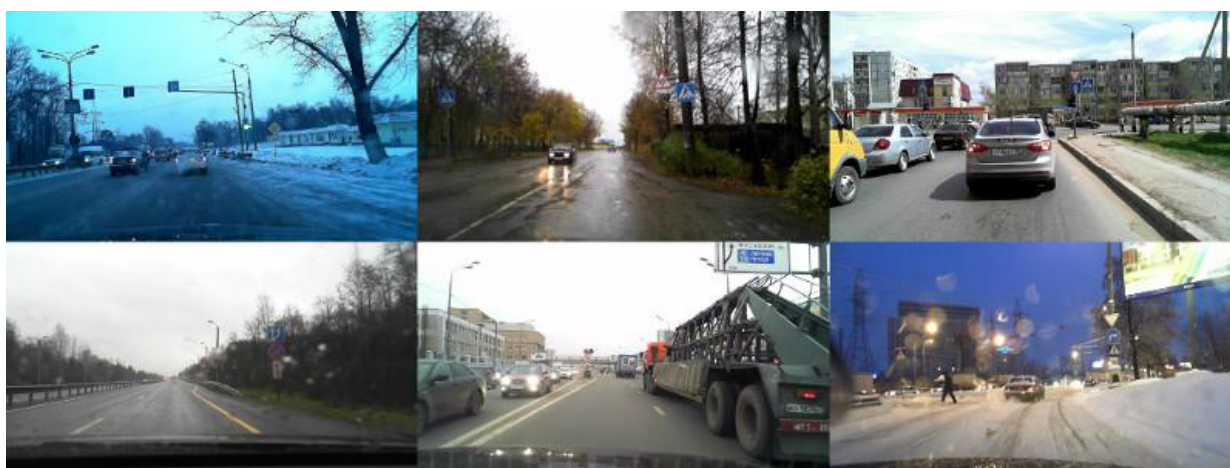


Рисунок 3.2. – Примеры кадров из базы изображений RTSD

### 3.2. Описание программного продукта

Для проведения описанного выше вычислительного эксперимента был разработан программный продукт на языке программирования Python. Данный язык был выбран, так как он является кроссплатформенным, обладает широким инструментарием для обработки статистических данных и работы с графикой, а также имеет высокую степень расширяемости с помощью пакетов.

#### 3.2.1. Калибровка цветовой маски

В первую очередь, следует описать модуль калибровки цвета для локализации объекта на изображении. В данной работе распознавание объектов рассматривается в

контексте распознавания дорожных знаков, следовательно, калибровать систему следует на обнаружение синего, красного и желтого цветов.

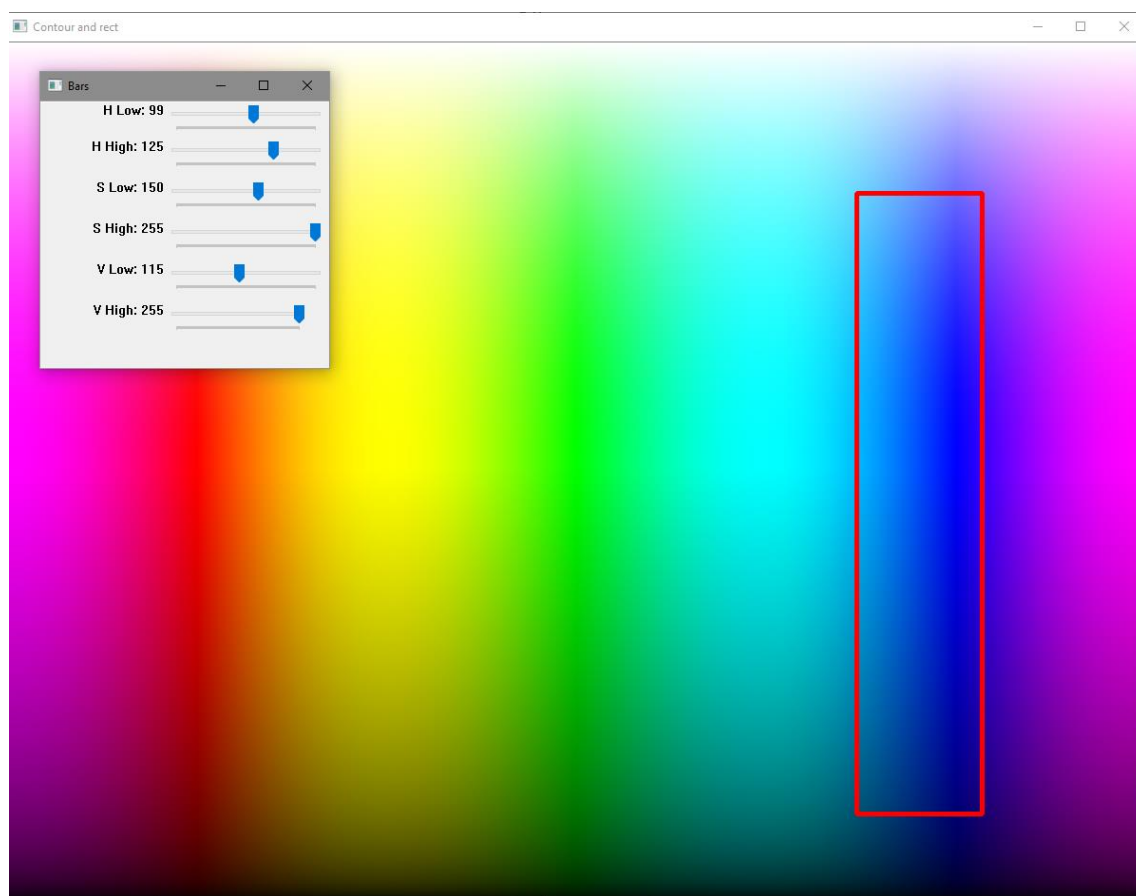


Рисунок 3.3. – Интерфейс модуля калибровки цвета

На рисунке 3.3 представлен интерфейс модуля калибровки цвета, настройками которого заданы границы допуски для синего цвета. Данный интерфейс включает в себя два окна – окно с цветовой палитрой, на котором динамически перемещается заданная область, выделяющая интересующий цветовой диапазон в пространстве HSV, и окно регуляторов, с помощью которых этот динамический цветовой диапазон задается. Окно регуляторов включает в себя следующие бегунки, значения которых предназначены для редактирования:

- H Low – регулятор нижнего порога Hue-компоненты внутри цветового пространства HSV
- H High– регулятор верхнего порога Hue-компоненты внутри цветового пространства HSV
- S Low– регулятор нижнего порога Saturation-компоненты внутри цветового пространства HSV
- S High– регулятор верхнего порога Saturation-компоненты внутри цветового пространства HSV

- V Low– регулятор нижнего порога Value-компоненты внутри цветового пространства HSV
- V High– регулятор верхнего порога Value-компоненты внутри цветового пространства HSV

Регулируя значения бегунков, исследователь видит диапазон выделяемого цвета на изображении цветовой палитры. Таким образом, функционал данного модуля необходим для обеспечения удобного способа калибровки цветового диапазона локализуемых объектов для нужд исследователя. Из рисунка 3.4 видно, что диапазон синего цвета составляет 99-125 для H-компоненты, 155-255 для S-компоненты, 115-255 для V-компоненты. Рассмотрим также процесс калибровки желтого цвета:

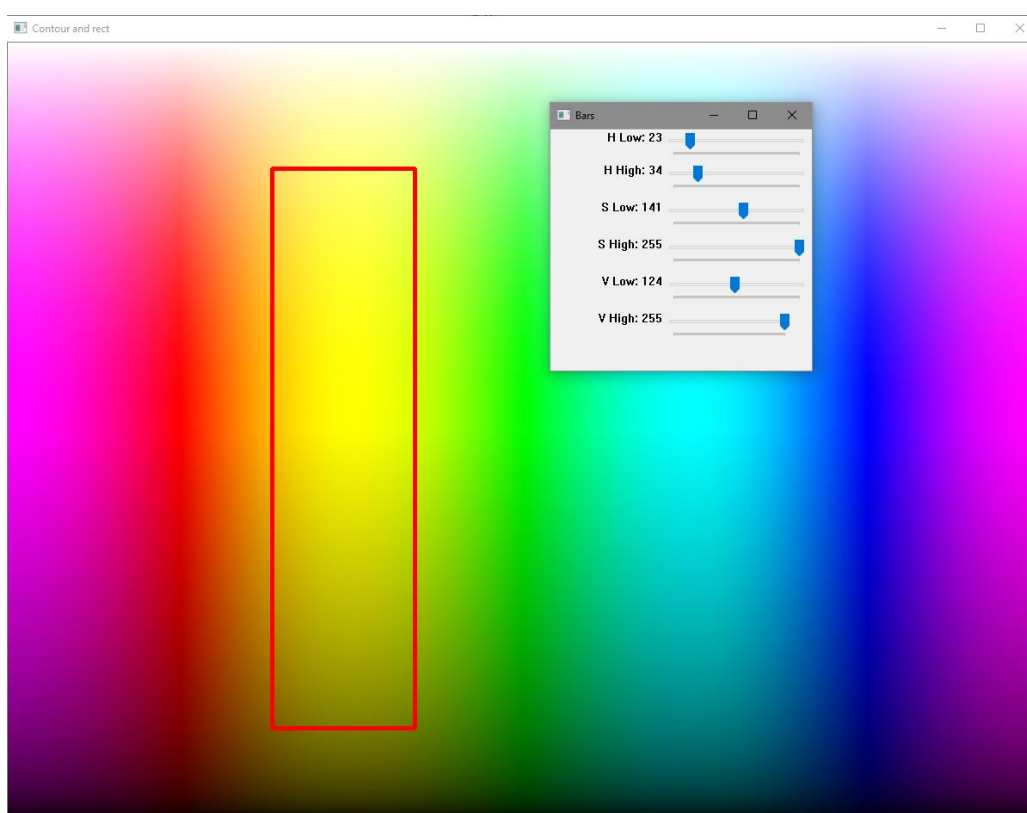


Рисунок 3.4. – Калибровка диапазона желтого цвета

Из рис. 3.4. видно, что диапазон значений для желтого цвета составляет 23-34 для H-компоненты, 141-255 для S-компоненты, 124-255 для V-компоненты.

Для красного цвета диапазон значений складывается из двух диапазонов. Это обусловлено тем, что H-компонента в цветовом пространстве HSV при нулевом значении означает ярко-красный цвет (рисунок 3.5), а для работы нам необходимо иметь как яркие, так и темные оттенки красного цвета внутри рассматриваемого диапазона. Поэтому было необходимо составить диапазон красного цвета, объединив два диапазона: первый диапазон значений для красного цвета составляет 0-10 для H-компоненты, 141-255 для S-

компоненты, 124-255 для V-компоненты, второй диапазон значений для красного цвета составляет 170-180 для H-компоненты, 141-255 для S-компоненты, 124-255 для V-компоненты.

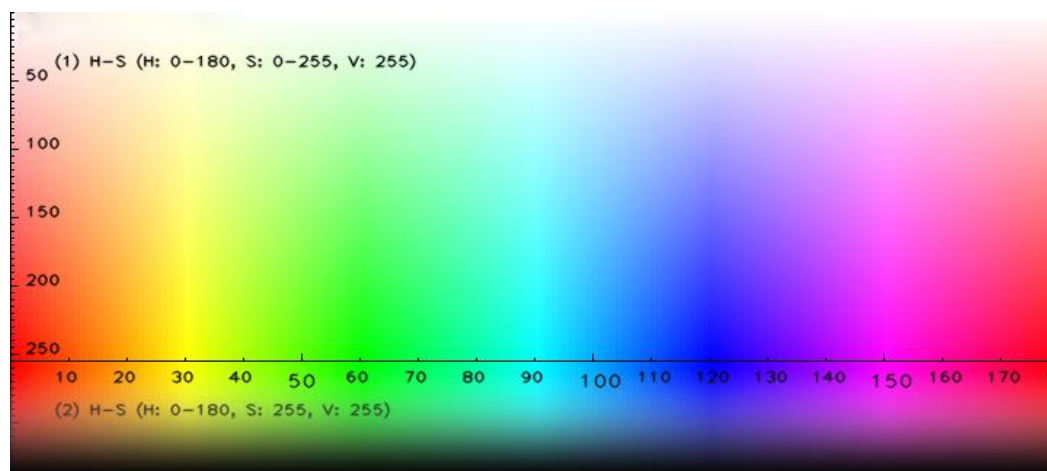


Рисунок 3.5. – Размеченный цветовой диапазон пространства HSV

После определения граничных значений интересующих цветовых диапазонов, сохраним конфигурацию каждого из них и для проверки выведем на экран комбинацию всех заданных цветовых диапазонов, отображенную на цветовой палитре:

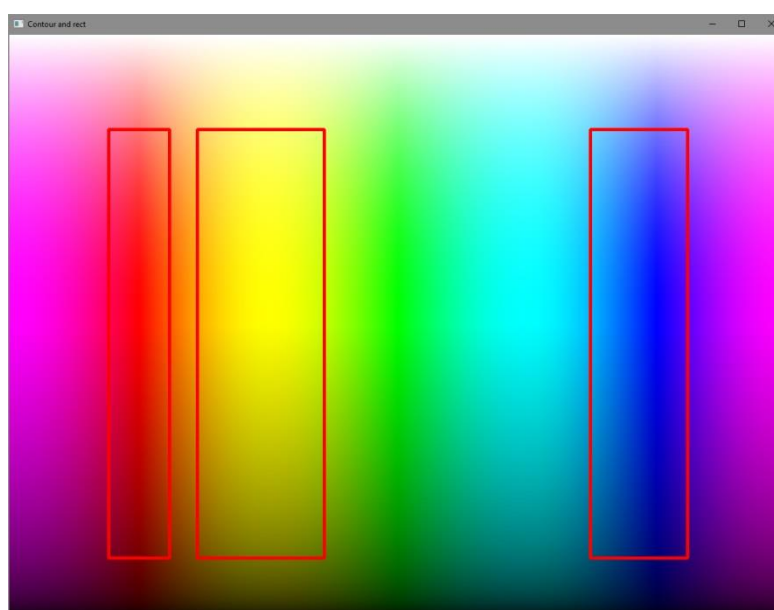


Рисунок 3.6. – Выделение полной комбинации целевых цветов на палитре

Как видно из рисунка 3.6, сохраненные конфигурации не искажаются при комбинировании и верно определяются на цветовой палитре. Следовательно, данные конфигурации пригодны к дальнейшему использованию с целью локализации дорожных знаков на изображении. Заданные конфигурации сохраняются и используются в дальнейшем либо сбрасываются и задаются каждый раз заново в зависимости от рода проводимого эксперимента на усмотрение исследователя.



### 3.2.2. Предварительная обработка входного изображения

Цель предварительной обработки изображения – избавление от шумов и сглаживание цветов на изображении. Данная операция позволяет значительно повысить точность на этапе локализации.

В первую очередь, к изображению применяется фильтр Гаусса, он позволяет избавиться от большинства шумов. Хотя это и приводит к некоторому размытию изображения, цветовые диапазоны сглаживаются и выравниваются – для этапа локализации эта операция существенно повышает точность детектирования. Параметры для фильтра Гаусса были подобраны эмпирически и представляют собой ядро размером 5 на 5 пикселей, применяемое к изображению в три последовательных прохода.

Следующим этапом предварительной обработки является наложение маски, фильтрующий заданный цветовой диапазон, который был получен с помощью модуля калибровки цветового пространства HSV ранее. Однако, в большинстве случаев применения только этих методов недостаточно – в нашем случае, наложение маски приведет не только к выделению синих, красных и желтых знаков на изображении, но так же и к выделению других регионов изображения, содержащих цвета заданного цветового диапазона. Таким образом, нам необходимо снизить погрешность модуля локализации путем устранения нежелательных регионов, выделенных цветовой фильтрацией по заданной маске, но так, чтобы случайно не отсеять интересующие нас регионы изображения, содержащие дорожные знаки.

В настоящей работе с этой целью применяются морфологические операции над выделенными регионами. Библиотека OpenCV, используемая для работы с изображениями в программной реализации системы, включает в себя ряд методов для выполнения морфологических преобразований, среди которых наиболее подходящими для нас будут методы Dilation и Erosion. Рассмотрим рисунок 3.7, на котором приведены три изображения: а) – оригинальный контур, полученный применением маски к изображению; б) – маска, к которой был применен метод dilate(); в) – маска, к которой был применен метод erode(). Таким образом, метод dilate() служит для увеличения, «наращивания» контура вокруг выделенного маской региона, а метод erode() – напротив, служит для уменьшения, «сужения» контура вокруг выделенного маской региона. В настоящей работе для устранения не интересующих нас регионов, будем применять сначала метод erode() в два последовательных прохода, а затем – dilate() в четыре последовательных прохода.

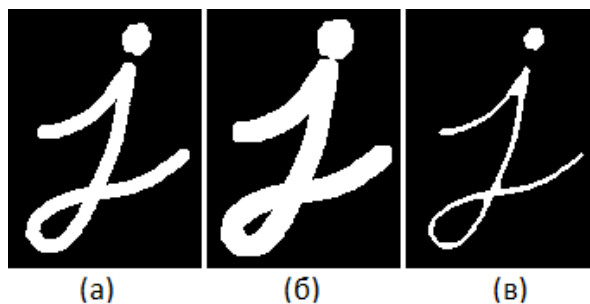


Рисунок 3.7. – Результат применения морфологических преобразований к изображению

После применения морфологических операций к изображению, выделим контурами все найденные области, и перейдем к их дальнейшей обработке.

### 3.2.3. Постобработка результатов локализации

Рассмотрим изображения до и после предварительной обработки. На изображении, прошедшем предварительную обработку, выделим контрастным цветом обнаруженные контуры. Изображения представлены на рисунке 3.8.



Рисунок 3.8 – Изображение до и после предварительной обработки



Как видно из изображения, прошедшего предварительную обработку, контурами были выделены интересующие нас ближайшие к камере дорожные знаки, но кроме этого есть и обведенные контурами нежелательные артефакты, от которых необходимо избавиться. В результате анализа возникающих артефактов на серии изображений, было выявлено, что в большинстве случаев площадь внутри контура артефактов значительно меньше площади внутри контура, окружающего интересующий нас регион на изображении, содержащий дорожный знак. Таким образом, первым этапом постобработки результатов локализации будет фильтрация контуров по величине описываемой ими площади.

Так как контур содержит множество точек, количество которых для каждого выделенного региона неодинаково, было решено пред фильтрацией каждый из контуров вписать в прямоугольник, для чего была использована функция `boundingRect()` из библиотеки `OpenCV`, позволяющая вписать данный контур в прямоугольник и получить координаты данного прямоугольника. Имея теперь уже прямоугольник, не составляет труда подсчитать его площадь и провести фильтрацию. Экспериментальным путем было подобрано подходящее значение пороговой площади прямоугольника, дающее наилучший результат по соотношению количества правильно отфильтрованных артефактов к количеству неверно отфильтрованных регионов, равное 1000 пикселей.

Кроме фильтрации, в рамках проводимого исследования на этапе постобработки результатов локализации к изображению применялся алгоритм Рамера-Дугласа-Пекера, который позволяет уменьшать число точек на кривой, которая аппроксимирована большей выборкой точек. Применение этого алгоритма позволяет фильтровать контуры по форме самих контуров. Таким образом, вначале задается ряд контуров, описывающих интересующие нас формы контуров, повторяющие формы дорожных знаков – круг, квадрат, треугольник. Затем, с помощью указанного алгоритма каждый из обнаруженных на изображении контуров аппроксимируется и проверяется на схожесть с одним из шаблонных контуров. Если наблюдается совпадение с одним из шаблонов, контур помечается как искомый и остается в выборке контуров, если совпадение ни с одним из шаблонов не наблюдается – контур помечается как не искомый и удаляется из выборки контуров. По результатам тестирования системы с включением в её функционал данного метода и исключением данного метода, были получены отличающиеся между собой результаты.

Анализ результатов показал, что внедрение данной функциональности отрицательно сказывается на скорости работы системы (время работы увеличивается на 5-10%), а точность локализации повышается (на 3-5%). По результатам анализа было

решено оставить реализацию данного функционала в системе, но сделать его включение опциональным на усмотрение исследователя.

Также в рамках исследования проводились эксперименты с изменением площади локализованных регионов. Предполагалось, что расширение границ вокруг прямоугольника, выделяющего дорожный знак на изображении, на 5-10% могло бы привести к улучшению точности на этапе распознавания и принятия решения. Была реализована функциональность, расширяющая границы каждого итогового прямоугольника на заданную в процентах величину. Таким же образом, как и с применением алгоритма Рамера-Дугласа-Пекера, был проведен ряд тестов с включением и отключением данной экспериментальной функциональности на серии изображений.

Анализ результатов показал, что даже незначительно увеличение площади прямоугольника, описывающего дорожный знак на изображении, например, на 5%, приводит к ухудшению точности распознавания и принятия решения на 10-20%. Дальнейшее увеличение площади прямоугольника еще сильнее влияет на точность и ухудшает её. На время работы системы данная функциональность сколь-нибудь существенного влияния не оказывает. По результатам анализа было принято решение исключить данную функциональность из разрабатываемой системы.

#### **3.2.4. Классификация обнаруженных объектов**

Рассмотрим детали процесса классификации обнаруженных знаков. Как было отмечено ранее, в базе имеется по 4 шаблона каждого из 156 дорожных знаков, взятых из Российской базы изображений автодорожных знаков.

Первый этап классификации является предварительным и проводится всего один раз, но без этого этапа никакая дальнейшая классификация не предоставляется возможной. Речь идет о получении глобального признакового описания со всех изображений-эталонов и сохранении этого описания в виде текстового документа на диске. Кроме самого описания, указывается так же номер класса, к которому принадлежит дорожный знак. Это необходимо для модуля оценки точности классификации и будет рассмотрено далее.

Российская база изображений автодорожных знаков, кроме самих изображений, включает также .csv-файл, в котором указаны координаты прямоугольников с дорожными знаками на предоставляемых изображениях, а так же номер класса для каждого знака. Именно в соответствии с этим номером класса были проставлены номера классов на предварительном этапе классификации, когда выполнялось сохранение глобального признакового описания изображений-шаблонов.

Далее будет рассмотрен основной этап классификации, выполняемый после того, как завершен этап локализации знаков на изображении и получены координаты регионов, содержащих локализованные знаки. Для каждого выделенного модулем локализации региона изображения вычисляется глобальное признаковое описание, а затем оценивается евклидово расстояние от данного описания до всех эталонных описаний, после чего выбирается такой элемент из массива эталонных дескрипторов, расстояние от которого до текущего анализируемого дескриптора минимально. Такой элемент указывает на эталон, который был распознан на данном изображении.

Для сопоставления исследуемого региона и эталонного изображения, используется контроль по определению принадлежности к одному из пронумерованных классов, а также контроль области, в которой была произведена локализация. Для такого контроля используются имеющиеся в Российской базе изображений автодорожных знаков классы и координаты регионов, в которых данные знаки расположены на анализируемых изображениях. В случае с контролем координат имеется допуск в 50 пикселей вокруг заданной области, таким образом, если локализатор допустит погрешность в координатах региона, содержащего дорожный знак, не превышающую 50 пикселей от истинного значения, проверка считается пройденной, а результат локализации – верным.

Для классификатора подобного допуска не предусмотрено – класс знака должен быть определен точно. Кроме этого, в системе имеется функционал, выводящий на экран само эталонное изображение распознанного знака в отдельном окне – данный функционал активно использовался при отладке и тестировании системы. Пример его использования приведен на рисунке 3.9.

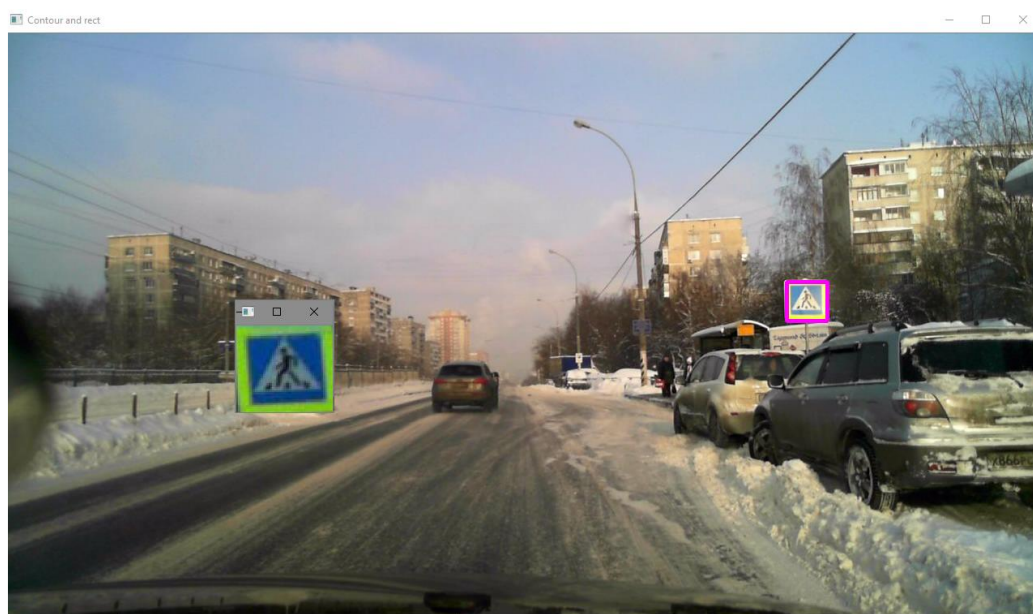


Рисунок 3.9. – Режим визуализации распознанного шаблона

Таким образом, осуществляется двойной контроль системы: как локализации, так и классификации.

### 3.3. Постановка вычислительного эксперимента

Для проведения вычислительного эксперимента из используемой базы изображений было случайным образом отобрано 10000 изображений, сделанных в разное время суток, при разной погоде, в разные сезоны года. Таким образом, выборка содержит как изображения, сделанные почти в идеальных условиях, когда знак отчетливо виден на изображении и ярко освещен, так и изображения, содержащие шумы и искажения естественного происхождения: блики и тени на дорожных знаках, темное время суток, плохие погодные условия, например, снег или дождь, а так же различного рода помехи, закрывающие собой части дорожных знаков: наледь, ветки деревьев, столбы, провода и тому подобное. Совокупность этих категорий изображений назовем выборкой изображений, сделанных в нормальных условиях.

Кроме того, для усложнения задачи и внесения дополнительной объективности в проведение эксперимента и исследование предлагаемого подхода, искусственно синтезируем еще две выборки изображений, взяв за основу выборку в нормальных условиях. Первая синтезированная выборка будет состоять из изображений, на которые дополнительно был добавлен шум различной степени. Вторая выборка будет включать в себя изображения, дорожные знаки на которых будут повернуты под различными углами. Каждая из выборок будет включать в себя 10000 изображений. Таким образом, итоговые выборки для проведения вычислительного эксперимента будут сформированы следующим образом:

- 1) Выборка изображений, сделанных в нормальных условиях
- 2) Выборка изображений с добавлением шума различной интенсивности:
  - а. ОСШ = 20дБ
  - б. ОСШ = 10дБ
  - с. ОСШ = 0дБ
- 3) Выборка изображений с поворотом искомых объектов:
  - а. Угол поворота = 15 градусов
  - б. Угол поворота = 30 градусов
  - с. Угол поворота = 45 градусов

Кроме того, будем замерять среднее время обработки одного изображения в системе, а так же проводить исследования с включением и отключением экспериментальной функциональности.

Так как оценивается и работа локализатора, и работа классификатора, рассмотрим возможные варианты результата обработки изображения отдельно для локализатора и отдельно для классификатора.

Для локализатора:

- 1) Объект был верно локализован на изображении, где он присутствует
- 2) Объект не был локализован на изображении, где он отсутствует
- 3) Объект был локализован на изображении в позиции, где он отсутствует
- 4) Объект не был локализован на изображении, где он присутствует

Для классификатора:

- 1) Объект был отнесен к классу, к которому он действительно принадлежит
- 2) Объект не был отнесен к классу, к которому он действительно не принадлежит
- 3) Объект был отнесен к классу, к которому он действительно не принадлежит
- 4) Объект не был отнесен к классу, к которому он действительно принадлежит

Как для локализатора, так и для классификатора, варианты 1 и 2 соответствуют корректной работе алгоритма. Варианты 3 и 4 соответствуют ошибкам 1-го (ложноположительное срабатывание) и 2-го рода (ложноотрицательное срабатывание) соответственно.

Оценка точности работы системы будет проводиться следующим образом: в первую очередь, будет оцениваться точность результатов работы локализатора с различными параметрами, без дальнейшей классификации. Цель – выявить такой набор параметров локализатора, при которых точность локализации максимальна. В дальнейшем, такой набор параметров локализатора будем называться оптимальными параметрами локализации. Далее будет проводиться оценка точности классификации, по следующему принципу: к локализатору будут применены оптимальные параметры локализации, и оценка классификации будет выполняться только при условии, что на этапе локализации не было допущено ошибок.

### **3.4. Анализ результатов вычислительного эксперимента**

В данном разделе приведены таблицы, отображающие результаты вычислительного эксперимента, проведенного с изображениями из каждой выборки.

В таблицах, содержащих результаты работы локализатора, отражены варьируемые параметры, такие как цветовой диапазон различных компонент в пространства HSV и количество итераций применения алгоритмов Erosion и Dilation к изображению, а также показатели точности локализации в процентах. Сначала рассмотрим именно эти таблицы.

В таблице 3.1 приведены три варианта установки цветового диапазона для каждого из целевых цветов в цветовом пространстве HSV. Первый набор описывает широкие цветовые границы, второй – более узкие, и третий – еще более узкие, чем во втором наборе.

№	Синий цвет			Желтый цвет			Красный цвет		
	H	S	V	H	S	V	H	S	V
1	95-130	140-255	100-255	19-37	130-255	110-255	160-20	135-255	115-255
<b>2</b>	<b>99-125</b>	<b>155-255</b>	<b>115-255</b>	<b>23-34</b>	<b>141-255</b>	<b>124-255</b>	<b>170-10</b>	<b>141-255</b>	<b>124-255</b>
3	105-115	165-255	125-255	25-31	150-255	136-255	175-5	152-255	135-255

Таблица 3.1. – Варианты установки границ цветового диапазона

Для каждого из этих наборов проводился вычислительный эксперимент, в котором варьировалось количество итераций алгоритмов Erosion и Dilation. В таблице 3.2 приведены результаты вычислительного эксперимента с локализатором, проведенного с изображениями первой выборки. В эту таблицу включены только такие комбинации параметров, при которых итоговая точность локализации составляла не менее 70%.

№	Итераций Erosion	Итераций Dilation	Точность
1	2	2	76.43%
1	3	4	79.64%
1	4	1	74.12%
2	3	4	89.78%
2	4	3	90.33%
<b>2</b>	<b>4</b>	<b>2</b>	<b>94.71%</b>
2	2	2	86.52%
3	2	3	83.46%
3	4	4	79.12%
3	3	4	81.83%

Таблица 3.2. – Результаты вычислительного эксперимента с локализатором для первой выборки

Из таблицы 3.2 видно, что наилучшей комбинацией параметров локализатора для изображений, снятых в нормальных условиях, является второй (средний) цветовой диапазон из трех представленных, и следующие параметры алгоритмов применения морфологических операций – 4 итерации для Erosion, 2 итерации для Dilation.

Результаты вычислительного эксперимента с локализатором, проведенного с изображениями второй выборки, приведены в таблице 3.3. В эту таблицу включены только такие комбинации параметров, при которых итоговая точность локализации составляла не менее 70%.

№	Итераций Erosion	Итераций Dilation	Точность		
			ОСШ = 20дБ	ОСШ = 10дБ	ОСШ = 0дБ
1	4	2	90.11%	85.21%	70.41%
2	2	2	89.81%	86.52%	71.90%
<b>2</b>	<b>5</b>	<b>3</b>	<b>91.26%</b>	<b>87.46%</b>	<b>72.91%</b>
3	4	4	88.28%	81.12%	72.14%
3	3	4	87.92%	84.61%	70.76%

Таблица 3.3. – Результаты вычислительного эксперимента с локализатором для второй выборки

Как видно из таблицы 3.3, добавление искусственного шума к изображению значительно ухудшает качество работы локализатора. С ростом значения отношения сигнала к шуму, точность локализации снижается. Однако, наилучший результат показывает второй набор значений цветового диапазона с пятью итерациями Erode и тремя итерациями Dilate.

Результаты вычислительного эксперимента с локализатором, проведенного с изображениями третьей выборки, приведены в таблице 3.4. В эту таблицу включены только такие комбинации параметров, при которых итоговая точность локализации составляла не менее 70%. УП означает угол поворота в градусах.

№	Итераций Erosion	Итераций Dilation	Точность		
			УП = 15°	УП = 30°	УП = 45°
2	4	3	90.33%	90.33%	90.33%
2	3	4	89.78%	89.78%	89.78%
<b>2</b>	<b>4</b>	<b>2</b>	<b>94.71%</b>	<b>94.71%</b>	<b>94.71%</b>
1	3	4	79.64%	79.64%	79.64%
3	2	3	83.46%	83.46%	83.46%

Таблица 3.4. – Результаты вычислительного эксперимента с локализатором для третьей выборки

Как видно из таблицы 3.4, поворот изображения искомого объекта на любой из заданных углов никак не влияет на качество работы локализатора. Для любого угла поворота объекта в кадре алгоритм выдал одинаковые результаты. Это обусловлено тем, что алгоритм локализации ориентируется только на цвет, а не на положение объекта в кадре. Стоит отметить, что результаты локализации с приведенными комбинациями параметров из таблицы 3.4 хорошо согласуются с результатами локализации из таблицы 3.2 – аналогичные наборы параметров выдают аналогичные друг другу результаты. Наилучший результат показывает второй набор значений цветового диапазона с четырьмя итерациями Erode и двумя итерациями Dilate. Именно такой набор параметров был принят за оптимальные параметры локализатора и был использован в дальнейших вычислительных экспериментах с классификатором.

Отдельно был проведен анализ изображений из первой выборки, для которых применение оптимальных параметров локализации привело к ошибке. У каждого из таких изображений есть отличительная особенность от всей остальной выборки – неправильно выставленный баланс белого на камере. К примеру, на изображении виден снег голубого цвета, хотя на самом деле цвет снега – белый. Подобного рода цветовое искажение приводит к ошибке локализации. При ручном исправлении баланса белого на таких снимках, локализатор выполняет свою задачу успешно.

В таблицах, содержащих результаты работы классификатора, отражены показатели точности в процентах, среднего времени обработки одного изображения в секундах, уровня накладываемого шума в децибелах, угол поворота изображения объекта в градусах, а так же отражено наличие либо отсутствие экспериментальной функциональности во время проведения эксперимента – АРДП (алгоритм Рамера-Дугласа-Пекера).

Результаты вычислительного эксперимента с классификатором, проведенного с изображениями из первой выборки, приведены в таблице 3.5.

АРДП	Точность	Время
Да	98.53%	0.34 сек.
Нет	96.07%	0.26 сек.

Таблица 3.5. – Результаты вычислительного эксперимента для первой выборки

Как видно из таблицы 3.5, применение алгоритма Рамера-Дугласа-Пекера увеличивает точность результатов, но при этом повышает среднее время обработки одного изображения системой.

Результаты вычислительного эксперимента с классификатором, проведенного с изображениями из второй выборки, приведены в таблице 3.6.



АРДП	ОСШ	Точность	Время
Да	20дБ	91.67%	0.38 сек.
	10дБ	87.32%	0.38 сек
	0дБ	81.12%	0.4 сек.
Нет	20дБ	89.71%	0.27 сек.
	10дБ	84.59%	0.28 сек.
	0дБ	78.97%	0.31 сек.

Таблица 3.6. – Результаты вычислительного эксперимента для второй выборки

Как видно из таблицы 3.6, с увеличением интенсивности накладываемого шума, снижается точность распознавания. При этом корреляция между результатами работы системы с включенным и выключенным АРДП остается такой же, как и для первой выборки.

Результаты вычислительного эксперимента с классификатором, проведенного с изображениями из третьей выборки, приведены в таблице 3.7.

АРДП	Угол поворота	Точность	Время
Да	15°	97.87%	0.35 сек.
	30°	97.34%	0.34 сек
	45°	96.91%	0.36 сек.
Нет	15°	95.17%	0.27 сек.
	30°	94.76%	0.24 сек.
	45°	94.23%	0.26 сек.

Таблица 3.7. – Результаты вычислительного эксперимента для третьей выборки

Как видно из таблицы 3.7, наличие угла поворота вплоть до 45° не оказывает существенного влияния на точность распознавания. Это говорит о том, что глобальное признаковое описание на основе теории активного восприятия является инвариантным и к масштабу, и к повороту объекта.

В таблице 3.8 представлены результаты работы известных методов локализации объектов на изображении, рассмотренных в главе 1:

Метод	Точность
GIST	93-96%
HOG	89-94%
CoHOG	95-98%
Моменты Ху	94-98%

Таблица 3.8 – Точность работы известных методов обнаружения объекта

Сравнивая данные из таблицы 3.8 и полученные результаты тестирования алгоритма в различных условиях (таблица 3.5, 3.6 и 3.7), можно сделать вывод о том, что разработанный метод имеет точность обнаружения объектов на изображении, вполне сопоставимую с точностью рассмотренных известных методов. По сравнению с некоторыми известными методами точность предложенного подхода к решению задачи оказывается выше.

### 3.5. Выводы по главе 3

В данной главе было рассмотрено описание реализованной программной системы и описание вычислительного эксперимента, проводимого с целью определить точность и время работы разрабатываемого метода, а также анализ его результатов и сравнение с наиболее популярными из существующих методов. Было приведено подробное описание базы изображений, которая использовалась для тестирования, то, как она была модифицирована для проведения наиболее полного тестирования. В главе также был описан программный продукт, реализующий алгоритм поиска объектов на изображении, рассмотрены результаты его работы.

При анализе результатов тестирования были выбраны параметры, дающие наибольшую точность и наименьшее время работы. Полученная точность при сравнении с точностью существующих методов показала, что в нормальных условиях разрабатываемый метод работает аналогично известным методам. Кроме того, включая дополнительный функционал в виде алгоритма Рамера-Дугласа-Пекера, предоставляется возможным повысить точность распознавания на 3-5%, однако замедлив среднее время обработки изображения на 10-15%

## Заключение

В ходе выполнения выпускной квалификационной работы было произведено исследование существующих методов для поиска объекта на изображении, а также исследование методов, применяемых на каждом из этапов поиска. В рамках этого обзора методы были проанализированы, выделены их достоинства и недостатки, изучены проблемы, с которыми сталкивались авторы методов и люди, использующие их.

Основываясь на полученной информации, был разработан новый алгоритм решения задачи поиска объекта на изображении, который заключается в новой комбинации подходов для решения задачи на разных ее этапах, а так же использовании нового способа формирования глобального признакового описания на этапе формирования признакового описания. В данной работе, таким подходом является использование инструментария теории активного восприятия.

Для реализации метода была разработана программа на языке Python, были найдены и подготовлены тестовые изображения разной сложности для проверки работоспособности системы. В результате проведения экспериментов на этих тестовых изображениях и их анализа, был сделан вывод о корректной работе метода при решении задачи обнаружения объекта на изображении.

Точность работы системы в нормальных условиях оказалось сопоставимой с точностью существующих методов. Точность работы на зашумленных изображениях оказалась достаточной высокой и приемлемой, а также разработанный метод может поддерживать случаи с двумя и более искомыми объектами на изображении, локализуя их все.

В дальнейшем планируется провести пробный запуск системы, выполнив её портирование на смартфон. Данный эксперимент позволит провести исследование, используя видеопоток с камеры смартфона в качестве входной информации для анализа.

Таким образом, можно сделать вывод, что разработанный метод можно использовать на практике. Задача, поставленная в начале выполнения выпускной квалификационной работы, была выполнена в полном объеме, а цели достигнуты.

## Список литературы

1. A. Oliva, A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope", Int. J. Comput. Vis., vol. 42, no. 3, pp. 145-175, 2001.
2. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005.
3. T. Watanabe, S. Ito, K. Yokoi, "Co-occurrence histograms of oriented gradients for pedestrian detection", PSIVT, 2009.
4. M. K. Hu, "Visual Pattern Recognition by Moment Invariants", IRE Trans. Info. Theory, vol. IT-8, pp.179–187, 1962
5. Y. Liu, H. Zhang, "Visual loop closure detection with a compact image descriptor", Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst., pp. 1051-1056, Oct 2012.
6. A. Murillo, G. Singh, J. Košecká, J. Guerrero, "Localization in urban environments using a panoramic gist descriptor", IEEE Trans. Robot., vol. 29, no. 1, pp. 146-160, Feb 2013.
7. M. Douze, H. Jégou, H. Singh, L. Amsaleg, C. Schmid, "Evaluation of GIST descriptors for web-scale image search", CIVR, July 2009, Art 19.
8. Zhang C. Computer Vision: What is a GIST descriptor?  
<https://www.quora.com/Computer-Vision-What-is-a-GIST-descriptor>
9. Южаков Г. Б. Алгоритм быстрого построения дескрипторов изображения, основанных на технике гистограмм ориентированных градиентов // ТРУДЫ МФТИ, 2013. № 3. С. 84-91.
10. Belkasim S.O., Shridhar M., Ahmadi M. Pattern recognition with moment invariants: a comparative study and new results// Pattern Recognition, 1991. Vol. 24. P. 1117-1138.
11. Dudani S.A., Breeding K.J., McGhee R.B. Aircraft identification by moment invariants// IEEE Trans. Comput., 1977. Vol. 26. P. 39-45
12. Wong R.Y., Hall E.L. Scene matching with invariant moments// Comput. Graphics Image Process., 1978. Vol. 8. P. 16-24.
13. Maitra S. Moment invariants// Proc. IEEE, 1979. Vol. 67. P. 697-699.
14. Hupkens T.M., de Clippeleir J. Noise and intensity invariant moments// Pattern Recognition, 1995. Vol. 16. P. 371- 376.
15. Wang L., Healey G. Using Zernike moments for the illumination and geometry invariant classification of multispectral texture// IEEE Trans. Image Process, 1998. Vol. 7. P. 196-203.

16. Flusser J., Suk T. Degraded image analysis: an invariant approach// IEEE Trans. Pattern Anal. Mach. Intell.,1998. Vol. 20. P. 590-603.
17. Абрамов Н. С., Фраленко В. П. Определение расстояний на основе системы технического зрения и метода инвариантных моментов // Информационные технологии и вычислительные системы. — 2012. — № 4. — С. 32-39
18. Роджерс Д., Адамс Дж. Математические основы машинной графики. – М.: Машиностроение, 1980. – 240 с.
19. Duda, R. O. and P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Comm. ACM, Vol. 15, pp. 11–15 (January, 1972)
20. A. Broggi, P. Cerri, P. Medici, P. Porta, and G. Ghisio, "Real time road signs recognition," Intelligent Vehicles Symposium, 2007 IEEE, pp. 981–986, June 2007.
21. A. D. L. Escalera, J. M. A. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," Image and Vision Computing, vol. 21, pp. 247–258, 2003.
22. X. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, "Recognition of traffic signs based on their colour and shape features extracted using human vision models," Journal of Visual Communication and Image Representation, vol. 17, no. 4, pp. 675–685, 2006.
23. Viola, Jones: Robust Real-time Object Detection, IJCV 2001 See pages 1,3
24. Утробин В. А. Элементы теории активного восприятия изображений // НГТУ, 2001. - 64 с.
25. R.A. Haddad and A.N. Akansu, "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 39, pp 723-727, March 1991
26. Travis, D. Effective Color Displays Theory and Practice // Academic Press, ISBN 0-12-697690-2. – 1991. – 328 p
27. Шахуро В., Конушин А. Российская база изображений автодорожных знаков. Компьютерная оптика, 2016